# PARAMETRIC REAL-TIME RECOGNITION of ISOLATED HAND GESTURES using WEARABLE SENSORS

Probabilistic Meta-Classification of American Sign Language Signs

English translation of the original dissertation

Originally published in German by Technische Universität Darmstadt

PHILIPP ACHENBACH, M.SC.

2026

**Disclaimer**
This English version was machine-translated using GPT-5.4 and has not been manually verified. It is provided solely as an accessibility aid to make the already published German dissertation available to a broader audience. In case of discrepancies, translation errors, or citation inconsistencies introduced by this translation, only the original German publication is authoritative.

**How to cite the original publication**
Please cite the original German dissertation rather than this translation. A preferred citation is:

Philipp Achenbach: *Parametrische Echtzeiterkennung von isolierten Handgesten mittels körpergetragener Sensorik.*
*Probabilistische Meta-Klassifikation amerikanischer Gebärden*. Dissertation,
Technische Universität Darmstadt, 2026.

URN: `https://nbn-resolving.org/urn:nbn:de:tuda-tuda-154735`
URL: `https://tuprints.ulb.tu-darmstadt.de/handle/tuda/15473`

This English translation should be used as a reading aid only.

## SUMMARY

This work presents *Parametric Framework for Sign Language Recognition (ParaSignRec)*, an approach for parametric real-time recognition of isolated hand gestures using body-worn sensors. While previous approaches largely treat signs holistically and rely on classifiers that recognize complete signs as individual classes, the approach presented here pursues a fundamentally different path. Signs are decomposed into their elementary, meaning-distinguishing components and classified at the parameter level by seven different specialized recognition approaches (for example handshape, palm orientation, and movements). A subsequent meta-classification identifies the executed sign through a weighted aggregation of these individual results and generates a ranking of all signs in the defined vocabulary.

This vocabulary, alongside the parameter-based recognition approaches and the meta-classification, represents a central result of this work. It is based on ASL-LEX, a lexical database for signs of American Sign Language (ASL), and includes 2,171 out of an original 2,723 signs. These are described by 28 different parameters, which are evenly distributed across the dominant and non-dominant hand. Significant optimizations of the underlying data were achieved through the introduction of multivalued parameter classes that take natural variations of signs into account, as well as through the addition of palm orientation as an additional parameter, which proved to be a relevant factor for improving classification results and increasing the size of the vocabulary.

The main advantages of the parametric approach lie in its structurally induced generalizability, since the individual recognition approaches are trained independently of one another on parameter-specific datasets rather than being based on a shared dataset. In addition, there is a significant reduction in the effort required for data collection: up to $5.39 \times 10^{20}$ new signs (-variants) can be integrated into the vocabulary without re-recording and complete retraining of the underlying recognition approaches, provided that the required classes already exist.

To capture the signs, a combination of data gloves and Inertial Measurement Units (IMUs) attached to the upper and forearm is used. The system achieves a top-1 accuracy of 80.41 % and can identify the sought sign among the first ten candidates in the ranking in 95.23 % of cases. The average processing time is 122.67 ms, so the classification of each isolated sign is performed significantly faster than its execution. The results thus show that a parameterized view of signs is not only possible but also useful, and that it enables probabilistic identification of signs from a limited, yet functionally sufficient and extensible, vocabulary of ASL in real time.

# ABSTRACT

This dissertation presents *Parametric Framework for Sign Language Recognition (ParaSign-Rec)*, an approach to parametric real-time recognition of isolated hand gestures using body-worn sensors. Whereas previous approaches have tended to view gestures holistically and rely on classifiers that recognize complete gestures as individual classes, the approach presented here takes a fundamentally different path. Gestures are broken down into their elementary, meaning-distinguishing components and classified using seven different specialized recognition approaches at the parameter level (e.g., hand shape, palm orientation, and movements). A subsequent meta-classification identifies the executed gesture by weighted aggregation of these individual results and generates a ranking of all gestures in the defined vocabulary.

This vocabulary represents a key result of this work, alongside the parameter-based recognition approaches and meta-classification. It is based on ASL-LEX, a lexical database for signs in American Sign Language (ASL), and contains 2,171 of originally 2,723 signs. These are described by 28 different parameters, which are distributed evenly between dominant and non-dominant hands. Significant optimizations of the data basis were achieved by introducing multi-valued parameter classes that take natural variations in signs into account, as well as by adding palm orientation as an additional parameter, which proved to be a relevant factor in improving classification results and increasing vocabulary size.

The main advantages of the parametric approach lie in its structural generalizability, as the individual recognition approaches are trained independently of each other on parameter-specific data bases rather than on a common data set. In addition, there is a significant reduction in the effort required for data collection: up to $5.39 \times 10^{20}$ new signs (and variants) can be integrated into the vocabulary without re-recording and completely retraining the underlying recognition approaches, provided that the required classes already exist.

A combination of data gloves and Inertial Measurement Units (IMUs) attached to the upper and lower arms is used to capture the gestures. The system achieves a top-1 accuracy of 80.41 % and can identify the desired gesture among the top ten candidates in the ranking in 95.23 % of cases. The average processing time is 122.67 ms, so that the classification of each isolated sign is much faster than its execution. The results thus show that a parameterized analysis of gestures is not only possible but also useful, enabling real-time probabilistic identification of gestures from a limited but functionally sufficient and expandable vocabulary of ASL.

# PUBLICATIONS AND TOOLS

## PREVIOUSLY PUBLISHED MATERIAL

This dissertation is based in part on material that has already been published in scientific journals and at academic conferences. An overview of the relevant publications used in this work is shown in Table 1. Care was taken to ensure that no text was taken directly from these publications. Graphics, especially those that illustrate central concepts or present specific results, were instead replicated or adapted and appropriately marked. Tables that show conceptual summaries or data analyses were also adapted in structure, but retained in terms of content.

Scientific research often arises through the collaborative effort of a team. This dissertation is devoted to an interdisciplinary field of research that encompasses electrical engineering, computer science, and linguistics in connection with sign language. The results of this work therefore represent a collective achievement by experts from the fields of computer science, engineering, and linguistics in the field of sign language. In recognition of this teamwork, the contributions of all co-authors and contributors, including their institutional affiliations, are made transparent. If no specific affiliation is stated upon first mention, the person concerned is either currently or previously employed at Technische Universität Darmstadt, specifically at the Multimedia Communications Lab (KOM) or the Serious Games group. In this dissertation, the use of the word *we* refers to the joint work and contributions of all those involved.

In the following sections, the chapters are discussed that contain both verbatim and paraphrased excerpts from my publications or collaborative projects. A comprehensive list of all my publications, including those that do not directly feed into this work, can be found in Appendix C.

Throughout the entire research and writing process, PD Dr.-Ing. Stefan Göbel assumed a supervisory role and made substantial contributions with regard to methodology, approach, and the final design of the work. Due to this continuous involvement, he is not mentioned separately in every individual contribution.

Chapter 2, BACKGROUND AND RELATED WORK includes a detailed overview of numerous notation systems in Section 2.2.1 that was created in collaboration with Yasemin Göksu and Timo Kullmann and was also published in our paper "Towards handshape identification for automatic gesture recognition using sign notation systems" [4]. The analysis and compilation of over 90 related works in Section 2.3 and Section A.2 was created in collaboration with Jakob Ohage and Silvia Zoraqi.

Chapter 4, RECOGNITION OF HANDSHAPE evaluates different concepts for handshape recognition that emerged as part of numerous theses: Section 4.2.1 examines the suitability of Electromyography (EMG) for handshape recognition and was created in collaboration with Lea Schott. The basis for this was provided by André M. Kleebe, Ute Lehmann and Jan Ulrich Schmitt with an EMG-controlled game, whose

findings we published together with Philipp Niklas Müller and Thomas Tregel in our paper "Flex your muscles: EMG-based serious game controls" [150]. Section 4.2.2 was created in collaboration with Marius Kempf and Philipp Niklas Müller and examines the suitability of Mechanomyography (MMG) for handshape recognition, whereas Section 4.2.3 was created in collaboration with Dennis Purdack and evaluates camera-based approaches.

As a fourth and most promising approach, the suitability of a data glove was examined. This was first done in different forms with a simple data glove model based on the recognition of rock-paper-scissors gestures. In collaboration with Tobias Wach (Section 4.2) and, building on this, with Sebastian Wolf (Section 4.2 and Section 4.3), a corresponding demonstrator was developed that was presented on numerous occasions, such as the Hessen Creative Economy Day 2023, and published in several papers ([6, 7]).

Together with Nico Kunz, these findings were used to record numerous sign handshapes with an improved data glove, which were used in Section 4.4 through Section 4.5 and as part of the training data for the final classifier for handshape recognition. This, in turn, is based on collaboration with Dennis Purdack and Sebastian Laux and is discussed in detail in Section 4.3 and Section 4.4 and was published together with Philipp Niklas Müller in our paper "Give Me a Sign: Using Data Gloves for Static Hand-Shape Recognition" [5].

A second demonstrator based on new handshape recordings and a third generation of data gloves was developed together with Wolfgang Brabänder and Sebastian Laux, with the associated data being recorded by Mateusz Ciesielski, Nicolas Liebig, Lena Todtenhöfer and Florian Wiegert. These data formed the second part of the training data for the final handshape classifier in Section 4.5.

Chapter 5, FURTHER SYSTEM COMPONENTS addresses the conception and implementation of further recognition approaches and, in this context, the development of custom hardware and recording tools. Felix Klose, Sarah Kuhlmann and Qilin Tan provided support in the development of the arm controller in Section 5.1, while Sarah Kuhlmann also made important contributions to the implementation of the recording and segmentation tool in Section 5.2. The conception of the calibration process in Section 5.2.3 took place under the advisory participation of Philipp Niklas Müller.

Further recognition approaches were developed in collaboration with Thanh Huan Hoang, Sarah Kuhlmann, Minh Pham, Rebekka Schiller, Maximilian Segers and Qilin Tan: Qilin Tan, Thanh Huan Hoang and Minh Pham addressed the recognition of the place of articulation in their respective theses and in doing so each continued the results and findings of the others to the outcome described in Section 5.3. Under the guidance of Sarah Kuhlmann, all training data for the design of threshold-based movement recognition in Section 5.4.1 were recorded, whereas Rebekka Schiller was responsible for the data recordings for Section 5.4.2 and contributed important insights into path movement recognition. Maximilian Segers provided support in the recognition of palm orientation in Section 5.5.

Chapter 6, EVALUATION OF THE OVERALL SYSTEM evaluates the overall system and also the sign dictionary and the individual parameter-based recognition approaches.

Both Elena Bock and Benedikt Hock as well as Alexandra Skogside addressed the MediaPipe tool as part of their theses, and the findings from this were incorporated into the design of the automatic annotation tool in Section 6.1.2.

In collaboration with Nicolai Kellerer, a tool for the automatic creation of lists of test signs to be recorded was developed, which fulfills the criteria mentioned in Section 6.1.4. These were recorded with the help of Philipp Kaul, Sebastian Laux, Minh Pham and Rebekka Schiller (see Section 6.2.1).

In addition, contact was also maintained with scholars from linguistics as well as representatives of the sign language community. Several email correspondences took place with Hochgesang [97] and Sehyr et al. [187], in which questions regarding *SignBank* and ASL-LEX were clarified.

Moreover, the DGS-Fabrik Frankfurt, in cooperation with Eintracht Frankfurt e.V., organized a German Sign Language course that enabled in-depth insights into sign language and clarified further questions.

### TOOLS USED

As part of this dissertation, various tools in the form of digital applications were used to support the scientific workflow. Literature management and organization were carried out with Zotero. For the development and implementation of custom software components, the source code editor Visual Studio Code and later Cursor as well as GitHub were used as version control systems. Cursor partially supported code completion and optimization suggestions through AI-based functions. This support was limited to technical assistance in the sense of an intelligent editor; the conceptual and functional design of the developed programs was carried out independently.

For the linguistic optimization of individual sections, language-based AI applications such as ChatGPT were used in part. This support was limited solely to assistance with wording. All substantive explanations, arguments, and assessments are the intellectual property of the author or are indicated by sources. The translation software DeepL was used on isolated occasions to translate excerpts of foreign-language technical literature and in part to translate the abstract of this work into English. Machine translations of external works served exclusively to aid understanding of complex passages; automated translations were not directly incorporated into the work.

All tools were used in compliance with the principles of good scientific practice and the applicable regulations of the department. Should the information about a tool have inadvertently remained incomplete, this was not intentional; all scientific content and assessments in this work are based on the author's independent achievement.

Table 1: Previously published material

| Chapter | [4] | [5] | [6] | [7] | [42] | [150] | [151] |
|---|---|---|---|---|---|---|---|
| Chapter 1: INTRODUCTION | | | | | | | |
| Introduction | ✓ | | | | | | |
| Motivation | ✓ | | ✓ | ✓ | | | |
| Chapter 2: BACKGROUND AND RELATED WORK | | | | | | | |
| Sign Language | ✓ | | | | | | |
| Parametrization of Signs | ✓ | | | | | | |
| Sign Recognition Using Body-Worn Sensors | | ✓ | ✓ | ✓ | | ✓ | |
| Chapter 3: PARASIGNREC: PARAMETRIC FRAMEWORK FOR SIGN RECOGNITION | | | | | | | |
| Introduction | | ✓ | | | | | |
| Parameter identification | ✓ | | | | | | |
| Parameter recognition | | ✓ | | ✓ | ✓ | | |
| Meta-classification using probabilistic models | | | | | | | |
| Chapter 4: RECOGNITION OF HANDSHAPE | | | | | | | |
| Introduction | | ✓ | | | | | |
| Selection of the hand parameters to be captured | ✓ | | | | | | |
| Evaluation of a system for capturing and analyzing handshapes | | ✓ | ✓ | ✓ | | ✓ | |
| Evaluation of a classification method for handshape recognition | | | | ✓ | | | |
| Evaluation of data processing for classification optimization | | ✓ | | ✓ | | | |
| Finalization of the overall system for handshape recognition | | | ✓ | ✓ | | | |
| Recognition of contact between thumb and finger | | | | ✓ | | | |
| Chapter 5: FURTHER SYSTEM COMPONENTS | | | | | | | |
| Recognition of the Sign Articulation Location | | | | | ✓ | | |
| Recognition of Hand Movements | | | | | | | ✓ |

CONTENTS

# 1 INTRODUCTION

Interpersonal communication is not limited to spoken language alone. Nonverbal forms of expression play a central role – especially where language reaches its limits. They include not only universally understandable gestures and body language that are used across national and linguistic boundaries, but, in the case of sign language, are also capable of fully replacing spoken language [71, 139]. Already 72 million deaf and hard-of-hearing people worldwide use one of the roughly 200 to 300 different sign languages, 60 of which have been studied [58, 223]. Integrating this complex language form into innovative technical applications offers not only new perspectives for interaction between humans and machines, but also holds great potential for promoting inclusion in our technologized world. Against the backdrop of legal requirements for accessibility – such as those under the Accessibility Strengthening Act (BFSG) [59] or EU Directive 2016/2102 [222] – it becomes clear that digital offerings, including websites and communication systems, must be accessible to all people. This dissertation makes a valuable contribution in this regard by investigating and developing technical approaches to integrating nonverbal communication. In doing so, it connects the technical fields of electrical engineering and computer science with the field of linguistics, each with its own particular challenges.

One sign language of particular relevance is American Sign Language (ASL), which, despite its widespread use[1], still poses challenges for those who do not master this language. The dissertation presented here aims to close this communication gap with the help of modern technology and to open up the possibility of translating signs, pursuing the approach of not recognizing *signs* as a whole, but rather decomposing them into their characteristic components and identifying them on that basis. The parameterization approach is essentially based on the insights of the linguist William C. Stokoe [205], who was the first to highlight and demonstrate the complexity and linguistic depth of ASL by reducing ASL signs to three basic, meaning-distinguishing form parameters [14, 71, 217]: handshape, the location of articulation, and movement of the hand. The orientation of the hand was subsequently added as a separate parameter by Robbin Battison [23], since in Stokoe it was part of the handshape [23, 217]. This property of parameterization, which was originally attributed exclusively to spoken language, revolutionized the understanding of sign languages and proved that ASL is an independent language. Subsequent work added additional parameters, with the parameters already listed always forming the basis (only the orientation of the hand is not considered in some few concepts or is mentioned in combination with the handshape) [4, 71].

The field of sign recognition is already being researched. Holistic approaches dominate here, focusing on a comprehensive view of signs (see Section A.2). The

1 ASL is the most widely used sign language [58, 223]

approach presented in this work differs fundamentally from this by identifying and interpreting the elemental components – the parameters of signs – instead. This means that there is no need to concentrate on a direct and often inflexible translation of whole signs. This approach promises increased adaptability and scalability. For Machine Learning (ML)-based recognition approaches, each subject to be recognized must be recorded, captured from different users with multiple repetitions. Instead of having to record hundreds to thousands of signs under these conditions, our approach requires only recordings of the individual parameters with a comparatively small number of classes. By combining these parameters, the sought-after sign can be identified precisely. The training effort for classification can thus be significantly reduced, although it requires a suitable data basis. Another advantage of this approach is that new signs composed of combinations of known parameters can be integrated into the vocabulary without fully retraining the system. In contrast, the holistic recognition approach requires constant updating and retraining of the models as the vocabulary of the living language continues to evolve.

An important aspect in this context is the availability of suitable sensors and hardware for the precise capture of complex signs. In particular, the increased attention paid to Virtual Reality (VR) and Augmented Reality (AR) in recent years has led to significant developments in the field of data gloves [61, 164]. Modern data gloves are now available at comparatively low cost and equipped with a variety of sensors. Most approaches use inertial sensors and flex sensors, which enable precise capture of movements and positions [54, 164]. More recent approaches capture the positions of the individual fingers through measurements of changes in the magnetic field around the data glove and can thus deliver even more precise results and detect the finest nuances of hand and finger movements [54]. There have also been advances in sensor technology for motion detection, such as Inertial Measurement Units (IMUs), which can be attributed to developments and increased demands in the area of VR/AR. Thanks to more powerful hardware, state-of-the-art sensors offer onboard sensor fusion, making it possible to obtain integrated and fused sensor values directly. Due to their compact design and good availability, these sensors are ideally suited for developing controllers that can be attached to the upper and lower arm and provide their movements and orientations via Bluetooth [227].

Beyond the technical aspects, one must not forget the relevance of this research in the field of assistive technologies: The aim of this dissertation is to develop an inclusive technology that enables the deaf community to communicate more effectively and to live without language barriers.

In summary, this work presents a parametric approach for recognizing isolated American signs from a vocabulary suitable for simple conversations. It does not view signs as a whole, but decomposes them into their meaning-distinguishing parameters and identifies them through a meta-classification by seven different specialized recognition approaches at the parameter level. The signs are captured in real time using mobile, body-worn sensors. The evaluation has shown that individual parameters contribute to the identification of signs with varying strength and that, in particular, the recognition of handshape and palm orientation has a significant influence on ac-

curacy and vocabulary size. At the same time, limitations of individual parameters as well as technical and conceptual challenges become apparent, which are systematically analyzed and contextualized over the course of the dissertation.

## 1.1 MOTIVATION

According to the World Health Organization, around 430 million people worldwide live with an impairment of the sense of hearing [233], which corresponds to approximately 5 % of the world population. The trend is rising; by 2050, every fourth person is expected to suffer from hearing problems [234], and every tenth person from complete hearing loss [233].

Everyday communication between deaf and hearing people is generally challenging. Some deaf people are able to lip-read and thus understand the person they are speaking with. Nevertheless, mutual communication is not ensured, as they cannot be understood, or can often only be understood with great difficulty. There is also no uniform sign language as the language of deaf people, independent of the presence of spoken language [71]. Worldwide, around 72 million deaf people speak over 200 to 300 different sign languages [58, 223]. Sign language is used in only about one in four families with deaf children, and 98 % of deaf people receive no education in sign language [228], with the situation in developing countries being particularly precarious [233].

All of these points lead to affected adults often being disadvantaged in their professional lives. Not only are those affected harmed; society is also harmed by annual worldwide costs amounting to 980 billion dollars. Of this, 57 % is borne by countries with low and middle income levels. This amount includes costs in the areas of health care (excluding the costs of hearing aids), education, society, and losses in productivity [233].

In view of the obvious challenges and needs of the deaf community, as described above, it becomes clear that more effective solutions for learning and using sign language are urgently needed. Everyday alternative solutions, such as the use of written text in direct conversations on the street or at the workplace, are cumbersome and reduce communication to a minimum. The need for an innovative solution that allows deaf people to make themselves understood in everyday life to hearing people is evident both from the perspective of those affected and from the perspective of society [4].

Such a solution could also help the direct relatives and social environment of those affected: with the aid of a portable sign language translator, it would, for example, be possible not only to translate signs into text or speech, but also to integrate it into a learning environment in order to learn sign language. However, in order to be usefully employed within a learning app, the vocabulary of this learning app would have to be of a similar scope to that of a dictionary.

To be able to help those affected, the desired solution must prove practical in everyday life. The use of cameras to capture signs enables a cost-effective and simple

application [28], but it requires an unrestricted view of the entire gesture space[2] of the sign language user, which presupposes a camera position independent of the sign language user. This, however, conflicts with the requirement of a mobile, everyday-suitable approach. In terms of the sign language user's privacy and that of their fellow human beings, the use of cameras should also be avoided.

This is where current developments in the field of AR and VR come in. Due to the increased attention these two areas have received in recent years [61], numerous novel concepts for Human-Computer Interaction (HCI) have emerged. Rather than the user learning to control machines, they use natural and instinctive means of communication, and the machine learns to understand them. Voice control enables us to operate devices with our voice as if we were delegating tasks to a colleague over the phone [162]. With the advent of smartphones and tablets, swipe and touch gestures also became established in our society as another example [6]. VR-based games, such as *Half-Life: Alyx*, also rely on the use of gestures to operate tools and pick up objects in order to increase immersion [7]. In addition to hand controllers, the emergence of novel data gloves in the field of VR now makes further innovative interaction possibilities conceivable. Through the precise capture of all finger configurations and movements as well as hand movements and orientation, modern data gloves are ideally suited to recording hand gestures. Simple hand gestures, such as the "thumbs up" gesture, could, for example, be used to silently issue commands within an emergency response simulation. Significantly more complex content could also be represented in the form of sign language. This could be because it is prescribed by the application (e.g. sign language learning software), or because this is necessary since sign language is the primary form of communication for the users (e.g. for deaf and hard-of-hearing people).

There are already numerous publications on sign recognition using data gloves (e.g. [73, 92, 145, 161, 170, 177, 181, 191, 219, 235, 236, 239] and many more, see Section A.2). So far, these have considered signs as a whole in their recognition approach. To compensate for the variance in the execution of signs between different individuals, every sign must be recorded in multiple instances for the training of machine learning classifiers in this holistic approach. If the vocabulary is later expanded, data must be collected not only for the new signs, but the entire classifier must also be retrained. The effort involved in data acquisition and training of the recognition system is therefore linearly dependent on the number of signs. In the case of a vocabulary comprising several thousand signs, this results in enormous effort. As part of this dissertation, a comprehensive analysis of existing approaches was carried out in Section A.2. It was shown that the approaches consider only a small vocabulary of up to 100 signs, which is not suitable for comprehensive everyday conversation. An alternative to this holistic (holistic) view of signs is the parametric recognition approach presented in this work. It breaks signs down into their fundamental components before the desired sign is then identified on this basis. Parametrization makes it possible to significantly reduce the effort required for data acquisition and training of the recognition systems. Rather than training each sign individually, the parametric approach focuses on identifying

2 The area around the user in which the sign is performed.

individual parameters and interprets the desired sign by combining the different parameters. This also makes it possible to add new signs to the approach subsequently without retraining, provided the individual classes are already known to the system. If this is not the case, then only the missing classes need to be re-recorded and the classifier of the corresponding parameter retrained.

Parameter recognition requires a comprehensive understanding of the structure and characteristics of signs. Especially in the field of sign notation, which transfers signs into written form, there are already numerous approaches to parametrizing signs [4, 79]. These often pursue different goals and, associated with this, a differing level of detail in the parametrization, which is reflected in the number of parameters and their possible classes [4, 79]. These findings make it possible to create the basis for a parametric, lexical database, such as the one we require for our project.

## 1.2 CHALLENGES

To develop a practical solution for overcoming the communication barrier already mentioned between sign language users and non-sign-language users, several aspects must be taken into account. These include the sensor technology employed, the recognition approach, and the relevant vocabulary of the sign language. In the following, the research challenges that are decisive for implementing a sign recognition system suitable for everyday use have been identified.

**Challenge 1:** *Extensive and flexible vocabulary*

A central criterion for the everyday usability of a sign recognition system is a sufficiently large and functional vocabulary that covers basic communication needs and enables simple everyday conversations. The extent to which such a vocabulary must be developed within the context of a sign language still needs to be specified more precisely in this work. Previous approaches using body-worn sensor technology have vocabularies of up to 100 signs [145], many of them are significantly smaller (on average 21.55 signs across more than 90 studies examined, see Section Section A.2). Since ASL is a continuously evolving language, there is also the requirement that the vocabulary be expandable and adaptable in order to ensure the continued relevance of the approach.

**Challenge 2:** *Creation of a parameterized data basis*

An ML-based recognition system requires a suitable data basis. In the case of a parameter-based approach, not only a large amount of training data to be recorded is needed, but also a concept for parameterizing the signs into their characteristic components. This includes both the definition of the parameters and the specification of the parameter values to be recognized, i.e., the classes. In addition, manual annotation (labeling) of the intended vocabulary, which is to comprise several hundred signs, represents a considerable challenge due to the enormous amount of work involved.

**Challenge 3:** *Everyday-suitable hardware*

To ensure that the proposed approach is well accepted, it is essential that it be designed to be equally robust, mobile, and user-friendly. The hardware should also be capable of capturing the required parameters comprehensively and as precisely as possible. A definition of cost constraints was deliberately omitted in the context of this study in order to enable an unbiased and broader development.

**Challenge 4:** *Real-time recognition*

For natural and fluid communication, the ability of real-time sign recognition is indispensable. A sign should be captured and recognized faster than the subsequent sign is performed. If this is not the case, the conversation can be adversely affected and the communication flow cannot be maintained. Therefore, a fast processing time of the recognition approach and low latency are crucial.

**Challenge 5:** *Reliable recognition approach*

The accuracy of sign recognition is crucial to the effectiveness of the system. This includes the correct recognition of the individual parameters and, based on this, the most effective possible interpretation of these results for identifying the sought-after signs, i.e., both quickly and precisely. The emergence of ambiguous or erroneous results must be taken into account in the design of the approach so that a sign can be returned in any case. Consideration of the variance in the execution of signs, for example among different users, is also essential for universal applicability.

## 1.3 RESEARCH OBJECTIVES AND CONTRIBUTION

The primary objective of this thesis is the conception, implementation, and evaluation of a system for recognizing signs of American Sign Language (ASL). This objective is divided into three sub-objectives that address the challenges named in Section 1.2 and are explained in more detail below.

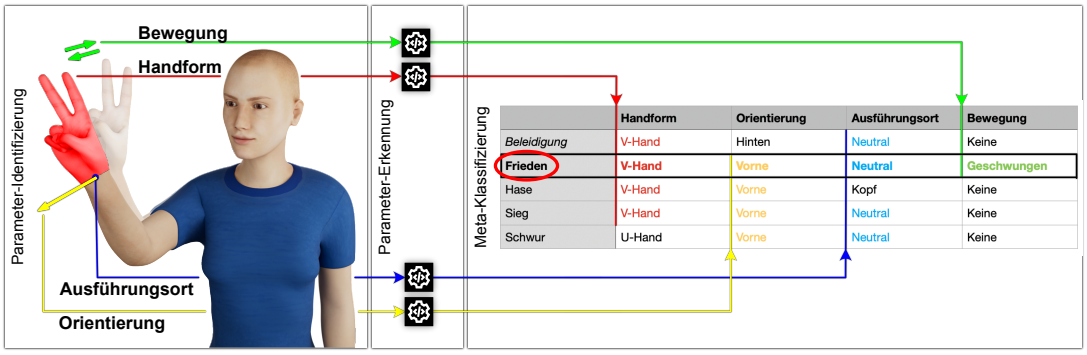



Figure 1: Parametric approach for recognizing isolated hand gestures using the example of a freely interpreted gesture

**Research Objective 1:** *Parameter Identification*

Before signs can be captured and recognized, it is important to investigate how signs are structured and which distinctive parameters they consist of. Only with these insights is it possible to parameterize signs and classify their parameters using suitable recognition approaches. With the aid of probabilistic models and a suitable database, the sought-after signs can then be identified from the recognized classes. This approach is schematically illustrated in Figure 1. Parameterization contributes to a significant reduction in the effort required for data collection for ML and makes it possible to keep the vocabulary both scalable and flexible. The steps required for parameter identification are therefore: *i)* identification of the characteristic parameters, *ii)* derivation of the possible classes for each parameter, and *iii)* development of an adequate, parameterized lexicon as a data basis for the subsequent parameter recognition.
This research objective addresses Challenges 1 and 2.

**Research Objective 2:** *Parameter Recognition*

After identifying the parameters to be recognized, the focus of this research objective is on the development of a robust, mobile, and user-friendly real-time parameter recognition system for signs in ASL. These sub-objectives are reflected both in the selection of suitable hardware for capturing the signs and in the design of the recognition approach. To increase scalability and flexibility, a parametric approach consisting of a variety of classifiers is pursued. These are used both to classify static and dynamic parameters, including those that build upon one another. The variation in sign execution by different users must also be taken into account. The objective therefore consists of: *i)* the selection of mobile and user-friendly sensors for capturing the parameters, and *ii)* implementation of the classifiers for the individual parameters, with *iii)* equally high accuracy and low classification time, taking into account variation in execution by different users.
This research objective addresses Challenges 3 to 5.

**Research Objective 3:** *Meta-classification*

Through meta-classification, the intended sign can be inferred from the captured and recognized parameters. An important component of this process is handling ambiguous or erroneous results, for example when the determined parameter configuration cannot be uniquely assigned to a specific sign. It is necessary to implement probabilistic models that, despite such inconsistencies, can return the most likely sign in order to improve the system's accuracy. The following steps are therefore required: *i)* implementation of a meta-classification using probabilistic models *ii)* taking into account a defined sign catalog, and *iii)* evaluation of the approach under realistic conditions.
This research objective addresses Challenge 5.

A central challenge is to achieve the highest possible recognition accuracy while maintaining a high level of detail, in order to cover an extensive vocabulary. This extends across all three research objectives and poses a particular difficulty: detailed

recognition often conflicts with accuracy, since an increasing level of detail also increases the number of distinguishable classes, which typically leads to a higher likelihood of confusion.

The approach also has the requirement of recognizing and translating signs quickly and efficiently. We define *real time* such that the processing time between the execution of a single sign and the output of its translation on commonly used hardware (comparable performance with e.g. *Apple MacBook Pro*) is less than the execution time of the subsequent sign. This is necessary in order to maintain the flow of communication. Ideally, each sign can be recognized faster than the fastest sign can be performed. We define the minimum requirement such that the average of all signs can be recognized faster than the average execution time of all signs. This requires low latency of the sensors used, fast data processing and classification of the individual parameters, but also efficient meta-classification and therefore extends over the last two of the three research objectives.

## 1.4 LIMITATIONS OF THE APPROACH

Due to the high prevalence and the good availability of data material and resources, this research work focuses on ASL. This decision makes it possible to draw on a rich data basis, which is of essential importance for the development and evaluation of a parametric recognition system. The widespread availability of numerous freely accessible sign language videos for ASL also plays an important role in a potential machine-based expansion of the data basis.

In addition to focusing on ASL, this work assumes the validity and observance of Battison's *rules of symmetry* and *rules of dominance* (see Section 2.1.2) and, consequently, the validity of the *Battison Sign Types*. This assumption makes it possible to apply simplifications based on symmetry properties or to supplement missing information in order to increase sign recognition accuracy.

To simplify the sign recognition process, it was decided in the course of this work to forego a complex automatic segmentation of signs. As a result, this work is limited to individual, isolated signs and not to entire sentences or complex sign structures. Thus, the context of the conversation, which may be essential for the meaning of individual signs, cannot be taken into account. The lack of segmentation also excludes signs with more than one morpheme[3] (the smallest meaningful language unit [40]) from the recognition approach.

The focus of this work is on body-worn sensors, so-called *wearables*. The decision to refrain from using cameras takes into account the aspects of mobility and privacy. Camera systems are limited in their mobility because they must capture the signing person in full, which requires a direct face-to-face arrangement. Moreover, in everyday use, cameras raise questions of privacy.

This decision, however, has the consequence that aspects of sign language that are strongly based on body language and facial expressions cannot be captured. These non-manual features signal questions and relative clauses on the one hand; on the

3 More than 90 % of the signs in an adequate vocabulary exhibit more than one morpheme [186].

other hand, adverbs in ASL are usually expressed through facial expression [138, 231]. There are already approaches that can capture facial expressions using electrodes placed on the face [117]; however, these are hardly practical for everyday use and are therefore not pursued further in this work. Nevertheless, this limitation is relativized by the fact that body language and facial expressions are equally understood by sign language users and non-sign language users and therefore do not have to be translated explicitly.

By foregoing cameras, some components of sign language, such as the use of pronouns by pointing at people or objects, cannot be captured. This also affects directional signs, such as SEND[4]. In these verbs, the direction in which they are performed indicates the interaction between the persons or objects involved. Here too, however, it can be argued that pointing at objects or the like is understood by everyone and therefore does not necessarily have to be part of the recognition approach.

In summary, this work focuses on the recognition of individual, isolated signs using a wearable-based approach. While this offers the advantage of increased mobility and privacy protection, it simultaneously entails the limitationthat more complex aspects of sign language that depend on facial expressions, body language, or visual context cannot be captured. However, this is countered by the argument that these aspects can also be understood by non-sign-language users and therefore do not need to be translated. Likewise, a lack of segmentation excludes complex signs from the vocabulary to be recognized.

## 1.5 STRUCTURE OF THIS DISSERTATION

Following this introduction to the topic and the fundamental research objectives of this dissertation, Chapter 2 addresses the background knowledge necessary for understanding this work. This includes specific information on the structure and organization of signs (Section 2.1) and the parametrization of signs (Section 2.2), as well as relevant research work in the field of sign recognition using body-worn sensors (Section 2.3).

Based on these foundations, Chapter 3 presents the concept of our *Parametric Framework for Sign Recognition Using Wearables* (ParaSignRec), which addresses real-time recognition of isolated signs using a parametric approach. It essentially comprises *parameter identification* and briefly outlines the approaches to *parameter recognition* and *meta-classification*, which are presented in detail in the further course of this work.

Beginning with Section 3.1, the focus is placed on identifying the individual parameters. This corresponds to the selection and definition of the parameters and their values that are necessary for a differentiated description of signs. The aim isto find a balanced trade-off between detail and comprehensibility in order to ensure efficient and precise real-time recognition. It is explained how the data basis is selected and supplemented with missing aspects of the parametrization.

Subsequently, in Section 3.2, the hardware requirements are defined and a preliminary selection is made. In addition, the concepts of the respective approaches to

4 `https://www.signingsavvy.com/sign/SEND/4458/1` (Last accessed on January 25, 2026)

real-time recognition are outlined in broad terms, which are elaborated and implemented in Chapter 4 and Chapter 5. Different classifiers and rule-based approaches are investigated and evaluated for each parameter. Depending on the results, the classes are adjusted in order to achieve optimal sign recognition.

The third core component comprises the *meta-classification*, whose concept is presented in Section 3.3 and which constitutes the conclusion of the sign recognition approach. It combines the recognized parameters as outputs of the individual components of our framework using probabilistic models in order to identify the intended sign. The implementation and detailed evaluation of the framework are carried out in Chapter 6.

Finally, a conclusion is drawn in Chapter 7 and an outlook on possible future work is given.

# 2

# BACKGROUND AND RELATED WORK

In this chapter, the background information necessary for understanding this thesis as well as related work are examined in more detail. It is essentially divided into *i)* the linguistic part on sign languages, their structure, and ways to notate them (Section 2.1), *ii)* the technical part on the general process of machine learning and the different classification methods (Section 2.2), and *iii)* the related work that has dealt with the topic of sign language recognition (Section 2.3).

## 2.1 SIGN LANGUAGE

Gestures, especially hand gestures, are found in all cultures, age groups, and fields of activity. Even people who are blind from birth use them instinctively [77]. Gestures are used to express feelings and thoughts, to provide spoken language with context (e.g., by pointing to an object while speaking), or even to replace spoken language entirely (e.g., by using the *thumbs-up* gesture to signal to the interlocutor that everything is fine, or by using sign languages) [71, 139].


Figure 2: *Kendon's continuum* according to McNeill, subdivided into *gesticulation* (blue), *emblems* (green), and *signs* (orange) [71, 139]

According to the gesture researcher Adam Kendon [114], gestures can be divided into three essential classes, which were supplemented by the gesture researcher McNeill [139] and mapped onto *Kendon's continuum* (see Figure 2). The classes are ordered according to how strongly the gestures are bound to the presence of (spoken) language. The classes at the left edge of the continuum are closely tied to spoken language, whereas the classes at the right edge no longer necessarily require spoken language [71, 139, 149]. The original classes of Kendon [114] are:

GESTICULATION or *speech-accompanying gestures* are hand movements accompanying speech whose meanings can be inferred from the context of what is said. They can be subdivided into *singular gestures* and *recurrent gestures*. The former are produced spontaneously and are closely tied to spoken language by extending it (for example, a shape depicted with the index finger that supports what is said). The latter are fixed in meaning and form for speech and cultural communities, but are not as lexicalized as emblems (for example, the gesture of a flat hand with the palm facing upward, used to present ideas or arguments) [71].

In addition, McNeill [139] names the class *pantomime*, which, however, is not further distinguished from *gesticulation* by Kendon [114] and is therefore not considered further here [71, 149].

EMBLEMS are gestures associated with a clear message. Spoken language is no longer strictly necessary here, although in-depth communication is only possible to a limited extent. One example is the VICTORY gesture (see Figure 4a) as an expression of *victory* or *joy* [71].

SIGN LANGUAGES are fully developed languages and replace spoken languages. Instead of the auditory modality, the visually spatial modality is used [71].

Based on the observations of Edward Amherst Ott and Adam Kendon, every gesture can be divided into essential movement segments following a strict hierarchy. Ott identifies the *i)* preparation phase, followed by the *ii)* meaning peak and finally the *iii)* retraction phases. Kendon expands this spectrum by adding a holding phase after the meaning peak and subdivides the preparation phase into rest position and preparation phase, as shown in Figure 3 [71].

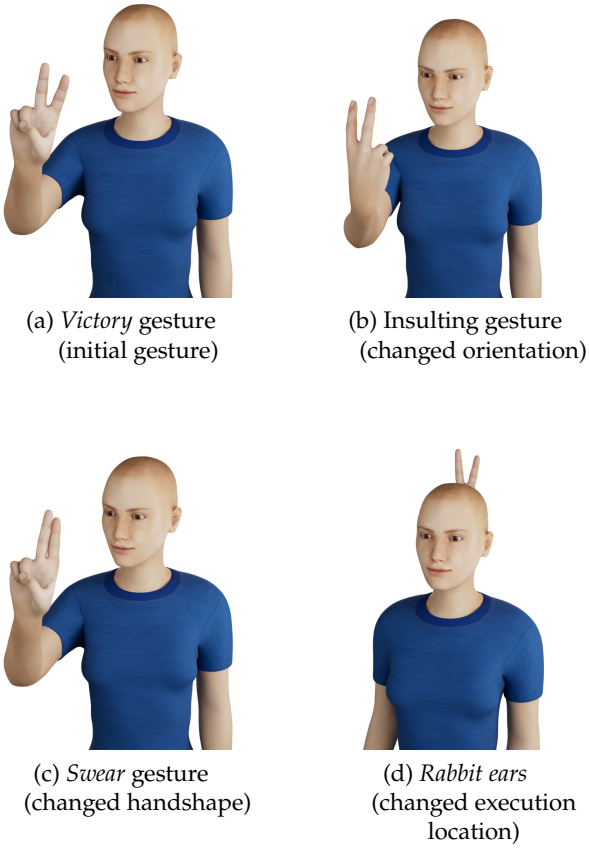
(a) Rest position (b) Preparation (c) Peak (d) Holding phase (e) Retraction

Figure 3: Five phases of a sign according to Kendon [71]

The most complex class of gestures is sign language. Deaf and hard-of-hearing people have the opportunity to communicate with one another through it. Today, around 72 million people worldwide use one of the approximately 200 to 300 different sign languages, of which only 60 have been studied. The most widely used sign language is ASL [58, 223]. The number of people who use ASL as their primary language is difficult to determine, as it has not yet been officially recorded. However, it is assumed that there are about 500.000 people in the United States [144, 196]. Worldwide, about 1 million people [228] speak ASL.

For a long time, sign language was not recognized as a fully fledged language due to the lack of linguistic research. It was said that sign language was less precise, flexible, and subtle than spoken language. It was not until 1960 that the linguist William C. Stokoe [205] was able to demonstrate that ASL is an independent language, since it uses a phonological or sublexical structure. This means that signs in ASL acquire different meanings due to differentiable structural elements (= parameters), as can be seen in Figure 4 [14]. This example shows how the meaning of the initial gesture (VICTORY) changes through the modification of one of its parameters. Rotating the palm from outward (away from the sign language user) to inward (toward the sign

language user), for example, changes the meaning of this gesture in some regions into an insult [71].



(a) *Victory* gesture (initial gesture)

(b) Insulting gesture (changed orientation)

(c) *Swear* gesture (changed handshape)

(d) *Rabbit ears* (changed execution location)

Figure 4: Effects of modified parameters on the meaning of signs (adapted from Fricke and Bressem [71])

As difficult as it is to quantify the number of sign languages and the number of sign language users, the number of signs is equally difficult to quantify. Analyzing different dictionaries of ASL, one arrives at a lexicalized vocabulary of up to 5,000 signs [52, 204, 206, 215, 224]. Here, the fact must not be overlooked that ASL is a living language, i.e., new signs are continually being added.

### 2.1.1 *Structure of Signs*

In linguistics, the following technical terms are used to describe the structure of words:

MORPHEMES are the smallest units of a language that carry meaning. A word consists of at least one morpheme, regardless of whether it is spoken aloud or signed. Words containing more than one morpheme are *morphologically complex* or *multimorphemic*, e.g. the English word "work-er" [11, 39, 187].

PHONEMES are the smallest units of a spoken language that distinguish meaning. They do not have their own meaning, but influence the meaning of what is said. A spoken morpheme consists of phonemes, e.g. the English words "fat" and "father", whose vowels differ only by the phonemes /æ/ and /ɑː/ [40, 67, 138].

GRAPHEMES are the smallest units of a written language that distinguish meaning; that is, to write down a morpheme, at least one grapheme is required. An example is provided by the English words "b-e-e" and "k-e-y", which have the same phoneme /iː/, but different graphemes [38, 67].

In the English language, we find representative examples of the difference between *graphemes* and *phonemes* in the words *i)* "amoeba" [əˈmiː.bə] *ii)* "believe" [bɪˈliːv] *iii)* "Caesar" [ˈsiː.zəʳ] *iv)* "he" [hiː] *v)* "key" [kiː] *vi)* "machine" [məˈʃiːn] *vii)* "people" [ˈpiː.pᵊl] *viii)* "see" [siː] *ix)* "sea" [siː] *x)* "seize" [siːz] , all of which have the same phoneme /iː/, i.e. are identical in pronunciation at this point, but are represented by different graphemes. A counterexample is provided by the English words *i)* "about" [əˈbaʊt] *ii)* "badly" [ˈbæd.li] *iii)* "father" [ˈfɑː.ðəʳ] *iv)* "made" [meɪd] *v)* "many" [ˈmen.i] *vi)* "village" [ˈvɪl.ɪdʒ], which in turn have the identical grapheme *a*, but different phonemes [4, 37, 67, 212].

In summary, a word has a complete meaning and consists of morphemes. Morphemes carry part of the meaning and, in spoken form, consist of phonemes and, in written form, of graphemes. Phonemes influence meaning and can in turn be represented by graphemes. Graphemes do not carry any meaning of their own [138].

This structure can be transferred to sign language, where the sign takes the role of the word. In contrast to morphemes in spoken language, morphemes in sign language do not consist of sounds, but of a variety of meaning-distinguishing parameters that function analogously to phonemes. Examples are handshape or the movement of the hand. William C. Stokoe [205] introduced the term *chereme*, analogous to the term *phoneme*, as the smallest meaning-distinguishing unit of a sign language. This was intended to denote a non-phonetic (non-sound-based) unit of sign language, whereas *phonemes* perform the same function for spoken language. Some linguists ([15, 49]) later adopted this term, but it was never able to establish itself, so that today the term *phonemes* is used as a collective term for both spoken and sign language and is therefore independent of the formal property of the linguistic signal, but rather refers to its structure [15, 71, 138, 147].

A distinction can be made between static and dynamic signs. Static signs contain spatial information, such as the aforementioned handshape, the orientation of the hand, or the location where the hand is executed. Dynamic signs additionally contain temporal information, such as the movement of the hand [170]. Movement can further be subdivided into *internal* (or *local*) movements and (external) *path movements*. Internal movements include, for example, movements of individual fingers, such as changes in flexion or abduction, or rotation of the wrist. External movements refer to a movement in which the entire hand changes position [41, 187, 225].

Some signs differ by only one parameter and are called *minimal pairs* ("minimal pairs") [11]. An example of a minimal pair is the aforementioned parameter changes

of the signs shown in Figure 4, in which, starting from the initial gesture for VICTORY, the orientation of the hand, the handshape, and the execution location of the sign were each changed in order to give the gesture a new meaning. The opposite are allophones, in which a parameter can take various values in order to represent the same sign without having a meaning-distinguishing function [166].

### 2.1.2 *Classification according to Battison*

Robbin Battison [23] created a typology of signs that is generally known as *Battison's sign types* (see Figure 5) [65]:

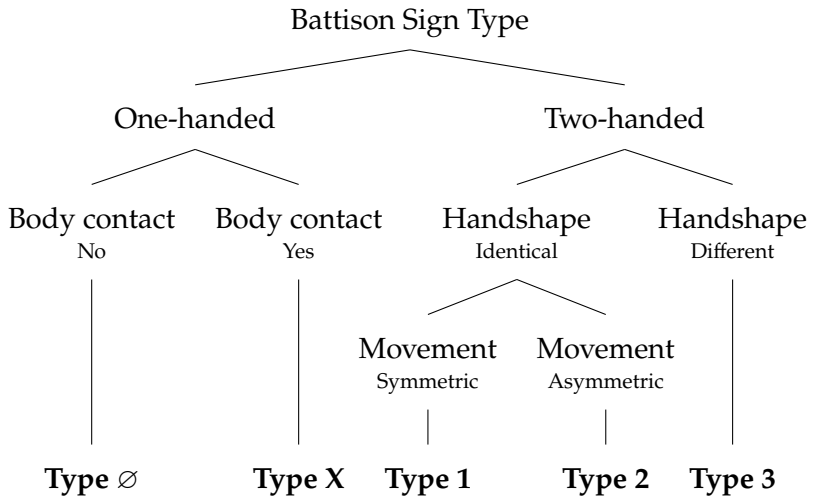



Figure 5: Classification of signs according to Battison (adapted from Battison [23])

TYPE ∅ includes signs that are performed with one hand and do not make contact with the body.

TYPE X includes signs that are performed with one hand and make contact with the body, excluding the other hand.

TYPE 1 includes signs that are performed with both hands simultaneously or alternately.

TYPE 2 includes signs that are performed with both hands, in which the dominant hand is active and the non-dominant hand rests passively, but both hands have the same handshape.

TYPE 3 includes signs that are performed with both hands, in which the dominant hand is active and the non-dominant hand rests passively, and both hands have different handshapes.

TYPE C includes all signs that cannot be clearly assigned to any other type.

(a) Type ∅/X: One-handed signs, e.g. THROW

(b) Type 1: Two active hands with the same handshape and movement (simultaneous or alternating), e.g. MEET

(c) Type 2: Two-handed signs with one active and one passive hand, both with the same handshape, e.g. REMEMBER

(d) Type 3: Two-handed signs with one active and one passive hand, both with different handshapes, e.g. WRITE

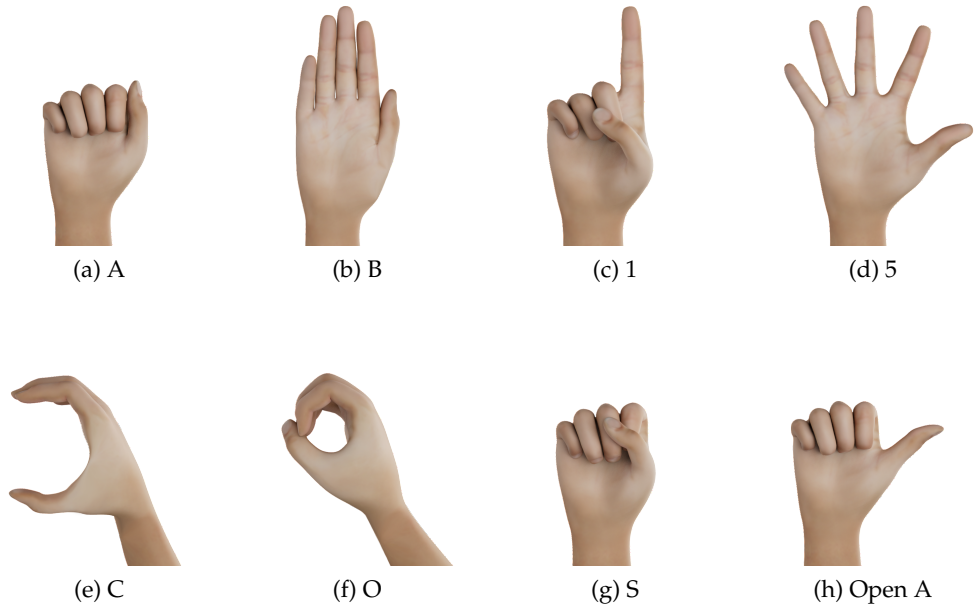
Figure 6: Examples of signs according to Battison's sign type [65]

His typology is based on the thesis that two-handed signs exhibit greater articulatory complexity than one-handed signs from an articulatory perspective, and that signs in which the hands act independently of one another require greater articulatory dexterity than signs with symmetric movement. Even though he originally examined lexical elements of ASL, his statements apply to every sign language. These also include two rules he formulated that are closely related to his typology [30, 65, 187]:

SYMMETRY RULE states that if both hands move simultaneously and independently of each other during articulation, they have the same handshape, the same movement (simultaneous or alternating), and identical or symmetric hand orientation.

DOMINANCE RULE states that if both hands have different handshapes, the dominant (active) hand moves and the non-dominant (passive) hand rests. The non-dominant hand takes on one of the basic handshapes (the *unmarked handshapes*, see Figure 7).

These so-called *unmarked handshapes* are those handshapes used by the non-dominant hand in signs of *Type 3*. They are the handshapes that children learn first and are used in all known sign languages. Depending on the source, four (Figure 7a to 7d, [93]), six (Figure 7a to 7f, [30] or Figure 7b to 7g, [166]), seven (Figure 7a to 7g, [23, 50, 186]) or eight (Figure 7a to 7h, [65]) different handshapes are named. These handshapes represent maximally different geometric basic forms (only A and S are similar) [23].



(a) A (b) B (c) 1 (d) 5

(e) C (f) O (g) S (h) Open A

Figure 7: Unmarked handshapes [23, 30, 65, 93]

Even though they are not explicitly distinguished (unlike A and OPEN A [65]), some sources ([50, 93, 203, 232]) use a different variant of B (compared with [23, 30, 65]). In this variant, the thumb is not extended, but instead rests closed in the palm, as in the letter B of the ASL fingerspelling alphabet[5] (see Figure 8b). Haluts et al. [86] also uses a variant with the thumb extended (see Figure 8c), however in the context of Israeli Sign Language (ISL).

### 2.1.3 *Comparison of Different Sign Languages*

In order to gain a sense of what differences exist between the various sign languages and to what extent the approach described in this thesis can be applied to other sign languages, the differences and similarities of various sign languages will be discussed here. Even though different morphemes are used in different sign languages, there are nevertheless considerable overlaps [214]. As already mentioned in Section 2.1.2, the so-called *unmarked handshapes* constitute a selection of handshapes that are used in most sign languages. However, it may happen that these handshapes stand for

5 The fingerspelling alphabet is used to spell names, proper nouns, and places and consists of 27 signs corresponding to the letters of the alphabet [4].

(a) B (b) Closed B (c) Open B

Figure 8: Different variants of the handshape B

different letters of the fingerspelling alphabet in different sign languages. Thus, the handshape for the ASL letter B corresponds to the letter D in Swedish Sign Language (STS) and to the letter V in Russian Sign Language (RSL). The letter C becomes S in both STS and RSL, and S becomes G in STS (this handshape is not defined in the fingerspelling alphabet of RSL). These are only a few examples, but they show that different languages may indeed work with similar handshapes, yet in some cases have completely different meanings [11]. German Sign Language (DGS), for example, overlaps with British Sign Language (BSL) and Austrian Sign Language (ÖGS), but above all with ASL. ASL derives from French Sign Language (LSF) and therefore shares some similarities with this language [58].

There is no international sign language, even though the languages are increasingly blending. ASL comes closest to an international sign language, as it is considerably more uniform than other sign languages. One example is the sign for SUNDAY: In northern Germany, a flat hand is moved downward over the chest toward the stomach, symbolically brushing over the fine Sunday clothes. In southern Germany, by contrast, the hands are placed together in front of the chest, as in prayer [58]. In ASL, the flat hand with spread fingers is held in front of the body and rotated in a circle, as if washing a window [127]. Nevertheless, there will always be different sign languages and dialects, since these are, as just shown, often shaped by the culture of the speakers.

Let us remain with DGS in order to highlight it as an extensive example. The structure of the signs is identical in form to the signs of ASL; likewise, the rules of symmetry and dominance defined by Battison apply. Here too, the four parameters handshape, orientation of the hand (and contact area), location where the sign is executed, and movement are used, as well as facial expression, mouth gesture, and mouth picture [58, 74]. Differences, however, are found in the classes used. According to different sources, DGS uses 30 handshapes[6] ([74, 166]) or 34 handshapes ([58, 90, 165]). In ASL, by contrast, 36 handshapes (for 4,500 signs [52]), 39 handshapes[7] ([18]), 40 handshapes (for 800 or 1,600 signs, [215, 216]), 60 handshapes (for 2,723 signs [187]), and even up to 150 handshapes [215] are named. These values, however, depend strongly on the vocabulary used. In the fingerspelling alphabet, for example, DGS has 31 different "letters", represented by 22 different handshapes. By comparison, ASL has

6 Depending on the source, either X or BENT L is used, which differ only in the position of the thumb.
7 The handshapes H and U differ only in orientation.

26 “letters” with 21 different handshapes. The discrepancy can largely be explained by the German umlauts; the letters and handshapes used are largely identical and differ only in the sign for T (see Figure 9) [4, 74].

(a) T in ASL (b) T in DGS

Figure 9: Different variants of the handshape T

An overview with the possibility of comparing all handshapes listed here can be found on our homepage `https://sign-parametrization.netlify.app/handshapes`.

## 2.2 PARAMETRIZATION OF SIGNS

The sign recognition approach presented in this thesis is necessarily dependent on the ability to parameterize signs. The existence of this feature has already been demonstrated in Section 2.1; therefore, existing solutions for parameterizing signs are to be presented in the following.

### 2.2.1 *Notation systems*

The possibilities for recording signs are diverse. This can take place in the form of videos; for static signs, photographs or drawings are possible. One can also describe signs in one’s own words or merely note their meaning.

This so-called *glossing* is the English technical term for representing signs by means of words of a spoken language. It makes it possible to describe signs with any desired level of detail, including non-manual features. As a result, the transcribed sign is already available in words of the target language and is easy to understand. The major disadvantage of *glossing*, however, is that there is no standardized method for transcribing signs into glosses. Different transcriptions by different people can lead to inconsistencies, which is particularly problematic in the area of machine processing and translation. *Glossing* is international, since every sign can be described, but still language-dependent, since the transcription is valid only for specific language pairs. For a translation into another spoken language, re-transcription may be necessary [79, 104].

All of these possibilities have in common that they are non-analytical representations of signs. Collections of these recordings cannot be sorted into subsets and analyzed. Notation systems attempt to solve this problem by annotating some (discrete) ele-

ments of the sign through systematic procedures (= *parameterization*), detached from their original linguistic meaning. However, the systems pursue different goals and thus different approaches to parameterization, especially with regard to the type and number of parameters [225].

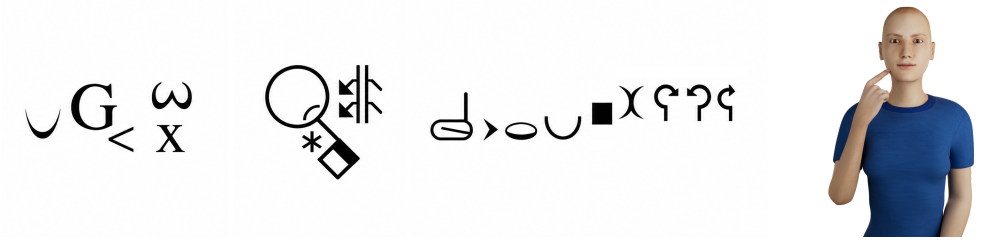
The *parameters* of a notation system are meaning-changing features of a sign and can have different configurations that describe the state of the parameter. Examples of parameters are the handshape or the place of articulation of the sign. Each sign is characterized by a combination of different parameters. To return to the terminology introduced in Section 2.1.1, the state of a parameter constitutes a *phoneme*, since any change alters the meaning of a sign. When the values of a parameter are recorded, this is referred to as a *symbol* within notation systems and as a *grapheme* in linguistics [4, 71, 95, 184]. For both parameters and symbols, different terms exist in the specialist literature. For example, Stokoe et al. [206] uses the term "aspects", Anderson et al., Battison [11, 23] the term "primes", and Johnson and Liddell Johnson and Liddell the term "features" instead of "parameters" ([15, 208]). Supalla et al. [208] uses the term "graphemes" instead of "symbols" ([15, 206]). The term *symbols* originates from their use within writing systems, where this (partly iconic) discrete set of signs enables a readable representation of signs. We therefore also use the term *symbols* in the context of notation systems, but prefer the term *classes* in the technical context.

The first approaches to sign notation were developed by Auguste Bébian [24] in 1825 and George Hutton [105] in 1869, neither of which are considered today. It was not until 1960 that William C. Stokoe's *Stokoe Notation* [205] published a notation system that still serves as the basis for numerous other notation systems today. It was designed with the aim of facilitating linguistic research in the field of sign language and was first published in 1965 in the book "A dictionary of American Sign Language on linguistic principles". In its original form, the notation system had three parameters, namely the *handshape* at the beginning of sign production, the *movement of the hand*and the *place of articulation of the sign*, that is, the (body) location at which the sign is produced. Non-manual features, such as facial expression or gaze direction, are not taken into account, although Stokoe did recognize that these constitute an integral part of a sign but wanted to prioritize the analysis of the basic aspects [138]. Later, these were expanded by Robbin Battison [23] with hand orientation as an additional parameter, from 55 to a total of 64 symbols. This expansion was also followed by other linguists (e.g. Kyle et al. [122]), which is why it can now be regarded as the standard [32, 138, 205, 206].

In contrast, the *SignWriting* system designed by Valerie Sutton in 1974 follows a pictorial and thus easily readable approach to the transcription of signs. For this purpose, *SignWriting* uses five parameters with a comparatively large symbol inventory of 672 pictographic symbols: handshape, movement, place of articulation, orientation, and non-manual features. Movements are divided into start and end positions as well as type of movement. What is special about *SignWriting* is its non-linear writing style from an expressive perspective (= from the sign language user's perspective) or receptive perspective (= from the conversation partner's perspective) and the free arrangement of symbols representing the actual relative positions in space. This en-

ables virtually unlimited combinations and introduces enormous complexity, which is why *SignWriting* can be used for every sign language as well as for one-handed and two-handed signs [79, 121, 138, 209].

The third of the best-known and internationally most widely used notation systems is *Hamburg Sign Language Notation System (HamNoSys)* [184]. It was developed in 1984 at the University of Hamburg and is based on the *Stokoe Notation*; however, like *SignWriting*, it was designed as an international notation system and not exclusively for ASL. HamNoSys aims to strike a compromise between complexity and user-friendliness and uses a total of five parameters and about $200 - 220$ iconic symbols. It describes signs in their initial configuration, consisting of the handshape, hand orientation, place of articulation of the sign, and non-manual features. On this basis, dynamic properties follow, describing how the initial configuration is modified. These include internal hand movements (e.g. finger flexion or spreading) and external hand movements (movement of the entire hand), which are performed in parallel or sequentially. As the final information, symmetry properties are available for two-handed signs. HamNoSys limits the description of non-manual features of a sign to a limited number of symbols in order to keep complexity low [79, 88, 89].



(a) Stokoe [206] (b) SignWriting [210] (c) HamNoSys [87] (d) Sign CANDY

Figure 10: Different notations for the sign CANDY

Another notation system, which will be discussed later, is the *Prosodic Model* developed by Diane Brentari [32] in 1998. It divides the parameters of a sign into inherent and prosodic features. Inherent parameters include, among other things, numerous features of the hand, such as the flexion of selected fingers, the place of articulation of the sign, and non-manual features. Prosodic features include, among other things, the type of movement and the movement path. Hand orientation depends on hand configuration and place of articulation. No symbols are used to describe the classes, as can also be seen in Figure 11. The system can therefore be represented with alphanumeric characters [32, 41, 100]. It is highly compatible with other prominent theories of sign language phonology, for example the *Hand Tier model* and the *Dependency Model* [41].

In addition to these notation systems, there are numerous others, many of which are international and designed for use with ASL, as listed in Table 2. For each notation system, in addition to the year of publication and the supported language, further information is provided on the parameters used, the possibility of machine processing,

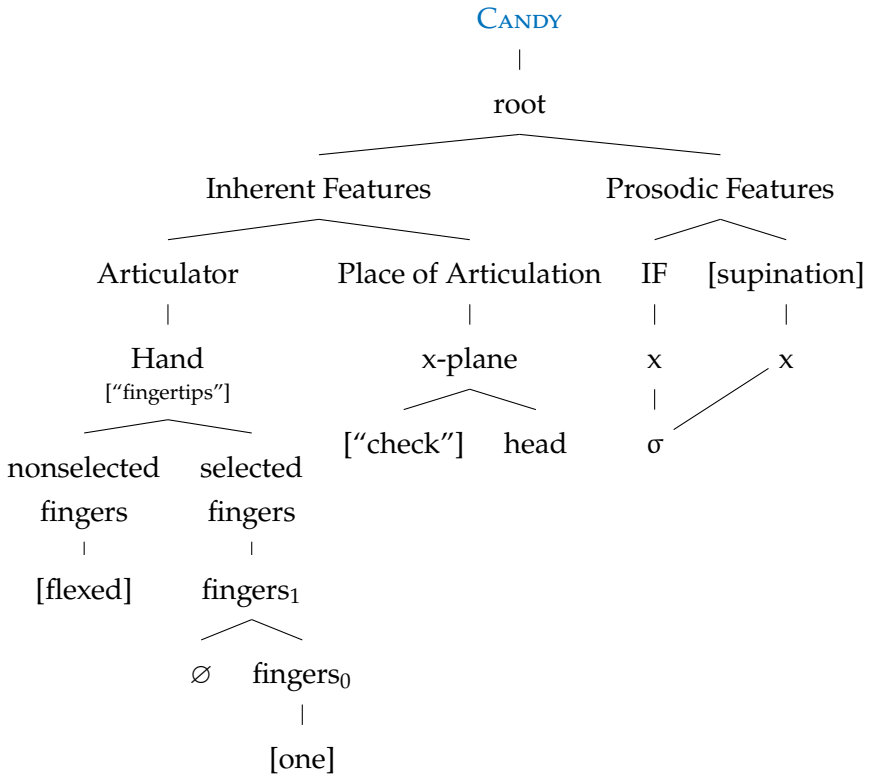



Figure 11: ASL sign CANDY annotated according to the *Prosodic Model* (adapted from Brentari [32])

and the availability of a freely accessible, extensive dictionary. Parameters used by only a minority of the notation systems examined here have been grouped under the column *Other*.

In addition to the supported sign language, differing priorities and goals as well as the current state of research influence the design of the various notation systems. Many of these are primarily aimed at supporting scientific research and analysis, including the *Stokoe Notation*, *HamNoSys*, *SignFont*, and *Sign Language IPA (SLIPA)*, to name only a few. Others, such as *SignWriting*, *si5s*, and *SignScript*, were specifically designed as scripts for public use and address the needs of a broad user base outside the academic context [17, 79]. Educational priorities are also in focus, as for example with the *ASL-phabet*, which aims to teach deaf children the basic principles of alphabet use, or *Signotation*, which is dedicated to promoting bilingual education as well as reading and writing competence among deaf pupils [67, 91, 208]. Personal motives also play a role in the design of notation systems. For example, *ASL Sign Jotting*

8 *Glossing* is not a standardized notation system and describes signs in its own words at any desired level of detail.
9 Subsequently supplemented by Robbin Battison [23] as a separate parameter.
10 Inferences about orientation can be drawn from the contact region and the execution location.
11 It was designed as a cross-linguistic handshape notation system and therefore, in this list of sign notation systems, is to be considered a special case [66].

Table 2: Different notation systems investigated in this work

| | Year | Language | Handshape | Execution location | Movement | Orientation | Non-manual | Other | Machine-readable | Availability |
|---|---|---|---|---|---|---|---|---|---|---|
| Glossing[8] [79, 104] | - | Int. | (✓) | (✓) | (✓) | (✓) | (✓) | (✓) | – | ✓ |
| Stokoe [32, 138, 205, 206] | 1965 | ASL | ✓ | ✓ | ✓ | ✓[9] | – | – | – | (✓) |
| SignWriting [79, 121, 138] | 1974 | Int. | ✓ | ✓ | ✓ | ✓ | ✓ | – | (✓) | (✓) |
| HamNoSys [79, 88, 89] | 1985 | Int. | ✓ | ✓ | ✓ | ✓ | ✓ | – | ✓ | (✓) |
| SignFont [79, 104] | 1987 | ASL | ✓ | ✓ | ✓ | (✓)[10] | ✓ | – | ✓ | – |
| ASL-phabet [67, 208] | 1990s | ASL | ✓ | ✓ | ✓ | – | – | ✓ | – | – |
| Prosodic Model [32, 41, 100] | 1998 | ASL | ✓ | ✓ | ✓ | – | ✓ | – | ✓ | – |
| si5s [15, 17, 49, 79] | 2003 | ASL | ✓ | ✓ | ✓ | ✓ | ✓ | – | – | – |
| SLIPA [17, 79] | 2005 | Int. | ✓ | ✓ | ✓ | – | ✓ | – | ✓ | – |
| PMHC[11] [66, 100] | 2008 | Int. | (✓) | – | – | – | – | – | ✓ | – |
| ASLSJ [17, 104] | 2009 | ASL | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | – |
| SignScript [79] | 2010 | ASL | ✓ | ✓ | ✓ | ✓ | ✓ | – | ✓ | – |
| SLPA [100, 109] | 2010 | Int. | ✓ | ✓ | ✓ | ✓ | ✓ | – | ✓ | – |
| Symbol Font for ASL [17] | 2013 | ASL | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | – |
| Signotation [91] | 2017 | ASL | ✓ | ✓ | ✓ | ✓ | ✓ | – | – | – |

*(ASLSJ)* was designed to enable the author to take notes on signs more quickly [104]. Only *SignWriting* is used today outside sign language research. None of the notation systems has established an accepted scientific standard, such as the International Phonetic Alphabet (IPA) for spoken language [79, 184, 225].

Not every notation system focuses on the notation of complete signs. Thus, out of the *Prosodic Model*, the *Prosodic Model Handshape Coding (PMHC)* specialized in the notation of complex handshapes [66, 100]. It is examined here as representative of a whole range of hand notation systems (including Hulst [103] and Sandler [182]), since the focus of this analysis is on sign language notation systems and the handshape, as their most important parameter ([49]), should not be neglected.

These differing requirements and target groups, which are to be served within the community of sign language users and scholars, are reflected in the selection of parameters and the level of detail of their discrete classes, the symbols: the more parameters and symbols the system uses, the greater the possible level of detail. In doing so, the systems consciously or unconsciously follow the aspiration of *identifiability* and *reproducibility* of the notated signs, which, however, stand in contrast to each other.

If *identifiability* is the focus, as few details of the sign as possible are capturedso that it can still be identified unambiguously. They are therefore more oriented toward people familiar with signs. Examples of this are *glossing* and *ASL-phabet*, which, with only three parameters and 32 symbols, exhibit many ambiguities in sign identification [67, 208]. In the case of *reproducibility*, the opposite is true, i.e. an attempt is made to capture all details of a sign in order to be able to reproduce it later as faithfully as possible, as with *SignWriting* (672 symbols), *HamNoSys* ($200 - 220$ symbols), *SLIPA* or *SignFont* (272 symbols), with the special feature that even the contact region of the hand is identified as a separate parameter. On the scale between these two extremes lie the *Stokoe Notation* (55 to 64 symbols), *ASLSJ*, *si5s* (at least 80 symbols), and probably also *SignScript* (131 symbols) [17, 79, 104, 225].

Another difference between the notation systems relevant to this work lies in the machine readability of the systems: *HamNoSys*, for example, has a linear, one-dimensional writing system with a fixed order of parameters, which have a defined set of symbols (even if the symbols are iconic) and can be easily processed on a computer. *SignFont*, *ASL-phabet*, *ASLSJ*, *SignScript*, and *SLIPA* also have this writing system, even if in the latter case it is not one-dimensional [4, 17, 67, 79, 225]. The opposite of this are *SignWriting* and *si5s*, both non-linear, two-dimensional, iconically arranged notation systems, in which the place of articulation (and in *si5s* also the orientation and movement) of the sign is graphically represented in relation to the drawn sign [17, 225]. *Signotation*, on the other hand, likewise uses iconic symbols to describe the parameters, but assigns them fixed places within a kind of tabular representation [91]. The fact that in these notation systems the informational content lies not only in the symbols themselves but also in their arrangement relative to one another makes machine processing more difficult. By contrast, many of the other notation systems were designed as fonts or have their own one (*SignFont*, *aslsj*, *SignScript*, *Symbol Font for ASL*), which greatly increases machine readability. Others can also be represented with Unicode or American Standard Code for Information Interchange (ASCII) (*SignWriting*, *Prosodic Model SLIPA*, *PMHC*, *ASLSJ*, *SignScript*, *Sign Language Phonetic Annotation (SLPA)*). *ASL-phabet* appears to have its own font, but it can only be viewed online and is therefore not processable [17, 79, 100, 104].

Of interest for this work is an extensive lexical collection of parameterized signs for machine processing. The notation systems presented here provide the prerequisite for this type of dictionary, but exhibit a major lack of availability. Most resources for the notation systems mentioned are no longer available due to their age or do not include a dictionary (e.g. *SignFont*, *si5s*, *SLIPA*, *ASLSJ*, and *SignScript*) [17, 104]. The only example known to us of an extensive dictionary is the book *A dictionary of American Sign Language on linguistic principles* [206], published in 1976 by Stokoe et al., with approximately 3,000 signs on 347 pages [137]. In addition, some resources are available online. For example, there are numerous resources for *SignWriting*, including translated excerpts from the Bible[12] as the largest available download [104]. Individual signs can also be looked up, but we were unable to find a way to export an extensive dictionary for our purposes. For *HamNoSys*, fonts, detailed documentation

12 `http://www.signbank.org/signpuddle2.0/export.php` (Last accessed on January 25, 2026)

including the handshapes used, and an online tool for annotating signs are provided[13], but no parameterizations for download. The *ASL-phabet* site[14] is based on *Adobe Flash*, which has no longer been supported since January 1, 2021 [15]. The documentation for *SLIPA* is available online[16], but here too no dictionary is available. *SLPA* offers tools for annotation [85, 218], but no dictionary could be found here either, nor for the two interdependent notation systems *Prosodic Model* and *PMHC*. *Symbol Font for ASL* offers on its site[17], in addition to detailed documentation and a comprehensive analysis of existing notation systems, some examples of the use of the notation system and an interactive course for learning it. However, the ASL online dictionary has been under construction for 11 years. *Signotation*, as a relatively new notation system, has its own Wordpress blog[18] and is available as a book [91], but without a dictionary (either online or as a printed version). *Glossing* is an exception as a non-standardized notation system. Although there are some available dictionaries or (video) databases with gloss annotations (see Section 2.2.2), these are only conditionally combinable with one another and machine-processable due to the lack of standardization.

### 2.2.2 *Parametric databases*

The preceding chapter examined various notation systems for ASL and for international use in depth. The result of this analysis confirmed that the parameterization of signs is not only feasible, but also represents an important instrument for capturing the complexity and diversity of ASL. This made it possible to show which parameters play a primary role in the description of signs. At the same time, however, it had to be noted that, although most notation systems describe a detailed methodological approach to the parameterization of signs, they generally do not provide a comprehensive data basis in the form of a parameterized sign dictionary.

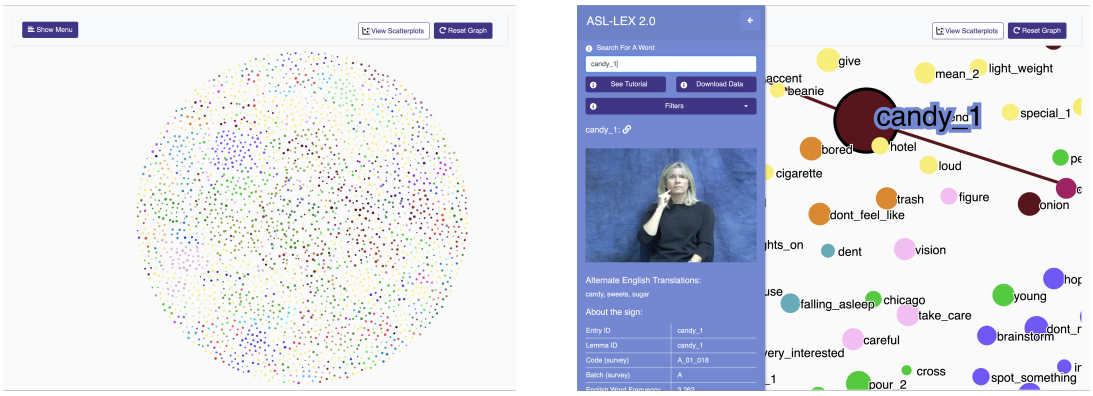
For this reason, the following chapter focuses on examining and evaluating existing parametric databases in the context of ASL. The aim of a parametric database is, similarly to a notation system, to describe each sign by means of a set of parameters so that it can be identified unambiguously. More extensive databases are therefore also based on the use of an already existing notation system [41].

#### ASL-LEX

Caselli et al. [41] recognized the lack of larger, freely accessible lexical databases for ASL and published the first version of the lexical database *ASL-LEX* for signs in ASL in 2016. The publication is intended to promote research in ASL and other sign

13 `https://web.dgs-korpus.de/hamnosys-97.html` (Last accessed on January 25, 2026)
14 `http://www.asl-phabet.com` (Last accessed on February 26, 2024)
15 `https://www.adobe.com/products/flashplayer/end-of-life.html` (Last accessed on February 26, 2024)
16 `https://dedalvs.com/slipa.html` (Last accessed on January 25, 2026)
17 `https://aslfont.github.io/Symbol-Font-For-ASL/` (Last accessed on January 25, 2026)
18 `https://signotation.com` (Last accessed on January 25, 2026)

languages, as the authors see a crucial role for lexical databases in understanding written and spoken languages.




(a) Overview (b) Detail view

Figure 12: *ASL-LEX* 2.0

The two screenshots in Figure 12 show the concrete structure of the platform interface and the forms of representation available there. The signs are visually represented on the platform as nodes (see Figure 12a), whereby the size of the nodes is directly related to the subjective frequency. Signs that exhibit a high phonological similarity are, upon selection, connected by a line and are located in direct proximity, which is highlighted by different colors. The signs may have at most one mismatching feature. By selecting a specific node, all information about this sign can be displayed (see Figure 12b), including numerous phonological properties. Users can filter and download the signs [41, 187].

*ASL-LEX* contains 993 signs from ASL in its original version and was expanded in 2020 by Sehyr et al. [187] in a second version by an additional 1.730 signs to a total of 2.723 signs. The vocabulary now comprises 1.279 nouns, 905 verbs, 300 adjectives, 118 subordinate classes (e.g., conjunctions), 50 adverbs, 42 names, 24 numbers, and five signs for which no lexical class could be determined. It includes a reference video for each sign with a single sign language user in front of a blue background, visible from the waist up. All videos were presented uniformly by the same person, who uses the right hand as the dominant hand. The videos are accompanied by information such as the video length as well as the beginning and end of the sign's meaning peak and further information about the sign itself. The most important lexical information can be divided into the following three categories:

*LEXICAL FREQUENCY* Lexical frequency is used in linguistics to understand language processing, that is, how the mental lexicon is encoded, structured, and processed. Words that are produced with high frequency are perceived and produced more quickly and efficiently than low-frequency words. This principle also applies to sign languages. *ASL-LEX* 2.0 provides average frequency values, their standard deviation, the Z-score, and the number of deaf participants who rated the sign [187].

ICONICITY In sign language, iconicity, or symbolicity, refers to an intended relationship between the form of the sign and its meaning. This relationship may depend on one's own linguistic, cultural, or sensory experience. *ASL-LEX* provides, among other things, a subjective rating of iconicity for individual signs by hearing, sign-naïve participants and, for part of the database, such a metric for deaf ASL sign users. *Sign Transparency*, by contrast, is a measure of the correspondence between form and meaning [186, 187].

PHONOLOGICAL PROPERTIES The phonological properties represent the parametrization approach based on the basic principles of Brentari's *Prosodic Model* [32] (see Section 2.2.1) and are therefore the most important category for this work. In the first version of *ASL-LEX*, six different phonological properties with the highest possible discriminative power are used to encode the first three morphemes of a given sign: *Battison's Sign Type* (cf. 2.1.2), *Major Location*, *Minor Location*, *Selected Fingers*, *Flexion*, and *Movement*. However, these properties could identify only 52 % of the 993 signs unambiguously. Therefore, nine additional properties were added in the second version: *Handshape of Dominant Hand*, *Handshape of Non-Dominant Hand*, *Second Minor Location*, *Repeated Movement*, *Abduction/Spread*, *Abduction/Spread Change*, *Flexion Change*, *Ulnar Rotation*, *Contact between Hand and Major Location*, *Thumb Position*, and *Thumb Contact with Selected Fingers*. These 17 properties and the description of all available morphemes of a sign instead of only the first three increased the unambiguous identification of the signs to about 70 %. All other signs are either homophonous or cannot be sufficiently differentiated by the specified phonological properties [41, 187].

A complete list of the 17 phonological properties together with descriptions as well as an overview of the associated classes can be found in Section A.1.2.

Interested readers can find an interactive version of this lexicon on the project website[19]. The database can be downloaded via the project homepage. However, the videos can only be viewed online and are made available only upon request for research purposes.

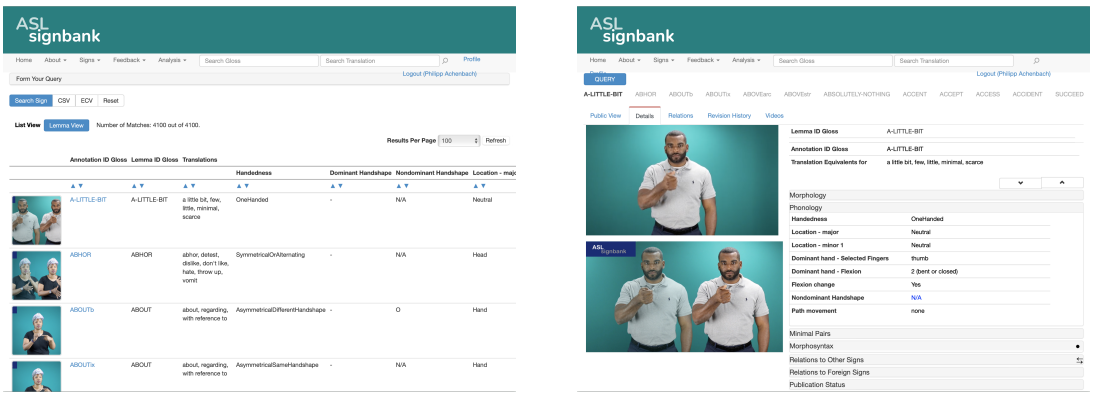
*SignBank*

*SignBank*[20] is a web-based open-source dictionary that was originally developed for Australian Sign Language (Auslan)[21]. It forms the basis for a range of new sign language dictionaries and corpora, including those for BSL[22], Dutch Sign Language (NGT)[23], Finnish Sign Language (FinSL)[24], and now also ASL[25]. In addition to the gloss and the language used (including dialect), example videos, relations to other signs,

19 `https://asl-lex.org` (Last accessed on January 25, 2026)
20 `https://github.com/Signbank` (Last accessed on January 25, 2026)
21 `https://auslan.org.au` (Last accessed on January 25, 2026)
22 `https://bslsignbank.ucl.ac.uk` (Last accessed on January 25, 2026)
23 `https://signbank.cls.ru.nl/` (Last accessed on January 25, 2026)
24 `https://signbank.csc.fi` (Last accessed on January 25, 2026)
25 `https://aslsignbank.haskins.yale.edu/` (Last accessed on January 25, 2026)

keywords, and further information can be stored for each sign, as well as phonological properties [43].

The *ASL SignBank*[25], as the most relevant branch of *SignBank* for this thesis, contains 2.857 publicly accessible signs of ASL and a total of 4.100 signs for registered users. It is part of the *Sign Language Acquisition Annotation Archiving and Sharing (SLAAASh)* project, which aims to systematize annotations in order to provide researchers with a broad basis of sign language corpora. The two screenshots in Figure 13 show the concrete structure of the user interface and the data fields visible there: for each sign, the platform displays the annotation and lemma ID, associated example videos, and possible translations (see Figure 13a). In addition, various phonological properties are described, including information on the handshape of the dominant and non-dominant hand at the start of the production, the location of the sign at the beginning and end of the production, and the type of movement, as exemplified in the detail view in Figure 13b. A complete list can be found in Section A.1.1. Since the phonological coding does not provide a complete phonological description of the signs, different entries within *ASL SignBank* may be coded identically [96, 98, 99].




(a) Overview

(b) Detail view

Figure 13: ASL SignBank

The lemmas used in *ASL SignBank* correspond as closely as possible to the lemmas of *ASL-LEX* in order to ensure compatibility between these two databases. Likewise, the phonological coding follows that of *ASL-LEX* and is thus also oriented toward the *Prosodic Model* by Brentari [32]. At present, at least 73 % of the *ASL-LEX* signs (1,989 entries) are included in SignBank. Signs are matched to one another if they share both meaning and form, although the latter may vary slightly (e.g., when a symmetrical sign is produced one-handed instead of two-handed) [96, 187]. Owing to its high compatibility with *ASL-LEX*, it is theoretically a sensible complement to it; however, it offers no download options[26] and only a limited range of phonological properties.

26 Through email correspondence with Hochgesang [97], it was confirmed that a download option is indeed planned, but that it is not functional for the time being.

*Word-Level American Sign Language (WLASL)*

WLASL is not, strictly speaking, a parametric database, but rather, according to Li et al. [126], the ("probably") largest database of sign videos for ASL. It consists of several datasets with up to 2,000 signs and 21,083 videos. The videos come from 20 different sources and were recorded by 119 different sign language users. Each video contains a single sign, recorded with at least three different sign language users from a (nearly) frontal view with different backgrounds. Each sign can have up to three variations, which may differ in their parameters, and at least five videos exist for each. WLASL is based, among other things, on the annotations of the first version of *ASL-LEX* and therefore also includes 814 of the 993 videos (approximately 82 %). In addition to the videos, the authors also provide pretrained ML models of a *3D Convolutional Network* video classifier, with which signs from the dataset containing 2,000 signs can be recognized with a moderate Top-5 accuracy of 57.31 % [126].

Neidle and Ballard [155] criticize an imprecise assignment of signs from different sources due to inconsistent glossing conventions (see Section 2.2.1). The sign videos were matched to one another on the basis of the glosses, which can lead to *i)* the same sign occurring under different glosses (e.g., WOMAN and LADY) and *ii)* different signs with different meanings appearing under the same gloss (e.g., CLOSE in the sense of "to close something" and "near") . The latter type of signs are differentiated through the use of different variations. This flawed approach also evidently leads to the comparatively poor recognition rates that Li et al. [126] achieved with their classification approach on the dataset with 2,000 signs. Through a revised annotation [156] of 19,672 videos based on the glossing conventions of the American Sign Language Linguistic Research Project (ASLLRP) Sign Bank, not only can added value be created through consistent annotation, but a bridge can also be built to ASLLRP-compliant datasets. Using 18,141 sign videos (1,449 signs with at least six occurrences), the recognition approach can thus be improved to a Top-5 accuracy of up to 94.64 % [126, 155, 157].

A few videos were not considered in the revised annotation. This was due to video quality, the hands not being visible in the video, poor quality in the performance of the sign, the presence of multiple signs within a video, and more. This too can increase the quality of the WLASL dataset [156].

*American Sign Language Linguistic Research Project (ASLLRP)*

In addition to the improved annotations of the WLASL videos just mentioned, the ASLLRP offers further datasets on ASL signs that follow its own glossing conventions. In total, 3,542 different signs are available with 44,012 sign videos in all, of which 23,452 are sign videos of isolated signs from 33 different sign language users. The data are distributed across three datasets *i)* Boston University American Sign Language Lexicon Video Dataset (ASLLVD) with 9,748 sign videos and 6 sign language users, *ii)* Rochester Institute of Technology (RIT) 11,801 sign videos and 12 sign language users, *iii)* DawnSignPress (DSP) with 1,903 sign videos and 15 sign language users. In addition to the gloss, further information is provided for each sign, including the

handshapes of the dominant and non-dominant hand at the beginning and end of the sign. By merging these data with those of WLASL, 1,480 signs with a total of 22,853 sign videos could be obtained, provided that at least six sign videos are available for each sign [157, 158].

## 2.3 SIGN RECOGNITION USING BODY-WORN SENSORS

For gesture and thus also sign recognition, the essential requirements are the capture of the performed gesture, a (pre-)processing of the recorded information, and subsequently the interpretation thereof. Figure 14 shows a schematic representation of the classification process for recognizing the parameter *hand shape*, as published by us previously [5, 7]. This chapter is intended to shed light on the background and possibilities for this process, as well as the possibility of combining the results of multiple processes via *meta-classification*.

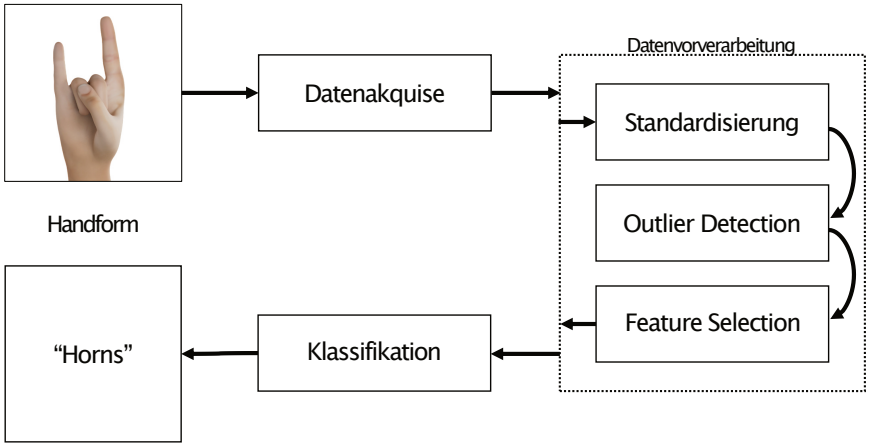



Figure 14: Schematic structure of an Machine Learning classification process for detecting hand shapes (adapted from [5])

As part of this, an extensive review of related publications on gesture recognition was conducted. A comparison of these publications is provided in Section A.2, grouped by differing complexity of the gestures captured. Table 3 gives an overview of this comparison with the number of recognizable gestures and the parameters considered per class. It can be stated in advance that the complexity of the gestures ranges from simple gestures, which are predominantly characterized by hand shape, to extensive signs, such as those of ASL with all their facets. Across all classes, the maximum number of recognizable gestures does not exceed 100, which was already criticized in the context of defining the challenge in Section 1.2 as too small a spectrum for a conversation. With an increasing number of parameters, the scope of the gestures also increases, both in complexity and in number. This depends particularly on the choice and complexity of the acquisition setup, which shapes the essential part of *data acquisition*.

Table 3: Overview of related work, grouped by gesture classes

| Gesture class | N | Gestures | ∅ | Hands | HF | Ori. | Bew. | Loc. | Sources |
|---|---|---|---|---|---|---|---|---|---|
| *Simple gestures* | 40 | 3-56 | 13,90 | 1 | ✓ | (✓) | (✓) | – | Table A.5 |
| *Finger alphabet* | 16 | 7-46 | 17,63 | 1 | ✓ | ✓ | (✓) | – | Table A.6 |
| *Simple signs* | 32 | 2-80 | 23,81 | 1-2 | ✓ | ✓ | ✓ | – | Table A.7 |
| *Extensive signs* | 5 | 50-100 | 78,00 | 1-2 | ✓ | ✓ | ✓ | ✓ | Table A.8 |
| | 93 | 2-100 | 21,40 | 1-2 | | | | | |

N = number of investigated works; HF = Handshape; Ori. = Orientation; Bew. = Movement; Loc. = Location of articulation of the sign;

### 2.3.1 *Data Acquisition*

For the capture and recording of hand gestures and signs, suitable hardware is required. As already discussed in Chapter 1, a camera-based approach was ruled out for various reasons, and the focus was placed on a mobile approach using body-worn sensors. Therefore, this chapter examines different approaches that meet these requirements.

#### *Tracking through motion*

Mechanical tracking methods capture movements of body parts using sensors attached to clothing or directly to the body [71].

A widely used example is the data glove. The related works examined in this thesis predominantly rely on self-developed data gloves, which are tailored to the respective project in varying degrees of complexity [1, 20, 33, 35, 36, 44, 47, 48, 63, 64, 92, 107, 111, 123, 134, 148, 152, 153, 170, 179, 181, 191, 195, 198, 229, 246]. In addition, many commercially available data gloves are used, such as the *5DT* [125], the *CyberGlove* [145, 161], the *DG5-VHand* [219], the *Senso Glove: DK2* [6, 7], or the *Manus Prime X* [5].

The data gloves differ mainly in the type and number of sensors used, as well as in the design and construction of the glove itself as the carrier of the sensors.

To capture fine finger movements as well as movements of the entire hand, IMUs are used in most data gloves, i.e., combinations of accelerometers and gyroscopes, and in some cases additionally a magnetometer [3, 6, 7, 82, 132, 141, 143, 152]. These are usually attached to the back of the hand to capture hand movements and orientation. To record information about individual finger movements, additional IMUs are attached to the fingers. However, this can restrict the finger's freedom of movement and, with prolonged use, lead to fatigue due to the additional weight [153]. As an alternative, the use of *flex sensors* (bending sensors) is therefore found [20, 48, 102, 111, 124, 153, 161, 176, 197, 207, 246], which are sometimes attached across the joints or along the entire finger. Changes in resistance of the piezoresistive material during flexion or extension make it possible to capture the movements. As a low-cost substitute for bending sensors, *tilt sensors* can also be used [195]. In addition to finger

movement, contact points between the fingers or with the hand can also be measured using capacitive touch sensors [1, 33].

Most data gloves use a combination of different sensors to capture hand shape and movement. A common combination is the use of bending sensors to determine hand shape and one or more IMUs to determine orientation and movement [33, 35, 36, 44, 60, 92, 107, 113, 134, 136, 148, 170, 175, 178, 179, 181, 191], sometimes supplemented by individual gyroscopes [106, 172, 239] or accelerometers [172, 195, 219, 239]. In a few cases, pressure or touch sensors are also added to detect contact between fingers, as is the case, for example, in distinguishing the signs of the finger alphabet for R, U, and V (see Figure 15) [47, 64, 188, 198].

(a) R in ASL (b) U in ASL (c) V in ASL

Figure 15: The handshapes R, U, and V in ASL differ only in the spread between the ring finger and the middle finger.

To obtain a more precise indication of position, some works also use position trackers such as the magnetic tracking system *Flock-of-Birds* [145, 161, 170, 177]. With these, it is possible to capture the three-dimensional position and orientation of the hand relative to the device's position, enabling these approaches to recognize even *extensive signs* (Table A.8). However, due to the system's design, mobile use of the approach is excluded.

The simplest setup of a data glove consists of fixing the sensors directly at the corresponding locations, such as the fingers and the back of the hand, using rubber bands or plastic rings [82, 123, 134]. In addition, it is possible to attach the sensors to or into gloves made of cotton or neoprene [12, 36, 44, 47, 63, 64, 73, 92, 102, 107, 111, 148, 152, 153, 170, 179, 181, 191, 198, 229]. This offers greater comfort when wearing, putting on, and taking off the data glove, and furthermore fewer restrictions on movement when performing gestures [44].

Another aspect is the connectivity of the data gloves. Some devices support onboard processing and classification of the data, i.e., directly on a microcontroller attached to the device [3, 12, 20, 107, 152, 191]. Alternatively, the data can be processed on an external device. In the studies examined, this was connected either wirelessly via Bluetooth or WiFi [19, 35, 48, 63, 68, 82, 106, 113, 123, 124, 170, 176, 179, 181, 246] or via cable [33, 44, 73, 111, 134, 195, 198].

Both the wired solution and additional hardware on the glove through the attachment of a microcontroller reduce wearing comfort and restrict freedom of movement. Likewise, factors in most prototype implementations, including exposed wiring or

bulky components, affect wearing comfort. Consequently, comparatively few of the examined approaches are free of these limitations [63, 92, 170, 179, 229, 246]. Different hand sizes also affect capture and complicate the development of universally usable hardware [12]. By minimizing these impairments, freedom of movement can be increased so that even complex gestures can be performed at high speed [19]. Commercial data gloves feature a compact design and high wearing comfort due to industrial manufacturing [6, 7, 145, 161, 219]. This explains why, in the studies examined, commercial data gloves were predominantly used for more complex gestures, whereas self-developed data gloves were used for less complex gestures.

However, this is reflected in a significantly higher price compared with self-developed solutions. For example, the prices of the current *5DT*[27] data glove are $5,495, that of the *CyberGlove*[28] is $13,750, the successor to the *Senso Glove: DK2*[29] costs $999 , and the *Manus Metagloves Pro*[30] as the company Meta's current data glove costs 5,677.88 $. In contrast, inexpensive self-developed solutions are available for around 30 to 50 [$] and could be reduced to $9 in serial production [188, 246].

The data gloves presented so far were either designed for specific research projects or have a broad range of applications and can be used in different domains, such as AR/VR or HCI. In addition, however, there are also data gloves that were developed specifically for sign recognition. Due to their (mostly) commercial background, little information about the technology used can be found, but for the sake of completeness they are still presented here:

SIGNALOUD was developed in 2016 by two students at the University of Washington and consists of two data gloves capable of recognizing individual words and phrases in ASL and outputting them through a speaker. The position of the hand and its movements are captured by motion sensors. The data are then sent to a computer via Bluetooth and processed there. Unfortunately, there is no information available about the supported vocabulary. The students were able to win a prize endowed with $10,000 [31], but at the same time they also drew considerable criticism from the Deaf community, as such a product does not address the full requirements of (American) Sign Language and a word-for-word translation likewise leads to an inaccurate translation [118].

BRIGHTSIGN was developed as a relatively low-cost communication option between sign language users and non-sign-language users. In contrast to *SignAloud*, *BrightSign* does not have a fixed vocabulary; instead, users record their own sign vocabulary so as not to be limited to one or a few sign languages. The translations can be displayed on a screen on the wrist, output via speaker, or sent to a smartphone. The accuracy of the data glove is given as 97 %. The

27 `https://5dt.com/5dt-data-glove-ultra/` (Last accessed on January 25, 2026)
28 `https://metamotion.com/hardware/Cyberglove-Sale.html` (Last accessed on January 25, 2026)
29 `https://senso.me` (Last accessed on January 25, 2026)
30 `https://www.knoxlabs.com/products/manus-metagloves-pro` (Last accessed on January 25, 2026)
31 `https://www.nbcnews.com/feature/n577636` (Last accessed on January 25, 2026)

data glove has been commercially available since 2018 for a price of just under 2,700 $[32] and was preordered approximately 700 times [118].

ACCELEGLOVE was developed as early as 2002 by Hernandez-Rebollar et al. [94] for ASL. However, the data glove is only capable of translating individual letters of the finger alphabet, not extensive signs. AcceleGlove also drew criticism from the Deaf community for the same reasons as *SignAloud* [159].

*Tracking through muscle activity*

In addition to the possibility of capturing and evaluating the movements of the hand and fingers, the activity of the muscles can also be used to capture gestures. Differences in the exertion of individual muscle groups make it possible to infer the shape of the hand and fingers. One advantage of this approach is that the respective sensors are attached to the arms, leaving the hands free of sensors and allowing for the greatest possible freedom of movement. Compared with data gloves, these sensors can be worn under a shirt and are thus hardly noticeable in public, which leads to greater acceptance [167].

There are three different approaches to capturing muscle activity:

ELECTROMYOGRAPHY (EMG) is used to measure myoelectric signals. These electrical signals are generated during muscle activity by changes in potential across the muscle membrane [150]. In *Intramuscular Electromyography (iEMG)*, the electrodes are inserted through the skin and placed invasively directly between the muscle tissue, whereas in Surface Electromyography (sEMG) the electrodes are placed non-invasively on the skin surface. Although this measurement method is less precise, it offers much greater accessibility and practicality. Among the related works examined as part of this dissertation, the electrodes were attached intramuscularly in only one case ([199]). Most works used electrodes integrated into a wristband, either with a commercially available solution [2, 13, 53, 62, 167] or a self-developed solution [10, 108, 200, 238], or they were placed directly on the arm [25–27, 130, 140, 183, 190, 199, 235, 236]. Placement directly on the arm has the disadvantage that accurate results require exact positioning depending on the participant [10]. Data quality can likewise be improved if conductivity between the arm and the electrodes is increased through measures such as shaving and disinfection, or the use of a special gel, which in turn can cause skin irritation in some participants [27, 108, 130, 140, 150, 190]. Disturbances and measurement inaccuracies can be caused by external influences, such as electrodes slipping, participant sweating, or interference signals. The partly overlapping muscle groups also influence one another [10, 26]. To minimize these disturbances, reference electrodes are often attached to a non-active extremity such as the elbow [130, 150].

32 https://www.forbes.com/sites/kittyknowles/2018/04/30/brightsign-incredible-sign-to-speech-smart-glove-startup-is-raising-1-million/ (Last accessed on January 25, 2026)

Looking at the works examined, it becomes apparent that purely EMG-based approaches are increasingly used for less complex gestures. Therefore, an IMU is often added in order to obtain additional information about hand movement and to capture more complex gestures as well [2, 108, 167, 235, 236].

MECHANOMYOGRAPHY (MMG) is another method for measuring muscle activity. In comparison with EMG, however, no electrical signals are measured here, but rather physical changes in the muscles during muscle contraction. These include accelerations, vibrations, and acoustic sounds, which are captured, among other things, with accelerometers, microphones, piezoelectric contact sensors, and displacement sensors. A key advantage of MMG is the high Signal-to-Noise Ratio (SNR) combined with low sensitivity to sensor placement, as well as low cost. One disadvantage is the still relatively low state of research, due to the novelty of this technology. This also means that commercially available MMG sensors are not yet available [70, 150, 213].

ELECTRICAL IMPEDANCE TOMOGRAPHY (EIT) represents another possibility to capture hand gestures. Zhang and Harrison [243] developed an arm controller consisting of paired electrodes that determines the internal impedance geometry of an arm by measuring the individual cross-sectional impedances. By injecting a small amount of current between two opposing electrodes, the electrical properties of the tissue in between are determined and the resulting voltages are measured. A tomographic image can then be generated from these measurements, which in turn provides information about the gestures performed. One advantage of the technology is the low-cost, compact, and energy-efficient electrodes that can be integrated into a wearable wristband [116, 243]. However, EIT is still not widely used in the field of gesture recognition.

In summary, the approaches presented here for gesture capture using muscle activity show great potential, but usually require elaborate and controlled data collection, for example through specific sensor placement. Against this backdrop, methods that enable the transfer of already existing data across different domains come into focus, as considered in the following section.

*Domain Transfer*

Recording data with human participants entails several difficulties. Depending on the sensors and recording devices used, on-site recording is indispensable in most cases, which increases the burden on participants. Likewise, the choice of hardware also affects the amount of time required to prepare the recording, for example through the complex attachment of sensors. Since the COVID-19 pandemic at the latest, hygienic factors have played an additional role, especially when using body-worn sensors and direct physical contact between people, such as between a recording supervisor and a participant.

Many researchers therefore make the data they collect publicly available in order to facilitate data acquisition for other researchers. Due to the highly specific sensor

technology used in sign recognition with body-worn sensors, a major lack of openly accessible databases in this area is noticeable. Only in the field of sign-language video databases are a large number of extensive collections available, some of which were already presented in Section 2.2.2 [131].

To address the shortage of data captured with specific hardware, a so-called *domain transfer* can be used, in which data from one domain (e.g., videos) are transferred into another domain (e.g., body-worn sensors). Liu et al. [131] present an approach in which information on finger and hand positions is extracted from openly accessible sign videos and converted into synthetic IMU data. In addition, various *data augmentation* techniques are presented that generate synthetic variations of the training data. During the training of the ML classifiers, these lead to greater generalizability and robustness with respect to user diversity.

An approach for the real-time extraction of information relevant to gesture recognition from body parts is provided by *MediaPipe* from Google [78, 133], which offers different solutions for extracting information from video material.

With the *Hand Landmark Detection*[33], 21 different 3D coordinates can be identified for each hand, even if they are partially occluded. Depending on the requirements, the coordinates are available either relative to the center of the hand or as absolute image coordinates. They enable detailed analysis of hand positions, movements, and orientations. The detection of individual palms is reported to have an accuracy of up to 95.70 %. The framework was trained with an estimated 30,000 real hand photos as well as several synthetic hand models [174, 242].

For capturing the entire body instead of a single hand, the *Pose Landmark Detection*[34] is available. It can be used to identify 33 3D coordinates of the entire body, including four coordinates for each hand and three coordinates for each arm, consisting of shoulder, elbow, and wrist. Here too, the coordinates are given either relative to the hip center or in absolute image coordinates.

The MediaPipe *Holistic*[35] framework integrates the two models just mentioned, as well as an additional model for face detection, into a multi-stage pipeline. In this way, the respective specialized components complement one another, yielding 543 landmarks of the hand, body, and face [81].

### 2.3.2 *Data Preprocessing*

The recorded sensor signals must be prepared for the classification process in the next step. Through *data preprocessing*, important information in the available data should be emphasized while redundant or misleading data are removed [172].

33 `https://developers.google.com/mediapipe/solutions/vision/hand_landmarker` (Last accessed on January 25, 2026)

34 `https://developers.google.com/mediapipe/solutions/vision/pose_landmarker` (Last accessed on January 25, 2026)

35 `https://developers.google.com/mediapipe/solutions/vision/holistic_landmarker` (Last accessed on January 25, 2026)

*Standardization*

A common and usually first performed data preprocessing step is the standardization of the data, i.e., scaling all data to a specified interval. For this purpose, statistical properties of the training data can also be used, such as subtracting the mean value $\mu$ for each feature $i$ of a dataset $X$ and subsequently dividing by the standard deviation $\sigma$:

$$Z_i = \frac{X_i - \mu_i}{\sigma_i} \tag{1}$$

Thus, all features follow a normal distribution with mean zero and unit variance. This procedure is recommended when using Support Vector Machines (SVMs) and linear models, as it ensures that the different features are on a uniform scale and are comparable. This leads to improved accuracy and efficiency of model training [5, 8, 168].

A frequently used alternative ([6, 7, 12, 27, 102, 123, 172, 183, 191]) is *min-max normalization*, in which the values $x$ of a feature $i$ are scaled to a fixed range, usually $[0, 1]$ or, as in the following example, $[-1, 1]$:

$$f_{\text{scaling}} = 2\frac{x - x_{\min}}{x_{\max} - x_{\min}} - 1 \tag{2}$$

Disturbing artifacts in the recordings can be removed using various filtering methods, whereby the studies examined mainly used band-pass ([10, 13, 190]) or low-pass filters [25, 63, 73].

When working with dynamic data from different sources, it is often necessary to synchronize them first. This can be useful when different sensors with different sampling rates are available or when there is a temporal offset [236]. At a particularly high sampling rate, it can also be useful to reduce the number of data points, since neighboring data points often contain similar information [219].

*Outlier Detection*

The goal of *Outlier Detection* is to cleanse the recorded data of outliers, i.e., data points that differ significantly from the other recorded data points. The cause of outliers can be various; for example, technical factors (e.g., sensor drift) can negatively affect the data, or they may already be performed incorrectly by the subject during acquisition. In both cases, outliers have a negative impact on the performance of ML classification, since the occurrence of outliers promotes so-called *overfitting*. In this case, the algorithm learns the data too precisely (including any errors) and thus loses the ability to generalize sufficiently accurately to new, unknown data. The opposite is *underfitting*, which corresponds to *overgeneralization*. In this case, the algorithm is not able to adequately capture and differentiate the underlying structure of the data [75].

Many approaches to *Outlier Detection* use clustering algorithms to group data points into clusters. Data points that do not belong to any cluster can thus be identified as outliers. One such approach is *Density-Based Spatial Clustering of Applications with*

*Noise (DBSCAN)*[36]. Starting with a random data point, additional data points within a fixed distance ε are searched for. The number of found data points must exceed a certain number of *minPoints* so that the original data point is marked as a *Core* point and all surrounding data points are grouped into a cluster (see Figure 16). This step is repeated for all data points in random order. Data points that have at least one Core point within the radius ε are *Border* points. All remaining points could not be assigned to any cluster and are *Noise* points, i.e., outliers.

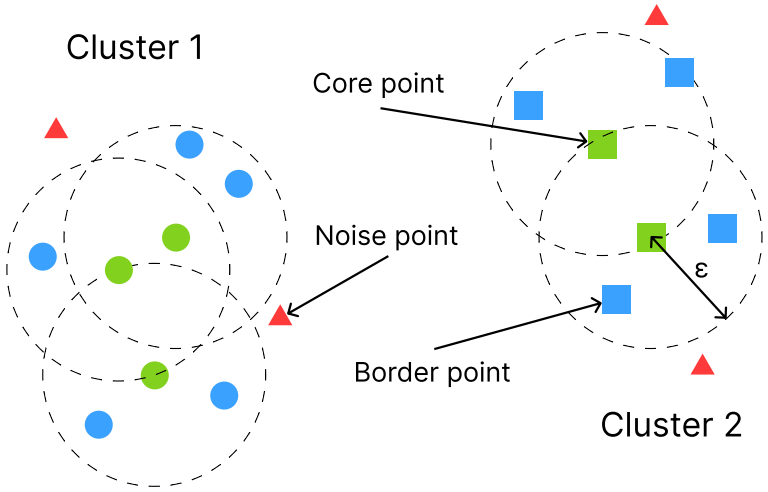



Figure 16: DBSCAN[36] categorizes data points into Core, Border, and Noise points

*Outlier Detection* is rarely used in other related works. Most often, outliers are removed from the test set, and a warning is issued directly during data acquisition in addition to the model's prediction ([163, 245]). This can prevent incorrect predictions. This is of particular relevance in fields such as medicine, where false predictions can lead to serious consequences. To improve data quality, outliers can also be identified in the training data in addition to the test data. However, the remaining data should then be standardized again according to the principle already described [5].

*Feature Selection*

To classify objects, a wide range of information about the objects to be classified is collected, for example using sensors. This information is also referred to as features. For the classification approach, some of these features may be particularly relevant, while others are redundant, noisy, or irrelevant. To clean up this data and improve the performance of the classification approach, *Feature Selection* is used. This is achieved by having *Feature Selection* select a subset of the features from the original dataset on the basis of redundancy and relevance, or by reducing the entire feature space in its dimension through mathematical transformation and thereby also reducing the amount of data [22, 183, 226]. Using handshape recognition as an example, the joint sequence relevant to this argument can be directly traced from the anatomy of the

36 `https://www.kdnuggets.com/2020/04/dbscan-clustering-algorithm-machine-learning.html` (Last accessed on January 25, 2026)

human hand shown in Figure 17. The figure shows the phalanges of the fingers as well as the successive joints *Metacarpophalangeal (MCP)*, *Proximal Interphalangeal (PIP)*, and *Distal Interphalangeal (DIP)* along the fingers. From precisely this serial arrangement visible in the figure, it can be inferred that the motion of the *DIP* joint is determined by that of the *PIP* joint and therefore cannot be moved independently. Consequently, only one of the two features contains important information about the handshape being performed [164].

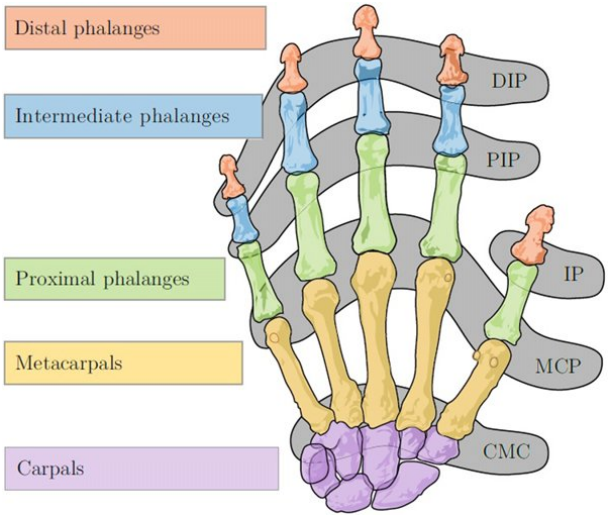



Figure 17: Joints and bones of the human hand [69]

For dynamic data, descriptive standard statistics are often used as features, including the mean or mean absolute deviation (*Mean Absolute Deviation (MAD)*), variance and standard deviation, minimum and maximum values, the root mean square (*Root Mean Square (RMS)*), or kurtosis. From the frequency domain, for example, peak and mean values of the power spectral density can serve as features, as well as other features that can further describe the characteristics of a signal (signal energy, area under the curve (*Area under the Curve*), and zero crossing (*Zero Crossing*)) or, specifically for EMG signals, the waveform length (*Waveform Length*) [35, 64, 161, 198, 200].

To obtain a suitable selection of features, there are various approaches to *Feature Selection*. Some methods examine the individual features for their information content and select only those with the highest information content for classification [236]. Other methods, such as Principal Component Analysis (PCA), are used to combine features from strongly correlated sensors [33, 172]. Genetic Algorithm (GA) is another algorithm used in the literature ([5, 125]) and is loosely based on evolutionary theory. The algorithm starts with a large number of randomly generated bit strings, the length of which is identical to the number of features. Each bit indicates whether the corresponding feature is discarded (= 0) or retained (= 1). Its mode of operation can be described in four phases. *i)* Models of a classifier with the selected features are now trained for each bit string and the results are evaluated (= *evaluation*). *ii)* The algorithm then selects some of these models for the next phase, with models of high accuracy being preferred (= *Selection*). *iii)* New feature combinations are formed from

the selected bit strings (= *Recombination*). *iv)* Each bit is now mutated with a low probability, i.e., individual bits are inverted (= *Mutation*). After the mutation, the evaluation phase begins again. The entire process is repeated until either a predefined number of iterations has been reached or the accuracy has not improved significantly over several iterations [230].

### 2.3.3 *Classification*

Subsequent to the *data preprocessing*, the actual *classification* of the data can be performed, i.e., the collected and processed data are now assigned to discrete classes. A key aspect in evaluating this classification lies in the generalizability of the model, i.e., a model's ability to reliably classify previously unseen data. In this work, we distinguish between *Within-User* and *Cross-User* approaches. The former use data from the same participants for training and testing, which often leads to higher accuracies with a less generalized approach. A popular example is *k-fold cross-validation* [76], in which participant data are mixed and randomly divided into $k$ equally sized groups. In turn, the model is trained with $k-1$ groups and the trained model is validated with the remaining group. This process is repeated $k$ times and an average accuracy across all runs is determined. In *Cross-User* approaches, by contrast, strict separation of participant data into training and test data is observed, which is often closer to the real application scenario. One example of this is Leave-One-Subject-Out Cross-Validation (LOSOCV). In this approach, the data of a single participant are used exclusively for testing and the data of all other participants for training. This procedure is repeated until each participant dataset has served once as the test dataset. Afterwards, the average of all probabilities is calculated here as well. This enables a result interpretation that is as representative and user-independent as possible [76]. In the following, some ML approaches used within the studies examined are presented.

#### *Support Vector Machine (SVM)*

The most widely used approach for solving the classification problem among the studies examined is the Support Vector Machine (SVM) [2, 5–7, 12, 25, 26, 44, 47, 64, 123, 132, 140, 148, 170, 172, 235, 236, 243]. With its help, data can be divided into two different classes. The data are mapped into a vector space and subsequently a linear hyperplane is sought that is capable of separating the data as effectively as possible according to the *max-margin* paradigm. In general, this leads to greater robustness and less overfitting on new, unknown data. Figure 18 illustrates this concept using two classes (circles and squares) and an optimal separating hyperplane (*OSH*) with normal vector $w$. The two margin hyperplanes H1 and H2 define the distance to both classes (*margin*). These elements make clear why maximizing the distance between the hyperplane and both classes can reduce the error rate on unseen data.

Since the data cannot always be optimally separated by a linear hyperplane, so-called *soft margins* were introduced, i.e., regions around the hyperplane in which an unambiguous assignment of the data is not possible [51]. If the data cannot

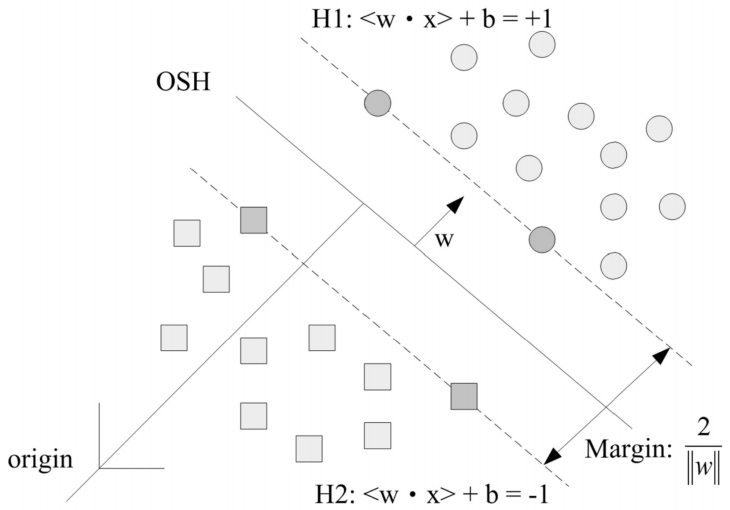



Figure 18: Optimal hyperplane (OSH) for separating data from two classes. The error rate of classifying unseen data can be minimized by maximizing the distance of the hyperplane to both classes [45].

generally be separated by a linear hyperplane, they can be transformed into a higher-dimensional vector space by applying the so-called *kernel trick*, in order to separate the data there with a hyperplane of greater complexity [55]. In a comparison by Hsu et al. [101] between four kernel functions (Linear, Polynomial, Radial Basis Function (RBF), Sigmoid), it was shown that RBF should be used in almost every given scenario, except when the number of features is significantly larger than the number of data points or when both quantities are very large. The kernel function $K\left(x_i, x_j\right)$ of RBF is defined as follows, with $\gamma$ as the kernel parameter and the training vectors $x$:

$$K\left(x_i, x_j\right) = \exp\left(-\gamma \|x_i - x_j\|^2\right), \gamma > 0 \tag{3}$$

For classification problems with more than two different classes, a combination of multiple SVMs must be used. Using One-versus-All (OvA), a separate SVM can be trained for each class $k$ to indicate whether the examined dataset belongs to that class. Alternatively, using One-versus-One (OvO), a SVM is trained for each pair of classes, i.e., $\frac{k(k-1)}{2}$ SVMs in total. The desired class is then determined by majority vote across all SVMs [5, 6, 72]. OvO is the significantly faster algorithm in comparison and also delivers equivalent or better results depending on the context [46, 142].

*Random Forest (RF)*

Another classifier used in many studies is the Random Forest (RF) [5, 7, 152, 153, 198]. It consists of a set of different, randomly generated decision trees (Decision Trees (DTs)) designed to achieve the classification objective. The individual DTs within an RF are usually trained not on the entire dataset but only on a subset. It is important that the (majority of) trees are trained on different subsets so that each tree focuses on part

of the training data while other parts are neglected. As a result, the individual DTs achieve worse results than a DT trained on all available data. However, the entirety of their results is combined, which usually makes them perform better together than a single DT. They also exhibit better generalizability and less *overfitting*, even when the number of trees is increased [5, 31, 55].

When using RFs, hundreds to thousands of DTs are often trained. By training such a large number of classifiers, they can also be used to analyze the data. For example, if noise is added to a single feature and the development of the accuracies of all DTs is observed, this can indicate the importance of that specific feature. [5, 55].

*Logistic Regression (LR)*

Some of the studies examined use Logistic Regression (LR) to solve the classification problem ([5, 7, 134, 199]). Like SVM, LR is designed to separate two classes from one another. This is achieved through a probability calculation used to determine which class the dataset belongs to. In most cases, this probability is calculated using the *logit* function, where $X_i$ represents the features of the data and $\beta_i$ the weights that must be computed during model fitting.

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right) = \beta_0 + \sum \beta_i * X_i, \tag{4}$$

The resulting function usually takes the form of a sigmoid. In general, a steeper slope leads to better predictions. In this way, there are fewer inputs with probabilities near 50 %, which are highly susceptible to noise. Probabilities near 0 % or 100 % are desirable, since in these cases the model is certain that the corresponding dataset can be assigned to one of the two classes [5, 7, 202].

If the data to be classified are to be assigned to more than two classes, LR applies a strategy similar to the one already presented for SVM. Due to the longer training times when using large amounts of data, OvA is usually used for LR instead of OvO [5].

*Template Matching*

Another approach among the studies examined is *Template Matching* [1, 36, 82, 83, 92, 102, 106, 113, 167, 183, 188, 191, 195]. In this procedure, a template is created for each class, for example the average of each feature of a class. A new, unknown dataset is then compared with the available templates and assigned to the class whose template is most similar, for example using the smallest Euclidean distance. The advantages of this approach are the low computational cost and the high interpretability of the data [36].

*Deep Learning*

Artificial neural networks are among the classification methods of *deep learning* and are used in many related works for gesture recognition [35, 102, 130, 161, 170, 175].

They consist of several layers of neurons inspired by biological neural networks. The neurons are interconnected by connections with different weights. The data are passed from one layer to the next and modified by the weighting of the connections. The output values for the entire network are located at the final layer. Training the neural network adjusts the weights.

In gesture classification, special variants of Artificial Neural Networks (ANNs), such as Feed-forward Neural Network (FNN) ([7, 63, 140, 239, 244]) and Back Propagation Neural Network (BPNN) ([125]), are also increasingly used. These are unidirectional neural networks and have no cycles. The neurons of a common layer receive their inputs from the neurons of the previous layer and forward their outputs to the neurons in the next layer. A specialization of the FNN is Convolutional Neural Network (CNN) [19, 53, 62, 178, 207, 229], in which features are extracted within the network by filtering the signals. In contrast stands Recurrent Neural Network (RNN) ([48, 111, 124]), in which data within the network can also flow in the opposite direction.

*Other methods*

In addition to the aforementioned methods for solving the classification problem, the literature examined contains numerous other established methods as well as the authors' own approaches. These include Linear Discriminant Analysis (LDA) ([3, 13, 108, 134, 145, 238]), Hidden Markov Model (HMM) ([73, 141, 190]), k-Nearest Neighbor (kNN) ([7, 176, 219]), Naive Bayes (NB) ([7, 10]) and others, each of which was used by one of the studies examined and is not mentioned further here for reasons of clarity.

### 2.3.4 *Meta-Classification*

In the present dissertation, a differentiated examination of two approaches to *meta-classification* is undertaken. What characterizes both approaches is that the so-called *meta-classifier* derives its predictions at the secondary level on the basis of the results of primary classifiers. These may include both ML-based and rule-based approaches. The difference between the two approaches lies in the type of input data (features) processed by the classifiers of the primary and secondary levels.

These input data (and thus also output data) can be identical for all primary classifiers, whereby the *meta-classifier* is able to aggregate their predictions and optimize its own prediction accuracies. This approach can increase the prediction accuracy and robustness of a system. However, the complete integration of all primary classifiers also increases computational effort and time. An example of such a *meta-classifier* is the Voting (Level 2) Meta-Classifier (VL2). In this case, the primary classifiers process the input data in order to predict the relevant classes. The voting classifier (*Voting Classifier*[37]) of the secondary layer aggregates these prediction probabilities to determine its own final prediction. A weighted average of the prediction probabilities, in which the weights are based on the optimized classifiers of the first layer, yields the

37 `https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.VotingClassifier.html` (Last accessed on January 25, 2026)

best results [5, 241]. For better differentiation between the two different approaches, we refer to classifiers of this type as *Aggregative Meta-Classifiers (AMKs)*.

The other type of *meta-classification* is referred to by us as Composite Meta-Classifier (KMK) and combines the results of multiple classifiers from different data to identify an overarching class. This method is applied in the present work to identify a specific sign from diverse parametric data. The primary classifier level therefore serves to classify the individual parameters, whereas the higher-level *meta-classifier* makes an overall prediction from these individual predictions and knowledge about the performance of the individual classifiers as well as the classes to be predicted (in this case signs). For example, a parameter that is correctly recognized with high probability can be weighted more heavily in *meta-classification* than other parameters. At the same time, analyzing parameter combinations makes it possible to exclude combinations that do not exist in the sign vocabulary and thus increase prediction accuracy.

This specific application of *meta-classification* underscores the innovative character of the parametric approach of this dissertation. In an extensive analysis of more than 90 publications, no comparable approach could be found. Only Kau et al. [113] pursue a similar approach by distinguishing 51 hand shapes of Taiwanese Sign Language (TSL) using two flex-sensor-based data gloves and comparing them with five two-handed, isolated signs. If no sign can be clearly identified, the orientation of the hand (six different values) and the movement trajectory (three values) are incrementally included in the recognition approach. A recognition accuracy of 94.56 % can be achieved.

Similarly to the approach presented in this work, Mohandes and Deriche [145] combine information from different systems at the decision level rather than at the sensor level. For this purpose, they use the evidences of two data gloves and an electromagnetic tracking system and achieve a classification accuracy of 98.10 % with the help of Dempster-Shafer Theory (DST) in recognizing 100 signs. This not only corresponds to the highest number of classes among all related works examined, but also to the highest accuracy in the area of *extensive signs* (cf. Table A.8). To this end, the systems are first evaluated separately and provide independent predictions about the same discrete sign space, i.e., they provide prediction probabilities for the same 100 sign classes. These prediction probabilities are interpreted as evidences and combined with the aid of DST such that matching evidences are reinforced and conflicting evidences are weakened. Data collection was conducted with a single participant, who performed each of the 100 signs 20 times, so that this is a purely *within-user* approach. On the one hand, this illustrates the increasing acquisition effort of a holistic approach compared to a parametric approach as the number of classes grows. On the other hand, it explains the higher recognition rates expected in comparison to *cross-user* scenarios.

An example from the camera-based field of sign recognition can be found in the work by Tavella et al. [214] on *WLASL-LEX*. Here, information on six phonological properties of the *ASL-LEX* (version 1) is extracted from videos in the WLASL database (see Section 2.2.2). Using various classification approaches, probabilities ranging from 62.30 % to 84.50 % could be achieved for recognizing the respective parameters, even

though these parameters were classified individually and not subsequently used to identify a sign.

# 3

# PARASIGNREC: PARAMETRIC FRAMEWORK FOR SIGN RECOGNITION

Building on the findings from Chapter 2 and the resulting research gap, our approach named *Parametric Framework for Sign Language Recognition (ParaSignRec)* for sensor-based sign recognition is now to be presented, taking into account the research objectives defined in Section 1.3. By parametrizing signs, it is intended to provide a sufficiently large vocabulary with high recognition accuracy in order to ensure seamless translation of everyday communication into sign language. Related work uses a holistic approach to sign recognition and can distinguish up to 100 signs with it.

In Section 2.3, various techniques from the field of classification have already been presented. Figure 19 shows a schematic representation of the resulting classification process that we want to use for our approach: In the course of *data acquisition*, the relevant characteristics of a sign are captured by suitable hardware. This can take place independently, and data or parts thereof can also be translated from other domains into the required sensor data via *domain transfer*. During *data preprocessing*, the data are standardized and can be cleaned of erroneous and irrelevant data. The remaining data are *classified* with the help of ML or by rule-based methods. Subsequently, the data and classification results, which have so far been considered in isolation, are used together with the specific knowledge of the target signs in the *sign lexicon* to identify the sought-after sign within the framework of *meta-classification*.

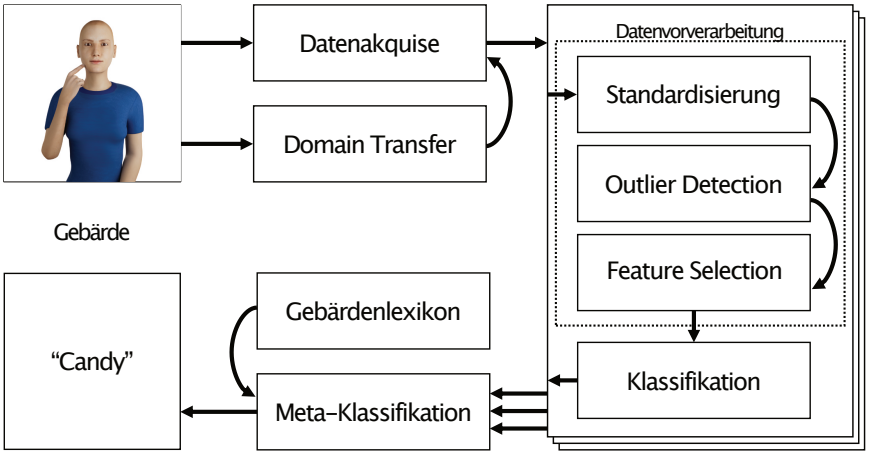



Figure 19: Schematic representation of the parametric framework *ParaSignRec* using the example of the sign Candy (adapted from Achenbach et al. [5])

Due to the simplifications made in Section 1.4 (adherence to the Battison Sign Types, manual segmentation, focus on body-worn sensors, omission of non-manual features,

pronouns, and other directional signs), the focus is on isolated signs using the example of American Sign Language. Only one-morpheme signs are considered, which means that a segmentation approach can be dispensed with. The focus on wearable sensors – so-called *wearables* – and practical everyday usability must be taken into account when selecting the hardware.

*ParaSignRec* consists of three essential components:
*i)* **Parameter identification** (Section 3.1) with the aim of obtaining precise findings on the parameters to be detected and their classes, and on this basis a data set in parameterized form with a sufficiently large vocabulary of signs; *ii)* **Parameter recognition** (Section 3.2), which includes the actual recognition approaches for detecting the individual parameters, consisting of the selection of suitable hardware for capturing the parameters, as well as the subsequent data processing and classification; and *iii)* **Meta-classification** (Section 3.3) to combine these individual results of the recognition approaches for identifying the sought-after signs.

## 3.1 PARAMETER IDENTIFICATION

For our plan of parametric sign recognition, an overview of the parameters of a sign to be captured and their possible classes is needed in addition to suitable recognition methods. With this knowledge, it then becomes possible to select suitable hardware with which the individual parameters can be captured and identified using specific recognition approaches. Moreover, this knowledge provides the basis for creating a data set suitable for meta-classification in the form of a parametrized sign dictionary or for building upon an existing data foundation. This dictionary serves, on the one hand, to identify the sought sign on the basis of the individual results of the recognition approaches; on the other hand, it helps increase the system's accuracy within the framework of meta-classification. This is achieved, for example, by taking symmetry properties into account or by excluding invalid combinations of classes of different parameters in the course of meta-classification.

Our approach to creating a parametrized sign dictionary is shown in Figure 20. The focus in this section is shown in blue and lies on the selection of a suitable data basis on the basis of relevant parameters. These parameters and classes must then be captured and recognized using suitable hardware and recognition approaches (gray areas), which represents the focus of the subsequent chapters 4 and 5.3 to 5.5. In this process, technical limitations in the design of the recognition approaches may lead to changes in the parameters to be captured and their classes. The final dictionary can therefore only be defined conclusively in Section 6.1 after taking these changes into account (green areas).

### 3.1.1 *Selection of relevant parameters*

In Section 2.2.1, numerous notation systems from sign language research were presented that, for the notation of signs, decompose them into their characteristic components. Table 2 provides an overview of all notation systems examined in this work,

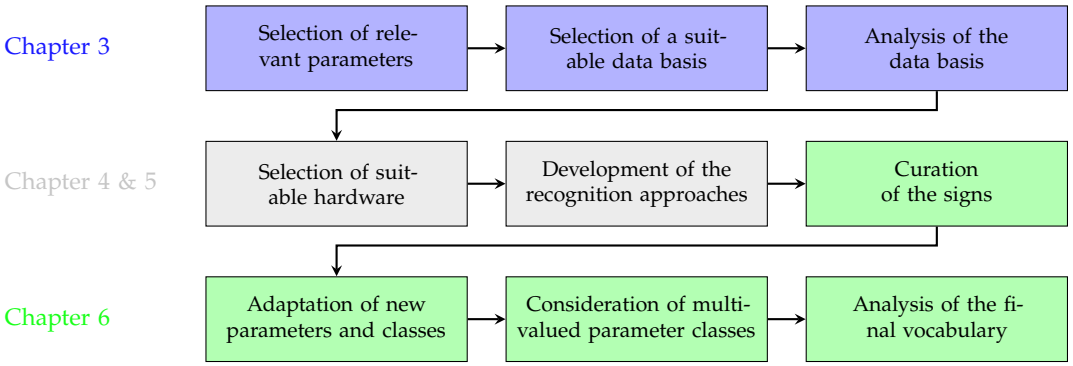



Figure 20: Conception and creation of a sign dictionary. The blue areas address the preparatory measures for the dictionary presented in this chapter (conception). Subsequently, in Chapter 4 and Section 5.3 to 5.5, suitable hardware is selected and the recognition approaches for these parameters and classes are implemented (gray areas), where structural adaptations to parameters and classes may arise. Based on these findings, the final dictionary is created in Section 6.1 (green area).

together with the parameters considered therein. It becomes clear that the parameters *handshape*, *sign location*, *hand movement*, *palm orientation*, as well as non-manual properties such as *facial expression*, are considered in most notation systems. Figure 21 illustrates an example gesture with these five parameters.

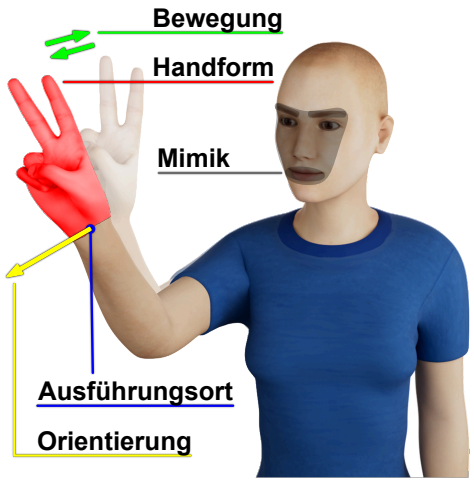



Figure 21: Identified parameters of a sign

In particular, the first three parameters – handshape, location, and movement – are used by all sign notation systems, while a large proportion also integrate non-manual properties. Palm orientation is used less frequently and explicitly in comparison with the other parameters. This is mainly because in some notation systems it is integrated into the symbols of the handshape and not represented separately from the handshape [17]. However, due to the ease with which the individual classes can be interpreted, it makes sense to consider palm orientation when selecting a data basis. From a technical

perspective, a simple form of palm orientation recognition can already be achieved by using an IMU on the back of the hand.

Technical detection of facial expression, on the other hand, poses a considerable challenge. As long as the use of cameras is ruled out for reasons of data protection, this is difficult to reconcile with our goal of everyday usability. Although experimental solutions exist in which sensors on the user's face allow conclusions to be drawn about facial expression [117], these prove to be of limited suitability for everyday use.

Our approach therefore focuses on the parameters *handshape*, *sign location*, and *hand movements* and can be supplemented by *palm orientation*. A suitable data basis in the form of a parametrized sign lexicon should therefore take these parameters into account.

### 3.1.2 *Selection of a suitable data basis*

Previous comparable approaches using body-worn sensors have a vocabulary of up to 100 signs [145], many of them are significantly smaller (on average 21,55 signs across more than 90 studies examined, see Section A.2). Everyday communication, however, requires a considerably larger vocabulary in order to cover a broad range of communicative needs. For instance, Nation [154] cites a minimum knowledge of approximately 3,000 to 5,000 words in order to understand and speak a language, whereby he considers a vocabulary of fewer than 1,000 words to be a limited but functionally sufficient vocabulary for simple tasks. In relation to sign language, ASL dictionaries with up to 5,000 signs can be found [52, 204, 206, 215, 224], which substantiate these figures.

To meet our requirements for a comprehensive and flexible vocabulary, we therefore seek a data basis that comprises at least 1,000 and ideally more than 3,000 words and provides them in parametrized form. As already mentioned, most notation systems do have a detailed, systematic approach to parametrizing signs, but they usually provide only exemplary examples rather than comprehensive collections of parametrized signs. The literature review in this work therefore went beyond the notation systems and also examined different lexical sign databases, which were presented in Section 2.2.2, including *ASL-LEX*, *ASL SignBank*, and *ASLLRP*. In particular, *ASL-LEX* stands out due to its good availability, the relatively large vocabulary of over 2,700 signs, and the numerous phonological properties. A disadvantage, however, is that only one sign video is stored in the database for each sign. *ASL SignBank* follows the phonological encoding of *ASL-LEX*, but offers no download option and only a limited set of phonological properties, which largely coincide with the phonological properties of *ASL-LEX*. Moreover, these are often incomplete and provided for only a fraction of the signs. *ASLLRP*, by contrast, contains only information on the handshape of the dominant and non-dominant hand at the beginning and end of the sign.

In our paper "Towards handshape identification for automatic gesture recognition using sign notation systems" [4], we examined eleven notation systems with regard to their ability to unambiguously represent 30 basic handshapes of ASL, including the fingerspelling alphabet. Notation systems with many symbols provide detailed

descriptions of the individual handshapes, but they make technical recognition more difficult because of the large number of classes the system has to distinguish. Too few symbols, on the other hand, carry the risk that not all relevant sign forms can be clearly differentiated. Ultimately, the two systems *SLIPA* and *Symbol Font for ASL* emerged as particularly promising, since, with a moderate number of 54 and 51 symbols, respectively, they were able to clearly differentiate the handshapes.

Comparing the results of this work with *ASL-LEX*, 27 of the 30 handshapes can be directly mapped to the 59 handshapes of *ASL-LEX*. The remaining three handshapes (*Flattened C*, *Modified X*, and *Curved 3*) can be described by the parameters of the *handshape configuration* (see Section 4.1.1). Since handshape is regarded as a central parameter in sign language [49] and *ASL-LEX* provides sufficient depth of detail for distinguishing and detecting it, we assume that *ASL-LEX* also provides sufficiently many pieces of information for the remaining parameters. In addition, *ASL-LEX* is relatively recent compared with other notation systems, which is advantageous for a living language, and, as already mentioned, it is also a suitable choice due to the lack of alternatives.

For our *ParaSignRec*, *ASL-LEX* provides a solid foundation. Only the parameter *palm orientation* as well as non-manual properties are not taken into account here (or in the other data sources). These must therefore be supplemented manually if needed.

### 3.1.3 *Analysis of the data basis*

After the selection of an apparently suitable data basis, this is now to be analyzed in detail so that subsequent simplifications or extensions can be made. In its second version, *ASL-LEX* consists of 2,723 signs, of which 1,067 are one-handed and the remaining are two-handed. The signs consist of up to six morphemes, with 90.38 % of the signs being monomorphemic (cf. Figure 22a). In total, the 2,723 signs consist of 2,401 different morphemes. 85.17 % of the morphemes occur in only one sign.

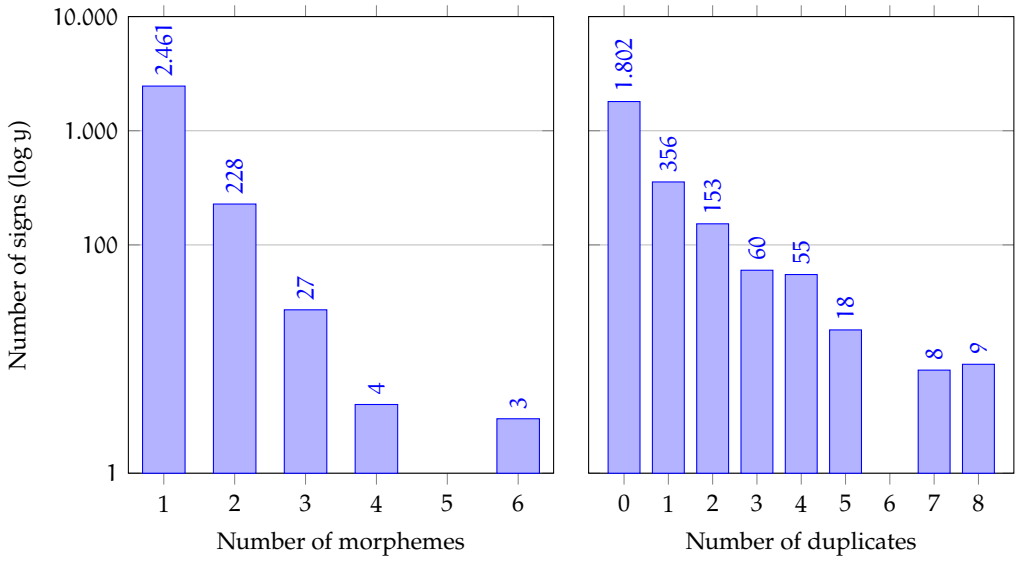
*ASL-LEX* has 16 phonological properties, which are shown in Table 4. These refer exclusively to the dominant hand. Only for handshape is information about the non-dominant hand available, in the form of NON-DOMINANT HANDSHAPE. However, we can make use of Battison's sign types to derive further information about the non-dominant hand from the symmetry properties of the signs. In total, there are four types of signs in *ASL-LEX*, modeled on Battison's sign types:

TYPE 0: One-handed signs (*OneHanded*, 39.52 %)

TYPE 1: Two-handed signs with identical handshape and symmetrical movement of both hands (*SymmetricalOrAlternating*, 36.32 %)

TYPE 2: Two-handed signs with identical handshape in which the non-dominant hand remains at rest (*AsymmetricalSameHandshape*, 5.95 %)

TYPE 3: Two-handed signs with different handshape in which the non-dominant hand remains at rest (*AsymmetricalDifferentHandshape*, 17.48 %)




(a) Number of morphemes in the 2,723 signs of *ASL-LEX*

(b) Number of duplicates among the 2,461 monomorphemic signs

Figure 22: Number of signs in *ASL-LEX* by number of morphemes and number of duplicates for monomorphemic signs (logarithmic scaling)

It is evident that signs of types 0 and 1 are the most frequent, at approximately 36.32 to 39.52 %. In addition, 2.90 % of the signs violate Battison's rules and therefore cannot be assigned to any of the types just mentioned.

Moreover, in *ASL-LEX* the Sign Type according to Battison is specified for each sign, although no further distinction is made for one-handed signs, which is why these will henceforth be referred to as *Type 0* for ease of reading. In comparison with the phonological properties, the sign types cannot be classified directly, but they provide valuable information about symmetry properties, such as handedness (one-handed or two-handed), the handshape of both hands (identical or different), palm orientation (symmetrical or asymmetrical), and the movements of both hands (symmetrical or asymmetrical). These can be used to enrich the system with additional information and thereby optimize the results of meta-classification.

Despite this wealth of information, *ASL-LEX* is not able to describe every sign unambiguously. Therefore, there are numerous duplicates within *ASL-LEX*, i.e., signs with the same phonological encoding but different meanings. A representative example is shown in Figure 23: all eight depicted signs have the same phonological encoding[38] and therefore cannot be distinguished in their phonological properties. However, it is clearly evident that the signs are different and thus also have different meanings. The only exception is the signs Street and Way, which are identical in execution, since they are synonymous and can therefore be regarded as one sign. This reduces the

38 Open B handshape of both hands, identical handshape configuration, no contact of the thumb with the selected finger, no internal movements, straight path movements of the hands without repetition, performed in neutral space, symmetrical behavior.

Table 4: Overview of the parameters of *ASL-LEX* (for more information, see Table A.2)

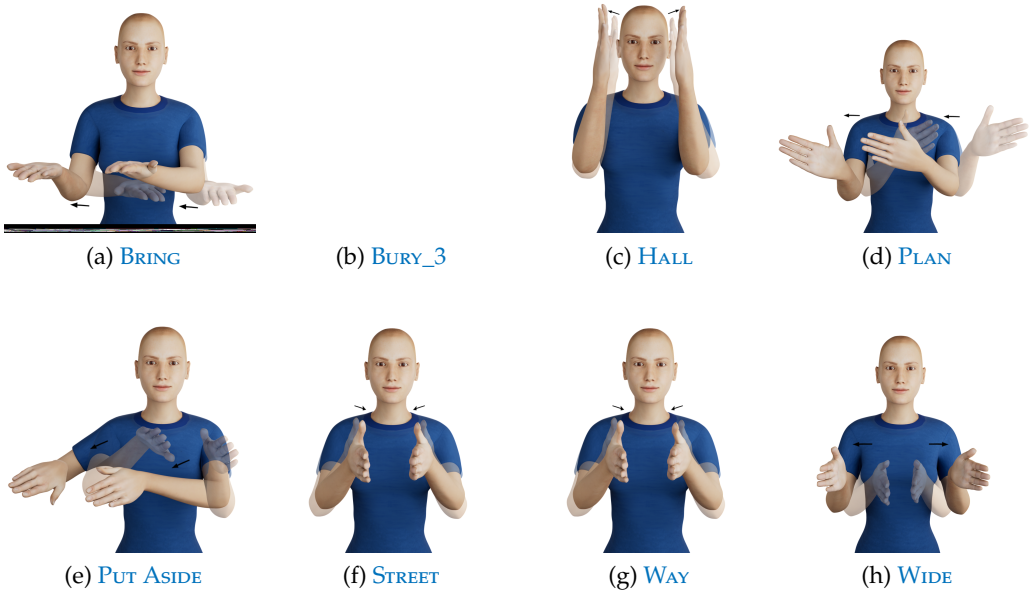
| Category | Parameter | Description |
|---|---|---|
| Handshape | HANDSHAPE | Dominant handshape |
| | NON-DOMINANT HANDSHAPE | Non-dominant handshape |
| | THUMB CONTACT | Does the thumb touch SELECTED FINGER? |
| Handshape-configuration | SELECTED FINGERS | Active or prominent fingers |
| | FLEXION | Flexion of the SELECTED FINGERS |
| | SPREAD | Spreading of the SELECTED FINGERS |
| | THUMB POSITION | Thumb position |
| Movements | PATH MOVEMENT | Type of hand movement |
| | ULNAR ROTATION | Is the wrist rotated? |
| | FLEXION CHANGE | Does the flexion change? |
| | SPREAD CHANGE | Does the spread change? |
| | REPEATED MOVEMENT | Is the movement repeated? |
| Location | MAJOR LOCATION | Primary location |
| | MINOR LOCATION | Secondary location |
| | SECOND MINOR LOCATION | Destination in path movement |
| | CONTACT | Does the hand touch MAJOR LOCATION? |

number of actual duplicates in this example to seven. Looking a little more closely at the remaining signs, one can in particular recognize differences in the movement directions of the two hands or their movement symmetry, as well as in the different palm orientations of both hands:

MOVEMENT DIRECTION AND SYMMETRY: BRING and PLAN are performed from one side to the other, HALL and STREET/WAY forward, BURY 3 downward, and PUT ASIDE laterally downward. WIDE is performed with an alternating movement in which both hands move outward.

PALM ORIENTATION: The palm faces inward at the beginning of the signs BURY 3, HALL, STREET/WAY, and WIDE, upward for BRING, sideways for PUT ASIDE, and one hand forward and one backward for PLAN. The extended fingers point forward in BRING, BURY 3, PUT ASIDE, STREET/WAY, and WIDE, upward in HALL, and sideways in PLAN. By the end of the sign, the palm orientation has in part changed depending on the direction of movement.

A detailed analysis shows that, based on the available phonological properties, there are 2,062 different feature combinations among 2,461 monomorphemic signs, and only 1,802 (73.22 %) of all 2,461 monomorphemic signs can be uniquely identified



(a) BRING (b) BURY_3 (c) HALL (d) PLAN

(e) PUT ASIDE (f) STREET (g) WAY (h) WIDE

Figure 23: Examples of different signs with identical phonological encoding[38].

(see Figure 22b). All other signs share the same phonological encoding with up to eight additional signs.

## 3.2 PARAMETER RECOGNITION

After the parameters to be captured have been identified, the focus in this section is on their detection. To this end, the requirements for the recognition approach and the hardware used are first specified. Suitable sensor technology must be selected for data acquisition that meets both the technical and general requirements. Only on this basis can the detection of the individual parameters and the design of the various ML classifiers be carried out in detail in the following chapters.

### 3.2.1 *Requirements for the Recognition Approach and Hardware*

For the capture and subsequent detection of signs, these must be recorded. Since no holistic recognition approach is pursued, we can consider all parameters to be captured and their measured quantities in isolation from one another when selecting suitable sensor technology. In addition, there are also general requirements for the hardware that are independent of the quantities to be measured. Fricke and Bressem define the requirements for the ideal recording procedure in their book "Gesten – gestern, heute, übermorgen" [71] as follows: *i)* High precision, i.e. sufficiently accurate capture of finger (-movements), but also reliable detection of contacts between the fingers themselves (e.g. contact between the thumb and fingertips, crossing of fingers), *ii)* as little data post-processing as possible, *iii)* low latency, *iv)* capture of the absolute (hand)

position in space and the relative orientation to the body, including contacts with one's own body, as well as with external bodies and objects, *v)* (if required) adaptability to individual physical characteristics (e.g. finger lengths).

On this basis and taking into account our research goals from Section 1.3, the general requirements shall first be defined. The approach should therefore be robust, mobile, and user-friendly, and thus offer a certain degree of suitability for everyday use. The system should be comfortable to wear and should not restrict movements while signing. Since our concept is an assistive aid system that represents significant added value for the user, price can be regarded as a secondary concern. For practical reasons and to preserve privacy, we focus on the use of wearable sensors and dispense with the use of cameras. As already explained in Section 1.4, the capture of pronouns and directional signs without the use of cameras, which would enable the capture of the absolute hand position, is difficult to realize. The detection of the individual parameters should also be implemented with a minimum of hardware. This ensures maximum flexibility and later gives us the possibility of replacing individual components with alternative solutions without fundamentally adapting the overall system.

To enable a comparable and meaningful evaluation of the classification results, the macro-$F_1$ score is used in this work as the harmonic mean of *precision* and *recall*. *Precision* is defined as the ratio of the number of correctly positive classifications (*True Positives*, TPs) to the total number of all cases classified as positive. *Recall*, on the other hand, denotes the ratio of the number of correctly positive classifications (TPs) to the sum of correctly positive classifications (TPs) and falsely negative classifications (*False Negatives*, FNs) [16].

Alternatively, in experiments with an approximately uniform class distribution, accuracy (*Accuracy*) is also used, since under these circumstances each instance is assigned to exactly one class and all misclassifications are weighted equally. *Accuracy* is defined as the ratio of the number of correct classifications (TPs and *True Negatives*, TNs) to the total number of all classifications.

The evaluation is carried out using macro averaging, since a class-based and not instance-based assessment is intended. This reduces the influence of differing class frequencies, as each class contributes equally to the overall result regardless of how often it occurs. In macro averaging, the metrics are first calculated separately for each class and then averaged with equal weighting. In contrast, micro averaging would result in frequently occurring classes dominating the overall result.

In order to recognize signs with their static poses and dynamic transitions, the hardware must be capable of accurately capturing the individual finger positions, as well as the orientation of the hand and the location where the sign is produced. The dynamic transitions can be derived from changes in these static poses. To better capture movements, the use of IMUs is also advisable. Any hardware should have a minimum sampling frequency of at least 24 to 40 Hz. This is derived from the maximum frequency of human movements, which depending on the source lies between 5 to 20 Hz [29, 120]. The frequency range relevant for sign recognition is therefore between 0 to 6 Hz [120] or 0 to 12 Hz [29]. To satisfy the Nyquist-Shannon sampling

theorem [160, 192], our hardware should therefore have a data rate of at least twice that amount ($f_s \geqslant 2 \cdot f_{max}$ with sampling rate $f_s$).

Sign detection is to take place in real time, which is why the hardware used must have low latency. In Section 1.3, it was already defined that our approach should be *real-time capable* in order to ensure fluid and seamless communication. We now wish to specify what our definition of *real time* is. As already explained, the processing time of our system on standard hardware[39] must always be shorter than the execution time of the respective subsequent sign. If this is not guaranteed, signs will accumulate and the flow of communication will be disrupted. Processing time includes the latency of the sensor technology or hardware, the duration of data preprocessing, the duration of parameter classification, and the subsequent meta-classification. The system latency is thus composed of $\sum t_{Hardware} + \sum t_{Datenvorverarbeitung} + \sum t_{Parameterklassifikation} + t_{Metaklassifikation}$. Since we do not have the means to measure the latency of the sensor technology precisely in the millisecond range, we are dependent on the manufacturers' specifications. Since we have already set the minimum sampling rate based on human movements to 24 Hz, the hardware used should be capable of processing this data in at least the time in which new data are acquired. The minimum latency for the hardware should therefore be $t_{Hardware} > \frac{1}{24}s \approx 41.67\,ms$. If the hardware latency cannot be determined unambiguously, we restrict ourselves to considering the classification duration of the individual parameters, the meta-classification duration and – insofar as carried out by our system – also data preprocessing.

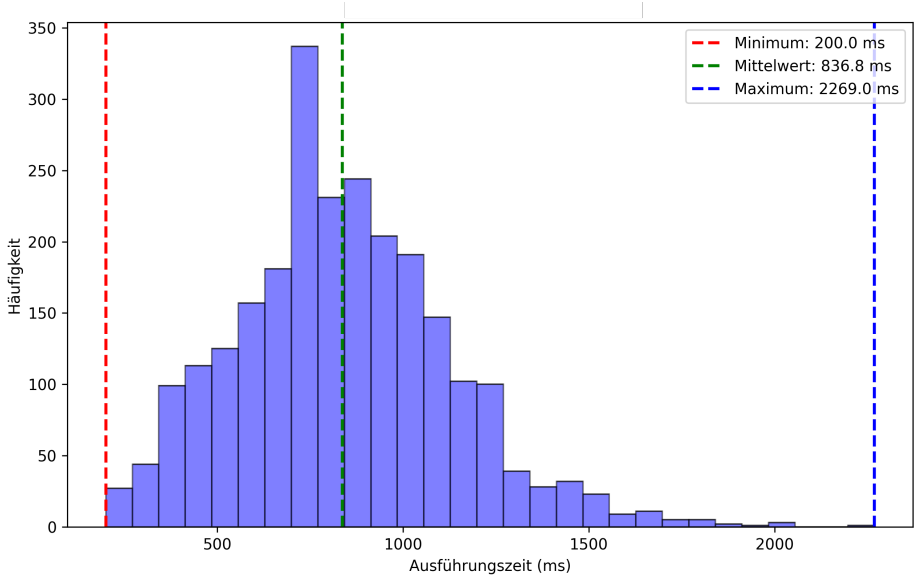



Figure 24: Execution times of the signs in *ASL-LEX*

Since there are significant differences in execution time within signs, our definition of real-time capability must be further specified accordingly. *ASL-LEX* provides the execution time of the video-recorded signs for each sign, with the onset and offset

39 for example, the current generation *Apple MacBook Pro*

of the sign production already subtracted. Figure 24 shows the execution times of the 2,461 monomorphemic signs in the *ASL-LEX*, with a minimum execution time of 200 ms up to a maximum execution time of 2,269 ms. The mean is 836.80 ms and marks the lower bound of our definition of real-time capability. Ideally, however, our system processes each sign faster than the shortest documented execution duration. Overall, we define the following four target levels for real-time capability with respect to the processing time of our approach compared to the execution time of signs. The processing time of a sign is ...

1. ... **always** faster than the **shortest** sign ($t_i < 200\,\text{ms}$).
2. ... **on average** faster than the **shortest** sign ($\bar{t} < 200\,\text{ms}$).
3. ... **always** faster than the **average** of all signs ($t_i < 836.80\,\text{ms}$).
4. ... **on average** faster than the **average** of all signs ($\bar{t} < 836.80\,\text{ms}$).

In addition to this runtime requirement, Table 5 finally shows the determined requirements for our system. These are grouped according to functional requirements for our overall system, technical requirements for the individual recognition approaches and their hardware, and user-oriented requirements for the usability of the approach.

Table 5: Comparison of the different system requirements

| **Functional** | **Technical** | **User-oriented** |
|---|---|---|
| Precise recognition of ...<br>• finger poses<br>• movements<br>• orientations<br>• (relative) position<br>• self-contacts<br>• external contacts<br>Real-time capable ($\bar{t} < 836.8\,\text{ms}$)<br>Minimal hardware setup<br>Adaptability to individual physical characteristics | Precise capture of the quantity sought<br>Robust<br>High sampling rate (> 24 Hz)<br>Low latency (< 41.7 ms)<br>No data post-processing necessary | Mobile<br>Freedom of movement<br>User-friendly<br>Privacy preservation<br>Availability |

### 3.2.2 *Selection of the Hardware*

The related works in Table A.7 and Table A.8 have shown that data gloves and EMG sensors appear to be suitable hardware for recognizing signs and, in particular, hand shape. Another, more experimental approach is MMG. Data gloves and MMG also have the advantage that they are often equipped with IMUs, which can provide us

with additional information about the movement as well as the orientation of the hand.

In particular, commercial products – if available[40] – often feature preprocessing of the data directly on the hardware and their own developer interfaces. These usually provide information on orientation in the form of quaternions for the back of the hand and for each finger. In addition, the angles between the individual finger segments can also be queried directly [5, 7]. This means that the complex calculation of quaternions based on IMU sensor data and using a *Kalman* filter can be dispensed with. The same applies to any sensor fusion when using multiple IMUs or when combining IMUs and flex sensors. Systematic and random errors such as noise and drift can also be reduced during data preprocessing.

By using commercial products, sufficient availability can also be ensured, and in general an appropriate degree of user-friendliness is also achieved, for example through simple attachment and handling of the sensor technology and comprehensive documentation.

If one considers the related works in Table A.8, which capture not only the hand shape, orientation, and movements of the hand but also the location where the sign is produced, it becomes apparent that mobile, camera-free approaches for capturing the execution location are not taken into account in the state of the art. It is therefore necessary to develop a separate solution. In earlier work, we have already shown how a whole-body pose can be reconstructed using *inverse kinematics* and stationary trackers [42]. We now apply this knowledge to use *direct kinematics* in the opposite direction for determining arm poses. Instead of estimating the angular positions of individual segments via their end-effector position, we determine the relative position of the hands with the aid of arm lengths and the orientations of the upper arm and forearm using IMUs. In addition to reliable sensor technology for capturing the hand parameters, we therefore also require sensor technology to capture the movement and orientation of the arms, for which IMUs are best suited.

## 3.3 META-CLASSIFICATION USING PROBABILISTIC MODELS

One of the central challenges of this work consists in realizing a reliable recognition approach for signs. This includes not only the recognition of the individual parameters by specialized classifiers, but, building on these results, also their most effective possible interpretation into a complete sign. We refer to this overall classification as compositional meta-classification (KMK), or simply *meta-classification*. In this process, both the properties of the individual classifiers and the interplay of these classifiers must be taken into account in order to predict the target signs as quickly and accurately as possible.

40 At the time of this work, we were not aware of any commercial products/sensors in the case of MMG

### 3.3.1 *Probabilistic model for class selection*

At the outset, the challenge arises of how a multitude of independent classifiers can be combined in order to identify the target sign from their predictions. An intuitive approach is a hierarchical structure in the form of a decision tree, as exemplarily shown in Figure 25. Each level of the decision tree represents a classifier for a single parameter, with each branch corresponding to one possible class. The leaves of the final level of the decision tree ultimately correspond to the possible signs and represent the result of the combination of all previously made classification decisions. The classifiers are evaluated sequentially and provide individual results in the form of an unambiguous class prediction. By following those predicted classes in the decision tree, one reaches the predicted sign.

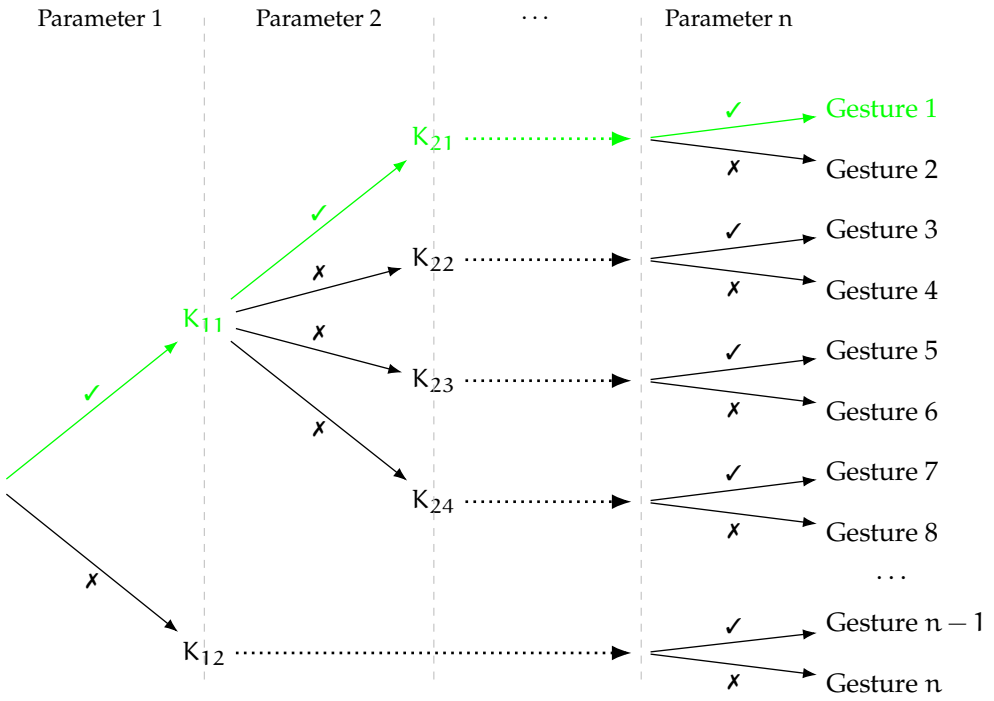



Figure 25: Multi-stage decision tree for the classification of gestures based on parametric features. Each branch represents a class of the corresponding classifier. The green path follows the individual predictions with *Gesture 1* as the target gesture.

The structure of the subsequent levels is independent of the respectively selected branch, so that structurally identical subtrees arise at each level, whose leaves, however, represent different combinations of class values and thus different signs. This structure makes it possible to retain all possible classification paths independently of previous decisions, but at the same time leads to a maximum number of branches in the decision tree.

Through the combination of the classifiers, however, their uncertainties also multiply. For example, the overall probability of a system of ten classifiers with 90 % accuracy each decreases to only 35 % ($0,9^{10} \approx 0,35$). This means that on average only

every third sign would be recognized correctly. The actual classifiers will, however, exhibit different numbers of classes and different accuracies, whereby in particular those with many classes are more prone to errors. In the described structure, even a single incorrect partial classification already leads to the correct sign no longer being predictable.

However, this disadvantage of hierarchical classification is offset by the advantage that with each additional classifier, the number of possible combinations of the different classes, and thus also the potentially distinguishable signs, increases exponentially. Let us assume in our example that each of the ten classifiers can distinguish between two classes; then there are already $2^{10} = 1.024$ possible combinations. If each classifier contained three classes, there would already be $3^{10} = 59.049$ different combinations – more than 21 times as many as there are signs in *ASL-LEX*. This large number of possible combinations fundamentally makes it possible to distinguish very many signs, but it entails the risk that incorrect classifications can lead to a multitude of invalid or misleading paths.

An obvious modification of the approach would therefore be to remove those paths in the decision tree that do not lead to a known or permissible sign. This can increase system accuracy, since incorrect class combinations are removed from the set of target combinations from the outset. However, one would then need a strategy for how to proceed if a classifier makes a prediction that has already been excluded. This problem particularly affects non-binary classifiers, since after the exclusion of one prediction, more than one alternative remains. It is obvious that providing numerous classifiers for all class combinations is impracticable. In addition, one fundamental problem would remain: the overall accuracy of the system is always limited by the *weakest classifier*. Once an incorrect path has been taken in the decision tree, it can no longer be corrected, even if the paths are heavily reduced. Figure 26 illustrates this challenge.

To counteract this, forcing an unambiguous decision should therefore be avoided. Instead, we propose a *probabilistic approach* to meta-classification by systematically taking prediction uncertainties into account. To this end, we make use of the possibility of obtaining from the classifiers used not only a hard class assignment, but also a *probability distribution* over all possible classes. In this way, among the remaining paths in the decision tree, the one with the highest probability can always be followed, as illustrated in Figure 27. This leads in every case to a possible prediction, but still with the restriction that an incorrect decision can no longer be corrected, provided that it leads to a valid class combination.

Against this background, it appears reasonable instead to calculate the *combined probability* of each potential sign on the basis of all relevant prediction probabilities of the parameters. The most promising sign can then be selected based on the highest score. The computational effort scales linearly both with the number $G$ of signs and with the number $P$ of parameters, resulting overall in a complexity of $\mathcal{O}(G \cdot P)$. Figure 28 shows the decision tree of the previous examples with this approach.

This linear scaling implicitly assumes that a sign has exactly one correct class per parameter. In practice, however, it may occur that multiple permissible classes are assigned to a sign for a parameter, for example due to uncertainties in classification

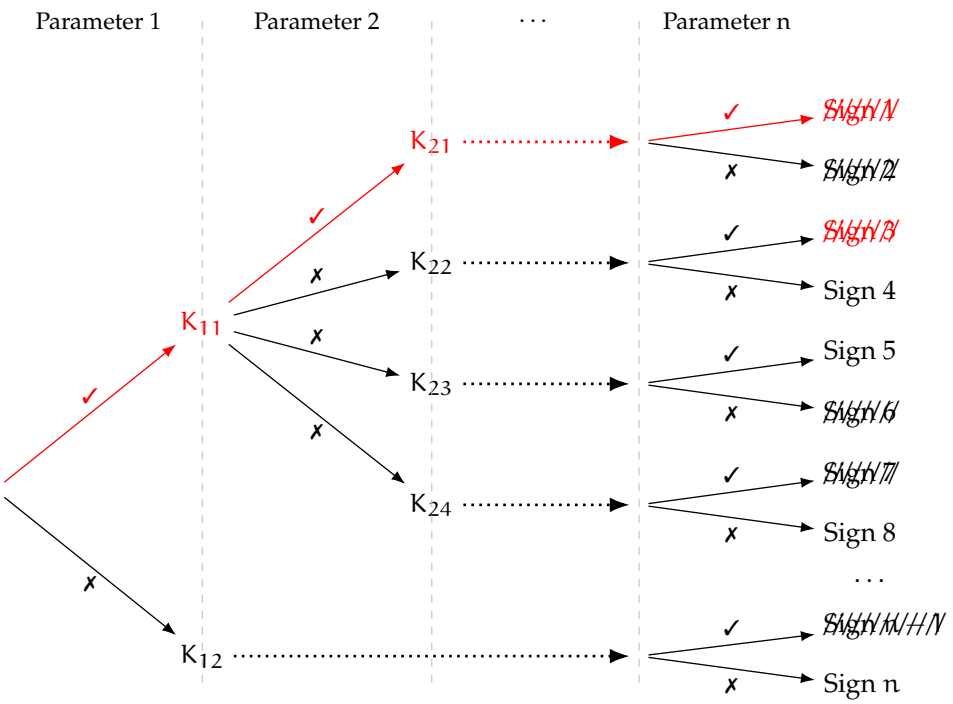



Figure 26: Decision tree from Figure 25 with signs excluded due to invalid class combinations. The red path follows the individual predictions, but it cannot determine a valid target sign.

or different variants. In this case, for a single sign, all combinations of the permissible classes across all parameters must be taken into account. The number of combinations to be evaluated can therefore grow exponentially with the number of parameters, leading to a considerable increase in computational effort. If each parameter on average allows $A$ permissible classes, then instead of one valid path, up to $A^P$ paths arise per sign. Since an aggregation over $P$ parameters is required for each path, this results overall in a complexity of $\mathcal{O}(G \cdot P \cdot A^P)$. Depending on the number of parameters and permissible classes, the computational effort thus quickly reaches a practically unmanageable scale.

Due to the fact that the subsequent subtrees are identical regardless of the selected class, however, it is not necessary to evaluate all possible paths explicitly for each sign: it suffices to consider, for each parameter, the permissible class with the highest score, as all alternative paths necessarily lead to a lower combined score. The computational effort is thereby reduced again to a linear scaling of $\mathcal{O}(G \cdot P)$ and is thus independent of the number $A$ of permissible classes per parameter. Figure 29 shows our example used with support for multiple sign variants.

### 3.3.2 *Combination of the individual classifiers and final prediction*

Now that it has been clarified in principle how a potential sign can be identified on the basis of a multitude of classifiers and which advantages arise from a probabilistic

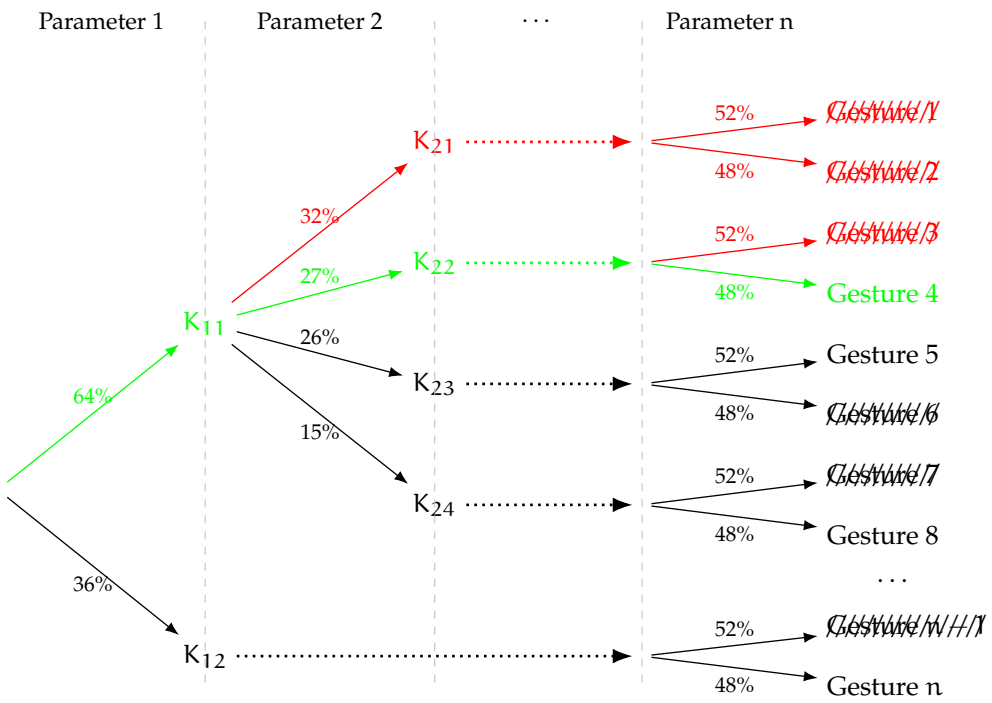



Figure 27: Decision tree from Figure 26 using probabilistic instead of discrete predictions. The marked path follows the highest accuracies until it can reach a valid target gesture with *Gesture 4*. The branches are sorted by their Vohersage accuracies and iterated sequentially, starting with Gesture 1.

approach, the next step describes the concrete calculation of a combined score that appropriately accounts for the contributions of the individual classifiers.

First, it is ensured that the prediction probabilities $P_n(c_i) \in [0, 1]$ of a classifier $n$ with $K$ classes $c_i$ are normalized:

$$P'_n(c_i) = \frac{P_n(c_i)}{\sum_{i=1}^{K} P_n(c_i)}. \tag{5}$$

This ensures that the sum of all prediction probabilities of a classifier always amounts to 100 %:

$$\sum_{i=1}^{K} P'_n(c_i) = 1, \tag{6}$$

The combined probability $S$ (*score*) of all $N$ classifiers can then be determined by

$$S(s_i) = \frac{\sum_{n=1}^{N} p_{s_i,n}}{N}, \quad \text{with} \quad p_{s_i,n} = \max_{(c_i) \in L_{s_i,n}} P'_n(c_i) \tag{7}$$

as the highest prediction probability of a permissible class of classifier $n$ for the target sign $s_i$, and $L_{s_i,n}$ as the set of permissible classes.

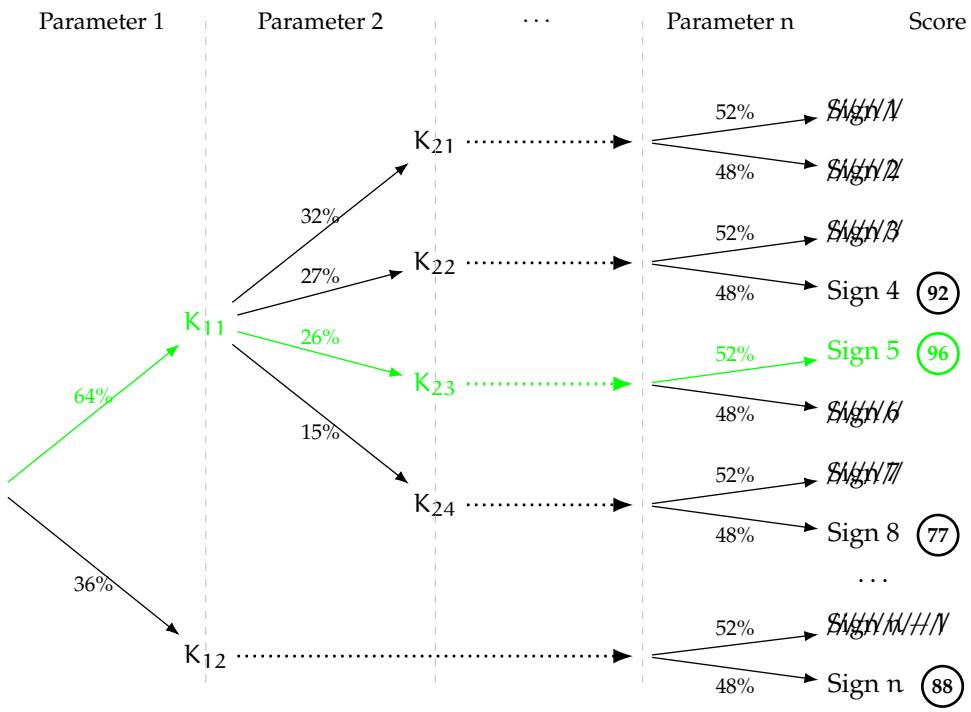



Figure 28: Decision tree from Figure 27 with determination of the target gesture via the calculation of the overall score. The green path leads to *Sign 5* as the target gesture with the highest combined probability (Score).

However, the classifiers exhibit different prediction accuracy and generalizability due to differing prerequisites – for example, the underlying features, the data quality, and the number of classes. In addition, each parameter has its own relevance, which must also be taken into account. A uniform treatment of all classifiers would implicitly assume their equivalence and could lead to less accurate or less relevant classifiers influencing the final prediction disproportionately. We therefore weight the classifiers and modify Equation 7 as follows:

$$S(s_i) = \frac{\sum_{n=1}^{N} \omega_n \, p_{s_i,n}}{\sum_{n=1}^{N} \omega_n} \tag{8}$$

The weights $\omega_i \in [0.5, 1.5]$ are deliberately chosen by us so that the influence of the strongest classifier exceeds that of the weakest by at most a factor of three and at the same time the weakest classifier still has at least half the influence on the result, as in an unweighted state. The concrete determination of the weight values is carried out later on the basis of a separate test data set with the aim of maximizing the accuracy of the final prediction as the result of meta-classification.

This weighted arithmetic mean achieves a robust combination of the individual classifiers that takes into account both different performance levels and varying relevance of the parameters and thus forms the basis of a reliable recognition approach for signs.

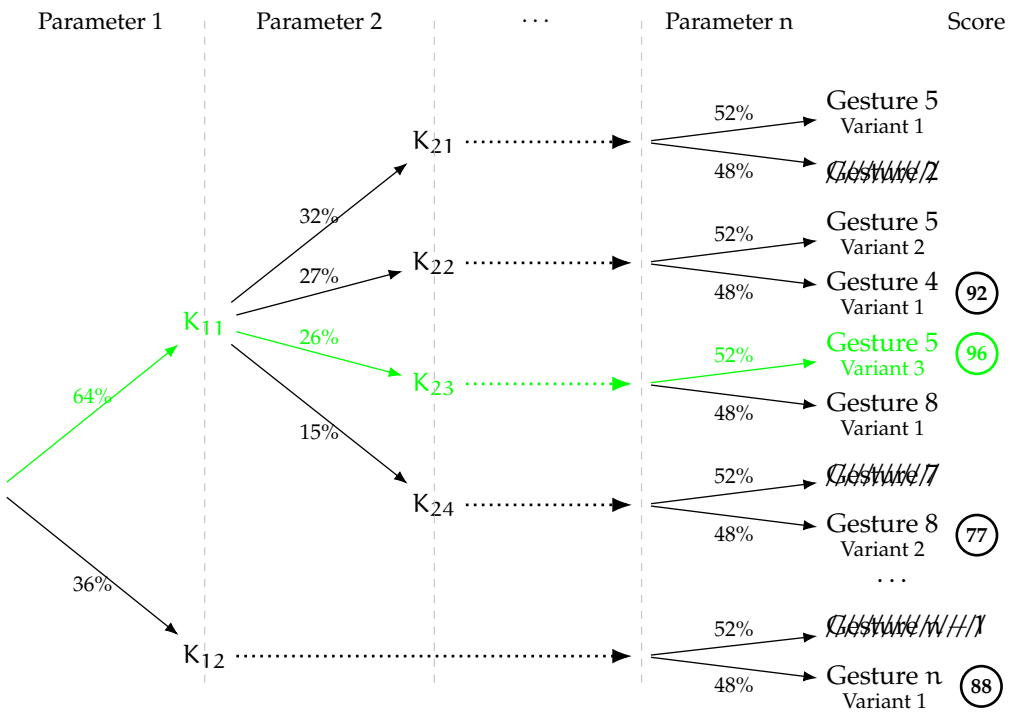



Figure 29: Decision tree from Figure 28 with support for multiple gesture variants by considering multiple possible correct classes per parameter. For each gesture, only the score of the most probable variant is computed. The green path leads to the third variant of *Gesture 5* as the target gesture with the highest combined probability (Score).

The arithmetic mean was chosen as the basis of this calculation because it represents the classical average and offers a balanced consideration of all classifiers, regardless of whether they provide high or low probabilities. In contrast, the geometric mean is multiplicatively linked, which would have the consequence that even a single misguided classification with a score of 0 leads to an irreparable overall result. The harmonic mean likewise exhibits high sensitivity to small values, so that parameter combinations with at least one very low individual result yield a correspondingly low overall probability. The quadratic mean places emphasis on larger individual results, so that even single very high scores can lead to an increased overall result [56, 173, 211].

# 4

# RECOGNITION OF HANDSHAPE

The most important parameter of a sign language is handshape [49]. Many signs can be distinguished only by handshape; this is also true for most signs of the fingerspelling alphabet, which is shown in Figure 30. Only two pairs of letters differ here by the movement of the hand (I ⇔ J and 1 ⇔ Z) and a further three by different palm orientations (K ⇔ P, G ⇔ Q and H ⇔ U) [5]. The 27 signs of the fingerspelling alphabet can therefore be described with 21 different handshapes. In the truest sense of the word, it is thus "obvious" that special importance should be attached to handshape detection.


Figure 30: American fingerspelling alphabet (adapted from [52])

In this chapter, we therefore aim to determine which handshape parameters are to be recognized and which concrete values they can assume. Suitable hardware is to be identified that is capable of capturing these parameters with sufficient accuracy. Based on the captured data, a classification process is to be developed that, on the one hand, ensures high accuracy in recognizing the parameters and, on the other hand, is sufficiently fast so that our system remains real-time capable.

## 4.1 SELECTION OF THE HAND PARAMETERS TO BE CAPTURED

For the ParaSignRec, we draw on the seven phonological properties of the *ASL-LEX* that describe *handshape*. These consist of *i)* HANDSHAPE and *ii)* NON-DOMINANT HANDSHAPE as unified combinations of the following properties, in order to describe the static handshape for both the dominant and the non-dominant hand. In contrast, the phonological properties *iii)* SELECTED FINGERS, as those fingers that move or exhibit a characteristic flexion or spread for the handshape, *iv)* FLEXION, as flexion, and *v)* SPREAD, as spread of the SELECTED FINGERS, as well as *vi)* THUMB POSITION, as information on whether the thumb is abducted or bent, are specified only for the dominant hand. All these properties have in common that they describe the sign at the beginning of sign production. Only *vii)* THUMB CONTACT indicates whether the thumb of the dominant hand touches one of the SELECTED FINGERS at any point during sign production. The full descriptions, including any exceptions, can be found in Section A.1.2, Table A.2.

Since the handshape changes in many cases during the production of a sign, it is described by the static handshape at the beginning of the sign and its dynamic change during the sign. This can occur through a change in flexion (FLEXION CHANGE) or spread (SPREAD CHANGE) of the individual fingers. In this section, the recognition of the static handshape at the beginning of the sign is the main focus; dynamic change is discussed in detail as part of movement recognition in Section 5.4.

### 4.1.1 *Handshape vs. handshape configuration*

The handshape description of the *ASL-LEX* is originally based on the *Prosodic Model* by Brentari [32], in which handshape is described by the information in SELECTED FINGERS and FLEXION and was later supplemented by SPREAD and THUMB POSITION. These pieces of information are also referred to as the *handshape configuration*. However, by focusing on the SELECTED FINGERS, only part of the fingers is described, namely those that move or are not fully extended or closed (if there is no movement of the individual fingers). This means that some handshapes cannot be described unambiguously and that a single handshape configuration can lead to different handshapes. For example, it is not possible to describe flexion separately for each finger, but only for a selection of fingers. Figure 31 shows examples that are excerpted in Table 6. The examples are always arranged in pairs and show, per column, two different handshapes with the same handshape configuration. To counter this problem, *ASL-LEX* identified 91 handshape configurations within the 2,723 signs used and subsequently manually grouped or separated them to obtain 58 different static handshapes[41] [187].

This issue is particularly relevant for our approach in that additional phonological properties of the *ASL-LEX* outside the handshape configuration are also based on the SELECTED FINGERS. These include the two dynamic properties FLEXION CHANGE and SPREAD CHANGE, which indicate whether the SELECTED FINGERS move or not. Like-

41 `https://osf.io/pu4bk` (Last accessed on January 25, 2026)
42 *i* index finger, *m* middle finger, *r* ring finger, *p* little finger and *t* thumb
43 No handshape from the *ASL-LEX*

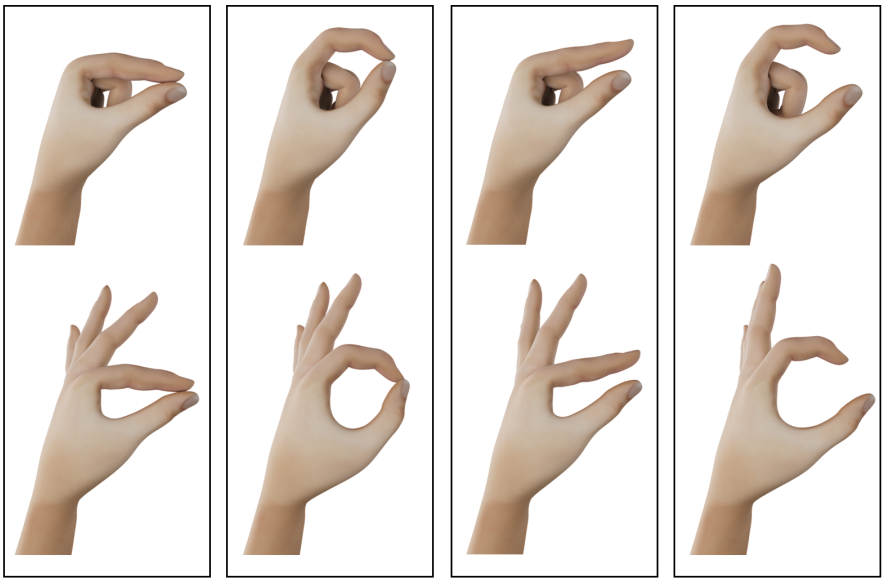


Figure 31: Pairs of handshapes that cannot be distinguished by the original phonological features of the *ASL-LEX* (handshape configuration, see Table 6) [187]

Table 6: Translation examples between handshape and handshape configuration [187].

| **Handshape** | | **Handshape configuration** | | | |
|---|---|---|---|---|---|
| | | Selected Fingers[42] | Flexion | Spread | Thumb Position |
|  | Baby Beak[43] | i | Flat | NA | Closed |
|  | Flat F[43] | i | Flat | NA | Closed |
|  | Baby O | i | Curved | NA | Closed |
|  | F | i | Curved | NA | Closed |
| ... | | ... | | | |

wise, the parameter THUMB CONTACT indicates whether the thumb touches one of the SELECTED FINGERS. It is neither part of the handshape configuration nor considered in the aggregated parameters (NON-DOMINANT) HANDSHAPE, since these describe the static handshape at the beginning of sign production, whereas contact between the thumb and the SELECTED FINGERS can occur at any time during production and not only initially.

As another criticism of the handshape configuration model, we note that the properties FLEXION and SPREAD are also not completely independent of one another. For example, the flexion values *Crossed* and *Stacked* are rather a combination of the flexion value *FullyOpen* and negative spread (adduction) (example R) or a particular preceding positioning of the thumb (examples P and STACKED 5), see Figure 32.

(a) R (b) P (c) Stacked 5

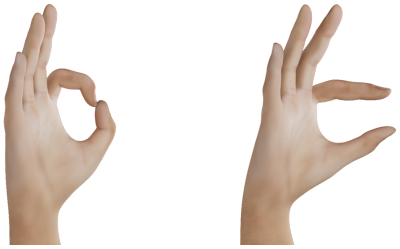
Figure 32: The handshapes R, P and STACKED 5 of the ASL.

### 4.1.2 *Selection of the parameters*

For handshape recognition, we therefore focus on recognizing the static handshapes for the dominant and non-dominant hand (HANDSHAPE and NON-DOMINANT HANDSHAPE) and ignore the phonological properties FLEXION, SPREAD and THUMB POSITION, since their information is already taken into account in the handshape. The information on SELECTED FINGERS is incorporated later in the consideration of the dynamic properties FLEXION CHANGE, SPREAD CHANGE and THUMB CONTACT. Since there are many signs that differ only in whether there is contact between the thumb and the SELECTED FINGERS (see example in Figure 33), the recognition of the thumb and the fingers for THUMB CONTACT must be very precise in order to distinguish between *contact* and *non-contact*.

Although *ASL-LEX* lists 58 different handshapes for the dominant hand, the two handshapes K[44] and P[45] are identical and are therefore adopted by us as K [4]. The handshapes FLAT H[46] and FLAT N[47] are also depicted identically; however, upon inquiry to Sehyr [185], we were informed that they differ in the position of the thumb (see Figure 34) and that the illustrations in the *ASL-LEX* therefore correspond, in both

44 `https://aslcdi.website/images/handshape_images/k.png` (Last accessed on January 25, 2026)
45 `https://aslcdi.website/images/handshape_images/p.png` (Last accessed on January 25, 2026)
46 `https://aslcdi.website/images/handshape_images/flat_h.png` (Last accessed on January 25, 2026)
47 `https://aslcdi.website/images/handshape_images/flat_n.png` (Last accessed on January 25, 2026)



(a) F (b) Open F

Figure 33: The handshapes F and OPEN F of the ASL differ only in the contact between the thumb and index finger.

cases, to a FLAT N hand. The complete handshape inventory therefore comprises 57 different handshapes.

(a) Flat H (b) Flat N

Figure 34: The handshapes FLAT H and FLAT N of the *ASL-LEX* differ only in the position of the thumb.

Moreover, in addition to the handshape at the beginning of sign production, it is advisable to also include the handshape at the end of sign production for both the dominant and non-dominant hand. This extension is particularly sensible because a large proportion of 89.68 % of the relevant information in the *ASL-LEX* can be derived from symmetry properties of the signs.

For the sake of simplicity, we begin with the derivation of the non-dominant final handshape, which can be derived entirely from the symmetry property of the sign:

For symmetric two-handed signs, the final handshape of the non-dominant hand corresponds to the final handshape of the dominant hand, whereas for asymmetric two-handed signs it corresponds to the initial handshape of the non-dominant hand, since the non-dominant hand does not move and therefore also exhibits no internal movements that could cause a change in handshape. For one-handed signs, by contrast, the final handshape of the non-dominant hand is negligible.

For the dominant hand, the following applies: if no internal movement in the form of a FLEXION CHANGE or SPREAD CHANGE occurs, the handshape remains unchanged throughout the entire sign and the final handshape of the dominant hand can be directly equated with the initial handshape. If, however, an internal movement occurs, manual annotation of the final handshape is required. This affects 562 of 2,723 signs in

the *ASL-LEX* (20.64 %).However, even in these cases it is apparent that the handshape often remains unchanged, so as a conservative assumption the final handshape of the dominant hand can likewise be equated with the initial handshape if no explicit annotation is available.

Taking the final handshape into account therefore actually yields new, independent information only in a comparatively small number of cases, which is why the number of duplicates in the vocabulary will probably be reduced only slightly. However, this gives us a more complete and consistent description of the sign structure through the additional parameters, which should increase the accuracy of our meta-classification and should lead to a more robust prediction.

## 4.2 EVALUATION OF A SYSTEM FOR CAPTURING AND ANALYZING HANDSHAPES

To capture these 57 different hand shapes, suitable hardware must now be selected. In Section 3.2, we identified EMG and MMG as two potentially suitable approaches for tracking hand shapes through muscle activity. Data gloves, whose tracking approach is based on capturing motion using IMU, flex sensors, and magnetic fields (see Section 2.3.1), offer a good alternative. Although excluded by our objectives, we also want to add a camera-based approach to this selection for comparison purposes.

This selection of hardware is now to be examined for its suitability for capturing static hand shapes in ASL. For this purpose, various preliminary studies took place in the period from November 2019 to July 2021. The outbreak of the COVID-19 pandemic at the beginning of 2020 had a strong impact on these studies and in particular made the collection of training and test data more difficult due to the applicable distancing and hygiene rules.

All preliminary studies share the fact that they evaluated the suitability of the respective hardware with regard to the capture of static hand shapes in the well-known game Rock-Paper-Scissors (RPS). We focused on RPS because at that time it was not yet clear which hand shapes would be relevant for ParaSignRec. The hand shapes in RPS are very similar to hand shapes in ASL and show a high degree of correspondence with the finger alphabet. For example, the hand shape Paper corresponds to the letter B in the finger alphabet, likewise Rock corresponds to the letters A or S depending on the execution, and Scissors to the letter V. The hand shapes of the variant with three to five hand gestures used by us are shown in Figure 35; the hand shapes of the 25-gesture variant can be found at `https://www.umop.com/rps25.htm` (last accessed on April 25, 2026). The closeness of the playful approach to the topic of our research group *Serious Games* also played a role.

To improve comparability of the approaches, in addition to using the same selection of hand shapes, care was also taken to ensure similar preprocessing and classification of the data. Thus, all trials were carried out using SVMs in the OvO procedure and prior grid-search hyperparameter optimization with tenfold cross-validation of the parameters C and $\gamma$ in the range specified in Table A.9 (column "Pre-Gridsearch"), since this has proven to be a well-suited algorithm for hand shape classification in related work (see Section 2.3.3). RBF was used as the kernel function, since the

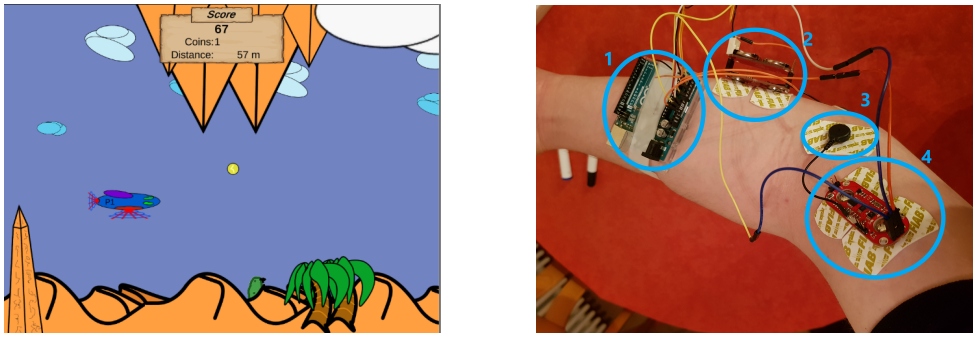
(a) Rock (b) Scissors (c) Matches (d) Paper (e) Well

Figure 35: Hand gestures of the 5-gesture variant of the rock-paper-scissors game.

number of features used in the approaches considered is overall to be regarded as low. With the exception of the preliminary EMG study, all data were scaled using a *Min-Max normalization*, whereby the minima and maxima of the training data were stored for each feature in order to scale the test data based on these values later on. For comparability of the results, accuracy (*accuracy*) was chosen, since we have an even class distribution and each class is to be prioritized to the same extent.

### 4.2.1 *Electromyography-based approach*

EMG offers the possibility of capturing muscle activity. In the context of *Serious Games*, we investigated in our publication "Flex your muscles: EMG-based serious game controls" [150] the usability of EMG as a game controller. By directly capturing the input data at the user's muscle, EMG has the advantage that – as long as the corresponding muscle is intact – it can also be used in the case of limited mobility, such as an injury or even amputation. By positioning it on the arm, EMG also promises greater and more natural freedom of movement for the hand than when using a keyboard and classic controllers. To realize an evaluable prototype, we investigated and compared six different EMG systems and developed a side-scrolling game that can optionally be controlled with a keyboard or EMG system and is shown in Figure 36.




(a) The spaceship must avoid obstacles, destroy crates, and collect coins

(b) EMG setup with Arduino microcontroller (1), MyoWare EMG sensors (2, 4), and reference diode (3)

Figure 36: Developed side-scroller with muscle-controlled spaceship [150]

In the game inspired by *Flappy Bird*, a spaceship moves through the level at constant speed and must avoid or destroy obstacles. The goal is to travel as far as possible while collecting as many coins as possible. In the process, the spaceship gradually sinks and must be readjusted by the player. Keyboard control is performed via one key to let the spaceship fly upward and another key to shoot. Alternatively, the game is controlled using *MyoWare* EMG sensors. The height of the spaceship can be controlled via the muscle tension of the *Biceps Brachii*, e.g., by flexing the forearm. Shooting can optionally happen automatically or by the player flexing the wrist and thereby activating the *Flexor Carpi Ulnaris*.

The game was evaluated empirically and could show that EMG is suitable for simple gesture control in certain scenarios and game types. With electrodes correctly attached, players typically needed less than a minute to be able to control the spaceship to a comparable extent as with keyboard-based input. An input delay compared to the keyboard was noticeable, but not large enough to disrupt the flow of the game. However, we assume that this will be a problem in faster and more reaction-intensive areas. Attaching the disposable adhesive electrodes proved to be uncomfortable and relatively time-consuming compared to conventional input devices. In addition, they require a clean, hairless contact area with the skin in order to achieve optimal results. The system proved to be susceptible to artifacts when additional electrodes were connected by cable, especially when the corresponding cables moved. Added to this are the specific disadvantages of surface EMG (*sEMG*) compared to intramuscular EMG. This is above all the limitation that with this method only muscle areas close to the skin surface can be captured, which are not covered by other muscle areas. Therefore, not all movements can be captured either. This also entails increased noise in the measurement signals.

Building on these findings, we have now investigated to what extent more complex hand gestures can be captured and differentiated using EMG. For this purpose, three RPS hand gestures (Scissors, Rock, and Paper; see Figure 35) were to be recognized based on four different configurations. The configurations differed in the number of sensors used and the muscle groups considered. A *Delsys Trigno Avanti* system with up to four wireless sEMG sensors and onboard data preprocessing was used. The sensors were attached to the *Flexor* and *Extensor Pollicis Longus* (thumb), *Extensor Digitorum* and *Flexor Digitorum Superficialis* (both fingers), with flexors responsible for flexion and extensors for extension of limbs. On the surface, only the muscle groups for all fingers (except the thumb) could be measured equally, which is why only the differentiation between the thumb and all other fingers is permissible, but this is sufficient for our strongly limited selection of hand gestures. Either three or five classes were classified, with two alternative manifestations of the classes Rock and Paper considered in each case. Classification was performed using SVMs in the configuration described above. Three different approaches to feature extraction were compared, including maximum, sum, and arithmetic mean of absolute values. The data were recorded at a data rate of 4 kHz and cleaned of a system-related offset. Using a *Sliding-Window procedure*, the arithmetic mean of one sensor was determined per second and the current window was segmented as a gesture when a defined threshold was exceeded or undershot.

The 100 recordings (five gestures each with 20 recordings) from one female participant were randomly split into training and test data and classified using tenfold cross-validation. The data were collected in 2020 in the running laboratory of the Institute of Sports Science at the Technical University of Darmstadt, where access was strongly restricted at that time due to the ongoing COVID-19 pandemic. This was also the reason why the data could only be collected from one female participant.

When distinguishing three gestures, an accuracy of approximately 96 to 100 % (depending on the sensors used) was achieved, whereas with five gestures approximately 84 to 93 % was achieved. The best results were obtained with the arithmetic mean and using all four sensors. The greatest challenge in collecting the training and test data was the placement of the sensors. This had to be validated using the manufacturer's software by checking with targeted movements whether the desired muscle activity was being detected.

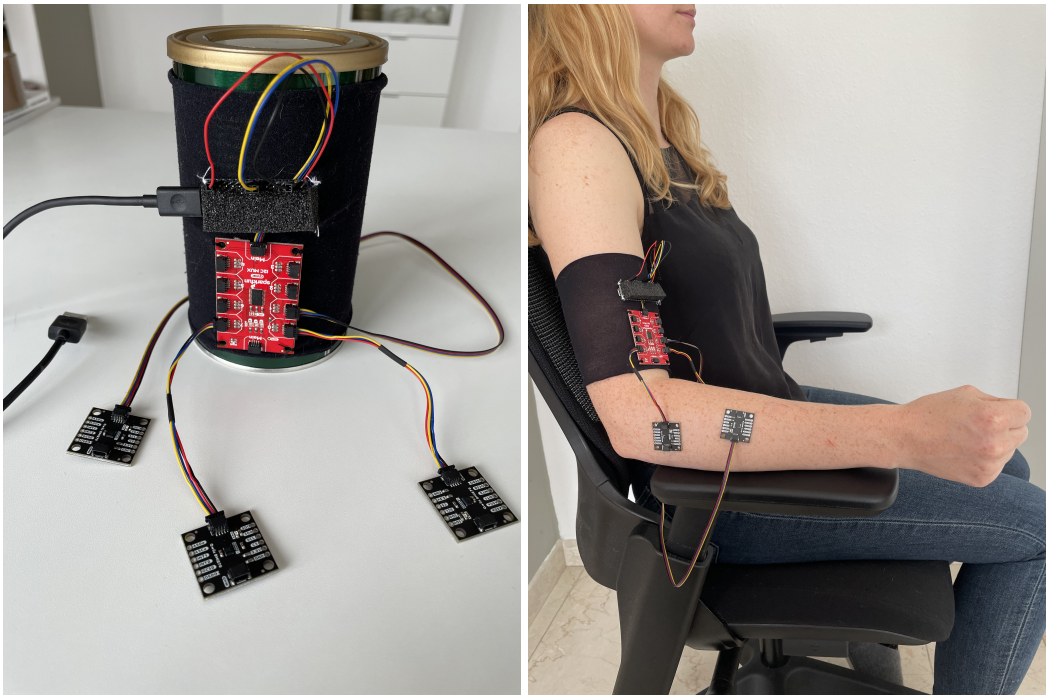
This study showed that hand gesture recognition using EMG is fundamentally possible, but it comes with many limitations. For example, a decision must be made whether to use sEMG, which is more convenient to attach and handle but can only measure the activities of superficial muscles, or whether to use an invasive method, which provides more information but can only be attached by medical personnel. In general, however, attaching the sensors is difficult and requires a certain anatomical understanding. A small and highly differentiated set of hand shapes can indeed be recognized with high accuracy, but in our experiment only movements of all fingers equally could be recognized, which can very quickly lead to limitations. Hand shapes that can only be distinguished by different finger positions therefore cannot be recognized. Due to the special circumstances, only data from one person could be recorded, i.e., the results provide no insight into the generalizability of the approach. It is to be expected that this would be significantly lower if training and test data came from different people, since each person generates different potential differences during a muscle contraction.

### 4.2.2 *Mechanomyography-based approach*

MMG is another approach for measuring muscle activity. However, it captures not the electrical signals but the mechanical signals of the muscles, e.g., muscle tremor. This makes the approach more robust against disruptive influences such as the conductivity of the skin. The approach is experimental and little researched, which is why no commercially available MMG systems exist. To evaluate the suitability of MMG as a technology for capturing hand gestures in the context of sign language recognition, we developed a MMG controller and conducted an experiment with a RPS variant consisting of five gestures (see Figure 35). The approach was inspired by Feng et al. [70], who classified various hand and finger movements of Chinese Sign Language using MMG.

For our experiment, the gestures were used for classification with SVMs with 50 repetitions each from four participants. The data were recorded using three IMUs that were attached directly to the relevant muscle groups on the participants' arms (see

Figure 37). The frequency range considered was in the range of 5 to 100 Hz, whereby according to *Nyquist-Shannon* [160, 192] the sampling rate should be at least 200 Hz, which is why the data were recorded at 400 Hz. The recordings were validated by spectrograms, which were able to show that the signals were in the desired frequency range of 5 to 100 Hz. The data were segmented based on abrupt signal changes using *Change Point Detection*. Due to (partly) varying sampling rates within the recordings, these were resampled to 333 Hz by interpolation. Subsequently, the recordings were filtered using a third-order Butterworth filter in the range of 5 to 100 Hz in order to reduce noise. For classification, three frequency-based and seven time-based features were extracted.



(a) In-house development of an MMG controller (b) Female participant with IMUs attached to the arm

Figure 37: Hardware setup of our MMG experiment

The interpolated data were randomly split into training and test data in a four-to-one ratio and were able to achieve a *Within-User* accuracy of approximately 83 %. When using unfiltered data, accuracy could be increased to up to 99 %. This suggests that the classification approach reacts more to low-frequency signals, such as skin deformation during tension and relaxation of a muscle, than to the muscle vibrations actually being investigated. Influences of the relatively rigid cables on the lightweight IMU sensors could also be observed. In addition, it became apparent that the segmentation approach could not clearly separate all gestures. As a result, some data contained parts of the transition movements between two gestures, which has a negative effect on classification performance. The features from the time and frequency domains are also not sufficient to adequately represent all information of the MMG signal.

Although the spectrograms could confirm that the developed setup is able to reliably capture MMG signals, it also became apparent here that the muscles responsible for finger movements are controlled by smaller muscle groups that are either located in the hand itself or in deeper layers of the forearm, which makes signal acquisition considerably more difficult. As a result, MMG is also limited in its recognition to a restricted set of highly differentiated hand shapes.

### 4.2.3 *Camera-based approach*

Although camera-based approaches are not to be considered in the conception of our system, we nevertheless conducted an experiment with a *Leap Motion* controller in order to obtain better comparability with the other approaches. The controller has two cameras for capturing the environment and three infrared LEDs for distance measurement and is shown in Figure 38 [9]. Three or five RPS hand shapes were recorded at 120 Hz by five people with 10 repetitions per hand shape: for the three-gesture variant, SCISSORS, ROCK, and PAPER from Figure 35; for the five-gesture variant, all gestures shown there. The preprocessing of the data and conversion of the image data into a 3D model of the hand including absolute position already takes place on the controller, except for the *Min-Max normalization*. The data from three random people were used for training, and those from the remaining two people for validation.

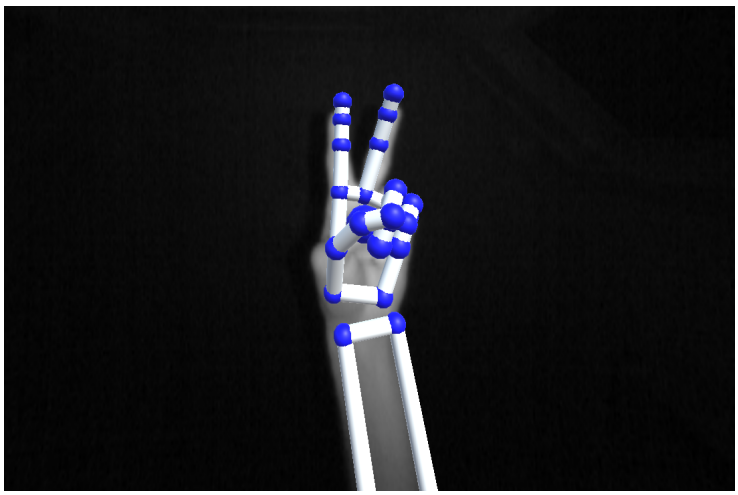


Figure 38: Data capture of a V handshape visualized using a Leap Motion controller

In the experiment, 100 % accuracy for three hand shapes and 95 to 100 % accuracy for five hand shapes could consistently be achieved, depending on the distribution of participants between test and training data. The hardware enabled free, natural movements and gesticulations due to the absence of body-worn sensors. However, the approach is only partially mobile, since the controller has a field of view of approximately 120 to 140 ° and must be placed statically at a distance of 20 to 80 cm below the hand in order to achieve the best results. The results are also strongly dependent on external influences such as light, background, or occlusions (occlusion). For example,

wearing gloves negatively affected the results, and the room had to be darkened during data collection and the participant's face should not be within the controller's field of view.

### 4.2.4 *Data glove-based approach*

As the last method within the scope of this preliminary study, the suitability of data gloves for recognizing hand shapes is to be examined. To this end, three different experiments were conducted with a commercially available *Senso Glove: DK2* data glove from the company *Senso Device Inc*[48]. The data glove was released in 2017 as a VR controller and can be operated completely wirelessly via Bluetooth Low Energy (BLE). It consists of a total of eight IMU sensors, one each attached to the middle phalanx of each finger, as well as to both phalanges of the thumb, on the back of the hand, and on the wrist. The sensors provide flexion and abduction information for each finger. Compared to the use of flex sensors – as with many other data gloves – abduction here is only estimated implicitly by the acceleration of the IMUs [34]. Due to the thumb saddle joint, the thumb has an additional degree of freedom and therefore requires an additional IMU to measure the flexion of the upper phalanx. Two magnetometers are attached to the back of the hand and the wrist and provide information on the orientation of the hand. The data are preprocessed on the motherboard attached to the back of the hand by cleaning the finger data of hand movements, i.e., no transformation from a global to a local coordinate system has to be performed. The data are provided at a data rate of 10 Hz via, among other things, a *Unity* plugin. The data glove is shown in Figure 39a. A comparison of all data gloves used in this work is shown in Table 18.

In our first experiment, which was published in the paper "Rock beats Scissor: SVM based gesture recognition with data gloves" [6], five hand shapes of the RPS game were again classified (see Figure 35). The hand shapes were recorded with ten repetitions each from eleven participants. In order to obtain as diverse, yet natural, data as possible for a more general approach, the rotation and position of the hand as well as of the body were not specified to the participants. Which hand (left or right) should be used was also not specified. Since only a single static frame was recorded as a snapshot of the hand, the participants were signaled by a countdown as well as by visual and haptic feedback when the respective hand shape was captured. From this, the 15 features shown in Table 7 were extracted. Since the finger values are independent of the hand's rotation and position due to the preprocessing of the data on the gloves, further information about the hand can be neglected.

Figure 40 shows the results of the hyperparameter optimization. It is evident that a large number of different hyperparameter combinations lead to a *Within-User* classification accuracy of 100 %. Before optimization, the data of the respective participant were removed from the dataset and then classified with these parameter combinations, which led to a *Cross-User* accuracy of 100 %. This validated that there is no *overfitting*

48 `https://senso.me` (Last accessed on January 25, 2026)


(a) Schematic structure of the Senso Glove DK2 data glove (adapted from Achenbach et al. [6])

(b) Participants test the RPS game at the Hessian Creative Economy Day 2022 in Frankfurt/Main (*Photo:* Jens Braune del Angel)

Figure 39: Hardware setup used with example application

to the training dataset and confirms the generalizability of our approach on unseen data from new users.

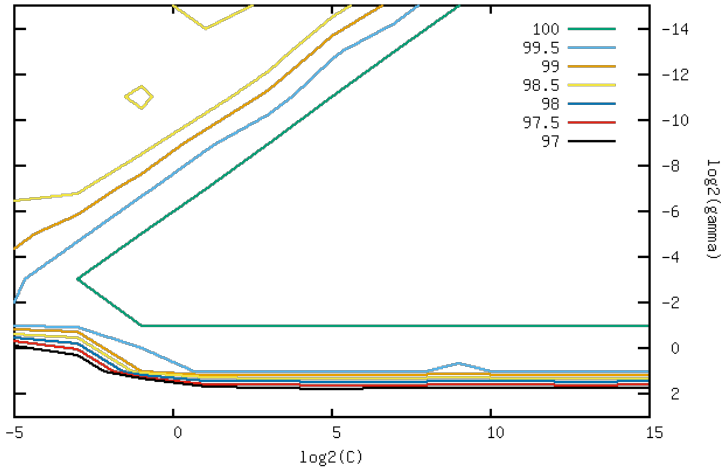



Figure 40: Results of the hyperparameter optimization as a function of C and $\gamma$

To check whether these results hold in practice, nine participants (five of whom had already participated in the previous experiment) were asked to play 20 rounds each of RPS with the data glove and an in-house development of the game (see Figure 39b) as freely as possible. The comparison between the hand shapes intended by the participant and those predicted by our model led to 98 out of 100 correctly classified hand shapes (100 %) for the participants who had already taken part in the first experiment. For the four participants who had not previously participated, 70 out

Table 7: Used sensor values of the Senso Glove SDKs (adapted from Achenbach et al. [7])

| Part of the hand | Type | Data type | Values | Used |
|---|---|---|---|---|
| Thumb | Rotation | Quaternion | 4 | ✓ |
| Thumb | Flexion | Float | 1 | ✓ |
| Thumb | Flexion/Abduction | Vector2 | 2 | ✓ |
| Index finger | Flexion/Abduction | Vector2 | 2 | ✓ |
| Middle finger | Flexion/Abduction | Vector2 | 2 | ✓ |
| Ring finger | Flexion/Abduction | Vector2 | 2 | ✓ |
| Little finger | Flexion/Abduction | Vector2 | 2 | ✓ |
| Palm | Rotation | Quaternion | 4 | (✓) |
| Palm | Position | Vector3 | 3 | ✗ |
| Palm | Orientation | Vector3 | 3 | ✗ |
| Wrist | Rotation | Quaternion | 4 | ✗ |
| Wrist | Orientation | Vector3 | 3 | ✗ |
| | | | | $\Sigma 15 - 19$ |

of 80 hand shapes (87.50 %) were correctly classified, resulting in an overall accuracy of 93.33 %.

In the follow-up paper "Paper beats Rock: Elaborating the best machine learning classifier for hand gesture recognition" [7], additional data were collected and, in addition to SVMs, other classifiers were evaluated, although only SVMs will be discussed here for better comparability. The dataset from the previous experiment was expanded by 20 participant datasets, so that 300 recordings are now available for each of the five hand shapes. In addition, another dataset with 25 different hand shapes (see `https://www.umop.com/rps25.htm`, last accessed on April 25, 2026) from nine participants with ten repetitions each was collected. Compared with the previous dataset, the orientation of the hand was now also recorded and considered in the classification, since, for example, the hand shapes Axe, Paper, and Water all have the same hand shape with different orientations of the hand. The number of features therefore increased from 15 to 19 (see Table 7).

Classification was performed with a LOSOCV, whereby the hyperparameters were re-optimized in each run. An average *Cross-User* accuracy of 96.70 % for five hand shapes and 81 % for 25 hand shapes could be achieved. If training and test data are not separated independently of the user, but instead randomly split and used in a four-to-one ratio (*Within-User*), the results improve in both datasets with the SVMs after 100 runs with fixed hyperparameters (the split of training and test data is randomized anew each time) to an average of 99.20 % and 99.50 %, respectively. Compared to the previous results, the accuracies here are each given as macro-$F_1$ score.

### 4.2.5 *Comparison and selection of hardware*

After carrying out and evaluating the various experiments, the hardware best suited for our approach is now to be selected on the basis of the requirements defined in Section 3.2.1. Table 5 shows the different requirements, divided into three categories: *i)* functional system requirements, which concern all requirements for the overall system, *ii)* technical system requirements for the respective recognition approaches and the hardware used therein, and *iii)* user-oriented system requirements for the usability of the system and the individual components . Since the first category refers to the overall approach, it is not relevant for the selection of individual components and can be neglected. We therefore focus on the technical and user-oriented requirements.

*Technical system requirements*

The technical requirements essentially comprise the following points:

PRECISE CAPTURE OF THE TARGET QUANTITY: The hardware should be able to recognize all necessary hand shapes precisely. Each of the approaches investigated was able to achieve high accuracies, albeit with a hand shape inventory of three to five hand shapes, whereas for our system we must differentiate between 57 hand shapes. Neither EMG nor MMG will be able to do this, since they can only capture the surface musculature and therefore cannot, for example, distinguish between the movements of the individual fingers, but only between the thumb and all other fingers as a whole. Data gloves and cameras, on the other hand, do not have this limitation and we assume that precise capture of far more than the five hand shapes already evaluated is possible, which has already been shown in the data glove example.

ROBUSTNESS: In addition to precise hand shape recognition, this should also be robust, i.e., not very susceptible to interference. Data gloves have the disadvantage that drift can occur due to the use of IMUs, which is why – even in our experiment – recalibration of the glove is occasionally necessary. Cameras are not affected by this, but are strongly influenced by light, background, and occlusion and have a limited field of view. EMG and MMG, on the other hand, are sensitive to movements and activities of other muscles, and EMG in particular is highly susceptible to interference from external influences such as skin properties and sweating.

HIGH SAMPLING RATE: As already explained, we require a sampling rate of at least 24 Hz in order to reliably capture human movements in the context of sign language. The requirements for EMG and MMG are higher, however, since here the sampling rate depends not on human movement but on muscle activity. For MMG, for example, we require a constant sampling rate of at least 200 Hz. Even though the version of the *Senso* data glove used by us only has a sampling rate of 10 Hz, the current version of the *Senso Glove* achieves a sampling rate of 400 Hz [189]. The *Leap Motion* controller comes standard at 120 Hz up to 240 Hz [220].

The *Delsys Trigno Avanti* system has a sampling rate of up to 4,370 Hz, depending on the configuration [57]. In our experiment with MMG, we were able to record data at a frequency of 400 Hz, but this varied partially during recording. The hardware examined is thus generally capable of capturing the data in full; only MMG exhibits a varying sampling rate, which could be problematic.

LOW LATENCY: As already explained in Section 3.2.1, we do not have the means to determine the exact latencies of the hardware examined. We therefore consider the manufacturers' specifications, which is not possible in the case of the in-house MMG controller. For the *Delsys Trigno Avanti* system, we could only find the specification that the "Inter-Sensor Delay" is less than one sampling period and the "Intra-Channel Delay" is less than one to two sampling periods. At a sampling rate of 4,370 Hz, this would be approximately 0.46 to 0.69 ms, although not when using Bluetooth ("the Bluetooth/BLE protocol does not guarantee latency") [57]. The current *Senso Glove* has a latency of 15 ms [189], whereas the *Leap Motion* controller has a camera latency of approximately 8 ms and the internal software processing requires less than one millisecond [220]. All latencies are below the required 41.67 ms and are therefore well suited for our approach.

NO DATA POSTPROCESSING NECESSARY: With the exception of MMG, the hardware used were commercial products with their own Software Development Kit (SDK) and onboard data preprocessing. Apart from normalization of the data, no further preprocessing of the data was therefore necessary. With MMG, the data also had to be filtered and resampled due to varying sampling frequency. In addition, an elaborate feature extraction of three frequency-based and seven time-based features took place. With EMG, statistical features were used, which are much easier to compute. With the data glove and camera, the normalized values were used directly as features.

*User-oriented system requirements*

The user-oriented requirements essentially comprise the following points:

MOBILE: All hardware examined was either wireless or could at least be attached to the body, so that use on the move would be conceivable. Only the use of a camera-based system proves difficult, since the camera would ideally have to be placed opposite the signer in order to capture the entire signing space, but then would only be partially mobile. For example, the investigated *Leap Motion* controller would have to be placed below the hand at a distance of 20 to 80 cm, which is almost impossible with a moving hand, especially when performing two-handed signs.

FREEDOM OF MOVEMENT: Freedom of movement was also mostly good; only when using data gloves can the placement of the sensors on the fingers negatively affect certain executions (e.g., crossing the fingers for R).

RESPECT FOR PRIVACY: All sensor technology examined with the exception of camera-based approaches measures the quantity to be captured of the person wearing it. Cameras also record the entire surroundings, which in everyday use is difficult to reconcile with the privacy of nearby persons.

USER-FRIENDLINESS AND AVAILABILITY: With the exception of MMG, the hardware examined was commercial products, which is why we assume a certain degree of user-friendliness and availability. Only the attachment of the EMG sensors can be regarded as user-unfriendly, since it is difficult to perform without prior knowledge and the help of another person. The same applies to MMG, which as an in-house development is also poorly available.

*Other relevant aspects*

In addition to the advantages and disadvantages of the individual approaches already mentioned, there are further aspects that could be relevant for our system and are therefore added here. For example, data gloves and camera-based approaches offer the advantage of being able to capture not only the hand shape but also movements and relative positions. In the example of the *Leap Motion* controller, absolute coordinates could even be returned, albeit due to the stationary use of the controller. MMG, on the other hand, has barely been researched yet and is to be regarded as very experimental.

*Selection of hardware*

The results of our investigation are shown in Table 8. The most important aspect is reliable and precise hand shape recognition, which is why the data glove is the clear preference, since cameras were excluded due to our objectives. EMG and MMG both have the serious disadvantage that only the surface musculature can be captured, which leads to considerable limitations in the hand shapes that can be captured. An invasive version of EMGs could also capture deeper musculature, but would meet with significantly lower user acceptance and user-friendliness would also decrease due to the elaborate setup.

Even taking the other requirements into account, it is clear that the data glove is the best alternative among the hardware examined. Only in the areas of robustness and freedom of movement were the requirements not fully satisfactorily met. By using a newer generation of data gloves, we hope, on the one hand, for less drift in the sensor data and, on the other hand, for a more compact sensor design to increase the freedom of movement of the individual fingers. At the time of writing this work, the *Metagloves Pro* by the *MANUS Technology Group* had just been launched as the successor to the *Prime X Metagloves* and *Quantum Metagloves* used by us later in this work[49]. These come with integrated sensors that promise even greater freedom of movement. In addition, according to the manufacturer, the sensor data are free of drift and do not require recalibration of the gloves.

49 `https://www.manus-meta.com/products/metagloves-pro` (Last accessed on January 25, 2026)

Table 8: Comparison of hardware approaches for hand gesture recognition

| | **Data glove** | **Camera-based** | **EMG** | **MMG** |
|---|---|---|---|---|
| **Technical system requirements** | | | | |
| Hand gesture recognition | Good | Good | Bad | Bad |
| Robustness | Medium | Medium | Bad | Medium |
| Sampling frequency | Good | Good | Good | Medium |
| Latency | Good | Good | Good | ? |
| Data processing | Good | Good | Medium | Bad |
| **User-oriented system requirements** | | | | |
| Mobility | Good | Bad | Good | Good |
| Freedom of movement | Medium | Good | Good | Good |
| Privacy | Good | Bad | Good | Good |
| Usability | Good | Good | Bad | Bad |
| Availability | Good | Good | Good | Bad |
| Further aspects | Movement and relative position sensing possible | Movement and absolute position sensing possible | | Experimental |

## 4.3 EVALUATION OF A CLASSIFICATION METHOD FOR HANDSHAPE RECOGNITION

After selecting data gloves for capturing static hand shapes, suitable classification methods must now be identified in order to convert the recorded data into hand shapes. In addition to high classification accuracy, the classification time must also be considered so that the overall approach remains capable of real-time operation. Since our models are trained offline, the training duration is not relevant for our project.

In our paper "Paper beats Rock: Elaborating the best machine learning classifier for hand gesture recognition" [7] we investigated, in addition to SVMs[50], seven other

50 `https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html` (Last accessed on January 25, 2026)

classification methods, including the conventional models DTs[51], RF[52], kNN[53], LR[54], NB[55], perceptron[56] and the deep learning model FNN[57]. Since DTs, kNN, NB and perceptron could not deliver convincing results, we want to focus on the remaining four classifiers. This is the same experiment under the same conditions, which has already been explained in Section 4.2.4, although this time the data of one participant for five hand shapes were removed due to poor quality. Table 9 shows the achieved accuracies and the average classification duration required for them.

Table 9: Cross-user accuracies (LOSOCV) and classification duration of a participant dataset for various ML models with 5 classes (29 participant datasets) and 25 classes (9 participant datasets, after cleaning erroneous participant data) (adapted from Achenbach et al. [7])

| **Model** | $F_1$**-score** (%) | | | | **Duration** (ms) | | | |
|---|---|---|---|---|---|---|---|---|
| | MW | Std. dev. | Min. | Max. | MW | Std. dev. | Min. | Max. |
| **5 classes** | | | | | | | | |
| SVM | 98.05 | 2.63 | 89.96 | 100 | 2.22 | 0.51 | 1.67 | 4.54 |
| RF | 98.47 | 2.41 | 89.66 | 100 | 8.48 | 9.19 | 1.08 | 29.51 |
| LR | 98.97 | 1.89 | 92.08 | 100 | 0.08 | 0.02 | 0.05 | 0.11 |
| FNN | 98.39 | 2.68 | 90 | 100 | 0.95 | 0.94 | 0.09 | 3.39 |
| **25 classes** | | | | | | | | |
| SVM | 81.05 | 6.62 | 74.84 | 94.27 | 270.30 | 78.25 | 159.28 | 359.50 |
| RF | 82.42 | 6 | 72.37 | 93.19 | 20.55 | 10.94 | 6.86 | 33.11 |
| LR | 81.89 | 5.72 | 75.96 | 90.21 | 0.12 | 0.03 | 0.08 | 0.16 |
| FNN | 80.53 | 7.59 | 71.58 | 95.75 | 6.77 | 4.02 | 1.91 | 13.98 |

Mean value (MW), standard deviation (Std. dev.), minimum (Min.), maximum (Max.)

It can be seen that for five different RPS hand shapes from 29 participants, all models were able to sufficiently distinguish the data with an $F_1$ score of 98.05 to 98.97 %, provided that it was a classification problem with fewer classes (5) than features (15).

51 https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html (Last accessed on January 25, 2026)

52 https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html (Last accessed on January 25, 2026)

53 https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html (Last accessed on January 25, 2026)

54 https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html (Last accessed on January 25, 2026)

55 https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.GaussianNB.html (Last accessed on January 25, 2026)

56 https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html (Last accessed on January 25, 2026)

57 https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html (Last accessed on January 25, 2026)

With at least one combination of training and test data, all classifiers achieve a result of 100 %. In the more advanced experiment with 25 different RPS hand shapes from nine participants, the influence of the higher number of classes (25) and features (19) becomes apparent. The models now achieve an average of 80.50 to 82.40 %. Comparing the results, one finds that RF and LR deliver very good results for both datasets with an average of 90.45 %, closely followed by SVMs (89.50 %) and FNN (89.45 %).

In contrast, when looking at the *Within-User* probabilities in Table 10 for a random 80 to 20 split of the training data of all participants (here, however, again with the previously excluded faulty participant data) and 100 repetitions, the models now achieve 98.40 to 99.20 % and 96.50 to 99.50 %, respectively, with SVMs delivering the best results. RF and FNN also deliver very good results, while LR shows a drop in classification accuracy, especially for the higher number of hand shapes. The fact that the classification results for 25 hand shapes are higher than for five hand shapes may be due, on the one hand, to the data quality and the number of participants, but also to the chosen hyperparameters, which were selected here before the experiment and not optimized.

Table 10: Within-User Accuracies (80:20 split, 100 repetitions) of different ML models for 5 classes (29 participant datasets) and 25 classes (10 participant datasets) (adapted from Achenbach et al. [7])

| **Model** | **5 Handshapes** (%) | | | | **25 Handshapes** (%) | | | |
|---|---|---|---|---|---|---|---|---|
| | MW | Stdabw. | Min. | Max. | MW | Stdabw. | Min. | Max. |
| SVM | 99.20 | 0.56 | 97.30 | 100 | 99.50 | 0.31 | 98.40 | 100 |
| RF | 98.70 | 0.59 | 97.30 | 100 | 99.30 | 0.38 | 98 | 100 |
| LR | 98.40 | 0.67 | 96.30 | 99.70 | 96.50 | 0.88 | 94.20 | 98.90 |
| FNN | 99.10 | 0.59 | 97.70 | 100 | 98.80 | 0.56 | 96.90 | 100 |

Mean value (MW), standard deviation (Stdabw.), minimum (Min.), maximum (Max.)

With regard to the performance of the individual approaches in Table 9, significant differences can be observed between the individual models in classification times. Since these depend heavily on the hardware used[58] and are strongly influenced by memory and processor utilization, they serve only as a rough guide in this case. The average classification times for all runs are given. Since the models were evaluated using the LOSOCV procedure, the number of runs corresponds to the number of participants, which is why the statistical significance with only nine participants is lower than with 29 participants. In any case, LR was the fastest classifier, whereas SVMs and RF were the slowest. It was noticeable that SVMs exhibited the greatest difficulties with regard to classification time as the number of classes increased, which, however, makes sense given the use of the OvO approach. Thus, the number of

58 An *Apple MacBook Pro* 14“, 2021 (`https://support.apple.com/de-de/111902`, last accessed on January 25, 2026), with an *Apple M1 Pro* processor (10-core CPU with 8 performance cores and 2 efficiency cores, 16-core GPU, 16-core Neural Engine, and 200 GB/s memory bandwidth) and 16 GB of shared memory, was used.

SVMs increases from ten to 300[59] when the number of hand shapes is quintupled. When considering the times, however, it must not be overlooked that these are the classification times of an entire participant dataset, i.e., 50 or 250 hand shapes, rather than the classification time of a single hand shape, as required for later sign recognition.

Overall, it can be said that RF and LR provide accurate results regardless of the number of classes, but LR requires significantly less time for prediction than RF. SVMs, on the other hand, also deliver accurate results, but are significantly slower compared to the aforementioned methods. In *Within-User* approaches, they achieve the highest accuracies, whereas RF also delivers very good results here and only slight weaknesses can be observed for LR. The FNN did not perform significantly better than the conventional models in comparison, even though it was partially able to classify faster. However, its design and optimization required substantially more effort.

## 4.4 EVALUATION OF DATA PROCESSING FOR CLASSIFICATION OPTIMIZATION

Building on these results, the thesis "Give Me a Sign: Using Data Gloves for Static Hand-Shape Recognition" [5] now investigated methods of data preprocessing and postprocessing in order to optimize classification accuracies and also classification times. The motivation for data preprocessing is to highlight important information in the available data while simultaneously removing some of the potentially present redundant or misleading data [172]. Among the data preprocessing methods, we consider an approach for the machine-based identification of anomalous data points (*Outlier Detection*), as well as an approach for the machine-based selection of the most relevant features of hand shapes (*Feature Selection*). An experimental attempt to artificially increase the amount of recorded data (*Data Augmentation*) in order to improve the generalizability of the approach was also examined.

In comparison to our previous experiments, the focus was now shifted away from RPS hand shapes toward actual hand shapes from the vocabulary of ASL. The American finger alphabet and the *ASL-LEX* served as the basis. From these two sources, a total of 58 different hand shapes were identified.

A new generation of data gloves was used for data acquisition in order to obtain more reliable data. The subsequent classification process was carried out with eight different configurations, each preceded by hyperparameter optimization, using SVM, RF, and LR from the previous experiment. These qualified best for our project in terms of accuracy and classification time. Due to the high optimization effort required for comparable results, the use of FNN was omitted in comparison to the previous experiment. For data postprocessing, a VL2 meta-classifier as an aggregative approach (AMK) for fusing the three classification results was evaluated.

59 $\frac{k(k-1)}{2} = \frac{5(5-1)}{2} = 10 \ll = \frac{25(25-1)}{2} = 300$

### 4.4.1 *Data Acquisition*

The different data preprocessing methods were evaluated on the basis of two different datasets: *i)* 27 hand shapes, consisting of 21 hand shapes of the ASL finger alphabet and six additional hand shapes, in order to also be able to sign the digits from zero to nine, and *ii)* 56 different hand shapes of the *ASL-LEX*, since at the time of this investigation it was still mistakenly assumed that FLAT H and FLAT N are identical hand shapes. Together, both datasets comprise a total of 58 different hand shapes, whereby all hand shapes of the finger alphabet except M and N are also part of the *ASL-LEX*.

Used were *Manus Prime X*[60] data gloves from the company *Manus Meta*, whose structure is shown in Figure 41. As with the previous data glove, this one is also based on the use of one 9-DoF IMU per finger, but additionally incorporates a 2D flex sensor. This now makes it possible to obtain information not only about flexion values (bending), but also about the abduction (spreading) of the fingers. An additional 9-DoF IMU for determining the orientation of the hand and as a reference point is attached to the back of the hand. The captured sensor values are fused and preprocessed internally. The accuracy of the values is stated as $\pm 2.5$ degrees. The data glove delivers the data with a latency of less than 5 ms and a data rate of 90 Hz. All relevant details of the data glove can also be found in Table 18.


(a) Green circles indicate the IMUs (b) Brown stripes indicate the flex sensors

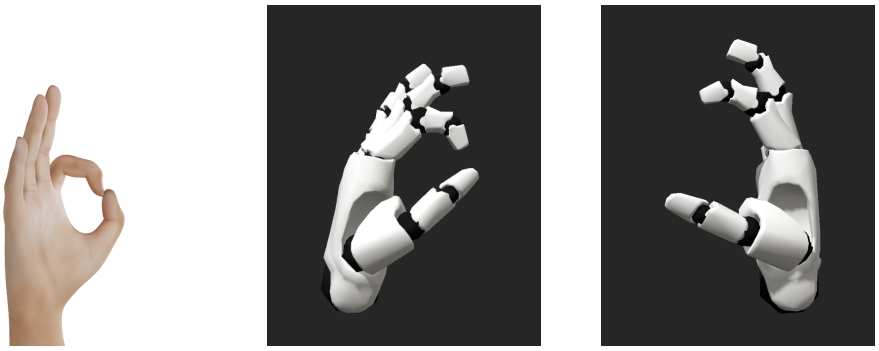
Figure 41: Manus Prime X data glove including sensor setup [5]

The hand shapes were recorded from 20 participants, each of the 58 hand gestures with three repetitions, resulting in a total of 3,480 recordings. For maximum mobility of the hands, the vibration motors on the gloves were removed. Before each experiment, the gloves were calibrated with the associated *Manus Core* SDK in order to eliminate

60 https://assets.website-files.com/61de97d15a7bb6441d9565c0/6273d8a29f527084a4fe036e_Datasheet%20-%20Prime%20X%20Haptics%20Glove.pdf (Last accessed on May 27, 2025)

possible drift in the IMUs and to ensure that differing hand sizes of the participants did not adversely affect the results.

Each participant sat at a table and was shown a sketch of the hand shape to be performed. By pressing the enter key, recording began, after which the participant had three seconds to perform the hand shape and then hold it for two seconds before returning the hand to the starting position on the table. This procedure was repeated three times for each hand gesture. Throughout the experiment, the participants were observed to ensure that the recording was carried out correctly. Incorrect recordings were repeated at the end of the experiment.



(a) Hand shape F

(b) Visualization of left data glove

(c) Visualization of right data glove

Figure 42: Quality difference between the left and right Manus Prime X data glove using the example hand shape F

Although all participants were right-handed, the recordings were performed exclusively with the left glove. This was because the hand shapes recorded with the right data glove exhibited a considerable quality difference compared with the hand shapes of the left hand, even after repeated calibration. An example using the hand shape F can be seen in Figure 42.

The actual static hand shape was extracted as a single keyframe from the middle of the two-second *Hold* phase. By using the new data glove, 20 rather than 15 features are now available for describing the hand shape – regardless of orientation. These consist exclusively of three *Stretch* and one *Spread* values per finger and thumb, for describing flexions of the individual finger segments and spreads of the fingers relative to one another, and are shown in Table 11.

### 4.4.2 *Data Processing Process*

The experiment was conducted on an *Apple MacBook Pro*[61] was conducted. After the hand shapes were recorded using the data glove, they were then scaled (cf. Equation 1) in order to obtain a normal distribution with mean zero and unit variance. To ensure

61 Apple MacBook Pro 16“, 2021 (`https://support.apple.com/kb/SP858`, last accessed on January 25, 2026), with an *Apple M1 Max* processor (10-Core CPU with 8 performance cores and 2 efficiency cores, 32-Core GPU, 16-Core Neural Engine, and 400 GB/s memory bandwidth) and 32 GB shared memory

Table 11: Used sensor values of the Manus Core SDKs for the Manus Prime X and Manus Quantum Metagloves data gloves, as well as the frequency of values discarded by Feature Selection (Feat.) (adapted from Achenbach et al. [5])

| Value | Finger | Joint | SDK Range | | Degree Range | | Discarded |
|---|---|---|---|---|---|---|---|
| | | | Min | Max | Min | Max | |
| Spread | Thumb | CMC | 0.0 | 1.0 | 5° | 50° | 29.38 % |
| | Index | MCP | −1.0 | 1.0 | −20° | 20° | 0 % |
| | Middle | MCP | −1.0 | 1.0 | −20° | 20° | 1.88 % |
| | Ring | MCP | −1.0 | 1.0 | −20° | 20° | 41.25 % |
| | Pinky | MCP | −1.0 | 1.0 | −20° | 20° | 15 % |
| Stretch | Thumb | CMC | 0.0 | 1.0 | −20° | 25° | 0 % |
| | Thumb | MCP | 0.0 | 1.0 | −20° | 45° | 0 % |
| | Thumb | IP | 0.0 | 1.0 | −15° | 80° | 0 % |
| | Index | MCP | 0.0 | 1.0 | 0° | 80° | 0 % |
| | Index | PIP | 0.0 | 1.0 | 0° | 100° | 6.25 % |
| | Index | DIP | 0.0 | 1.0 | 0° | 90° | 5.62 % |
| | Middle | MCP | 0.0 | 1.0 | 0° | 80° | 1.88 % |
| | Middle | PIP | 0.0 | 1.0 | 0° | 100° | 9.38 % |
| | Middle | DIP | 0.0 | 1.0 | 0° | 90° | 10 % |
| | Ring | MCP | 0.0 | 1.0 | 0° | 80° | 3.12 % |
| | Ring | PIP | 0.0 | 1.0 | 0° | 100° | 22.50 % |
| | Ring | DIP | 0.0 | 1.0 | 0° | 90° | 13.75 % |
| | Pinky | MCP | 0.0 | 1.0 | 0° | 80° | 5.62 % |
| | Pinky | PIP | 0.0 | 1.0 | 0° | 100° | 18.75 % |
| | Pinky | DIP | 0.0 | 1.0 | 0° | 90° | 22.50 % |

high-quality data, an *Outlier Detecion* using DBSCAN was performed in the next step. In order to reduce the amount of training data and increase the speed of training and classification, *Feature Selection* in the form of a GA was subsequently applied. Both methods are described in Section 2.3.1.

Based on the outlier-cleaned data reduced to its essential features, the hyperparameters of the individual ML models were now optimized. To this end, a grid search was first carried out over the original parameter range. Both datasets were optimized for each combination of our data preprocessing methods using tenfold cross-validation, resulting in a total of 8 different hyperparameter configurations. Based on these results, a narrower but more precise search range was defined. Both grid-search ranges,

as well as the most frequently determined hyperparameter combinations in the later course, can be found in Section A.3, Table A.9, and Table A.10.

In the following, both datasets were evaluated for each combination of data preprocessing methods in the LOSOCV procedure. In each of the 320 runs, a grid search with fivefold *HalvingGridSearch*[62] cross-validation within the newly defined search range was first applied to the training dataset, consisting of 19 participant datasets. The algorithm is based on the *Successive Halving* procedures and iteratively selects the best hyperparameter combinations. The parameter combinations are initially evaluated with a small amount of resources and the most promising parameter combinations are then retained. Further resources are now added step by step and the process is repeated until the most promising parameter combination has been determined. This procedure enables comparable results to conventional grid-search cross-validation, but can significantly reduce computational effort.

For data postprocessing in this experiment, a VL2 meta-classifier was used. It sits on the eponymous second level (Level 2) and determines its own prediction by weighting the results of the first-level classifiers (cf. Section 2.3.4), in our case SVM, RF, and LR.

To compare the performance of the classifiers, the mean and standard deviation of each LOSOCV cross-validation were stored for both the accuracy and the duration of classifying a participant dataset. Due to the homogeneous class distribution and prioritization, other metrics such as the F-score were neglected.

The pipeline is generic and can be applied to any type of static data. However, it is recommended to adapt the various parameters of the pipeline, such as the hyperparameters of the classifiers, to the new data.

### 4.4.3 *Results*

Table 12 and Table 13 show the average accuracies and classification times of a participant dataset in comparison. The *Score* gives the mean of all four individual results. In order to be able to compare the classification times of the different classifiers with one another, these were first normalized for each classifier as follows:

$$s = \frac{t_{max} - t}{t_{max} - t_{min}} \tag{9}$$

This yields a *Score* of 100 % for the shortest time and 0 % for the longest. The classification times of LR are so low that we cannot take them into account, as they are too strongly influenced by external factors such as processor runtime fluctuations.

Regardless of the data preprocessing methods used, 27 hand shapes can be classified with an accuracy of 89.51 to 91.91 %, and 56 hand shapes with an accuracy of 85.39 to 87.50 %. The VL2 meta-classifier consistently achieved the highest average accuracy in both datasets (91.50 % for 27 hand shapes and 86.59 % for 56 hand shapes). With respect to the different data preprocessing methods, the highest accuracy values were achieved without the use of data preprocessing (with the exception of normaliza-

62 `https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.HalvingGridSearchCV.html` (Last accessed on January 25, 2026)

Table 12: Cross-User accuracies (LOSOCV) of different ML models considering various data preprocessing methods for 27 and 56 hand shapes (adapted from Achenbach et al. [5]). The score corresponds to the mean of all four classification accuracies.

| **Classes** | **Preprocessing** | | **Models** (%) | | | | **Score** (%) |
|---|---|---|---|---|---|---|---|
| | Out. | Feat. | SVM | RF | LR | VL2 | |
| 27 | ✗ | ✗ | 90.80 | 91.23 | 89.63 | 91.60 | 90.82 |
| | ✗ | ✓ | 90.37 | 91.05 | 89.51 | 91.85 | 90.69 |
| | ✓ | ✗ | 90.80 | 90.43 | 89.81 | 91.91 | 90.74 |
| | ✓ | ✓ | 90.43 | 90.37 | 89.81 | 91.11 | 90.43 |
| 56 | ✗ | ✗ | 86.10 | 87.14 | 85.42 | 87.44 | 86.53 |
| | ✗ | ✓ | 85.71 | 86.61 | 85.63 | 87.11 | 86.26 |
| | ✓ | ✗ | 85.68 | 86.96 | 85.39 | 87.50 | 86.38 |
| | ✓ | ✓ | 85.54 | 86.46 | 85.39 | 87.11 | 86.12 |

Outlier Detection (Out.), Feature Selection (Feat.)

tion), and the lowest accuracy values were obtained when using *Outlier Detection* and *Feature Selection*. The values are in the range of 90.43 to 90.82 % for 27 hand shapes and in the range of 86.12 to 86.53 % for 56 hand shapes. Only the VL2 meta-classifier can achieve slightly more accurate results for both datasets when using *Outlier Detection* than without data preprocessing.

The results should be interpreted with caution, as the differences are marginal, which may also be attributable to the choice of hyperparameters.

Table 13: Classification times of different ML models when classifying a participant dataset, considering different data preprocessing methods for 27 and 56 classes (adapted from Achenbach et al. [5]). The score takes into account normalized classification times, except LR.

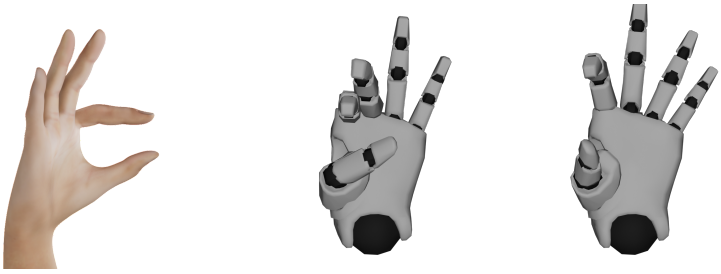
| **Classes** | **Preprocessing** | | **Models** (ms) | | | | **Score** (%) |
|---|---|---|---|---|---|---|---|
| | Out. | Feat. | SVM | RF | LR | VL2 | |
| 27 | ✗ | ✗ | 2.78 | 16.58 | 0.11 | 21.10 | 36.48 |
| | ✗ | ✓ | 2.64 | 14.44 | 0.13 | 18.20 | 84.37 |
| | ✓ | ✗ | 2.68 | 16.32 | 0.13 | 20.26 | 55.66 |
| | ✓ | ✓ | 2.51 | 20.14 | 0.12 | 23.66 | 33.33 |
| 56 | ✗ | ✗ | 18.25 | 44.27 | 0.37 | 72.97 | 0.38 |
| | ✗ | ✓ | 17.59 | 38.52 | 0.44 | 62.97 | 95.32 |
| | ✓ | ✗ | 18.25 | 40.35 | 0.34 | 67.49 | 41.02 |
| | ✓ | ✓ | 17.49 | 42.68 | 0.34 | 70.42 | 51.04 |

Outlier Detection (Out.), Feature Selection (Feat.)

The differences in classification times are generally more significant than those in accuracy. The time difference was measured immediately before and after the call to the classifiers' *predict* function and comprises the classification of an entire participant dataset, i.e., up to 168 recordings (corresponding to 56 hand shapes with three repetitions), before any *Outlier Detection* was carried out. It is obvious that classification times also increase strongly with the number of hand shapes. In the case of the VL2 meta-classifier, the classification times of the Level 1 classifiers are included, which is why it understandably takes the longest time for classification (18.20 to 23.66 ms for 27 hand shapes and 62.97 to 72.97 ms for 56 hand shapes). The use of *Outlier Detection* and/or *Feature Selection* leads in most cases to an improvement in classification time. The fastest runs were those performed with only *Feature Selection*.

With regard to *Outlier Detection*, we found that detecting only very few outliers led to the best results. The DBSCAN algorithm counts all data points as outliers that are farther than the maximum distance $\varepsilon$ from data points within a cluster (*CorePoints*). The Euclidean distance $d$ of a data point from the origin is determined by $d = \sqrt{\sum_{i=1}^{n} x_i^2}$, with $n$ as the number of features $x_i$. Using standardized features (mean $\mu = 0$ and unit variance $\sigma = 1$), the origin of the coordinate system lies at the center of the cluster. If one assumes that the variance of new data is distributed evenly across all features, this means $\sigma_1 = \sigma_2 = \cdots = \sigma_n$. The maximum distance of a data point from the center of the cluster is now to be calculated so that this data point is still counted as part of the cluster. If this distance is represented by a hypothetical data point that lies exactly one standard deviation from the mean in each dimension, then this point has a distance of $d_{max} = \sqrt{\sum_{i=1}^{n} (\mu_i - \sigma_i)^2} = \sqrt{\sum_{i=1}^{n} (-1)^2} = \sqrt{n}$ from the center of the cluster. With 20 features, a data point should therefore not be farther than $\sqrt{20} =\approx 4.47$ from the center of the cluster, which is why we use this value as the DBSCAN parameter value for $\varepsilon$. The *minPoints* parameter was set to 51, since there were a total of 57 samples for each gesture in the training set (19 participants with three repetitions) and at most 10 % of the data were to be outliers. The choice of these parameters led to an average of 3.25 outliers (0.21 %) per run in 1,539 samples (27 hand shapes) and 5.8 outliers (0.18 %) per run in 3,192 samples (56 hand shapes). Figure 43 shows an example of an outlier compared with the originally intended hand shape OPEN F and a hand shape not classified as an outlier. Both the thumb and the middle finger do not correspond to the intended hand shape.

Table 11 shows the relative number of features discarded by *Feature Selection*. It is worth mentioning that, on average, 2.069 features were discarded per run. The range extended from no discarded feature (31.25 % of the runs) to six discarded features (3.13 % of the runs). The table makes it clear that the thumb, index finger, and middle finger are the fingers most important for classification. Not a single one of the three *Stretch* values of the thumb was discarded, and features of the index finger were discarded least often in total and, when they were, only information of the upper extremities (DIP and PIP). Looking at the *Stretch* values of the joints, the most informative joint for classification is the MCP. As already mentioned, DIP and PIP are discarded most often, with usually only one of the two values being discarded. This is because it is anatomically not possible to move the upper phalanx (DIP) independently



(a) Visualization (b) Outlier (c) Inlier

Figure 43: Outlier (curved thumb and middle finger) of the hand shape *Open F* compared with an inlier and a visualization of the same hand shape

of the middle phalanx (PIP) without external influence [164]. For the *Spread* values, those of the ring finger are discarded most often (more than 40 %), since it hardly moves independently of the other fingers. If, for example, one places the hand flat on the table, it is much more difficult to raise the ring finger than the other fingers. This dependency was apparently recognized by our *Feature Selection*.

### 4.4.4 *Discussion*

In general, the achieved accuracy values are at a comparable level for all classifiers. By aggregating the results of all classifiers, the VL2 meta-classifier can almost exclusively achieve the highest accuracy values. It thus combines the advantages of the other classifiers, but consequently also requires the longest classification time. However, this is relativized by the fact that in a real-time scenario normally only one hand shape needs to be recognized at a time. In the experiment conducted, however, the classification time was given for the classification of up to 82 or 168 hand shapes. The VL2 meta-classifier therefore represents a useful addition to the classifiers already evaluated.

With regard to data preprocessing, the use of *Feature Selection* is advantageous for time-critical applications, as it delivers the fastest classifications. In the case of the VL2 meta-classifier, however, the highest accuracies for both datasets were achieved with *Outlier Detection*, which was also able to achieve the second-highest accuracies on average. However, these advantages cannot be summed, which is why it is not necessarily sensible to use both methods simultaneously. All achieved accuracies were nevertheless close to one another, which is why the temporal aspect should play a greater role in the selection of data preprocessing methods here.

The study also examines the individual predictions. It is noticeable that there are conspicuously many confusions or incorrect predictions in three areas:

The first area consists of signs that differ essentially in the position of the thumb, such as M, N, T, as shown in Figure 44, but also S and Closed E. On closer inspection of the visualized data, it becomes apparent that the position of the thumb is not captured precisely enough by the data glove. This weakness is also evident in other IMU-based data gloves [7].

(a) M, participant 5 (b) N, participant 5 (c) T, participant 2

Figure 44: Incorrect recordings of the hand shapes M, N and T

The second area consists of signs that differ essentially through a change in abduction, that is, the *Spread* values. These include, for example, R, H, and V, but also 4 and Closed B.

The third area includes signs that are differentiated by the flexion of the fingers (*Stretch* values). This includes in particular *Curved* and *Bent* flexions, as in Curved/Bent L and Curved/Bent 1. The differences between the two flexions are often marginal, and the data gloves used also do not have sensors for each finger segment, so that the differences often cannot be captured accurately.

Since the data acquisition was monitored, i.e., it was verified whether the hand shapes were performed correctly, the other two types of errors listed are mainly attributable to the data gloves or the classification methods. Based on our results and the visualizations shown (see Figure 44), however, it is clear that the limiting factor is the hardware used and not the classifiers.

***Note** on Data Augmentation*

In the original experiment, an experimental approach to *Data Augmentation*, adapted from Liu and Ostadabbas [129], was also investigated. The approach generates synthetic data through targeted transformations of existing data. This is intended to help reduce *Overfitting* and improve the robustness as well as the generalizability of the models. In the experiment, the minimum and maximum value were determined for each feature, and this value range was extended in both directions by half a standard deviation in order to further increase generalizability. Within this range, new synthetic data points were then generated at random and the amount of existing data was doubled. Figure 45 shows two synthetically generated Horns hand shapes as an example, together with an original recording for comparison. It is evident that the technique is suitable for artificially expanding the dataset.

With regard to the achieved accuracies and classification times, however, the investigated approach was not convincing. Generalizability could also not be increased, which was particularly noticeable for LR. The runs with *Data Augmentation* performed worst in terms of both accuracies and classification times, which is why we do not discuss this approach further in this thesis. One reason for this could be that the

(a) Original recording (b) Augmented hand shape 1 (c) Augmented hand shape 2

Figure 45: Two examples of data augmentation using the hand shape HORNS

recorded data already contain sufficient information for the classifiers to differentiate the hand shapes. It is assumed that the method achieves better results on dynamic data and in the field of *Deep Learning*, where approaches to *Data Augmentation* are already in use [221].

## 4.5 FINALIZATION OF THE OVERALL SYSTEM FOR HANDSHAPE RECOGNITION

Als last step, the final classification process for the recognition of static handshapes of ASL for our system is now to be defined and evaluated on the basis of the insights gained. In doing so, we orient ourselves to the results of the dataset with 56 handshapes, as this is almost identical to our final handshape dataset.

A VL2 meta-classifier is to be used as the classifier, with SVM, RF and LR as first-level classifiers, since this was able to achieve higher accuracies in the previous experiment than the individual classifiers. We consider the higher classification time to be uncritical, as it referred to the classification of an entire participant dataset with up to 168 handshapes, whereas our system will classify only two handshapes (dominant and non-dominant hand) over one or fewer frames at the same time.

We expect a further increase in overall accuracy through the use of a new data glove with alternative sensor technology, which is more robust than the drift-prone IMU. The *Manus Quantum Metagloves*[63] data gloves from *Manus Meta* represent such an alternative. They are based on the measurement of electromagnetic fields (EMF) and largely dispense with the use of IMUs or flex sensors. Only on the back of the hand is there a 9-DoF IMU for measuring absolute hand orientation, rotation, and acceleration, together with a 3-DoF EMF coil that serves as the source. Another 3-DoF EMF coil is attached to each fingertip, as well as to the thumb. These act as detectors and capture the three-dimensional EMF generated by the source. Through nine data combinations (3x3 axes) at each detector, fluctuations generated by movement of the respective finger can thus be detected. From these data, the 6-axis position and rotation of each finger can be determined. Using inverse kinematics and a calibration procedure, the missing values of the individual phalanges can then be inferred [34]. In comparison

63 `https://cdn.prod.website-files.com/6617b472be5db819f4dbdf45/675814b5d00f538badc017a3_Datasheet%20-%20Quantum%20Metgloves%20(10-12-2024).pdf` (Last accessed on June 10, 2025)

with IMU- and flex-sensor-based data gloves, it yields comparable results for (quasi-)static handshapes in terms of repeatability, reproducibility, and reliability for closed handshapes (e.g., fist / A) and better results for open, flat handshapes (e.g., flat hand / B). A slight correlation with the participant's hand size can be observed, which is probably due to the inverse kinematics approach, which cannot exactly account for all hand sizes. In dynamic scenarios, the data glove has better repeatability and similar results to flex-sensor-based data gloves, but better results than pure IMU-based data gloves [180]. Another advantage for our work is that the data gloves use the same SDK as the *Manus Prime X* data gloves, which means that the recorded data are compatible with one another and can be used jointly for training and testing classifiers. However, care should be taken to ensure that the data from the different data gloves achieve similar results on the shared dataset, i.e., that compatibility is truly ensured and that there are no quality differences.

The classification process shall be able to recognize all handshapes of the *ASL-LEX*. In addition, it should be investigated whether it is sensible to merge the handshapes Curved L and Bent L, Curved 1 and Bent 1, and Curved V and Bent V. Because of only a marginal difference in finger flexion, significantly more errors occur in execution here, which leads to poorer recognition accuracies. Since the *Manus Quantum Metagloves* can only determine the position of the fingertips and the values of the individual phalanges are estimated via an inverse kinematics approach, it is also not possible to reliably distinguish between flexion of the lower (MCP), middle (PIP), and upper (DIP) phalanges. This gives us a total of 54 handshapes instead of 57 handshapes.

Since the results regarding the optimal data preprocessing in the previous experiment were inconclusive, we again investigate at this point which combination of data preprocessing methods leads to the most accurate and fastest results. The data are in any case scaled with a *Min-Max normalization*, and the influence of the combinations of an *Outlier Detection* and a *Feature Selection* on our overall system is examined. As before, *Outlier Detection* is performed in the form of DBSCAN, whereas instead of a machine-learning-based *Feature Selection*, based on the previous results we manually remove the information for the DIP joint of the fingers as well as for the splay of the ring finger, thereby reducing the number of features from 20 to 15 (cf. Table 11). In contrast to previous experiments, no hyperparameter optimization is performed; instead, the most frequent combinations from the previous experiment with 56 handshapes are used for the respective combination of the employed data preprocessing methods (see Section A.3, Table A.10).

### 4.5.1 *Procedure*

For the implementation of our final experiments, handshapes of the *ASL-LEX* were once again recorded, including the same 56 handshapes as in the previous experiment with an additional distinction between the handshapes Flat H and Flat N. In total, 38 participant datasets were recorded under the same conditions as before, 14 of them, as with the *Manus Prime X*, with the left hand only and twelve with both hands synchronously (nine of these were participants who had previously taken part

in both recordings). Instead of the two-dimensional sketches of the *ASL-LEX* [186] used previously, participants were shown the same three-dimensional visualizations as in Table A.4 and, in addition, 3D models that they could freely move in order to become familiar with the handshape. The visualizations and models of all recorded handshapes can also be viewed at `https://sign-parametrization.netlify.app/handshapes/asl_lex`.

Due to the compatibility of the data with the *Manus Prime X* data glove, it is worthwhile to train our model both with the previously recorded data of the *Manus Prime X* data glove and with the new data from the *Manus Quantum Metaglove* data glove. We therefore obtain 58 participant datasets, 46 of them for the left hand and 12 for the right hand.

### 4.5.2 *Results*

Table 14 shows the results of our evaluation with all 58 participant datasets and 54 handshapes. In addition to the average accuracies and classification times of the LOSOCV procedure across all runs or participant datasets, the average results are also shown separately by handedness and data glove model of the individual participant datasets. For better comparability, we have also chosen accuracy as the metric for evaluating the model here. As before, *Feature Selection* leads to the lowest classification times, whereas *Outlier Detection* yields the highest accuracies almost exclusively. It can likewise be observed that it is not sensible here either to use both data preprocessing methods in combination. Since the time savings when using *Feature Selection* compared to classification without data preprocessing amount to 11.38 to 16.20 %, whereas the accuracy with *Outlier Detection* can be improved by at most only 0.40 % (and even deteriorates in one case), we choose a data preprocessing with *Feature Selection* for our final approach.

The differences between the individual hand and glove configurations are marginal in terms of accuracy and independent of the data preprocessing. However, it must be noted that only twelve participant datasets of the right hand are available compared to 46 participant datasets of the left hand, i.e., the influence of the left hand on the results is greater than that of the right hand. Likewise, 20 participant datasets are available for the *Manus Prime X* data glove, whereas 38 participant datasets are available for the *Manus Quantum Metaglove* data glove. The fact that the right-hand data were recorded exclusively with the *Manus Quantum Metaglove*, whereas the left-hand data were recorded with both models, may also correlate with the result. Regardless, this shows us that compatibility between the different data glove models, but also between the different handednesses, is ensured, since no configuration reacts significantly worse or better to the model.

With regard to the classification times, it can be observed that data from the *Manus Prime X* data glove are classified somewhat more slowly (194.79 ms) than those from the *Manus Quantum Metaglove* data glove (177.97 ms). The times refer to the classification of an entire participant dataset.

Table 14: Evaluation results of the various data preprocessing configurations across all participant datasets (LOSOCV) and influence of handedness and the data glove model

| **Data glove** | **Hand** | **N** | **Data preprocessing** | | | |
|---|---|---|---|---|---|---|
| | | | Without | Out. | Feat. | Both |
| **Accuracy** (%) | | | | | | |
| Prime X | left | 20 | 86.61 | 86.75 | 86.58 | 86 |
| Quantum Metagloves | left | 26 | 86.03 | 86.37 | 86.12 | 85.94 |
| | right | 12 | 87.18 | 86.94 | 87.23 | 87.14 |
| | both | 38 | 86.40 | 86.55 | 86.47 | 86.32 |
| **All** | | **58** | **86.47** | 86.62 | **86.51** | **86.21** |
| **Classification time** (ms) | | | | | | |
| Prime X | left | 20 | 222.91 | 222.06 | 194.79 | 308.19 |
| Quantum Metagloves | left | 26 | 201.41 | 204.34 | 180.83 | 303.14 |
| | right | 12 | 199.61 | 186.20 | 171.78 | 262.11 |
| | both | 38 | 200.84 | 198.61 | 177.97 | 290.18 |
| **All** | | **58** | **208.45** | **206.70** | 183.77 | ***296.39*** |

N = number of participant datasets; Out. = outlier detection; Feat. = feature selection

If the models are trained exclusively with data from the *Manus Prime X* data glove, average accuracies of 88.92 to 89.42 % are obtained depending on the chosen data preprocessing. In contrast, a model trained exclusively with data from the *Manus Quantum Metagloves* achieves accuracies of 87.20 to 87.64 %. The *Manus Prime X* data glove therefore tends to achieve slightly better results. On the one hand, this could be due to the quality of the data, since there is a difference of up to 10.15 percentage points between the minimum accuracy of the *Manus Prime X* data glove and that of the *Manus Quantum Metaglove*. On the other hand, we did not perform hyperparameter optimization and used parameter combinations that were optimized with the data of the *Manus Prime X*, which in turn may have resulted in an advantage for this model. The results are shown in Table 15.

The accuracies using *Feature Selection* are 89.25 % and 87.26 %, depending on the data glove, and are thus somewhat above the accuracy of 86.51 % when using both data gloves. Although this shows that the data from the different data glove models are compatible, there still appear to be differences, which is why the separate models perform somewhat better than the joint one. With regard to better generalizability

Table 15: Cross-user accuracies (LOSOCV) of different data preprocessing configurations depending on the data glove

| DVV | **Manus Prime X** (%) | | | | **Manus Quantum Metagloves** (%) | | | |
|---|---|---|---|---|---|---|---|---|
| | MW | Stdabw. | Min. | Max. | MW | Stdabw. | Min. | Max. |
| Ohne | 89.42 | 0.70 | 73.89 | 99 | 87.58 | 0.66 | 63.74 | 98.25 |
| Out. | 88.92 | 0.72 | 72.22 | 99.44 | 87.64 | 0.56 | 67.25 | 98.25 |
| Feat. | 89.25 | 0.70 | 73.33 | 98.89 | 87.26 | 0.65 | 64.33 | 98.25 |
| Beide | 89.06 | 0.69 | 73.33 | 98.89 | 87.20 | 0.64 | 64.33 | 97.66 |

DVV = Data Preprocessing; MW = Mean Value; Stdabw. = Standard deviation; Min. = Minimum; Max. = Maximum

and due to the very small difference in accuracy, we nevertheless use the data from both glove models for our final model.

Table 16 shows the accuracies of the models that were trained exclusively with data from the left or right hand of the *Manus Quantum Metaglove*. Comparing these results, one sees that the accuracies for the left hand are between 1.37 and 2.72 percentage points better than for the right hand. The maximum accuracies achieved for the left hand are consistently 98.25 %, whereas those for the right hand vary between 92.98 to 95.32 %. We therefore assume that 98.25 % represents the actually attainable upper bound for the left hand and that the remaining 1.75 % difference is due to erroneous data. The results for the right hand may be disadvantaged by the significantly smaller amount of available training and test data.

Table 16: Cross-user accuracies (LOSOCV) of different data preprocessing configurations as a function of handedness

| DVV | **Left** (%) | | | | **Right** (%) | | | |
|---|---|---|---|---|---|---|---|---|
| | MW | Stdabw. | Min. | Max. | MW | Stdabw. | Min. | Max. |
| Without | 86.89 | 0.56 | 70.18 | 98.25 | 84.21 | 0.54 | 67.25 | 92.98 |
| Out. | 86.51 | 0.55 | 69 | 98.25 | 85.14 | 0.54 | 70.18 | 95.32 |
| Feat. | 86.93 | 0.56 | 70.76 | 98.25 | 84.21 | 0.59 | 67.25 | 94.15 |
| Both | 86.57 | 0.59 | 69 | 98.25 | 84.89 | 0.49 | 71.35 | 95.32 |

DVV = Data Preprocessing; MW = Mean Value; Stdabw. = Standard deviation; Min. = Minimum; Max. = Maximum

All these results refer to the classification of 54 handshapes without differentiation between “Bent” and “Curved” handshapes. In order to better assess the results, we reclassified the model with manual *Feature Selection* using all data available to us and differentiating between “Bent” and “Curved” handshapes, and were able to achieve an accuracy of 84.34 %, thus, as expected, lower than without differentiation (86.51 %

accuracy). The difference in the $F_1$ score is even more noticeable (81.64 % compared to 84.20 %).

As already mentioned, several handshapes were simplified and merged for the training of the final handshape classifier, such as the "Bent" and "Curved" handshapes. In order to be able to train the classifier with a maximum amount of data, all available data were used, regardless of class balance. This means that some classes have an above-average number or a below-average number of data points. An overview of all recorded handshapes and those used for classification can be found in Section A.3, Table A.11.

Since three repetitions were recorded for each handshape from each participant, the following exceptions arise:

"BENT" AND "CURVED" HANDSHAPES: Originally, handshapes from 58 participants were recorded for both the labels BENT 1, BENT L and BENT V, as well as CURVED 1, CURVED L and CURVED V. These were later merged, which is why there is a double number of handshapes for these labels. Since the ASL fingerspelling alphabet was also recorded for the *Manus Prime X* data glove and the letter X is identical to BENT 1, 20 additional participant data were added here.

ORIENTATION-INDEPENDENT HANDSHAPES: For the *Manus Prime X* data glove, the ASL fingerspelling alphabet was also recorded taking the correct orientation into account, which is why for the handshapes K and Q there are still 20 additional participant data each for the letters P (identical to the handshape K) and G (identical to the handshape Q).

INCORRECT REPRESENTATION IN *ASL-LEX*: As already mentioned, it was initially not clear in our experiments what differences exist between FLAT H and FLAT N. Since both handshapes were represented identically in *ASL-LEX*, they were also recorded identically with the *Manus Prime X* data glove as FLAT N. This leads to the FLAT N handshape having double the number of recordings from the *Manus Prime X* data glove, whereas for FLAT H only the handshapes from the *Manus Quantum Metagloves* are available.

From these exceptions arises a class ratio (most frequent to least frequent class) of $408/114 \approx 3.58$. Although only seven of the 54 classes are affected, we nevertheless want to investigate whether one can still speak of a balanced class distribution and whether accuracy is the correct metric for evaluating our model. Table 17 shows the average classification performance across all classified handshapes in the form of a *classification report*. The full report can be viewed in Section A.3, Table A.12.

The average values for *Accuracy*, *Macro*-$F_1$, and *Weighted*-$F_1$ are extremely close to one another at 86.40 to 86.51 %. This is a strong indication that the classifier does not systematically favor or disadvantage any class and that the class distribution does not cause a dominant bias. Looking at the classification performance at the level of individual classes, the particularly low $F_1$ *score* of CURVED H with less than 60 % stands out. It is the only handshape for which more than 20 % of the predictions are assigned to a specifically different class (CURVED V) with 30.46 %. This indicates a pronounced

Table 17: Aggregated average values of the classification report for the final hand pose classifier from Table A.12

| Label | Precision | Recall | $F_1$-Score | Support |
|---|---|---|---|---|
| accuracy | 86.52 % | 86.52 % | 86.52 % | 10,038 |
| macro avg | 86.76 % | 86.36 % | 86.43 % | 10,038 |
| weighted avg | 86.58 % | 86.52 % | 86.42 % | 10,038 |

risk of confusion. Since the two handshapes differ only in the splay of the index and middle fingers, this confusion is plausibly justified. The misclassification can also be observed in Figure 46.

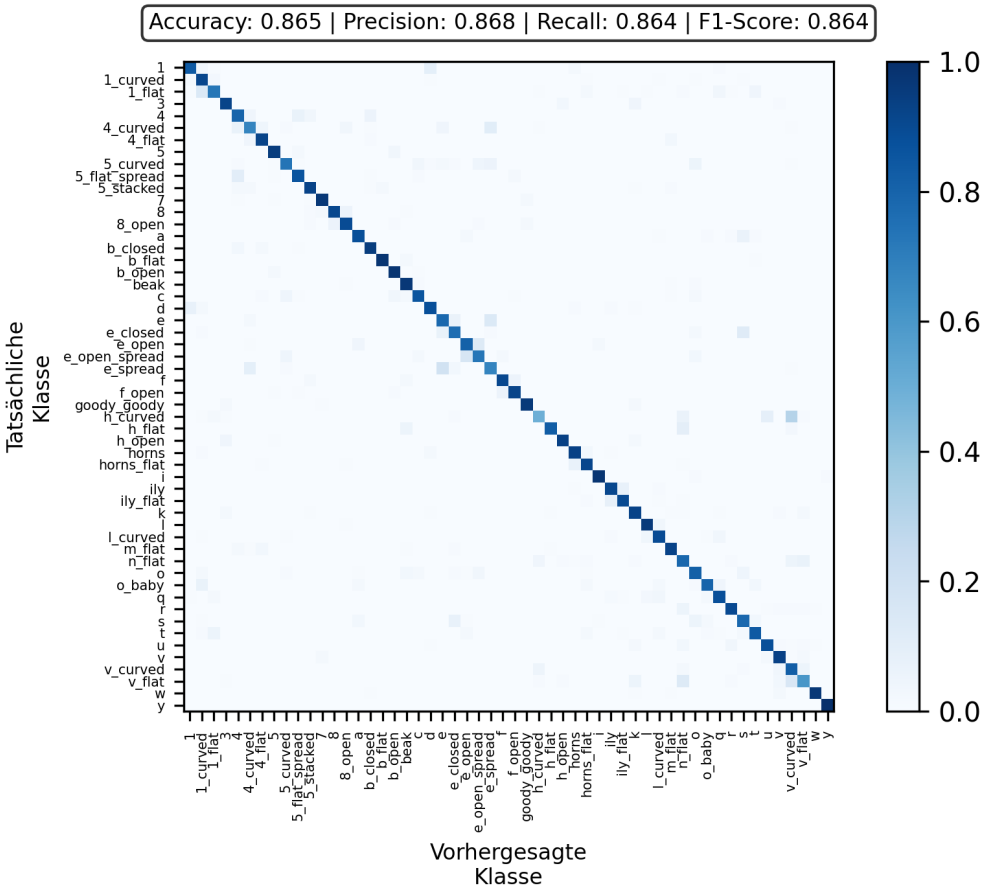



Figure 46: Confusion matrix of the final handshape recognition

### 4.5.3 *Conclusion on handshape recognition*

After completing our investigations, it was decided to use a VL2 meta-classifier with SVM, RF and LR as first-level classifiers for the classification of static handshapes. In addition to *Min-Max normalization*, we manually excluded five of the 20 available features in order to speed up the classification process. The trained model is compatible with both the *Manus Prime X* data gloves and the *Manus Quantum Metagloves* and supports both hands. The originally 57 handshapes of the *ASL-LEX* were reduced to 54 handshapes by merging the *Curved* and *Bent* handshapes, which are prone to confusion.

Table 18: Comparison of the technical specifications of the data gloves investigated within the scope of this work

| **Specifications** | **Senso Glove: DK2** (DK3) [6, 7, 34, 110] | **Manus Prime X**[60] | **Manus Quantum Metagloves**[64 65] [180] |
|---|---|---|---|
| Publication | 2017 (2020) | 2021 | 2022 |
| Sensor technology | IMU-based | IMU-/Flexsensor-based | EMF-based |
| Sensor type | 8x IMUs<br>2x Magnetometer | 6x 9-DoF IMUs ($\pm$2.50°)<br>5x 2-DoF Flex sensors | 1x 9-DoF IMUs<br>6x 3-DoF EMF-coils |
| Sampling rate | 10 Hz (400 Hz) | 90 Hz | 120 Hz |
| Signal latency | $\approx$ 20 ms (15 ms) | $<$ 5 ms | $\leqslant$ 7.50 ms |
| Battery life | 4.80 h (1.50 to 4 h) | 5 h (swappable) | 3 h (swappable) |
| Charging time | 1 h (1 h) | 2 h | 2 h |
| Weight | 90 g (300 g) | 60 g | 138 g |
| Connectivity | USB, Bluetooth | USB-C, Prop. 2.40 GHz protocol | USB-C, Bluetooth Low Energy (BLE) 5 |
| Finger features | **Thumb**:<br>Quat. (4), Flex. (2), Spread<br>**Finger**:<br>Flex. (2) | **Thumb**:<br>Quat. (4), Pos. (3), Flex. (3), Spread<br>**Finger**:<br>Quat. (4), Pos. (3), Flex. (3), Spread | **Thumb**:<br>Quat. (4), Pos. (3), Flex. (3), Spread<br>**Finger**:<br>Quat. (4), Pos. (3), Flex. (3), Spread |
| Hand features | **Palm**:<br>Quat. (4), Pos. (3), Ori. (3)<br>**Wrist**:<br>Quat. (4), Ori. (3) | **Palm**:<br>Quat. (4), Pos. (3) | **Palm**:<br>Quat. (4), Pos. (3) |

Proprietary (Prop.), Quaternion (Quat.), Position (Pos.), Orientation (Ori.), Flexion (Flex.), Spread (Spread)

In total, three data gloves from two different manufacturers, each using different sensor technologies, were examined in the experiments listed above. Table 18 shows a comparison of the gloves with all information relevant to our system. What all gloves have in common is that they capture handshape too imprecisely to provide truly accurate results. At least every tenth handshape is misclassified with these gloves, even though simplifications to the handshapes have already been made. Regardless of sensor technology, the data gloves have only one sensor per finger, which means that the positions of the individual finger segments must be estimated. This leads to difficulties precisely in the case of flexions of the *Curved* and *Bent* handshapes when distinguishing between them. In addition, the *Senso Glove: DK2* determines finger spread via IMUs and does not measure it explicitly using flex sensors, for example, which would yield significantly more robust results. Since handshape is the most important parameter, even more precise data gloves should be used in the future that, in addition to the five finger spreads, support at least one additional sensor per finger for more precise capture of flexion. This would make it possible to track the MCP and PIP-/DIP joints, since the latter cannot be moved independently of one another. The data glove would therefore need to have at least ten sensors for measuring the flexion of the individual finger segments, five sensors for measuring finger spread, and one IMU for measuring the orientation of the hand. According to Caeiro-Rodríguez et al. [34], for example, the *VMG 35 Haptic* (IMUs and flex sensors) and the *SenseGlove DK1* (IMUs and rotary encoders) would be suitable for this purpose. Both provide 23 DoF with three flexion values plus one spread value per finger as well as the orientation of the hand, although in the case of the *SenseGlove DK1* this is achieved with the aid of an external tracker.

***Note** on the detection of contact between thumb and finger*

In addition to the limitations of the data gloves mentioned above, difficulties in the precise detection of thumb posture were also observed across all models. Even in the first experiment with the *Senso Glove: DK2*, a strong drift of the thumb could be observed [7]. Even when using the *Manus Prime X* data glove, imprecise detection of the thumb posture could be identified, as shown in Figure 47. This primarily affects handshapes in which the tip of the thumb should touch a specific part of the hand or other fingers, such as F (fingers are extended, the tips of the thumb and index finger touch), M/N/T of the fingerspelling alphabet (hand forms a fist, thumb lies between two adjacent fingers depending on the letter), or S (as before, but with the thumb lying on the fingers).

The only parameter of the *ASL-LEX* relating to handshape that has not yet been addressed at this point (cf. Section 4.1.2) is THUMB CONTACT. Since this parameter addresses exactly this problem – recognizing whether the thumb touches a particular other finger – and since at the outset we already required "very precise" recognition in

64 https://cdn.prod.website-files.com/6617b472be5db819f4dbdf45/675814b5d00f538badc017a3_Datasheet%20-%20Quantum%20Metgloves%20(10-12-2024).pdf (Last accessed on May 27, 2025)

65 https://docs.manus-meta.com/3.0.0/Software/Skeletons/ (Last accessed on May 29, 2025)

(a) Left data glove (b) Right data glove

Figure 47: Imprecise capture of the thumb posture with the Manus Prime X for handshape F

order to distinguish between *contact* and *non-contact*, we exclude this parameter from our ParaSignRec at this point.

# 5 FURTHER SYSTEM COMPONENTS

After the conception and implementation of the handshape recognition approach as the most important parameter of a sign language, this section now presents additional essential system components of the ParaSignRecs. In order to capture arm movements and poses, additional sensors are required alongside the data gloves; moreover, a strategy for calibration and sensor fusion of the different sensor systems is needed in the context of a suitable data acquisition procedure. In addition, further recognition approaches for detecting the *articulation location* of the sign, the *hand orientation*, and the external and internal *hand movements* are presented.

## 5.1 ARM CONTROLLER

As already mentioned in Section 3.2.2, in order to capture further relevant parameters of signs – such as arm movements or the articulation location of the sign based on arm poses – additional sensor technology is required in addition to the use of data gloves. For this purpose, two *arm controllers* were designed in the context of this work, as shown in Figure 48.

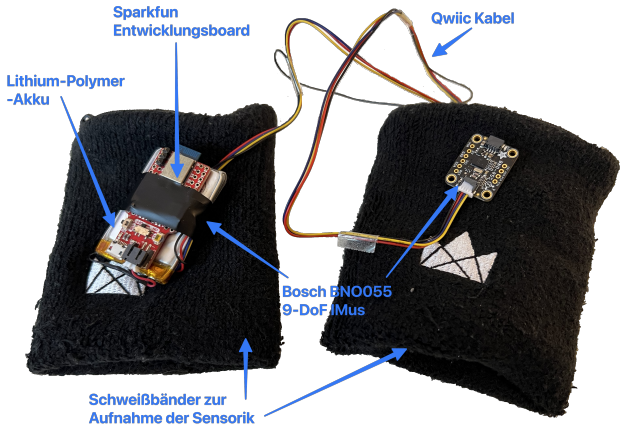



Figure 48: Components of the arm controller

The arm controllers essentially consist of a *SparkFun Pro nRF52840 Mini – Bluetooth Development Board*[66] and two *Bosch BNO055*[67] IMUs. The IMUs combine a 3-DoF accelerometer, a 3-DoF gyroscope, and a 3-DoF magnetometer. Using an integrated

66 `https://www.sparkfun.com/sparkfun-pro-nrf52840-mini-bluetooth-development-board.html` (Last accessed on January 25, 2026)

67 `https://www.bosch-sensortec.com/products/smart-sensor-systems/bno055/` (Last accessed on January 25, 2026)

sensor fusion unit, directly usable absolute orientation values in the form of quaternions, Euler angles, or vectors can be provided at a sampling rate of 100 Hz. The IMUs are each mounted on a breakout board from *Adafruit* and connected to the development board via the *Qwiic* interface. This features an ARM Cortex-M4 microcontroller with Floating Point Unit (FPU), 1 MB of internal flash memory and 256 kB of RAM, as well as BLE 5 for wireless communication with the computer. The setup is powered by a Lithium Polymer (LiPo) battery (3.70 V, 1,100 mA h), ensuring a compact and portable measurement setup that is attached to the upper and lower arm using two sweatbands with zipper pockets for housing the hardware.

This setup also fulfills the hardware requirements defined in Section 3.2.1: By being attached directly to the arms and due to the low total weight of 100.90 g and a maximum depth of less than 2.50 cm per unit (consisting of two sweatbands including electronics), the arm controllers only slightly restrict the participants' movements and can be flexibly attached to the upper and lower arm. The acceleration data and quaternions as the quantities to be recorded are processed onboard at a data rate of up to 100 Hz and can be used directly. The hardware is readily available and, thanks to the use of the Qwiic interface, can be easily connected with the appropriate cables without soldering. Moreover, the hardware captures only the quantities to be measured in the form of movements of the participant wearing it and therefore provides good protection of the privacy of uninvolved third parties.

## 5.2 DATA ACQUISITION PROCESS

After the hardware components have been defined, the data acquisition process is described here. Data acquisition comprises both the collection of training and test data for the development and adaptation of the individual recognition approaches as well as the recording of the final signs for the evaluation process of the meta-classification. The included data processing ensures that data from different sensors can be used together and are compatible. It includes calibration of the data as well as interpolation in order to merge the data despite differing data rates. Overall, special emphasis was placed on ensuring a high data quality. At the same time, however, the process is intended to provide an image of the actual execution of signs that is as realistic as possible, so that minor variations or slight irregularities are not deliberately eliminated, but retained as part of the natural variability.

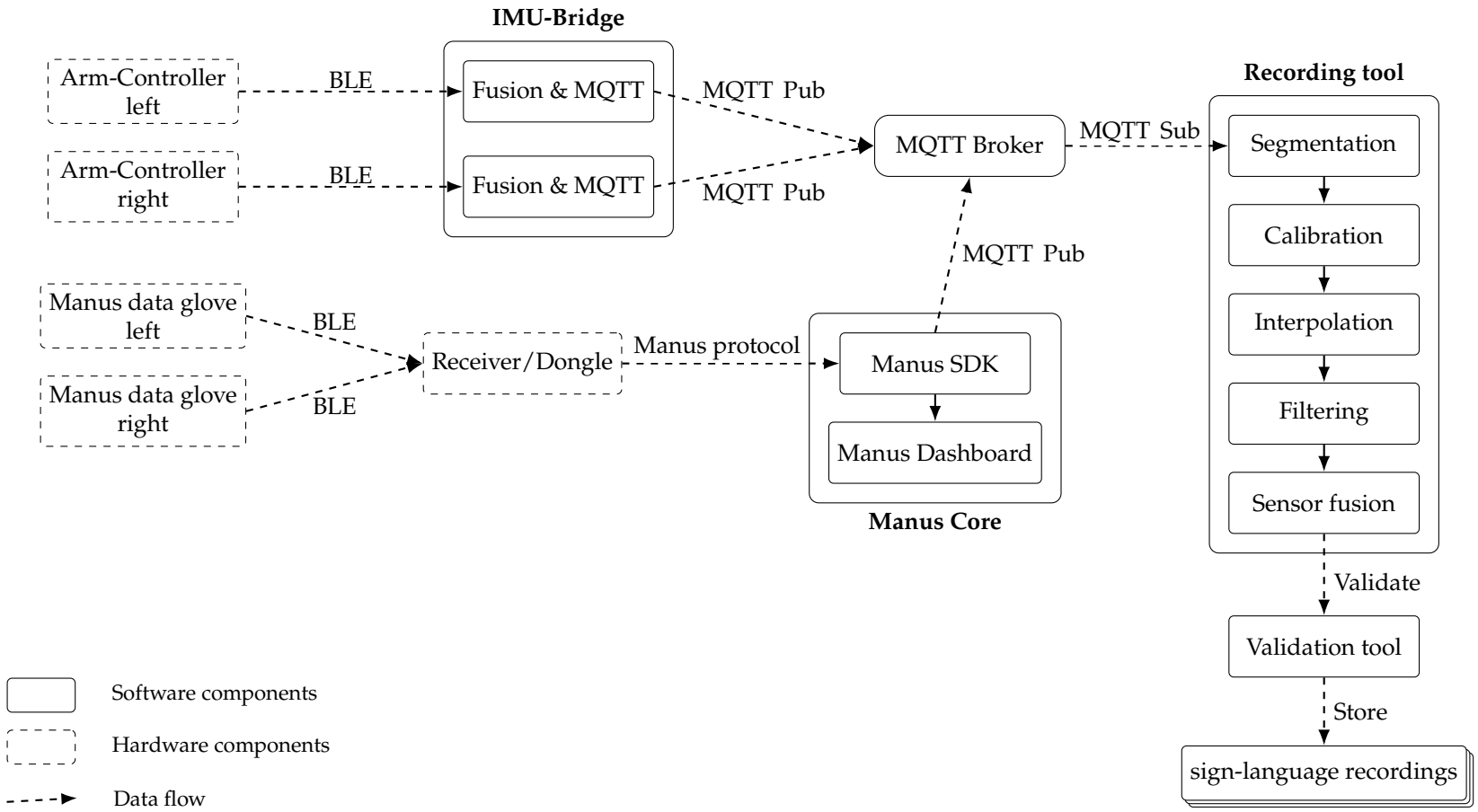



Figure 49: System architecture of the sign-language recording tool

Figure 49 shows the system architecture of our data acquisition and pre-processing. Both data gloves as well as both arm controllers are connected to the acquisition computer via the appropriate interfaces and integrated into the system. Using the Message Queuing Telemetry Transport (MQTT) protocol, data from different sources are made available to the data acquisition tool. Either the *Manus Prime X* or the *Manus Quantum Metagloves* can be used as data gloves, whose data are compatible with one another.

### 5.2.1 *Data Collection*

The following first describes the data acquisition process from an organizational and procedural perspective.

In addition to the participant, at least one other person was always involved in the data collection. This person supervised the recording and ensured correct execution. Before recording, each participant was provided with an overview of all signs to be recorded, including example videos, so that the signs could be practiced. In addition, the recording supervisor practiced the signs with the participant before each recording and sometimes also demonstrated the sign during recording so that the participant could imitate it. Neither the recording supervisor nor the participant had any prior knowledge of ASL or any other sign language.

After putting on the data gloves and arm controllers, the data gloves were first calibrated using the *Manus Core* SDK. Subsequently, the necessary processes for the data acquisition tool were started, i.e., our self-developed *IMU-Bridge* as well as the MQTT broker and the actual acquisition tool. The *IMU-Bridge* serves to receive the data of the IMUs from both arm controllers via BLE and publish them unprocessed to the MQTT broker. It provides the current timestamp, the rotation of the upper and lower arm as a quaternion, as well as the acceleration data corrected for gravitational force for each of the three axes of the lower arm at 30 times per second. Through the *Manus Core* SDK, the current timestamp, the 20 flexion and spread values of the individual fingers and the thumb shown in Table 11, as well as the rotation of the back of the hand in the form of a quaternion, are likewise provided via MQTT.

Each recording started in a so-called *T-pose*, i.e., the participant stood in an upright position with the arms extended sideways and the palms facing downward. From there, the arms were moved into the *I-pose* by rotating the shoulder joint, so that they now pointed downward, with the palms facing inward toward the body. These two poses are shown in Figure 50 and serve as reference poses for subsequent orientation calibration of the data gloves and arm controllers.

Starting from the I-pose, the signs were then recorded, either as multiple repetitions of a single sign or as several random signs in succession, but with a maximum of four signs or repetitions. Afterwards, the next recording was started again, beginning with the T-pose.

The recordings were carried out while standing. Each recording session was filmed with a camera so that the recordings could subsequently be segmented and the correct execution of the sign could be validated afterwards.

(a) T-Pose (b) I-Pose

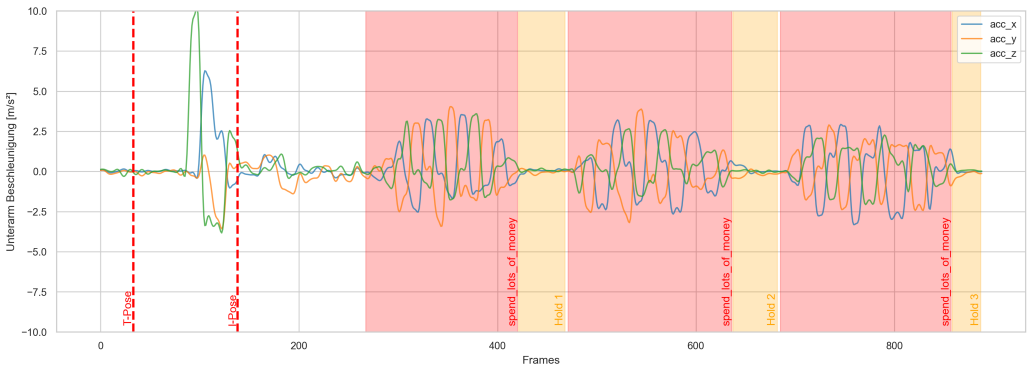
Figure 50: T-pose and I-pose for quaternion calibration

### 5.2.2 *Segmentation*

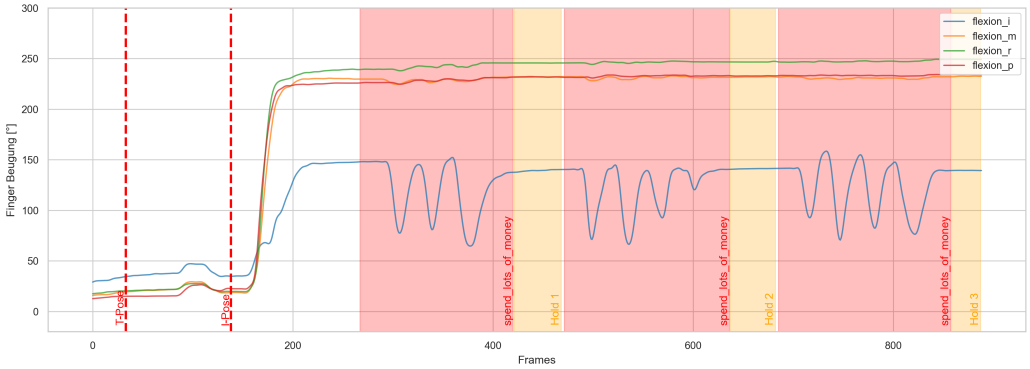
Following data acquisition, the recordings were manually segmented using the video recordings and the timestamps available for both the video and the data streams. Using a self-developed interface, the actual sign apex according to Kendon (see Figure 3 in Section 2.1) as well as the subsequent optional hold phase were selected for each sign based on the video recordings. Figure 51 shows four different sensor values of a segmented recording with three sign repetitions (red) and optional hold phases (yellow). The timestamps of the T-pose calibration and I-pose calibration are marked as a red vertical line.

### 5.2.3 *Calibration*

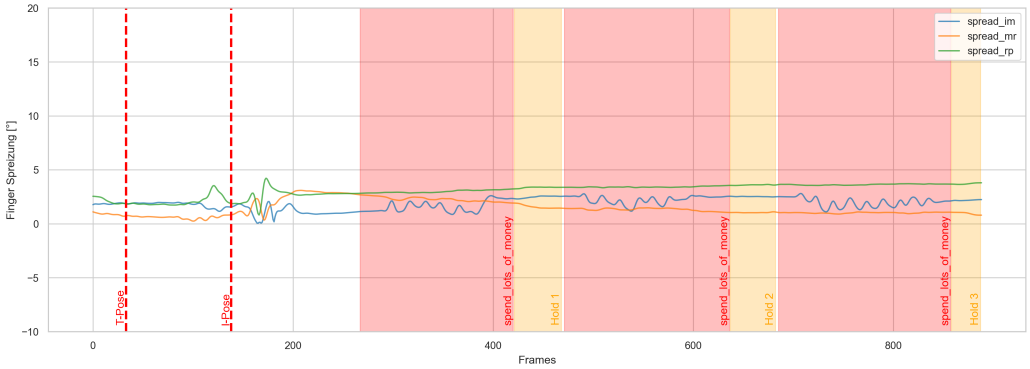
Based on this manual segmentation, the individual signs were extracted from the recordings and processed. First, the orientation calibration of the quaternions is carried out using a two-stage reference pose system that employs the already described T-pose and I-pose. This two-stage calibration approach ensures precise orientation calibration for both the data gloves and the arm controllers, enabling the quaternions from the different reference frames to be transformed into a common system through suitable coordinate system transformations. It also ensures correct alignment of the arms, independent of sensor placement.



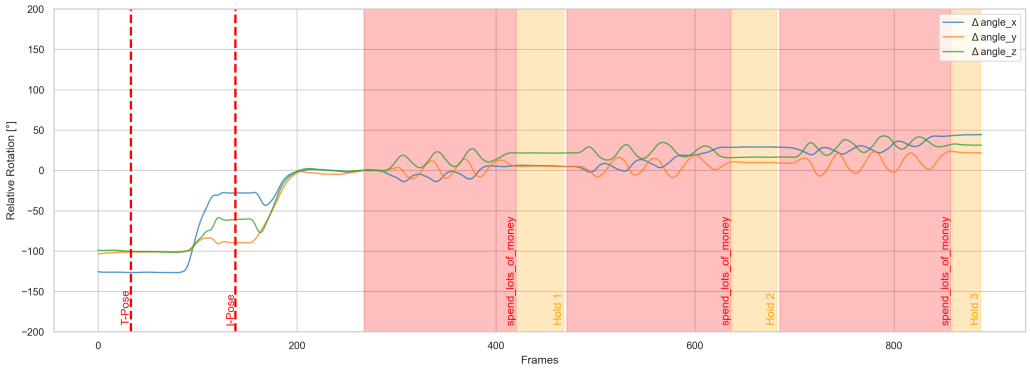

(a) Acceleration of the lower arm




(b) Flexion values of the fingers




(c) Spread values between the finger pairs




(d) Relative rotation of the hand

Figure 51: Example recording of three repetitions of the sign SPEND LOTS OF MONEY.

First, the raw quaternions are normalized and mapped to a consistent orientation of the quaternion representation in order to avoid sign flips. The quaternions at the time of the T-pose and I-pose are extracted from the time series data ($q_{t,ist}$, $q_{i,ist}$). The T-pose serves as the primary reference orientation, so the quaternion $q_{t,ist}$ is now mapped to the desired reference orientation $q_{t,soll}$. The mapping is performed by forming a relative rotation quaternion and then applying it to the raw data. For this purpose, the following difference is formed:

$$q_{rel,t} = q_{t,ist}^{-1} \cdot q_{t,soll} \tag{10}$$

This difference is now applied to $q_{t,ist}$ and to all subsequent quaternions:

$$q_{calibrated} = q_{raw} \cdot q_{rel,t} \tag{11}$$

Using this procedure, the quaternions of the upper and lower arm can be mapped to the T-pose; however, it cannot yet be ensured that the rotation about the longitudinal axis of the extended arm corresponds to the actual rotation. By using the I-pose as a secondary reference pose, an additional axis alignment is enabled that can solve this problem. To do this, the difference between T-pose and I-pose is first calculated, which ideally corresponds to a 90° rotation of the shoulder joint in the frontal plane:

$$q_{rel,i} = q_{t,soll}^{-1} \cdot q_{i,tmp} = q_{t,soll}^{-1} \cdot q_{i,ist} \cdot q_{rel,t}. \tag{12}$$

This relative quaternion $q_{rel,i}$ is then converted into a rotation vector $\vec{r}$[68], where the direction of $\vec{r}$ represents the axis of rotation and the length $|\vec{r}|$ the angle of rotation. The goal is to align this axis of rotation with the y-axis of our coordinate system[69], so that only the pure rotation about the y-axis remains.

To do this, the normalized axis $\vec{n}$ is first determined:

$$\vec{n} = \frac{\vec{r}}{|\vec{r}|}. \tag{13}$$

The alignment axis $\vec{a}$ is obtained as the cross product with the y-axis:

$$\vec{a} = \vec{n} \times \vec{y}, \quad \vec{y} = (0, 1, 0)^T. \tag{14}$$

The rotation angle $\theta$ between $\vec{n}$ and $\vec{y}$ is determined via the dot and cross product:

$$\cos(\theta) = \vec{n} \cdot \vec{y}, \quad \sin(\theta) = |\vec{a}|, \quad \theta = \arctan 2\left(|\vec{a}|, \vec{n} \cdot \vec{y}\right) \tag{15}$$

The desired alignment quaternion $q_{align}$ is then:

$$q_{align} = \left(\cos\left(\tfrac{\theta}{2}\right), \vec{a}_{norm} \cdot \sin\left(\tfrac{\theta}{2}\right)\right), \tag{16}$$

68 https://docs.scipy.org/doc/scipy-1.7.0/reference/generated/scipy.spatial.transform.Rotation.as_rotvec.html (Last accessed on January 25, 2026)

69 Right-handed coordinate system with x-axis in the direction of the extended right arm in T-pose and z-axis pointing upward.

where $\vec{a}_{norm} = \frac{\vec{a}}{|\vec{a}|}$ is the normalized rotation axis.

Using this quaternion, the sensor data are now transformed analogously to Equation 11, so that the T-pose is preserved, but the axis of rotation for the movement into the I-pose is correctly aligned and represents a pure rotation about the y-axis.

### 5.2.4 *Sensor Fusion*

After the sensor data have been transformed into a common coordinate system, they are now to be merged in the temporal domain. Since the recorded signs, however, have different sampling rates (cf. Figure 52), they must first be interpolated to a uniform signal of 50 Hz using the function `interp1d`[70] from the *SciPy* package with the parameters `kind=linear` and `fill_value=extrapolate`.

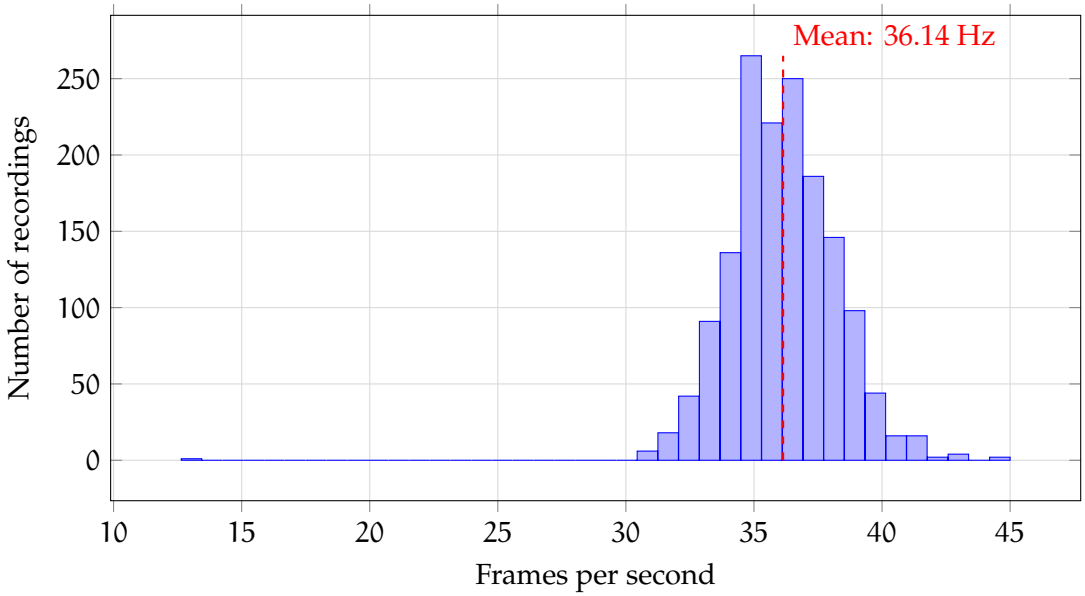



Figure 52: Average number of frames per second with a mean value of 36.14 Hz and a standard deviation of 2.09 Hz when evaluating 1,544 recorded test gestures

To suppress noise and high-frequency interference, a Butterworth (low-pass) filter with a cutoff frequency of 6 Hz and 4th order was subsequently applied to the interpolated data, taking [29, 120] into account. This interference also includes the muscle tremor in the range of 5 to 100 Hz already described in Section 4.2.2.

For the fusion of data streams from different sources, the latest common start point and the earliest common end point were first determined within the signals based on the timestamps. Since all signals were already available with identical sampling frequency, the overlapping range could be extracted and merged into a uniform signal. Based on the curves shown in Figure 51, it is clearly visible that the sensor data are synchronized and that our timestamp-based synchronization approach works.

70 `https://docs.scipy.org/doc/scipy/reference/generated/scipy.interpolate.interp1d.html` (Last accessed on January 25, 2026)

### 5.2.5 *Quality Assurance Process*

In order to ensure a high data quality – especially when using the data to train our recognition approaches – each individual sign recording was subsequently reviewed and subjectively checked for correct execution, annotation, segmentation, and data quality. For this purpose, the video recording of the execution and the graphical plots for the data provided by the right data glove were examined in a self-developed interface. These include *i)* the acceleration values of the lower arm, *ii)* the flexion values of the index finger, middle finger, ring finger, and little finger, *iii)* the spread values between the index finger/middle finger, middle finger/ring finger, and ring finger/little finger pairs, and *iv)* the rotation of the back of the hand. The latter were transformed from quaternions into Euler angles for better readability. All data are presented as a relative change to the first available data point of each sign execution in order to make changes easier to perceive. A screenshot of the tool is shown in Figure 53.

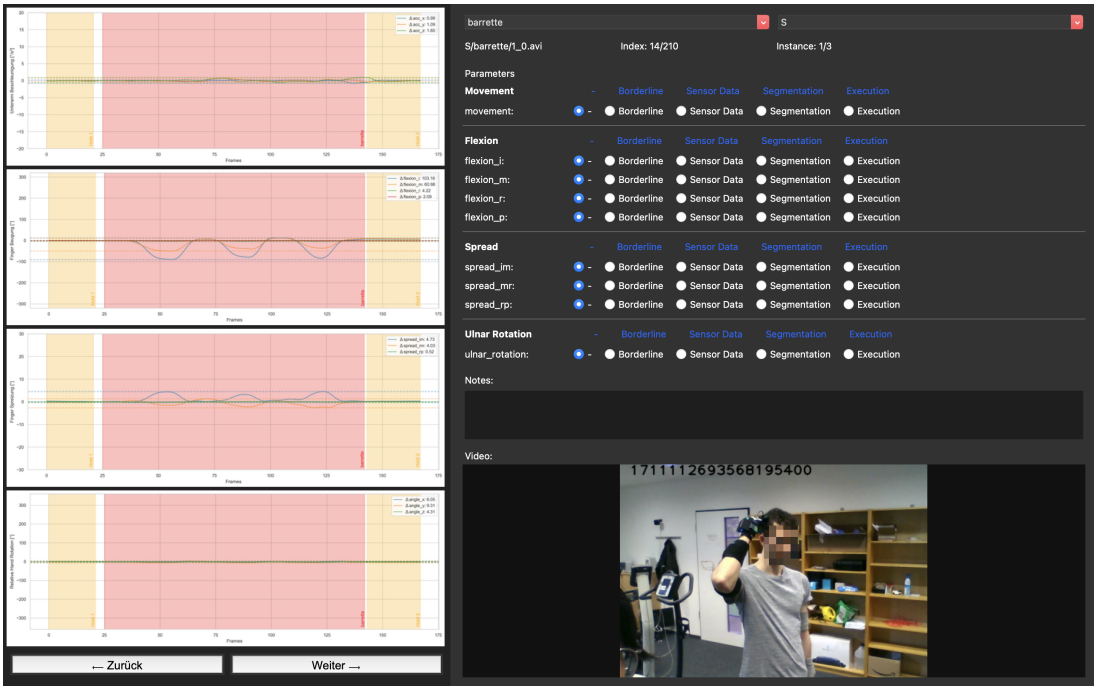



Figure 53: Self-developed tool for validating sign recordings

During validation, four different sources of error were examined:

INCORRECT EXECUTION: The participant executed the sign or parts of the sign incorrectly.

INCORRECT ANNOTATION OF THE DATA: The sign was annotated incorrectly by *ASL-LEX*, i.e., the participant executed the sign correctly based on the example video from *ASL-LEX*, but the stored information about the sign is incorrect.

INCORRECT SENSOR DATA: Both execution and annotation of the data are correct, but the data do not match expectations.

INCORRECT SEGMENTATION: Relevant parts of the sign execution lie outside the segmentation or the segmentation includes data that do not belong to the actual sign execution.

The data were then examined to determine whether a perceivable change in the data as movement also matched the intended sign or whether the opposite was the case. If a movement was detected although the sign did not intend this type of movement, or if no movement was detected although a movement should have been present, the video recording was used to check whether the sign was executed incorrectly or whether the sign in *ASL-LEX* was assigned to the wrong class, and the corresponding parameter of the sign execution was marked accordingly. If neither of these cases applied, we assumed incorrect sensor data. Minor movements were ignored at the user's discretion during the review, as a certain degree of freedom in movements is natural, especially when the sign exhibits other types of movement. Furthermore, it was checked whether the movements under investigation also lay within the segmented sign window or whether relevant parts lay outside the segmentation. However, it must be noted that it is certainly possible that incorrect executions were not recognized in the video recordings and these executions were then labeled as incorrect sensor data.

If a sign execution was marked as incorrect, the affected parameters were excluded from further processing. In the case of incorrectly annotated data, however, the sign executions were still used and only the data basis was corrected instead.

The tool also offers the possibility to create notes and to mark sign recordings as "borderline". This can be used if a slight deviation from the expected data can be identified, but it should not yet be described as an outlier or as incorrect. However, these two pieces of information were not used further.

## 5.3 RECOGNITION OF THE SIGN ARTICULATION LOCATION

Another meaning-distinguishing parameter for recognizing signs is the location of articulation. This often contributes to the meaning of the sign; for example, signs with cognitive meanings are articulated near the head, whereas signs expressing feelings are articulated near the heart. According to Tennant et al. [215], about 75 % of signs are articulated at head and chest height, since they are easy to see there. This observation also roughly matches the vocabulary of *ASL-LEX*, for which this applies to about 71.17 % of signs (after excluding the *Other* classes). Due to the good visibility, all signs are located within a defined space, the signing space, which roughly corresponds to a rectangle and is shown in Figure 54. This extends in height from hip level upwards to just above the head, in depth from the body to a maximum of 30 cm in front, and in width rarely more than 30 cm to the left or right beyond the upper body. The hand does not necessarily need to be in contact with the body, but can also be near a body part. When speaking before a large audience, the signing space becomes proportionally larger and the signs can be presented in an exaggerated manner. In the opposite case,

the signing space can become smaller for faster signs or for reasons of discretion [186, 215].


Figure 54: Signing space (adapted from [215])

Recognizing the location of articulation of a sign without using an external observer – as required by our research objectives – poses a particular challenge. All related work (cf. Table A.8) that also considered the location of articulation dispensed with a mobile approach and used external observers such as stationary magnetic trackers. The approach we propose, however, builds on earlier work. In that work, we already showed how a full-body pose can be reconstructed using *inverse kinematics* [42]: stationary trackers were used to determine the positions of the forearms, lower legs, and hips, and based on these and specific knowledge about human anatomy, the angular positions or poses of the body parts in between, such as upper arm, upper body, and thigh, were estimated. This time, however, we take the opposite approach and determine the relative position of the hands with respect to the signer using *direct kinematics*. To this end, the orientations of the upper arm and forearm are captured using IMUs, the pose is derived from them, and based on that, the relative coordinate of the hand is determined using the normalized arm lengths and trigonometric functions. This, in turn, can then be translated into a relative location of articulation, i.e., a discrete class.

The aim of our approach is therefore to estimate the membership of the current hand position in these discrete classes. To do so, the position of the hand must first be determined. Depending on how accurate the chosen approach is, the classes to be captured can then be determined and test data recorded. With this, a strategy can then be devised to assign the position of the hand to the classes using a distance-based similarity score.

### 5.3.1 *Determining the hand position*

To determine the hand position relative to the signer, we use the calibrated quaternions of the arm controllers and data gloves. These are converted via Equation 19 into the rotation matrices $R(q_{\text{Oberarm},i})$, $R(q_{\text{Unterarm},i})$, and $R(q_{\text{Hand},i})$ for time $i$. Using these and the normalized lengths of the respective limbs, starting from a fixed shoulder

coordinate $\vec{s}_0$, the positions $\vec{e}_i$ (elbow), $\vec{w}_i$ (wrist), and $\vec{h}_i$ (center of the hand) can be determined:

$$\begin{aligned}\vec{e}_i &= \vec{s}_0 + R(q_{\text{Oberarm},i})\,[\ell_{\text{Oberarm}}, 0, 0]^\top,\\ \vec{w}_i &= \vec{e}_i + R(q_{\text{Unterarm},i})\,[\ell_{\text{Unterarm}}, 0, 0]^\top,\\ \vec{h}_i &= \vec{w}_i + R(q_{\text{Hand},i})\,[\ell_{\text{Hand}}, 0, 0]^\top\end{aligned} \tag{17}$$

To normalize the lengths, the ratio of the length of the respective limb to the total length of upper arm, forearm, and palm is formed, where the length of the palm corresponds to the distance from the wrist to the beginning of the index finger. We use this point as the reference point for determining the hand position, since the fingers can adopt different poses depending on the sign (for example, closed, as in a fist/*A* handshape, or extended, as in a flat hand/*B* handshape) and we want to determine the position independently of the finger poses.

For the collection of the training data, the relevant lengths are recorded individually for each participant and used in the approach. In order to keep the approach as universal as possible and avoid an elaborate setup, the evaluation instead uses the average values of all participants in the data collection. It would indeed be possible to further increase the accuracy of the approach by taking into account all individual body proportions of the subjects, but this is deliberately omitted in favor of the universality of the approach due to the considerable effort involved.

The entire process of hand position determination is carried out for each of the two arms, using a different fixed shoulder coordinate in each case. The distance between these two coordinates is determined using the average shoulder width according to Springs et al. [201], which is also normalized by the total length of upper arm, forearm, and center of the hand. According to the reference, shoulder width is approximately 41.90 cm ± 2.30 cm to 48.20 cm ± 2.60 cm, depending on gender and origin, where in our approach the approximate mean value of 46.40 cm is used independent of the user's gender.

### 5.3.2 *Selection of the locations of articulation*

The *ASL-LEX* has four phonological properties that describe the *sign location of articulation* of the dominant hand and can be sketched as follows (for the full descriptions, including any exceptions, see Table A.2):

MAJOR LOCATION is the primary location of articulation of the dominant hand relative to the body of the sign language user, i.e., *Head*, *Body*, *Arm*, *Hand*, *Neutral*, and *Other*.

MINOR LOCATION is the secondary location of articulation of the dominant hand with respect to the MAJOR LOCATION at the beginning of the articulation. It usually consists of nine classes per primary location of articulation (with the exception of *Neutral* and *Other* classes).

SECOND MINOR LOCATION is the secondary location of articulation of the dominant hand with respect to the MAJOR LOCATION at the end of the articulation, with the same possible classes as for MINOR LOCATION.

CONTACT indicates whether the dominant hand had contact with the MAJOR LOCATION at any time during the articulation, i.e., touched it.

If one considers the *ASL-LEX*, one sees that all information of the phonological property MAJOR LOCATION (primary location of articulation) is contained in MINOR LOCATION (secondary location of articulation), since these are fixedly assigned to one another: each MAJOR LOCATION (with the exception of *Neutral*) can be subdivided into ten different MINOR LOCATIONS (including *Other*). A MINOR LOCATION can only be assigned to one MAJOR LOCATION. The location of articulation at the beginning of the sign is therefore determined exclusively by MINOR LOCATION and at the end of the sign by SECOND MINOR LOCATION, which is why MAJOR LOCATION can be neglected without losing information about the sign. SECOND MINOR LOCATION always has the same primary location of articulation as MINOR LOCATION, which is why a dependency exists here as well. After excluding the unspecified *Other* classes (cf. Section 6.1.1), nine different classes per primary location of articulation remain for each secondary location of articulation. The location of articulation *Neutral* is an exception and is not further subdivided. This results in $4 \cdot 9 + 1 = 37$ locations of articulation, which in alphabetical order are as follows (the percentages indicate the absolute frequency of the class(es) within the *ASL-LEX* after excluding *Other* [186]):

ARM (9 subclasses, 2.67 % frequency): *ArmAway* (0.53 %), *ElbowBack* (0.15 %), *ElbowFront* (0 %), *ForearmBack* (0.26 %), *ForearmFront* (0.11 %), *ForearmUlnar* (0.56 %), *UpperArm* (0.49 %), *WristBack* (0.41 %), *WristFront* (0.15 %)

BODY (9 subclasses, 9.53 % frequency): *BodyAway* (1.66 %), *Clavicle* (1.73 %), *Hips* (0.41 %), *Neck* (0.68 %), *Shoulder* (0.79 %), *TorsoBottom* (0.26 %), *TorsoMid* (0.56 %), *TorsoTop* (3.28 %), *Waist* (0.23 %)

HAND (9 subclasses, 24.70 % frequency): *FingerBack* (2 %), *FingerFront* (1.54 %), *FingerRadial* (3.54 %), *FingerTip* (2.41 %), *FingerUlnar* (0.68 %), *HandAway* (3.61 %), *Heel* (0.49 %), *Palm* (7.91 %), *PalmBack* (2.49 %)

HEAD (9 subclasses, 25.90 % frequency): *CheekNose* (3.61 %), *Chin* (3.80 %), *Eye* (1.62 %), *Forehead* (4.07 %), *HeadAway* (6.74 %), *HeadTop* (0.45 %), *Mouth* (4.74 %), *UnderChin* (0.60 %), *UpperLip* (0.26 %)

NEUTRAL (1 subclass, 37.20 % frequency): *Neutral* (37.20 %)

Each primary location of articulation (with the exception of *Neutral*) has an *Away* class, for example *BodyAway*, for signs that were articulated near this primary location of articulation but did not touch it ("The minor location 'away' was used to describe signs that were produced near, but not in contact with, a location." [187]). For the SECOND MINOR LOCATION, the class *NA* is additionally used when the secondary location of articulation of the sign is identical throughout the entire articulation ("If the

location was constant throughout the sign, then there was no second location (i.e., second minor location was coded as NA)." [187]). It can therefore be replaced by the respective value of MINOR LOCATION.

If one considers the 37 possible locations of articulation, it becomes apparent that they differ greatly in their proximity to one another and in their size. For example, the so-called *philtrum length*, as the distance between the nose and upper lip, is according to the source only about 1.30 to 2.05 cm in women and 1.40 to 2.15 cm in men depending on age[71] [240]. Another study found values of 12.45 mm (± 2.24 mm) in women and 14.30 mm (± 2.15 mm) in men[72] [237]. The entire face, by contrast, has a height of about 10.60 cm (± 0.60 cm) to 12.70 cm (± 0.60 cm), measured from the chin to the deepest point of the bridge of the nose, and a width between the zygomatic arches of 12.90 cm (± 0.60 cm) to 14.30 cm (± 0.50 cm), depending on gender and origin [201].

Our approach to determining the hand position is based, with regard to the lengths used, on simplified and generalized anatomical assumptions. In addition, the quaternions provided by the sensors are based on the integration of acceleration and angular velocity data, which makes them fundamentally susceptible to drift and noise errors. Furthermore, the hand position is derived from the positions of elbow, wrist, and hand center, so the errors of the individual partial estimates accumulate.

Lin et al. [128] investigated the accuracies of several IMUs in a study, including the *Bosch BNO055* used by us in the arm controller. They determined a static error deviation of 0.03 to 0.71° after 24 hours at rest and a dynamic error deviation of up to 4.61° during a 360° rotation of the sensor, averaged over 20 repetitions. According to Springs et al. [201], the upper arm length (including the shoulder) is approximately 32.20 cm (± 1.70 cm) to 36.60 cm (± 2.10 cm), depending on gender and origin. The forearm length (including the outstretched hand), on the other hand, is slightly longer at 44.20 cm (± 2.50 cm) to 48 cm (± 2 cm), of which the outstretched hand accounts for a length of approximately 17.90 cm (± 1 cm) to 19.20 cm (± 0.80 cm). Assuming that a tall person uses our system, whose upper arm, forearm, and hand together are approximately 84.60 cm long, the absolute error for a 360° rotation would be about $\tan(4.61°) \cdot 84.60\,\text{cm} \approx 6.82\,\text{cm}$. This would amount to at least the 3.17-fold of the philtrum length or the 0.54-fold of the face height. Since the error may be at most half the distance between two locations to distinguish between them, it would not be possible to distinguish between any locations on the face. Strictly speaking, one would have to subtract the length of the index fingers from the arm length, but one would still in no case be able to distinguish between, for example, *CheekNose* and *UpperLip*, since the philtrum length is far too short. For a maximum philtrum length of 2.15 cm, the correct location of articulation would already be indistinguishable at an error of $\alpha = \arctan(\frac{2.15\,\text{cm}}{2} \cdot \frac{1}{84.60\,\text{cm}}) \approx 0.73°$, which corresponds roughly to the static error deviation of the sensor. It should be noted that the *Bosch BNO055* we used was able to achieve the most accurate results among the tested consumer electronics sensors in comparison.

71 Measured on 2,500 healthy participants from Switzerland between the ages of 0 and 97, with both genders represented equally.

72 Measured on 200 men and 200 women of Nepalese origin between the ages of 17 and 25.

Other IMU-based approaches, such as that of Shen et al. [193], in which the pose of the arm is recognized with the help of an IMU-based smartwatch, also arrive at a non-negligible error magnitude of 7.90 cm for the elbow and 9.20 cm for the wrist.

Since a measurement error in the upper single-digit centimeter range must be assumed, it seems sensible to merge locations of articulation that are close together spatially. Accordingly, all secondary locations of articulation are assigned to their respective primary locations of articulation. An exception is the primary location of articulation *Body*, since it has comparatively large secondary locations of articulation, which are merged as follows:

- *Neck* → *Head*, since the neck is closer to the head than, for example, to the shoulders.
- *Clavicle* → *Shoulder*, due to its proximity to the shoulders.
- *Shoulder* → *Shoulder*
- *TorsoTop* → *TorsoTop*
- *TorsoMid* → *TorsoTop*, together with *TorsoTop* it defines the chest area.
- *TorsoBottom* → *TorsoBottom*
- *Waist* → *TorsoBottom*, together with *TorsoBottom* it defines the abdominal area.
- *Hips* → *TorsoBottom*, together with *TorsoBottom* it defines the abdominal area.
- *BodyAway* → *Neutral*, since the neutral location of articulation is approximately at the same position, only slightly farther away from the body.

For the same reason, the secondary locations of articulation *WristFront* and *WristBack* are no longer assigned to the primary location of articulation *Arm*, but to the class *Hand*. The location of articulation *Arm* is also further differentiated by treating *UpperArm* as its own class, while the remaining secondary locations of the arm are jointly assigned to the class *LowerArm*. We thus obtain the following eight locations of articulation with their frequencies within the *ASL-LEX*: *Head* (26.58 %), *Shoulder* (2.52 %), *TorsoTop* (3.84 %), *TorsoBottom* (0.90 %), *UpperArm* (0.49 %), *LowerArm* (1.62 %), *Hand* (25.26 %), *Neutral* (38.86 %). The groupings are shown in Figure 55.

All of these locations of articulation refer to the position of the dominant hand. However, through Battison's sign types, we are able to obtain the missing information about the non-dominant hand. Due to the symmetry properties, for signs of type 1 (symmetrical signs) this takes the same location of articulation as the dominant hand. In asymmetrical signs (types 2 and 3), the non-dominant hand is in the neutral space in 713 out of 736 cases (28.97 % of all monomorphemic signs of the *ASL-LEX*), since it represents the location of articulation of the dominant hand. The remaining 23 signs must be manually annotated, with the majority of 18 signs being articulated in the neutral space, three signs at the upper torso, and two signs at the head or at the. For one-handed signs of type 0, by contrast, the non-dominant hand is outside the signing space and, for example when signing while standing, is left hanging at the side of the body, as can be seen in Figure 56. Shown is a three-dimensional visualization of the sensor data of the dominant (green) and non-dominant (red) arms during the

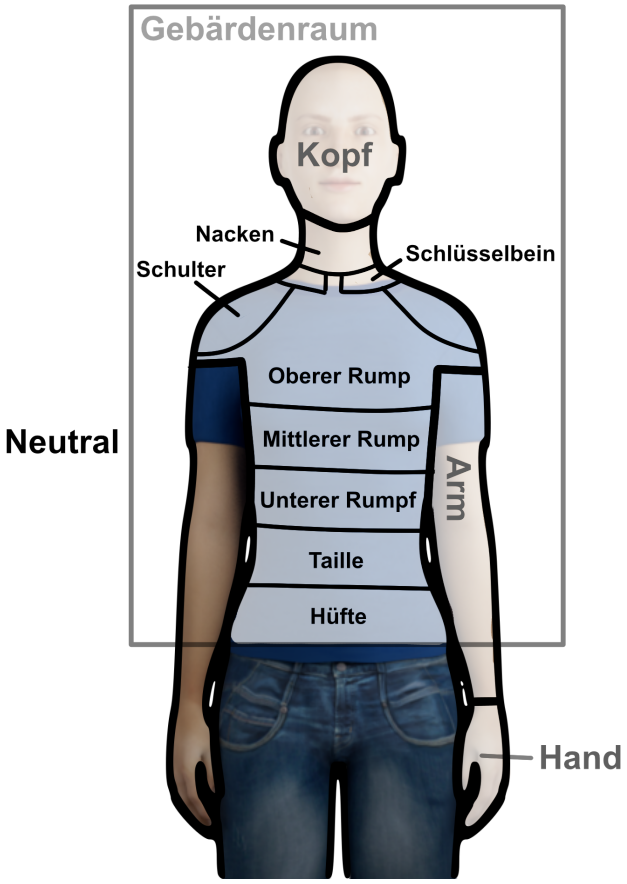
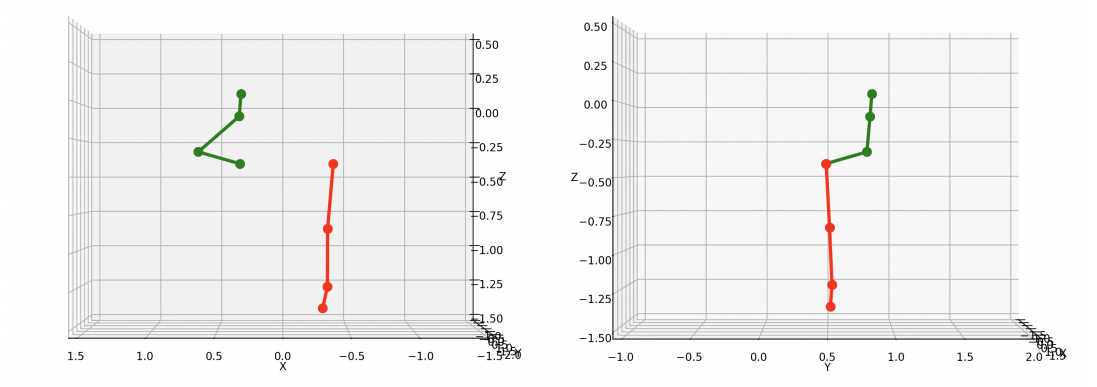



Figure 55: Grouped locations of articulation of the *ASL-LEX* except for *Hand* (adapted from [186])

execution of the one-handed sign BARRETTE. In this sign, the dominant hand is at the temple and the non-dominant hand is at the side of the body. This therefore requires a new class *None*, which is reserved exclusively for the non-dominant hand and is used for one-handed signs when the non-dominant hand is outside the signing space.

### 5.3.3 *Collection of the training data*

For assigning hand positions to locations of articulation and for the purpose of evaluating the approach, the positions of 18 locations of articulation were recorded from a total of six male and four female subjects. Some secondary locations of articulation were deliberately recorded in addition, although they had already been grouped previously or did not exist in this form before (e.g., *Face*). These locations of articulation are assigned to the primary locations of articulation later in the process, but they provide more variance in the training data. An overview of the recorded (secondary) locations of articulation and their assignment to the primary locations of articulation can be seen in Table 19.




(a) Frontal view (xz plane)

(b) Side view (yz plane)

Figure 56: Characteristic pose of the non-dominant arm (red) compared with the dominant arm (green) using the example of the one-handed sign BARRETTE, visualized quaternions of the arm controllers and data gloves

During data collection, two people were always present: a subject who performed the different poses and a recording person who operated the tool and monitored the execution. Before each recording, the hardware was put on, consisting of one arm controller each on the upper arm and forearm, as well as the *Manus Quantum Metagloves* data gloves. For calibration of the sensors, a T-pose as in Section 5.2 was performed and the data gloves were calibrated via the *Manus Core* SDK. An I-pose calibration was not used at this point; instead, alignment of the body with magnetic north was ensured, because the IMUs of the arm controllers align their orientation, among other things, with the magnetic field. This can be compensated for by the later use of I-pose calibration.

Before the start of each recording, the lengths of the upper arm and forearm and the distance from the wrist to the beginning of the index finger were measured individually for each participant. These values were stored in the recording tool before data acquisition in order to calculate the position of the hand according to Equation 17. Table 20 and Table 21 show the measured as well as normalized lengths for each participant, separated by gender.

Each participant was asked to touch the respective location of articulation ten times in succession with the center of the dominant hand, with the hand position varying slightly to generate a spatial spread of the data points. For each data point, the positions of elbow, wrist, and center of the hand were recorded for both arms or hands. The non-dominant hand was located outside the signing space, as in the execution of a one-handed sign. Only for signs with the primary location of articulation *Arm* or *Hand* was the non-dominant hand located in the neutral space, since it represented the location of articulation of the dominant hand.

After recording ten data points for a location of articulation, a renewed calibration in T-pose was performed to avoid systematic drift before proceeding to the next location of articulation. Thus, 100 data points could be recorded for each of the 18 locations of articulation, whereby for the recordings of *ArmAway* and *HandAway* care was taken

Table 19: Assignment of the recorded locations to the locations of articulation

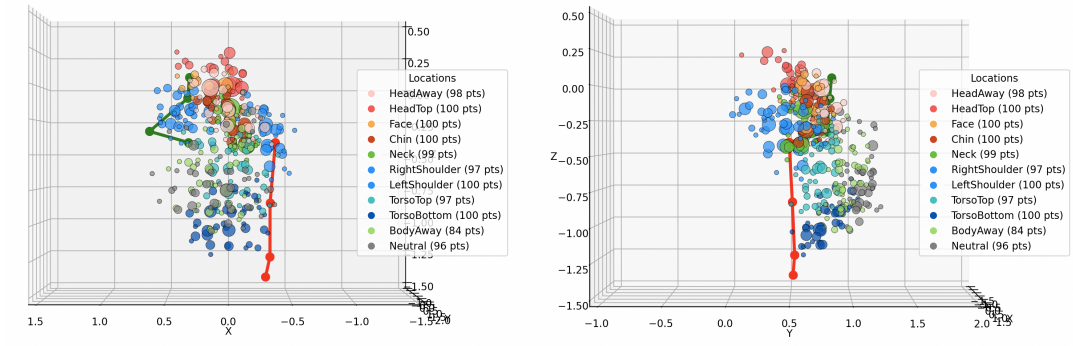
| Secondary location | Primary location | Count |
|---|---|---|
| HeadTop | Head | 100 |
| Face | Head | 100 |
| Chin | Head | 100 |
| HeadAway | Head | 100 |
| Neck | Head | 100 |
| RightShoulder | Shoulder | 100 |
| LeftShoulder | Shoulder | 100 |
| TorsoTop | TorsoTop | 100 |
| TorsoBottom | TorsoBottom | 100 |
| BodyAway | Neutral | 100 |
| UpperArm | Arm | 100 |
| Elbow | Arm | 100 |
| LowerArm | Arm | 100 |
| Wrist | Arm | 100 |
| ArmAwayAbove | Arm | 50 |
| ArmAwayBelow | Arm | 50 |
| Hand | Hand | 100 |
| HandAwayAbove | Hand | 50 |
| HandAwayBelow | Hand | 50 |
| Neutral | Neutral | 100 |
| None | None | 1,100 |

to ensure that half of the recordings were taken above the respective body part and the other half below the respective body part. From the non-dominant hand, only the 1,100 recordings outside the signing space (location of articulation *None*) were further used. The 700 recordings in the neutral space were excluded, however, because there the non-dominant hand occupied a comparatively narrow spatial region in order to serve as the location of articulation of the dominant hand. Their inclusion would have weighted this specific area within the neutral space disproportionately strongly, whereas the recordings of the dominant hand are distributed much more evenly in this area.

Figure 57 shows, in the already familiar visualization of the execution of the sign BARRETTE, the recorded positions of the dominant hand center after performing an *outlier detection* relative to the dominant (green) and non-dominant (red) arm. Spatially close data points were aggregated into larger markers, whose size is proportional to the number of associated data points.

Table 20: Anthropometric data of the subjects – absolute lengths in cm

| Code | Gender | Age | Upper arm | Forearm | Palm | Total |
|---|---|---|---|---|---|---|
| MP | m | 22 | 28.50 | 25 | 10 | 63.50 |
| SL | m | 25 | 31 | 29 | 11 | 71 |
| LK | m | 22 | 30 | 27 | 10 | 67 |
| LO | m | 22 | 31 | 27 | 10 | 68 |
| PN | m | 22 | 29 | 25.50 | 10 | 64.50 |
| MK | m | 23 | 26 | 23 | 9 | 58 |
| PK | m | 25 | 28.50 | 25 | 10 | 63.50 |
| **Average** | **m** | 23 | 29.14 | 25.93 | 10 | 65.07 |
| IP | w | 20 | 27 | 24 | 9.50 | 60.50 |
| SS | w | 21 | 28 | 25 | 9 | 62 |
| EK | w | 21 | 29 | 26 | 10 | 65 |
| RS | w | 27 | 28 | 24 | 10 | 62 |
| **Average** | **w** | 22.25 | 28 | 24.75 | 9.63 | 62.38 |
| **Total** | | 22.55 | 28.70 | 25.45 | 9.86 | 63.90 |




(a) Frontal view (xz plane)

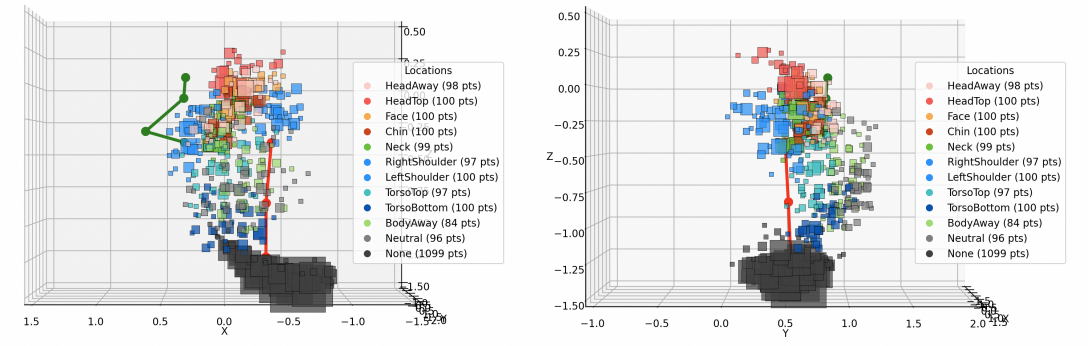
(b) Side view (yz plane)

Figure 57: Recorded positions of the dominant hand for the primary locations of articulation *Head*, *Body*, and *Neutral* relative to the dominant (green) and non-dominant (red) arm after performing outlier detection. Nearby data points were aggregated into larger markers, whose size corresponds to the number of data points.

Figure 58 shows the positions considered for the recognition approach of the non-dominant hand center, which essentially result from the mirrored and transferred positions of the dominant hand and are supplemented by recorded positions of the class *None*.

Table 21: Anthropometric data of the subjects – normalized lengths

| Code | Gender | Age | Upper arm / total | Forearm / total | Palm / total |
|---|---|---|---|---|---|
| MP | m | 22 | 0.45 | 0.39 | 0.16 |
| SL | m | 25 | 0.44 | 0.41 | 0.16 |
| LK | m | 22 | 0.45 | 0.40 | 0.15 |
| LO | m | 22 | 0.46 | 0.40 | 0.15 |
| PN | m | 22 | 0.45 | 0.40 | 0.16 |
| MK | m | 23 | 0.45 | 0.40 | 0.16 |
| PK | m | 25 | 0.45 | 0.39 | 0.16 |
| **Average** | **m** | 23 | 0.45 | 0.40 | 0.15 |
| IP | w | 20 | 0.45 | 0.40 | 0.16 |
| SS | w | 21 | 0.45 | 0.40 | 0.15 |
| EK | w | 21 | 0.45 | 0.40 | 0.15 |
| RS | w | 27 | 0.45 | 0.39 | 0.16 |
| **Average** | **w** | 22.25 | 0.45 | 0.40 | 0.15 |
| **Total** | | 22.55 | 0.45 | 0.40 | 0.15 |




(a) Frontal view (xz plane)

(b) Side view (yz plane)

Figure 58: Recorded positions of the non-dominant hand for the primary locations of articulation *Head*, *Body*, *Neutral*, and *None* relative to the dominant (green) and non-dominant (red) arm after performing DBSCANs. Nearby data points were aggregated into larger markers, whose size corresponds to the number of data points.

### 5.3.4 *Determining the location of articulation*

In order to infer the corresponding (primary) location of articulation from the estimated hand-center coordinates, a distance-based similarity score is used for class as-

signment. As a preparatory step, outliers are removed from the training data using DBSCAN. Since at least five recordings per participant were taken for each secondary location of articulation, all points are removed that are farther than one palm length from other points with at most three neighboring points ($\varepsilon = 0.15, \text{min\_samples} = 5$). This affects approximately 0.02 % of the recorded positions. In addition, all recorded points of the dominant hand are copied into the coordinate system of the non-dominant hand and mirrored. Then, starting from the position of the hand center, all distances $d_i$ to the recorded training data are calculated. These distances are then normalized using the formula $s_i = 1 - \frac{d_i}{r}$, whereby only training points within a radius of $r = 0,5$ are considered (since the system is normalized, a value of 1 corresponds to the distance from the upper arm to the hand center) in order to reduce the mathematical effort. The score can thus also be derived directly from the distance: points with $d_i \geqslant r$ receive a score of 0, points at the exact reference position of the hand center receive a score of 1. All points in between receive a linearly scaled share of the score.

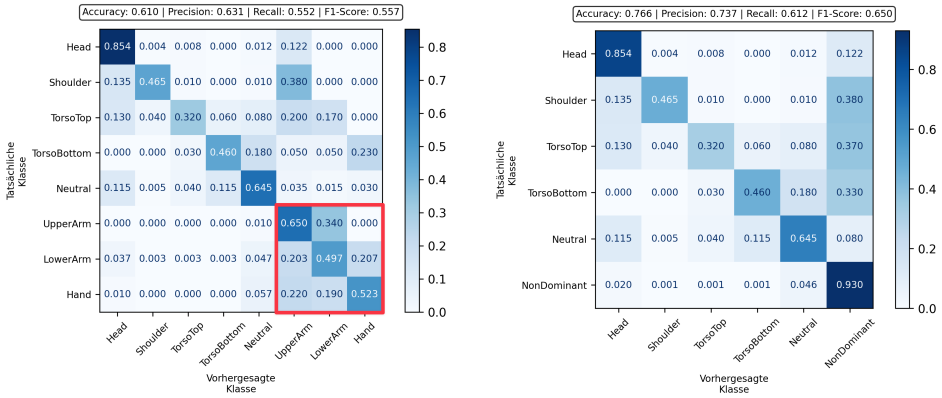
For each secondary location of articulation, the mean score from these scores is then determined, considering only those points that belong to the location of articulation. The score of the primary location of articulation is then determined as the maximum mean score of all associated secondary locations of articulation. For *Shoulder*, for example, this would be the highest score among *RightShoulder* and *LeftShoulder*, since these areas are not contiguous and therefore must be captured separately. In order for the sum of all scores of the primary locations of articulation to equal 1, these are then normalized.

The approach assumes that the locations of articulation are static and do not change over time or between different people, for example due to differing body proportions. To reduce person-specific differences, the locations of articulation are derived from recordings of ten different test subjects, resulting in a general and universal body model. With regard to temporal invariance, it can be stated that the head and upper body remain largely stable relative to the signing person – especially within a communication situation in which the interlocutors do not move significantly in space. However, this assumption does not apply to the non-dominant arm and the non-dominant hand, since during signing they are moved along depending on the sign and can therefore change position. We therefore compute the scores for these dynamic locations of articulation not from the distances to their recorded, static positions, but rather from the actual positions of shoulder, elbow, wrist, and center of the hand according to Equation 17 in real time. For the arm, the distances to shoulder, elbow, and wrist are taken into account, and for the hand the distances to wrist and hand center.

### 5.3.5 *Evaluation of the results*

For the evaluation of our approach, a LOSOCV cross-validation was performed across all participant data. The recorded positions of one participant were compared with the recorded positions of the other participants according to the method described above, and the score was calculated. The positions recorded for the primary locations

of articulation *Arm* and *Hand* were also taken into account, since in addition to the position of the hand center, the positions of elbow and wrist were also recorded and the pose of the non-dominant arm can be reconstructed from this data. Figure 59a shows the aggregated result of all ten runs in the form of a confusion matrix.




(a) Upper arm, forearm, and hand separated (red outline)

(b) Upper arm, forearm, and hand combined into *NonDominant*

Figure 59: Confusion matrices of the classifier for determining the location of articulation of signs for the dominant hand

The $F_1$ score for recognizing eight different locations of articulation is approximately 55.70 % across all ten runs. It is noticeable that confusion occurs particularly with the classes of the non-dominant arm and the non-dominant hand, both within these classes and with other classes, whereby other classes are often recognized as a class of arm or hand, but not the other way around. Since the confusion within these classes is also particularly high, it makes sense to merge the classes *UpperArm*, *LowerArm*, and *Hand* into a new class *NonDominant*. The results of this aggregation are shown in Figure 59b and, as expected, are significantly better, with an average $F_1$ score of about 65 %.

The reason for this high number of confusions with classes of the non-dominant arm and the non-dominant hand is certainly that dynamic reference positions are used here instead of static ones, as is the case for the other classes. This can lead to overlaps of these locations of articulation with other static regions, for example when the non-dominant hand is in the neutral space. One example where this is clearly visible is the sign BRA, in which the execution of both hands starts in the middle of the chest and ends at the outer edge. The recognition approach, due to the proximity to the other arm, incorrectly identifies the location of articulation at the beginning of the sign as *NonDominant* rather than as *TorsoBottom*.

However, since this should only be the case when the actual location of articulation of the dominant hand is the non-dominant arm or the non-dominant hand, this is less significant. To prevent the non-dominant arm from being recognized when it is in fact a one-handed sign, one could exclude it from the capture, for example by setting

the score for the affected locations of articulation to 0. This is, however, not necessary, since these cases are automatically removed as part of the meta-classification.

Another factor contributing to the inaccurate results of the approach is the already mentioned high error rate. Added to this are the differing body proportions of the users, which are difficult to account for in the approach since, for an ideal result, all body parts would have to be measured. Among the participants in our experiment, participant *SL* had the smallest ratio between upper arm and forearm and hand (normalized values, $0.437 : 0.563$, see Table 21). For a pose in which the upper arm is stretched forward and the forearm with outstretched hand is rotated by $90°$, the center of the hand would have the relative coordinate $(0.437, 0.563)$ with respect to the shoulder. If one instead considered participant *LO*, who has the largest ratio of upper arm to forearm and hand, his hand center would be at coordinate $(0.456, 0.544)$ for the same pose. The deviation between the two participants would be $\sqrt{(0.456-0.437)^2+(0.544-0.563)^2} = 0.027$, i.e., for the same pose, despite normalized lengths, a deviation of 1.91 cm would arise in addition to the already mentioned deviation (assuming 71 cm total arm length as the longest total length in the participant group, see Table 20).

Alignment toward magnetic north as compensation for the missing I-pose calibration or a poorly performed I-pose calibration can also make a difference of a few degrees, which, over the full arm length, can lead to deviations of several centimeters.

***Note** on recognizing contact between body parts*

The phonological property CONTACT of the *ASL-LEX* is binary and indicates whether the dominant hand touches the location of articulation at any point during the execution of the sign. However, the approach we chose for determining the location of articulation has an error rate in the upper single-digit centimeter range in favor of a mobile approach, so it would hardly be possible to reliably capture this value. We therefore dispense with capturing the phonological property CONTACT.

## 5.4 RECOGNITION OF HAND MOVEMENTS

Hand movements play a decisive role in sign language. While static properties can be captured comparatively well by individual snapshots – for example with handshape recognition in Chapter 4 – dynamic aspects are more complex. In addition, they also take into account the change of these individual static snapshots. They are subdivided into external movements of the hand – translational and rotational movements of the entire hand – and internal or finger-intrinsic movements. The aim of this chapter is therefore to explain how these hand movements can be detected with the help of data gloves and our IMU-based arm controllers.

The *ASL-LEX* features five phonological properties that describe *hand movements* and can be outlined as follows (for the full descriptions, including any exceptions, see Table A.2):

FLEXION CHANGE is an internal movement and indicates whether finger flexion changes during execution.

SPREAD CHANGE indicates whether finger spreading changes during execution.

ULNAR ROTATION is an external movement and indicates whether the wrist is rotated about the longitudinal axis during execution or remains at rest.

PATH MOVEMENT is an external movement and indicates the type of hand movement in three-dimensional space (e.g., from one position in signing space to another).

REPEATED MOVEMENT indicates whether at least one of the movements mentioned here is repeated (unchanged).

FLEXION CHANGE, SPREAD CHANGE, and ULNAR ROTATION each indicate whether or not there is a change in the reference quantity. They do not indicate the extent or manner in which this movement is present[73]. PATH MOVEMENT, on the other hand, is significantly more complex, as it provides not only the information of whether there is movement of the hand, but also the type of path this movement follows. This can be one of six different movement paths, for example a straight-line movement or a curved one. The class *BackAndForth* is not considered further here, as it originates from a previous version of the *ASL-LEX* and cannot be translated unambiguously into the classes of the current version, or overlaps with several current classes. The class *Other* is likewise not considered, as it is not defined in more detail [41, 187]. The phonological property REPEATED MOVEMENT does not, strictly speaking, represent a movement type of its own, but rather provides information about the movements present.

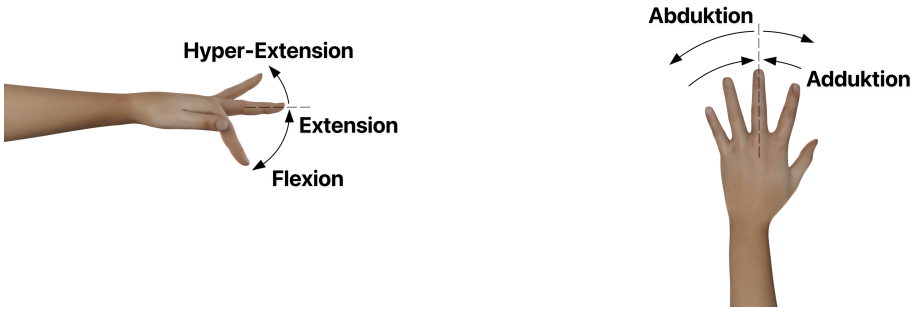
Overall, the majority of signs exhibit at least one type of movement. Only 155 (5.69 %) of 2,723 signs in the *ASL-LEX* have no movement. If signs with more than one morpheme are excluded, as we do for *ParaSignRec*, there are only 38 (1.54 %) of 2,461 signs.

### 5.4.1 *Threshold-based movement recognition*

Finger movements can refer to individual fingers or to all fingers collectively, with the thumb excluded because of its anatomical peculiarities. In the *ASL-LEX*, these movements occur in the form of a change in flexion (FLEXION with flexion/extension and extension of the fingers, e.g. PHOTOGRAPHY, see Figure 60a) and a change in spreading (SPREAD with abduction/spreading and adduction/closing of the fingers, e.g. SCISSORS, see Figure 60b).

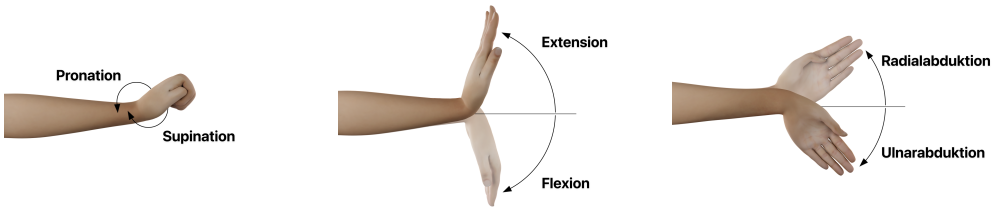
By contrast, the *ASL-LEX* marks signs with wrist rotation when the wrist rotates during sign execution (e.g. in APPLE or RADIO). This rotation corresponds to pronation/supination and is shown in Figure 61a. It is explicitly stated that other wrist movement possibilities are incorporated into the parameter PATH MOVEMENTS, e.g. in YES (flexion/extension, see Figure 61b) and VITAMINS (radial abduction/ulnar abduction, Figure 61c) – both as straight-line movement [112, 187].

73 For FLEXION CHANGE, the additional information is given that the change does not have to be sufficient for the signs to be classified into a different FLEXION category [186].




(a) Flexion/Extension (b) Abduction/Adduction

Figure 60: The finger movements relevant for flexion (flexion/extension) and spreading (abduction/adduction) [169]




(a) Pronation/Supination (b) Flexion/Extension (c) Radial abduction/Ulnar abduction

Figure 61: Degrees of freedom of the wrist (adapted from [84, 112])

For our recognition approach, we now need to develop suitable classifiers that are capable of classifying these movements. *ASL-LEX* specifies whether one of these movements is present or not, which is why these are binary classifiers. For the finger-intrinsic movements, SELECTED FINGERS indicates which of the four fingers perform the movement. The data glove provides, for each finger, a value for the spread of the MCP joint, one value each for the flexion of the MCP, PIP, and DIP joints, and a quaternion describing the rotation of the back of the hand. Due to the manageable complexity, threshold-based approaches are therefore suitable: a comparison value $v(x)$ is generated from the available sensor data and compared with a threshold $\tau$. If the comparison value lies above the threshold – i.e., $v(x) \geqslant \tau$ – then the class *Movement* is present; otherwise, the class *No Movement*. This type of classification is transparent and usually efficient, provided that no complex calculations are required to determine the comparison value.

However, since we pursue a probabilistic meta-classification, we do not require a hard class assignment, but instead use continuous *scores* with $s(x) \in [0, 1]$ as the basis for evaluation. The highest score indicates the most likely prediction, and the sum of all scores within a classifier equals 1, accordingly $s_{pos}(x) = 1 - \sum s_{neg}(x)$ applies, with $s_{pos}(x)$ as the score for the correct class and $s_{neg}(x)$ as the scores for the incorrect classes. Since these are binary classifiers, a score below 0.50 would mean that no movement is present, whereas a score above 0.50 would indicate movement. The further the score is from 0.50, the more certain the decision. A score of 1, for example, would mean that a movement is present with 100 % certainty. In addition to determining a threshold, it is therefore also necessary to define those ranges $L \leqslant \tau \leqslant U$ in which a movement or no movement is present with certainty. For $v(x) \leqslant L$ we set $s(x) = 0$ ("certainly negative") and for $v(x) \geqslant U$ we have $s(x) = 1$ ("certainly positive"). The range $L < v(x) < \tau$ and $\tau < v(x) < U$ is linearly interpolated using Equation 18, with $v(x) = \tau$ continuing to yield 0.50.

$$s(x) = \frac{v(x) - L}{2(\tau - L)}, v(x) \leqslant \tau \qquad s(x) = \frac{v(x) - \tau}{2(U - \tau)} + \frac{1}{2}, v(x) > \tau \tag{18}$$

These upper and lower bounds therefore have no influence on the direct classification result, since this depends only on the threshold. However, they do influence the result of our meta-classification, because an overly restrictive interpretation can cause almost all classes to be evaluated binarily, thereby significantly reducing the influence of the relevant parameter on the meta-classification. Conversely, an overly permissive interpretation can lead to a residual probability for the losing class even when the result is clear.

We therefore set the lower bound so that it is equal to or below the lowest value of the positive instances (signs with movement), whereas the upper bound is equal to or above the highest value of the negative instances (signs without movement). Machine-based detection of outliers is deliberately omitted so as not to exclude future valid signs performed under different conditions merely because of the specific selection of training signs. However, the data were manually validated prior to training using the process described in Section 5.2.5, so the presence of extreme values in the test data can be ruled out.

The optimal threshold is finally determined by evenly subdividing the range between upper and lower bound into 100 steps (thresholds). For each of these thresholds, the measured values are separated into instances of the positive class and instances of the negative class, and the respective $F_1$ score is calculated. The threshold with the highest $F_1$ score is selected.

*Reference signs for threshold determination*

To determine the optimal threshold and the bounds, recordings of real signs are to be examined. To this end, 38 signs[74] of the *ASL-LEX* were recorded and processed from 14 different participants under the conditions specified in Section 5.2. In each case, three to four different signs were recorded consecutively in random order, with each sequence repeated two to four times. Ten sequences with a total of 38 signs were recorded from each participant. The only exceptions are participant 13 with 34 signs and participant 14 with three signs. This gives a total of 1,123 sign recordings, with 27 to 34 recordings per sign.

Unfortunately, the arm controllers were not yet available in their current form at this time, which is why only the data from the two *Manus Prime X* data gloves were recorded.

All sign recordings were validated before further use for training the threshold-based recognition approaches using the process described in Section 5.2.5. For each sign, the eight parameters for FLEXION CHANGE (4x), SPREAD CHANGE (3x), and ULNAR ROTATION (1x) of the right hand were examined, i.e. 8,984 data records in total. The data from the left hand were not examined in order to keep the effort low and because 18 of the 38 signs were one-handed. Therefore, only data from the right hand were used for training.

Overall, 71.10 % of the sign executions were found to be error-free, i.e. none of the eight parameters was marked as erroneous. A total of 1,483 errors were found in 324 sign executions, of which 926 were erroneous sensor data, 508 erroneous executions, and 49 erroneous segmentations. Conversely, this means that approximately 1.32 errors were recorded per sign execution.

Figure 62 shows the distribution of error types across each individual sign. It becomes clear that there is a strong dependence of the error types on the signs and that they are not distributed evenly across all signs. If one also considers the quality of the individual parameters for each sign in Figure 63, this thesis is substantiated. Signs such as MEMORIZE, NOT, or LOVE exhibit few execution errors because they are distinctive and easy to execute.

WELCOME 1, UNIVERSE, and TIRED, by contrast, perform worst of all signs, which is attributable to the below-average SPREAD values. What these signs have in common is that the fingers are fully or at least predominantly open and a rather quick hand movement is performed, which apparently leads to incorrectly detected SPREAD differences. In the sign THANK YOU, for example, a *B* hand (flat hand) moves from the chin straight forward. Although the fingers do not move independently of the hand, spread changes of up to 5 to 8° are detected. Spread changes in this range should be clearly visible, but they were not discernible in the video recordings, otherwise it would also have been an erroneous execution. This behavior can also be observed

74 LOVE (34), STRESS (33), THANK YOU (32), NOT (32), MEMORIZE (32), DARK 1 (31), COMMENT (31), PRETTY (31), KEEP (31), SURPRISE (31), SORRY (31),OPINION 2 (30), BARELY (30), JACKET (30), SEE (30), TEAM (30), TIRED (30), REMOVE (29), UNIVERSE (29), SCRIPT (29), PLEASE (29), PROGRAM (29), HOSPITAL (29), HURRY (29), BRIGHT (29), PERCENT (28), DROP (28), ANGEL (28), THROW AWAY (28), CARROT (28), PANTS 2 (28), FILL OUT (28), WELCOME 1 (28), UNIQUE (28), UNDERSTAND (28), TWO OF US (28), SHY (27), CUTE 2 (27)

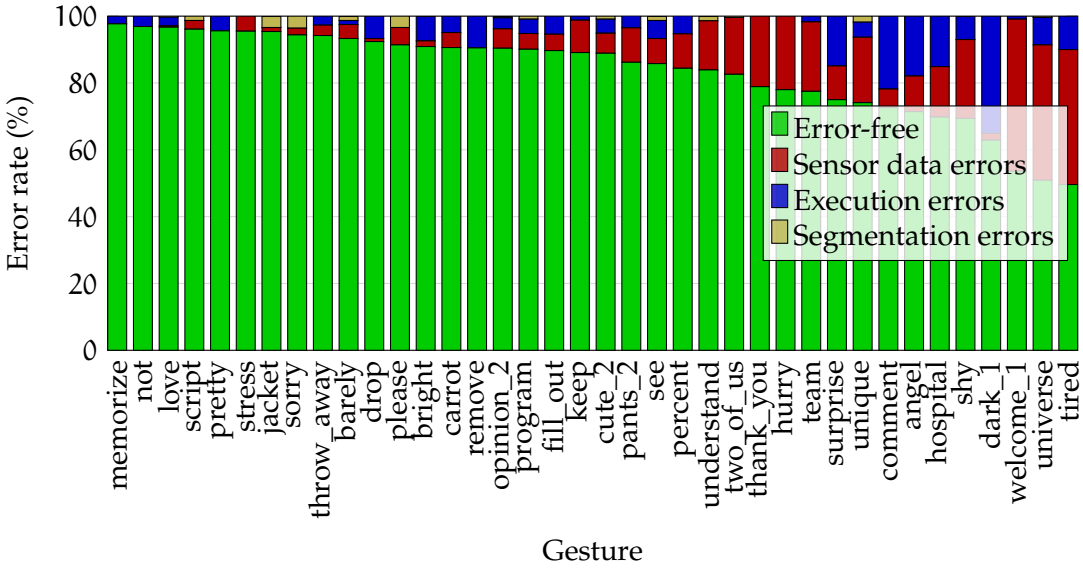



Figure 62: Distribution of error types across all gestures

in the signs UNIQUE, HURRY, and ANGEL. If the hand is closed, as in TEAM, for example, this phenomenon can also be observed, even though spreading is anatomically impossible with a closed hand.

Erroneous SPREAD values could also be observed in some signs, arising from contact with other parts of the body. These include signs such as TIRED, JACKET, SORRY, SHY, and SCRIPT. Among these signs there are also hands that are (predominantly) completely closed and spreading should not be possible.

For participant 3, we also observed that in the sign UNDERSTAND in three out of four recordings the sign was executed correctly (the hand forms a fist at the beginning, then the index finger is fully extended), but the maximum FLEXION change was only 9.09 to 22.71° and is barely perceptible. The correctly captured sign was recognized with 173.25°. The behavior did not change when we revalidated the data without applying the *Butterworth* filter in a test run. There is evidently an error in the data gloves here.

For the same sign, with participant 5, a change in the FLEXION value at the little finger of 40.52° was erroneously detected in parallel with the flexion of the index finger, which was likewise not discernible in the videos.

Signs such as ANGEL show only a very small change in the FLEXION value, which could likewise arise from quick hand movements or when a body part is touched.

The sign most frequently executed incorrectly is DARK 1, in which participants often mistakenly changed finger spreading. In addition, the sign in the *ASL-LEX* example video starts from the side rather than frontally, which is why a wrist rotation about the longitudinal axis of up to 90° was often incorrectly performed here as well. Such hand rotation was also often incorrectly performed in the signs PRETTY, BRIGHT, SURPRISE, and REMOVE.

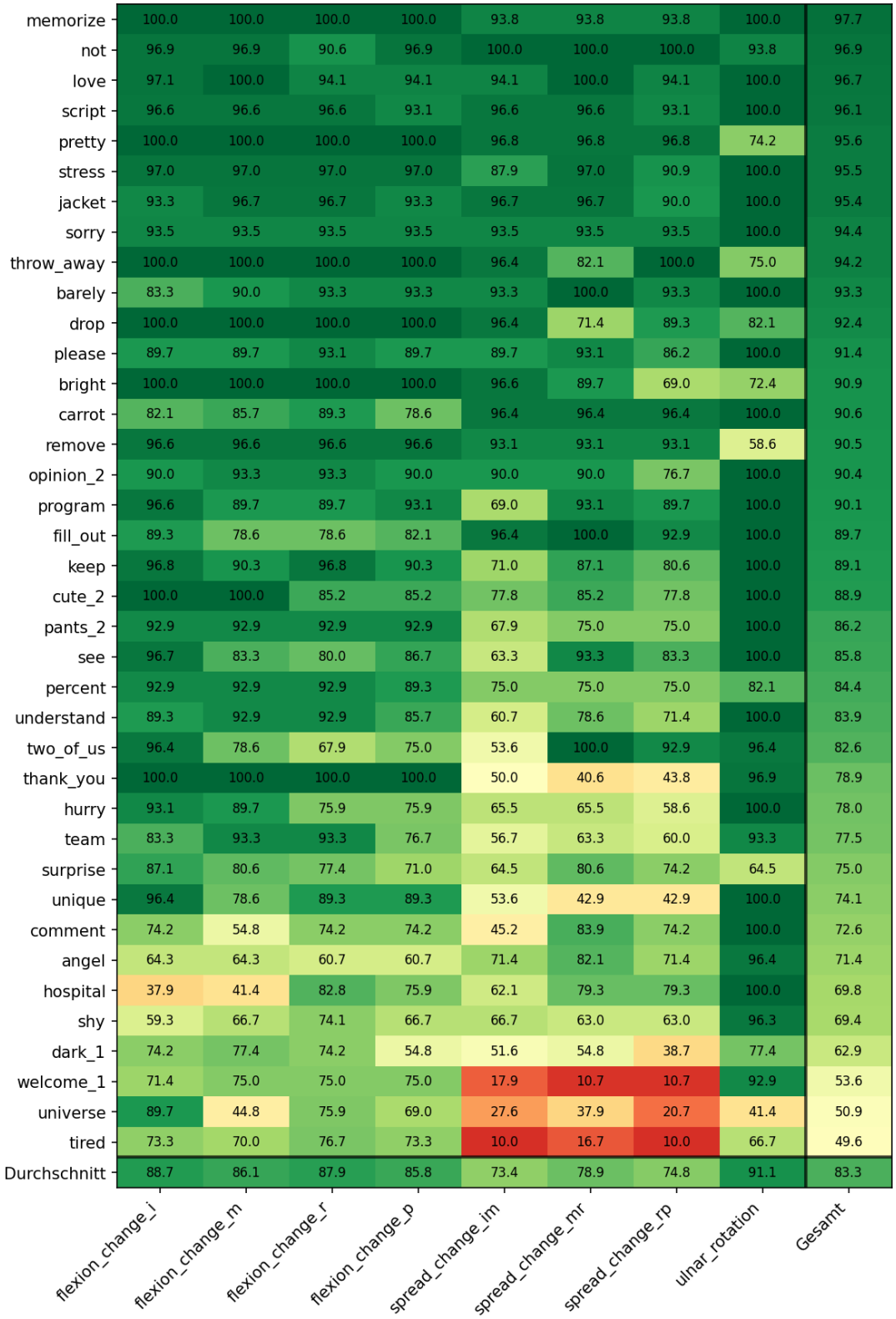



Figure 63: Quality of the individual parameters for each sign

For some signs, involuntary movements were observed that are also difficult to avoid anatomically. In the sign COMMENT, for example, the middle finger was often mistakenly moved as well, which led to an incorrect FLEXION value for the middle finger and incorrect SPREAD differences between the index and middle fingers. This can also be seen in the *ASL-LEX* example video. In general, it can be observed that a detected *Spread Change* is often correlated with a *Flexion Change*. In addition to COMMENT, this can also be observed for UNIQUE and ANGEL.

Some signs, in turn, were so pictorial in their execution that the participants had their own interpretation of how they should be performed: in the sign HOSPITAL, the "cross gesture" was often indicated by a path movement instead of a change in finger flexion. In the sign ANGEL, by contrast, the suggested "flapping of wings" often resulted from a movement of the wrist rather than from a change in finger flexion. It was also clearly observable that, among participants who performed the "flapping of wings" slowly, the SPREAD values were correct, whereas with faster execution there was a high error rate.

Figure 64 shows the distribution of error types across participants. Here too it becomes clear that the number of execution errors depends on the participant. There are therefore participants for whom it is easier to imitate signs from an example video than for others. The high sensor data errors of participants 3, 12, and 13 may on the one hand be due to anatomical reasons (e.g. hand size), but also to the execution of the signs. In all three participants, a noticeably faster execution of the sign movements can be observed than in the remaining participants. This, in turn, can – as already explained – lead to errors in the SPREAD values.

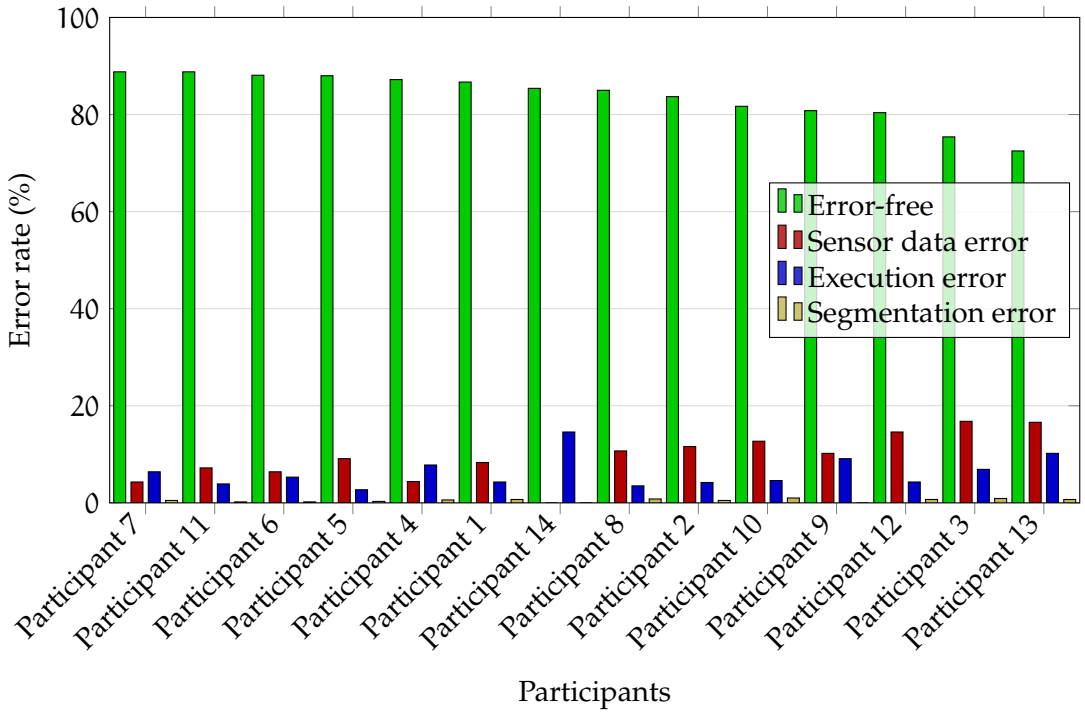



Figure 64: Distribution of error types across participants

Segmentation errors are largely independent of individual signs, participants, and parameters and generally occur rather rarely, since a sign execution is usually followed by a *Hold* phase in which no change can be perceived, which simplifies the selection of a suitable endpoint.

Figure 65 finally shows the distribution of error types across the eight parameters. Overall, 83.50 % of the parameters were recognized without error. Most errors were sensor data errors (10.30 %), with the majority attributable to the SPREAD values (58.20 % of all sensor data errors, or roughly three to four times as often as with the other parameters). Execution errors are found in approximately 5.70 % of the parameters, whereas segmentation errors at 0.50 % are rather negligible.

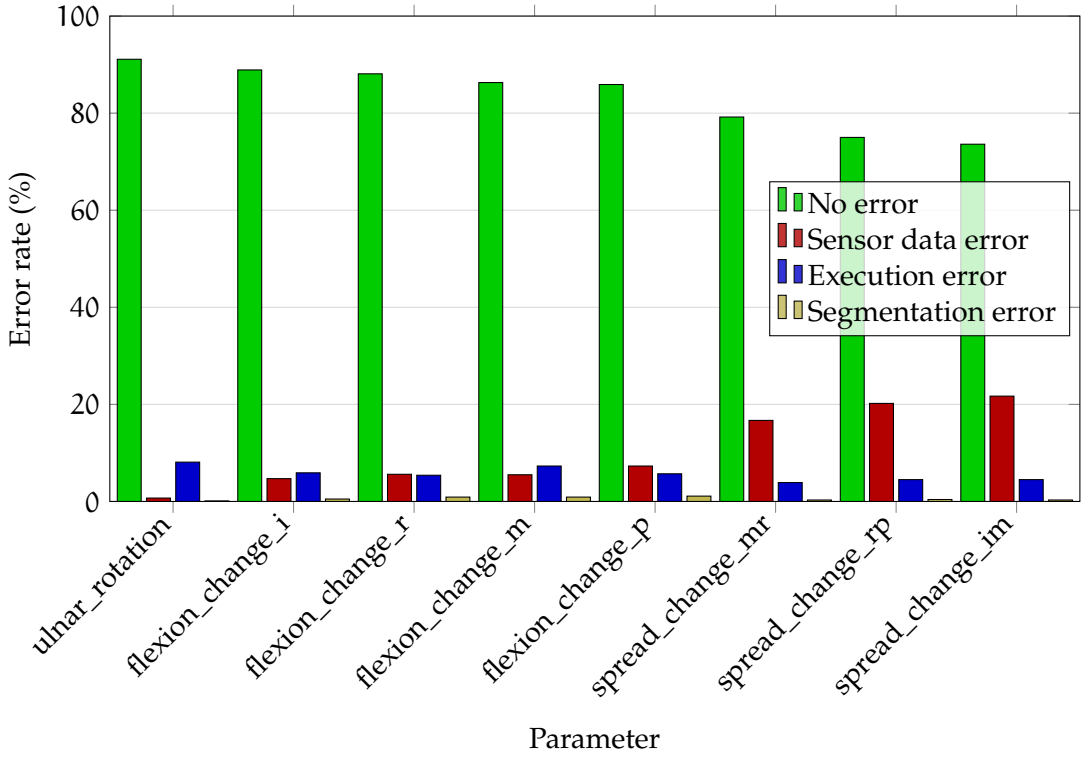



Figure 65: Distribution of error types across the eight parameters

In addition, a total of four signs (10.53 %) had erroneous annotations in the *ASL-LEX*, and as a result five parameters (1.64 %) had to be re-annotated.

Based on these findings and the recordings cleaned of erroneous data, a threshold-based approach for detecting changes in finger flexion and spreading as well as wrist rotation is now to be implemented and evaluated in the following.

*Finger flexion changes*

The flexion (bending) and extension (straightening) of the fingers occur at the three joints MCP, PIP, and DIP and are considered primary movements at the MCP. The range of motion is approximately 90° for the index finger during active flexion (without

external influences) and increases slightly up to the little finger. Maximum extension lies between 10° for the index finger and 30° for the little finger [112, 115].

(a) Closed fingers for handshape S (*FullyClosed*)

(b) Flexed fingers for handshape FLAT 4 (*Flat*)

(c) Flexed fingers for handshape CURVED 5 (*Curved*)

(d) Extended fingers for handshape B (*FullyOpened*)

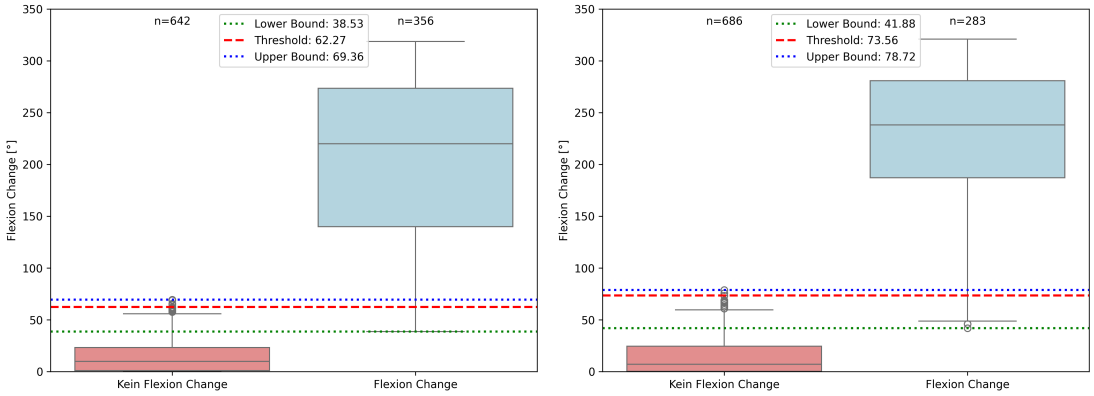
Figure 66: Different finger flexions

*ASL-LEX* defines a FLEXION value and a corresponding FLEXION CHANGE value for each sign. The FLEXION value indicates the flexion of the SELECTED FINGERS at the beginning of the sign and is described by the three joints MCP, PIP, and DIP. It can therefore take different values: The scale begins with *FullyClosed*, when all finger joints are fully flexed, and ends with *FullyOpened* (all joints extended). If only the MCP joints are flexed, the flexion is called *Flat*; the opposite (DIP and PIP joints are flexed) is called *Bent*. An even flexion of all three joints is a *Curved* flexion. *Stacked* and *Crossed* represent two special types of flexion, in which the selected fingers exhibit different flexions (as in the letter K of the finger alphabet) or the fingers are crossed (as in the letter R of the finger alphabet). FLEXION is *NA* when the thumb (*T*) is the only SELECTED FINGER. Figure 66 shows different examples of finger flexions. In handshape S, all fingers are closed into a fist and the corresponding FLEXION value is *FullyClosed*. On the other end of the scale is handshape B, in which all fingers are open and the corresponding FLEXION value is *FullyOpened*. The handshapes FLAT 5 and CURVED 5 represent intermediate stages of these flexion values. In a FLEXION CHANGE, the fingers change their flexion, for example to move from handshape S via CURVED 5 to handshape B.

In *ASL-LEX*, the value for FLEXION CHANGE can be either *False* (no change in flexion), *True* (change in flexion present), or *NA*. The latter value poses a problem for our recognition approach, as it contains a definitional gap here. This is always the case when FLEXION also has the value *NA*, i.e., when the thumb is the only SELECTED FINGER. According to the definition of SELECTED FINGERS, these are the fingers that *i)* move, *ii)* have a flexion value other than *FullyClosed* and *FullyOpened*, or *iii)* are fully extended (*FullyOpened*) (in that order). Since FLEXION CHANGE is a movement, the SELECTED FINGERS must match the moved fingers. Since there are also cases in which the thumb is the only SELECTED FINGER and a change in flexion (FLEXION CHANGE = *True*) is present, it follows that we can change all *NA* values to *False*.

It is therefore also possible to determine for each finger individually via the SELECTED FINGERS whether a FLEXION CHANGE is present or not. Instead of one binary value per hand, we now obtain four binary values, i.e., 16 different combinations of parameters.

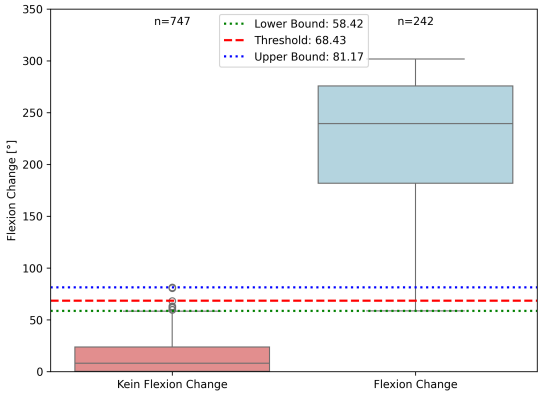
Since there are flexions that result from a change in the *MCP* joint (*Flat*) and flexions that result from a change in the *PIP* and *DIP* joints (*Bent*), as well as flexions that result from a change in all joints equally (*Curved*), the most sensible approach for determining a FLEXION CHANGE is to use the sum of all three joint-angle changes.



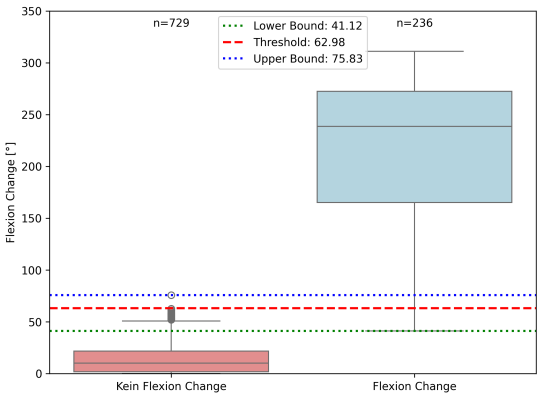

(a) Index finger




(b) Middle finger




(c) Ring finger




(d) Little finger

Figure 67: Distribution of flexion change values (°) across the two classes with thresholds and boundary values marked

In Figure 67, the distribution of the flexion values across the two classes *Flexion Change present* (*True*) and *Flexion Change not present* (*False*) is shown. The flexion values are obtained from the difference between the maximum and minimum sum of all three joint angles of an individual finger during the entire sign execution. The value ranges of the data gloves listed in Table 11 lie, depending on the joint, between 0 to 80° and 0 to 100°, resulting in a range of 0 to 270° for the sum. However, we had to determine that a newer version of the *Manus Core* SDKs also delivers values beyond these limits, which is why the value range shown in the boxplots is slightly higher. The lower

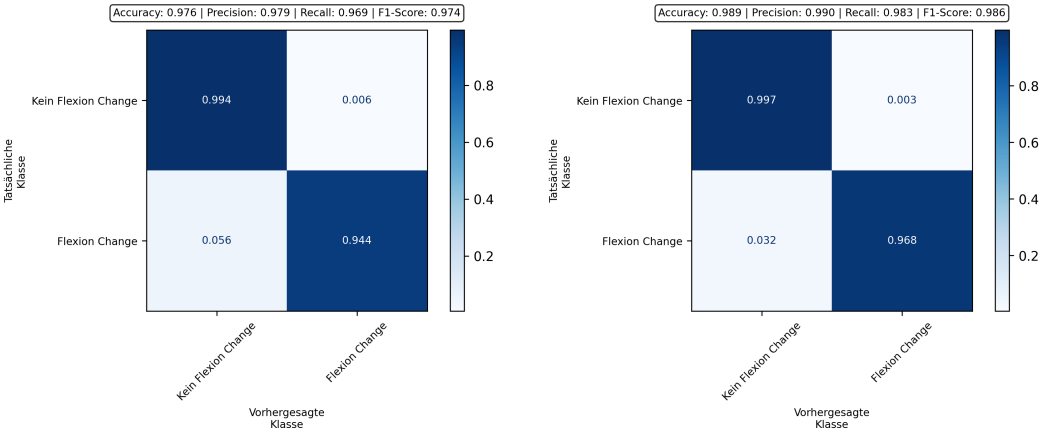
and upper thresholds correspond to the minimum of the positive and the maximum of the negative instances. The threshold lies between these two thresholds and is determined empirically by stepwise maximization of the $F_1$ score (cf. Section 5.4.1). The thresholds for all four classifiers are at a relatively uniform level, with less than 20° difference between the lower thresholds and less than 12° difference between the upper thresholds of the individual fingers (38.53 to 58.42° for the lower threshold, 69.36 to 81.17° for the upper threshold, 62.27 to 73.56° for the threshold value). Here, the index finger has the lowest thresholds and the ring finger the highest thresholds, whereas the middle finger has the highest threshold value. This roughly corresponds to the movement ranges mentioned at the beginning, where the index finger has the smallest range of motion and this then increases slightly for each finger.

The corresponding confusion matrices of the threshold-based classification, including the most important metrics, are shown in Figure 68. It is evident that all four classifiers yield similar results. All four have a very high $F_1$ score of 97.40 to 99.50 %, with the index finger showing the lowest value and the ring finger the highest. The individual class probabilities are also close for all four classifiers (99.40 to 99.90 % for correct recognition of a non-present flexion change (TN) and 94.40 to 99.20 % for correct recognition of a present flexion change (TP)), with the classifiers most frequently misclassifying an actually present flexion change as no flexion change (FN, 0.80 to 5.60 %).

*Finger spreading changes*

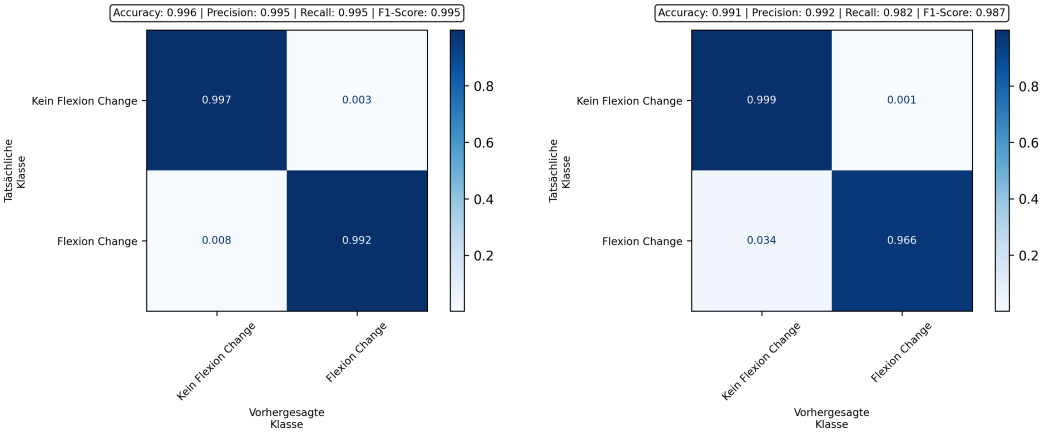
In addition to the abduction and extension of the individual fingers, lateral movement of the fingers is also possible. If the fingers move apart and are spread, this is referred to as *abduction*. The opposite process is called *adduction*. Figure 60 illustrates these movements. The middle finger forms the axis of the hand, i.e., the movements are to be regarded as symmetric about it. If the MCP joints are flexed, neither abduction nor adduction is possible. Depending on the source, the range of motion for abduction lies between 20 to 30° and for adduction approximately at 0°, with the index finger exhibiting the greatest lateral mobility in both directions [112, 115, 169, 171]. Considering the fingers individually, the maximum lateral mobility (abduction and adduction) lies at 0 to 35.20° between the index finger and middle finger, 0 to 25.70° between the middle finger and ring finger, and 0 to 28.40° between the ring finger and little finger [80]. The range of motion for men is generally larger than for women. These figures refer to active movements, that is, those without external influence.

*ASL-LEX* defines a Spread value and a corresponding Spread Change value for each sign. Both values refer to the Selected Fingers and indicate whether a spread or a change in spread is present, but not to what extent. Figure 69 shows two different examples of finger spreading. In handshape B, all fingers are together and the corresponding Spread value is *False*. In handshape 5, by contrast, all fingers are spread and the corresponding Spread value is *True*. In a Spread Change, the fingers change their spread, for example in order to transition from handshape B to handshape 5. It is easy to see that the outer fingers cover a greater distance between these two states and the middle finger hardly moves at all.




(a) Index finger

(b) Middle finger




(c) Ring finger

(d) Little finger

Figure 68: Confusion matrices for the detection of flexion change by finger

(a) Closed fingers for handshape B

(b) Spread fingers for handshape 5

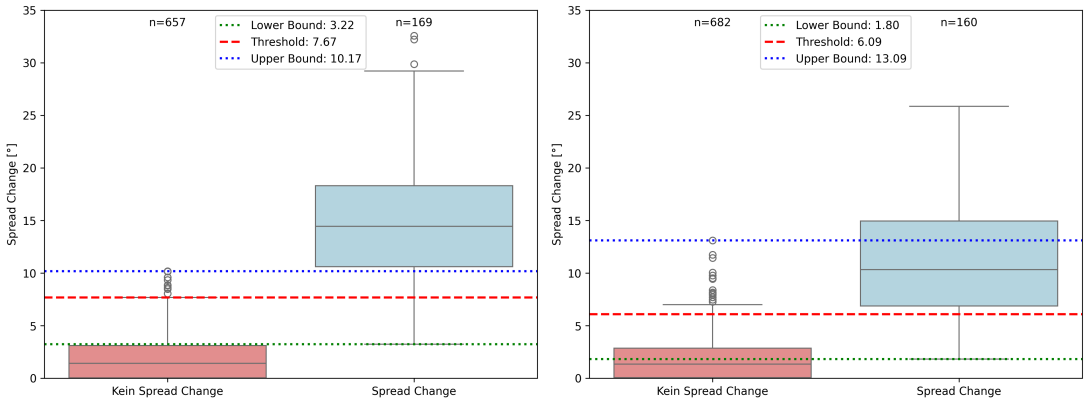
Figure 69: Different spreads of the fingers

Analogous to FLEXION CHANGE, *ASL-LEX* also has undefined fields for the values of SPREAD CHANGE, for which a rule for resolution must be found: A spread of the fingers can, by definition, only be present with more than one SELECTED FINGER. According to *ASL-LEX*, these must also be adjacent. Conversely, this means that SPREAD always has the value *NA* if the SELECTED FINGERS consist of only one finger or of non-adjacent fingers. In addition, FLEXION must not have the values *FullyClosed* or *Stacked*. The former, because the fingers form a fist and a spread is therefore anatomically not possible; the latter, because *Stacked* essentially represents an overlap of the thumb with the index and middle fingers (see K and STACKED 5 in Table A.4) and thus already includes a spread of these fingers.

If, for the reasons just mentioned, an undefined spread (SPREAD = *NA*) is present, SPREAD CHANGE can assume the values *True*, *False*, and *NA*. For our recognition approach, we must replace *NA*. This affects 1,136 of the original 2,723 signs in *ASL-LEX*, and is therefore not negligible. We therefore define that no change in spread is present in the following cases: with only one selected finger (925 signs), with the thumb as the selected finger (156 signs), and with non-adjacent fingers (19 signs). Of the remaining 36 signs, we assume that no change in spread is present for all signs with a FLEXION of *Flat* or *FullyClosed* and without FLEXION CHANGE, since the MCP joint must not be flexed (11 signs). The same applies to signs with FLEXION *Stacked* and without FLEXION CHANGE (12 signs). Thirteen signs remain, but only two of them are monomorphemic (DOG and THINK PONDER) and were manually annotated. These two signs do not exhibit a change in spread. There are therefore 194 out of 2,723 signs (7.12 %) with a change in spread in *ASL-LEX*.

For our recognition approach, we now consider the fingers pairwise rather than finger by finger as before. This is due to the fact that the middle finger hardly moves and that a spread always requires at least two adjacent fingers. Three combinations arise from four fingers. The study by Gracia-Ibáñez et al. [80] has shown that all three combinations can be expected to exhibit distinctive movements. The movement between the index finger and middle finger has a somewhat larger range of motion than the other two combinations and is at the same time the combination that occurs in every sign with a change in spread (100 % in 194 signs). The other two combinations occur (always together) in 170 signs (87.63 %) with a change in spread. Since using both classifiers does not yield any additional information about the signs, it is sufficient to consider only the spread between the index and middle fingers, as well as one of the two remaining classifiers, even though the recorded data have already been examined for all three changes in spread. Between the ring finger and little finger there is greater mobility than between the middle and ring fingers, which is why this classifier is chosen.

In Figure 70, the distribution of the determined values across the two classes *Spread Change present* (*True*) and *Spread Change not present* (*False*) is shown. The values always refer to two adjacent fingers and are derived from the maximum range of the difference between both spread values during the entire sign production. The value ranges of the data gloves listed in Table 11 lie, depending on the finger, between $-20$ to $20°$, which leads to a maximum difference per finger pair of $40°$. As before, the two thresholds




(a) Index finger/Middle finger

(b) Ring finger/Little finger

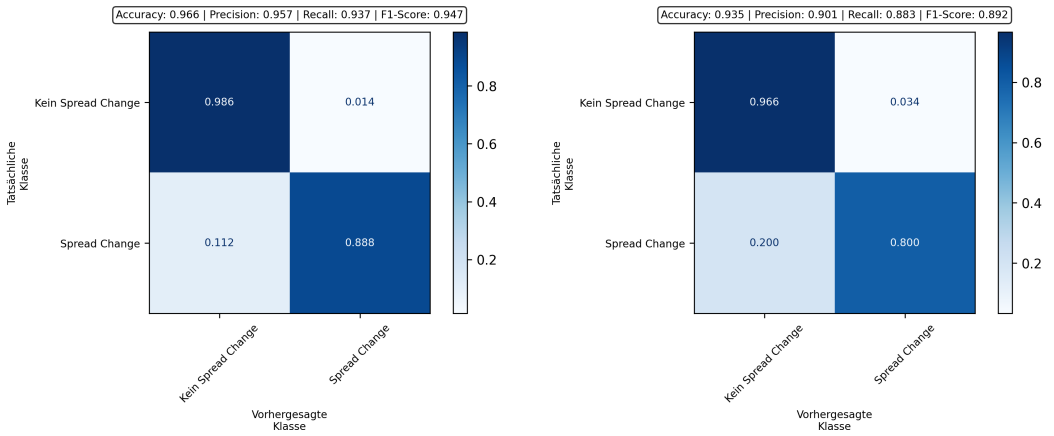
Figure 70: Distribution of the differences in spread values (°) across the two classes with threshold and boundary values indicated

as well as the threshold value are determined according to the procedure described in Section 5.4.1. The threshold value is at a similar level for both classifiers, at 7.67° compared with 6.09°. The thresholds for the ring finger/little finger classifier exhibit a larger range (1.80 to 13.09°) than those of the index finger/middle finger classifier (3.22 to 10.17°), although, as mentioned at the outset, the maximum lateral mobility between the index and middle fingers is the highest. The narrower this band, the greater the classifier's discriminative power, i.e., the better it can distinguish between the two classes. It follows that the index finger/middle finger classifier can achieve a better separation of the classes, which can also be attributed precisely to the larger range of motion.

This is also reflected when considering the corresponding confusion matrices in Figure 71. The classifier for the index finger/middle finger can achieve an $F_1$ score of 94.70 %, approximately 5.50 percentage points higher than the classifier for the ring finger/little finger (89.20 %). The ability to correctly detect the absence of a flexion change (TN) is roughly the same for both classifiers (98.60 % versus 96.60 %); however, the classifier for the ring finger/little finger has considerably greater difficulty correctly detecting an actual change in spread as well (TP, 88.80 % versus 80 %).

*Wrist movements*

Besides finger-intrinsic movements, wrist rotation recognition is also a suitable threshold-based recognition approach. Here, only rotation around the longitudinal axis of the hand is meant (pronation/supination, cf. Figure 61), since other possible wrist movements are explicitly included in the parameter PATH MOVEMENTS. Supination and pronation describe the rotation of the wrist around the longitudinal axis of the hand. From a neutral position – the palm is vertical, the thumb points upward – supination means rotating the palm "upward" by up to 90°. Rotating the palm "downward" is




(a) Index finger/Middle finger

(b) Ring finger/Little finger

Figure 71: Confusion matrices for detecting the change in spread by finger pair

pronation and can amount to a maximum of 85°. This yields a maximum range of motion of 175° [112].

The specification for ULNAR ROTATION is clearly defined by the *ASL-LEX* for each of the 2,723 signs, so that we do not need to resolve any definitional gaps here. However, our detailed examination of the data recordings at the beginning of this chapter showed that there are significant differences on the part of *ASL-LEX* in the assessment of ULNAR ROTATION for the signs examined. Even though, according to *ASL-LEX*, only a rotation of the wrist around the longitudinal axis is meant, for some signs (e.g., WELCOME 1 or SWEATER) the movement of the arm is also taken into account in the assessment. As contrasting examples, ANGEL, TIRED, and DARK 1 should be mentioned, in which similar arc-like movements of the arm are performed that result in a rotation of the hand, but explicitly without ULNAR ROTATION. It is not possible to implement a recognition approach under these conditions, which is why the data must be annotated again. Either the arm movement is included in the assessment of ULNAR ROTATION (absolute rotation of the hand), or one restricts the annotation to the movement of the wrist alone (relative rotation of the hand). In our approach, we pursue the absolute rotation of the hand, i.e., including arm movement, and re-annotate all signs used by us.

In addition, for signs such as PLAY, SENTENCE, PARANOID, or EASTER, there are sometimes very small rotations of the wrist (more of a "shaking"), yet they are annotated with ULNAR ROTATION. Since these changes are difficult to detect, they were also re-annotated by us (without ULNAR ROTATION). Of 250 examined signs, 23 changes in ULNAR ROTATION were made (9.20 %).

Figure 72 shows the distribution of the determined rotation values across the two classes *Ulnar Rotation present* (*True*) and *Ulnar Rotation absent* (*False*). For this purpose, for each frame $i$, the quaternions of the hand orientation provided by the data gloves were converted into a normalized unit quaternion $q_i$ and the corresponding rotation matrix $R_i$. To determine the rotation matrix, Equation 19 is applied [194].

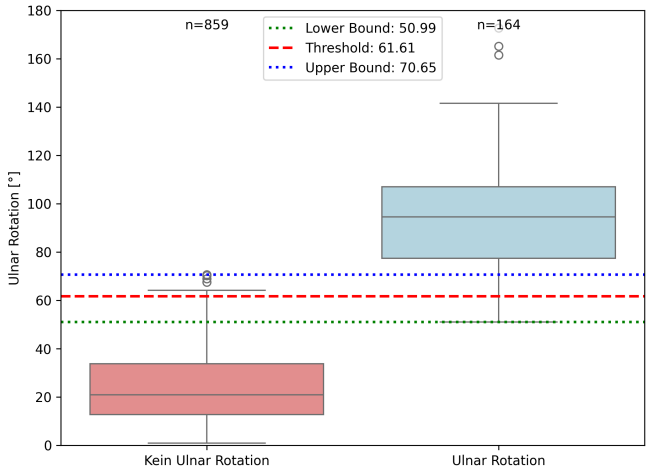



Figure 72: Distribution of hand rotation values (°) across the two classes with threshold and boundary values marked

$$R(q) = \begin{bmatrix} 1-2(y^2+z^2) & 2(xy+wz) & 2(xz-wy) \\ 2(xy-wz) & 1-2(x^2+z^2) & 2(yz+wx) \\ 2(xz+wy) & 2(yz-wx) & 1-2(x^2+y^2) \end{bmatrix} \tag{19}$$

The local rotational change effective between two successive frames is then determined from the difference of the two quaternions as $q_{rel,i} = q_i^{-1} \cdot q_{i+1}$. $q_{rel,i}$ is subsequently converted into a rotation vector and translated into the world coordinate system with the aid of the rotation matrix $R_i$ ($\vec{r}_i^{welt} = R_i \cdot \vec{r}_i^{lokal}$). To obtain the change in rotation relative to the initial state, the vector is then transformed into the reference coordinate system of the first frame ($i = 0$) via $\vec{r}_i^{ref} = R_0^T \cdot \vec{r}_i^{welt}$. From this transformed vector, the change in the angle around the longitudinal axis of the hand can finally be determined by considering the corresponding axis component; the greatest rotational change results from the difference between the maximum and minimum angle. The classifier thresholds are at 50.99° and 70.65° with a threshold value of 61.61°. This leads to a very good $F_1$ score of 97.30 %, i.e., within the test dataset the signs could be distinguished almost error-free with respect to the presence of a wrist rotation.

*Assessment of generalizability*

The threshold-based recognition approaches for FLEXION CHANGE, SPREAD CHANGE, and ULNAR ROTATION were developed on the basis of the validated and cleaned right-hand data. In order to assess the generalizability of the approaches, they were then applied to the left-hand data. These data sets had not previously been manually checked and therefore constitute unseen instances by which the transferability and robustness of the classifiers can be evaluated. It should be noted that, in the validated right-hand data, there was an error rate (erroneous sensor data, execution errors, and

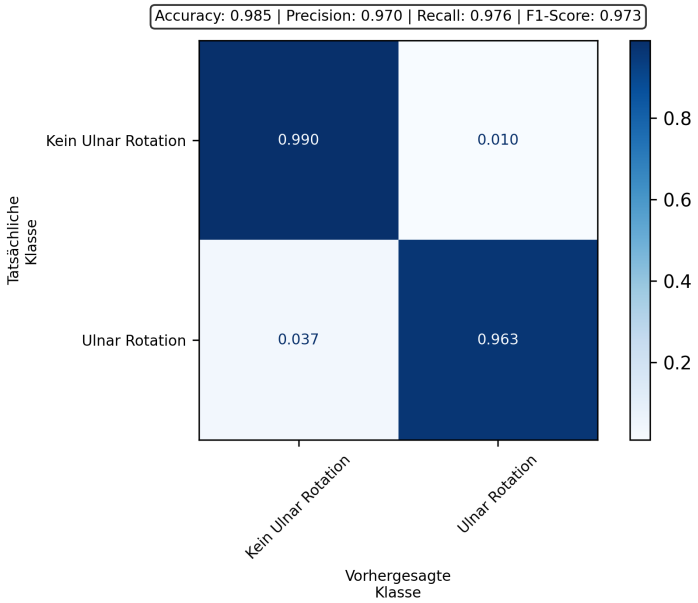



Figure 73: Confusion matrices of the threshold-based classifier for wrist rotation recognition

segmentation errors) of approximately 11 to 14.10 % for Flexion Change, 25 to 26.40 % for Spread Change[75], and 8.90 % for Ulnar Rotation. In addition, 18 of the 38 signs are one-handed, i.e. there is no Flexion Change, Spread Change, or Ulnar Rotation in these data, which is why the number of negative instances is significantly higher than for the right hand.

The results for Ulnar Rotation are shown as an example in Figure 74, while the results of the other classifiers are in Figure 75. Starting with the classifiers for Flexion Change in Figure 75a to Figure 75d, one still sees good results with an $F_1$ score between 89.80 to 91.80 %, i.e. about 7.60 to 7.70 percentage points lower than for the training data, depending on the finger. In particular, the correct recognition of absent changes (TN) was able to maintain its high recognition rate, whereas the correct recognition of present changes (TP) is 10.50 to 16.60 percentage points lower. For the two classifiers for Spread Change, the change compared with the training data is significantly larger than for Flexion Change, at 20.90 percentage points and 16.80 percentage points respectively. In particular, for the True Positive-Rate (TPR) of the spreading change between ring and little finger one can see a clear deterioration of 20.40 percentage points, whereas the True Negative-Rate (TNR) lost only 1.30 percentage points in accuracy. For the classifier for index and middle finger, this difference is less drastic, at 9.90 percentage points and 5.40 percentage points respectively. In the recognition of Ulnar Rotation, the $F_1$ score has dropped by 18.70 percentage points, with the TPR also dropping by 27.30 percentage points, compared with 3.80 percentage points

75 The error rate for spreading between the middle and ring fingers was not considered, as this classifier was not used further.

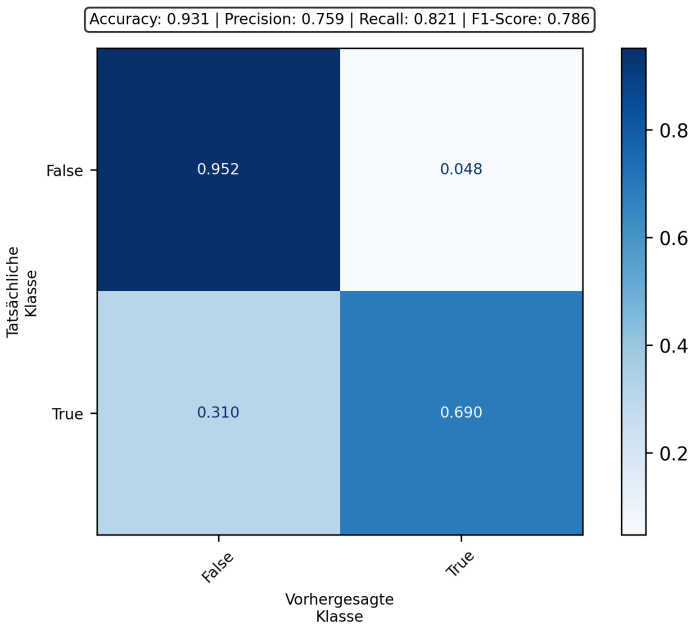



Figure 74: Confusion matrices of threshold-based wrist rotation recognition on unseen data from the left hand

for the TNR. The behavior is therefore similar to that observed in the recognition of SPREAD CHANGE.

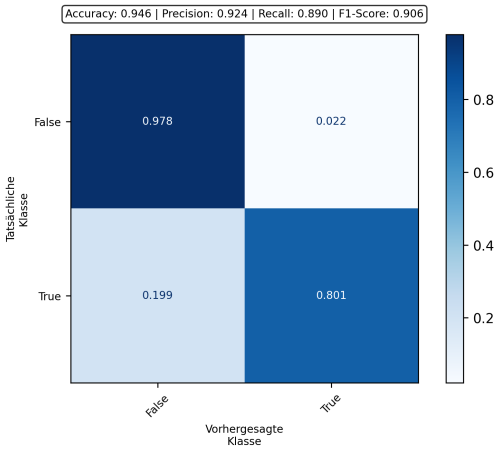
If the above-mentioned error rates are taken into account, which were particularly high for SPREAD CHANGE, and the fact that a large proportion of instances show no change and that the TNR is higher than the TPR, then these results are valid and comprehensible for FLEXION CHANGE and SPREAD CHANGE. Only for ULNAR ROTATION does the question arise whether the existing classifier exhibits *overfitting*, i.e. whether the approach is too strongly optimized for the training data set. This will be shown in the course of evaluating the meta-classifier in Section 6.3.

### 5.4.2 *Path movements of the hand*

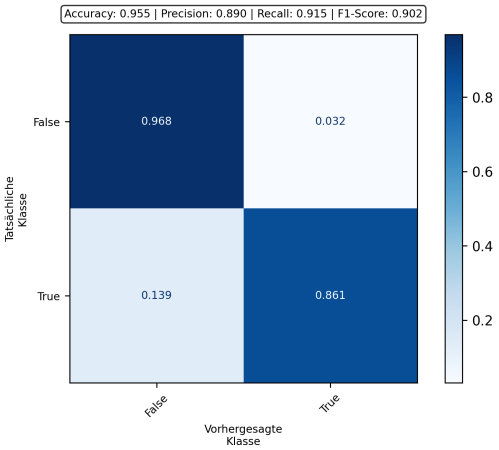
A large proportion of 84.69 % of the defined signs in *ASL-LEX* exhibit a path movement, i.e., the hand moves through three-dimensional space. PATH MOVEMENT denotes the movement path of the hand as the type of movement, irrespective of other properties such as its speed. It is composed of the movement of the forearm and the movement of the wrist (extension/flexion and radial abduction/ulnar abduction).

*ASL-LEX* divides PATH MOVEMENTS into the classes *Circular*, *Curved*, *Straight*, *X-Shaped*, *Z-shaped*, *None*, *BackAndForth* and *Other*, whereby the last two classes have already been excluded by us. The three classes *Circular*, *Curved*[76] and *Straight* originally stem from the approach by Brentari, while *X-Shaped* and *Z-shaped* were added by Sehyr

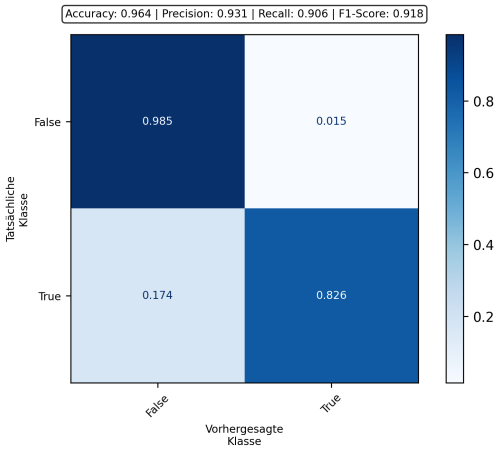
76 under the designation *Arc*



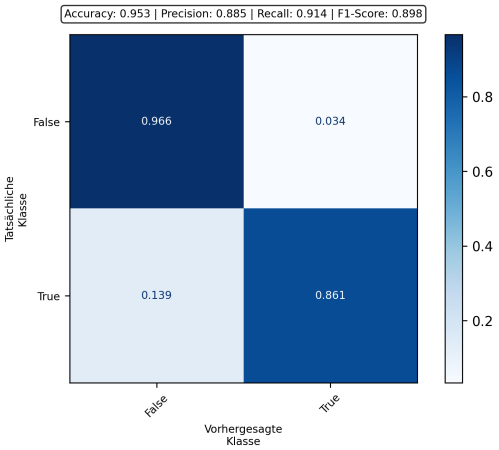

(a) Flexion Change index finger



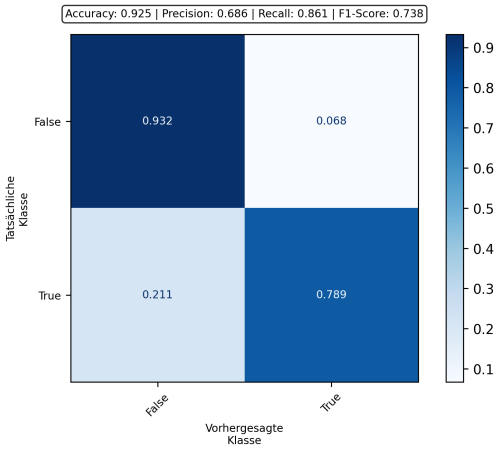

(b) Flexion Change middle finger



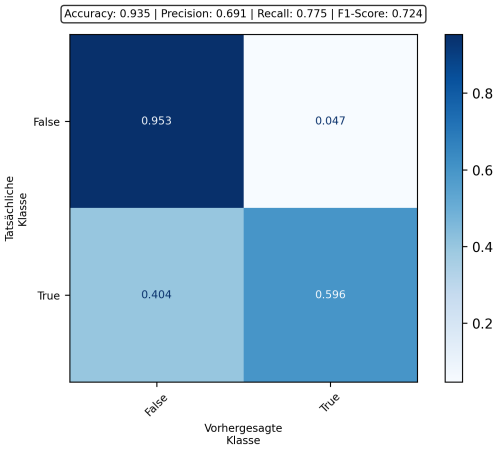

(c) Flexion Change ring finger




(d) Flexion Change little finger




(e) Spread Change index finger/middle finger




(f) Spread Change ring finger/little finger

Figure 75: Confusion matrices of the threshold-based classifiers on unseen data from the left hand

et al. and describe more complex path movements modeled on the letters "X" (e.g., SWITZERLAND) and "Z" (e.g., DOLPHIN).

In order to capture these path movements of the hand, it is advisable not to examine the movement of the hand, since through ULNAR ROTATION it can perform movements that are independent of PATH MOVEMENT (pronation/supination). Instead, the approach focuses on the movements of the forearm or the wrist. However, there are signs that are characterized by a path movement (e.g., YES or VITAMIN with a *Straight* movement or BASKETBALL with a *Curved* movement), in which the arm is at rest and the movement is performed entirely from the wrist. A converse example can be found in signs such as CANDY 1, where PATH MOVEMENT is *None*, but a ULNAR ROTATION is present, i.e., a movement takes place that can also be measured at the forearm.

In addition to this challenge, it should be noted that the main arm joints (shoulder and elbow joints) and the wrist are rotational joints, so that translational movements of the wrist are always the result of combined rotations of several joints. This can make it difficult to differentiate between a *Straight* and a *Curved* movement based on the forearm data.

Another complicating factor is the fact that both *X-Shaped* and *Z-shaped* movements can be regarded as composite *Straight* or *Curved* movements. We expect that, in particular for repeated movements, such a distinction will be difficult to detect. This also applies to the classes *Circular* and *Curved*.

There are also differences within the signs with respect to the distance covered, as for example in EYEGLASSES, where the movement path is very short, or, conversely, in CASTLE. Especially for short movements, it must be assumed that it will become increasingly difficult to distinguish between a *Curved* and a *Straight* movement.

Likewise, when examining the sign example videos of *ASL-LEX*, it becomes apparent that the signs also differ in terms of direction (cf. Figure 23) and, to a lesser extent, in speed (e.g., slow as in SNOW compared with fast as in SCARF). Even though this information is not part of the PATH MOVEMENT classes, it is nevertheless part of the movements to be captured and must be taken into account, as it enormously expands the variability of the individual classes. For a reliable recognition approach based on an ML classifier, this implies an enormous amount of training data to be captured.

In summary, the following challenges in the design of a path-movement recognizer for *ASL-LEX* classes can be outlined:

- Independent hand movements apart from arm movements are possible.
- Inconsistent annotation of the signs with respect to the reference point of the path movement.
- Translational movements arise from rotational joints, which can result in slight bends in the movement path.
- *X-Shaped* and *Z-shaped* movements are combinations of other classes.
- *Circular* and *Curved* are difficult to distinguish, especially when a repeated movement is present.

- Movement paths can be very short, making it difficult to distinguish between *Curved* and *Straight* movements.
- Classes have internal variability, e.g., different directions and speed.

*Recognition approach*

Due to the challenges just mentioned, no high recognition rates are expected. The path movements of the signs are too diverse to reduce them to a few classes, which are also annotated inconsistently. Nevertheless, we want to attempt to design an ML-based approach that differentiates the various path movements as well as possible. For the previous approach to handshape recognition, we used static features. Since we want to continue using our findings from this approach, we convert the dynamic time-series data of the path movements into static features for machine classification with the help of the Python library Time Series Feature Extraction Library (TSFEL)[77], which we have already used in a previous work and in which we were able to show that a variety of different features in the time and frequency domain can be used for human activity recognition [151].

TSFEL was developed in 2019 by Barandas et al. [21] and aims to centralize and optimize the development of feature extraction processes for time-series data. To this end, it provides a collection of feature extraction routines with which more than 65 features from the temporal, statistical, spectral, and fractal domains can automatically be extracted:

TEMPORAL FEATURES capture temporal changes and patterns in the data in order to identify trends, cycles, and correlations and to better understand dynamic processes and future developments.

STATISTICAL FEATURES describe the data using metrics such as mean, variance, skewness, and kurtosis, thus providing a summary representation of its distribution, dispersion, and shape.

SPECTRAL FEATURES analyze the frequency domain of the data using transformations such as Fourier or wavelet analyses in order to uncover hidden periodicities, harmonics, and oscillatory patterns.

FRACTAL FEATURES quantify the complexity and self-similarity of data across different scales, for example through the fractal dimension, and thus capture irregular, complex structures of natural systems. This domain was only subsequently[78] added on February 2, 2024.

Building on the findings of our previous work [151], in this paper we use features from the temporal and statistical domains of the forearm acceleration data as well as the Euler angles derived from the quaternions for all three axes, since these were

77 `https://tsfel.readthedocs.io/en/latest/` (Last accessed on January 25, 2026)
78 `https://github.com/fraunhoferportugal/tsfel/pull/144` (Last accessed on January 25, 2026)

able to achieve the highest recognition rates across all classifiers used. In addition, we observed that the generation of the spectral features took considerably longer than that of the statistical and temporal features. Fractal features were not available in the version used at the time and were only introduced later. In version 0.1.9 available to us now, 20 different statistical and 15 different temporal features are extracted from the time series.

*Training data*

For training the ML model, typical, characteristic movements of the dominant arm were selected for each of the six movement types (*Circular*, *Curved*, *Straight*, *X-Shaped*, *Z-Shaped*, or *None*) on the basis of a sample of sign videos from *ASL-LEX*. These were intended to represent the respective category well and to be clearly assignable in order to avoid overlaps between the classes. In addition, care was taken when selecting the signs to ensure that the movements were simple enough that even untrained persons could perform them correctly. Signs whose movement results only from a wrist rotation (e.g., GIFT) were annotated as *None*. Representative movements of all six movement types from 38 signs[79] of *ASL-LEX* were recorded and processed from ten participants under the conditions specified in Section 5.2. None of the participants had prior knowledge of a sign language. Before recording began, a T-pose calibration was performed. The movements were demonstrated by the recording supervisor and practiced. Afterwards, they were repeated three times and recorded separately for each arm, so that approximately 60 data recordings per sign were available after exclusion of erroneous recordings. 16 selected signs[80] were recorded once more with a repeated movement, i.e., for these signs there are approximately 120 recordings. The movement classes *X-Shaped* and *Z-Shaped* were not considered here, as they do not occur with a repeated movement in *ASL-LEX*. In total, we therefore arrive at 3,407 sign recordings with 57 to 120 recordings per sign. An example recording of the sign CAMEL with three repetitions is shown in Figure 76.

79 BANANA (60 recordings, *Z-Shaped*, list 1), BOWTIE (60 recordings, *None*, list 4), CAMEL (60 recordings, *Z-Shaped*, list 2), CANDY 1 (117 recordings, *None*, list 3), CATHOLIC (63 recordings, *X-Shaped*, list 1), CLOUD 1 (60 recordings, *Circular*, list 1), CROSS (59 recordings, *X-Shaped*, list 2), DINOSAUR 2 (60 recordings, *Z-Shaped*, list 4), DOLPHIN (57 recordings, *Z-Shaped*, list 3), DRESSER (60 recordings, *Z-Shaped*, list 1), DROWN (60 recordings, *Z-Shaped*, list 2), ECLIPSE (120 recordings, *Straight*, list 2), GIFT (60 recordings, *None*, list 2), HOSPITAL (60 recordings, *X-Shaped*, list 4), ISLAND (120 recordings, *Circular*, list 4), JELLYFISH (60 recordings, *Z-Shaped*, list 3), KNEEL (120 recordings, *Straight*, list 4), LAPTOP (119 recordings, *Curved*, list 4), LESBIAN (120 recordings, *Straight*, list 3), MONKEY (60 recordings, *None*, list 2), NEXT (120 recordings, *Curved*, list 3), OPEN BOOK (120 recordings, *Curved*, list 1), ORAL (120 recordings, *Circular*, list 3), PARANOID (60 recordings, *None*, list 4), PRISON (120 recordings, *Straight*, list 1), SCHEDULE (60 recordings, *X-Shaped*, list 3), SCREWDRIVER (117 recordings, *None*, list 1), SIREN (120 recordings, *Circular*, list 1), STARS (119 recordings, *Straight*, list 2), SUNSET (120 recordings, *Curved*, list 2), SWEATER (120 recordings, *Curved*, list 4), SWORD (60 recordings, *Z-Shaped*, list 3), TREE (60 recordings, *None*, list 4), VALUE (120 recordings, *Curved*, list 3), WANDER (60 recordings, *Z-Shaped*, list 1), WEDDING (120 recordings, *Curved*, list 1), WOLF (120 recordings, *Curved*, list 2), WORLD (120 recordings, *Circular*, list 2)

80 CANDY 1, ECLIPSE, ISLAND, KNEEL, LAPTOP, LESBIAN, NEXT, OPEN BOOK, ORAL, PRISON, SCREWDRIVER, SIREN, STARS, SUNSET, SWEATER, VALUE, WEDDING, WOLF, WORLD

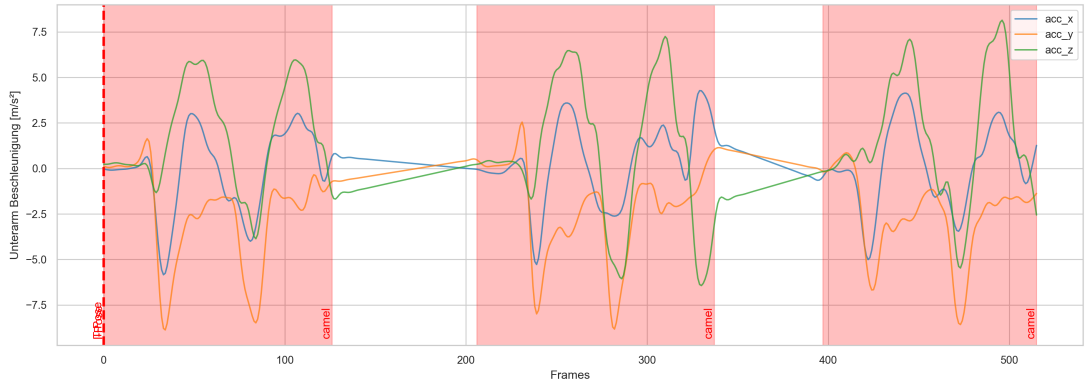
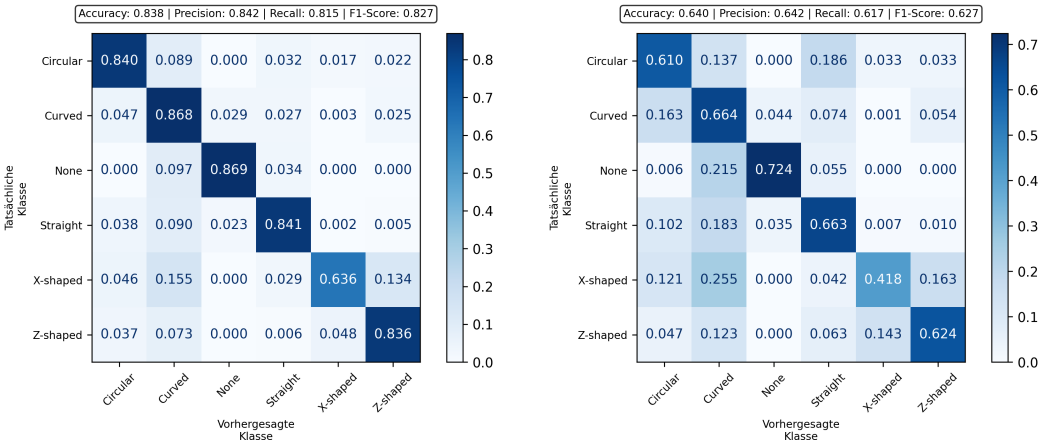



Figure 76: Arched *Z-Shaped* movement of the sign CAMEL with three repetitions

*Results*

Given the difficult conditions, we first want to examine how generalizable a possible ML model would be. It is to be expected that generalizability with respect to new participant data should be significantly higher than generalizability with respect to new signs. To investigate this hypothesis, we conduct two different grid search cross-validations in the LOSOCV setting with the ML classifiers and parameter ranges already used in Table A.9, column "Gridsearch Bereich". For both validations, we use the same data and partition them into four approximately equal subsets, once separated by participants – i.e., one participant always appears in only one subset – and once by signs. The results of this experiment for an VL2 classifier based on the optimized classifiers are shown in Figure 77 and are intended to demonstrate how well the model responds to new participants on the one hand and how well it responds to new signs on the other. As expected, the $F_1$ score of 82.70 % of the participant-independent investigation is significantly higher (cf. Figure 77a) than the $F_1$ score of 62.70 % of the sign-independent investigation (cf. Figure 77b).

One of the main goals of our parametric approach is not to have to include every sign to be captured during training, but instead only a representative selection of signs per class. This presupposes that the approaches used exhibit good generalizability with respect to new, unseen signs. In handshape recognition, this problem was secondary due to the high number of classes and the associated level of detail, since 15 features were used to differentiate 54 classes. In path movement recognition, by contrast, six classes are distinguished on the basis of 270 features – i.e., many different movements are grouped within one class – and correspondingly more training data are required. It is therefore sensible to proceed on the basis of the sign-independent approach by further using the parameter values determined via the grid search. Moreover, it is advisable to merge classes, since the classes *Circular*, *X-Shaped* and *Z-Shaped* can be regarded as sequences of *Curved* and *Straight*, and there may be confusions with repeated movements. We therefore define *Circular* and *Z-Shaped* as *Curved* and *X-Shaped* as *Straight*, and now obtain the results shown in Figure 78 with an $F_1$ score of 88.90 %, i.e., an increase of 26.20 percentage points compared with the six-class model.




(a) Participant-separated investigation (b) Sign-separated investigation

Figure 77: Results of the grid-search cross-validations in the LOSOCV setting

It is evident that the greatest confusion is that linear movements are recognized as curved movements. In future, one could therefore annotate movements with a low degree of curvature and short movements as *Straight* in order to reduce confusions, and continue to annotate only clearly visible bent movements as *Curved*. Furthermore, the class *Circular* could be reintroduced and curved, repeated movements could be annotated as *Circular*, whereas curved, non-repeated movements would continue to be annotated as *Curved*. In addition, more training data should be collected in order to obtain greater variance in the training data. A uniform solution must also be found as to whether hand movements are part of the path movements or not. The signs must accordingly be re-annotated and the hand movement taken into account in the classifier. A movement of the hand could then be classified as *Curved*, even if the arm remains stationary.

### 5.4.3 *Repeated movements*

The parameter *Repeated Movement* indicates whether a repeating movement is present and refers to all types of movements. This poses us with particular challenges, since it is not explicitly defined which type of movement is considered to be repeating. As an example, consider the sign BUBBLES, in which a repeating movement is present for the flexion of the fingers, but not for the hand path movement or for the change in finger abduction/spread, which are also present. The sign OCTOPUS, by contrast, likewise includes flexion and changes in finger abduction/spread, as well as a hand path movement; however, here the repeating movement refers to all of these parameters. It is similarly unclear whether a repeating movement constitutes a simple repetition (e.g., sign MARCH) or multiple repetitions (e.g., sign JUGGLE).

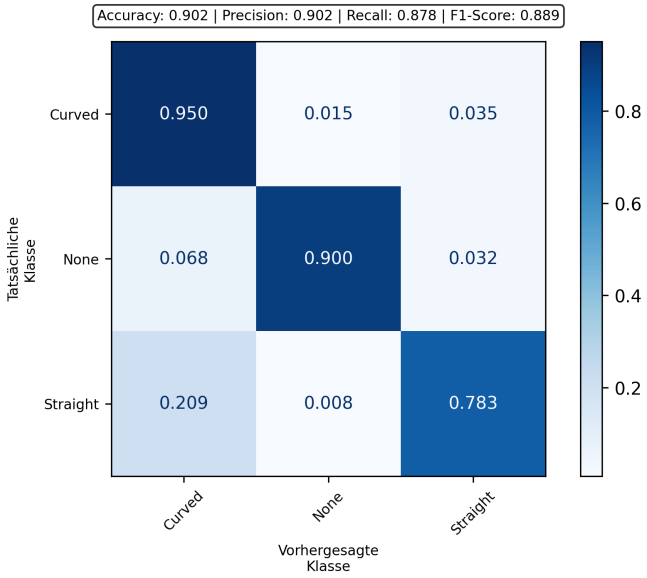



Figure 78: Results of path movement recognition with three classes

Especially with regard to the ML classifier used for the path movements, this may lead to difficulties in defining the classes. Thus, a *Curved* movement executed repeatedly or an imprecisely executed repeating *Straight* movement could be recognized as a *Circular* movement. Extending the class space is also challenging, since it is not clearly defined how many times the individual classes can be repeated.

For these reasons, we have decided not to further consider the parameter *Repeated Movement* within the scope of *ParaSignRec*.

## 5.5 RECOGNITION OF HAND ORIENTATION

Another important parameter for describing signs is *palm orientation*. In contrast to the parameters *handshape*, *sign location*, and *hand movement*, palm orientation is less frequently used explicitly by notation systems and is often integrated into the parameter of *handshape*. In our study, this was the case in nine of the 14 notation systems examined (cf. Section 3.1, Table 2, excluding *Glossing*). In another comparison, eleven notation systems that explicitly or implicitly mention the orientation of the hand (or parts of it such as the palm or extended fingers) contained between 5 and 144 classes [17]. This shows that the parameter has a certain relevance for the contrastive description of signs and can also be designed with varying degrees of detail.

Palm orientation is not yet considered by the *ASL-LEX*, although it can be annotated and captured in a straightforward way due to the easy interpretability of the individual classes and the use of IMU-based data gloves[81]. In order to reduce the existing

81 Regardless of the data glove model used, an IMU was always attached to the back of the hand, via which the current orientation of the hand can be determined.

duplicates in our *ParaSignRec*, it makes sense to complement this with additional information regarding movement or palm orientation compared with *ASL-LEX*, see Section 3.1.3. Since palm orientation at the end of sign production results from the palm orientation at the beginning and the movement performed thereafter, this contains information both about palm orientation and about movement direction and symmetry. In addition, the selected time points (beginning and end of the sign) are clearly defined. Our goal is therefore to complement the palm orientations of both hands at the beginning and end of signs as additional parameters. This would also provide additional information for identifying the sign sought, with the help of which less significant parameters can be compensated for.

### 5.5.1 *Selection of the classes of palm orientations*

For the design of a suitable classifier for recognizing palm orientation, a selection of classes representing the palm orientations must first be determined. In theory, the hand can adopt any orientation depending on the arms, although some positions feel uncomfortable and others more natural. Therefore, some palm orientations are hardly used in sign language, whereas others are very common [17]. The finer the distinctions to be made between palm orientations, the more classes are required and the more precisely signs can be distinguished on the basis of palm orientation, thereby reducing duplicates. However, as the number of classes increases, so does the risk of signs being performed incorrectly, since it will be difficult, especially for laypersons, to set the palm orientation to within a few degrees while simultaneously taking into account the other parameters such as handshape and movements. Likewise, our recognition approach will become more error-prone, since a larger number of classes generally leads to lower recognition accuracy and the amount of annotated data per class available for training is also reduced, because a fixed number of signs is distributed across more different classes.

To represent the orientation of a hand in three-dimensional space, at least two pieces of information are required, for example the direction in which the fingers point and the direction in which the palm faces. It is assumed here that the fingers are extended and together with the palm and the thumb extended to the side form an orthogonal coordinate system, as shown in Figure 79. This means that the fingers, the palm, and the thumb are at right angles to one another. The coordinate system for the right hand is right-handed, and the coordinate system for the left hand is left-handed.

As mentioned at the outset, the notation systems examined in this work differ in some cases greatly in the number of classes for palm orientation. Stokoe, for example, specifies only six finger directions, with one positive and one negative direction defined for each three-dimensional axis. The six symbols for these directions can be supplemented with one of two additional symbols to indicate the palm orientation. HamNoSys extends this further with 12 classes for the diagonal directions of the fingers as well as eight classes for the direction of the palm (*Inside*/*Outside*, *Up*/*Down*, and the diagonals in between). *SignScript* specifies only the direction of the palm and omits the class *Outside*. *Symbol Font for ASL* gives six directions for both fingers and

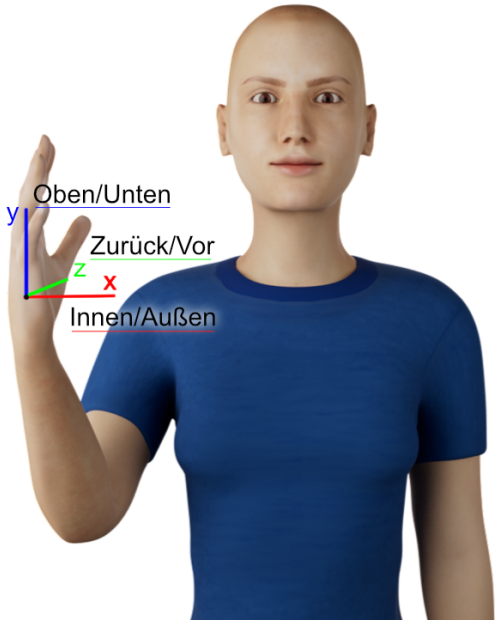



Figure 79: Coordinate system of palm orientation from the signer's perspective

palm and thus arrives at 24 different classes, since fingers and palm are orthogonal to each other and thus twelve of the 36 possible combinations are ruled out. Some systems choose to include only one aspect of hand orientation and infer the missing information from context. Other notation systems are much more detailed in their representation of palm orientation or focus on wrist flexion [17].

We orient ourselves toward the approach of *Symbol Font for ASL* in order to describe the orientation of the hand along each main axis in positive and negative directions with 24 classes in steps of 90°. Using the coordinate system shown in Figure 79 as an example, this would yield a designation of the finger direction as *Inside*/*Outside*, *Up*/*Down*, and *Back*/*Front*, depending on which axis segment is closest to the extended finger. In the next step, among the remaining four axis segments, we determine the direction in which the palm faces. The combination of these two pieces of information yields the desired class of palm orientation, e.g. *UpInside* for a vertically held hand with fingers pointing upward and palm to the side, as shown in the coordinate-system example. Diagonal directions are not taken into account in order to keep the number of classes manageable.

### 5.5.2 *Implementation of the recognition approach*

For our recognition approach, we model palm orientation as a combination of two orthogonal directions: *i)* finger direction of the extended hand and *ii)* palm direction.

The six canonical spatial directions are $\{\texttt{Out}, \texttt{In}, \texttt{Front}, \texttt{Back}, \texttt{Up}, \texttt{Down}\}$ with unit vectors $\{\pm\vec{e}_x, \pm\vec{e}_y, \pm\vec{e}_z\}$. For the right hand, $\texttt{Out} = +\vec{e}_x$, $\texttt{In} = -\vec{e}_x$; for the left hand these are swapped. Combined, this yields 24 different classes such as *UpFront* (fingers

point upward, palm faces forward) or *OutDown* (fingers point to the side, away from the body, palm faces downward).

To assign the hand orientation classes, at each time point of the data sequence we consider the normalized quaternion of the back of the hand, which is provided via the *Manus Core* SDK and calibrated within the scope of our system (cf. Section 5.2.3). It now describes the relative orientation of the hand to the viewer, is converted into a rotation matrix, and applied with $R(q) \cdot \vec{e}$ to the two local hand axes $\vec{e}_{\text{Finger}} = [1, 0, 0]^\top$ and $\vec{e}_{\text{Palm}} = [0, 0, -1]^\top$.

The cosine is now determined between each of the two local hand axes $\vec{e}_{\text{Hand}} \in \{\vec{e}_{\text{Finger}}, \vec{e}_{\text{Palm}}\}$ and each of the six canonical spatial directions $\vec{e}_{\text{Space}} \in \{\pm\vec{e}_x, \pm\vec{e}_y, \pm\vec{e}_z\}$. This is done using the cosine function

$$\cos(\theta) = \frac{\vec{e}_{\text{Hand}}^{\top} \cdot \vec{e}_{\text{Space}}}{|\vec{e}_{\text{Hand}}| \cdot |\vec{e}_{\text{Space}}|} \in [-1, 1] \tag{20}$$

which simplifies, due to the normalized unit vectors $|\vec{e}_{\text{Hand}}| = |\vec{e}_{\text{Space}}| = 1$, to $\cos(\theta) = \vec{e}_{\text{Hand}}^{\top} \cdot \vec{e}_{\text{Space}}$. By normalizing the function to the range $[0, 1]$ and adding an exponent $\gamma > 0$ for a sharper distribution of the scores, we obtain the following formula:

$$f_{\text{score}}(\vec{e}_{\text{Hand}}, \vec{e}_{\text{Space}}, \gamma) = \left( \frac{\vec{e}_{\text{Hand}}^{\top} \vec{e}_{\text{Space}} + 1}{2} \right)^{\gamma} \in [0, 1] \tag{21}$$

With this formula, we obtain a score between 0 and 1 for each of the twelve combinations of hand axis and spatial direction.

In the final step, the combined class score is calculated for each of the 24 valid classes by forming the product of the two individual results and normalizing over the sum of all class scores.

Empirical tests with the data gloves showed that this approach delivers error-free results. This is due in particular to the fact that the complexity of the approach is manageable and the measurements of the *Manus Prime X*[82] data gloves exhibit only small error rates of about $\pm 2.50°$, whereas the approach chosen by us predicts an incorrect class only from a deviation of 45°. Larger deviations arise mainly from inaccurate calibrations, in which the reference poses are not performed exactly and the calibrated orientation of the sensors therefore has an offset.

Due to these very good results and the fact that correct annotation of signs with respect to orientations is still lacking at this point, we refrain here from a detailed evaluation and will provide this in the context of meta-classification. Since the exponent $\gamma$ significantly influences both the sharpness of the distribution and the results of meta-classification, it is examined in more detail below.

82 The error rates of the *Manus Quantum Metaglove* data gloves could unfortunately not be determined.

### 5.5.3 *Selection and influence of the sharpening parameter $\gamma$*

In general, for a classification, a score greater than $1/n$ is required for the correct prediction of a class, where $n$ denotes the number of possible classes. By contrast, within meta-classification, the absolute magnitude of the score is decisive. Scores that lie too close together cause the discriminative power of the classifier within meta-classification to decrease and its influence on the overall result to be weakened accordingly. Therefore, suitable sharpening of the scores is necessary in order to regulate the distribution of probabilities and correctly represent neighborhoods.

Without additional sharpening of the scores ($\gamma = 1$), the score of a perfectly executed orientation would amount to only 16.70 %, compared with 8.30 % for the neighboring classes. In the course of meta-classification, the influence of a correct orientation on the overall result would therefore be small, although the reliability of the classifier is to be regarded as high.

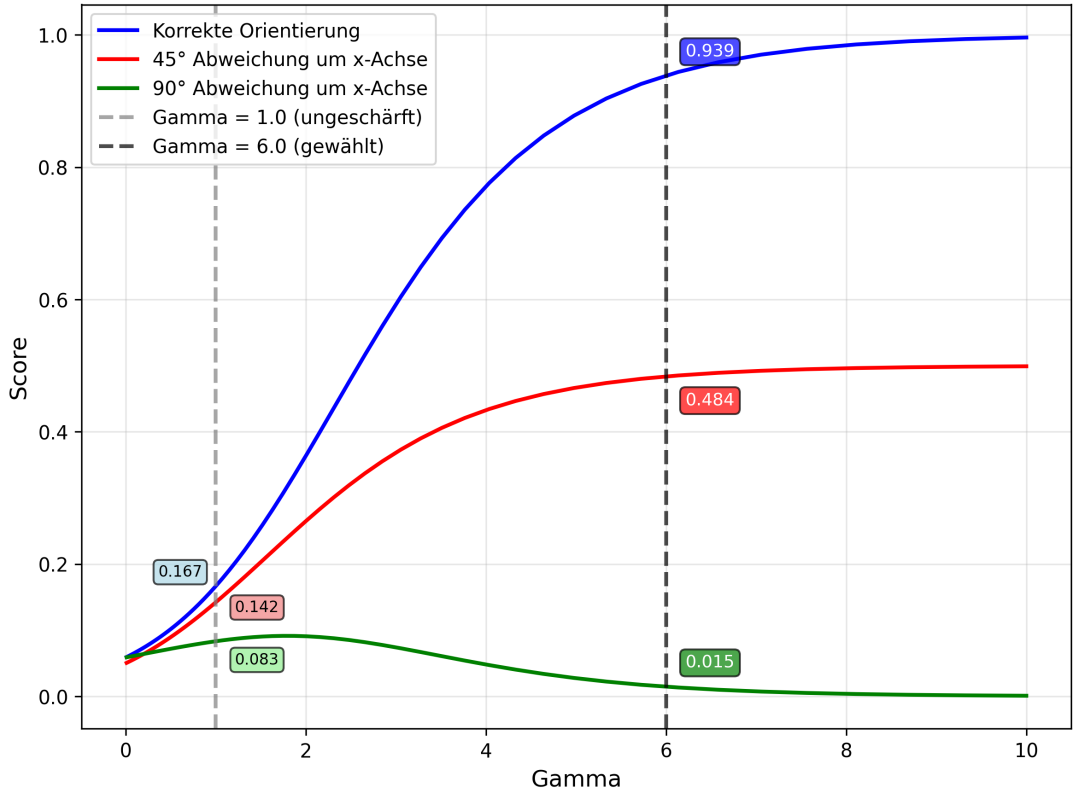



Figure 80: Predicted score as a function of $\gamma$ for the correctly executed orientation as well as two erroneous executions

For the analysis, the influence of $\gamma$ on three different executions of the same class was examined: *i)* the correctly executed orientation, *ii)* an execution deviating by 45° around the x-axis (exactly between two classes), *iii)* and an execution corresponding to a neighboring class (90° deviation around the x-axis). Figure 80 shows the change in the scores of these three executions as a function of the sharpening parameter $\gamma$. The

aim is to determine a value for $\gamma$ at which the correct class receives a clearly higher score, while neighboring classes reach nearly zero, but are not completely excluded. Ideally, an orientation lying exactly between two classes should be assigned a score of about 0.50. Our choice therefore fell on $\gamma = 6$, which best satisfies these conditions.

# 6 EVALUATION OF THE OVERALL SYSTEM

Goal of this chapter is to evaluate the parametric approach to sign language recognition that we have advocated. This approach should be able to reliably and in real time to recognize a vocabulary of signs that is sufficient for simple conversations. The requirements regarding real-time capability and the precise scope of this vocabulary have already been stated in Section 1.2 and specified in Section 3.2.1 and Section 3.1.2, respectively. For a systematic analysis of the overall system, the evaluation is carried out in several steps:

First, it is necessary to define the final vocabulary based on the recognition approaches for the individual parameters that were designed and implemented in Chapters 4 and 5. In this context, a selection of signs to be included should also be made for the actual evaluation process.

The individual recognition approaches have already been evaluated using data recorded specifically for this purpose, and their suitability is now to be tested under realistic deployment conditions with real sign-language data. For this, the framework conditions of the test data acquisition are described. The recognition approaches are then considered in isolation and, based on the test data recordings, evaluated with respect to their respective recognition performance.

Building on these individual results, the weighting of these components is then designed in order to identify the target signs with the highest possible accuracy within the context of meta-classification. Finally, the evaluation of the meta-classification is carried out, which integrates the previously investigated components and reflects the performance of the overall system.

## 6.1 FINALIZATION AND EVALUATION OF THE SIGN DICTIONARY

After, in Section 3.1, we have already selected a data basis for the lexicon of the *ParaSignRec* using the *ASL-LEX*, this chapter addresses the specific design of the lexicon. This is defined primarily by which parameters, together with which classes, are used through the implementation of the different recognition approaches from the preceding chapters. Due to different reasons, on the one hand some classes of existing parameters have already been consolidated, or entire parameters such as Major Location, Contact, or the parameters of the handshape configuration have been excluded. This leads to a reduction in sign diversity and, consequently, to a smaller lexicon of class combinations. On the other hand, new parameters such as Hand Orientation have been added, along with numerous pieces of information regarding the non-dominant hand. This in turn enables a higher number of class combinations and, associated with this, a higher number of signs.

### 6.1.1 *Curation of the Data Basis*

In order to meet the requirements of our system, it is also necessary to curate the individual signs. For instance, for the further use of the *ASL-LEX* within our *ParaSignRec*, the sign inventory needs to be reduced by those signs that contain more than one morpheme, which simplifies our approach because we do not require elaborate procedures to segment individual morphemes. Overall, this excludes 262 signs, which corresponds to 9.62 % of the original sign inventory.

Furthermore, all signs are excluded that exhibit *Other* values in at least one phonological feature, since these are not uniquely defined and would lead to problems in the implementation of the recognition approaches. This concerns the parameters Major Location, Minor Location, Second Minor Location and Path Movement and excludes 133 signs (4.88 %).

Since we make use of the rules by Battison (see Section 2.1.2) in order to achieve higher recognition accuracy from the symmetry properties of the signs, we additionally exclude all signs that violate these rules. This concerns all signs with the Sign Type *DominanceViolation* and *SymmetryViolation* (79 signs, 2.90 %) as well as all signs with Non-dominant Handshape = *DominanceConditionViolation* (70 signs, 2.57 %). Moreover, by definition, signs with the SignType *SymmetricalOrAlternating* or *AsymmetricalSameHandshape* may not exhibit different handshapes of both hands, which leads to the exclusion of an additional 20 signs (0.73 %).

In addition, 147 further signs (5.40 %) were excluded due to additional inconsistencies. One example is the use of old class names from the first version of the *ASL-LEX*, which cannot be transferred unambiguously to the new version. Since, for the conception and implementation of individual classifiers, we rely on the sign videos provided by *ASL-LEX*, five signs were excluded for which no sign videos were available. We also need these videos for the sign recording within the scope of the evaluation in order to validate the signs that were performed.

All measures and the number of signs affected thereby are shown in Table 22. Some signs were excluded due to multiple measures listed here. After carrying out these measures, 2,171 (79.73 %) remain in the sign inventory of the *ParaSignRec*.

In addition to the measures mentioned above for simplifying the data basis, some signs were also corrected manually, for example in cases of incorrect or ambiguous sign IDs, or incorrectly annotated data. One example is the original handling of *ASL-LEX* signs that are produced on the non-dominant arm (Major Location = *Arm*): In order not to be forced to specify a single handshape for the non-dominant hand in this case, the sign type was defined as *one-handed* and annotated with an "relaxed handshape" (*Lax*) ("This rule alleviated pressure to code the non-dominant handshape in these signs, which was often in a lax position" [187]). However, this is not correct; these signs belong to the *two-handed*, asymmetrical signs (types 2 and 3). Since the *Lax* handshape is not defined in greater detail, each of the 65 affected signs must be annotated manually by specifying a non-dominant handshape for each sign. This determines whether the sign is of type 2 (same handshape as the dominant hand)

Table 22: Simplifications performed on the *ASL-LEX* dataset. Signs can be excluded multiple times for different reasons. Only the first morpheme of a sign is considered.

| **Reason for exclusion** | **Signs** |
|---|---|
| **Signs of the *ASL-LEX*** | **2.723** |
| *Simplifications* | |
| More than 1 morpheme | −262 |
| Without sign video | −5 |
| *Insufficient definition* | |
| *Other* values in MAJOR LOCATION | −1 |
| *Other* values in MINOR LOCATION | −67 |
| *Other* values in SECOND MINOR LOCATION | −20 |
| *Other* values in MOVEMENT | −45 |
| *Violation of Battison's rules* | |
| SIGN TYPE = *DominanceViolation* | −63 |
| SIGN TYPE = *SymmetryViolation* | −16 |
| NON-DOMINANT HANDSHAPE = *DominanceConditionViolation* | −70 |
| SIGN TYPE = *SymmetricalOrAlternating* / *AsymmetricalSameHandshape* and dominant ≠ non-dominant handshape | −20 |
| *Inconsistent annotation* | |
| FLEXION CHANGE = *NA* when FLEXION ≠ *NA* | −39 |
| SPREAD CHANGE = *NA* and SPREAD ≠ *NA* | −16 |
| REPEATED MOVEMENT = *True* without any movement occurring (FLEXION CHANGE = *False*/*NA*, SPREAD CHANGE = *False*/*NA*, PATH MOVEMEMENT = *None* and ULNAR ROTATION = *False*) | −20 |
| PATH MOVEMEMENT = *BackAndForth* (original class from *ASL-LEX* v1, which was replaced by other classes in v2) | −72 |
| **Remaining signs for the *ParaSignRec*** | **2,171** |

or type 3 (different handshape as the dominant hand – one of the seven[83] *Unmarked Handshapes*). Alternatively, one could allow multiple handshapes at the same time – the seven *Unmarked Handshapes* and the dominant hand's handshape – in order not to have to commit to a single handshape and sign type.

83 OPEN A is not considered in the *ASL-LEX* vocabulary, which is why here the variant with seven *Unmarked Handshapes* (see Section 2.1.2) is used.

### 6.1.2 *Adapting New Parameters and Classes*

Due to the design and characteristics of the individual recognition approaches in Chapter 4 and Section 5.3 to 5.5, the original phonological properties of the *ASL-LEX*, on which our *ParaSignRec* is based, have in part been fundamentally changed. From the originally 16 phonological properties of the *ASL-LEX*, of which 15 were related to the dominant hand and one to the non-dominant hand, 28 parameters have been developed, with 14 parameters per hand. Particularly noteworthy in this context is the exclusion of all four phonological properties for the handshape configuration, as well as the extension by four parameters for palm orientation, which did not exist previously. Likewise, all parameters now contain information on the dominant and non-dominant hands, and all static parameters contain information at the beginning and end of sign articulation, which had previously been provided exclusively for the place of articulation by *Minor Location* and *Second Minor Location*. The number of classes has also changed fundamentally, in particular due to the exclusion of some vaguely defined classes (e.g., *NA* or *Other*), or due to the fusion of the original SELECTED FINGERS and FLEXION/SPREAD CHANGE into six new parameters. However, some classes had to be consolidated strictly because of technical limitations of the recognition approaches, such as for movements (from originally eight classes to three) and for places of articulation (from originally 39 classes to six).

All the aforementioned changes can be seen in Table 23. If one forms the Cartesian product of all classes of the *ASL-LEX*, one obtains a theoretical number of $2.79 \times 10^{13}$ combination possibilities for representing a sign. The Cartesian product of all classes of the *ParaSignRec* yields, with $5.39 \times 10^{20}$, a multiple of that; it therefore contains substantially more information for representing a sign than the *ASL-LEX*.

Most of the information newly added in the *ParaSignRec* was derived using Battison's rules on symmetry properties. Therefore, it is also evident that—even while adhering to Battison's rules—not every conceivable parameter combination results in a valid sign. The handshapes of the non-dominant hand, for instance, comprise essentially the same selection as those of the dominant hand. Computationally, therefore, there are with the classes of the ParaSignRec $54 \cdot 54 = 2.916$ combination possibilities for the two hands. However, these can be substantially reduced by reference to Battison's *Rule of Symmetry* and *Rule of Dominance* (see Section 2.1.2). According to these rules, the sign types mentioned above can be divided into three categories, which yield different numbers of possible combinations:

ONE-HANDED SIGNS: *Type 0* is restricted to the use of the dominant hand. The combination possibilities reduce to 54 handshapes of the dominant hand.

SIGNS WITH IDENTICAL HANDSHAPE: For *Type 1* and *Type 2*, the handshapes of the dominant and non-dominant hand are identical. The combination possibilities also reduce to 54 pairs of handshapes.

SIGNS WITH DIFFERENT HANDSHAPES: For *Type 3*, both handshapes differ; however, according to Battison's *Rule of Dominance*, the non-dominant hand can assume

Table 23: Overview and comparison of all parameters and classes of *ASL-LEX* and of *ParaSignRec* with the number of their combinations as a Cartesian product.

| **Category** | **Phonological property / parameter** | ***ASL-LEX*** | | ***ParaSignRec*** | |
|---|---|---|---|---|---|
| | | d | nd | d | nd |
| Handshape | HANDSHAPE (beginning of the execution) | 58 | 53 | 54 | 54 |
| | HANDSHAPE (end of the execution) | – | – | 54 | 54 |
| | THUMB CONTACT | 3 | – | – | – |
| Handshape-configuration | SELECTED FINGERS | 12 | – | – | – |
| | FLEXION | 8 | – | – | – |
| | SPREAD | 3 | – | – | – |
| | THUMB POSITION | 2 | – | – | – |
| Movements | PATH MOVEMENT | 8 | – | 3 | 3 |
| | ULNAR ROTATION | 2 | – | 2 | 2 |
| | FLEXION CHANGE (four fingers) | 3 | – | $2^4$ | $2^4$ |
| | SPREAD CHANGE (two finger pairs) | 3 | – | $2^2$ | $2^2$ |
| | REPEATED MOVEMENT | 2 | – | – | – |
| Place of execution | MAJOR LOCATION | 6 | – | – | – |
| | MINOR LOCATION | 39 | – | 6 | 6 |
| | SECOND MINOR LOCATION | 39 | – | 6 | 6 |
| | CONTACT | 2 | – | – | – |
| Orientation | ORIENTATION (beginning of the execution) | – | – | 24 | 24 |
| | ORIENTATION (end of the execution) | – | – | 24 | 24 |
| **Number of combination possibilities** | | $2.79 \times 10^{13}$ | | $5.39 \times 10^{20}$ | |

only one of seven basic handshapes. The combination possibilities therefore reduce to $54 \cdot 7 = 378$ handshapes.

Figure 81 provides an overview of the various combination possibilities. It can be seen that the combinatorial diversity of Type 1 and 2, compared to Type 0, remains unchanged. The only exception is Type 3. Instead of 2,916 combination possibilities, there are therefore only $54 + 54 + 378 - 7 = 479$ (16.43 %) combination possibilities when one takes into account that the seven *Unmarked Handshapes* are represented in the repertoire of the 54 handshapes and that "no handshape" constitutes no separate class for one-handed signs.

The same is true for the other parameters. Table 24 shows which parameters provide information about the non-dominant hand that we can derive from the *ASL-LEX* based on the sign types. This concerns exclusively two-handed signs and, essentially, is limited to signs of Type 1, since by definition symmetric execution is present here. In this case, the non-dominant hand is symmetric (with movements executed simulta-

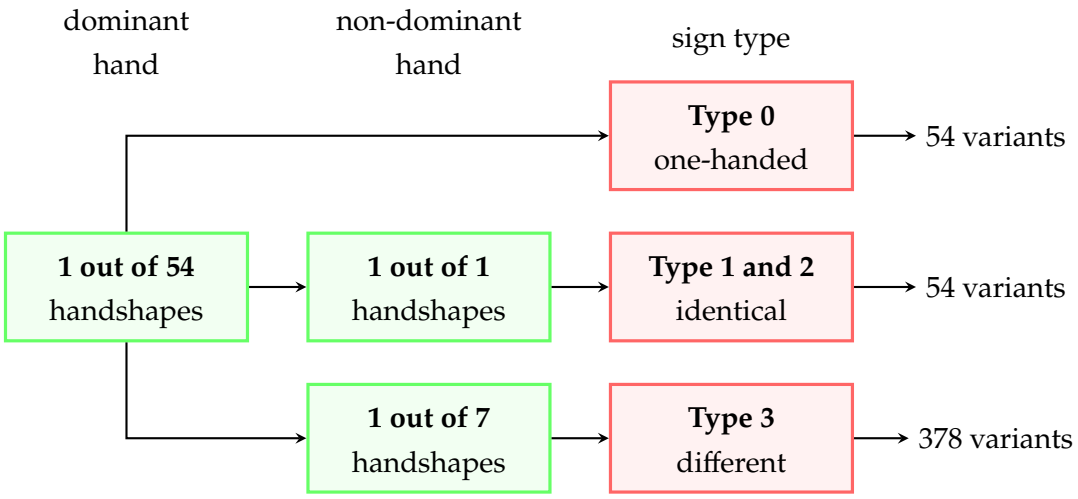



Figure 81: Combination variants of the dominant and non-dominant handshape depending on the sign type

neously or alternately) with respect to the dominant hand. For palm orientation, this means that the non-dominant hand must be oriented identically or mirrored relative to the dominant hand. From a dominant palm orientation follow four possible non-dominant palm orientations; e.g., from *FrontDown* one obtains *FrontDown*, *FrontUp*, *BackDown*, or *BackUp*. Since all remaining parameters are direction-independent and therefore only specify, for example, movement paths but not their directions, it makes no difference for these recognition approaches whether the movements of the two hands are carried out simultaneously or alternately. For the handshape of the non-dominant hand, explicit information is available with NON-DOMINANT HANDSHAPE that we can also take into account for Type 2 and 3. In addition, we implicitly have the information that the non-dominant hand rests in an asymmetric sign, i.e., no movements occur.

Since we thus have actual independence between the information about the dominant hand and the non-dominant hand only for the place of articulation and orientation of Type 2 and Type 3, most of the information derived from symmetry properties helps only to a limited extent, in that it increases the number of possible combinations and thereby the vocabulary. However, this additional information can be used to increase the accuracy of the system. Instead of classifying only the dominant hand, the non-dominant hand can also be classified to obtain two rather than one result for each parameter in the evaluation. Thus, in cases of ambiguous results, the second result may under certain circumstances provide the decision.

Table 25 shows the updated number of combination possibilities of the *ParaSignRec* while adhering to Battison's rules. Although the theoretical number of combination possibilities in the *ParaSignRec* is $5.39 \times 10^{20}$, the actual vocabulary of valid signs is therefore with $2.88 \times 10^{13}$ far below that value, but still slightly above the theoretical combination possibilities of the *ASL-LEX* with $2.79 \times 10^{13}$ signs. The number of valid combination possibilities of the *ASL-LEX* was not determined in detail by us, because

Table 24: Information content of the *ASL-LEX* regarding the non-dominant hand for various sign types according to Battison

| **Category** | **Phonological property / Parameter** | **Battison Sign Type** 0 | 1 | 2 | 3 |
|---|---|---|---|---|---|
| Handshape | HANDSHAPE (beginning) | – | = / ✓ | = / ✓ | ✓ |
| | HANDSHAPE (end) | – | = | = | ~ |
| Movement | PATH MOVEMENT | – | = | ~ | ~ |
| | ULNAR ROTATION | – | = | ~ | ~ |
| | FLEXION CHANGE (four fingers) | – | = | ~ | ~ |
| | SPREAD CHANGE (two fingers) | – | = | ~ | ~ |
| Place of articulation | MINOR LOCATION | – | = | ✗ | ✗ |
| | SECOND MINOR LOCATION | – | = | ✗ | ✗ |
| Orientation | ORIENTATION (beginning) | – | ~ | ✗ | ✗ |
| | ORIENTATION (end) | – | ~ | ✗ | ✗ |

–: Not relevant, =: Symmetric to the dominant hand,
✗: Not defined, ~: Implicitly defined, ✓: Explicitly defined

here, additionally, many dependencies exist within the individual parameters, such as the congruent representation possibilities of the hand (handshape vs. handshape configuration), the dependency between MAJOR LOCATION, MINOR LOCATION, and SECOND MINOR LOCATION, or between the SELECTED FINGERS and FLEXION/SPREAD CHANGE, and not more specifically defined classes such as *NA* in FLEXION/SPREAD CHANGE. It is nevertheless likely that it exhibits approximately the same ratio between theoretical and actual class combinations as in the *ParaSignRec*.

Finally, Figure 82 shows the class distribution of sign types according to Battison for the *ASL-LEX* (only monomorphemic signs) in green compared to the *ParaSignRec* in red. The category *Other* includes signs that violate Battison's rules and therefore cannot be assigned to any of the types mentioned above. The distribution has remained roughly unchanged, with the changes resulting exclusively from the curation mentioned at the beginning: the exclusion of some signs, and the change of the sign type to Type 2 or Type 3 (asymmetric) for 65 one-handed signs of Type 0 with MAJOR LOCATION = *Arm*.

*Adding Hand Orientation*

As the palm orientation, as the only parameter, cannot be derived from the original data of *ASL-LEX*, these data must be re-annotated. Since manual annotation of 2,171 gestures would require a considerable amount of time, while at the same time an example video is available for each gesture (and, through the combination with ASLLRP

Table 25: Valid combination possibilities of the classes of the *ParaSignRec* taking into account the rules according to Battison

| **Category** | **Phonological property / Parameter** | **Battison Sign Type** 0 | 1 / 2 | 3 |
|---|---|---|---|---|
| Handshape | HANDSHAPE (Beginn) | 54 | $54 \cdot 1$ | $54 \cdot 7$ |
| | HANDSHAPE (Ende) | 54 | $54 \cdot 1$ | $54 \cdot 1$ |
| Movements | PATH MOVEMENT | 3 | $3 \cdot 1$ | $3 \cdot 1$ |
| | ULNAR ROTATION | 2 | $2 \cdot 1$ | $2 \cdot 1$ |
| | FLEXION CHANGE (vier Finger) | $2^4$ | $2^4 \cdot 1^4$ | $2^4 \cdot 1^4$ |
| | SPREAD CHANGE (zwei Fingerp.) | $2^2$ | $2^2 \cdot 1^2$ | $2^2 \cdot 1^2$ |
| Place of execution | MINOR LOCATION | 5 | $6 \cdot 5$ | $6 \cdot 5$ |
| | SECOND MINOR LOCATION | 5 | $6 \cdot 5$ | $6 \cdot 1$ |
| Orientation | ORIENTATION (Beginn) | 24 | $24 \cdot 4$ | $24 \cdot 24$ |
| | ORIENTATION (Ende) | 24 | $24 \cdot 4$ | $24 \cdot 1$ |
| **Valid combination possibilities** | | $1.61 \times 10^{10}$ | $9.29 \times 10^{12}$ | $1.95 \times 10^{13}$ |
| | | | | $2.88 \times 10^{13}$ |

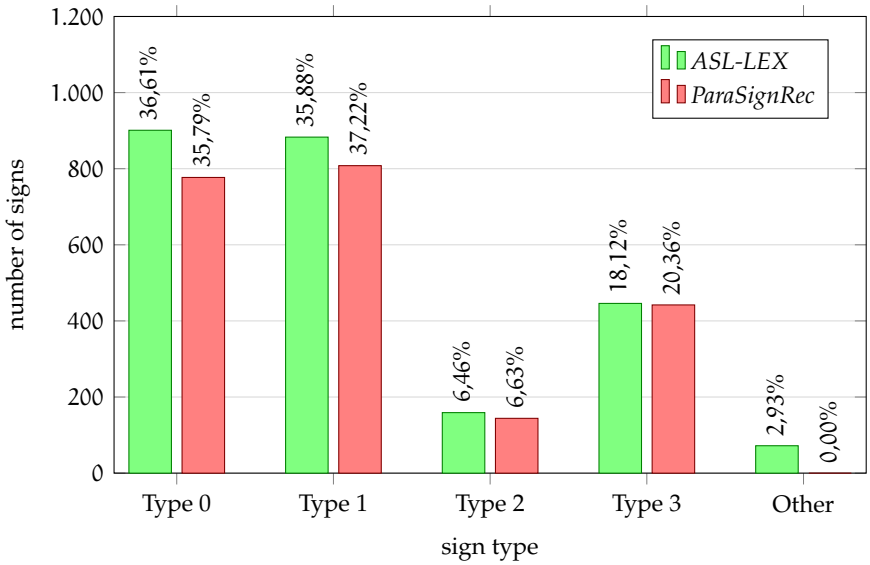



Figure 82: Class distribution according to Battison's sign types for monomorphemic signs of the *ASL-LEX* and the *ParaSignRec* in comparison

and WLASL, theoretically a large number of additional videos for a large proportion of the gestures), it is therefore useful to automate the annotation using an image processing tool such as *MediaPipe* (cf. Figure 2.3.1). In particular, the orientation of the hand can thereby be easily captured and evaluated in the individual frames of the

videos; here, the palm orientation at the beginning and end of the gesture execution is especially suitable, since these time points are clearly defined.

At this point, the focus of the annotation is not primarily on achieving maximal accuracy for the automatically obtained palm orientations. Rather, it serves first as a necessary step to fully define the gesture dictionary, because this information is not contained in the original data and, otherwise, corresponding gaps would exist in the palm-orientation parameters. This completeness is also a prerequisite for the subsequent meta-classification built on top of it.

Moreover, the initial annotation enables basic statistical analyses of the dictionary, for example with regard to the class distribution within individual parameters or the occurrence of redundant entries. After successful conceptual validation of the approach, in a subsequent step the increased effort can be applied to manually annotate all 2,171 gestures and thereby create a reliable data basis. For the actual evaluation of the meta-classification, however, the test gestures to be included are in each case annotated manually independently, in order to rule out possible errors of the automated annotation and to ensure a valid assessment. These data can then also be used to validate the results of this annotation approach.

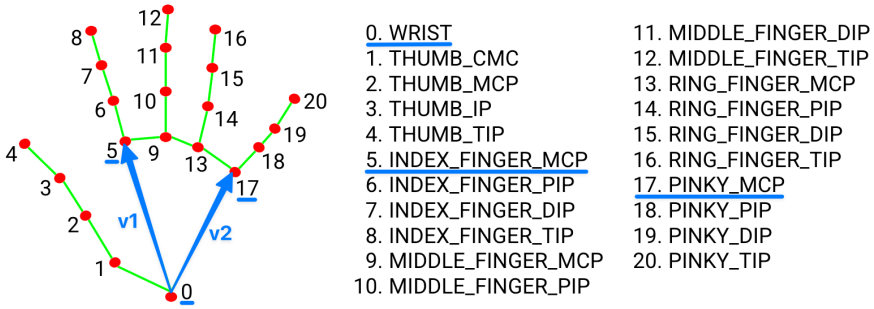



Figure 83: MediaPipe Hand Landmarks (adapted from [78])

Using the *Hand Landmarks Detection* of *MediaPipe*, from the frames of the sign-gesture videos the 2D image coordinates of various hand landmarks are extracted. Based on the SIGN ONSET and SIGN OFFSET values from *ASL-LEX*, the videos are first cropped around the moment of peak meaning as the actual execution of the gesture. For the beginning of the gesture execution, the first 20 % of the frames are considered, and for the end of the gesture execution the last 20 %. To this end, vectors are formed between the individual landmarks, which are used as hand axes for fingers and the palm interior to compute the scores. Two auxiliary vectors are formed, starting from the wrist (*Landmark 0*) to the root of the index finger (*Landmark 5*) and to the root of the little finger (*Landmark 17*), see Figure 83. These approximately span the plane of the palm interior, so that its cross product defines a normal vector in the direction of the palm interior. The finger-direction vector lies within this plane and runs from the wrist to the midpoint between *Landmark 5* and *Landmark 17*. In normalized form, the vectors thus obtained form the basis for the score computation described in Section 5.5.2.

The selection of the final prediction is made via Equation 21, however without sharpening the scores, i.e., with $\gamma = 1$. For each of the 24 classes, the score is

determined and averaged over all frames. The class with the highest score represents the predicted palm orientation.

Since the hands of one-handed gestures are usually not visible in the sign-gesture videos, the palm orientation for the non-dominant hand at the beginning and end of the gesture execution was set to *DownIn*. This corresponds to the I-pose holding position in the assumption that the hand remains in this pose unless it is required to articulate a gesture.

In this way, palm orientation could be estimated for 8,542 parameters of 2,162 gestures, considering both hands at the beginning and end of the gesture execution. For 106 parameters, however, no classes could be determined, because, for instance, the non-dominant hand was occluded by the dominant hand (e.g., ADVERTISE or BOW 1). In these cases, the classes at the beginning of the gesture execution were instead equated with those at the end of the gesture execution, and the classes of the non-dominant hand were equated with those of the dominant hand, and vice versa. These gestures must be checked manually, but then have valid classes and can be used in the context of the meta-classification.

Nine additional gestures had to be annotated manually, because for example the video segments were annotated incorrectly (e.g., BACON) or the videos had an incorrect resolution (e.g., CURTSEY 1), and thus no usable information regarding the palm orientation was available.

### 6.1.3 *Consideration of Multi-Label Parameter Classes*

In the design and implementation of the individual recognition approaches in the preceding chapters, it has become clear that a unambiguous class assignment is not equally suitable for all parameters. In particular for the handshape, but also for hand orientation, cases were observed in which multiple classes had to be considered equally correct. This ambiguity is not to be understood as an error of the annotation or the recognition approaches, but rather arises from the natural variability in sign production.

While *ASL-LEX* generally assumes an unambiguous assignment of one class per phonological feature or treats ambiguities within a parameter as their own, separate classes (e.g., *Lax* as a non-dominant handshape), the analyses in Chapter 4 and Chapter 5 have shown that this assumption cannot be maintained for the *ParaSignRec* in all cases. Instead, for the same parameter, multiple plausible classes can exist, as was also hinted at in Section 3.3.1.

To account for this and at the same time increase the robustness of the system, so-called multi-valued parameter classes are introduced, with the property that a sign can be assigned to multiple classes of a parameter simultaneously. They differ from *multi-label* classification in that the multiple assignment is a property of the sign representation and not the result of a classification process.

An example of this is the class selection limited to 90° steps for hand orientation, even though the hand can in principle take on any orientation and thus orientations can lie between multiple classes. Independently of our considerations, Kita et al. [119]

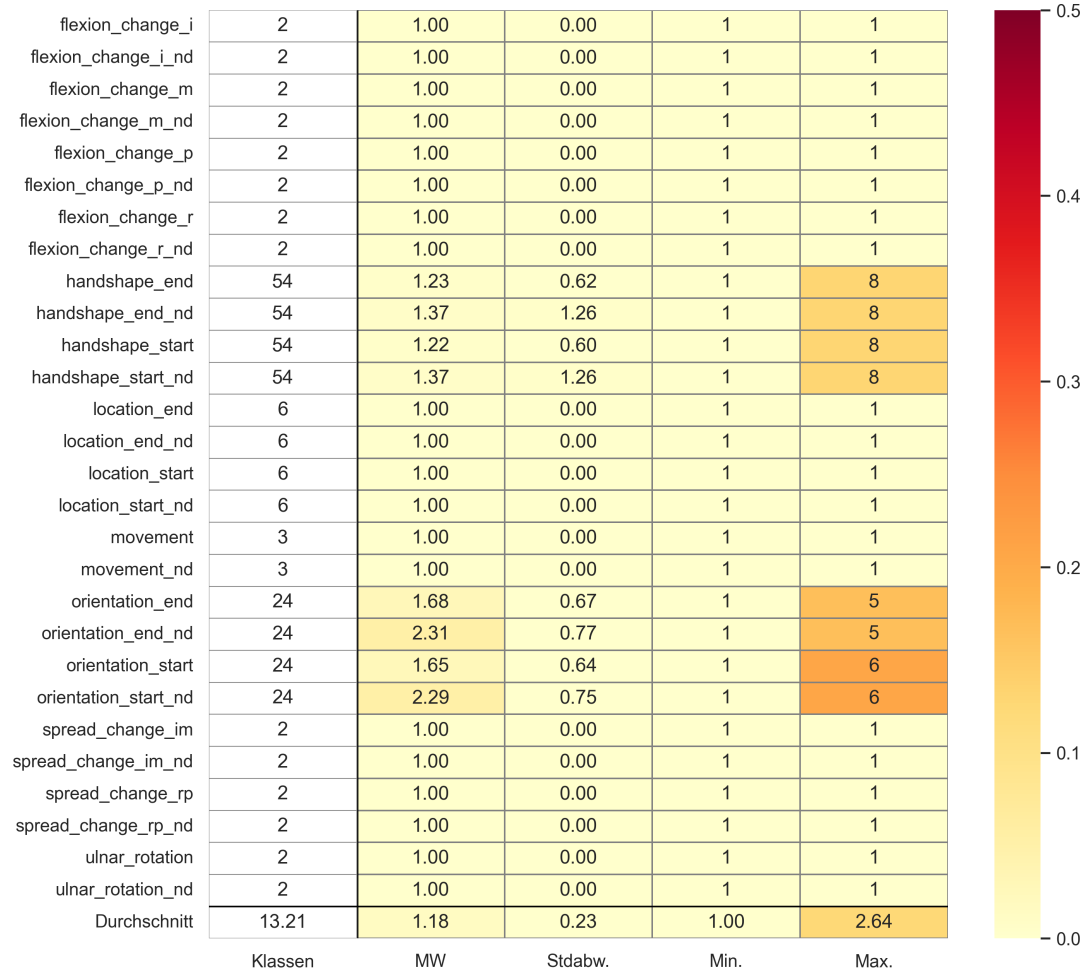
arrived at the same findings with respect to the applied logic, the classes used, and also the possibility of defining multiple correct classes. They use an almost identical system to describe hand orientation as we do, with 24 classes that describe the combined finger and palm-side direction at the beginning and end of the sign production.

To extend our machine annotation of hand orientation from Section 6.1.2 with these multi-valued parameter classes, we no longer select only the class with the highest score, but every class with a score $> 0,0705$. This score allows for a deviation of about 46° of the finger-and-palm axis or 67.50° deviation on one of the two axes, provided that the other axis shows no deviation. Predicted classes from alternative video sources are only considered if they are larger than the score and appear in at least two-thirds of the alternatives in order to limit the number of classes. If no class reaches the score and no class can be determined from alternative video sources, the class with the highest score is chosen. For one-handed signs, hand orientation for the non-dominant hand is supplemented by the two classes *DownBack* and *DownFront*. We maintain our assumption that the non-dominant hand remains in the I-pose for one-handed signs, but we grant it more freedom with a maximum deviation of the palm-side surface of less than 135° (fingers always point downward). This corresponds to a natural posture of a hanging arm.

Even when considering other data sources, such as the databases mentioned in Section 2.2.2 alongside *ASL-LEX*, it becomes apparent that the same sign can exhibit slightly varying handshapes across different sources. For example, the sign CAT has the handshape *F* in *ASL-LEX*, the handshapes *Open 8*, *L*, and *Open F* in ASLLVD – with the latter also used in DSP and RIT – and in RIT it additionally includes the handshape *Q*, i.e., a total of five different but valid handshapes across four sources. The basis for this are the revised annotations of the WLASL signs by Neidle et al. [156], with the help of which the signs of *ASL-LEX* can be mapped to the signs of ASLLVD (v7), DSP (v5) and RIT (v5) using the *main entry* labels. This procedure is also recommended by the authors ("Since the gloss labels are consistent, the main entries from the two datasets can be merged for purposes of sign recognition research." [156]) and applies equally to all three datasets of ASLLRP. In this way, the signs can be supplemented both with the dominant and with the non-dominant handshape, and both at the beginning and at the end of the sign production. For all 748 signs that also exist in *ASL-LEX*, this means that, on average, 1.78 handshapes can be added for the beginning of the sign production and 1.68 handshapes for the end of the sign production for the dominant hand. The values for the non-dominant hand are slightly lower at 1.12 and 1.06, respectively, which can be explained by the missing handshapes in 301 one-handed signs.

In addition, there are special cases introduced by particular classes such as the non-dominant handshape *Lax*. In the *ParaSignRec*, this cannot be treated as an independent, unambiguously interpretable handshape, but is instead replaced by the combination of the seven *Unmarked Handshapes* and the dominant handshape, in order to satisfy Battison's sign types Type 2 (same handshape) and 3 (different handshape) equally and to avoid the need for manual annotation. For one-handed signs, furthermore, by definition the non-dominant hand is inactive, so in principle any handshape is permis-

sible here. These aspects also argue against strictly unambiguous class assignment and motivate the introduction of multi-valued classes.



| | Klassen | MW | Stdabw. | Min. | Max. |
|---|---|---|---|---|---|
| flexion_change_i | 2 | 1.00 | 0.00 | 1 | 1 |
| flexion_change_i_nd | 2 | 1.00 | 0.00 | 1 | 1 |
| flexion_change_m | 2 | 1.00 | 0.00 | 1 | 1 |
| flexion_change_m_nd | 2 | 1.00 | 0.00 | 1 | 1 |
| flexion_change_p | 2 | 1.00 | 0.00 | 1 | 1 |
| flexion_change_p_nd | 2 | 1.00 | 0.00 | 1 | 1 |
| flexion_change_r | 2 | 1.00 | 0.00 | 1 | 1 |
| flexion_change_r_nd | 2 | 1.00 | 0.00 | 1 | 1 |
| handshape_end | 54 | 1.23 | 0.62 | 1 | 8 |
| handshape_end_nd | 54 | 1.37 | 1.26 | 1 | 8 |
| handshape_start | 54 | 1.22 | 0.60 | 1 | 8 |
| handshape_start_nd | 54 | 1.37 | 1.26 | 1 | 8 |
| location_end | 6 | 1.00 | 0.00 | 1 | 1 |
| location_end_nd | 6 | 1.00 | 0.00 | 1 | 1 |
| location_start | 6 | 1.00 | 0.00 | 1 | 1 |
| location_start_nd | 6 | 1.00 | 0.00 | 1 | 1 |
| movement | 3 | 1.00 | 0.00 | 1 | 1 |
| movement_nd | 3 | 1.00 | 0.00 | 1 | 1 |
| orientation_end | 24 | 1.68 | 0.67 | 1 | 5 |
| orientation_end_nd | 24 | 2.31 | 0.77 | 1 | 5 |
| orientation_start | 24 | 1.65 | 0.64 | 1 | 6 |
| orientation_start_nd | 24 | 2.29 | 0.75 | 1 | 6 |
| spread_change_im | 2 | 1.00 | 0.00 | 1 | 1 |
| spread_change_im_nd | 2 | 1.00 | 0.00 | 1 | 1 |
| spread_change_rp | 2 | 1.00 | 0.00 | 1 | 1 |
| spread_change_rp_nd | 2 | 1.00 | 0.00 | 1 | 1 |
| ulnar_rotation | 2 | 1.00 | 0.00 | 1 | 1 |
| ulnar_rotation_nd | 2 | 1.00 | 0.00 | 1 | 1 |
| Durchschnitt | 13.21 | 1.18 | 0.23 | 1.00 | 2.64 |



Figure 84: Number of classes per parameter and their occurrences for 2,171 signs in the *ParaSignRec* after adding multi-valued parameter classes for handshapes and hand orientations. The color scale visualizes the range of the number of classes per parameter.

Figure 84 shows the number of classes per parameter and their occurrences for 2,171 signs in the *ParaSignRec* after adding multi-valued parameter classes for handshapes and hand orientations. Not considered here are the non-dominant handshapes for one-handed signs, since these can in principle take on any handshape and would therefore be distributed uniformly across all classes.

The 28 parameters shown have on average 13.21 classes, of which 1.18 classes per sign are used. However, this considers only the four parameters of handshape (average 1.30) and hand orientation (average 1.98) that use multiple classes. The number of classes for hand orientation of the non-dominant hand is slightly higher (average 2.30) than for the dominant hand (average 1.67), which is due to the use of the three replacement classes in one-handed signs. Similarly, the handshapes of the non-dominant hand

have on average 0.15 more classes than those of the dominant hand. This can be explained by substituting the *Lax* handshapes with the seven *Unmarked Handshapes* and the dominant handshape.

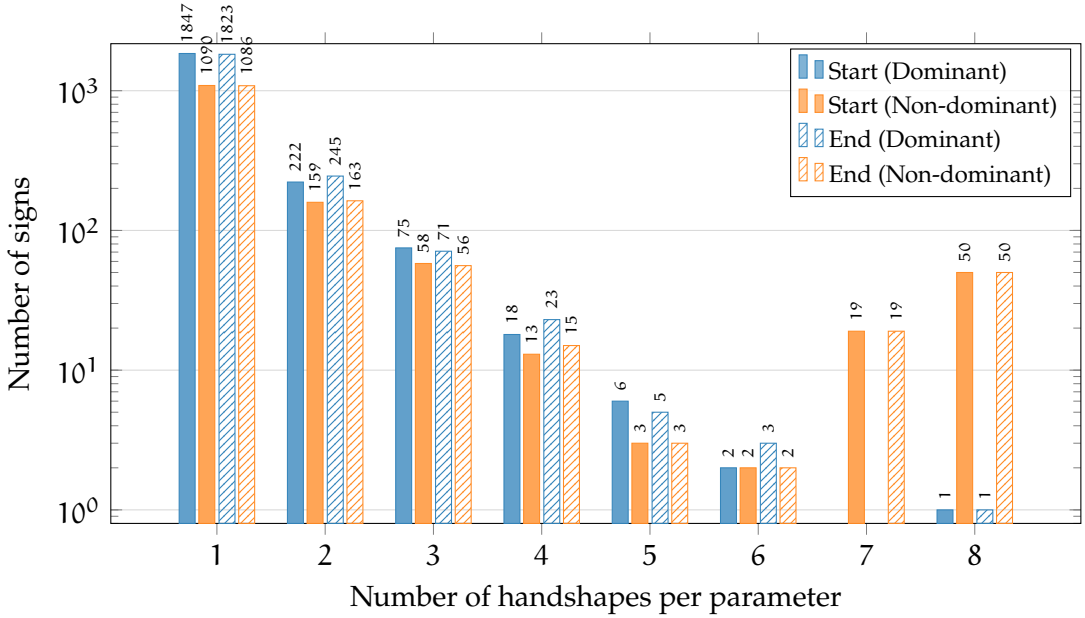



Figure 85: Distribution of the number of handshapes per sign at the beginning and end of sign production (logarithmic scaling)

The distribution of the number of occurring handshapes can be viewed in Figure 85. The sign GUITAR has the highest number of handshapes for both hands at the beginning and end of the sign with eight handshapes. The handshapes[84] of the dominant hand all belong to ASLLVD, except for *Spread Open E*, as the original handshape from *ASL-LEX*. The handshapes of the non-dominant hand consist of the seven *Unmarked Handshapes* together with the original dominant handshape of *ASL-LEX* and replace the *ASL-LEX* class *Lax*. This occurs for 49 additional signs.

Figure 86 shows in comparison the distribution of the number of occurring palm orientations. The signs with the most palm orientations at the beginning of sign production is SHOOT 2 with six classes[85]. The signs with the most palm orientations at the end of sign production is SWEATER with five classes[86]. Our video-based annotation tool was able to identify two different classes for each hand at both signs using the *ASL-LEX* videos, and in total one or two additional classes from alternative video sources, of which ten or seven were available. Due to the symmetrical behavior of both signs, the number and selection of classes for the dominant and non-dominant hand were matched, so that both include the same class selection.

By taking into account multi-valued parameter classes, the computation of the used metrics must also be adapted, since for a parameter several classes may now

84 *5*, *Curved 5*, *Spread Flat 5*, *A*, *Beak*, *Spread Open E*, *Spread E* and *F*
85 *DownFront*, *DownIn*, *InBack*, *UpBack*, *UpFront* and *UpIn*
86 *FrontUp*, *InUp*, *DownFront*, *InBack* and *InFront*

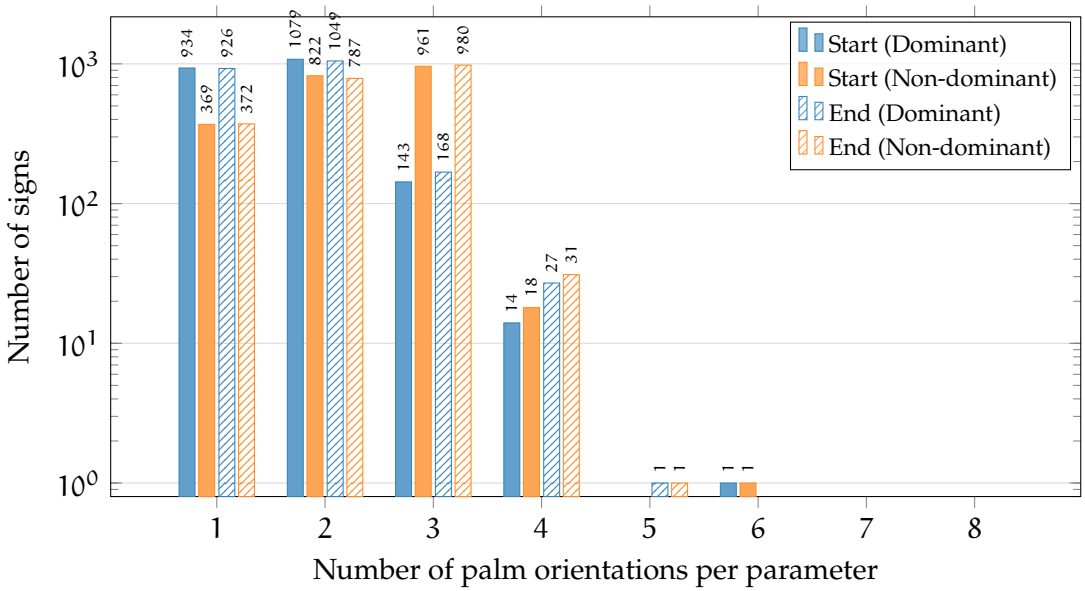



Figure 86: Distribution of the number of palm orientations per sign at the beginning and end of sign production (logarithmic scaling)

be considered correct, while the classifier continues to predict exactly one class. If one of the permissible classes is predicted, this prediction is counted as correct. The other permissible classes in this case are not counted as errors, but instead as TNs, i.e., "not applicable". If no permissible class is predicted, the prediction is counted as an error; that is, the first listed permissible class is counted as FN and the predicted, impermissible class as False Positive (FP). The remaining permissible classes are still treated as TNs.

The subsequent calculations of Accuracy, Precision, Recall and the $F_1$-score are performed unchanged as described in Section 3.2.1, with overall values determined via macro-averaging over the classes present in the data.

### 6.1.4 *Selection Procedure and Manual Annotation of the Test Signs*

In order to perform a final evaluation of our recognition approaches and of the meta-classifier as a connecting element for identifying the target sign, some signs must be included as test data. In an ideal scenario, all signs of the *ParaSignRec* would be recorded multiple times by different individuals in order to compensate for differences in the execution of signs across different signers as well as for errors in the recordings, such as noise and drift in the sensor data. However, given a sign vocabulary of 2,171 signs, this would require an enormous effort, which is why a sample-based evaluation is carried out. To obtain the most informative result possible, as many classes as possible should be represented in the sample. Moreover, the selection of signs should be random, and the number of signs to be recorded should be minimized in order to reduce the effort of the recording process as well.

For the selection of the test signs, initially random signs were chosen from the vocabulary for the test dataset until all classes present in the vocabulary for the individual parameters were covered. Afterwards, the dataset was iteratively reduced by removing redundant signs, i.e., signs for which each of their classes was already present in the test dataset.

For classes that occurred multiple times in the test dataset, it was iteratively checked whether individual signs could be replaced by more suitable alternatives that either covered multiple relevant classes at once or made it possible to remove other redundant signs in a later step.

This strategy was repeated until the dataset could no longer be reduced. Since the selection of signs was random at the beginning, the algorithm was executed multiple times to determine the smallest possible test dataset.

Since, unfortunately, we did not receive support from the deaf community during the recording of the test data, we had to ensure in the selection of signs to be recorded that they could be reproduced as well as possible by laypeople. To this end, we made use of the *phonological complexity* of the *ASL-LEX* for the individual signs. Phonological complexity is described by Morgan et al. [146]: For each of the following seven categories, one point is awarded if the respective condition for the first morpheme of a sign is met; otherwise the value remains unchanged. The categories include *i)* two-handed signs with different handshape (corresponding to *Type 3* of the Battison Sign Types), *ii)* violation of Battison's *rule of symmetry or dominance*, *iii)* other selected finger groupings (see *Selected Fingers* in Section A.1.2) than all (index, middle, ring, and little fingers) or only the index finger, *iv)* other finger flexions (see *Flexion* in Section A.1.2) than fully extended or fully closed, *v)* stacked or crossed fingers, *vi)* other movement paths than straight and *vii)* more than one type of movement (e.g., change in finger flexion or spread and movement of the hand) .

In order to obtain the widest possible range of signs while, at the same time, not setting the complexity too high for the reasons mentioned above, we limited it to a maximum value of 4, whereas the highest value in the dataset was 6. Figure 87 shows that this decision excludes 19 signs, i.e., 0.88 % of the *ParaSignRec* vocabulary, at an average complexity of 1.93. The most complex sign is KITE 2. Relative to the original *ASL-LEX*, there are 39 signs, i.e., 1.26 % of the vocabulary, with an average of 2.01, whereas the test signs are at 2.14 and thus even exhibit the highest complexity.

Overall, using this procedure, three lists with different signs were created so that each class is used in at least three signs (provided it occurs at least three times in the signs of the *ParaSignRec* vocabulary). To prevent signs from appearing in multiple lists, the selected signs were removed from the pool of signs to be chosen after each list creation, before creating the next list.

The selection of the test signs was made relatively early in the course of this work for several strategic reasons. At that time, many later findings about parameters and classes were not yet available. Therefore, the selection was based on the parameters and classes of the *ASL-LEX*, and as a result, in particular, the classes of hand orientation were not taken into account. Nevertheless, all 24 classes are represented among

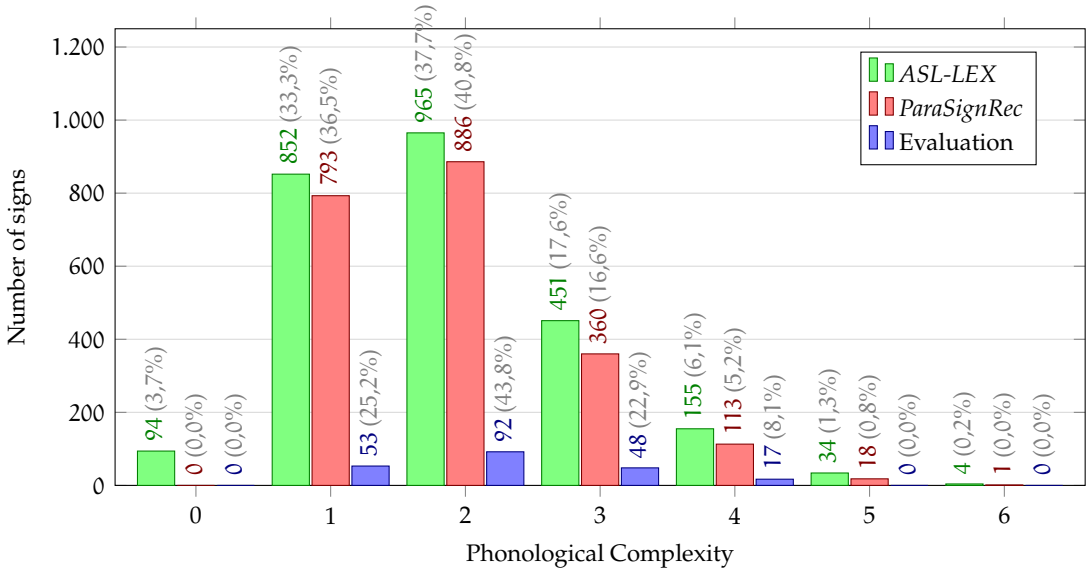


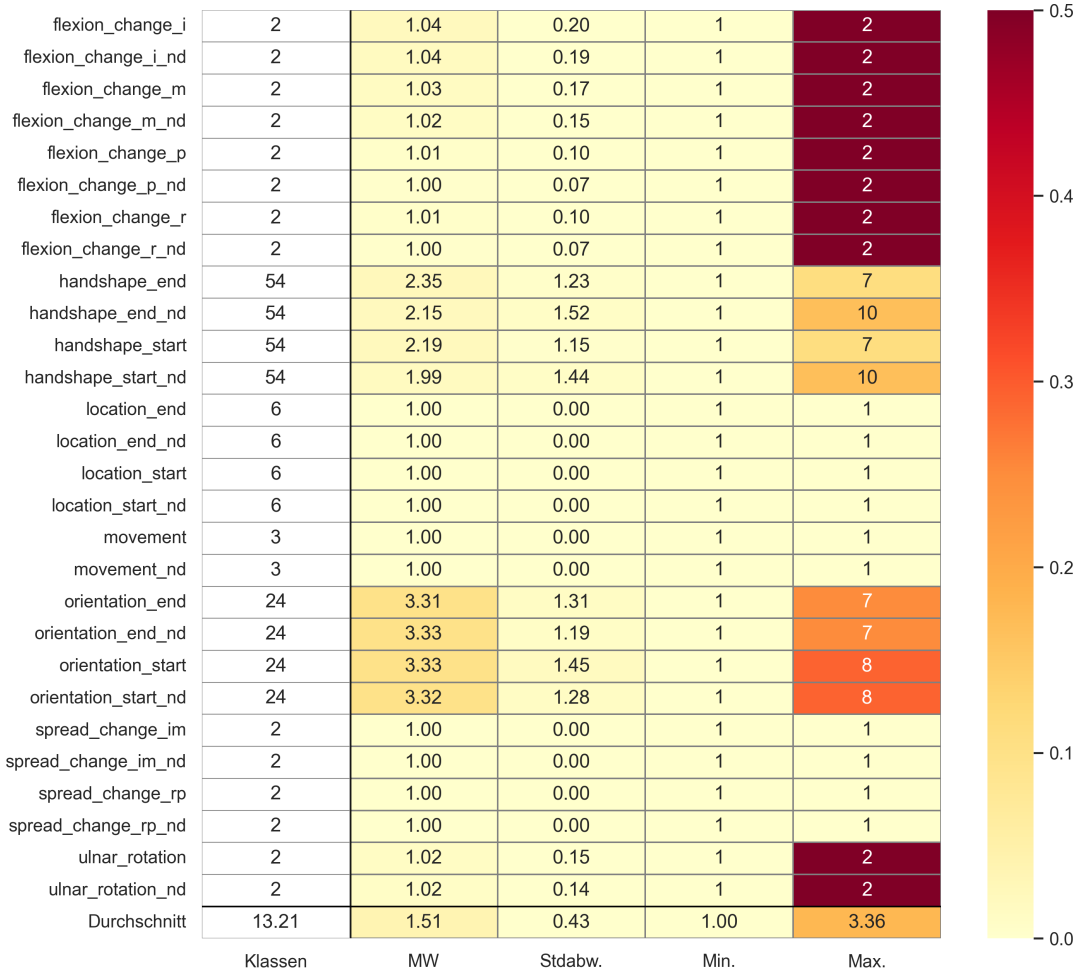

Figure 87: Number of signs by phonological complexity, in *ASL-LEX*, in ParaSignRec and in the evaluation test signs

the signs in the evaluation; however, four classes appear in only three of the four parameters.

In addition, each selected sign was manually annotated, i.e., all parameters were validated based on the sign video and adjusted if necessary. In particular, handshapes and hand orientations were supplemented with possible classes or validated if they already existed. For the movements as well, some special cases were added in which a movement is simultaneously allowed, but also the absence of movement corresponds to a valid prediction. This is the case for ten signs for the values of FLEXION CHANGE, such as for the signs ADMIRE or HALF HOUR, because it is not clear whether the change in finger orientation occurs via a change in finger flexion or via a movement of the wrist. For the signs DIG, ERUPT 2, and three other signs, however, it is not unambiguously recognizable whether a ULNAR ROTATION is present or not.

Figure 88 shows the updated class distribution from Section 6.1.3 for the 210 test signs. In addition to the already mentioned special cases regarding FLEXION CHANGE and ULNAR ROTATION, in particular the number of hand orientations increased on average from 1.98 to 3.32. This corresponds to approximately a doubling of the dominant hand, or to roughly one additional class for the non-dominant hand. For handshapes, an increase from 1.30 to 2.17 can be observed, which in turn means an increase of about one class for the dominant hand and about +0.7 classes for the non-dominant hand. The average of all class occurrences increased from 1.18 to 1.51.

The substantially higher number of hand orientations in the manual annotation arises from the fact that, unlike automatic annotation, sign variants are also taken into account. While automatic methods capture only the state observed in the considered frame, manual annotation allows for a contextual interpretation of the sign and thus an assessment of which features are relevant for meaning.



| | Klassen | MW | Stdabw. | Min. | Max. |
|---|---|---|---|---|---|
| flexion_change_i | 2 | 1.04 | 0.20 | 1 | 2 |
| flexion_change_i_nd | 2 | 1.04 | 0.19 | 1 | 2 |
| flexion_change_m | 2 | 1.03 | 0.17 | 1 | 2 |
| flexion_change_m_nd | 2 | 1.02 | 0.15 | 1 | 2 |
| flexion_change_p | 2 | 1.01 | 0.10 | 1 | 2 |
| flexion_change_p_nd | 2 | 1.00 | 0.07 | 1 | 2 |
| flexion_change_r | 2 | 1.01 | 0.10 | 1 | 2 |
| flexion_change_r_nd | 2 | 1.00 | 0.07 | 1 | 2 |
| handshape_end | 54 | 2.35 | 1.23 | 1 | 7 |
| handshape_end_nd | 54 | 2.15 | 1.52 | 1 | 10 |
| handshape_start | 54 | 2.19 | 1.15 | 1 | 7 |
| handshape_start_nd | 54 | 1.99 | 1.44 | 1 | 10 |
| location_end | 6 | 1.00 | 0.00 | 1 | 1 |
| location_end_nd | 6 | 1.00 | 0.00 | 1 | 1 |
| location_start | 6 | 1.00 | 0.00 | 1 | 1 |
| location_start_nd | 6 | 1.00 | 0.00 | 1 | 1 |
| movement | 3 | 1.00 | 0.00 | 1 | 1 |
| movement_nd | 3 | 1.00 | 0.00 | 1 | 1 |
| orientation_end | 24 | 3.31 | 1.31 | 1 | 7 |
| orientation_end_nd | 24 | 3.33 | 1.19 | 1 | 7 |
| orientation_start | 24 | 3.33 | 1.45 | 1 | 8 |
| orientation_start_nd | 24 | 3.32 | 1.28 | 1 | 8 |
| spread_change_im | 2 | 1.00 | 0.00 | 1 | 1 |
| spread_change_im_nd | 2 | 1.00 | 0.00 | 1 | 1 |
| spread_change_rp | 2 | 1.00 | 0.00 | 1 | 1 |
| spread_change_rp_nd | 2 | 1.00 | 0.00 | 1 | 1 |
| ulnar_rotation | 2 | 1.02 | 0.15 | 1 | 2 |
| ulnar_rotation_nd | 2 | 1.02 | 0.14 | 1 | 2 |
| Durchschnitt | 13.21 | 1.51 | 0.43 | 1.00 | 3.36 |



Figure 88: Number of classes per parameter and their occurrences for 210 manually annotated signs of the *ParaSignRec*. The color scale visualizes the range of the number of classes per parameter.

This can be illustrated by the sign TREE. For this sign, it is crucial that the dominant arm, with extended fingers, is oriented upward, whereas the exact hand orientation carries less weight, because during execution the hand rotates around the arm axis. An automatic annotation method therefore recognizes only the upward-directed arm posture together with a single, frame-specific hand orientation. A manual annotation, in contrast, would consider multiple permissible hand orientations and assign them to the sign, for example, three of six possible orientations.

The automatic annotation approach, on the other hand, is limited to the orientation observed in the respective frame and can at most summarize cases that are not unambiguous, such as diagonal orientations. Furthermore, the influence of alternative video sources is likely to be rather small, since alternative hand orientations are considered only if they can be identified in at least two thirds of the video sources.

The lists are provided in Section A.4, Table A.13 to Table A.19, and contain a total of 210 signs, i.e., 9.67 % of the signs in the *ParaSignRec*, split into 67, 77, and 66 signs.

The tables include, for each sign, all classes of all 28 parameters, as well as the sign type and its complexity.

The manual annotation of hand orientation now also makes it possible to evaluate the approach for automatic hand orientation annotation presented in Section 6.1.2. To this end, the automatically determined hand orientation labels for each of the 210 signs are compared with the manually annotated label sets, separately by hand (dominant/non-dominant) and position (start/end). The accuracy is computed across all signs as the ratio of the number of correctly predicted labels (those that occur in both sets) to the total number of all predicted labels. This also takes into account the fact that manual annotation includes variants and alternatives of the signs, and that a missing label from these should not lead to a penalty for the automatic annotation.

Accordingly, for the annotation of hand orientation at the beginning of sign execution, an accuracy of 80.90 % is achieved for the dominant hand and 82.60 % for the non-dominant hand. For the annotation of hand orientation at the end of sign execution, the accuracy is 78.80 % for the dominant hand and 87.30 % for the non-dominant hand. A negative effect on recognition could be caused by the camera perspective of the sign videos, which is identical across all videos, but does not exactly correspond to the canonical coordinate directions used. Potentially, an ML-based approach could yield more robust results than the rule-based approach, provided that a sufficient amount of annotated data is available.

### 6.1.5 *Class Frequencies and Diversity of Combinations*

After the manual annotation of 210 signs, a final look at the final vocabulary of the *ParaSignRec* should be taken. We will examine how many of the 2,171 can be identified unambiguously, how the number of duplicates has developed compared to the *ASL-LEX*, and what influence the use of multi-valued parameter classes has on this. To this end, we first take a look at how the actually used classes are distributed across the parameters of the *ParaSignRec*.

Figure 89 shows the normalized distribution of the classes used in *ParaSignRec* within each parameter. Each parameter is represented by a colored bar, with parameters of the dominant hand shown in blue and those of the non-dominant hand in orange. In the upper plot, the classes are shown as points, whose positions on the ordinate correspond to their distribution factors $q_i$, with $i \in \{1, \ldots, K\}$. The distribution factor $q_A$ of a class $A$ is defined as

$$q_A = \frac{N_A}{N/K},$$

where $N_A$ is the number of labels of class $A$, $N$ is the total number of all labels, and $K$ is the number of classes. It describes the ratio of the actual to the optimal number of labels per class, where the optimal number corresponds to an ideal uniform distribution. A

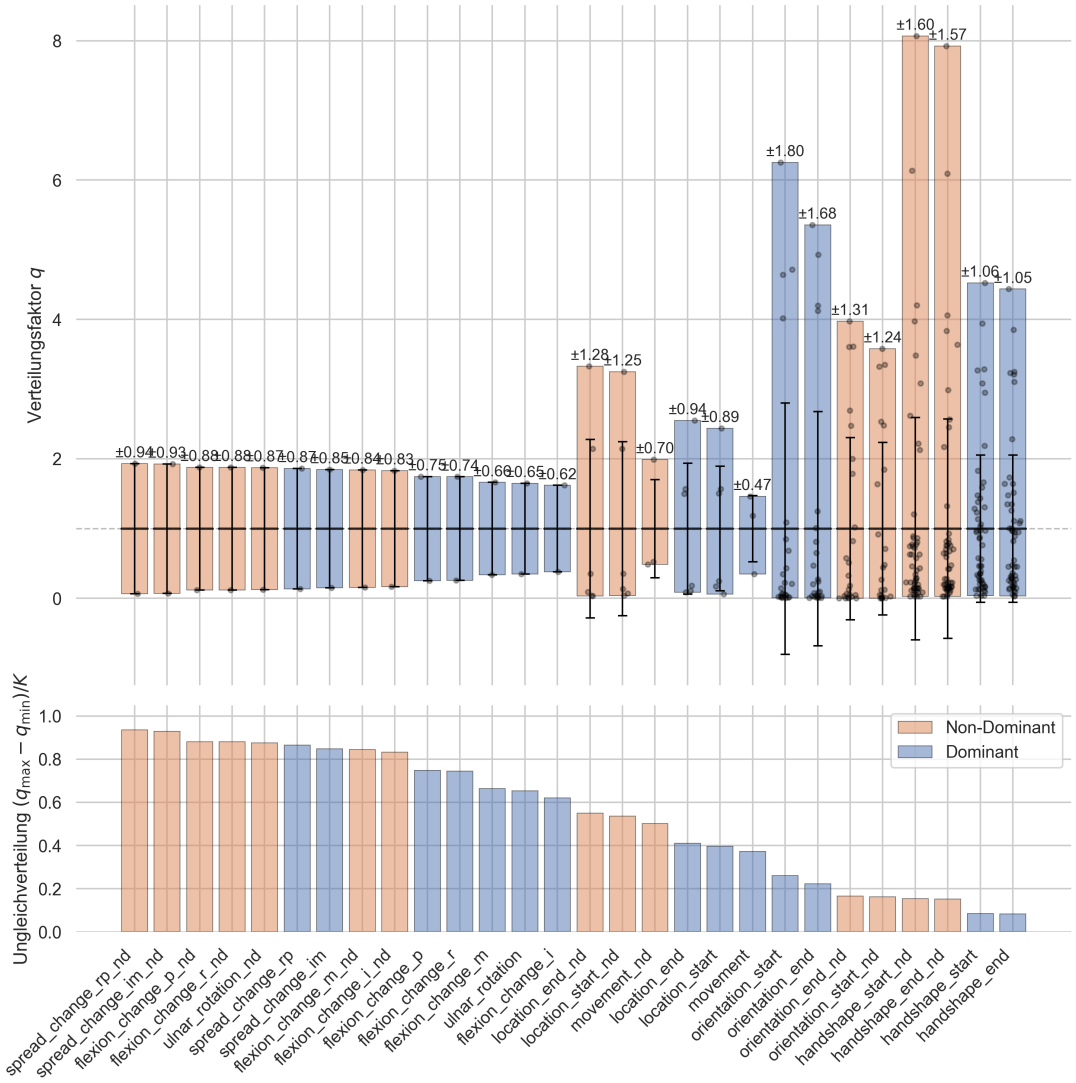



Figure 89: Normalized class distribution of the parameters in *ParaSignRec* for the dominant (blue) and non-dominant (orange) hand. In the upper plot, the classes $q_A$ are shown as points, the bars indicate the range $[q_{min}, q_{max}]$, and the error bars represent the standard deviation at the normalized mean value $q = 1$. In the lower plot, for direct comparison, the span normalized by the number of classes K, $(q_{max} - q_{min})/K$, is shown for each parameter.

value of $q_A = 1$ corresponds to a uniformly distributed class, whereas $q_A > 1$ and $q_A < 1$ indicate over- or under-representation, respectively:

$$q_A \begin{cases} = 1, & \text{uniformly distributed class} \\ > 1, & \text{over-represented class} \\ < 1, & \text{under-represented class} \end{cases}$$

Additionally, the mean and standard deviation of the distribution factors within a parameter are shown as black error bars, where the mean is always at $y = 1$ due to the normalization.

The diagram thus enables a visual identification of those parameters in which a higher degree of class imbalance occurs. The larger a bar relative to its number of classes, the more strongly the class distribution of the corresponding parameter deviates from a uniform distribution. Therefore, the sizes of the bars of the individual parameters can only be compared directly if they have the same number of classes.

For this reason, in the lower plot, for direct comparability, the normalized spans $\frac{q_{max}-q_{min}}{K}$ of the individual parameters are shown, sorted in descending order by inequality, i.e., first the parameters with the highest inequality and last the parameters with the most uniformly distributed classes.

For our approach, a class distribution that is as uniform as possible within each parameter is desirable, because it ensures that all classes contribute in comparable proportions to the identification of the target sign within the meta-classification. Uneven class distributions, on the other hand, cause some classes to dominate while others are insufficiently represented, thereby reducing the separability of the parameter.

In the most extreme case, this would imply the exclusive use of a single class within a parameter. Such a parameter provides no discriminative information for the meta-classification, because no variation between classes exists. In Figure 89, such a parameter would be characterized by a bar that extends from the value $y = 0$ to $y = K$ in the upper plot, or from $y = 1$ in the lower plot.

For the binary movement classifiers (Flexion Change, Spread Change and Ulnar Rotation), the highest degree of class imbalance can be observed, since the stationary classes clearly dominate. This effect is particularly pronounced for the non-dominant hand, because by definition it can only play an active role in symmetric signs. As a result, the standard deviation here is on average 0.16 higher, with a mean value of 0.86 compared to the dominant hand. The parameters for handshape and palm orientation show a relatively low degree of imbalance, which is attributable to the high number of classes.

For the parameters of the non-dominant handshape, a considerable over-representation of the seven *Unmarked Handshapes* (*1*, *5*, *A*, *Closed B*, *C*, *O* and *S*) can be observed. Regardless of the time of execution of the sign, these handshapes constitute 49.58 to 49.62 % of the handshapes used, which is why they all fall among the ten most frequently used handshapes of these two parameters out of 52. The most frequently used non-dominant handshape is *Closed B*, with 280 occurrences at the beginning of sign execution and 281 occurrences at the end of sign execution, followed by *Open B* with 213 and 216 occurrences, respectively, and *1* with 146 and 144 occurrences, respectively. Even in this representation, the non-dominant handshape in one-handed signs is not considered, since it is not explicitly defined.

Also for the dominant handshapes, the *Unmarked Handshapes* are among the 15 and 13 most frequent classes, respectively (among the 8 and 9 most frequent classes without considering the handshape *O*), and each account for 34.07 to 34.49 % of the

handshapes, underscoring their role as basic handshapes. The most frequent handshape is, however, *Open B*, with a share of approximately 8.50 %. Compared to the non-dominant hand, the class distribution is less strongly pronounced here, because the distances between individual classes are significantly smaller.

For the palm orientation of the dominant hand, the four classes *UpIn*, *UpBack*, *UpFront* and *InBack* dominate, together occurring in 81.79 % of the signs at the beginning of execution and in 77.56 % at the end of execution. For the non-dominant hand, primarily the three classes for one-handed signs dominate, each accounting for a total of about 46.60 %. Next come the four dominant classes of the dominant hand, with shares totaling 38.67 % and 37.35 %, respectively. Again, only seven out of a total of 24 classes account for the majority of 85.27 % and 83.95 % of the occurrences.

The preceding analysis shows that parameters with a pronounced imbalance reduce the effective separability within the meta-classification, because frequent classes are weighted disproportionately strongly. A direct consequence of this reduced discriminative ability is the increased occurrence of signs that cannot be distinguished unambiguously from one another. High classification accuracy is only meaningful if the underlying parameter space allows sufficiently distinct differentiation between individual signs.

An extensive vocabulary alone is not a sufficient criterion: if multiple signs share identical parameter combinations, they cannot be distinguished in principle by the approach. In such a case, the classification performance of the meta-classification is usually higher, because one only needs to choose among the available alternatives and the influence of the underlying individual classifiers becomes smaller. The resulting accuracy then primarily reflects the small number of available alternatives rather than the discriminative capability of the parameters used.

Against this background, it is necessary to analyze the structure of the final vocabulary to determine to what extent signs can be identified unambiguously by the combination of their parameter classes, and how frequently duplicates occur in the vocabulary.

Determining the number of duplicates for uni-valued parameter classes can be done straightforwardly by comparing the unique combinations of all parameter classes. For a sign $g \in G$, there are exactly as many duplicates as there are other signs in the vocabulary $G = \{g_1, g_2, \ldots, g_{|G|}\}$ that share the same class for every parameter. A sign can be identified unambiguously if it has no duplicates, and it cannot be identified unambiguously if it has duplicates.

For multi-valued parameter classes, however, each sign cannot have a unique combination of all parameter classes; instead, it can have any set $V(g) = \{v_{g,1}, v_{g,2}, \ldots, v_{g,k_g}\}$ of combinations, hereafter referred to as *sign variants*. Each sign variant by itself may therefore have its own duplicates, and it is called *exclusive* if it belongs to exactly one single sign. This means that a sign with multi-valued parameter classes can be identified as long as it has exclusive sign variants. A sign $g$ is considered *strongly identifiable* if all of its variants are exclusive, and thus it is identifiable regardless of which of its variants is actually executed or recognized in a given case. It is considered *weakly identifiable* if it has both exclusive and non-exclusive variants. In this case, $g$ can only be

identified once one of its exclusive variants is executed or recognized. Identifiability is therefore variant-dependent and not always guaranteed.

The size of our unique vocabulary therefore lies between the number of strongly identifiable signs as a conservative lower bound on identifiability—since each of the counted signs remains uniquely identifiable across all allowed variants—and the number of weakly identifiable signs as a less conservative upper bound, because the mere presence of a single exclusive variant is sufficient. The actual identifiability then depends on which variant is realized in the given context. It holds that every strongly identifiable sign is also weakly identifiable, and the number of identifiable signs does not exceed the total number of signs in the vocabulary.

Table 26 shows the development of the number of identifiable signs in *ParaSignRec* while accounting for all parameter and class changes performed, starting from the original 2,171 of the *ParaSignRec* as listed in Table 22 for the simplifications of the *ASL-LEX* dataset, and before any other changes to the vocabulary were applied. The individual rows are not independent of each other and therefore the effects of the listed measures should not be understood as additive; instead, each effect is relative to the initial configuration of the first row. The left result column represents the lower bound of the number of identifiable signs (strongly identifiable), and the right result column represents the upper bound (weakly identifiable).

Since the measures are partly difficult to distinguish, the table should be viewed more as an estimate. For example, by adding alternative handshapes using WLASL, handshapes are also added that occur until the end of sign execution and should actually be considered separately. Nevertheless, the first row of the table corresponds to the actual number of unambiguously identifiable signs in the *ASL-LEX* after simplifications, and the last row represents the actual range of identifiable signs in *ParaSignRec*.

Overall, it can be seen that the number of strongly identifiable signs has decreased by about 33.49 % compared to the initial situation, whereas the number of weakly identifiable signs has increased by 9.43 %. This is mainly due to the high number of variants, which on the one hand increases the risk of having duplicates, but on the other hand also increases the chances of being exclusive. It also becomes apparent that discrepancies between strongly identifiable signs and weakly identifiable signs arise only through the addition of multi-valued parameter classes. Considering symmetry properties in the absence of new information leads to rather negligible changes in the vocabulary. Most of the loss in the size of the identifiable vocabulary can be traced back to the reduction of the execution locations from initially 37 to ultimately six classes, where the consolidation of the movement classes—including the disregard of the Repeated Movements—also led to a substantial loss. The largest improvements in vocabulary size come from adding palm orientation, resulting in up to 184 additional strongly identifiable signs and 344 additional weakly identifiable signs. Due to the dominance of four classes with 77.56 to 81.79 % of occurrences for the dominant hand, and seven classes with 83.95 to 85.27 % for the non-dominant hand, the positive impact on the development of the number of identifiable signs is likely somewhat smaller than would be expected for 24 classes.

Table 26: Development of the number of identifiable signs in the *ParaSignRec* while taking into account all changes made to parameters and classes, subdivided into strongly identifiable signs (left results column) and weakly identifiable signs (right results column)

| Chapter | Measure | ΔVocabulary | |
|---|---|---|---|
| – | **2,171 cleaned signs of the *ASL-LEX*** | **1,538** | **1,538** |
| 4 | *Hand form* | | |
| 4.1.1 | Removal of the hand form configuration | −6 | −6 |
| 4.1.2 | Consideration of the hand form (end, dominant) | ±0 | ±0 |
| 4.1.2 | Consideration of the hand form (end, non-dominant) | ±0 | ±0 |
| 4.5 | Removal of THUMB CONTACT | −6 | −6 |
| 4.5, 4.5.3 | Merging curved and bent hand forms | ±0 | ±0 |
| 6.1.3 | Addition of alternative hand forms using WLASL | −4 | +72 |
| 6.1.4 | Manual annotation of hand forms of the test signs | +4 | +28 |
| 5.3 | *Place of execution* | | |
| 5.3.2 | Removal of MAJOR LOCATION | ±0 | ±0 |
| 5.3.2 | Consideration of the place of execution (start, non-dominant) | ±0 | ±0 |
| 5.3.2 | Consideration of the place of execution (end, non-dominant) | ±0 | ±0 |
| 5.3.5, 5.3.2 | Consolidation and optimization of the places of execution | −323 | −323 |
| 5.3.5 | Removal of CONTACT | −85 | −85 |
| 5.4 | *Movement* | | |
| 5.4.1 | Fusion of Flexion Change and Selected Fingers | −2 | −2 |
| 5.4.1 | Fusion of Spread Change and Selected Fingers | −6 | −6 |
| 5.4.2 | Consolidation and optimization of the movement classes | −50 | −50 |
| 5.4.3 | Removal of REPEATED MOVEMENT | −151 | −151 |
| 6.1.2 | Consideration of the movement (non-dominant) | +5 | +5 |
| 6.1.2 | Consideration of the Flexion Change (non-dominant) | ±0 | ±0 |
| 6.1.2 | Consideration of the Spread Change (non-dominant) | ±0 | ±0 |
| 6.1.2 | Consideration of the Ulnar Rotation (non-dominant) | +3 | +3 |
| 6.1.4 | Manual annotation of the movements of the test signs | +2 | ±0 |
| 5.5 | *Orientation* | | |
| 5.5 | Consideration of orientation (start, dominant) | +184 | +344 |
| 5.5 | Consideration of orientation (end, dominant) | +184 | +344 |
| 5.5 | Consideration of orientation (start, non-dominant) | +122 | +270 |
| 5.5 | Consideration of orientation (end, non-dominant) | +118 | +269 |
| 6.1.4 | Manual annotation of the orientation of the test signs | −6 | +1 |
| – | **2,171 signs of the *ParaSignRec*** | **1,023** | **1,683** |

As a result, it can be stated that now 47.12 % of the signs are strongly identifiable and 77.52 % are weakly identifiable, whereas originally it was 70.84 %. By comparison, the monomorphemic signs of the *ASL-LEX* have 73.22 % unambiguously identifiable signs (cf. Section 3.1.3). Of the originally eight duplicates in the example shown in Figure 23, only up to three duplicates remain (BURY 3, PUT ASIDE, WIDE or PLAN, STREET, WAY). All signs were previously not identifiable and are now weakly identifiable, with up to eight additional duplicates depending on the variant[87]. Only the sign HALL is still not identifiable.

The entire sign dictionary, including the corrected and extended signs, can be found on our project homepage `https://sign-parametrization.netlify.app/dictionary`.

## 6.2 EVALUATION OF PARAMETER-BASED RECOGNITION APPROACHES

Before evaluating the meta-classification, the 28 parameters are first evaluated in isolation. These can be assigned to the four categories handshape recognition, sign execution location recognition, hand movement recognition, and palm orientation recognition, with an even distribution between the dominant and non-dominant hands in each case.

The results obtained in the following are compared with the results from Chapters 4 and 5, in which the individual recognition approaches were conceived and implemented. In those chapters, however, training and design of the respective approaches were often based on isolated sensor data that each represented only a specific parameter and not the complete sign execution. The present evaluation therefore also serves to investigate to what extent the individual approaches can be transferred to real sign data and how adaptive they are under realistic operating conditions. To this end, the recording conditions for the test data are first explained.

### 6.2.1 *Recording of the evaluation data*

For the evaluation of the overall system, consisting of the individual recognition approaches as well as the meta-classification built upon them, it is necessary to include a suitable test dataset. The goal is to achieve the highest possible class coverage while ensuring sufficient variability within each of the individual signs.

As described in Section 6.1.4, a total of 210 signs were selected for this purpose, which were distributed across three lists. For the evaluation, each of these signs was recorded by at least three different people, with three repetitions each. This procedure is intended to ensure a certain variability in the sign executions and a sufficiently high number of repetitions, in order to be able to exclude, if needed, erroneous or unsuitable recordings without impairing the statistical significance of the dataset.

87 BABY, CORNER, DOES NOT MATTER, MOTIVATION 1, PLAN, SEAL, STREET, WAY, WHATEVER: Symmetric signs, *Open B* handshape in the neutral space with fingers pointing inward and the palm facing backward, straight-line movement, no movement of the fingers or the wrist

The data recording as well as its post-processing were carried out according to the procedures described in Section 5.2. The test dataset was recorded by a total of four participants, consisting of three men and one woman. None of the participants was an active user of a sign language. All participants, however, had basic knowledge in the field of sign language due to their work at the chair or in the context of academic research, in particular with regard to movement execution, place of articulation, and handshape.

These specific skills proved advantageous during the recording process, as they enabled the participants to validate one another. If necessary, immediate feedback could be provided to improve the execution, by checking the correct performance of the signs. This mutual validation, which was expanded compared to earlier recording sessions by involving an additional person, helped further increase the consistency and quality of the recorded data and, at the same time, supported the participants in familiarizing themselves with the correct execution of the signs.

For data gloves, *Manus Quantum Metagloves* were used. If there were deviations between the expected classes and the example sign execution shown in the reference videos, the signs were executed accordingly with the defined classes and the video was not considered further. The three repetitions per sign were recorded continuously and directly one after another.

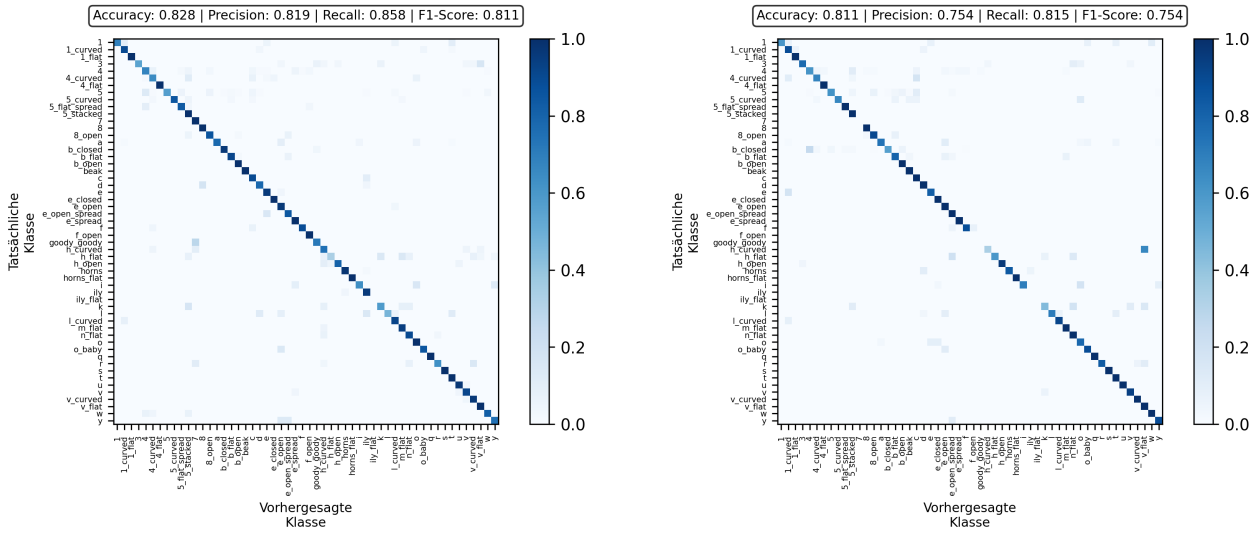
As part of quality assurance, compared to the earlier approach, decisions were no longer made at the level of individual parameters; instead, the respective sign recording was completely excluded as soon as an erroneous parameter was identified. This decision is based on the assumption that reliable recognition of a sign is only possible on the basis of a complete and consistent dataset. The focus of quality assurance was on minimizing human errors, i.e., incorrect executions and segmentations. Recordings with faulty sensor data, in contrast, were not excluded, since the proposed approach is intended to compensate for technical inaccuracies to a certain extent, whereas we can assume correct system usage in human-related aspects. Erroneous annotations were corrected already within the scope of the manual annotation in Section 6.1.4.

In total, each of the 210 signs was recorded up to nine times; however, 133 recordings were subsequently removed from the dataset due to faulty executions, and four recordings due to faulty segmentations. From participant *P1*, the final dataset contains 559 recordings of 188 signs, and from participant *P2* 590 recordings of 202 signs. Since participant *P3.1* was no longer available for later recordings, he was replaced by an independent participant *P3.2*. From participant *P3.1* and *P3.2*, there are 592 recordings of 201 signs. Overall, this provides 1.741 valid recordings of 208 signs, corresponding to a share of 92.12 % error-free sign executions. This represents a clear improvement over the 71.10 % achieved in Section 5.4.1 and over 86.50 % without considering faulty sensor data. This improvement is likely due to the more extensive validation during the recording process and to the participants' prior knowledge. A detailed overview of the recorded signs by participant is provided in the appendix in Table A.22.

### 6.2.2 *Evaluation of handshape recognition*

Figure 90 and Figure 91 show the confusion matrices of handshape recognition at the beginning and at the end of sign execution. The macro-$F_1$ score lies, depending on hand and position, between 74.20 to 83.50 %. The dominant hand is recognized significantly more reliably, with an average macro-$F_1$ score of 82.30 %, than the non-dominant hand, with an average of 74.80 %.

The cause of this difference cannot be determined unambiguously on the basis of the available data. One possible explanation lies in the participants' handedness, since the dominant hand, due to its leading role in everyday life, may also be used more skillfully in sign execution.

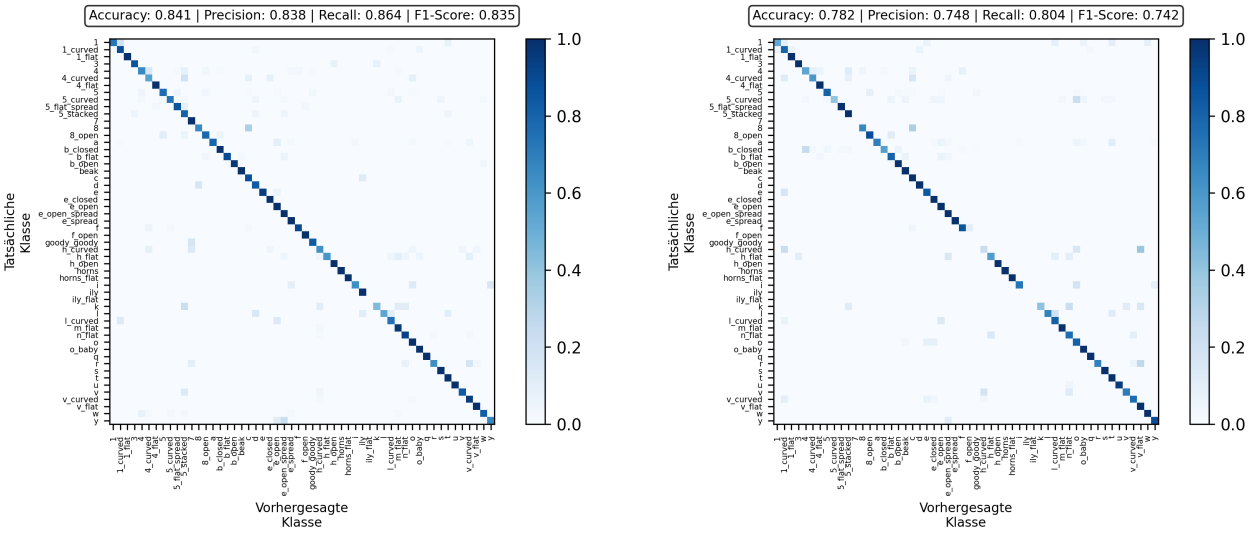



(a) Dominant

(b) Non-dominant

Figure 90: Confusion matrices of handshape recognition at the beginning of sign execution for the dominant and non-dominant hand

The results achieved are, as expected, below the $F_1$ score of 86.40 % measured under idealized conditions, which was determined as part of the final training of handshape recognition. This difference is due in particular to changed boundary conditions during data acquisition, above all the fact that handshape recognition now takes place in the context of dynamic sign executions. Even though only a single frame is still considered and thus, formally, a static classification is performed, the extracted handshape may originate from a transition phase in which the handshape is currently changing. These sometimes marginal variations within sign execution have a direct impact on classification performance, especially since many classes are close to one another and differ only in details.




(a) Dominant

(b) Non-dominant

Figure 91: Confusion matrices of handshape recognition at the end of sign execution for the dominant and non-dominant hand

Another aspect is that the handshapes in *ASL-LEX* represent only an approximation of the actual handshapes, which are not always identical to the annotated handshapes recorded for training.

In addition, during data acquisition the participants must concentrate on the complete sign execution, so that the handshape is only one of several parameters to be performed simultaneously. By contrast, the focus in the previous experiment was exclusively on correct handshape execution, which led to more consistent training data. This aspect affects a large number of parameters, since the recognition approaches were conceived and implemented in isolation from one another and were often trained with data collected specifically for that purpose.

An improvement in the results should actually also be supported by the fact that in the present scenario several handshapes can be considered correct, whereas the original handshape recognition only considered one correct class and was evaluated accordingly. Under this assumption, a higher $F_1$ score would generally be expected.

Another influencing factor arises from the hardware and software configuration used. During the present evaluation, only the *Manus Quantum Metagloves* in conjunction with a newer version of the *Manus Core* SDK were used. Compared with the previously used SDK version, the sensor data delivered by the new version are no longer strictly limited to a fixed value range, but exhibit continuous values that, in extreme cases, can exceed the previously known bounds. This creates value ranges that were not taken into account during model training. Downgrading to the earlier SDK version was not possible, nor could a new data collection followed by retraining of the models be carried out.

Since the vocabulary of the ParaSignRec has been significantly reduced compared with *ASL-LEX*, not all originally defined handshapes need to be classified anymore. Specifically, by excluding the sign CALIFORNIA (multi-morphemic sign), the handshape *Flat ILY* is no longer used for the dominant hand. For the non-dominant hand, the handshape *Goody Goody* is also no longer used, so that the handshape repertoire here comprises only 52 handshapes.

A natural approach would be to reduce the handshape classifier to these handshapes and provide separate models for the dominant and non-dominant hands. This would simplify classification and is likely to lead to better classification results. Moreover, there would be no risk of predicting handshapes that are not considered in the current vocabulary.

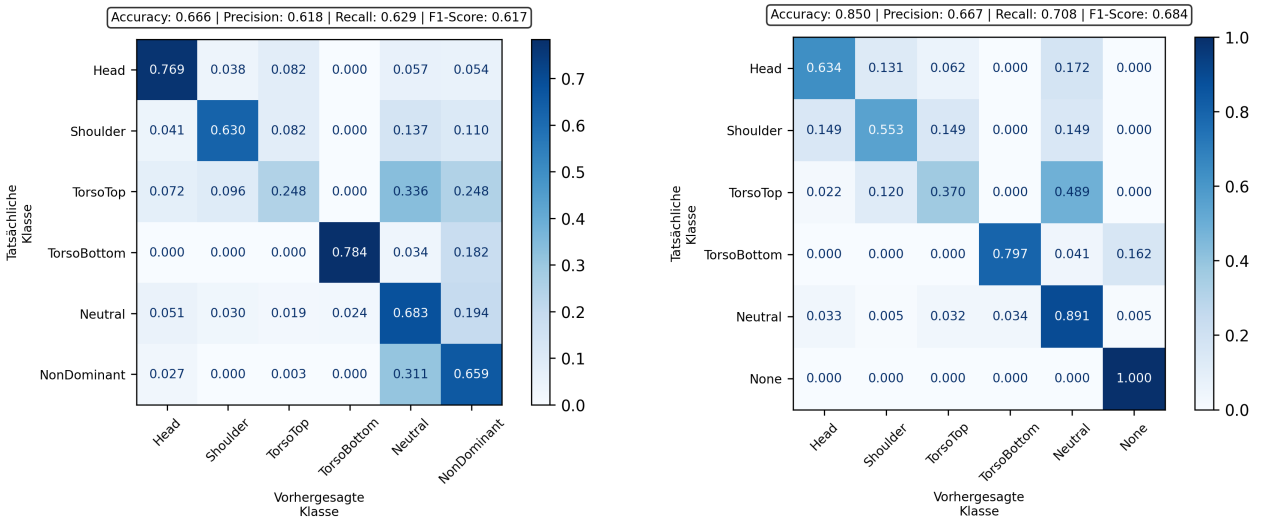
However, since one of the central advantages of our approach lies in its flexible extensibility, the more extensive model with 54 handshapes is retained for both the dominant and non-dominant hand. This makes it possible to add further signs with these 54 handshapes to the vocabulary later without costly retraining of the models. Since, for the prediction of the target sign in the context of meta-classification, only the permissible class with the highest score is considered for each parameter, it is not disadvantageous for the meta-classification if parameters exhibit more classes than are needed in the vocabulary.

The non-dominant hand is either undefined, assumes the same handshape as the dominant hand, or one of seven *Unmarked Handshapes* of *ASL-LEX*. Since these handshapes are also included among the 54 handshapes of the dominant hand, the same class set results for the dominant and non-dominant hand. Therefore, no distinction is made here either by using separate models.

Finally, it should be noted that the achieved results could be further improved by several measures. A key factor is the hardware used: more precise sensors with higher resolution, separate acquisition of individual finger phalanges, and lower susceptibility to electromagnetic interference would increase the quality of the input data and thus positively affect classification performance. In addition, retraining the models with more realistic handshapes, extracted directly from complete sign executions rather than based on static hand poses, appears sensible. Expanding the training dataset would further increase the robustness of handshape recognition, or at least improve generalizability. Furthermore, it would need to be evaluated whether it would be advantageous for classification performance to include again the features previously excluded from classification.

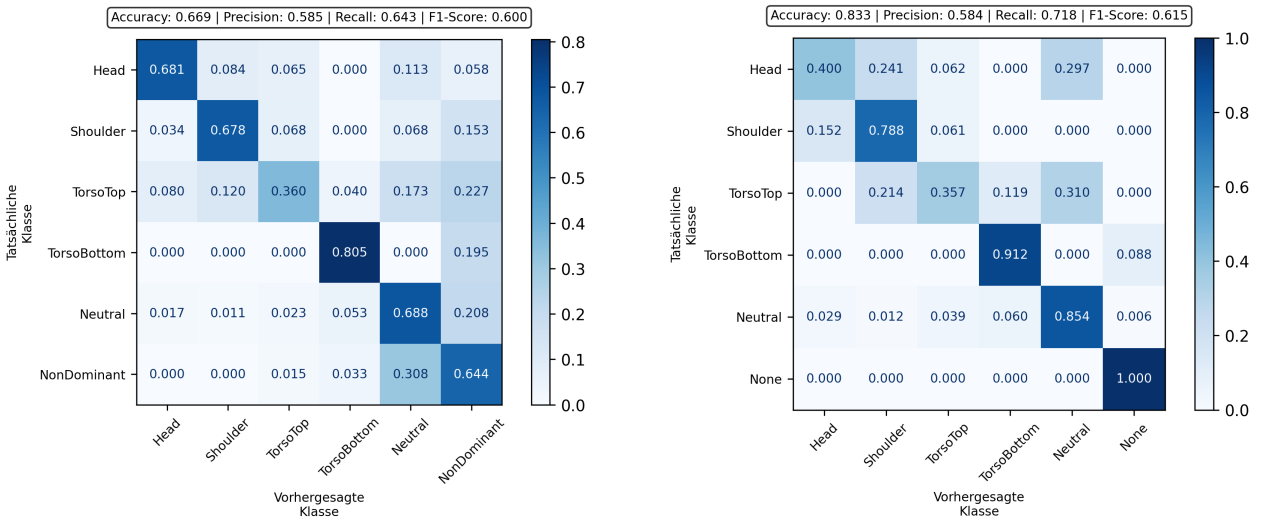
### 6.2.3 *Evaluation of sign execution location recognition*

Figure 92 and Figure 93 show the confusion matrices of sign execution location recognition at the beginning and at the end of sign execution. The achieved macro-$F_1$ score varies in the range of 60 to 68.40 % and represents the system's ability to differentiate six execution locations for the dominant and non-dominant hand. The dominant hand is recognized somewhat less reliably, with an average $F_1$ score of 60.85 %, than the non-dominant hand, with an average of 64.95 %.




(a) Dominant (b) Non-dominant

Figure 92: Confusion matrices of sign execution location recognition at the beginning of sign execution for the dominant and non-dominant hand




(a) Dominant (b) Non-dominant

Figure 93: Confusion matrices of sign execution location recognition at the end of sign execution for the dominant and non-dominant hand

This difference of 4.10 percentage points results in particular from the slightly different class composition. Both hands consider the execution locations *Head*, *Shoulder*,

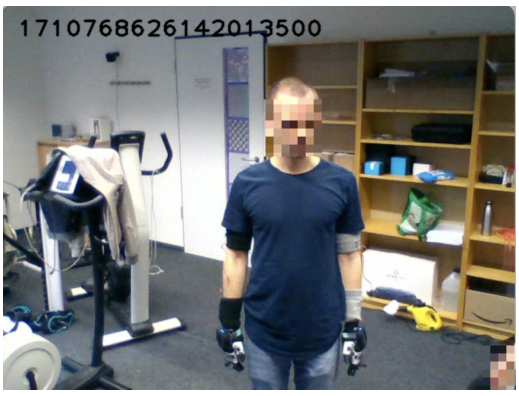
*TorsoTop*, *TorsoBottom*, and *Neutral*, but differ in the additional classes *NonDominant* for the dominant hand and *None* for the non-dominant hand.

The class *NonDominant* describes the execution location of the dominant hand in relation to the non-dominant arm, that is, situations in which the dominant hand is located at or in the immediate vicinity of the non-dominant arm. This class is the only dynamic variant of the execution location, as it results from the relative position of both arms. Accordingly, overlaps occur here with static execution locations such as *TorsoTop*, *TorsoBottom*, and *Neutral*, which leads to an increased number of misclassifications.

The class *None*, by contrast, is used for one-handed signs when the non-dominant hand is outside the signing space. The results achieved here are error-free, since there are no competing classes with which confusions could occur.

In the context of recording and evaluating static reference positions in Section 5.3.5, an $F_1$ score of 65 % was achieved for the dominant hand, which roughly corresponds to the results obtained in the present evaluation for the non-dominant hand. The position deviations in the upper single-digit range analyzed there in detail, as well as differing body proportions of the users, have a noticeable effect on classification performance. This effect is particularly relevant for the class *NonDominant*, since position errors of both hands overlap here and thus add up.

In addition, this effect is reinforced by the fact that an incorrect I-pose calibration was performed for some recordings. In these cases, the palm was not aligned perpendicular to the floor during calibration, but showed rotations around various axes. Figure 94 shows two exemplary examples. Incorrect calibration leads to a constant systematic deviation between the actual and assumed palm orientation and thus has a direct effect on position determination.




(a) Incorrect I-pose calibration with wrists rotated about the longitudinal axis (incorrect pronation/supination)


(b) Incorrect I-pose calibration with wrists rotated about the transverse axis (incorrect flexion/extension) and vertical axis (incorrect ulnar/radial abduction)

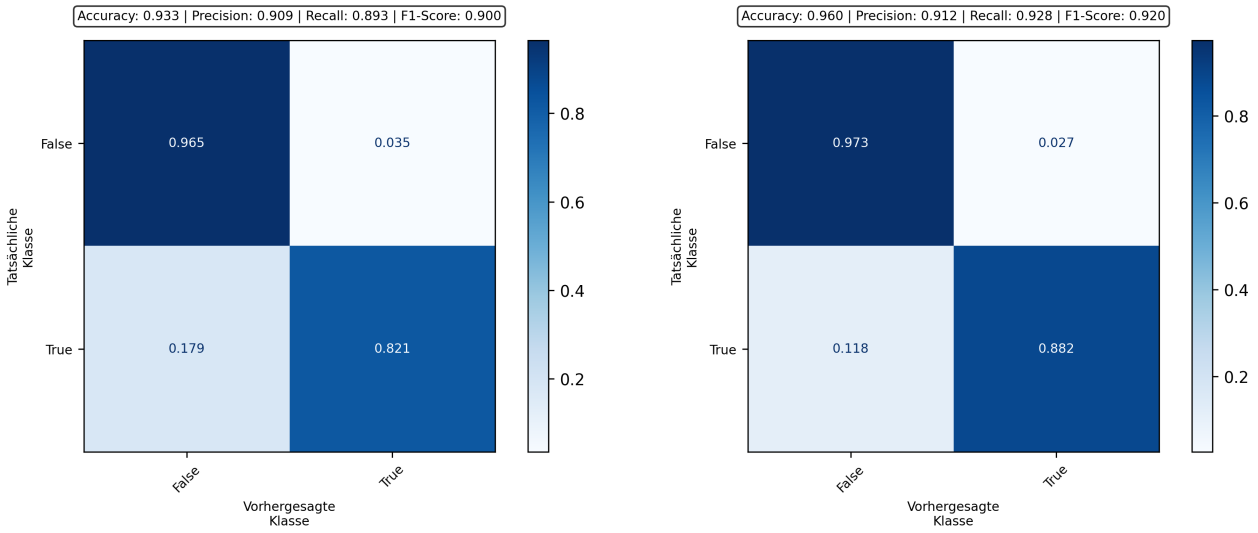
Figure 94: Two examples of incorrectly performed I-pose calibrations

Finally, it should be noted that the results of sign execution location recognition could be improved by several measures. One possibility is the use of higher-quality IMUs with industrial standards, such as the *Xsens MTi-300*, which has a significantly

lower angular error deviation than the *Bosch BNO055* we used [128]. While this showed a mean angular deviation of about 4.61°, the *Xsens MTi-300* achieved values of up to 0.54°. In our example from Section 5.3.2, the position error would thus be reduced from 6.82 cm to about 0.80 cm. Regardless of the hardware, however, the correct execution of the I-pose calibration as well as a new recording of the reference positions using both a T-pose and an I-pose constitutes the simplest and at the same time most effective measure for improving the results.

### 6.2.4 *Evaluation of flexion and spread change recognition*

Figure 95 to Figure 98 show the confusion matrices of flexion change recognition (FLEXION CHANGE) for the individual fingers of both hands. The achieved macro-$F_1$ scores lie in the range of 88.90 to 93.10 %. For the dominant hand, an average $F_1$ score of 90.43 % is obtained, while for the non-dominant hand a slightly higher value of 92.10 % is achieved.

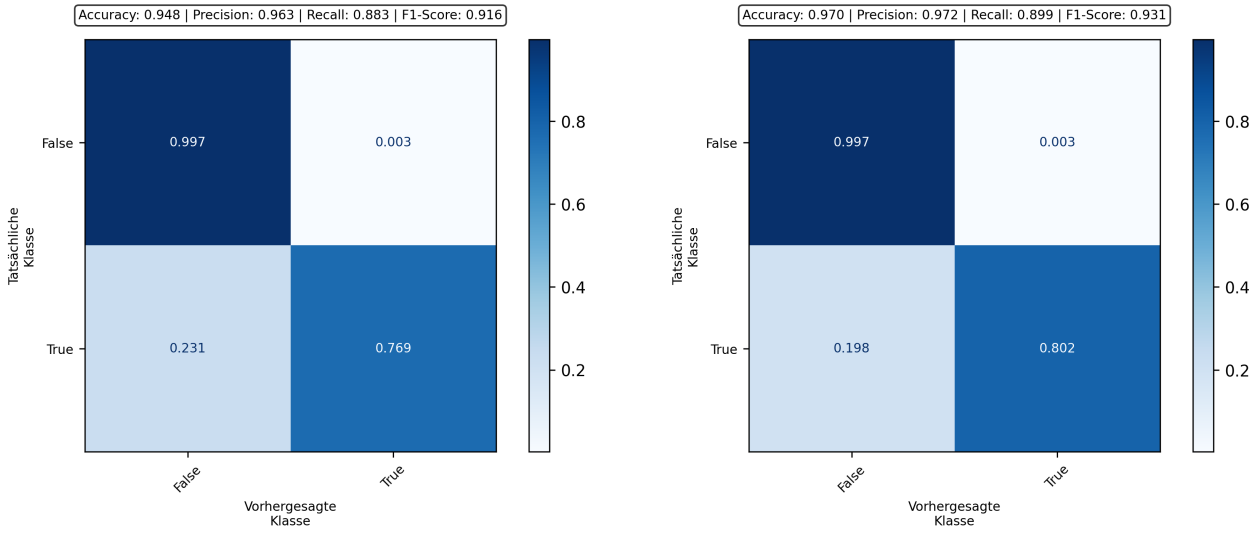



(a) Dominant

(b) Non-dominant

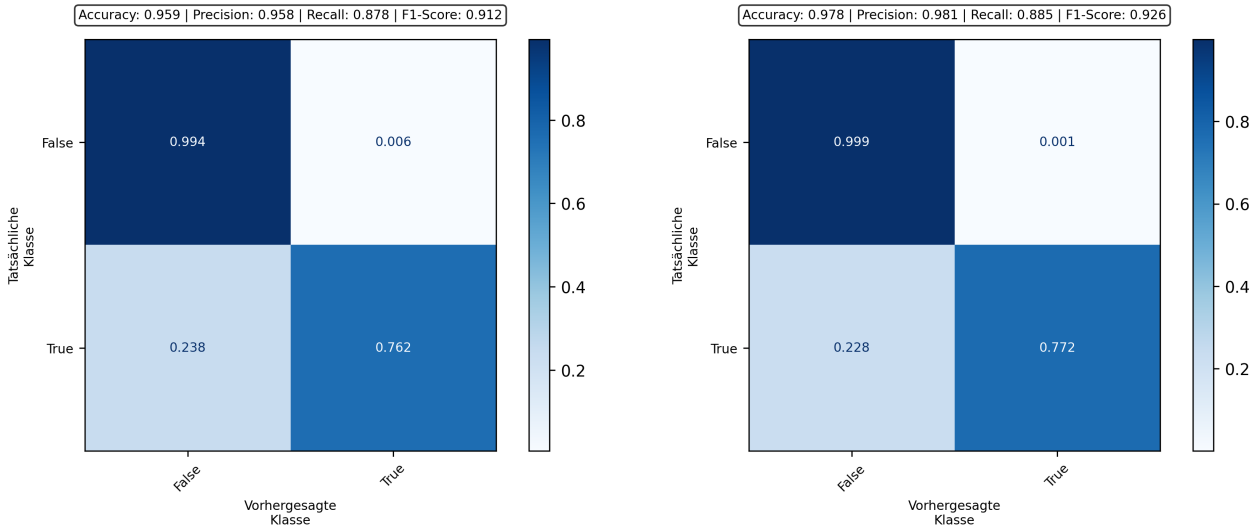
Figure 95: Confusion matrices of flexion change recognition (FLEXION CHANGE) for the index finger of the dominant and non-dominant hand

This difference is mainly attributable to the different class distribution between the dominant and non-dominant hand. Correct recognition of the state "no movement" succeeds with 98.89 % considerably more reliably than recognition of actual movement with 78.21 %. Since the proportion of motionless flexion parameters on the non-dominant hand with 92.98 % is significantly higher than on the dominant hand with 84.71 %, this effect positively influences the achieved results for the non-dominant hand.




(a) Dominant (b) Non-dominant

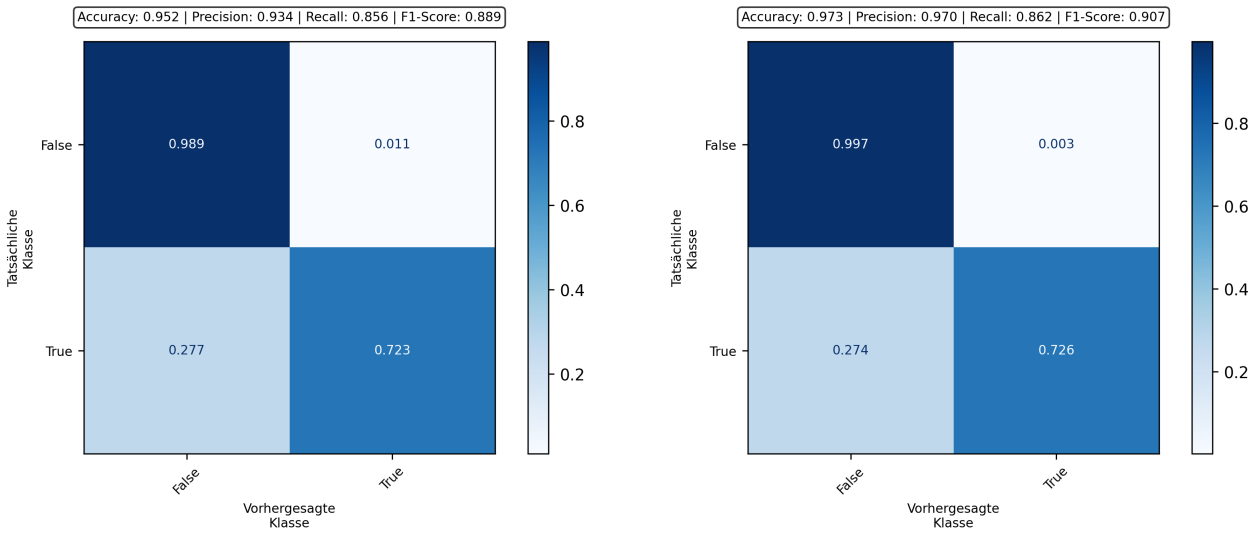
Figure 96: Confusion matrices of flexion change recognition (Flexion Change) for the middle finger of the dominant and non-dominant hand




(a) Dominant (b) Non-dominant

Figure 97: Confusion matrices of flexion change recognition (Flexion Change) for the ring finger of the dominant and non-dominant hand

At the level of individual fingers, the middle finger is recognized most reliably, with an average $F_1$ score of 92.35 %, whereas the little finger shows the lowest recognition performance, with an average of 89.80 %.




(a) Dominant

(b) Non-dominant

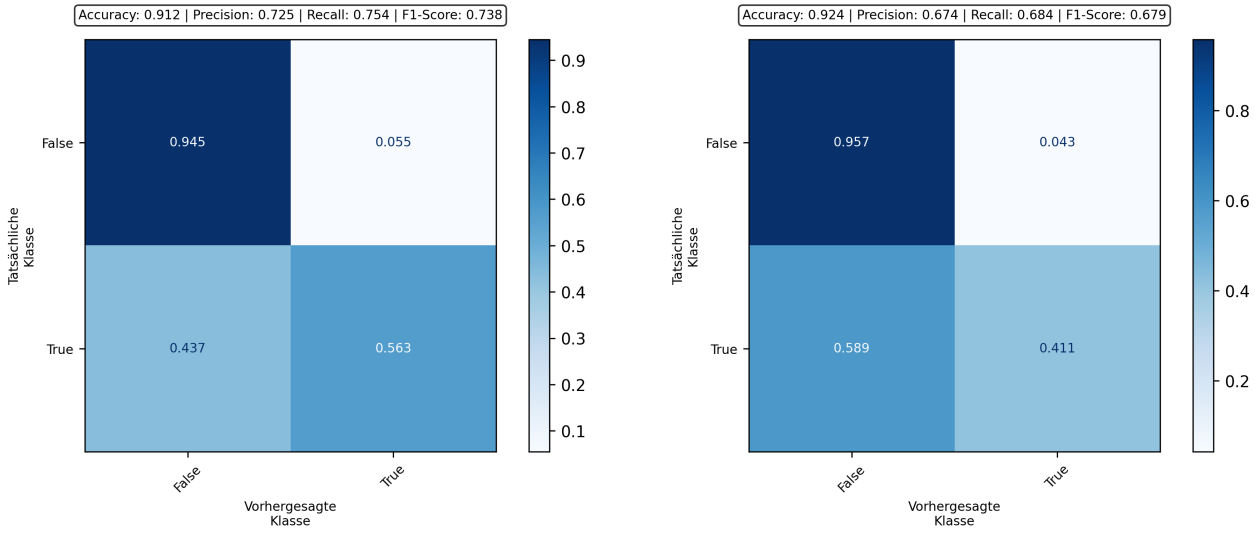
Figure 98: Confusion matrices of flexion change recognition (FLEXION CHANGE) for the little finger of the dominant and non-dominant hand

In the conception of the approach in Section 5.4.1, an average $F_1$ score of 98.55 % for the dominant hand was achieved under idealized conditions, with all fingers exhibiting comparably high values. The decrease of about 8.13 percentage points observed in the present evaluation scenario may be related to the fact that the original threshold values were determined with the *Manus Prime X* data gloves and the earlier SDK version with fixed value ranges.

Figure 99 and Figure 100 show the confusion matrices of spread change recognition (SPREAD CHANGE) for the respective finger pairs of both hands. The achieved macro-$F_1$ scores lie in the range of 67.90 to 82.70 % and thus exhibit a considerably larger spread than in flexion change recognition. For the dominant hand, an average $F_1$ score of 78.25 % is obtained, while for the non-dominant hand a lower value of 73.60 % is achieved.

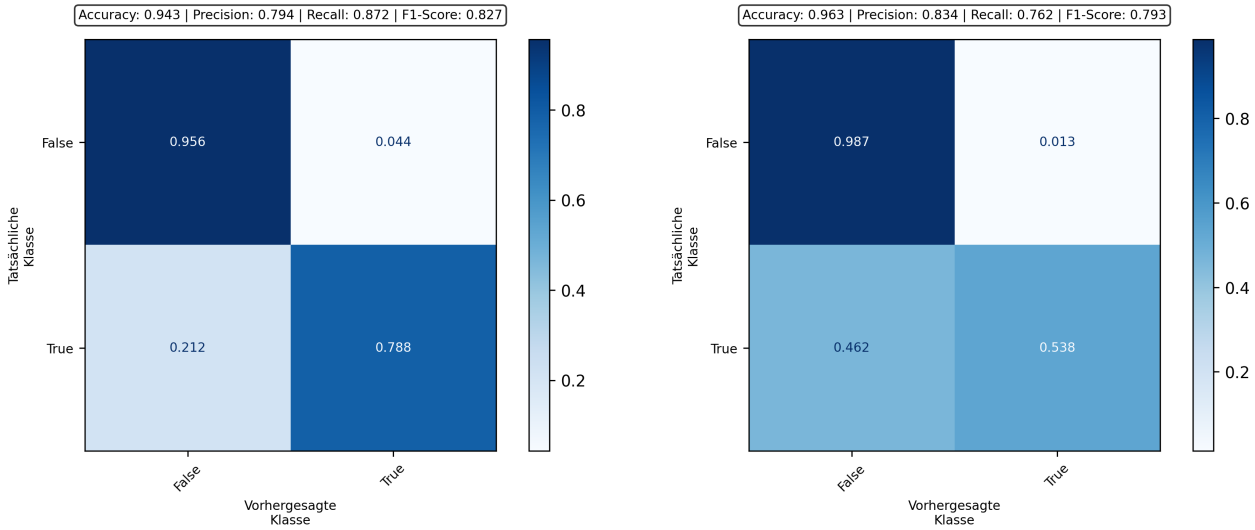
In the conception of the approach in Section 5.4.1, an $F_1$ score of 94.70 % for the finger pair index finger/middle finger and 89.20 % for the finger pair ring finger/little finger was achieved under idealized conditions, i.e., 20.90 percentage points and 6.50 percentage points less than in the present evaluation scenario for the dominant hand.

Notably, the recognition performance for the correct detection of the moved class is low, with an average of only 57.50 %. This figure is even only 50.40 % if the finger pair ring finger/little finger of the dominant hand is not taken into account, for which recognition performs significantly better at 78.80 %. Analogous to flexion change, it is also evident here that recognition of the state “no movement” succeeds very reliably,




(a) Dominant (b) Non-dominant

Figure 99: Confusion matrices of spread change recognition (Spread Change) for the finger pair index finger/middle finger of the dominant and non-dominant hand




(a) Dominant (b) Non-dominant

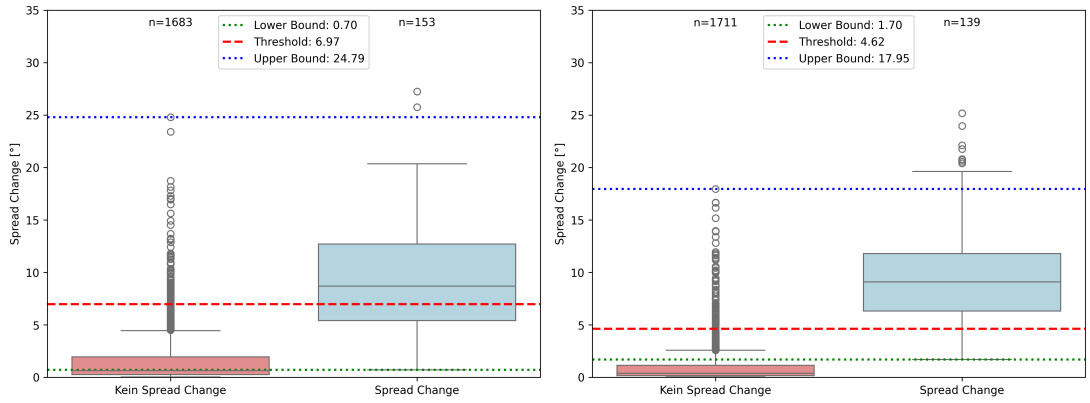
Figure 100: Confusion matrices of spread change recognition (Spread Change) for the finger pair ring finger/little finger of the dominant and non-dominant hand

with an average of 96.13 %, whereas distinguishing actual spread changes is much more error-prone.

To analyze the causes of the low true-positive rate more closely, Figure 101 shows the distribution of spread value differences for the recorded data of the dominant hand. Regardless of whether the finger pair is index finger/middle finger or ring finger/little finger, the two classes can in principle still be well separated: with the threshold values shown, theoretical maximum $F_1$ scores of 75.69 % for index/middle finger and 84.81 % for ring/little finger could be achieved.

Since the thresholds affect only the results of the meta-classification, the focus here is on the threshold values. Compared with the threshold values from training of the approach, these were originally 7.67° and 6.09° for the dominant hand and, in the present evaluation, amount to 6.97° and 4.62°. For the finger pair index finger/middle finger, the differences are only small, which is why the deviation between the optimal and achieved result at 1.89 percentage points is comparatively small and more likely attributable to data quality than to the choice of thresholds. The same applies to the finger pair ring finger/little finger, with a difference of 2.11 percentage points below the theoretically optimal value. The comparison with the non-dominant hand yields similar results.

The causes of the lower sensor data quality are to be found in particular in simultaneous arm movements or pronounced flexion changes. In such cases, measurement noise and coupling effects between the fingers have a strong impact on the spread values. This effect may have been further amplified by the switch to the SDK version with continuous sensor data, but it primarily points to hardware-related limitations.




(a) Index finger/middle finger

(b) Ring finger/little finger

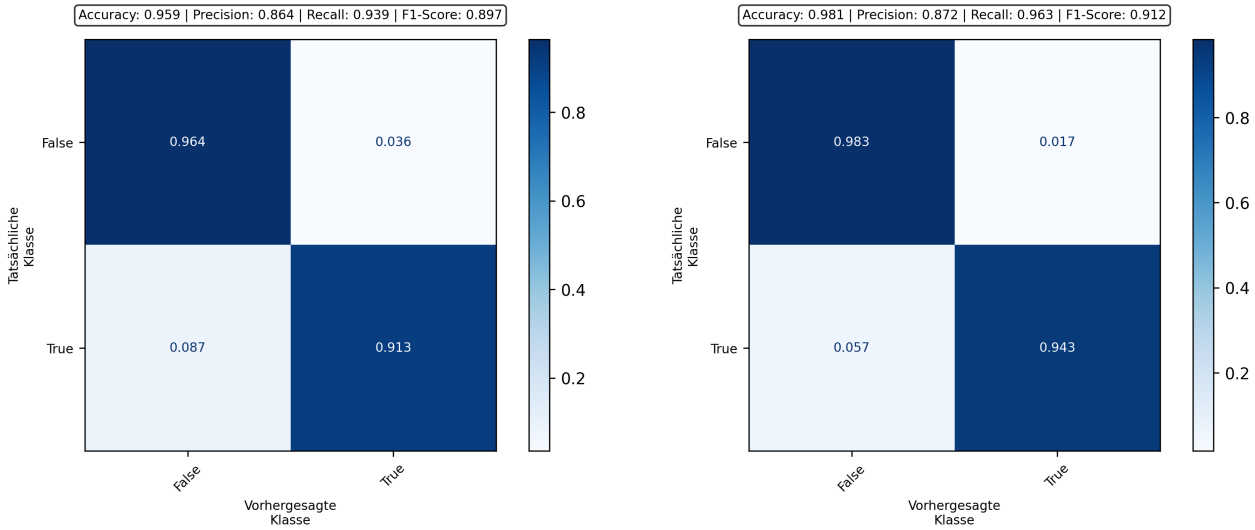
Figure 101: Distribution of the spread value differences (°) for the dominant hand across the two classes, with threshold and boundary values plotted

In conclusion, both the results of flexion change recognition and those of spread change recognition could be further improved, in particular through better hardware. One conceivable approach would be the use of additional specialized sensors, such as separate flex sensors on the individual finger phalanges and between the respective finger pairs. This would make it possible to capture the respective finger positions accurately, which had previously been estimated using an inverse kinematics approach. Finger spread could then also be determined more precisely and independently of

external influences such as movements. In addition, an adaptive classification or regression approach could be developed that determines the threshold as a function of hand or arm movement, for example by using higher threshold values during rapid movements. However, such an approach would require a considerably larger data basis as well as a variety of differently executed signs.

### 6.2.5 *Evaluation of wrist rotation recognition*

Figure 102 shows the confusion matrices of wrist rotation recognition (ULNAR ROTATION) for both hands. The achieved macro-$F_1$ scores lie in the range of 89.70 to 91.20 % and are therefore at a similar level. This also applies to the true-positive rate and the false-positive rate, which, with average values of 92.80 % and 97.35 % respectively, are both very high and comparable to the values achieved during training in Section 5.4.1. For the dominant hand, an $F_1$ score of 89.70 % is obtained, while for the non-dominant hand a slightly higher value of 91.20 % is obtained.



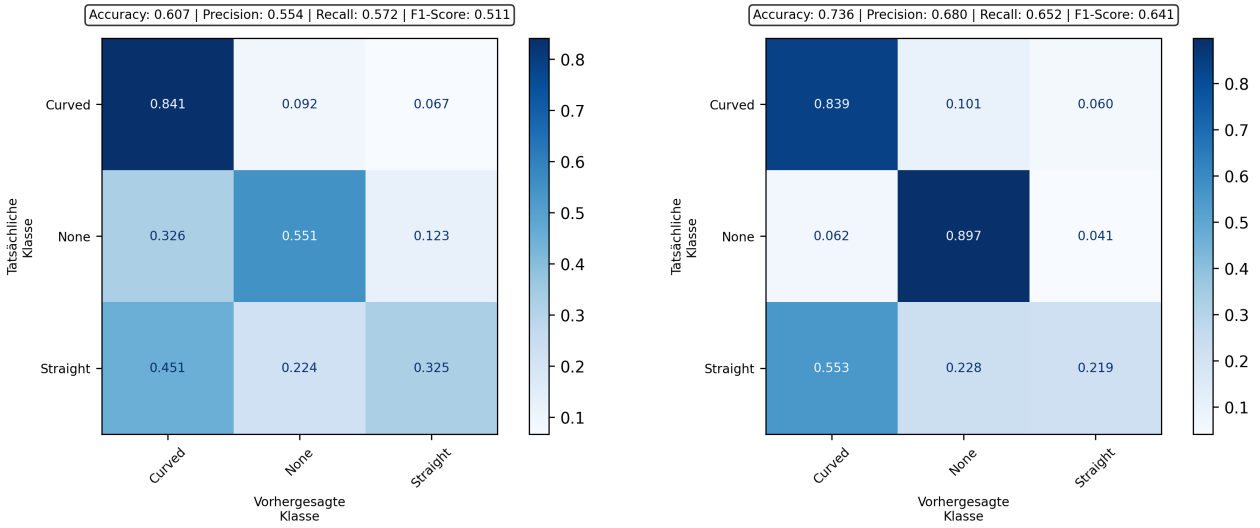

(a) Dominant (b) Non-dominant

Figure 102: Confusion matrices of wrist rotation recognition (ULNAR ROTATION) for the dominant and non-dominant hand

### 6.2.6 *Evaluation of path movement recognition*

Figure 103 shows the confusion matrices of path movement recognition (PATH MOVEMENT) for both hands. The achieved macro-$F_1$ scores for distinguishing three movement classes lie in the range of 51.10 to 64.10 %, below a level at which a positive influence on meta-classification would be expected. Nevertheless, we deliberately de-

cided to retain the parameter in the system, as it is relevant for sign recognition from a prospective point of view.




(a) Dominant

(b) Non-dominant

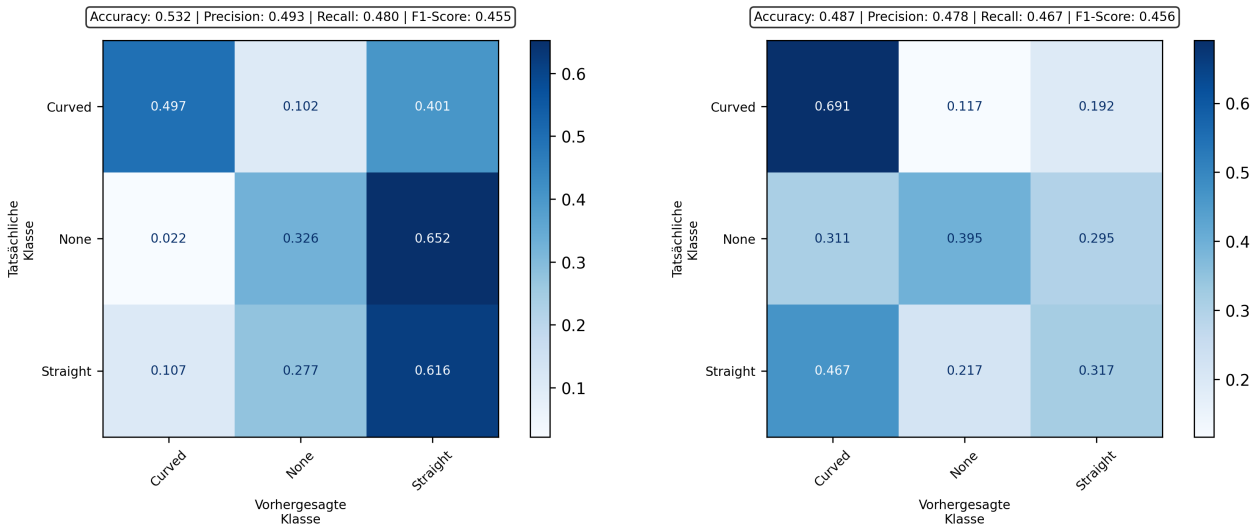
Figure 103: Confusion matrices of path movement recognition for the dominant and non-dominant hand and RF classifier

In particular, correct classification of straight movements proves problematic. With recognition rates of 32.50 % for the dominant hand and 21.90 % for the non-dominant hand, these values are below the *baseline accuracy* of 33.33 %, which results from a purely random class assignment when there are three equally probable classes. By contrast, recognition of curved movements succeeds much more reliably, at 84.10 % and 83.90 %, respectively. Classification of absent movement performs very well for the non-dominant hand, with 89.70 %, whereas it is considerably worse for the dominant hand, with only 55.10 %.

Classification was performed exclusively using the RF classifier employed in VL2, since the VL2 classifier achieved an even lower performance of about 45.50 to 45.60 % macro-$F_1$ score, as can be seen in Figure 104.

Compared with the results of classifier training given in Section 5.4.2, the values achieved here are significantly worse by 24.80 to 37.80 percentage points. Even in the training scenario, the recognition of straight movements posed the greatest challenge, which is further exacerbated under realistic evaluation conditions.

The reasons for this lie primarily in the definition and characteristics of the movement classes used. On the one hand, many signs exhibit only very short or weakly pronounced path movements, making it difficult to distinguish straight and curved movements in the sensor data. On the other hand, unintended co-movements of the hand often occur even when it is formally annotated as stationary. In addition, many




(a) Dominant

(b) Non-dominant

Figure 104: Confusion matrices of path movement recognition for the dominant and non-dominant hand and VL2 classifier

movements result from rotational joint movements of the shoulder or elbow, which means that some straight movements nevertheless have a curved trajectory.

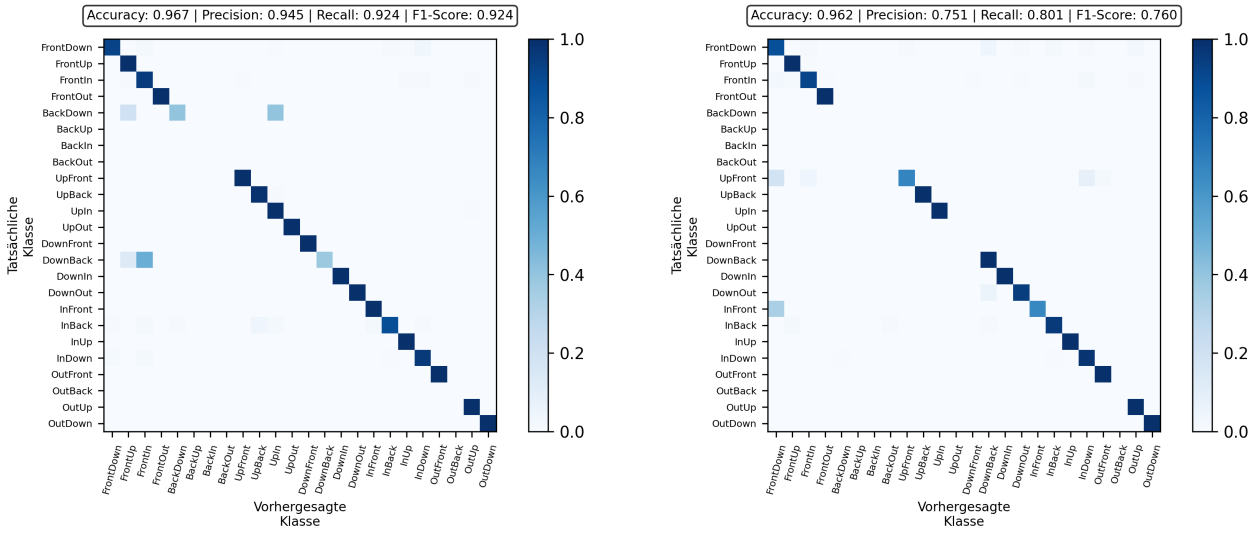
An additional complicating factor is the annotation of wrist movements, which are partially defined as Ulnar Rotation but are also partially included in path movements, although this is not followed consistently. The underlying approach, however, currently considers only arm movements in its present form. Furthermore, repetitive or composite movements lead to a strong overlap of several movement types, further reducing the discriminative power of the classes.

Several approaches are suitable for improving path movement recognition. In principle, redesigning the movement classes, for example by a finer subdivision or by separating short and pronounced movements, would potentially make the model more robust. A natural extension in this context would be the explicit consideration of movement direction, as this provides additional discriminative information. Regardless of this, it is to be expected that a significant improvement in results would come from retraining the models with a larger and more diverse data basis in order to adequately account for the high variability of real sign movements.

### 6.2.7 *Evaluation of palm orientation recognition*

Figure 105 and Figure 106 show the confusion matrices of palm orientation recognition at the beginning and at the end of sign execution for distinguishing 24 different palm orientations. The achieved macro-$F_1$ score lies in the range of 92.50 to 94.90 %, although the non-dominant hand at the beginning of sign execution with 76 % constitutes an

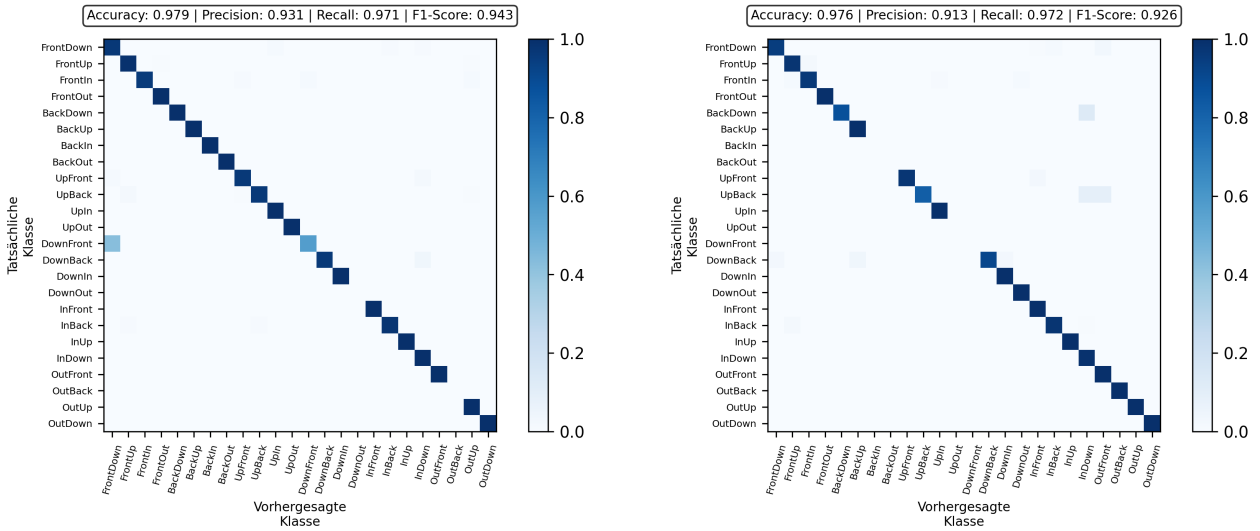
exception, even though the associated confusion matrix visually exhibits an almost perfect diagonal.




(a) Dominant

(b) Non-dominant

Figure 105: Confusion matrices of palm orientation recognition at the beginning of sign execution for the dominant and non-dominant hand




(a) Dominant

(b) Non-dominant

Figure 106: Confusion matrices of palm orientation recognition at the end of sign execution for the dominant and non-dominant hand

This apparent discrepancy is due to the handling of multi-valued parameter classes. For palm orientation, several orientations per instance are permissible. In the evaluation considered here, however, in the case of success only the predicted class is counted as correct, while all other permissible orientations are counted as TNs, i.e., "not applicable". As a result, individual classes can exhibit a strongly reduced or even zero-valued $F_1$ score despite overall very good separability.

For better illustration, Figure 107 shows additional OvA confusion matrices for all 24 classes of palm orientation, based on the same data as the aggregated confusion matrix, but from a binary perspective per class.

The classes *BackUp*, *BackIn*, *UpOut*, and *OutBack* contain only TNs, since no instances of these classes are included in the dataset or, for the reasons mentioned above, they are counted as TNs. Accordingly, they are not included in the confusion matrix from Figure 105b, since the macro averaging is determined only on the basis of classes occurring in the data. The classes *OutFront* and *OutUp* do show below-average classification results; however, the comparatively lower classification result of the aggregated confusion matrix is primarily attributable to the classes *BackDown*, *BackOut*, and *DownFront*. Since these classes contain only one or two instances each, individual misclassifications lead to an $F_1$ score of 0 %. With only 20 considered classes that actually influence the aggregated macro-$F_1$ score, the underperformance of three or five classes affects the overall result accordingly strongly.

In addition, the classification performance of palm orientation is also impaired by the incorrect I-pose calibrations, as already shown in Figure 94. Incorrect alignment of the palms during calibration leads to a systematic shift in the measured orientations and directly affects the assignment to the discrete orientation classes. Since the underlying classification models otherwise exhibit high separability, correct execution of the I-pose calibration is by far the most effective measure for further improving the results. Apart from that, palm orientation recognition overall proves to be very robust and well suited for downstream meta-classification.

## 6.3 EVALUATION OF META-CLASSIFICATION

The evaluation of the meta-classification corresponds to the evaluation of the entire *ParaSignRec* approach, since it combines the results of all preceding classifiers. The aim of the meta-classification is to generate an ordered list of possible signs on the basis of these individual results and to assign each sign a probability indicating how likely it is to correspond to the performed sign.

To make these predictions as precise as possible, the weights according to which the individual parameter-based recognition approaches enter into the calculation of the resulting probability score are first optimized.

### 6.3.1 *Optimization of the class weights*

To optimize the class weights, *List 1* of the three available lists of sign recordings from Section 6.1.4 is first evaluated using our unweighted meta-classification, i.e.,

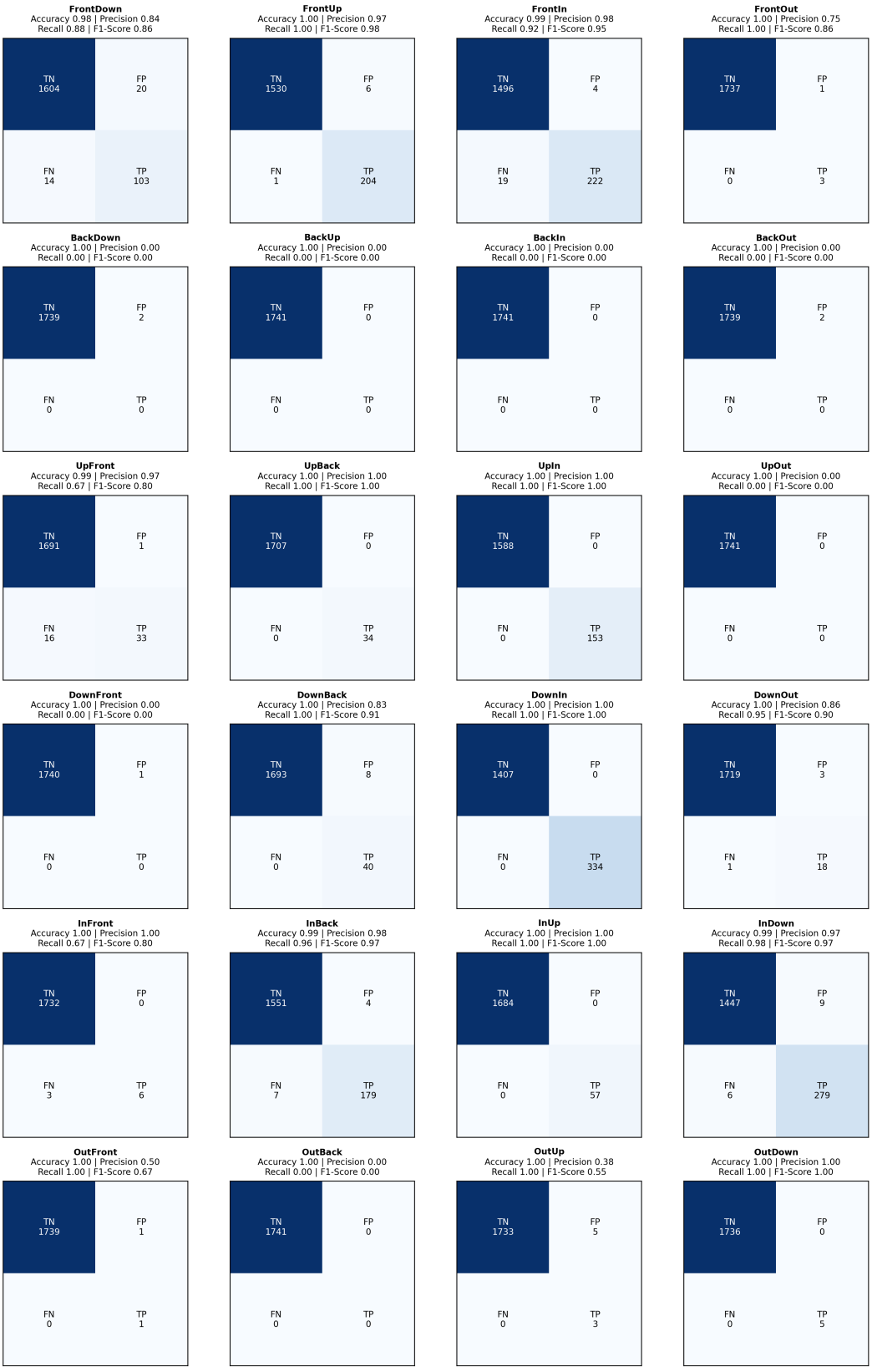



Figure 107: OvA confusion matrices for all 24 classes of palm orientation for the non-dominant hand at the beginning of sign execution

all parameter-based recognition approaches are taken into account with an identical weight of 1. Under this assumption, the meta-classification achieves a Top-1 accuracy of 73.90 %. The normalized scores per sign with average rank and unweighted, average score are shown in Figure 108.

The scores of the individual parameter-based recognition approaches were normalized for presentation in order to ensure comparability across all parameters. The individual parameters have different scales as well as different ranges between minimum and maximum score.

The aggregated score of a sign recording is computed as the weighted arithmetic mean of the corresponding parameter scores of the true class (cf. Equation 8). If a parameter has multiple true classes, the highest score among these true classes is used. Without prior normalization, parameters with a larger range would dominate the score calculation.

To avoid this, the individual parameter scores $p$ for each sign recording are normalized to the value range $[0, 1]$ using $\tilde{p} = \frac{p - p_{worst}}{p_{best} - p_{worst}}$, where the former worst score is then 0 and the best score is 1. Through this normalization, all parameters contribute equally to the calculation of the aggregated score, regardless of their original scale.

Building on this reference value, a weighted optimization of the individual parameter pairs, consisting of dominant and non-dominant hand, is carried out using a three-stage grid search in which all weighting combinations of all parameter pairs are iterated. In the first stage, a coarse search with the values {0.7, 1.0, 1.3} is performed to identify a promising value range. This is followed by a second stage with a finer search in the range of $\pm 0.15$ around the respective best value. In the third stage, a further refinement is then carried out with a search range of $\pm 0.05$.

This procedure selects the weights so that they cover the range from 0.5 to 1.5 in steps of 0.05. This range was deliberately chosen in order, on the one hand, to ensure that the influence of a parameter is at most halved in the worst case and thus no relevant information is completely suppressed. On the other hand, the maximum influence of a parameter is limited so that it is at most three times the influence of the weakest parameter. In this way, individual parameters are prevented from dominating the meta-classification.

The choice of a multi-stage grid search is due to the high number of possible parameter combinations. Even though the dominant and non-dominant hand are assigned the same weight for each parameter, there are a total of 14 parameter pairs to be weighted and three possible values per parameter pair, resulting in $3^{14} = 4.782.969$ combinations per stage of the grid search, i.e., 14,348,907 combinations in total. A direct grid search over the entire range with the same resolution would lead to $3^{21} \approx 1.05 \times 10^{10}$ combinations – i.e., 729 times as many as the three-stage grid search – and is therefore not computationally practical. The chosen multi-stage procedure greatly reduces the search complexity, but it comes with the risk of finding only a local optimum and not necessarily determining the globally best weighting.

After completing the weight optimization, the meta-classification is performed again on the same sign list, the results of which are shown in Figure 109. The Top-1

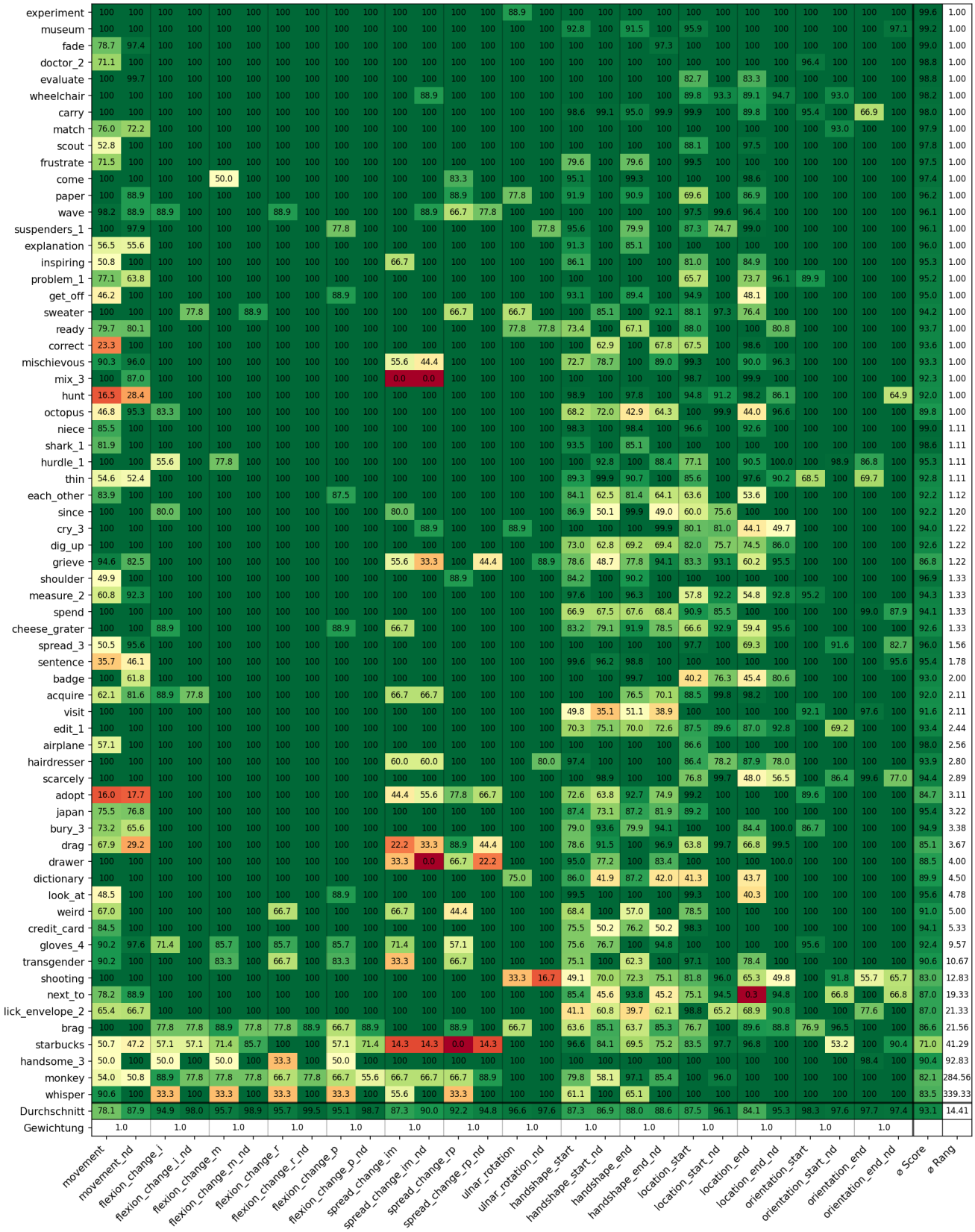



Figure 108: Parameter heatmap of the normalized scores per sign with average rank and unweighted, average score

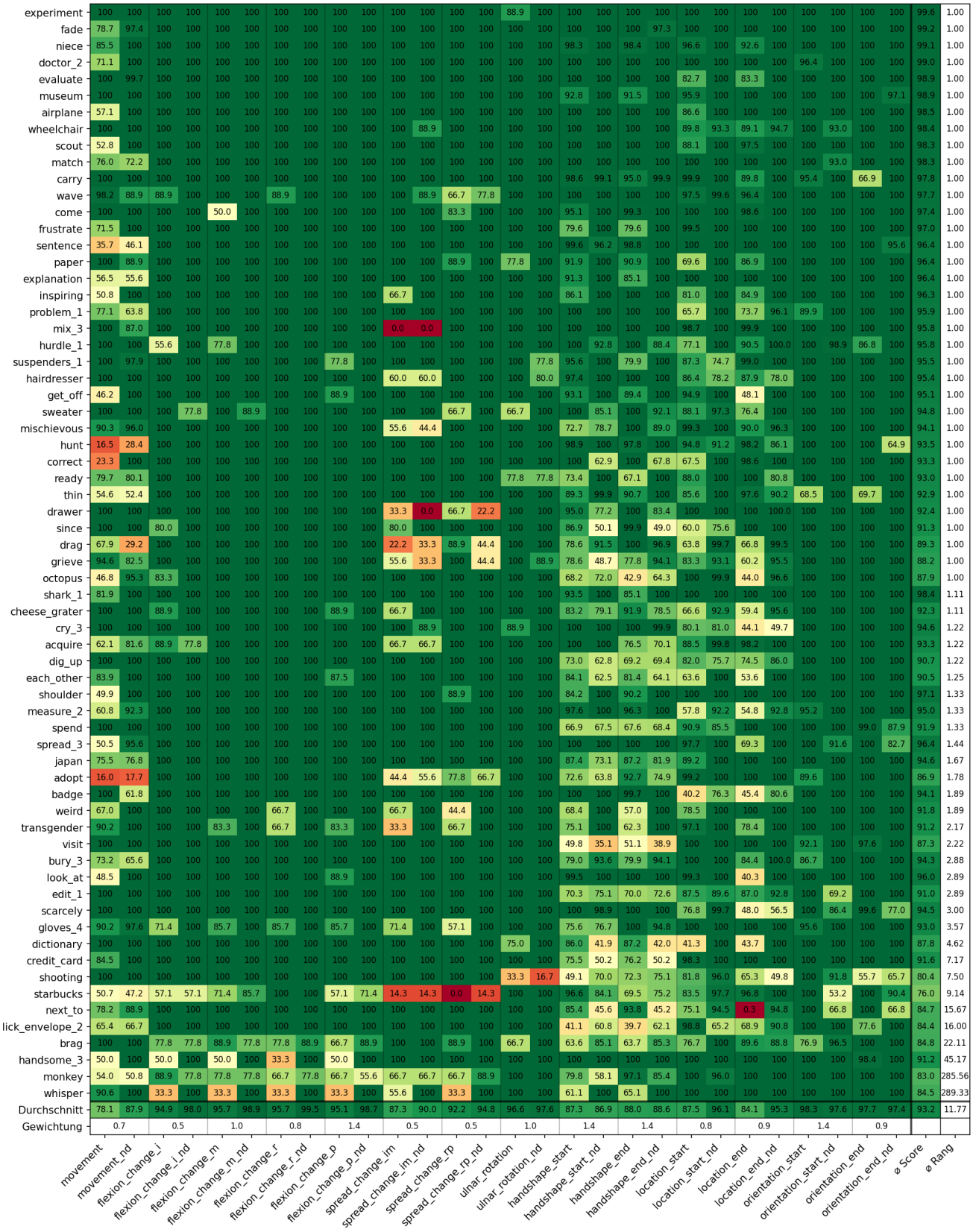



Figure 109: Parameter heatmap of the normalized scores per sign with average rank and weighted, average score

accuracy can thereby be increased to 81.85 %, which represents an improvement of 7.95 percentage points compared to the unweighted initial configuration.

*Evaluation of the class weights*

For a better interpretation of the optimized class weights, we now analyze, on the one hand, how the weighting changes when an alternative sign list is used and to what extent the resulting weights differ. On the other hand, we examine the influence of a particular weighting on other lists of sign recordings. Table 27 shows the determined weights for all three lists as well as for the optimization over all sign recordings, each together with the resulting classification results of the meta-classification.

Table 27: Parameter weights for the meta-classification determined by means of grid search using different selections of test utterances and the highest accuracies achieved for the entire vocabulary of the *ParaSignRec*

| **Weight** | **All** | **List 1** | **List 2** | **List 3** | **Min.** | **Max.** | **MW** | **Std. dev.** |
|---|---|---|---|---|---|---|---|---|
| flexion_change_i | 0.50 | 0.50 | 1 | 0.50 | 0.50 | 1 | 0.62 | 0.22 |
| flexion_change_m | 0.55 | 1 | 0.70 | 0.60 | 0.55 | 1 | 0.71 | 0.17 |
| flexion_change_p | 1.10 | 1.40 | 1 | 1 | 1 | 1.40 | 1.12 | 0.16 |
| flexion_change_r | 0.70 | 0.75 | 1 | 1 | 0.70 | 1 | 0.86 | 0.14 |
| handshape_end | 1.50 | 1.45 | 1.50 | 1.10 | 1.10 | 1.50 | 1.39 | 0.17 |
| handshape_start | 1.40 | 1.40 | 1.40 | 1.40 | 1.40 | 1.40 | 1.40 | 0 |
| location_end | 0.55 | 0.90 | 0.90 | 0.85 | 0.55 | 0.90 | 0.80 | 0.15 |
| location_start | 1.15 | 0.75 | 1.10 | 1.40 | 0.75 | 1.40 | 1.10 | 0.23 |
| movement | 0.60 | 0.70 | 0.65 | 0.65 | 0.60 | 0.70 | 0.65 | 0.04 |
| orientation_end | 1.40 | 0.95 | 1.40 | 1.25 | 0.95 | 1.40 | 1.25 | 0.18 |
| orientation_start | 1.35 | 1.40 | 0.90 | 0.85 | 0.85 | 1.40 | 1.12 | 0.25 |
| spread_change_im | 0.50 | 0.50 | 0.85 | 0.55 | 0.50 | 0.85 | 0.60 | 0.15 |
| spread_change_rp | 0.50 | 0.50 | 0.90 | 1 | 0.50 | 1 | 0.72 | 0.23 |
| ulnar_rotation | 0.60 | 1 | 0.85 | 1 | 0.60 | 1 | 0.86 | 0.16 |
| **Accuracy** | **81.85 %** | **81.85 %** | **80.64 %** | **83.68 %** | **80.64 %** | **83.68 %** | **82.01 %** | **1.09 %** |

The highest optimized Top-1 accuracy is achieved with *List 3* and amounts to 83.68 %. *List 2* achieves the lowest result at 80.64 %, but overall remains close to the other results. Optimization over all sign recordings yields a Top-1 accuracy of 81.85 % and is therefore at a comparable level to the optimization based on *List 1*.

Considering the weights of the individual parameter pairs, consisting of dominant and non-dominant hand, an average weight of 0.94 results, which is close to the default weight of 1. The mean range between minimum and maximum weight per parameter pair is 0.39, with the largest range of 0.65 occurring for the recognition of the execution location at the beginning of sign execution. This parameter pair therefore shows the strongest dependence on the choice of sign list, with an below-average relevance of 0.75 for *List 1* up to a very high relevance of 1.40 for *List 3*.

The parameter pairs for handshape recognition at the beginning of sign execution as well as the path movement exhibit the smallest differences between the four weighting

variants, with ranges of 0 and 0.10, respectively. This indicates that handshape consistently has a very high influence on the meta-classification, whereas path movement has only low relevance regardless of the list used.

Table 28: Parameter weights for the meta-classification determined via grid search using different selections of test gestures, grouped by detection approach

| Weight | All | List 1 | List 2 | List 3 | Min. | Max. | MW | Std. dev. |
|---|---|---|---|---|---|---|---|---|
| flexion_change | 0.71 | 0.91 | 0.93 | 0.78 | 0.71 | 0.93 | 0.83 | 0.11 |
| handshape | 1.45 | 1.43 | 1.45 | 1.25 | 1.25 | 1.45 | 1.40 | 0.10 |
| location | 0.85 | 0.83 | 1 | 1.13 | 0.83 | 1.13 | 0.95 | 0.14 |
| movement | 0.60 | 0.70 | 0.65 | 0.65 | 0.60 | 0.70 | 0.65 | 0.04 |
| orientation | 1.38 | 1.18 | 1.15 | 1.05 | 1.05 | 1.38 | 1.19 | 0.14 |
| spread_change | 0.50 | 0.50 | 0.88 | 0.78 | 0.50 | 0.88 | 0.67 | 0.19 |
| ulnar_rotation | 0.60 | 1 | 0.85 | 1 | 0.60 | 1 | 0.86 | 0.19 |

If the parameters are grouped according to their respective recognition approaches and their means are calculated, the weights shown in Table 28 result. With an average range of 0.28, these exhibit considerably lower variance between minimum and maximum values than the ungrouped parameters. This shows that although larger differences can occur locally within individual recognition approaches, the weighting of the recognition approaches at the global level remains consistent across the different lists.

The recognition approaches for path movement and spread change are weighted least strongly, with average weights of 0.65 and 0.67, respectively, whereas handshape recognition and palm orientation have the highest weights at 1.40 and 1.19, respectively, and thus exert the greatest influence on the classification result. This result is consistent with the observations from the evaluation of the individual classifiers in Section 6.2, where path movement and spread change achieved the lowest classification performance, whereas handshape and palm orientation not only showed high accuracies, but also provided the greatest amount of information for sign identification with 54 and 24 classes, respectively.

Recognition of flexion changes achieves good classification results in the individual evaluation, but comprises only two classes and, despite twice the number of parameters compared to handshape and palm orientation, provides only limited information content. Nevertheless, its influence on the meta-classification is rather low, with an average weight of 0.83. This can be explained in particular by the strongly unbalanced class ratio: for index and middle fingers, only 12.29 to 13.67 % of the instances belong to moving classes, and for ring finger and little finger only 9.28 to 9.39 %. Thus, in about nine out of ten cases there is no movement, so that despite good individual classification only a small contribution is made to identifying the sought sign.

To investigate the generalizability of the determined weights, they were cross-validated with different selections of sign recordings. For this purpose, the three

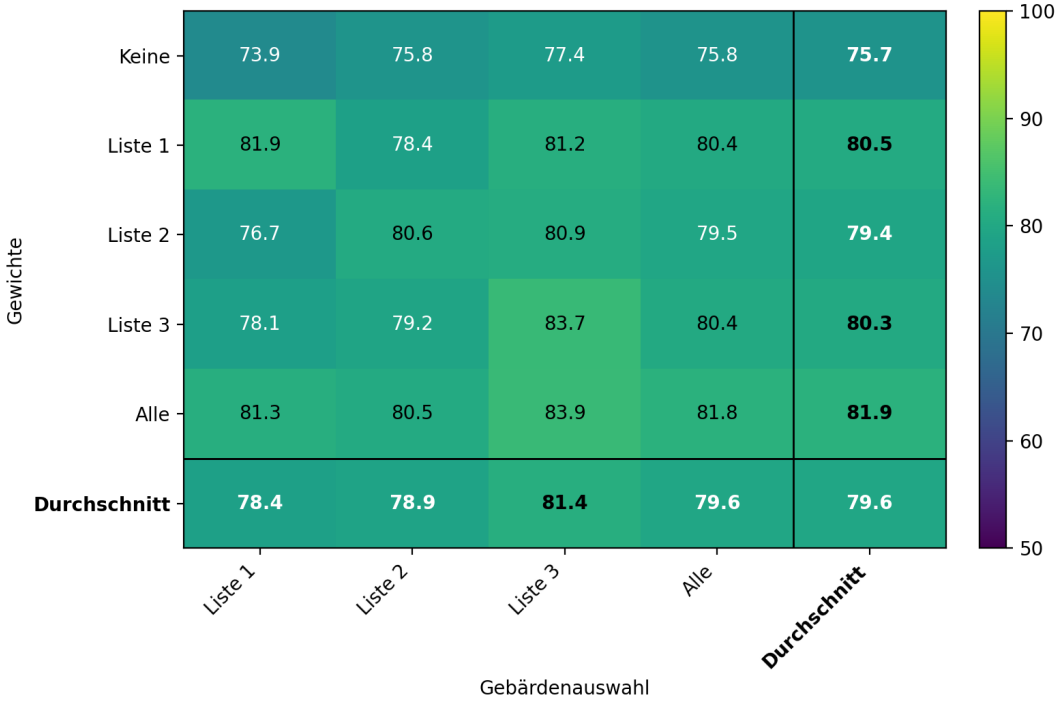



Figure 110: Top-1 accuracy of the meta-classification for different combinations of weighting variants and sign lists

previously defined lists of sign recordings and all sign recordings were classified using the four weighting variants shown in Table 27, as well as an unweighted configuration, and the respective Top-1 accuracy was determined. The results are shown in Figure 110.

Overall, the achieved accuracies are all in the range of 73.90 to 83.90 %, i.e., within a window of ten percentage points. Across all lists, the weighted approaches achieved average accuracies between 79.40 to 81.90 %, thus lying close together, which indicates good generalizability of the weights. As expected, the highest values were achieved with the weights optimized on all sign recordings. The lowest values were achieved with the weights from *List 2*.

The sign recordings of *List 3* achieved the highest accuracies across all weighting variants, with an average of 81.40 %, compared to 78.40 to 79.60 % for the other lists. It is notable that *List 3*, using the globally optimized weights, achieves a higher result than with the weights optimized on *List 3* itself. This is a direct consequence of the three-stage grid search used, in which the risk was deliberately taken of finding only a local optimum and not necessarily determining the globally best weighting.

Due to the good generalizability of the determined class weights, the subsequent evaluation of the meta-classification is carried out on the basis of all available sign recordings using the weights optimized on *List 1*. The signs from *List 1* are not excluded, in order to take into account a larger number of recordings as well as increased variance in the data.

### 6.3.2 *Results of the weighted meta-classification*

This section presents the results of the weighted meta-classification, which reflect the performance of the overall system. Building on the previously carried out optimization of the class weights, all available sign recordings are evaluated using the weights determined on the basis of *List 1*.

The aim of this evaluation is to investigate the recognition performance of the complete *ParaSignRec* approach under realistic conditions in which all parameter-based recognition approaches jointly contribute to the decision-making process. In particular, it is analyzed to what extent the individual parameters lead to an improved ranking of the predicted signs and how reliably the sought sign can be identified. The development of the identifiable sign vocabulary as a function of the parameters included is also taken into account.

Table 29: Top-k accuracies and average recognition time of the meta-classification for different gesture selections

| **Gestures** | **N** | **Ø Time** | **Top-k accuracies** | | | | |
|---|---|---|---|---|---|---|---|
| | | | **Top-1** | **Top-2** | **Top-3** | **Top-5** | **Top-10** |
| **All** | 1,741 | 122.67 ms | 80.41 % | 88.17 % | 91.84 % | 93.74 % | 95.23 % |
| **List 1** | 540 | 124.10 ms | 81.85 % | 86.67 % | 90.37 % | 92.22 % | 94.44 % |
| **List 2** | 625 | 123.51 ms | 78.40 % | 89.12 % | 92.48 % | 94.40 % | 95.84 % |
| **List 3** | 576 | 129.30 ms | 81.25 % | 88.54 % | 92.53 % | 94.44 % | 95.31 % |
| **Average** | 870.50 | 124.90 ms | 80.48 % | 88.13 % | 91.81 % | 93.70 % | 95.21 % |

N = Number of gesture recordings that were available for the respective evaluation

Table 29 shows the results of the meta-classification weighted according to *List 1* for different sign selections. In addition to the Top-k accuracies, the average recognition times per sign are also given. To calculate the runtime metrics, outliers were removed, i.e., any runtime that lies at least 1.50 standard deviations above the mean of the measured recognition times. This is intended to account for fluctuations in runtime due to background processes or other system-related influences. All runtime metrics refer to these cleaned values unless explicitly stated otherwise.

Regardless of the sign selection considered, the sought sign can be correctly identified in about four out of five cases; when considering all sign recordings, this corresponds to a Top-1 accuracy of 80.41 %. If the ranking list is considered further, an additional 10 to 20 % of the signs can be identified if the sought sign is not necessarily searched for in first place, but for example in second or tenth place of the ranking list. This behavior is particularly plausible when context information is additionally used to identify the sought sign, for example previously used signs within a conversation. Under this assumption, an average of 95.21 % of the signs can be correctly identified within the Top-10 predictions, which corresponds to about 19 out of 20 signs.

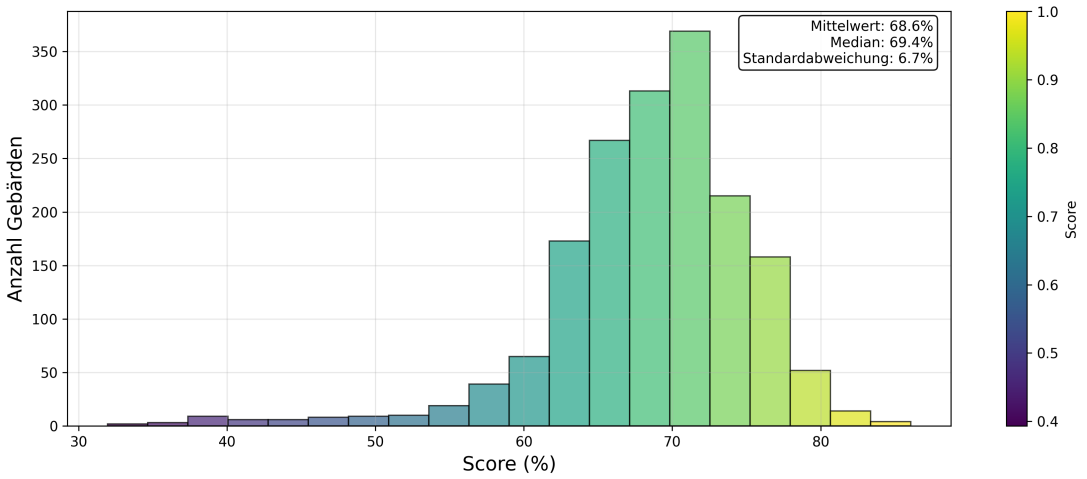



Figure 111: Distribution of the scores across all sign recordings

In the following, only the results considering all sign recordings are discussed. Figure 111 shows the distribution of the achieved scores across all 1,741 sign recordings. This exhibits an approximately bell-shaped structure, supported by the fact that mean and median, at 68.60 % and 69.40 %, respectively, are close together. The standard deviation is 6.70 %, with the majority of 90 % of the scores lying in the range of approximately 57.37 to 77.28 % (5th and 95th percentile). The lowest score is 31.96 %, the highest 86.08 %.

The score range of the respective ranks is shown in Figure 112. Overall, there are 1,400 sign recordings with a rank of 1, all of which have a score in the range of 48.67 to 86.08 %, with an average of 69.86 %. As the rank increases, the average score

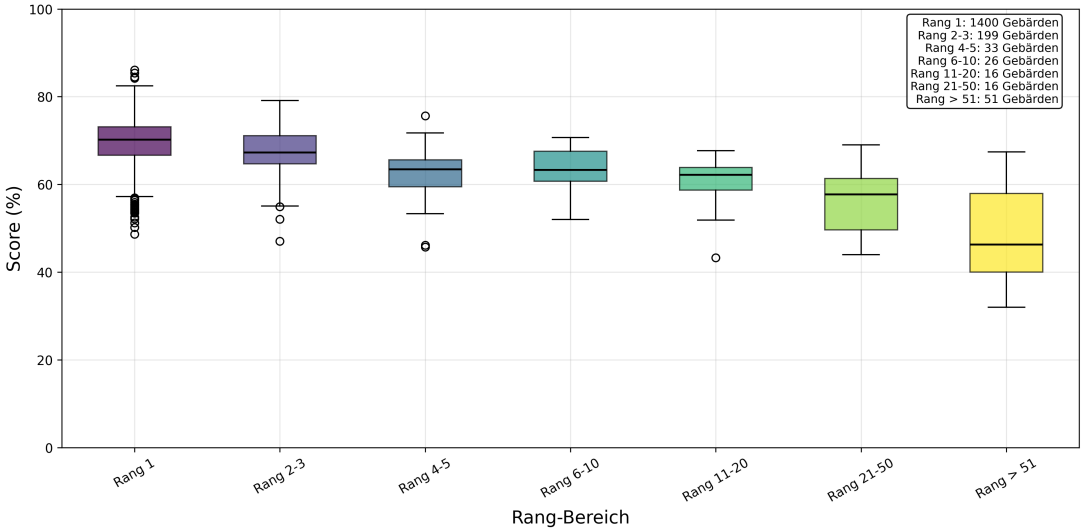



Figure 112: Score distributions as a function of the achieved rank

decreases as expected: for rank 2 it is 68.02 %, for rank 3 67.05 %, for rank 5 63.73 %, and for rank 10 60.19 %.

The results show that higher scores tend to lead to better ranks, but there is no unambiguous assignment between score and rank. For example, sign recordings with a score of 50 % (± 1 % tolerance) occupy ranks between 1 and 77, with an average rank of 35.80. The achieved rank therefore does not depend on the score alone, but largely on the score distribution of all signs in the vocabulary as well as on the similarity between competing signs.

Figure 113 shows the relationship between achieved score and attained rank for all sign recordings. Overall, the rank range extends from 1 to 1,678 with an average rank of 22.77 and a vocabulary of 2,171, which thus also represents the theoretically highest possible rank.

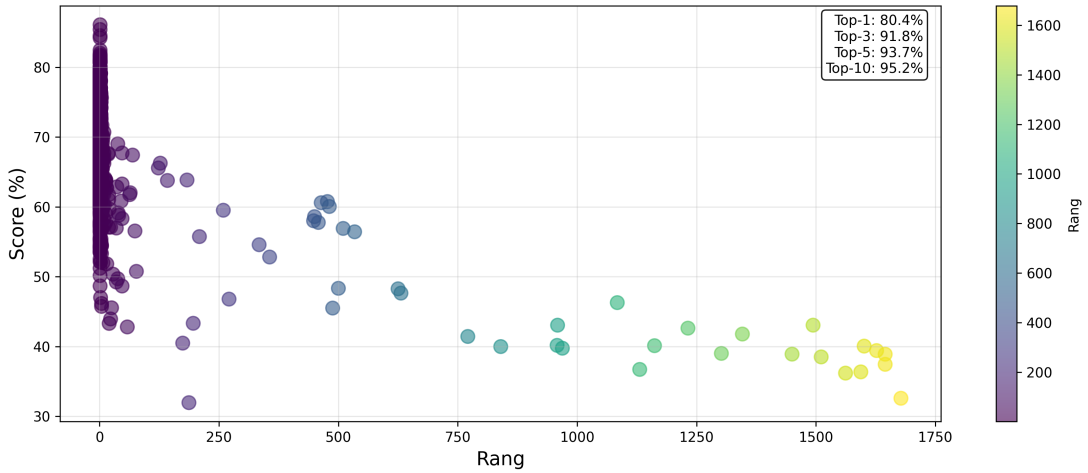



Figure 113: Relationship between score and rank of the predicted signs

Due to the problem of not unambiguously distinguishable signs described in Section 6.1.5, the actual dimension of the ranking list is generally below the number of signs in the vocabulary and varies depending on the respective recording: for each sign recording, a score is calculated for each sign in the vocabulary on the basis of the respective best parameter classifications. Since different signs can be mapped to identical combinations of parameter classes, identical score combinations may occur multiple times.

In the present evaluation, the minimum actual dimension of the meta-classifier is 1,807, the maximum 1,976. The mean is 1,900.40 with a standard deviation of 34.22. Figure 114 shows the distribution of these effective dimensions across all sign recordings.

If the achieved Top-1 accuracies are considered separately for the four sign types according to Battison, as shown in Figure 115, comparable classification results in the range of 77.42 to 80.97 % are obtained for the first three types. Only type 3, i.e., the asymmetric signs with different handshapes, achieves a somewhat higher classification result at 87.45 %.

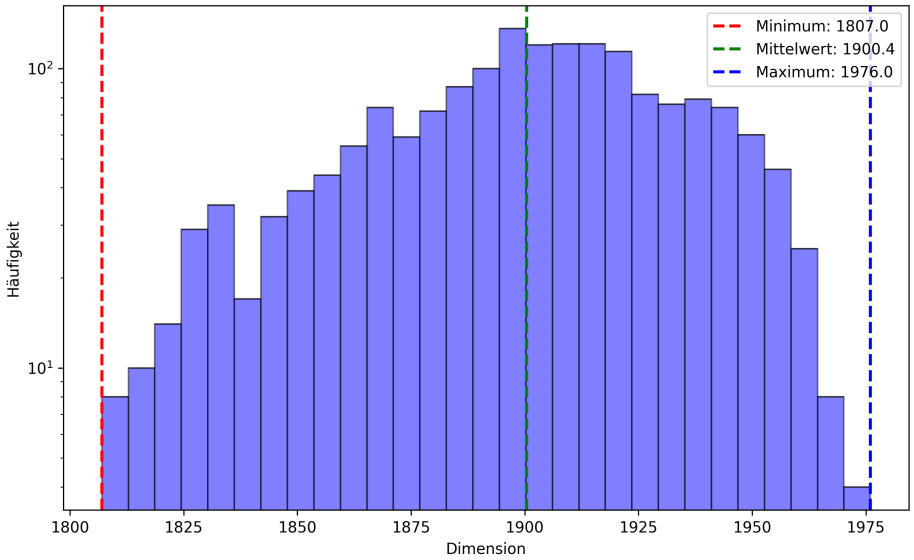



Figure 114: Effective dimension of the ranking list of the meta-classifier

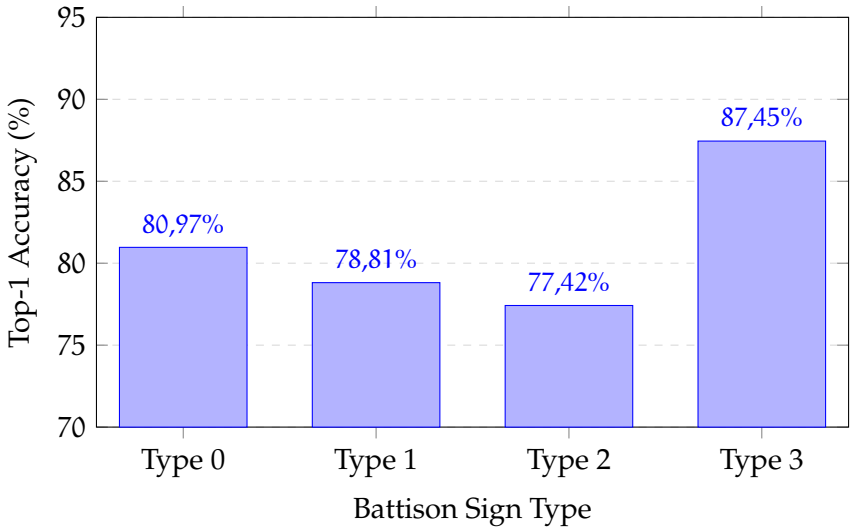



Figure 115: Top-1 accuracy by Battison sign types

### 6.3.3 *Analysis of recognition times*

Another aspect of the evaluation concerns the recognition times required to classify all parameters of a sign recording and, on the basis of these individual results, identify

the performed sign with the help of the meta-classification. All experiments carried out in this chapter were performed on an *Apple MacBook Pro*[88].

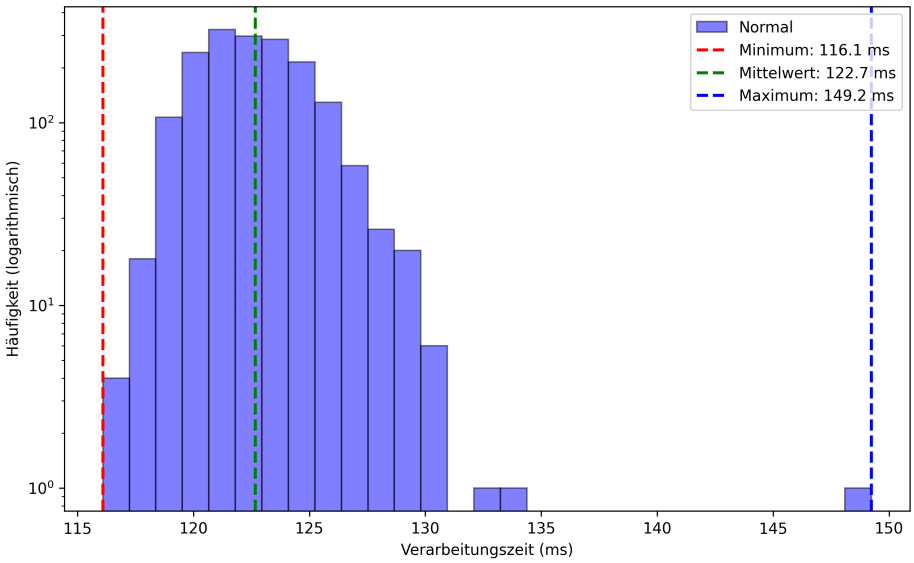


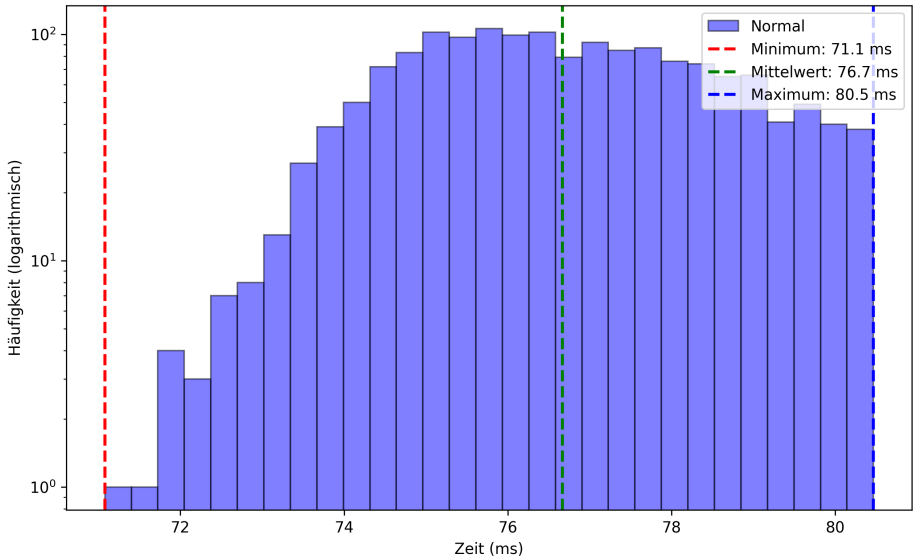

Figure 116: Distribution of the total recognition times per sign recording

Figure 116 shows the distribution of the total recognition times over all sign recordings examined. On average, the time required for the complete identification of a sign is 122.70 ms. The minimum measured runtime is 116.10 ms, the maximum 149.20 ms. 90 % of all runtimes lie between the 5th and 95th percentile in the range of 119.21 to 126.85 ms. In total, seven outliers in the range of 868.97 to 876.27 ms were identified by the procedure mentioned at the beginning and excluded from the evaluation, since these are likely attributable to external influences such as background processes.

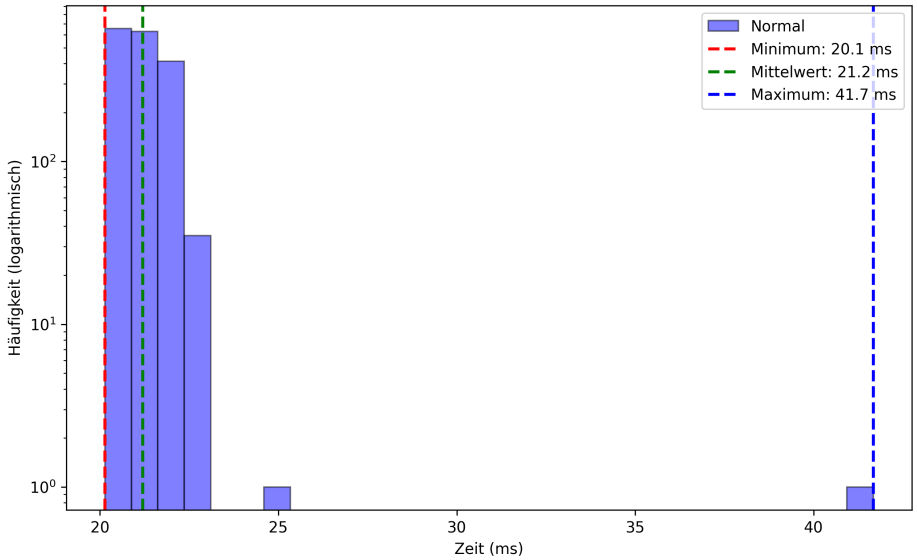
In Section 3.2.1, the requirements for real-time capability of the approach were defined and divided into four target levels, where the times achieved here, at 116.10 to 149.20 ms, are all below the threshold of 200 ms and thus satisfy the highest target level. Even when considering the outliers, the approach reaches the second target level with an average recognition time of 125.70 ms. This requires that the average recognition time should be faster than the shortest sign in its execution in order to prevent a backlog of sign recognition. For the first target level, it would also have been necessary that every recognition time be faster than the shortest sign in its execution.

If one differentiates between the times attributable to the overall parameter-based recognition approaches and the times required by the actual meta-classification when processing the individual results to calculate the scores for each sign in the vocabulary,

88 Apple MacBook Pro 14“, 2024 (`https://support.apple.com/de-de/121553`, last accessed on January 25, 2026), with *Apple M4 Pro* processor (14-core CPU with 10 performance cores and 4 efficiency cores, 20-core GPU, 16-core Neural Engine, and 410 GB/s memory bandwidth) and 48 GB shared memory




(a) Run times of the parameter-based recognition approaches




(b) Run times of the meta-classification

Figure 117: Run times of the parameter-based recognition approaches and the meta-classification

then Figure 117a and Figure 117b show that the share of the parameter-based recognition approaches, at 71.08 to 80.47 ms, is significantly higher than that of the actual meta-classification. This requires only 20.14 to 41.68 ms, with an average of 21.20 ms.

In addition, further runtimes arise that cannot be directly assigned to individual recognition approaches or to the meta-classification. These include, among others, overhead for data management or score normalization. These runtimes are obtained as the difference between the total recognition time and the previously considered shares, and are shown in Figure 118. They lie in the range of 23.02 to 30.83 with an average value of 24.41 ms, and are therefore of a similar order of magnitude as the meta-classification itself.

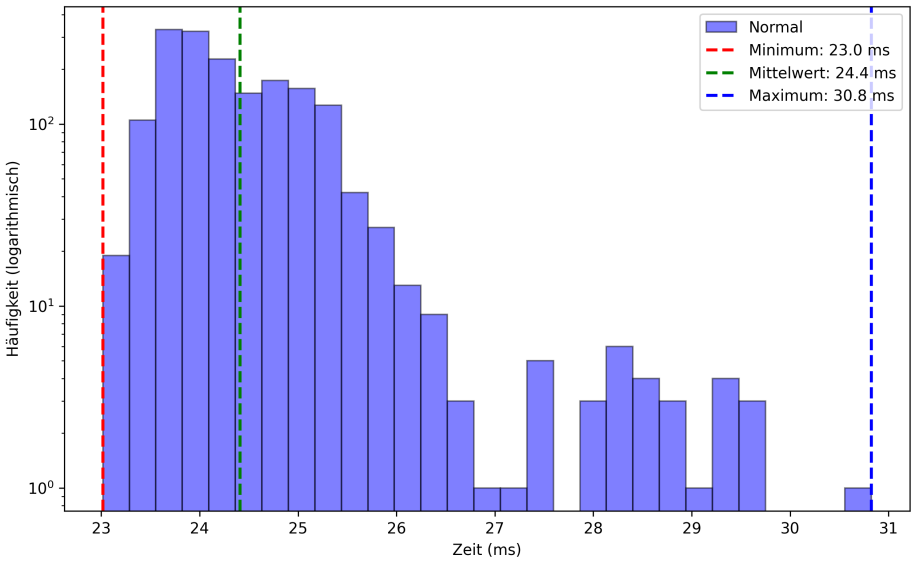



Figure 118: Other runtime components

### 6.3.4 *Influence of the parameter-based recognition approaches*

The previous analysis shows that the total runtime of the meta-classification is determined largely by the parameter-based recognition approaches. The next step is therefore to investigate what temporal share the individual parameter recognitions have in the total runtime. To this end, the mean time expenditure of the individual parameter recognitions is considered, which is shown in Figure 119. The runtimes of the dominant and non-dominant hand are combined in each case.

Figure 119 makes clear that in particular the ML-based approaches of handshape recognition and path movement recognition exhibit significantly longer average runtimes than the rule-based recognition approaches. While the latter require only 0.97 ms on average, the runtimes of handshape recognition are between 24.04 to 24.09 ms and are thus the highest. Path movement recognition follows with an average runtime of 17.74 ms.

Among the rule-based approaches, recognition of the sign execution location takes the longest, at 4.28 to 4.30 ms. This is due to the calculation of distances between the

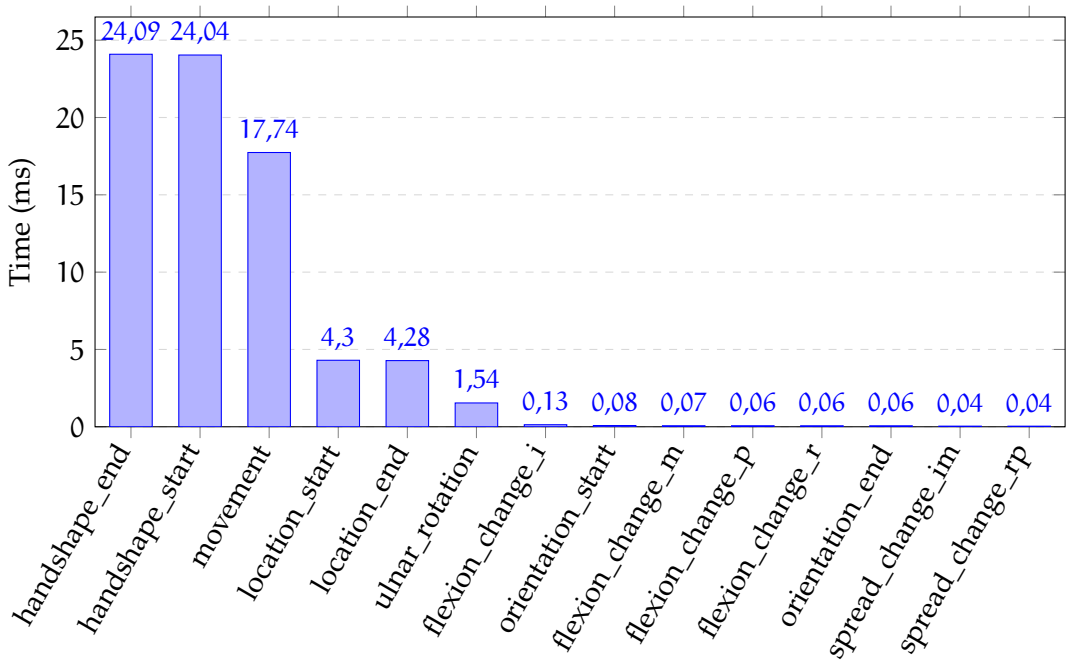



Figure 119: Average run times of the parameter-based recognition approaches per sign recording

current hand position and all reference positions. Recognition of a wrist movement also has an above-average runtime of 1.54 ms, which is attributable to the computationally expensive calculation and transformation of the rotational change across all frames of a sign recording. With an average of 79.50 frames per sign recording (see Figure 120), this effort accumulates accordingly.

In addition to the influence of the individual parameter-based recognition approaches on the runtime of sign recognition, the following also examines the contribution of the respective parameters to recognition accuracy and to the size of the effectively identifiable sign vocabulary. To this end, the results were evaluated again and one individual parameter was removed from the meta-classification in each case, one after the other. These results of this analysis are shown in Figure 121.

It can be seen that removing palm orientation has the greatest negative influence on Top-1 accuracy, followed by handshape recognition. Path movement recognition, on the other hand, makes the smallest positive contribution to Top-1 accuracy, and removing it even leads to a slight improvement in the results. This behavior is consistent with the previously determined weights of these parameters. Similar effects are observed for Top-10 accuracy, but in a weaker form. Due to the small number of classes in path movement recognition, its removal no longer has any influence here.

If one considers the influence of the individual parameters on the size of the effectively identifiable sign vocabulary, the recognition of path movement, palm orientation, and, to a lesser extent, wrist rotation show a clear effect. The influence of handshape recognition on the vocabulary is, accordingly, low, because the handshape parameters are present both at the beginning and at the end of sign execution and

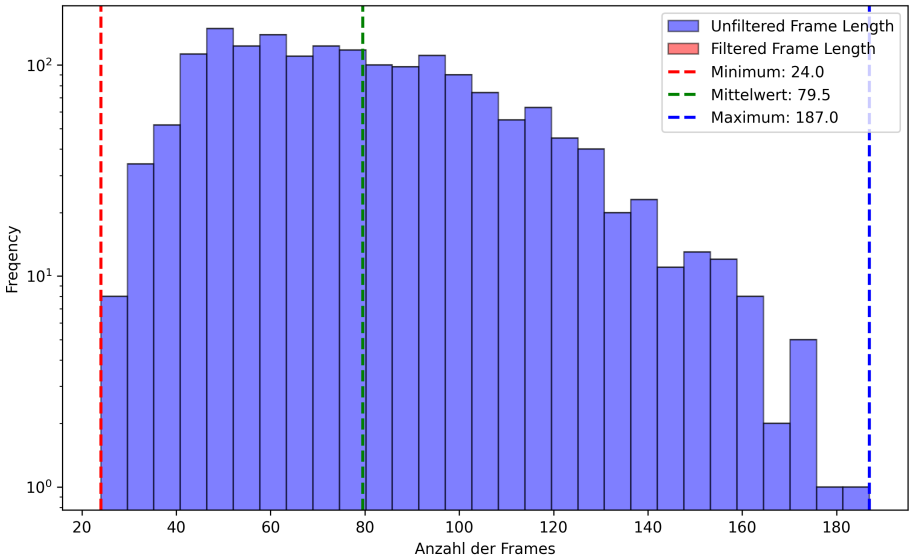



Figure 120: Number of frames per sign

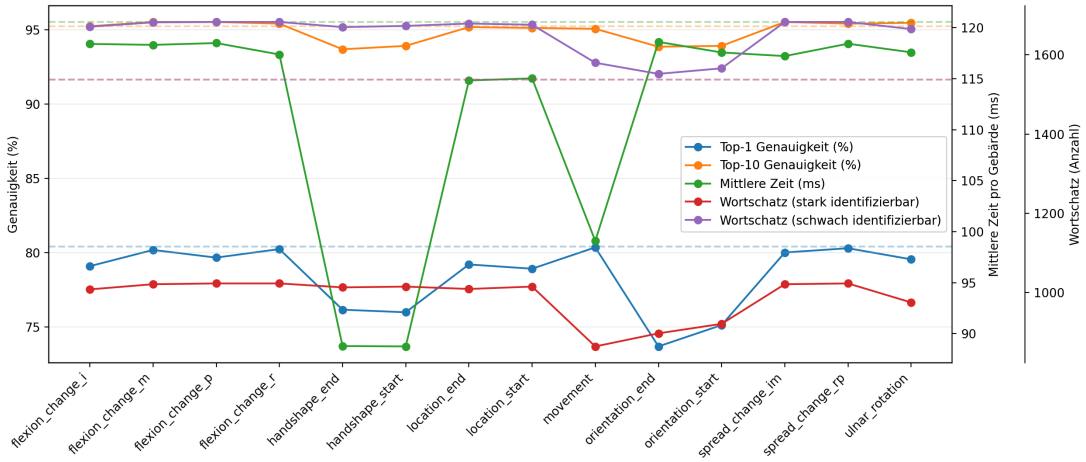



Figure 121: Influence of missing parameters on classification results, runtimes, and vocabulary size

often do not differ: for the dominant hand, 71.43 % of the handshape classes at the beginning and end of sign execution are identical; for the non-dominant hand, the figure is 78.10 %. If one of these parameters is removed, the other can compensate for a large part of the missing information. For palm orientation, this redundancy is considerably lower, at 29.52 % for the dominant hand and 50.48 % for the non-dominant hand, since it often takes a passive role. Its influence on the vocabulary is therefore correspondingly greater. In the reverse case, the influence of path movement is higher,

because, apart from wrist rotation, it is the only parameter that exists only once for the dominant and non-dominant hand.

The evaluation also confirms our assumption that including handshape at the end of sign execution increases the accuracy of the system, while no significant change is observed in the scope of the distinguishable vocabulary.

Since up to this point the influence of individual parameters has been considered, the next step is to analyze the influence of entire recognition approaches. For this purpose, the parameters are grouped according to their respective recognition approaches and the experiment is repeated. The results are shown in Figure 122. In particular for parameters that differ only in their temporal position within sign execution (e.g., handshape at the beginning and end of sign execution), dependencies become apparent depending on whether a movement, and thus a change, occurred or not.

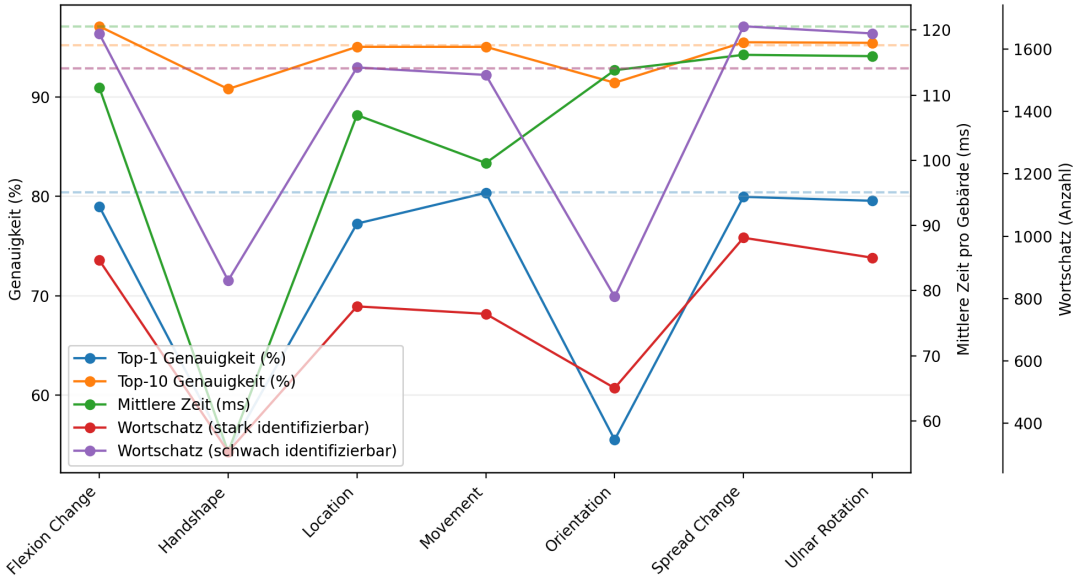



Figure 122: Influence of missing recognition approaches on classification results, runtimes, and vocabulary size

Figure 122 makes clear that removing handshape recognition has the greatest influence on the size of the identifiable sign vocabulary, followed by palm orientation. Since these parameters are particularly characterized by the use of polysemous parameter classes, it is noticeable that the strongly and weakly identifiable vocabularies become much more similar once one of the two parameters is missing. If only unambiguous parameter classes were available, the strongly identifiable vocabulary would be identical to the weakly identifiable vocabulary.

### 6.3.5 *Vocabulary dependence*

As a final experiment, it is investigated how the presented approach behaves with a varying number of target signs, i.e., with different sizes of the sign vocabulary. To this end, 25 random signs were first selected from the *ParaSignRec* vocabulary and

evaluated. Subsequently, the vocabulary was expanded step by step, doubling the number of target signs each time, up to a size of 1,600 signs. For comparison, the complete vocabulary was also included in the consideration and evaluated again.

When selecting the target signs, care was taken to choose signs from the 210 recorded test signs first and only then additional, non-recorded signs. All signs of a smaller vocabulary are fully contained in the respective larger vocabulary sizes.

In this experiment, not only the quality of the data determines the classification result, but also the number and selection of target signs. This is reflected in particular in the average rank of the predicted signs, since the number of competing signs also increases with the number of target signs.

Table 30: Top-k accuracies and average rank of the meta-classification for different sign selection sets

| **Signs** | | **Top-k accuracies** | | | | | **Ø rank** |
|---|---|---|---|---|---|---|---|
| **Number** | **Share** | **Top-1** | **Top-2** | **Top-3** | **Top-5** | **Top-10** | |
| **25** | 100 % | 99.06 % | 100 % | 100 % | 100 % | 100 % | 1.01 |
| **50** | 100 % | 96.48 % | 97.89 % | 98.83 % | 99.30 % | 99.30 % | 1.13 |
| **100** | 100 % | 90.26 % | 94.06 % | 95.84 % | 96.56 % | 97.39 % | 2.01 |
| **200** | 100 % | 84.35 % | 91.64 % | 94.32 % | 95.58 % | 97.31 % | 3.41 |
| **400** | 52 % | 82.31 % | 90.12 % | 93.34 % | 94.83 % | 96.32 % | 6.02 |
| **800** | 26 % | 81.56 % | 89.03 % | 92.59 % | 94.26 % | 95.81 % | 10.34 |
| **1.600** | 13 % | 80.76 % | 88.57 % | 92.07 % | 93.97 % | 95.35 % | 18.04 |
| **2,171** | 9.60 % | 80.41 % | 88.17 % | 91.84 % | 93.74 % | 95.23 % | 22.77 |

Share = share of the signs included from the respective vocabulary

Table 30 shows the classification results for the vocabulary sizes examined. For each dataset, it is also stated what proportion of recorded signs there is relative to the vocabulary. Since the recorded test signs were prioritized when selecting the target signs, this proportion is 100 % for vocabulary sizes of 25 to 200 signs each and then decreases gradually as the vocabulary size increases.

As expected, the Top-k accuracies and the average rank decrease as the number of signs in the vocabulary increases. For a small vocabulary of 25 and 50 signs, the Top-1 accuracy is very high at 99.06 % and 96.48 %, respectively, and stabilizes from a vocabulary of 400 signs onward at a level of a little more than 80 %. Across all vocabulary sizes, a very high Top-10 accuracy of 95.23 to 100 % is achieved.

However, the average rank provides a more differentiated view. In particular, for the complete vocabulary there is a high Top-10 accuracy, but at the same time the average rank of 22.77 lies clearly outside the Top-10. If the proportion of correctly identified signs at rank 1 (80.41 %) is excluded, the remaining recordings yield an average rank of 109.65, which, in relation to a vocabulary of 2,171 signs, can still be classified as comparatively favorable.

Table 31: Number of strongly and weakly identifiable signs in the vocabulary and average recognition times of the meta-classification for different sign selections

| Signs | | ΔVocabulary | | Ø Time | | |
|---|---|---|---|---|---|---|
| **Count** | **Share** | **strong** | **weak** | **Parameter** | **Meta** | **Total** |
| **25** | 100 % | 23 | 24 | 72.43 ms | 0.29 ms | 76.84 ms |
| **50** | 100 % | 40 | 47 | 73.85 ms | 0.55 ms | 78.47 ms |
| **100** | 100 % | 80 | 97 | 73.85 ms | 0.98 ms | 79.10 ms |
| **200** | 100 % | 137 | 192 | 76.03 ms | 1.94 ms | 83.23 ms |
| **400** | 52 % | 283 | 376 | 75.40 ms | 3.89 ms | 86.60 ms |
| **800** | 26 % | 512 | 719 | 76.10 ms | 8.12 ms | 95.59 ms |
| **1.600** | 13 % | 853 | 1,318 | 76.80 ms | 15.65 ms | 110.96 ms |
| **2,171** | 9.60 % | 1,023 | 1,683 | 77.84 ms | 21.94 ms | 123.87 ms |

Share = share of the recorded signs from the respective vocabulary;
Ø Time Parameter/Meta/Total = average run times of all parameter-based recognition approaches / average run times of the actual meta-classification / total time

Table 31 shows the development of the recognition times and the proportion of strongly and weakly identifiable signs as a function of vocabulary size. The runtimes remain largely stable up to a vocabulary size of about 100 signs, because in this range the meta-classification accounts for only a small proportion of the total runtime. As the vocabulary size increases, the time required by the meta-classification rises significantly, from 0.29 ms for small vocabularies to 21.94 ms for the complete vocabulary. The runtimes of the parameter-based recognition approaches, in contrast, remain largely constant across all vocabulary sizes and lie in the range of 72.43 to 77.84 ms. Other runtime components are not shown separately in these values, but are part of the total runtime.

Since in particular the Top-k accuracies and the average rank depend strongly on the specific selection of target signs, an additional experiment was conducted in which 25 signs were randomly selected 100 times each from the vocabulary of recorded signs and evaluated using the meta-classification. The results of this evaluation are shown in Table 32.

The results show that even with an identical vocabulary size, considerable variance in the results can occur. Top-1 accuracy varies between 84.24 % and 100 %, with a mean of 93.68 %. The average rank is very low at 1.39 and shows a small standard deviation. These results emphasize that the performance of the approach is influenced not only by size, but also, and decisively, by the composition of the sign vocabulary.

Table 32: Average classification results of 100 randomized sign vocabularies with 25 signs each

| **Metric** | **Minimum** | **Maximum** | **Mean** | **Std. dev.** |
|---|---|---|---|---|
| Top-1 accuracy | 84.24 % | 100 % | 93.68 % | 2.98 % |
| Top-2 accuracy | 88.18 % | 100 % | 96.46 % | 2.52 % |
| Top-3 accuracy | 89.16 % | 100 % | 97.11 % | 2.42 % |
| Top-5 accuracy | 92.57 % | 100 % | 97.83 % | 1.95 % |
| Top-10 accuracy | 93.12 % | 100 % | 98.29 % | 1.71 % |
| ∅ Rank | 1.00 | 2.36 | 1.39 | 0.32 |

# 7 CONCLUSION AND OUTLOOK

This chapter summarizes the central results of the present work, places them in the context of the challenges and objectives from Chapter 1, and provides an outlook on possible extensions and future research directions.

## 7.1 CONCLUSION

This work addresses the design, implementation, and evaluation of an approach for parametric real-time recognition of isolated hand gestures in the form of ASL signs using body-worn sensors, referred to as *ParaSignRec*. In contrast to holistic methods, in which signs are considered and classified as a whole, the signs in this approach are decomposed into their characteristic, meaning-distinguishing parameters and classified at this level.

Through this parametrization, the effort required for data collection can be significantly reduced, since new signs can be integrated into the vocabulary without retraining the underlying models from scratch. Moreover, there is a structurally induced generalizability, as the individual recognition approaches were trained largely independently on parameter-specific data foundations rather than being based on a shared dataset, as in holistic methods.

The signs are captured by means of a combination of data gloves and IMUs attached to the upper arm and forearm. The actual identification of the signs is based on the aggregation of the recognized parameters through a probabilistic meta-classification, which computes a weighted score for each sign in the dictionary and ranks the signs accordingly.

A central objective and at the same time the foundation of this work was the identification and systematic description of the parameters that distinguish signs semantically. To this end, a comprehensive analysis of existing notation systems was first carried out in order to identify the essential parameters of a sign and their possible classes in handshape, palm orientation, location of articulation, path movement of the sign, wrist rotation, and finger-intrinsic changes, such as flexion or spreading. Based on these requirements, various sign databases were examined in order to obtain a suitable data foundation, which was found with *ASL-LEX*. This was analyzed, curated, and extended in detail.

Starting from initially 2,723 signs described by 16 parameters and exhibiting $2.79 \times 10^{13}$ theoretical combination possibilities, the sign dictionary of *ParaSignRec* comprises a vocabulary of 2,171 signs described by 28 parameters and allowing $5.39 \times 10^{20}$ possible combinations. A major achievement in this context is the addition of palm orientation, which was not previously considered in *ASL-LEX* and proved to be a key factor for increasing the vocabulary size as well as improving classification results. In

addition, the system also supports multi-valued parameter classes that are not defined unambiguously but take into account a certain natural variability of signs.

The results show that signs can be described by a manageable set of characteristic parameters whose classes enable reliable parametrization. The chosen parametrization proved suitable for identifying signs probabilistically and for making the sign vocabulary flexible and scalable. The developed parametrized sign dictionary comprises over 2,000 signs, of which more than 1,000 signs are highly identifiable and thus exhibit unambiguous parameter combinations, and therefore constitutes a limited, yet functionally sufficient and extendable vocabulary of ASL. Thus, the objective of parameter identification as a basis for scalable and extensible sign recognition was achieved.

Building on the successful identification of the relevant parameters, the next central objective of this work was the development of suitable recognition approaches for the individual parameters. The aim was to develop robust, mobile, and user-friendly methods that support real-time sign recognition.

For this purpose, different hardware concepts were first analyzed. A combination of data gloves and IMUs attached to the forearm and upper arm proved to be a suitable setup, enabling sufficiently accurate capture of the relevant motion and orientation data.

The 28 parameters identified in total are distributed across the dominant and non-dominant hand and are recognized by seven specialized, ML-based as well as rule-based classifiers. These were designed with particular emphasis on high generalizability in order to compensate for variations in sign execution by different users. In the evaluation, they were able to achieve high recognition accuracy while maintaining a low classification time.

The results show that the chosen sensor setup, in combination with the developed recognition approaches, is fundamentally suitable for reliably capturing both static and dynamic sign parameters. Thus, the objective of fast and robust parameter recognition was achieved; it is also mobile and user-friendly and enables real-time sign identification. Nevertheless, individual parameters still exhibit optimization potential with regard to robustness when processing real sign recordings.

As the final central objective of this work, the development of a meta-classification was pursued that combines the individual parameter-based recognition approaches in such a way that the target sign can be reliably identified in real time from the developed sign vocabulary. A probabilistic approach was implemented that enables robust sign predictions even in the presence of ambiguous or partially erroneous parameter configurations while also requiring little computational effort. This is particularly important with regard to the score computation for multi-valued parameter classes, as these can lead to a large number of possible sign variants.

By aggregating the individual classifiers using weighted probabilistic models and focusing on the respective best-ranked classes, an ordered ranking of possible signs could be generated. In this ranking, the correct sign is in first place in 80.41 % of cases and among the top ten candidates in 95.23 % of cases. The complete processing of a sign recording requires an average of 122.67 ms, of which approximately 76.70 ms

are attributable to the overall parameter-based recognition approaches and 21.20 ms to the actual meta-classification.

The evaluation under realistic conditions therefore confirms that the developed meta-classifier can reliably handle ambiguities and uncertainties in the recognized parameter configurations and correctly identifies the target sign within the defined target times in about four out of five cases. Thus, the objective of meta-classification, namely to effectively combine the recognized parameters and enable reliable identification of the target sign, was successfully achieved.

Related work was able, in the best case, to classify up to 100 signs with a *within-user* accuracy of 98.10 % [145]. In a comparable scenario with the same number of signs, the approach presented in this work achieves an accuracy of 90.26 %. While this value is 7.84 percentage points below the cited result, it represents a *cross-user* accuracy and thus a significantly more generalizable result. Furthermore, the data came from an experienced sign language user, whereas the signs used in this work were recorded by three laypersons.

The results of this work show that a parametric approach to sign recognition is not only feasible in principle, but also offers clear advantages over holistic methods, especially as vocabularies grow. From a certain vocabulary size onward, the effort required for data collection is significantly reduced. In addition, new signs can be integrated into the system without retraining the underlying recognition approaches.

In the evaluation, the recognition of handshape and palm orientation proved to be the most semantically significant parameters, whereas the recognition of path movement and finger spreading exhibited the greatest potential for improvement. Further limitations are partly attributable to hardware-related constraints.

Moreover, the calibration process constitutes a critical factor and should be better monitored and validated in the future. Although the constructed sign vocabulary forms a solid foundation, it is not yet fully mature. In particular, the class definitions of path movement exhibit significant conceptual weaknesses, which had a substantial negative effect on the achieved classification results.

The number of considered articulation locations was reduced significantly due to the limited accuracy of the sensor setup, which causes potentially relevant information for sign identification to be lost. More precise capture of these articulation locations could contribute both to expanding the sign vocabulary and to improving classification results. In addition, existing annotations should be validated, as numerous inconsistencies were identified in the *ASL-LEX* data foundation on the one hand, and, on the other hand, palm orientation was predominantly annotated automatically and handshape at the end of sign execution was in some cases equated with handshape at the beginning of sign execution.

## 7.2 OUTLOOK

Finally, an outlook on possible extensions of the presented approach and possible future research work is given here.

Future work could first further examine and revise the class definitions of individual parameters, especially path movements. A finer and alternative class definition of path movements could lead to better classification results. In addition, a more detailed selection of articulation locations offers the potential to increase the informational content of the parametrization in order to reduce duplicates and further improve the quality of the dictionary. In this context, manual annotation of additional signs would also be useful in order to further expand the sign vocabulary.

At the level of the recognition approaches, the sensor setup used constitutes an essential starting point for future improvements. More precise capture, for example through more accurate IMUs, could improve the distinction of articulation locations and thus enable finer class differentiation. Likewise, improved data gloves, for example through additional flex sensors between the fingers and along the finger flexions, could contribute to more reliable spread values and directly measured rather than estimated flexion values. In addition, the thumb could be given a greater role as the most versatile finger, since in the current approach it is considered only in handshape recognition and, for example, not in finger-intrinsic movements.

For better prediction of the target sign within the framework of meta-classification, it would be advisable to take the lexical frequency of the signs into account. This is given in *ASL-LEX* and describes how often a particular sign occurs in a language [187]. Frequently used signs could thus be given higher priority, which would be particularly useful in realistic conversational situations. However, for the evaluation carried out in this work, this approach is only partially applicable, since the signs were selected randomly independently of their actual language use and therefore exhibit a uniform frequency distribution.

A similar approach could not only consider the data of the current sign in isolation, but also include previously recognized signs in the decision. A context-based extension of the meta-classification could narrow down the candidates under consideration and improve prediction.

However, the approach presented in this work is currently limited to monomorphemic, isolated signs and would have to be extended for such context-based recognition. By adding automatic segmentation of continuous sign streams, it would then be possible, on the one hand, to recognize whole sentences instead of individual isolated signs. On the other hand, multimorphemic signs could be differentiated by no longer applying the existing approach to sign recognition but to morpheme recognition and supplementing it with an aggregation layer above it. This then derives the target sign from several recognized morphemes, analogous to the meta-classification presented here.

In the area of *data augmentation*, the experiments already carried out could also be extended to dynamic parameters and it could be investigated to what extent artificially generated variations of existing data contribute to increasing the robustness and generalizability of the ML-based classifiers. Conversely, external data sources, such as freely available sign video databases, could be used to transfer them into our domain. From these data, additional training data could either be extracted or annotations

derived, as was already exemplarily implemented in this work for palm orientation. In this way, for example, new signs could also be integrated into the vocabulary.

Finally, the parametric approach is not fundamentally limited to ASL. By separating parametrization and sign dictionary, the system can in principle be transferred to other sign languages, provided that the respective signs can be described by the parameter classes available in the system or that these are expanded accordingly. Thus, the presented approach offers a promising basis for cross-linguistic, scalable sign recognition.

# BIBLIOGRAPHY


[1] Kalpattu S. Abhishek, Lee Chun Fai Qubeley, and Derek Ho. "Glove-based hand gesture recognition sign language translator using capacitive touch sensor." In: *2016 IEEE International Conference on Electron Devices and Solid-State Circuits (EDSSC)*. 2016, pp. 334–337. DOI: 10.1109/EDSSC.2016.7785276.

[2] Joao Gabriel Abreu, Joao Marcelo Teixeira, Lucas Silva Figueiredo, and Veronica Teichrieb. "Evaluating Sign Language Recognition Using the Myo Armband." en. In: *2016 XVIII Symposium on Virtual and Augmented Reality (SVR)*. Gramado, Brazil: IEEE, June 2016, pp. 64–70. ISBN: 978-1-5090-4149-7. DOI: 10.1109/SVR.2016.21. URL: http://ieeexplore.ieee.org/document/7517255/ (Last accessed on Dec. 4, 2023).

[3] Huda Abualola, Hanin Al Ghothani, Abdulrahim Naser Eddin, Nawaf Almoosa, and Kin Poon. "Flexible gesture recognition using wearable inertial sensors." en. In: *2016 IEEE 59th International Midwest Symposium on Circuits and Systems (MWSCAS)*. Abu Dhabi, United Arab Emirates: IEEE, Oct. 2016, pp. 1–4. ISBN: 978-1-5090-0916-9. DOI: 10.1109/MWSCAS.2016.7870143. URL: http://ieeexplore.ieee.org/document/7870143/ (Last accessed on Jan. 7, 2024).

[4] Philipp Achenbach, Yasmin Göksu, Timo Kullmann, Thomas Tregel, and Stefan Göbel. "Towards handshape identification for automatic gesture recognition using sign notation systems." en. In: *Proceedings of the 8th European Conference on Social Media: A Virtual Conference Hosted By University of Central Lancaster, UCLan, Cyprus 1-2 July 2021*. University of Central Lancashire, Cyprus: Academic Conferences International Limited, July 2021. ISBN: 978-1-914587-01-6. DOI: 10.34190/ESM.21.100.

[5] Philipp Achenbach, Sebastian Laux, Dennis Purdack, Philipp Niklas Müller, and Stefan Göbel. "Give Me a Sign: Using Data Gloves for Static Hand-Shape Recognition." en. In: *Sensors* 23.24 (Dec. 2023), p. 9847. ISSN: 1424-8220. DOI: 10.3390/s23249847. URL: https://www.mdpi.com/1424-8220/23/24/9847 (Last accessed on Jan. 6, 2024).

[6] Philipp Achenbach, Philipp Niklas Muller, Tobias Alexander Wach, Thomas Tregel, and Stefan Gobel. "Rock beats Scissor: SVM based gesture recognition with data gloves." en. In: *2021 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*. Kassel, Germany: IEEE, Mar. 2021, pp. 617–622. ISBN: 978-1-6654-0424-2. DOI: 10.1109/PerComWorkshops51409.2021.9430962. URL: https://ieeexplore.ieee.org/document/9430962/ (Last accessed on Jan. 3, 2024).

[7] Philipp Achenbach, Dennis Purdack, Sebastian Wolf, Philipp Niklas Müller, Thomas Tregel, and Stefan Göbel. "Paper Beats Rock: Elaborating the Best Machine Learning Classifier for Hand Gesture Recognition." en. In: *Serious Games*. Ed. by Heinrich Söbke, Pia Spangenberger, Philipp Müller, and Stefan Göbel. Vol. 13476. Series Title: Lecture Notes in Computer Science. Cham: Springer International Publishing, 2022, pp. 229–245. ISBN: 978-3-031-15324-2 978-3-031-15325-9. DOI: 10.1007/978-3-031-15325-9_17. URL: https://link.springer.com/10.1007/978-3-031-15325-9_17 (Last accessed on Jan. 3, 2024).

[8] Shawkat Ali and Kate A Smith-Miles. "Improved Support Vector Machine Generalization Using Normalized Input Space." In: *AI 2006: Advances in Artificial Intelligence*. Springer, 2006, pp. 362–371. URL: https://link.springer.com/content/pdf/10.1007/11941439.pdf?pdf=button.

[9] All About Circuits. *Teardown Tuesday: Leap Motion Controller*. Tech. rep. All About Circuits, Sept. 2019. URL: https://www.allaboutcircuits.com/news/teardown-tuesday-leap-motion-controller/ (Last accessed on May 18, 2025).

[10] Christoph Amma, Thomas Krings, Jonas Böer, and Tanja Schultz. "Advancing Muscle-Computer Interfaces with High-Density Electromyography." en. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Seoul Republic of Korea: ACM, Apr. 2015, pp. 929–938. ISBN: 978-1-4503-3145-6. DOI: 10.1145/2702123.2702501. URL: https://dl.acm.org/doi/10.1145/2702123.2702501 (Last accessed on Feb. 14, 2024).

[11] Catherine Anderson, Bronwyn Bjorkman, Derek Denis, Julianne Doner, Margaret Grant, and Nathan Sanders. "Essentials of Linguistics, 2nd edition." en. In: (2022).

[12] Derek W. Orbaugh Antillon, Christopher R. Walker, Samuel Rosset, and Iain A. Anderson. "Glove-Based Hand Gesture Recognition for Diver Communication." en. In: *IEEE Transactions on Neural Networks and Learning Systems* (2022), pp. 1–13. ISSN: 2162-237X, 2162-2388. DOI: 10.1109/TNNLS.2022.3161682. URL: https://ieeexplore.ieee.org/document/9760170/ (Last accessed on Nov. 28, 2023).

[13] Chris Wilson Antuvan and Lorenzo Masia. "An LDA-Based Approach for Real-Time Simultaneous Classification of Movements Using Surface Electromyography." en. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 27.3 (Mar. 2019), pp. 552–561. ISSN: 1534-4320, 1558-0210. DOI: 10.1109/TNSRE.2018.2873839. URL: https://ieeexplore.ieee.org/document/8649712/ (Last accessed on Feb. 13, 2024).

[14] David F. Armstrong, Michael A. Karchmer, and John Vickrey Van Cleve, eds. *The Study of Signed Languages: Essays in Honor of William C. Stokoe*. eng. 1st edition edition. OCLC: 1100860759. Washington, D.C: Gallaudet University Press, 2011. ISBN: 978-1-56368-187-5.

[15] Robert W Arnold. "A proposal for a written system of American Sign Language." en. In: (2007).

[16] Vincent Van Asch. "Macro- and micro-averaged evaluation measures." en. In: (2013).

[17] *ASL Font*. English. Publication Title: ASL Font: Symbol Font for ASL. May 2013. URL: http://aslfont.github.io/ (Last accessed on May 4, 2021).

[18] ASLdeafined.com. *40 ASL handshapes*. English. Publication Title: 40 ASL handshapes. Jan. 2024. URL: ASLdeafined.com (Last accessed on Jan. 31, 2024).

[19] Emmanuel Ayodele, Tianzhe Bao, Syed Ali Raza Zaidi, Ali M. A. Hayajneh, Jane Scott, Zhi-Qiang Zhang, and Des McLernon. "Grasp Classification With Weft Knit Data Glove Using a Convolutional Neural Network." en. In: *IEEE Sensors Journal* 21.9 (May 2021), pp. 10824–10833. ISSN: 1530-437X, 1558-1748, 2379-9153. DOI: 10.1109/JSEN.2021.3059028. URL: https://ieeexplore.ieee.org/document/9353548/ (Last accessed on Nov. 20, 2023).

[20] Debasish Bal, Asif Mohammed Arfi, and Sujoy Dey. "Dynamic Hand Gesture Pattern Recognition Using Probabilistic Neural Network." en. In: *2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS)*. Toronto, ON, Canada: IEEE, Apr. 2021, pp. 1–4. ISBN: 978-1-6654-4067-7. DOI: 10.1109/IEMTRONICS52119.2021.9422496. URL: https://ieeexplore.ieee.org/document/9422496/ (Last accessed on Nov. 20, 2023).

[21] Marília Barandas, Duarte Folgado, Letícia Fernandes, Sara Santos, Mariana Abreu, Patrícia Bota, Hui Liu, Tanja Schultz, and Hugo Gamboa. "TSFEL: Time Series Feature Extraction Library." In: *SoftwareX* 11 (2020). Publisher: Elsevier, p. 100456.

[22] Richard G Baraniuk, Volkan Cevher, and Michael B Wakin. "Low-dimensional models for dimensionality reduction and signal recovery: A geometric perspective." In: *Proceedings of the IEEE* 98.6 (2010). Publisher: IEEE, pp. 959–971.

[23] Robbin Battison. "Lexical borrowing in American sign language." Publisher: ERIC. PhD thesis. 1978.

[24] Auguste Bébian. *Mimographie, ou essai d'écriture mimique propre á régulariser le langage des sourds-muets*. L. Colas, 1825.

[25] S. Benatti, B. Milosevic, F. Casamassima, P. Schonle, P. Bunjaku, S. Fateh, Q. Huang, and L. Benini. "EMG-based hand gesture recognition with flexible analog front end." en. In: *2014 IEEE Biomedical Circuits and Systems Conference (BioCAS) Proceedings*. Lausanne, Switzerland: IEEE, Oct. 2014, pp. 57–60. ISBN: 978-1-4799-2346-5. DOI: 10.1109/BioCAS.2014.6981644. URL: http://ieeexplore.ieee.org/document/6981644/ (Last accessed on Feb. 13, 2024).

[26] Simone Benatti, Filippo Casamassima, Bojan Milosevic, Elisabetta Farella, Philipp Schonle, Schekeb Fateh, Thomas Burger, Qiuting Huang, and Luca Benini. "A Versatile Embedded Platform for EMG Acquisition and Gesture Recognition." en. In: *IEEE Transactions on Biomedical Circuits and Systems* 9.5 (Oct. 2015), pp. 620–630. ISSN: 1932-4545, 1940-9990. DOI: 10.1109/TBCAS.2015.2476555. URL: http://ieeexplore.ieee.org/document/7303979/ (Last accessed on Jan. 17, 2024).

[27] Simone Benatti, Fabio Montagna, Victor Kartsch, Abbas Rahimi, Davide Rossi, and Luca Benini. "Online Learning and Classification of EMG-Based Gestures on a Parallel Ultra-Low Power Platform Using Hyperdimensional Computing." en. In: *IEEE Transactions on Biomedical Circuits and Systems* 13.3 (June 2019), pp. 516–528. ISSN: 1932-4545, 1940-9990. DOI: 10.1109/TBCAS.2019.2914476. URL: https://ieeexplore.ieee.org/document/8704957/ (Last accessed on Jan. 19, 2024).

[28] Rasel Ahmed Bhuiyan, Abdul Kawsar Tushar, Akm Ashiquzzaman, Jungpil Shin, and Md Rashedul Islam. "Reduction of gesture feature dimension for improving the hand gesture recognition performance of numerical sign language." en. In: *2017 20th International Conference of Computer and Information Technology (ICCIT)*. Dhaka: IEEE, Dec. 2017, pp. 1–6. ISBN: 978-1-5386-1150-0. DOI: 10.1109/ICCITECHN.2017.8281833. URL: http://ieeexplore.ieee.org/document/8281833/ (Last accessed on Jan. 4, 2024).

[29] Felix Bigand, Elise Prigent, Bastien Berret, and Annelies Braffort. "How Fast Is Sign Language? A Reevaluation of the Kinematic Bandwidth Using Motion Capture." en. In: *2021 29th European Signal Processing Conference (EUSIPCO)*. Dublin, Ireland: IEEE, Aug. 2021, pp. 711–715. ISBN: 978-90-827970-6-0. DOI: 10.23919/EUSIPCO54536.2021.9616097. URL: https://ieeexplore.ieee.org/document/9616097/ (Last accessed on Mar. 1, 2025).

[30] Penny Boyes Braem. *Einführung in die Gebärdensprache und ihre Erforschung*. de. 3., überarb. aufl. Internationale Arbeiten zur Gebärdensprache und Kommunikation Gehörloser 11. Hamburg: Signum-Verl, 1995. ISBN: 978-3-927731-10-3.

[31] Leo Breiman. "Random Forests." In: *Machine Learning* 45 (Oct. 2001), pp. 5–32.

[32] Diane Brentari. *A prosodic model of sign language phonology*. en. Language, speech, and communication. Cambridge, Mass: MIT Press, 1998. ISBN: 978-0-262-02445-7.

[33] JanFizza Bukhari, Maryam Rehman, Saman Ishtiaq Malik, Awais M. Kamboh, and Ahmad Salman. "American Sign Language Translation through Sensory Glove; SignSpeak." en. In: *International Journal of u- and e-Service, Science and Technology* 8.1 (Jan. 2015), pp. 131–142. ISSN: 20054246. DOI: 10.14257/ijunesst.2015.8.1.12. URL: http://article.nadiapub.com/IJUNESST/vol8_no1/12.pdf (Last accessed on Nov. 20, 2023).

[34] Manuel Caeiro-Rodríguez, Iván Otero-González, Fernando A. Mikic-Fonte, and Martín Llamas-Nistal. "A Systematic Review of Commercial Smart Gloves: Current Status and Applications." en. In: *Sensors* 21.8 (Apr. 2021), p. 2667. ISSN: 1424-8220. DOI: 10.3390/s21082667. URL: https://www.mdpi.com/1424-8220/21/8/2667 (Last accessed on May 27, 2025).

[35] Alexandre Calado, Vito Errico, and Giovanni Saggio. "Toward the Minimum Number of Wearables to Recognize Signer-Independent Italian Sign Language With Machine-Learning Algorithms." en. In: *IEEE Transactions on Instrumentation and Measurement* 70 (2021), pp. 1–9. ISSN: 0018-9456, 1557-9662. DOI: 10.1109

/TIM.2021.3109732. URL: https://ieeexplore.ieee.org/document/9528301/ (Last accessed on Jan. 9, 2024).

[36] Alexandre Calado, Paolo Roselli, Vito Errico, Nathan Magrofuoco, Jean Vanderdonckt, and Giovanni Saggio. "A Geometric Model-Based Approach to Hand Gesture Recognition." en. In: *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 52.10 (Oct. 2022), pp. 6151–6161. ISSN: 2168-2216, 2168-2232. DOI: 10.1109/TSMC.2021.3138589. URL: https://ieeexplore.ieee.org/document/9672762/ (Last accessed on Jan. 11, 2024).

[37] Cambridge University Press. *Cambridge Dictonary*. Englisch. Online Dictonary. URL: https://dictionary.cambridge.org/de/ (Last accessed on May 8, 2021).

[38] Cambridge University Press. *Definition of Grapheme*. English. Online Dictonary. URL: https://dictionary.cambridge.org/de/worterbuch/englisch/grapheme (Last accessed on May 8, 2021).

[39] Cambridge University Press. *Definition of Morpheme*. Englisch. Online Dictonary. URL: https://dictionary.cambridge.org/de/worterbuch/englisch/morpheme (Last accessed on May 8, 2021).

[40] Cambridge University Press. *Definition of Phoneme*. Englisch. Online Dictonary. URL: https://dictionary.cambridge.org/de/worterbuch/englisch/phoneme (Last accessed on May 8, 2021).

[41] Naomi K. Caselli, Zed Sevcikova Sehyr, Ariel M. Cohen-Goldberg, and Karen Emmorey. "ASL-LEX: A lexical database of American Sign Language." en. In: *Behavior Research Methods* 49.2 (Apr. 2017), pp. 784–801. ISSN: 1554-3528. DOI: 10.3758/s13428-016-0742-0. URL: http://link.springer.com/10.3758/s13428-016-0742-0 (Last accessed on Aug. 24, 2023).

[42] Polona Caserman, Philipp Achenbach, and Stefan Gobel. "Analysis of Inverse Kinematics Solutions for Full-Body Reconstruction in Virtual Reality." en. In: *2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH)*. Kyoto, Japan: IEEE, Aug. 2019, pp. 1–8. ISBN: 978-1-7281-0300-6. DOI: 10.1109/SeGAH.2019.8882429. URL: https://ieeexplore.ieee.org/document/8882429/ (Last accessed on Jan. 6, 2024).

[43] Steve Cassidy, Onno Crasborn, Henri Nieminen, Wessel Stoop, Micha Hulsbosch, Susan Even, Erwin Komen, and Trevor Johnston. "Signbank: Software to Support Web Based Dictionaries of Sign Language." en. In: *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)* (2018).

[44] Xiaoshi Chen, Li Gong, Liang Wei, Shih-Ching Yeh, Li Da Xu, Lirong Zheng, and Zhuo Zou. "A Wearable Hand Rehabilitation System With Soft Gloves." en. In: *IEEE Transactions on Industrial Informatics* 17.2 (Feb. 2021), pp. 943–952. ISSN: 1551-3203, 1941-0050. DOI: 10.1109/TII.2020.3010369. URL: https://ieeexplore.ieee.org/document/9146885/ (Last accessed on Nov. 26, 2023).

[45] Yen-Ting Chen and Kuo-Tsung Tseng. "Multiple-angle Hand Gesture Recognition by Fusing SVM Classifiers." In: *2007 IEEE International Conference on Automation Science and Engineering*. 2007, pp. 527–530. DOI: 10.1109/COASE.2007.4341729.

[46] Chih-Wei Hsu and Chih-Jen Lin. "A comparison of methods for multiclass support vector machines." en. In: *IEEE Transactions on Neural Networks* 13.2 (Mar. 2002), pp. 415–425. ISSN: 10459227. DOI: 10.1109/72.991427. URL: http://ieeexplore.ieee.org/document/991427/ (Last accessed on May 25, 2025).

[47] Xianzhi Chu, Jiang Liu, and Shigeru Shimamoto. "A Sensor-Based Hand Gesture Recognition System for Japanese Sign Language." en. In: *2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech)*. Nara, Japan: IEEE, Mar. 2021, pp. 311–312. ISBN: 978-1-6654-1875-1. DOI: 10.1109/LifeTech52111.2021.9391981. URL: https://ieeexplore.ieee.org/document/9391981/ (Last accessed on Apr. 1, 2024).

[48] Wei-Chieh Chuang, Wen-Jyi Hwang, Tsung-Ming Tai, De-Rong Huang, and Yun-Jie Jhang. "Continuous Finger Gesture Recognition Based on Flex Sensors." en. In: *Sensors* 19.18 (Sept. 2019), p. 3986. ISSN: 1424-8220. DOI: 10.3390/s19183986. URL: https://www.mdpi.com/1424-8220/19/18/3986 (Last accessed on Nov. 27, 2023).

[49] Adrean Clark. *How to write American Sign Language*. ASLwrite, 2012.

[50] Frances Conlin-Luippold. "Deaf children's understanding of the language of motion and location in ASL." In: 2015. URL: https://api.semanticscholar.org/CorpusID:62965942.

[51] Corinna Cortes and Vladimir Vapnik. "Support-vector networks." en. In: *Machine Learning* 20.3 (Sept. 1995), pp. 273–297. ISSN: 0885-6125, 1573-0565. DOI: 10.1007/BF00994018. URL: http://link.springer.com/10.1007/BF00994018 (Last accessed on May 25, 2025).

[52] Elaine Costello. *Random House Webster's Concise American Sign Language Dictionary*. Random House, 1999.

[53] Ulysse Cote-Allard, Cheikh Latyr Fall, Alexandre Drouin, Alexandre Campeau-Lecours, Clement Gosselin, Kyrre Glette, Francois Laviolette, and Benoit Gosselin. "Deep Learning for Electromyographic Hand Gesture Signal Classification Using Transfer Learning." en. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 27.4 (Apr. 2019), pp. 760–771. ISSN: 1534-4320, 1558-0210. DOI: 10.1109/TNSRE.2019.2896269. URL: https://ieeexplore.ieee.org/document/8630679/ (Last accessed on Feb. 15, 2024).

[54] Christopher Michael Jason Cox, Ben Hicks, James Gopsill, and Chris Snider. "From haptic interaction to design insight: An empirical comparison of commercial hand-tracking technology." en. In: *Proceedings of the Design Society* 3 (July 2023), pp. 1965–1974. ISSN: 2732-527X. DOI: 10.1017/pds.2023.197. URL: https://www.cambridge.org/core/product/identifier/S2732527X23001979/type/journal_article (Last accessed on Dec. 6, 2023).

[55] Padraig Cunningham, Matthieu Cord, and Sarah Jane Delany. "Supervised Learning." In: *Machine Learning Techniques for Multimedia*. Springer, 2008, pp. 21–49. URL: https://link.springer.com/chapter/10.1007/978-3-540-75171-7_2.

[56] Miguel De Carvalho. "Mean, What do You Mean?" en. In: *The American Statistician* 70.3 (July 2016), pp. 270–274. ISSN: 0003-1305, 1537-2731. DOI: 10.1080/00031305.2016.1148632. URL: https://www.tandfonline.com/doi/full/10.1080/00031305.2016.1148632 (Last accessed on Apr. 15, 2025).

[57] Delsys Inc. *Trigno Wireless Biofeedback System User's Guide*. MAN-031-1-7. Natick, MA, USA: Delsys Incorporated, 2023. URL: https://delsys.com/downloads/USERSGUIDE/MAN-031-1-7%20Trigno%20Wireless%20Biofeedback%20System.pdf (Last accessed on May 18, 2025).

[58] *Deutsche Gebärdensprache lernen*. ger. 1. Auflage. Berlin: Discendum Linguarum, 2020. ISBN: 979-8-6921-8548-8.

[59] Bundesrepublik Deutschland. *Gesetz zur Umsetzung der Richtlinie (EU) 2019/882 des Europäischen Parlaments und des Rates über die Barrierefreiheitsanforderungen für Produkte und Dienstleistungen und zur Änderung anderer Gesetze*. de. June 2025. URL: https://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&jumpTo=bgbl121s1387.pdf (Last accessed on July 20, 2025).

[60] Yongfeng Dong, Jielong Liu, and Wenjie Yan. "Dynamic Hand Gesture Recognition Based on Signals From Specialized Data Glove and Deep Learning Algorithms." en. In: *IEEE Transactions on Instrumentation and Measurement* 70 (2021), pp. 1–14. ISSN: 0018-9456, 1557-9662. DOI: 10.1109/TIM.2021.3077967. URL: https://ieeexplore.ieee.org/document/9424596/ (Last accessed on Nov. 27, 2023).

[61] M. Druga. "Pokémon GO: Where VR and AR have gone since its inception." In: *IEEE Potentials* 37.1 (Jan. 2018), pp. 23–26. ISSN: 0278-6648. DOI: 10.1109/MPOT.2017.2675498.

[62] Yu Du, Yongkang Wong, Wenguang Jin, Wentao Wei, Yu Hu, Mohan Kankanhalli, and Weidong Geng. "Semi-Supervised Learning for Surface EMG-based Gesture Recognition." en. In: *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*. Melbourne, Australia: International Joint Conferences on Artificial Intelligence Organization, Aug. 2017, pp. 1624–1630. ISBN: 978-0-9992411-0-3. DOI: 10.24963/ijcai.2017/225. URL: https://www.ijcai.org/proceedings/2017/225 (Last accessed on Feb. 15, 2024).

[63] Shengshun Duan, Yucheng Lin, Chenyu Zhang, Yinghui Li, Di Zhu, Jun Wu, and Wei Lei. "Machine-learned, waterproof MXene fiber-based glove platform for underwater interactivities." en. In: *Nano Energy* 91 (Jan. 2022), p. 106650. ISSN: 22112855. DOI: 10.1016/j.nanoen.2021.106650. URL: https://linkinghub.elsevier.com/retrieve/pii/S2211285521009010 (Last accessed on Jan. 15, 2024).

[64] Debeshi Dutta, Srinivasan Aruchamy, Soumen Mandal, and Soumen Sen. "Post-stroke Grasp Ability Assessment Using an Intelligent Data Glove Based on Action Research Arm Test: Development, Algorithms, and Experiments." en. In: *IEEE Transactions on Biomedical Engineering* 69.2 (Feb. 2022), pp. 945–954. ISSN: 0018-9294, 1558-2531. DOI: 10.1109/TBME.2021.3110432. URL: https://ieeexplore.ieee.org/document/9531064/ (Last accessed on Jan. 11, 2024).

[65] Petra Eccarius and Diane Brentari. "Symmetry and dominance: A cross-linguistic study of signs and classifier constructions." en. In: *Lingua* 117.7 (July 2007), pp. 1169–1201. ISSN: 00243841. DOI: 10.1016/j.lingua.2005.04.006. URL: https://linkinghub.elsevier.com/retrieve/pii/S0024384106000210 (Last accessed on Nov. 16, 2023).

[66] Petra Eccarius and Diane Brentari. "Handshape coding made easier." en. In: (2008).

[67] Sarah Etta Eiffert. "Samuel Supalla and the ASL-phabet." en. In: (Dec. 2012).

[68] Bin Fang, Fuchun Sun, Huaping Liu, and Chunfang Liu. "3D human gesture capturing and recognition by the IMMU-based data glove." en. In: *Neurocomputing* 277 (Feb. 2018), pp. 198–207. ISSN: 09252312. DOI: 10.1016/j.neucom.2017.02.101. URL: https://linkinghub.elsevier.com/retrieve/pii/S0925231217314054 (Last accessed on Dec. 4, 2023).

[69] Thomas Feix. *Anthropomorphic hand optimization based on a latent space analysis*. na, 2011.

[70] Wanjun Feng, Chunming Xia, Yue Zhang, Jing Yu, and Wendu Jiang. "Research on Chinese Sign Language Recognition Methods Based on Mechanomyogram Signals Analysis." en. In: *2019 IEEE 4th International Conference on Signal and Image Processing (ICSIP)*. Wuxi, China: IEEE, July 2019, pp. 46–50. ISBN: 978-1-7281-3660-8. DOI: 10.1109/SIPROCESS.2019.8868884. URL: https://ieeexplore.ieee.org/document/8868884/ (Last accessed on Mar. 4, 2024).

[71] Ellen Fricke and Jana Bressem. *Gesten - gestern, heute, übermorgen. Vom Forschungsprojekt zur Ausstellung*. Deutsch. Chemnitz: Universitätsverlag Chemnitz, Dec. 2020. URL: https://nbn-resolving.org/urn:nbn:de:bsz:ch1-qucosa2-339590.

[72] Mikel Galar, Alberto Fernández, Edurne Barrenechea, Humberto Bustince, and Francisco Herrera. "An overview of ensemble methods for binary classifiers in multi-class problems: Experimental study on one-vs-one and one-vs-all schemes." In: *Pattern Recognition* 44.8 (Aug. 2011), pp. 1761–1776. URL: https://www.sciencedirect.com/science/article/pii/S0031320311000458?casa_token=0fnyMYtHKPkAAAAA:B-cKxtQDzK6SqxEFW4AUnuEGbhreTjOlptLchlXT-QBjf1clVNDuO9_BERRLidJWltYZNAmibjM.

[73] Jakub Galka, Mariusz Masior, Mateusz Zaborski, and Katarzyna Barczewska. "Inertial Motion Sensing Glove for Sign Language Gesture Acquisition and Recognition." en. In: *IEEE Sensors Journal* 16.16 (Aug. 2016), pp. 6310–6316. ISSN: 1530-437X, 1558-1748, 2379-9153. DOI: 10.1109/JSEN.2016.2583542. URL:

http://ieeexplore.ieee.org/document/7497574/ (Last accessed on Jan. 21, 2024).

[74] *Gebärdensprache lernen - Wörterbuch: 400 praktische Vokabeln - der illustrierte Grundwortschatz der Deutschen Gebärdensprache (DGS) inkl. Fingeralphabet - Gebärdensprache lernen für Anfänger*. ger sgn. 1. Auflage. Berlin: Discendum Linguarum, 2021. ISBN: 979-8-5009-0535-2.

[75] Benyamin Ghojogh and Mark Crowley. *The Theory Behind Overfitting, Cross Validation, Regularization, Bagging, and Boosting: Tutorial*. en. arXiv:1905.12787 [cs, stat]. May 2019. URL: http://arxiv.org/abs/1905.12787 (Last accessed on Feb. 3, 2023).

[76] Davoud Gholamiangonabadi, Nikita Kiselov, and Katarina Grolinger. "Deep Neural Networks for Human Activity Recognition With Wearable Sensors: Leave-One-Subject-Out Cross-Validation for Model Selection." en. In: *IEEE Access* 8 (2020), pp. 133982–133994. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2020.3010715. URL: https://ieeexplore.ieee.org/document/9144538/ (Last accessed on May 25, 2025).

[77] Susan Goldin-Meadow. "The role of gesture in communication and thinking." en. In: *Trends in Cognitive Sciences* 3.11 (Nov. 1999), pp. 419–429. ISSN: 13646613. DOI: 10.1016/S1364-6613(99)01397-2. URL: https://linkinghub.elsevier.com/retrieve/pii/S1364661399013972 (Last accessed on Sept. 19, 2023).

[78] Google LLC. *MediaPipe*. English. Documentation. 2023. URL: https://developers.google.com/mediapipe (Last accessed on Aug. 28, 2023).

[79] Lalit Goyal. "Review and Comparison of Writing Notations of Sign Language." In: *International Journal of Engineering Sciences (IJoES)* (2015).

[80] Verónica Gracia-Ibáñez, Margarita Vergara, Joaquín L. Sancho-Bru, Marta C. Mora, and Catalina Piqueras. "Functional range of motion of the hand joints in activities of the International Classification of Functioning, Disability and Health." en. In: *Journal of Hand Therapy* 30.3 (July 2017), pp. 337–347. ISSN: 08941130. DOI: 10.1016/j.jht.2016.08.001. URL: https://linkinghub.elsevier.com/retrieve/pii/S0894113016301284 (Last accessed on July 28, 2025).

[81] Ivan Grishchenko and Valentin Bazarevsky. "Mediapipe holistic—simultaneous face, hand and pose prediction, on device." In: *Retrieved June* 15 (2020), p. 2021. URL: https://blog.research.google/2020/12/mediapipe-holistic-simultaneous-face.html.

[82] Florian Grützmacher, Johann-Peter Wolff, and Christian Haubelt. "Exploiting thread-level parallelism in template-based gesture recognition with dynamic time warping." en. In: *Proceedings of the 2nd international Workshop on Sensor-based Activity Recognition and Interaction*. Rostock Germany: ACM, June 2015, pp. 1–6. ISBN: 978-1-4503-3454-9. DOI: 10.1145/2790044.2790050. URL: https://dl.acm.org/doi/10.1145/2790044.2790050 (Last accessed on Dec. 30, 2023).

[83] Xiaopei Guo, Zhiquan Feng, Changsheng Ai, Yingjun Li, Jun Wei, Xiaohui Yang, and Kaiyun Sun. "A Novel Method for Data Glove-Based Dynamic Gesture Recognition." en. In: *2017 International Conference on Virtual Reality and Visualization (ICVRV)*. Zhengzhou, China: IEEE, Oct. 2017, pp. 43–48. ISBN: 978-1-5386-2636-8. DOI: 10.1109/ICVRV.2017.00018. URL: https://ieeexplore.ieee.org/document/8719173/ (Last accessed on Jan. 1, 2024).

[84] HackMotion. *Wrist Action in the Golf Swing*. Mar. 2025. URL: https://hackmotion.com/wrist-action-in-golf-swing/.

[85] Kathleen Hall, Scott Mackie, Michael Fry, and Oksana Tkachman. "SLPAnnotator: Tools for Implementing Sign Language Phonetic Annotation." In: Aug. 2017, pp. 2083–2087. DOI: 10.21437/Interspeech.2017-636.

[86] Neta Haluts, Massimiliano Trippa, Naama Friedmann, and Alessandro Treves. "Professional or Amateur? The Phonological Output Buffer as a Working Memory Operator." en. In: *Entropy* 22.6 (June 2020), p. 662. ISSN: 1099-4300. DOI: 10.3390/e22060662. URL: https://www.mdpi.com/1099-4300/22/6/662 (Last accessed on Jan. 31, 2024).

[87] *HamNoSys Platfom independent input palette*. 2024. URL: https://www.sign-lang.uni-hamburg.de/hamnosys/input/.

[88] Thomas Hanke. "Auf dem Wege zu Language Resources für Gebärdensprachen." de. In: (2001). URL: https://api.semanticscholar.org/CorpusID:63361557.

[89] Thomas Hanke. "HamNoSys – Representing Sign Language Data in Language Resources and Language Processing Contexts." en. In: (May 2004).

[90] Anke Hannig. "Erstellung eines systematischen Wortschatzes für die automatische Gebärdenspracherkennung." de. In: ().

[91] Shelly Hansen. *Signotation: ASL writer's guide : how to write sign language using the 5 parameters*. eng. OCLC: 1037353211. Place of publication not identified: ASLiSH, 2017. ISBN: 978-0-9987186-1-3.

[92] Endi Sailul Haq, Devit Suwardiyanto, and Miftahul Huda. "Indonesian Sign Language Recognition Application For Two-Way Communication Deaf-Mute People." en. In: *2018 3rd International Conference on Information Technology, Information System and Electrical Engineering (ICITISEE)*. Yogyakarta, Indonesia: IEEE, Nov. 2018, pp. 313–318. ISBN: 978-1-5386-7082-8. DOI: 10.1109/ICITISEE.2018.8720982. URL: https://ieeexplore.ieee.org/document/8720982/ (Last accessed on Jan. 21, 2024).

[93] Jonathan Henner, Leah C. Geer, and Diane Lillo-Martin. "Calculating Frequency of Occurrence of ASL handshapes." en. In: *LSA Annual Meeting Extended Abstracts* 4 (May 2013), p. 16. ISSN: 2377-3367. DOI: 10.3765/exabs.v0i0.764. URL: http://journals.linguisticsociety.org/proceedings/index.php/ExtendedAbs/article/view/764 (Last accessed on Jan. 30, 2024).

[94] Jose L. Hernandez-Rebollar, Nicholas Kyriakopoulos, and Robert W. Lindeman. "The AcceleGlove: a whole-hand input device for virtual reality." en. In: *ACM SIGGRAPH 2002 conference abstracts and applications*. San Antonio Texas: ACM, July 2002, pp. 259–259. ISBN: 978-1-58113-525-1. DOI: 10.1145/1242073.1242272. URL: https://dl.acm.org/doi/10.1145/1242073.1242272 (Last accessed on Mar. 7, 2024).

[95] Julie Hochgesang. *Introduction to Stokoe Notation*. 2007.

[96] Julie Hochgesang. "SLAASh ID Glossing Principles, ASL Signbank and Annotation Conventions, Version 3.2." In: (2022). Artwork Size: 0 Bytes Publisher: figshare, 0 Bytes. DOI: 10.6084/M9.FIGSHARE.12003732.V4. URL: https://figshare.com/articles/online_resource/SLAASh_ID_Glossing_Principles_ASL_Signbank_and_Annotation_Conventions_Version_3_0/12003732/4 (Last accessed on Feb. 27, 2024).

[97] Julie Hochgesang. *Download possibility of SignBank*. July 2023.

[98] Julie Hochgesang, Onno Crasborn, and Diane Lillo-Martin. "ASL Signbank." In: (2019). Artwork Size: 0 Bytes Publisher: figshare, 0 Bytes. DOI: 10.6084/M9.FIGSHARE.9741788.V1. URL: https://figshare.com/articles/online_resource/ASL_Signbank/9741788/1 (Last accessed on Feb. 27, 2024).

[99] Julie Hochgesang and Diane Lillo-Martin. "Sign Language Acquisition, Annotation, Archiving and Sharing (SLAAASh)." In: (2021). Publisher: figshare. DOI: 10.6084/M9.FIGSHARE.C.5236682.V2. URL: https://figshare.com/collections/Sign_Language_Acquisition_Annotation_Archiving_and_Sharing_SLAAASh_/5236682/2 (Last accessed on Feb. 27, 2024).

[100] Julie A Hochgesang. "Using design principles to consider representation of the hand in some notation systems." In: *Sign Language Studies* 14.4 (2014). Publisher: JSTOR, pp. 488–542.

[101] Chih-Wei Hsu, Chih-Chung Chang, and Chih-Jen Lin. *A Practical Guide to Support Vector Classification*. 2003.

[102] Xin'an Huang, Qi Wang, Siyao Zang, Jiaxin Wan, Guang Yang, Yongqing Huang, and Xiaomin Ren. "Tracing the Motion of Finger Joints for Gesture Recognition via Sewing RGO-Coated Fibers Onto a Textile Glove." en. In: *IEEE Sensors Journal* 19.20 (Oct. 2019), pp. 9504–9511. ISSN: 1530-437X, 1558-1748, 2379-9153. DOI: 10.1109/JSEN.2019.2924797. URL: https://ieeexplore.ieee.org/document/8744608/ (Last accessed on Jan. 2, 2024).

[103] Harry van der Hulst. "The composition of handshapes." In: *Trondheim Work. Papers* 23 (1995), pp. 1–17.

[104] Jessica Jeanne Hutchinson. "Analysis of Notation Systems for Machine Translation of Sign Languages." Bachelor of Science. Rhodes, USA: Rhodes University, Nov. 2012.

[105] George Hutton. "The practicability and advantages of writing and printing natural signs." In: *American Annals of the Deaf and Dumb* 14.3 (1869). Publisher: JSTOR, pp. 157–182.

[106] Sohal Islam, Showni Rudra Titli, Kazi Arham Kabir, Md Abdullah Al Hossain, and Md. Altaf Hossain. "Improving Real-time Hand Gesture Recognition System for Translation: Sensor Development." en. In: *2022 17th Annual System of Systems Engineering Conference (SOSE)*. Rochester, NY, USA: IEEE, June 2022, pp. 254–259. ISBN: 978-1-6654-9623-0. DOI: 10.1109/SOSE55472.2022.9812664. URL: https://ieeexplore.ieee.org/document/9812664/ (Last accessed on Feb. 4, 2024).

[107] Abhishek B. Jani, Nishith A. Kotak, and Anil K. Roy. "Sensor Based Hand Gesture Recognition System for English Alphabets Used in Sign Language of Deaf-Mute People." en. In: *2018 IEEE SENSORS*. New Delhi: IEEE, Oct. 2018, pp. 1–4. ISBN: 978-1-5386-4707-3. DOI: 10.1109/ICSENS.2018.8589574. URL: https://ieeexplore.ieee.org/document/8589574/ (Last accessed on Jan. 21, 2024).

[108] Shuo Jiang, Bo Lv, Weichao Guo, Chao Zhang, Haitao Wang, Xinjun Sheng, and Peter B. Shull. "Feasibility of Wrist-Worn, Real-Time Hand, and Surface Gesture Recognition via sEMG and IMU Sensing." en. In: *IEEE Transactions on Industrial Informatics* 14.8 (Aug. 2018), pp. 3376–3385. ISSN: 1551-3203, 1941-0050. DOI: 10.1109/TII.2017.2779814. URL: https://ieeexplore.ieee.org/document/8141959/ (Last accessed on Feb. 17, 2024).

[109] Robert E. Johnson and Scott K. Liddell. "A Segmental Framework for Representing Signs Phonetically." en. In: *Sign Language Studies* 11.3 (Mar. 2011), pp. 408–463. ISSN: 1533-6263. DOI: 10.1353/sls.2011.0002. URL: https://muse.jhu.edu/article/425452 (Last accessed on Jan. 24, 2024).

[110] Alexey Kachan. *Request from [senso.me]*. Englisch. June 2025.

[111] Takahiro Kanokoda, Yuki Kushitani, Moe Shimada, and Jun-ichi Shirakashi. "Gesture Prediction Using Wearable Sensing Systems with Neural Networks for Temporal Data Analysis." en. In: *Sensors* 19.3 (Feb. 2019), p. 710. ISSN: 1424-8220. DOI: 10.3390/s19030710. URL: http://www.mdpi.com/1424-8220/19/3/710 (Last accessed on Jan. 3, 2024).

[112] I. A. Kapandji. *The physiology of the joints*. eng. 6th ed., English ed. Edinburgh ; New York: Churchill Livingstone, 2007. ISBN: 978-0-443-10350-6 978-0-7020-3942-3 978-0-7020-2959-2.

[113] Lih-Jen Kau, Wan-Lin Su, Pei-Ju Yu, and Sin-Jhan Wei. "A real-time portable sign language translation system." en. In: *2015 IEEE 58th International Midwest Symposium on Circuits and Systems (MWSCAS)*. Fort Collins, CO, USA: IEEE, Aug. 2015, pp. 1–4. ISBN: 978-1-4673-6558-1. DOI: 10.1109/MWSCAS.2015.7282137. URL: http://ieeexplore.ieee.org/document/7282137/ (Last accessed on Jan. 21, 2024).

[114] Adam Kendon. "How gestures can become like words." In: *Crosscultural Perspectives in Nonverbal Communication* (Jan. 1988).

[115] Kenhub GmbH. *Metacarpophalangeal (MCP) joints*. Publisher: Kenhub. URL: https://www.kenhub.com/en/library/anatomy/metacarpophalangeal-mcp-joints.

[116] Talha Ali Khan and Sai Ho Ling. "Review on Electrical Impedance Tomography: Artificial Intelligence Methods and its Applications." en. In: *Algorithms* 12.5 (Apr. 2019), p. 88. ISSN: 1999-4893. DOI: 10.3390/a12050088. URL: https://www.mdpi.com/1999-4893/12/5/88 (Last accessed on Mar. 8, 2024).

[117] Chunghwan Kim, Ho-Seung Cha, Junghwan Kim, HwyKuen Kwak, WooJin Lee, and Chang-Hwan Im. "Facial Motion Capture System Based on Facial Electromyogram and Electrooculogram for Immersive Social Virtual Reality Applications." en. In: *Sensors* 23.7 (Mar. 2023), p. 3580. ISSN: 1424-8220. DOI: 10.3390/s23073580. URL: https://www.mdpi.com/1424-8220/23/7/3580 (Last accessed on Dec. 18, 2023).

[118] Keith Kirkpatrick. "Technology for the deaf." en. In: *Communications of the ACM* 61.12 (Nov. 2018), pp. 16–18. ISSN: 0001-0782, 1557-7317. DOI: 10.1145/3283224. URL: https://dl.acm.org/doi/10.1145/3283224 (Last accessed on Mar. 7, 2024).

[119] Sotaro Kita, Ingeborg Van Gijn, and Harry Van Der Hulst. "The non-linguistic status of the Symmetry Condition in signed languages: Evidence from a comparison of signs and speech-accompanying representational gestures." en. In: *Sign Language & Linguistics* 17.2 (Dec. 2014), pp. 215–238. ISSN: 1387-9316, 1569-996X. DOI: 10.1075/sll.17.2.04kit. URL: http://www.jbe-platform.com/content/journals/10.1075/sll.17.2.04kit (Last accessed on Jan. 9, 2024).

[120] Chemuttaai C Koech. "A kinematic analysis of sign language." en. In: (Jan. 2007).

[121] Sussane König. "SignWriting." In: *Wörterbücher zur Sprach- und Kommunikationswissenschaft (WSK) Online*. Ed. by Stefan J. Schierholz and Laura Giacomini. Berlin, Boston: De Gruyter, 2013. URL: https://www.degruyter.com/database/WSK/entry/wsk_id_wsk_artikel_artikel_21755/html (Last accessed on Jan. 22, 2024).

[122] Jim G Kyle, James Kyle, Bencie Woll, G Pullen, and F Maddix. *Sign language: The study of deaf people and their language*. Cambridge university press, 1988.

[123] Boon Giin Lee and Su Min Lee. "Smart Wearable Hand Device for Sign Language Interpretation System With Sensors Fusion." en. In: *IEEE Sensors Journal* 18.3 (Feb. 2018), pp. 1224–1232. ISSN: 1530-437X, 1558-1748, 2379-9153. DOI: 10.1109/JSEN.2017.2779466. URL: http://ieeexplore.ieee.org/document/8126796/ (Last accessed on Jan. 21, 2024).

[124] Minhyuk Lee and Joonbum Bae. "Real-Time Gesture Recognition in the View of Repeating Characteristics of Sign Languages." en. In: *IEEE Transactions on Industrial Informatics* 18.12 (Dec. 2022), pp. 8818–8828. ISSN: 1551-3203, 1941-0050. DOI: 10.1109/TII.2022.3152214. URL: https://ieeexplore.ieee.org/document/9716737/ (Last accessed on Jan. 12, 2024).

[125] Dong-Jie Li, Yang-Yang Li, Jun-Xiang Li, and Yu Fu. "Gesture Recognition Based on BP Neural Network Improved by Chaotic Genetic Algorithm." In: *International Journal of Automation and Computing* 15.3 (2018), pp. 267–276. URL: https://link.springer.com/article/10.1007/s11633-017-1107-6.

[126] Dongxu Li, Cristian Rodriguez Opazo, Xin Yu, and Hongdong Li. *Word-level Deep Sign Language Recognition from Video: A New Large-scale Dataset and Methods Comparison*. en. arXiv:1910.11006 [cs]. Jan. 2020. URL: http://arxiv.org/abs/1910.11006 (Last accessed on Jan. 14, 2024).

[127] lifeprint.com. *American Sign Language: "Sunday"*. English. Jan. 2024. URL: https://www.lifeprint.com/asl101/pages-signs/s/sunday.htm (Last accessed on Jan. 31, 2024).

[128] Zhirong Lin, Yongsheng Xiong, Houde Dai, and Xuke Xia. "An Experimental Performance Evaluation of the Orientation Accuracy of Four Nine-Axis MEMS Motion Sensors." en. In: *2017 5th International Conference on Enterprise Systems (ES)*. Beijing: IEEE, Sept. 2017, pp. 185–189. ISBN: 978-1-5386-0936-1. DOI: 10.1109/ES.2017.37. URL: http://ieeexplore.ieee.org/document/8119388/ (Last accessed on Oct. 25, 2025).

[129] Shuangjun Liu and Sarah Ostadabbas. "A Semi-supervised Data Augmentation Approach Using 3D Graphical Engines." en. In: *Computer Vision – ECCV 2018 Workshops*. Ed. by Laura Leal-Taixé and Stefan Roth. Vol. 11130. Series Title: Lecture Notes in Computer Science. Cham: Springer International Publishing, 2019, pp. 395–408. ISBN: 978-3-030-11011-6 978-3-030-11012-3. DOI: 10.1007/978-3-030-11012-3_31. URL: http://link.springer.com/10.1007/978-3-030-11012-3_31 (Last accessed on June 8, 2022).

[130] Xilin Liu, Jacob Sacks, Milin Zhang, Andrew G. Richardson, Timothy H. Lucas, and Jan Van Der Spiegel. "The Virtual Trackpad: An Electromyography-Based, Wireless, Real-Time, Low-Power, Embedded Hand-Gesture-Recognition System Using an Event-Driven Artificial Neural Network." en. In: *IEEE Transactions on Circuits and Systems II: Express Briefs* 64.11 (Nov. 2017), pp. 1257–1261. ISSN: 1549-7747, 1558-3791. DOI: 10.1109/TCSII.2016.2635674. URL: https://ieeexplore.ieee.org/document/7769179/ (Last accessed on Feb. 15, 2024).

[131] Yilin Liu, Shijia Zhang, and Mahanth Gowda. "When Video meets Inertial Sensors: Zero-shot Domain Adaptation for Finger Motion Analytics with Inertial Sensors." en. In: *Proceedings of the International Conference on Internet-of-Things Design and Implementation*. Charlottesvle VA USA: ACM, May 2021, pp. 182–194. ISBN: 978-1-4503-8354-7. DOI: 10.1145/3450268.3453537. URL: https://dl.acm.org/doi/10.1145/3450268.3453537 (Last accessed on Mar. 12, 2024).

[132] Danling Lu, Yuanlong Yu, and Huaping Liu. "Gesture recognition using data glove: An extreme learning machine method." en. In: *2016 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. Qingdao, China: IEEE, Dec. 2016, pp. 1349–1354. ISBN: 978-1-5090-4364-4. DOI: 10.1109/ROBIO.2016.7866514. URL: http://ieeexplore.ieee.org/document/7866514/ (Last accessed on Jan. 3, 2024).

[133] Camillo Lugaresi, Jiuqiang Tang, Hadon Nash, Chris McClanahan, Esha Uboweja, Michael Hays, Fan Zhang, Chuo-Ling Chang, Ming Guang Yong, Juhyun Lee, Wan-Teh Chang, Wei Hua, Manfred Georg, and Matthias Grundmann. *MediaPipe: A Framework for Building Perception Pipelines*. en. arXiv:1906.08172 [cs]. June 2019. URL: http://arxiv.org/abs/1906.08172 (Last accessed on Feb. 2, 2024).

[134] Granit Luzhnica, Jorg Simon, Elisabeth Lex, and Viktoria Pammer. "A sliding window approach to natural hand gesture recognition using a custom data glove." en. In: *2016 IEEE Symposium on 3D User Interfaces (3DUI)*. Greenville, SC, USA: IEEE, Mar. 2016, pp. 81–90. ISBN: 978-1-5090-0842-1. DOI: 10.1109/3DUI.2016.7460035. URL: http://ieeexplore.ieee.org/document/7460035/ (Last accessed on Jan. 3, 2024).

[135] Na Lv, Xiaohui Yang, Yan Jiang, and Tao Xu. "Sparse decomposition for data glove gesture recognition." en. In: *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*. Shanghai: IEEE, Oct. 2017, pp. 1–5. ISBN: 978-1-5386-1937-7. DOI: 10.1109/CISP-BMEI.2017.8302114. URL: http://ieeexplore.ieee.org/document/8302114/ (Last accessed on Jan. 4, 2024).

[136] Wenchao Ma, Junfeng Hu, Jun Liao, Zhencheng Fan, Jianjun Wu, and Li Liu. "Finger Gesture Recognition Based on 3D-Accelerometer and 3D-Gyroscope." en. In: *Knowledge Science, Engineering and Management*. Ed. by Christos Douligeris, Dimitris Karagiannis, and Dimitris Apostolou. Vol. 11775. Series Title: Lecture Notes in Computer Science. Cham: Springer International Publishing, 2019, pp. 406–413. ISBN: 978-3-030-29550-9 978-3-030-29551-6. DOI: 10.1007/978-3-030-29551-6_36. URL: http://link.springer.com/10.1007/978-3-030-29551-6_36 (Last accessed on Jan. 13, 2024).

[137] Jane Maher. *Seeing Language in Sign: The Work of William C. Stokoe*. eng. Paperback edition. OCLC: 827787902. Washington, Chicago: Gallaudet University Press Chicago Distribution Center [distributor], Apr. 2010. ISBN: 978-1-56368-243-8.

[138] Joe Martin. "A Linguistic Comparison: Two Notation Systems for Signed Languages." en. In: (Feb. 2000). URL: https://www.signwriting.org/archive/docs1/sw0032-Stokoe-Sutton.pdf.

[139] David McNeill. "Hand and Mind: What Gestures Reveal About Thought." In: *Bibliovault OAI Repository, the University of Chicago Press* 27 (Aug. 1992). DOI: 10.2307/1576015.

[140] R. Meattini, S. Benatti, U. Scarcia, D. De Gregorio, L. Benini, and C. Melchiorri. "An sEMG-Based Human–Robot Interface for Robotic Hands Using Machine Learning and Synergies." en. In: *IEEE Transactions on Components, Packaging and Manufacturing Technology* 8.7 (July 2018), pp. 1149–1158. ISSN: 2156-3950, 2156-3985. DOI: 10.1109/TCPMT.2018.2799987. URL: https://ieeexplore.ieee.org/document/8301546/ (Last accessed on Feb. 13, 2024).

[141] Pedro Miguel Melo. "Gesture Recognition for Human-Robot Collaborative Assembly." en. In: (July 2018).

[142] Jonathan Milgram, Mohamed Cheriet, and Robert Sabourin. ""One Against One" or "One Against All": Which One is Better for Handwriting Recognition with SVMs?" en. In: (Oct. 2006).

[143] Hongrui Min, Chao Chen, Shixin Huang, Xiaorui Tian, Yongkui Yang, and Zheng Wang. "Highly-accurate gesture recognition based on ResNet with low-budget data gloves." en. In: *Proceedings of the 3rd International Conference on Advanced Information Science and System*. Sanya China: ACM, Nov. 2021, pp. 1–6. ISBN: 978-1-4503-8586-2. DOI: 10.1145/3503047.3503153. URL: https://dl.acm.org/doi/10.1145/3503047.3503153 (Last accessed on Jan. 4, 2024).

[144] Ross E Mitchell, Travas A Young, Bellamie Bachleda, and Michael A Karchmer. "How Many People Use ASL in the United States? Why Estimates Need Updating." en. In: *Sign Language Studies* 6.3 (2006), pp. 306–335. ISSN: 1533-6263. DOI: 10.1353/sls.2006.0019. URL: http://muse.jhu.edu/content/crossref/journals/sign_language_studies/v006/6.3mitchell.html (Last accessed on Feb. 3, 2023).

[145] M. Mohandes and M. Deriche. "Arabic sign language recognition by decisions fusion using Dempster-Shafer theory of evidence." en. In: *2013 Computing, Communications and IT Applications Conference (ComComAp)*. Hong Kong: IEEE, Apr. 2013, pp. 90–94. ISBN: 978-1-4673-6044-9 978-1-4673-6043-2 978-1-4673-6042-5. DOI: 10.1109/ComComAp.2013.6533615. URL: http://ieeexplore.ieee.org/document/6533615/ (Last accessed on Jan. 4, 2024).

[146] H Morgan, R Novogrodsky, and W Sandler. "Phonological complexity and frequency in the lexicon: A quantitative cross-linguistic study." In: *13th Theoretical Issues in Sign Language Research (TISLR 13)* (2019).

[147] Sara Morrissey. "Data-Driven Machine Translation for Sign Languages." en. PhD thesis. Dublin, Ireland: Dublin City University, Apr. 2008.

[148] Taha Müezzinoğlu and Mehmet Karaköse. "An Intelligent Human–Unmanned Aerial Vehicle Interaction Approach in Real Time Based on Machine Learning Using Wearable Gloves." en. In: *Sensors* 21.5 (Mar. 2021), p. 1766. ISSN: 1424-8220. DOI: 10.3390/s21051766. URL: https://www.mdpi.com/1424-8220/21/5/1766 (Last accessed on Jan. 4, 2024).

[149] Cornelia Müller. "Gesture and Sign: Cataclysmic Break or Dynamic Relations?" In: *Frontiers in Psychology* 9 (2018). ISSN: 1664-1078. DOI: 10.3389/fpsyg.2018.01651. URL: https://www.frontiersin.org/articles/10.3389/fpsyg.2018.01651.

[150] Philipp Niklas Müller, Philipp Achenbach, André Mihca Kleebe, Jan Ulrich Schmitt, Ute Lehmann, Thomas Tregel, and Stefan Göbel. "Flex Your Muscles: EMG-Based Serious Game Controls." en. In: *Serious Games*. Ed. by Minhua Ma, Bobbie Fletcher, Stefan Göbel, Jannicke Baalsrud Hauge, and Tim Marsh. Vol. 12434. Series Title: Lecture Notes in Computer Science. Cham: Springer

International Publishing, 2020, pp. 230–242. ISBN: 978-3-030-61813-1 978-3-030-61814-8. DOI: 10.1007/978-3-030-61814-8_18. URL: https://link.springer.com/10.1007/978-3-030-61814-8_18 (Last accessed on Jan. 6, 2024).

[151] Philipp Niklas Müller, Alexander Josef Müller, Philipp Achenbach, and Stefan Göbel. "IMU-Based Fitness Activity Recognition Using CNNs for Time Series Classification." en. In: *Sensors* 24.3 (Jan. 2024), p. 742. ISSN: 1424-8220. DOI: 10.3390/s24030742. URL: https://www.mdpi.com/1424-8220/24/3/742 (Last accessed on Mar. 9, 2025).

[152] Chaithanya Kumar Mummadi, Frederic Philips Peter Leo, Keshav Deep Verma, Shivaji Kasireddy, Philipp Marcel Scholl, and Kristof Van Laerhoven. "Real-time Embedded Recognition of Sign Language Alphabet Fingerspelling in an IMU-Based Glove." en. In: *Proceedings of the 4th International Workshop on Sensor-based Activity Recognition and Interaction*. Rostock Germany: ACM, Sept. 2017, pp. 1–6. ISBN: 978-1-4503-5223-9. DOI: 10.1145/3134230.3134236. URL: https://dl.acm.org/doi/10.1145/3134230.3134236 (Last accessed on Jan. 4, 2024).

[153] John Nassour, Hojjat Gharaei Amirabadi, Saddam Weheabby, Abbas Al Ali, Heinrich Lang, and Fred Hamker. "A Robust Data-Driven Soft Sensory Glove for Human Hand Motions Identification and Replication." en. In: *IEEE Sensors Journal* 20.21 (Nov. 2020), pp. 12972–12979. ISSN: 1530-437X, 1558-1748, 2379-9153. DOI: 10.1109/JSEN.2020.3001982. URL: https://ieeexplore.ieee.org/document/9115656/ (Last accessed on Jan. 7, 2024).

[154] I. S. P. Nation. *What Should Every ESL Teacher Know?* eng. 2nd print. OCLC: 910271237. Seoul: Compass Publishing, 2014. ISBN: 978-1-59966-266-4.

[155] Carol Neidle and Carey Ballard. *Why alternative gloss labels will increase the value of the WLASL dataset*. 2022. URL: https://open.bu.edu/handle/2144/45441.

[156] Carol Neidle, Carey Ballard, and Dimitris Metaxas. *Revised Gloss Labels for Signs from the WLASL Dataset – Version 2*. en. 2024.

[157] Carol Neidle, Augustine Opoku, Carey Ballard, Konstantinos M Dafnis, Evgenia Chroni, and Dimitri Metaxas. "Resources for Computer-Based Sign Recognition from Video, and the Criticality of Consistency of Gloss Labeling across Multiple Large ASL Video Corpora." en. In: (June 2022).

[158] Carol Neidle, Augustine Opoku, and Dimitris Metaxas. "ASL Video Corpora & Sign Bank: Resources Available through the American Sign Language Linguistic Research Project (ASLLRP)." en. In: ().

[159] Erez Nusem, Karla Straker, and Cara Wrigley. *Design Innovation for Health and Medicine*. en. Singapore: Springer Nature Singapore, 2020. ISBN: 978-981-15-4361-6 978-981-15-4362-3. DOI: 10.1007/978-981-15-4362-3. URL: https://link.springer.com/10.1007/978-981-15-4362-3 (Last accessed on Mar. 7, 2024).

[160] H. Nyquist. "Certain Topics in Telegraph Transmission Theory." In: *Transactions of the American Institute of Electrical Engineers* 47.2 (Apr. 1928), pp. 617–644. ISSN: 0096-3860. DOI: 10.1109/T-AIEE.1928.5055024. URL: http://ieeexplore.ieee.org/document/5055024/ (Last accessed on May 13, 2025).

[161] Cemil Oz and Ming C. Leu. "American Sign Language word recognition with a sensory glove using artificial neural networks." en. In: *Engineering Applications of Artificial Intelligence* 24.7 (Oct. 2011), pp. 1204–1213. ISSN: 09521976. DOI: 10.1016/j.engappai.2011.06.015. URL: https://linkinghub.elsevier.com/retrieve/pii/S0952197611001230 (Last accessed on Jan. 8, 2024).

[162] Rohit Pahwa, Harion Tanwar, and Dr Sachin Sharma. "Speech Recognition System: A review." en. In: *International Journal of Future Generation Communication and Networking* 13.3 (2020).

[163] Sameera Palipana, Dariush Salami, Luis A Leiva, and Stephan Sigg. "Pantomime: Mid-air gesture recognition with sparse millimeter-wave radar point clouds." In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5.1 (2021), pp. 1–27. URL: https://dl.acm.org/doi/abs/10.1145/3448110.

[164] Mingzhang Pan, Yingzhe Tang, and Hongqi Li. "State-of-the-Art in Data Gloves: A Review of Hardware, Algorithms, and Applications." In: *IEEE Transactions on Instrumentation and Measurement* (2023), pp. 1–1. ISSN: 1557-9662. DOI: 10.1109/TIM.2023.3243614.

[165] Chrissostomos Papaspyrou. *Gebärdensprache und universelle Sprachtheorie: Versuch einer vergleichenden generativ-transformationellen Interpretation von Gebärdensprache und Lautsprache sowie der Entwurf einer Gebärdenschrift*. Signum-Verlag, 1990.

[166] Chrissostomos Papaspyrou. *Grammatik der deutschen Gebärdensprache aus der Sicht gehörloser Fachleute*. Signum, 2008.

[167] Prajwal Paudyal, Ayan Banerjee, and Sandeep K.S. Gupta. "SCEPTRE: A Pervasive, Non-Invasive, and Programmable Gesture Recognition Technology." en. In: *Proceedings of the 21st International Conference on Intelligent User Interfaces*. Sonoma California USA: ACM, Mar. 2016, pp. 282–293. ISBN: 978-1-4503-4137-0. DOI: 10.1145/2856767.2856794. URL: https://dl.acm.org/doi/10.1145/2856767.2856794 (Last accessed on Jan. 8, 2024).

[168] Fabian Pedregosa, Gael Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, and David Cournapeau. "Scikit-learn: Machine Learning in Python." en. In: *MACHINE LEARNING IN PYTHON* (Oct. 2011).

[169] Esteve Peña Pitarch, Karim Abdel-Malek, and Anas Al Omar Mesnaoui. "Virtual Human Hand: Grasping Strategy and Simulation." en. PhD thesis. Universitat Politècnica de Catalunya, Jan. 2008. DOI: 10.5821/dissertation-2117-94313. URL: https://hdl.handle.net/2117/94313 (Last accessed on July 28, 2025).

[170] Francesco Pezzuoli, Dario Corona, and Maria Letizia Corradini. "Recognition and Classification of Dynamic Hand Gestures by a Wearable Data-Glove." en. In: *SN Computer Science* 2.1 (Feb. 2021), p. 5. ISSN: 2662-995X, 2661-8907. DOI: 10.1007/s42979-020-00396-5. URL: http://link.springer.com/10.1007/s42979-020-00396-5 (Last accessed on Jan. 8, 2024).

[171] Physiotutors. *Wrist/Hand Active Range of Motion*. Publisher: Physiotutors. URL: https://www.physiotutors.com/wiki/wrist-hand-active-range-of-motion/.

[172] Pawel Plawiak, Tomasz Sosnicki, Michal Niedzwiecki, Zbislaw Tabor, and Krzysztof Rzecki. "Hand Body Language Gesture Recognition Based on Signals From Specialized Glove and Machine Learning Algorithms." en. In: *IEEE Transactions on Industrial Informatics* 12.3 (June 2016), pp. 1104–1113. ISSN: 1551-3203, 1941-0050. DOI: 10.1109/TII.2016.2550528. URL: http://ieeexplore.ieee.org/document/7448427/ (Last accessed on Feb. 3, 2023).

[173] Feng Qi. "Generalized abstracted mean values." en. In: *Journal of Inequalities in Pure and Applied Mathematics* 1 (May 2000), Article 4.

[174] Dennis Reimer, Iana Podkosova, Daniel Scherzer, and Hannes Kaufmann. "Evaluation and improvement of HMD-based and RGB-based hand tracking solutions in VR." en. In: *Frontiers in Virtual Reality* 4 (July 2023), p. 1169313. ISSN: 2673-4192. DOI: 10.3389/frvir.2023.1169313. URL: https://www.frontiersin.org/articles/10.3389/frvir.2023.1169313/full (Last accessed on Feb. 2, 2024).

[175] Brunna Carolinne Rocha Silva, Geovanne Pereira Furriel, Wesley Calixto Pacheco, and Junio Santos Bulhoes. "Methodology and comparison of devices for recognition of sign language characters." en. In: *2017 18th International Scientific Conference on Electric Power Engineering (EPE)*. Kouty nad Desnou: IEEE, May 2017, pp. 1–6. ISBN: 978-1-5090-6406-9. DOI: 10.1109/EPE.2017.7967322. URL: https://ieeexplore.ieee.org/document/7967322/ (Last accessed on Jan. 10, 2024).

[176] Paul D. Rosero-Montalvo, Pamela Godoy-Trujillo, Edison Flores-Bosmediano, Jorge Carrascal-Garcia, Santiago Otero-Potosi, Henry Benitez-Pereira, and Diego H. Peluffo-Ordonez. "Sign Language Recognition Based on Intelligent Glove Using Machine Learning Techniques." en. In: *2018 IEEE Third Ecuador Technical Chapters Meeting (ETCM)*. Cuenca: IEEE, Oct. 2018, pp. 1–5. ISBN: 978-1-5386-6657-9. DOI: 10.1109/ETCM.2018.8580268. URL: https://ieeexplore.ieee.org/document/8580268/ (Last accessed on Jan. 21, 2024).

[177] David Rozado, Francisco B. Rodriguez, and Pablo Varona. "Extending the bioinspired hierarchical temporal memory paradigm for sign language recognition." en. In: *Neurocomputing* 79 (Mar. 2012), pp. 75–86. ISSN: 09252312. DOI: 10.1016/j.neucom.2011.10.005. URL: https://linkinghub.elsevier.com/retrieve/pii/S0925231211006230 (Last accessed on Jan. 9, 2024).

[178] G. Saggio, P. Cavallo, A. Fabrizio, and S. O. Ibe. "Gesture recognition through HITEG data glove to provide a new way of communication." en. In: *Proceedings of the 4th International Symposium on Applied Sciences in Biomedical and Communication Technologies*. Barcelona Spain: ACM, Oct. 2011, pp. 1–5. ISBN: 978-1-4503-0913-4. DOI: 10.1145/2093698.2093711. URL: https://dl.acm.org/doi/10.1145/2093698.2093711 (Last accessed on Jan. 13, 2024).

[179] Giovanni Saggio, Pietro Cavallo, Mariachiara Ricci, Vito Errico, Jonathan Zea, and Marco E. Benalcázar. "Sign Language Recognition Using Wearable Electronics: Implementing k-Nearest Neighbors with Dynamic Time Warping and Convolutional Neural Network Algorithms." en. In: *Sensors* 20.14 (July 2020), p. 3879. ISSN: 1424-8220. DOI: 10.3390/s20143879. URL: https://www.mdpi.com/1424-8220/20/14/3879 (Last accessed on Jan. 10, 2024).

[180] Giovanni Saggio, Luca Pietrosanti, I Lee, Bor-Shing Lin, et al. "Quasi-Static and Dynamic Measurement Performances of an Electromagnetic Field-Based Sensory Glove Termed Manus Quantum Metaglove." In: *Bor-Shing, Quasi-Static and Dynamic Measurement Performances of an Electromagnetic Field-Based Sensory Glove Termed Manus Quantum Metaglove* (Apr. 2025). DOI: 10.2139/ssrn.5208853.

[181] Neven Saleh, Mostafa Farghaly, Eslam Elshaaer, and Amr Mousa. "Smart glove-based gestures recognition system for Arabic sign language." en. In: *2020 International Conference on Innovative Trends in Communication and Computer Engineering (ITCE)*. Aswan, Egypt: IEEE, Feb. 2020, pp. 303–307. ISBN: 978-1-7281-4801-4. DOI: 10.1109/ITCE48509.2020.9047820. URL: https://ieeexplore.ieee.org/document/9047820/ (Last accessed on Jan. 17, 2024).

[182] Wendy Sandler. "Representing handshapes." In: *International review of sign linguistics* 1 (2013). Publisher: Psychology Press, pp. 115–158.

[183] Christos Sapsanis, George Georgoulas, Anthony Tzes, and Dimitrios Lymberopoulos. "Improving EMG based classification of basic hand movements using EMD." en. In: *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Osaka: IEEE, July 2013, pp. 5754–5757. ISBN: 978-1-4577-0216-7. DOI: 10.1109/EMBC.2013.6610858. URL: http://ieeexplore.ieee.org/document/6610858/ (Last accessed on Feb. 14, 2024).

[184] Constanze Schmaling. "Gebärdenschrift." In: *Wörterbücher zur Sprach- und Kommunikationswissenschaft (WSK) Online*. Ed. by Stefan J. Schierholz. Berlin, Boston: De Gruyter, 2013. URL: https://www.degruyter.com/database/WSK/entry/wsk_id_wsk_artikel_artikel_13149/html (Last accessed on Nov. 6, 2023).

[185] Zed Sevcikova Sehyr. *Questions about ASL-LEX*. Sept. 2022.

[186] Zed Sevcikova Sehyr, Naomi Caselli, Ariel Cohen-Goldberg, and Karen Emmorey. "ASL-LEX 2.0 Project: A database of lexical and phonological properties for 2723 signs in American Sign Language." In: (Aug. 2022). Publisher: Open Science Framework. DOI: 10.17605/OSF.IO/ZPHA4. URL: https://osf.io/zpha4/ (Last accessed on Aug. 26, 2023).

[187] Zed Sevcikova Sehyr, Naomi Caselli, Ariel M Cohen-Goldberg, and Karen Emmorey. "The ASL-LEX 2.0 Project: A Database of Lexical and Phonological Properties for 2,723 Signs in American Sign Language." en. In: *The Journal of Deaf Studies and Deaf Education* 26.2 (Mar. 2021), pp. 263–277. ISSN: 1081-4159, 1465-7325. DOI: 10.1093/deafed/enaa038. URL: https://academic.oup.com/jdsde/article/26/2/263/6142509 (Last accessed on Aug. 24, 2023).

[188] Harini Sekar, R Rajashekar, Gosakan Srinivasan, Priyanka Suresh, and Vineeth Vijayaraghavan. "Low-cost intelligent static gesture recognition system." en. In: *2016 Annual IEEE Systems Conference (SysCon)*. Orlando, FL, USA: IEEE, Apr. 2016, pp. 1–6. ISBN: 978-1-4673-9519-9. DOI: 10.1109/SYSCON.2016.7490642. URL: http://ieeexplore.ieee.org/document/7490642/ (Last accessed on Jan. 10, 2024).

[189] Senso Devices Inc. *Senso Devices – The Best Controller for AR/VR*. 2025. URL: https://senso.me/.

[190] Ala Shaabana, Joey Legere, Jun Li, Rong Zheng, Martin V. Mohrenschildt, and Judith M. Shedden. "Portable Electromyography: A Case Study on Ballistic Finger Movement Recognition." en. In: *IEEE Sensors Journal* 19.16 (Aug. 2019), pp. 7043–7055. ISSN: 1530-437X, 1558-1748, 2379-9153. DOI: 10.1109/JSEN.2019.2908312. URL: https://ieeexplore.ieee.org/document/8676288/ (Last accessed on Feb. 15, 2024).

[191] Hina Shaheen and Tariq Mehmood. "Talking Gloves: Low-Cost Gesture Recognition System for Sign Language Translation." en. In: *2018 IEEE Region Ten Symposium (Tensymp)*. Sydney, Australia: IEEE, July 2018, pp. 219–224. ISBN: 978-1-5386-6989-1. DOI: 10.1109/TENCONSpring.2018.8692069. URL: https://ieeexplore.ieee.org/document/8692069/ (Last accessed on Jan. 20, 2024).

[192] C.E. Shannon. "Communication in the Presence of Noise." In: *Proceedings of the IRE* 37.1 (Jan. 1949), pp. 10–21. ISSN: 0096-8390. DOI: 10.1109/JRPROC.1949.232969. URL: http://ieeexplore.ieee.org/document/1697831/ (Last accessed on May 13, 2025).

[193] Sheng Shen, He Wang, and Romit Roy Choudhury. "I am a Smartwatch and I can Track my User's Arm." en. In: *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services*. Singapore Singapore: ACM, June 2016, pp. 85–96. ISBN: 978-1-4503-4269-8. DOI: 10.1145/2906388.2906407. URL: https://dl.acm.org/doi/10.1145/2906388.2906407 (Last accessed on May 17, 2024).

[194] Ken Shoemake. "Animating rotation with quaternion curves." In: *Proceedings of the 12th Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH '85. New York, NY, USA: Association for Computing Machinery, 1985, pp. 245–254. ISBN: 0-89791-166-0. DOI: 10.1145/325334.325242. URL: https://doi.org/10.1145/325334.325242.

[195] Ahmad Zaki Shukor, Muhammad Fahmi Miskon, Muhammad Herman Jamaluddin, Fariz Bin Ali@Ibrahim, Mohd Fareed Asyraf, and Mohd Bazli Bin Bahar. "A New Data Glove Approach for Malaysian Sign Language Detection." en. In: *Procedia Computer Science* 76 (2015), pp. 60–67. ISSN: 18770509. DOI: 10.1016/j.procs.2015.12.276. URL: https://linkinghub.elsevier.com/retrieve/pii/S1877050915037771 (Last accessed on Jan. 10, 2024).

[196] SIL International. *Ethnologue Languages of the World - American Sign Language*. English. Tech. rep. SIL International, 2021. URL: https://www.ethnologue.com/language/ase/23 (Last accessed on Feb. 4, 2021).

[197] Miguel Simão, Pedro Neto, and Olivier Gibaru. "Improving novelty detection with generative adversarial networks on hand gesture data." en. In: *Neurocomputing* 358 (Sept. 2019), pp. 437–445. ISSN: 09252312. DOI: 10.1016/j.neucom.2019.05.064. URL: https://linkinghub.elsevier.com/retrieve/pii/S0925231219307702 (Last accessed on Jan. 10, 2024).

[198] Thiago Simoes Dias, Jose Jair Alves Mendes Junior, and Sergio Francisco Pichorim. "An Instrumented Glove for Recognition of Brazilian Sign Language Alphabet." en. In: *IEEE Sensors Journal* 22.3 (Feb. 2022), pp. 2518–2529. ISSN: 1530-437X, 1558-1748, 2379-9153. DOI: 10.1109/JSEN.2021.3136790. URL: https://ieeexplore.ieee.org/document/9656150/ (Last accessed on Jan. 13, 2024).

[199] Lauren H. Smith, Todd A. Kuiken, and Levi J. Hargrove. "Evaluation of Linear Regression Simultaneous Myoelectric Control Using Intramuscular EMG." en. In: *IEEE Transactions on Biomedical Engineering* 63.4 (Apr. 2016), pp. 737–746. ISSN: 0018-9294, 1558-2531. DOI: 10.1109/TBME.2015.2469741. URL: http://ieeexplore.ieee.org/document/7214240/ (Last accessed on Feb. 15, 2024).

[200] Wei Song, Anhe Wang, Yang Chen, Shuo Bai, Qingquan Han, Zhonghang Lin, Nan Yan, Deng Luo, Yiqiao Liao, Milin Zhang, Zhihua Wang, and Xiang Xie. "Design of a Flexible Wearable Smart sEMG Recorder Integrated Gradient Boosting Decision Tree Based Hand Gesture Recognition." en. In: *IEEE Transactions on Biomedical Circuits and Systems* 13.6 (Dec. 2019), pp. 1563–1574. ISSN: 1932-4545, 1940-9990. DOI: 10.1109/TBCAS.2019.2953998. URL: https://ieeexplore.ieee.org/document/8903245/ (Last accessed on Feb. 6, 2024).

[201] Webb Associates (Yellow Springs, Ohio) Anthropology Research Project, Space Administration Scientific, and Technical Information Office. *Anthropometric source book*. National Information Service, 1978.

[202] Natchalee Srimaneekarn, Anthony Hayter, Wei Liu, and Chanita Tantipoj. "Binary response analysis using logistic regression in dentistry." In: *International Journal of Dentistry* 2022 (2022). URL: https://www.hindawi.com/journals/ijd/2022/5358602/.

[203] Markus Steinbach and Roland Pfau. "Grammaticalization of auxiliaries in sign languages." en. In: *Visible Variation*. Ed. by Pamela M. Perniss, Roland Pfau, and Markus Steinbach. Mouton de Gruyter, Aug. 2007, pp. 303–340. ISBN: 978-3-11-019578-1. DOI: 10.1515/9783110198850.303. URL: https://www.degruyt

er.com/document/doi/10.1515/9783110198850.303/html (Last accessed on Jan. 31, 2024).

[204] Martin LA Sternberg and Marin Sternberg. *American Sign Language Dictionary-Flexi*. HarperResource, 1998.

[205] W Stokoe. *Sign language structure. 1978*. Linstok Press, Silver Spring, MD, 1960.

[206] William C Stokoe, Dorothy C Casterline, and Carl G Croneberg. *A dictionary of American Sign Language on linguistic principles*. Linstok Press, 1976.

[207] Subramanian Sundaram, Petr Kellnhofer, Yunzhu Li, Jun-Yan Zhu, Antonio Torralba, and Wojciech Matusik. "Learning the signatures of the human grasp using a scalable tactile glove." en. In: *Nature* 569.7758 (May 2019), pp. 698–702. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-019-1234-z. URL: https://www.nature.com/articles/s41586-019-1234-z (Last accessed on Feb. 6, 2024).

[208] Samuel Supalla, Cecile McKee, and Jody Cripps. "An Overview on the ASL-phabet." In: *Gloss Institute's Monograph Series* 1 (2014), pp. 1–18.

[209] Valerie Sutton. *Lessons in sign writing: textbook*. en. 4th ed. OCLC: 900638280. La Jolla (CA): Deaf Action Committee for Sign Writing, 2014. ISBN: 978-0-914336-55-6.

[210] Valerie Sutton. *SignPuddle Online*. 2024. URL: https://www.signbank.org/signpuddle2.0/searchword.php?ui=1&sgn=4&sid=1507,2557,9069&sTrm=candy&type=any&sTxt=&sSrc=&.

[211] Stanislav Sykora. "Mathematical Means and Averages: Basic Properties." eng. In: *Stan's Library* Volume III (July 2009). DOI: 10.3247/sl3math09.001. URL: https://doi.org/10.3247/sl3math09.001 (Last accessed on Apr. 15, 2025).

[212] Adam Szczegielniak. "Phonetics: The Sounds of Language." en. In: ().

[213] I Talib, K Sundaraj, and C K Lam. "Choice of Mechanomyography Sensors for Diverse Types of Muscle Activities." en. In: 10.1 (2018).

[214] Federico Tavella, Viktor Schlegel, Marta Romeo, Aphrodite Galata, and Angelo Cangelosi. "WLASL-LEX: a Dataset for Recognising Phonological Properties in American Sign Language." en. In: *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. Dublin, Ireland: Association for Computational Linguistics, 2022, pp. 453–463. DOI: 10.18653/v1/2022.acl-short.49. URL: https://aclanthology.org/2022.acl-short.49 (Last accessed on Jan. 15, 2024).

[215] Richard A Tennant, Marianne Gluszak, and Marianne Gluszak Brown. *The American sign language handshape dictionary*. Gallaudet University Press, 1998.

[216] Richard A. Tennant and Marianne Gluszak Brown. *The American Sign Language handshape starter: a beginner's guide*. Washington, D.C: Gallaudet University Press, 2002. ISBN: 978-1-56368-130-1.

[217] Gudny Bjork Thorvaldsdottir. "The Beginnings of Phonetic and Phonological Coding in the Signs of Ireland Digital Corpus: the Representation of Handshapes." en. In: (2010). Publisher: Dublin Institute of Technology. DOI: 10.21427/D7316R. URL: https://arrow.dit.ie/itbj/vol11/iss1/4/ (Last accessed on Jan. 9, 2024).

[218] Oksana Tkachman, Kathleen Hall, Andre Xavier, and Bryan Gick. "Sign Language Phonetic Annotation meets Phonological CorpusTools: Towards a sign language toolset for phonetic notation and phonological analysis." In: *Proceedings of the Annual Meetings on Phonology* 3 (June 2016). DOI: 10.3765/amp.v3i0.3667.

[219] Noor Tubaiz, Tamer Shanableh, and Khaled Assaleh. "Glove-Based Continuous Arabic Sign Language Recognition in User-Dependent Mode." en. In: *IEEE Transactions on Human-Machine Systems* 45.4 (Aug. 2015), pp. 526–533. ISSN: 2168-2291, 2168-2305. DOI: 10.1109/THMS.2015.2406692. URL: http://ieeexplore.ieee.org/document/7061411/ (Last accessed on Jan. 12, 2024).

[220] Ultrahaptics Ltd. *Leap Motion Controller – Technical Datasheet*. The West Wing, Glass Wharf, Bristol, BS2 0EL, United Kingdom: Ultraleap (formerly Leap Motion), 2023. URL: https://www.inforabreu.pt/cte/I_1450.pdf (Last accessed on May 18, 2025).

[221] Terry T. Um, Franz M. J. Pfister, Daniel Pichler, Satoshi Endo, Muriel Lang, Sandra Hirche, Urban Fietzek, and Dana Kulić. "Data augmentation of wearable sensor data for parkinson's disease monitoring using convolutional neural networks." en. In: *Proceedings of the 19th ACM International Conference on Multimodal Interaction*. Glasgow UK: ACM, Nov. 2017, pp. 216–220. ISBN: 978-1-4503-5543-8. DOI: 10.1145/3136755.3136817. URL: https://dl.acm.org/doi/10.1145/3136755.3136817 (Last accessed on May 30, 2022).

[222] Europäische Union. *Richtlinie (EU) 2016/2102 des Europäischen Parlaments und des Rates vom 26. Oktober 2016 über den barrierefreien Zugang zu den Websites und mobilen Anwendungen öffentlicher Stellen*. de. Dec. 2016. URL: https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:32016L2102 (Last accessed on July 20, 2025).

[223] United Nations News. *Sign language protects 'linguistic identity and cultural diversity' of all users, says UN chief*. Tech. rep. United Nations, Sept. 2019. URL: https://news.un.org/en/story/2019/09/1047012 (Last accessed on Apr. 26, 2021).

[224] Clayton Valli. *The Gallaudet dictionary of American sign language*. Gallaudet University Press, 2006.

[225] Harry Van Der Hulst and Rachel Channon. "Notation systems." en. In: *Sign Languages*. Ed. by Diane Brentari. 1st ed. Cambridge University Press, May 2010, pp. 151–172. ISBN: 978-0-521-88370-2 978-0-511-71220-3 978-0-521-71003-9. DOI: 10.1017/CBO9780511712203.009. URL: https://www.cambridge.org/core/pro

duct/identifier/CBO9780511712203A019/type/book_part (Last accessed on Nov. 6, 2023).

[226] B. Venkatesh and J. Anuradha. "A Review of Feature Selection and Its Methods." en. In: *Cybernetics and Information Technologies* 19.1 (Mar. 2019), pp. 3–26. ISSN: 1314-4081. DOI: 10.2478/cait-2019-0001. URL: https://www.sciendo.com/article/10.2478/cait-2019-0001 (Last accessed on Mar. 9, 2024).

[227] Zhengyang Wang and Jason Gu. "Finger Tracking for Human Computer Interface Using Multiple Sensor Data Fusion." en. In: *2022 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*. Halifax, NS, Canada: IEEE, Sept. 2022, pp. 319–323. ISBN: 978-1-6654-8432-9. DOI: 10.1109/CCECE49351.2022.9918205. URL: https://ieeexplore.ieee.org/document/9918205/ (Last accessed on Dec. 18, 2023).

[228] Sophia Waterfield. *ASL Day 2019: Everything You Need To Know About American Sign Language*. English. Tech. rep. Newsweek, Apr. 2019. URL: https://www.newsweek.com/asl-day-2019-american-sign-language-1394695 (Last accessed on Apr. 30, 2021).

[229] Feng Wen, Zhongda Sun, Tianyiyi He, Qiongfeng Shi, Minglu Zhu, Zixuan Zhang, Lianhui Li, Ting Zhang, and Chengkuo Lee. "Machine Learning Glove Using Self-Powered Conductive Superhydrophobic Triboelectric Textile for Gesture Recognition in VR/AR Applications." en. In: *Advanced Science* 7.14 (July 2020), p. 2000261. ISSN: 2198-3844, 2198-3844. DOI: 10.1002/advs.202000261. URL: https://onlinelibrary.wiley.com/doi/10.1002/advs.202000261 (Last accessed on Feb. 5, 2024).

[230] Darrell Whitley. "A genetic algorithm tutorial." In: *Statistics and Computing* (1994), pp. 65–85. URL: https://link.springer.com/article/10.1007/BF00175354.

[231] William Newell and Frank Caccamise. *American Sign Language Grammer*. 6th ed. Washington School for the Deaf (WSD) and National Technical Institute for the Deaf, May 2008.

[232] Joshua T. Williams and Sharlene D. Newman. "Impacts of Visual Sonority and Handshape Markedness on Second Language Learning of American Sign Language." en. In: *Journal of Deaf Studies and Deaf Education* 21.2 (Apr. 2016), pp. 171–186. ISSN: 1081-4159, 1465-7325. DOI: 10.1093/deafed/env055. URL: https://academic.oup.com/jdsde/article-lookup/doi/10.1093/deafed/env055 (Last accessed on Jan. 31, 2024).

[233] World Health Organization. *Deafness and hearing loss*. English. Tech. rep. World Health Organization, Apr. 2021. URL: https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss (Last accessed on Apr. 26, 2021).

[234] World Health Organization. *WHO: 1 in 4 people projected to have hearing problems by 2050*. English. Tech. rep. World Health Organization, Apr. 2021. URL: https://www.who.int/news/item/02-03-2021-who-1-in-4-people-projected-to-have-hearing-problems-by-2050 (Last accessed on Mar. 2, 2021).

[235] Jian Wu, Lu Sun, and Roozbeh Jafari. "A Wearable System for Recognizing American Sign Language in Real-Time Using IMU and Surface EMG Sensors." en. In: *IEEE Journal of Biomedical and Health Informatics* 20.5 (Sept. 2016), pp. 1281–1290. ISSN: 2168-2194, 2168-2208. DOI: 10.1109/JBHI.2016.2598302. URL: http://ieeexplore.ieee.org/document/7552525/ (Last accessed on Feb. 15, 2024).

[236] Jian Wu, Zhongjun Tian, Lu Sun, Leonardo Estevez, and Roozbeh Jafari. "Real-time American Sign Language Recognition using wrist-worn motion and surface EMG sensors." en. In: *2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*. Cambridge, MA, USA: IEEE, June 2015, pp. 1–6. ISBN: 978-1-4673-7201-5. DOI: 10.1109/BSN.2015.7299393. URL: http://ieeexplore.ieee.org/document/7299393/ (Last accessed on Feb. 14, 2024).

[237] Sanjay Kumar Yadav, Banshi Krishan Malla, Ashok Kumar Srivastava, Ram Prasad Timsina, Nityanand Srivastava, and Alok Kumar. "Anthropometric study of philtrum (face) and other nasal parameters in Nepal." en. In: *International Journal of Modern Anthropology* 2.11 (Sept. 2018), p. 163. ISSN: 1737-8176, 1737-7374. DOI: 10.4314/ijma.v2i11.8. URL: https://www.ajol.info/index.php/ijma/article/view/177213 (Last accessed on Nov. 9, 2025).

[238] Geng Yang, Jia Deng, Gaoyang Pang, Hao Zhang, Jiayi Li, Bin Deng, Zhibo Pang, Juan Xu, Mingzhe Jiang, Pasi Liljeberg, Haibo Xie, and Huayong Yang. "An IoT-Enabled Stroke Rehabilitation System Based on Smart Wearable Armband and Machine Learning." en. In: *IEEE Journal of Translational Engineering in Health and Medicine* 6 (2018), pp. 1–10. ISSN: 2168-2372. DOI: 10.1109/JTEHM.2018.2822681. URL: https://ieeexplore.ieee.org/document/8356006/ (Last accessed on Jan. 19, 2024).

[239] Guan Yuan, Xiao Liu, Qiuyan Yan, Shaojie Qiao, Zhixiao Wang, and Li Yuan. "Hand Gesture Recognition using Deep Feature Fusion Network based on Wearable Sensors." en. In: *IEEE Sensors Journal* (2020), pp. 1–1. ISSN: 1530-437X, 1558-1748, 2379-9153. DOI: 10.1109/JSEN.2020.3014276. URL: https://ieeexplore.ieee.org/document/9158332/ (Last accessed on Jan. 12, 2024).

[240] Andreas Zankl, Lukas Eberle, Luciano Molinari, and Albert Schinzel. "Growth charts for nose length, nasal protrusion, and philtrum length from birth to 97 years." en. In: *American Journal of Medical Genetics* 111.4 (Sept. 2002), pp. 388–391. ISSN: 0148-7299, 1096-8628. DOI: 10.1002/ajmg.10472. URL: https://onlinelibrary.wiley.com/doi/10.1002/ajmg.10472 (Last accessed on Nov. 9, 2025).

[241] Piero Zappi, Clemens Lombriser, Thomas Stiefmeier, Elisabetta Farella, Daniel Roggen, Luca Benini, and Gerhard Tröster. "Activity recognition from on-body sensors: accuracy-power trade-off by dynamic sensor selection." In: *Wireless Sensor Networks: 5th European Conference, EWSN 2008, Bologna, Italy, January 30-February 1, 2008. Proceedings*. Italy, Springer, Jan. 2008, pp. 17–33. URL: https://link.springer.com/chapter/10.1007/978-3-540-77690-1_2.

[242] Fan Zhang, Valentin Bazarevsky, Andrey Vakunov, Andrei Tkachenka, George Sung, Chuo-Ling Chang, and Matthias Grundmann. *MediaPipe Hands: On-device Real-time Hand Tracking*. en. arXiv:2006.10214 [cs]. June 2020. URL: http://arxiv.org/abs/2006.10214 (Last accessed on Feb. 2, 2024).

[243] Yang Zhang and Chris Harrison. "Tomo: Wearable, Low-Cost Electrical Impedance Tomography for Hand Gesture Recognition." en. In: *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*. Charlotte NC USA: ACM, Nov. 2015, pp. 167–173. ISBN: 978-1-4503-3779-3. DOI: 10.1145/2807442.2807480. URL: https://dl.acm.org/doi/10.1145/2807442.2807480 (Last accessed on Jan. 12, 2024).

[244] Yangyang Zhang, Ying Huang, Xuehu Sun, Yunong Zhao, Xiaohui Guo, Ping Liu, Caixia Liu, and Yugang Zhang. "Static and Dynamic Human Arm/Hand Gesture Capturing and Recognition via Multiinformation Fusion of Flexible Strain Sensors." en. In: *IEEE Sensors Journal* 20.12 (June 2020), pp. 6450–6459. ISSN: 1530-437X, 1558-1748, 2379-9153. DOI: 10.1109/JSEN.2020.2965580. URL: https://ieeexplore.ieee.org/document/8955861/ (Last accessed on Jan. 12, 2024).

[245] Yi Zhang, Yue Zheng, Kun Qian, Guidong Zhang, Yunhao Liu, Chenshu Wu, and Zheng Yang. "Widar3.0: Zero-Effort Cross-Domain Gesture Recognition With Wi-Fi." In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44.11 (2021), pp. 8671–8688. URL: https://ieeexplore.ieee.org/abstract/document/9516988?casa_token=XeAejjKN4YsAAAAA:UbTdf5l-Vjs0M0yFkGlojtQReRECDp2xcroYj_YCEVr_Mu1OIVrFpNUHt3zNZQ5Mpw8q7CuMDQ.

[246] Zhihao Zhou, Kyle Chen, Xiaoshi Li, Songlin Zhang, Yufen Wu, Yihao Zhou, Keyu Meng, Chenchen Sun, Qiang He, Wenjing Fan, Endong Fan, Zhiwei Lin, Xulong Tan, Weili Deng, Jin Yang, and Jun Chen. "Sign-to-speech translation using machine-learning-assisted stretchable sensor arrays." en. In: *Nature Electronics* 3.9 (June 2020), pp. 571–578. ISSN: 2520-1131. DOI: 10.1038/s41928-020-0428-6. URL: https://www.nature.com/articles/s41928-020-0428-6 (Last accessed on Mar. 7, 2024).

# APPENDIX

## A.1 NOTATION SYSTEMS AND SIGN DATABASES

### A.1.1 *Phonological Properties of the ASL SignBank*

Table A.1: Phonological Properties of the ASL SignBank

| Phonological Property | Description |
|---|---|
| HANDEDNESS | Corresponds to SIGN TYPE of *ASL-LEX*. |
| WEAK DROP | Indicates whether a sign originally produced with two hands is represented with only one hand in a specific context. |
| WEAK PROP | Indicates whether a sign originally produced with one hand is represented with two hands, either symmetrically or alternating, in a specific context. |
| LOCATION - MAJOR | Corresponds to the MAJOR LOCATION of the *ASL-LEX*. |
| LOCATION - MINOR 1 | Corresponds to the MINOR LOCATION of the *ASL-LEX*, plus *ElbowFront* but without *ArmAway*, *BodyAway*, *HandAway* and *HeadAway*. *WristFront* and *WristBack* are assigned to the hand rather than the arm. |
| LOCATION - MINOR 2 | Corresponds to the SECOND MINOR LOCATION of the *ASL-LEX* and can take the same values as LOCATION - MINOR 1. |
| DOMINANT HAND - SELECTED FINGERS | Corresponds to the SELECTED FINGERS of the *ASL-LEX*. |
| DOMINANT HAND - FLEXION | Corresponds to a combination of the FLEXION and THUMB POSITION values of the *ASL-LEX*. |
| DOMINANT HAND - ABDUCTION CHANGE | Corresponds to the SPREAD CHANGE of the *ASL-LEX*. |
| DOMINANT HAND - FLEXION CHANGE | Corresponds to the FLEXION CHANGE of the *ASL-LEX*. |
| NONDOMINANT HANDSHAPE | One of seven *Unmarked Handshapes*, if the sign is categorized as *AsymmetricalDifferentHandshape*, otherwise *N/A*. |
| PATH MOVEMENT | Corresponds to the PATH MOVEMENT of the first version of the *ASL-LEX* (*BackAndForth*, *Circular*, *Curved*, *Straight*) [41]. |

### A.1.2 *Phonological Properties of* ASL-LEX

In the following, the phonological properties of *ASL-LEX* (see Table A.2) as well as the associated classes (see Table A.3) are presented [41, 186, 187]. The number of classes indicates the classes actually used.

Table A.2: Phonological properties of the *ASL-LEX*

| Phonological property | Description |
|---|---|
| HANDSHAPE | Unique combination, automatically generated from the values for SELECTED FINGERS, FLEXION, SPREAD, THUMB POSITION and subsequently manually corrected. In the process, some handshapes were merged and others separated. Sehyr et al. [187] point out that static handshapes are not ideal for representing dynamic signs, and that the resulting handshapes may not correspond to the user's intuition. |
| SELECTED FINGERS | *Selected Fingers* are *i)* the fingers that move; otherwise *ii)* the fingers that are neither fully open nor closed. If no fingers can still be clearly identified, then they are *iii)* those fingers that are extended. |
| FLEXION | The type of flexion of the SELECTED FINGERS at the beginning. The scale begins with *FullyClosed*, when all finger joints are fully flexed, and ends with *FullyOpened* (all joints extended). If only the MCP joints are flexed, the flexion is referred to as *Flat*; the opposite (DIP and PIP joints are flexed) as *Bent*. A uniform flexion of all three joints is a *Curved* flexion. *Stacked* and *Crossed* denote two special types of flexion in which the selected fingers exhibit different flexions (as in the letter K of the manual alphabet) or the fingers are crossed (as in the letter R of the manual alphabet, both see Table A.4). FLEXION is *NA* if the thumb (*T*) is the only SELECTED FINGER. |
| FLEXION CHANGE | Indicates whether the flexion of the SELECTED FINGERS changes during the articulation. *NA* means that FLEXION is also *NA*. |
| SPREAD | Indicates whether the SELECTED FINGERS are spread at the beginning. This requires more than one finger to be selected and these must lie next to one another; otherwise *NA* is assigned, as is the case when FLEXION is defined as *Stacked* or *FullyClosed*. If FLEXION is defined as *Crossed*, there is no spreading. |
| SPREAD CHANGE | Indicates whether the spread of the SELECTED FINGERS changes during the articulation. *NA* means that SPREAD (abduction) is also *NA*, unless FLEXION is *FullyClosed* (completely closed, e.g. handshape *S*). |
| THUMB POSITION | The position of the thumb at the beginning can be closed, i.e. it rests on the palm or on fully flexed fingers, or open, if it merely touches them (e.g. handshape *A*) or has no contact. If the thumb is a SELECTED FINGERS, the position is *open*. |
| THUMB CONTACT | Indicates whether the tip of the thumb or the thenar eminence (not the lateral side of the thumb) touches a SELECTED FINGERS at any time during the articulation. |

Table A.2: Phonological properties of the *ASL-LEX* (continued)

| Phonological property | Description |
|---|---|
| SIGN TYPE | Classification according to *Battison's Sign Types* (cf. 2.1.2), but without further differentiation of one-handed signs. If one hand touches the other arm, the sign is considered *one-handed*. |
| PATH MOVEMENT | Indicates the type of hand movement in three-dimensional space. |
| REPEATED MOVEMENT | Indicates whether the movement(s) (PATH MOVEMENT, ULNAR ROTATION or FLEXION/SPREAD CHANGE) repeat(s) unchanged. |
| MAJOR LOCATION | The articulation location of the sign relative to the signer's body or the *neutral space*. |
| MINOR LOCATION | The articulation location of the sign with respect to the MAJOR LOCATION at the beginning of articulation. According to Brentari's *Prosodic Model* [32], each MAJOR LOCATION is subdivided into eight MINOR LOCATION, therefore there is a strong dependence between the two values. If the MAJOR LOCATION is the *neutral space*, this also applies to the MINOR LOCATION. |
| SECOND MINOR LOCATION | If a path movement (PATH MOVEMENT) is executed, a second MINOR LOCATION is defined by the location that the dominant hand has as its target location and touches. If no contact takes place, the value *Away* is set. |
| CONTACT | Indicates whether the dominant hand had contact with MAJOR LOCATION at any time during the articulation. If a symmetrical sign is articulated in the neutral space, CONTACT is given if there is contact between both hands. |
| NON-DOMINANT HANDSHAPE | According to Battison (cf. 2.1.2), i.e. for an asymmetrical sign one of seven *Unmarked Handshapes* (*B*, *A*, *S*, *1*, *C*, *O*, or *5*), otherwise as the handshape of the dominant hand. Violations of the *Symmetry Rule* are labeled *SymmetryViolation*, violations of the *Dominance Rule* *DominanceViolation*. |
| ULNAR ROTATION | Indicates whether the wrist is rotated during the articulation. |

Table A.3: Phonological classes of the *ASL-LEX*

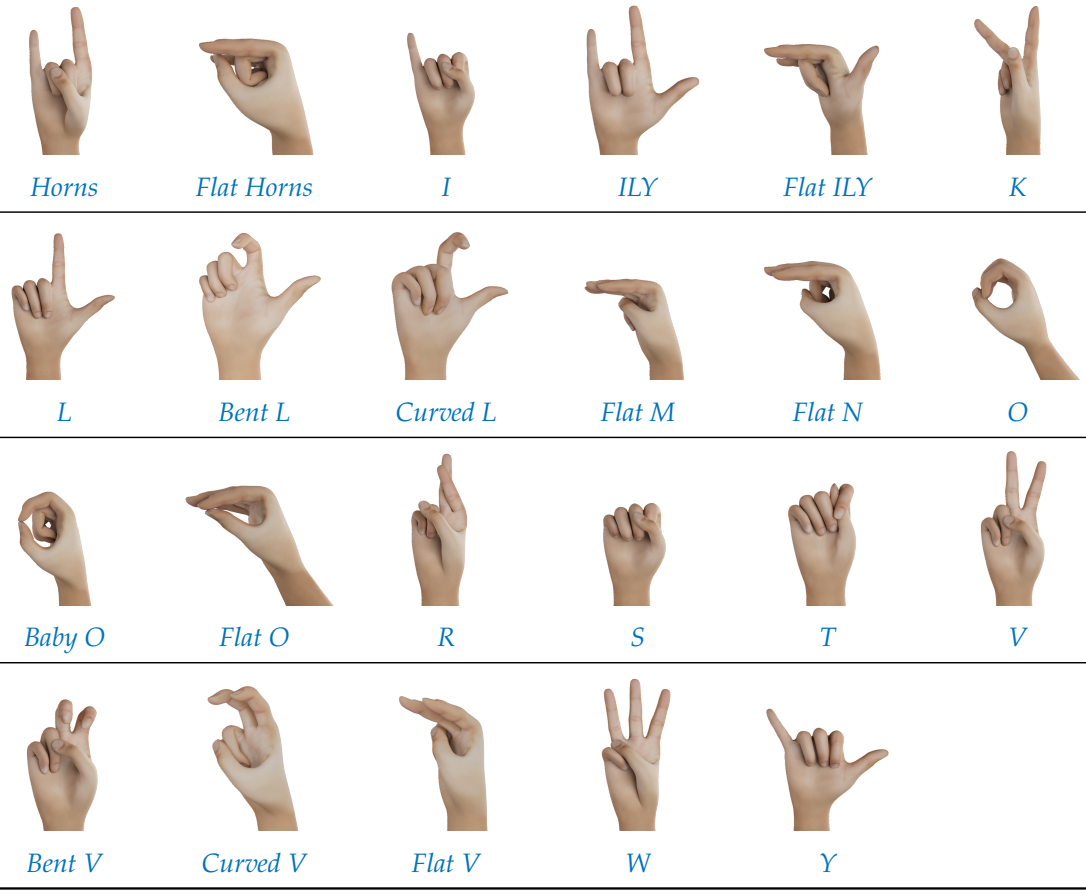
| Phonological property | Classes |
|---|---|
| Handshape (58) | *1*, *Bent 1*, *Curved 1*, *Flat 1*, *3*, *4*, *Curved 4*, *Flat 4*, *5*, *Curved 5*, *Flatspread 5*, *Stacked 5*, *7*, *8*, *Open 8*, *A*, *Closed B*, *Flat B*, *Open B*, *C*, *D*, *E*, *Closed E*, *Open E*, *Spread Open E*, *Spread E*, *F*, *Open F*, *G*, *Goody Goody*, *H*, *Curved H*, *Flat H*, *Open H*, *Horns*, *Flat Horns*, *I*, *ILY*, *Flat ILY*, *K*, *L*, *Bent L*, *Curved L*, *Flat M*, *Flat N*, *O*, *Baby O*, *Flat O*, *P*, *R*, *S*, *T*, *V*, *Bent V*, *Curved V*, *Flat V*, *W*, *Y* |
| Selected Fingers (12) | *i*, *im*, *imp*, *imr*, *imrp*, *ip*, *m*, *mr*, *mrp*, *p*, *r*, *t* |
| Flexion (8) | *Bent*, *Crossed*, *Curved*, *Flat*, *FullyClosed*, *FullyOpen*, *NA*, *Stacked* |
| Flexion Change (3) | *NA*, *False*, *True* |
| Spread (3) | *NA*, *False*, *True* |
| Spread Change (3) | *NA*, *False*, *True* |
| Thumb Position (2) | *Closed*, *Open* |
| Thumb Contact (3) | *NA*, *False*, *True* |
| Sign Type (7) | *AsymmetricalDifferentHandshape*, *AsymmetricalSameHandshape*, *DominanceViolation*, *NA*, *OneHanded*, *SymmetricalOrAlternating*, *SymmetryViolation* |
| Path Movement (8) | *BackAndForth*, *Circular*, *Curved*, *None*, *Other*, *Straight*, *X-Shaped*, *Z-shaped* |
| Repeated Movement (2) | *False*, *True* |
| Major Location (6) | *Arm*, *Body*, *Hand*, *Head*, *Neutral*, *Other* |
| Minor Location (39) | *ArmAway*, *BodyAway*, *CheekNose*, *Chin*, *Clavicle*, *ElbowBack*, *ElbowFront*, *Eye*, *FingerBack*, *FingerFront*, *FingerRadial*, *FingerTip*, *FingerUlnar*, *ForearmBack*, *ForearmFront*, *ForearmUlnar*, *Forehead*, *HandAway*, *HeadAway*, *HeadTop*, *Heel*, *Hips*, *Mouth*, *NA*, *Neck*, *Neutral*, *Other*, *Palm*, *PalmBack*, *Shoulder*, *TorsoBottom*, *TorsoMid*, *TorsoTop*, *UnderChin*, *UpperArm*, *UpperLip*, *Waist*, *WristBack*, *WristFront* |
| Second Minor Location (39) | *ArmAway*, *Away*, *BodyAway*, *CheekNose*, *Chin*, *Clavicle*, *ElbowBack*, *Eye*, *FingerBack*, *FingerFront*, *FingerRadial*, *FingerTip*, *FingerUlnar*, *ForearmBack*, *ForearmFront*, *ForearmRadial*, *ForearmUlnar*, *Forehead*, *HandAway*, *HeadAway*, *HeadTop*, *Heel*, *Hips*, *Mouth*, *NA*, *Neck*, *Neutral*, *Other*, *OtherAway*, *Palm*, *PalmBack*, *Shoulder*, *TorsoBottom*, *TorsoMid*, *TorsoTop*, *UnderChin*, *UpperArm*, *UpperLip*, *Waist* |
| Contact (3) | *NA*, *False*, *True* |
| Non-dominant Handshape (53) | *1*, *Bent 1*, *Curved 1*, *Flat 1*, *3*, *4*, *Curved 4*, *Flat 4*, *5*, *Curved 5*, *Flatspread 5*, *Stacked 5*, *8*, *Open 8*, *A*, *B*[89], *Closed B*, *Curved B*, *Flat B*, *Open B*, *C*, *E*, *Open E*, *Spread Open E*, *F*, *G*, *H*, *Flat H*, *Open H*, *Horns*, *Flat Horns*, *I*, *ILY*, *L*, *Bent L*, *Curved L*, *Flat M*, *O*[89], *Baby O*, *Flat O*, *P*, *R*, *S*, *V*, *Bent V*, *Curved V*, *Flat V*, *W*, *Y*, *NA*, *Lax*, *DominanceConditionViolation*, *SymmetryViolation* |
| Ulnar Rotation (2) | *False*, *True* |

### A.1.3 *Handshapes of* ASL-LEX

Table A.4: 59[90] Handshapes of *ASL-LEX* (dominant and non-dominant hand) [186].

| | | | | | |
|---|---|---|---|---|---|
|  *1* |  *Bent 1* |  *Curved 1* |  *Flat 1* |  *3* |  *4* |
|  *Curved 4* |  *Flat 4* |  *5* |  *Curved 5* |  *Flatspread 5* |  *Stacked 5* |
|  *7* |  *8* |  *Open 8* |  *A* |  *Open A* |  *B* |
|  *Closed B* |  *Flat B* |  *Open B* |  *C* |  *D* |  *E* |
|  *Closed E* |  *Open E* |  *Spread Open E* |  *Spread E* |  *F* |  *Open F* |
|  *G* |  *Goody Goody* |  *H* |  *Curved H* |  *Flat H* |  *Open H* |

89 According to *ASL-LEX*, these handshapes do not correspond to those of the finger alphabet (as with the dominant hand), but originate from the Battison inventory.



| Horns | Flat Horns | I | ILY | Flat ILY | K |
|---|---|---|---|---|---|
| L | Bent L | Curved L | Flat M | Flat N | O |
| Baby O | Flat O | R | S | T | V |
| Bent V | Curved V | Flat V | W | Y | |

90 The handshapes for K and P are identical and differ only in the orientation of the hand.

## A.2 RELATED WORK

The following provides an overview of the works examined in Section 2.3 in the field of hand gesture recognition with body-worn sensing, grouped by the complexity of the gestures to be recognized and sorted by the number of gesture classes recognized. Where information from the works was not clearly discernible, it was marked with "-". Categories that are satisfied only under certain constraints, for example only for a fraction of the gestures, are marked with "(✓)" or, in the opposite case, with "(✗)". The tables comprise the following columns:

- **Source**: Source of the investigated work
- **Year**: Time of publication of the source
- **Hardware**: System used for gesture capture
- **Classifier**: Algorithm that delivered the best classification result
- **Within-User**: Classification result where training and test data were collected from the same individuals
- **Cross-User**: Classification result where training and test data were collected from different individuals
- **Metric**: Metric in which the result is reported: Accuracy (Acc.), F1-Score (F1), Precision (Prec.), Sensitivity (Sens.), Task Success Rate (TSR)
- **Features**: Number of features used
- **Gesture Type**: Type of gestures recognized
- **Hands**: Number of hands recognized
- **Handshape (HF)**: Consideration of hand shape
- **Orientation (Ori.)**: Consideration of palm orientation
- **Movement (Bew.)**: Consideration of hand movement
- **Location**: Consideration of hand position
- **Dynamic (Dyn.)**: Consideration of dynamic behavior as a change in hand shape, orientation, or position
- **Mobile**: Consideration of the method's mobility (the gesturing person can move freely)
- **Gestures**: Number of gesture classes to be recognized
- **True Negative (TN)**: Number of participants in the data collection
- **Samples**: Number of data points in the dataset with which the classifier was trained and tested

Table A.5: Overview of body-worn sensing-based recognition approaches for simple gestures

| Source | Year | Hardware | Classifier | Within-User | Cross-User | Metric | Features | Gesture Type | Hands | HF | Ori. | Bew. | Loc. | Dyn. | Mobile | Gestures | TN | Samples |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Kanokoda et al. [111] | 2019 | Data glove (3 Flex) | RNN | 90,80% | - | Acc. | 60 | Finger gestures | 1 | ✓ | ✗ | ✗ | ✗ | ✓ | (✓) | 3 | 8 | 400.000 |
| Lv et al. [135] | 2017 | Data glove | SVM | 93,19% | - | Acc. | 33 | Finger gestures | 1 | ✓ | ✗ | ✗ | ✗ | ✓ | ✓ | 4 | - | 315 |
| Chuang et al. [48] | 2019 | Data glove (5 Flex) | RNN | - | 97,27% | Acc. | - | Finger gestures | 1 | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | 4 | 11 | 4.488 |
| Benatti et al. [25] | 2014 | sEMG (8 electrodes) | SVM | 92,36% | - | Acc. | 4 | Finger gestures | 1 | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | 5 | 1 | 50 |
| Grützmacher et al. [82] | 2015 | Data glove (3 IMUs) | Template Matching | 66,70% | - | F1 | 5 | Finger gestures | 1 | ✓ | ✗ | ✗ | ✗ | ✓ | (✓) | 5 | 20 | 160 |
| Meattini et al. [140] | 2018 | sEMG (8 electrodes) | SVM | 96,30% | - | TSR | 8 | Finger gestures | 1 | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | 5 | 4 | 384 |
| Ma et al. [136] | 2019 | Data glove (4 Flex, 1 IMU) | FNN | 94,30% | - | Acc. | 104 | Finger gestures | 1 | ✓ | ✗ | ✗ | ✗ | ✓ | ✓ | 5 | 10 | 500 |
| **Achenbach et al. [6]** | **2021** | **Data glove (7 IMUs)** | **SVM** | **98,00%** | **87,50%** | **Acc.** | **15** | **Finger gestures** | **1** | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | **5** | **11** | **550** |
| **Achenbach et al. [7]** | **2022** | **Data glove (7 IMUs)** | **SVM** | **99,20%** | **97,60%** | **Acc.** | **15** | **Finger gestures** | **1** | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | **5** | **30** | **1.500** |
| Sapsanis et al. [183] | 2013 | sEMG (2 electrodes) | Linear Classifier | 89.21% | - | Acc. | 8 | Grasps | 1 | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | 6 | 5 | 900 |
| Ayodele et al. [19] | 2020 | Data glove (conductive fibers) | CNN | 88,27% | 75,37% | Acc. | - | Grasps | 1 | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | 6 | 5 | 750 |
| Benatti et al. [26] | 2015 | sEMG (8 electrodes) | SVM | 89,20% | - | Acc. | 8 | Finger gestures | 1 | ✓ | (✓) | ✗ | ✗ | ✓ | ✓ | 7 | 4 | 112 |
| Sundaram et al. [207] | 2019 | Data glove (Flex array) | CNN | 89,40% | - | Acc. | - | Grasps | 1 | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | 8 | - | 4.326 |
| Zhang and Harrison [243] | 2015 | EIT (8 channels) | SVM | 96,60% | 60,90% | Acc. | 406 | Finger gestures | 1 | ✓ | (✓) | ✗ | ✗ | ✗ | ✓ | 8 | 10 | 22.000 |
| Yang et al. [238] | 2018 | sEMG (8 electrodes) | LDA | 99,85% | - | Acc. | 6 | Finger gestures | 1 | ✓ | (✗) | ✗ | ✗ | ✗ | ✓ | 9 | 3 | 81 |
| Huang et al. [102] | 2019 | Data glove (3 Flex) | Template Matching | - | 82,50% | Acc. | 3 | CSL words | 1 | ✓ | ✗ | ✗ | ✗ | ✓ | ✓ | 9 | 4 | 20.180 |
| Li et al. [125] | 2018 | Data glove (14 Flex) | GA + BPNN | 98,67% | - | Acc. | 14 | Finger gestures | 1 | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | 10 | 10 | 3.000 |
| Guo et al. [83] | 2017 | Data glove | Template Matching | 98,00% | - | Acc. | 15 | Finger gestures | 1 | ✓ | ✗ | ✗ | ✗ | ✓ | ✓ | 10 | 10 | 500 |
| Liu et al. [130] | 2017 | sEMG (4 electrodes) | ANN | 94,00% | - | Acc. | 8 | Surface finger gestures | 1 | ✓ | ✗ | ✗ | ✗ | ✓ | ✓ | 10 | 4 | - |
| Rosero-Montalvo et al. [176] | 2018 | Data glove (5 Flex) | kNN | 85,00% | - | Acc. | 5 | Numbers | 1 | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | 10 | 10 | 5.100 |
| Huang et al. [102] | 2019 | Data glove (10 Flex) | ANN | - | 98,50% | Acc. | 10 | Numbers | 1 | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | 10 | 4 | 22.000 |
| Bal et al. [20] | 2021 | Data glove (3 Flex) | PNN | 92,70% | - | Acc. | - | BdSL letters | 1 | ✓ | ✗ | ✗ | ✗ | ✓ | ✓ | 10 | 88 | 17.600 |
| Benatti et al. [27] | 2019 | sEMG (8 electrodes) | HDC | 85,00% | - | Acc. | 8 | Finger gestures | 1 | ✓ | (✓) | ✗ | ✗ | ✗ | ✓ | 11 | 10 | 660 |
| Zhou et al. [246] | 2020 | Data glove (5 Flex) | PCA + SVM | 98,63% | - | Acc. | - | ASL letters/numbers | 1 | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | 11 | 4 | 660 |
| Song et al. [200] | 2019 | sEMG (32 electrodes) | GBDT | 90,07% | - | Acc. | 9 | Finger gestures | 1 | ✓ | (✓) | ✗ | ✗ | ✗ | ✓ | 12 | 7 | 840 |
| Antillon et al. [12] | 2021 | Data glove (5 Force) | SVM | 97,00% | 98,00% | Acc. | 5 | Finger gestures | 1 | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | 13 | 24 | 12.480 |
| Min et al. [143] | 2021 | Data glove (6 IMUs) | ResNet | 99,20% | - | Acc. | 6 | Finger gestures | 1 | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | 13 | - | 6.500 |
| Nassour et al. [153] | 2020 | Data glove (14 Flex) | RF | 88,50% | - | Acc. | 14 | Finger gestures | 1 | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | 15 | - | 2.250 |
| Chen et al. [44] | 2020 | Data glove (10 Flex, 10 Force) | SVM | 93,32% | - | Acc. | 5 | Finger gestures | 1 | ✓ | ✗ | ✗ | ✗ | ✗ | (✓) | 16 | 8 | 14.4000 |
| Lee and Bae [124] | 2022 | Data glove (10 Flex) | RNN | - | 94,21% | Acc. | 10 | ASL letters | 1 | ✓ | ✗ | ✗ | ✗ | ✓ | ✓ | 17 | 13 | 1.072 |
| Feng et al. [70] | 2019 | MMG (4 IMUs) | SVM | 95,38% | - | Acc. | 26 | CSL actions | 1 | ✓ | ✗ | ✓ | ✗ | ✓ | ✓ | 18 | 6 | 32.400 |
| Saggio et al. [178] | 2011 | Data glove (15 Flex) | EDC | 95,00% | 91,00% | Acc. | 10 | Finger gestures | 1 | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | 20 | 9 | 180 |
| Duan et al. [63] | 2022 | Data glove (conductive fibers) | FNN | 98,10% | - | Acc. | 5 | Single letters/numbers | 1 | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | 20 | - | 3.000 |
| Amma et al. [10] | 2015 | sEMG | NB | 89,80% | 75,00% | Acc. | 168 | Finger gestures | 1 | ✓ | (✓) | ✗ | ✗ | ✗ | (✓) | 27 | 5 | 7.250 |
| Du et al. [62] | 2017 | sEMG | CNN | 96,90% | - | Acc. | 168 | Finger gestures | 1 | ✓ | (✓) | ✗ | ✗ | ✗ | (✓) | 27 | 5 | 6.750 |
| Shaabana et al. [190] | 2019 | sEMG (8 electrodes) | GM-HMM | 92,43% | - | Prec. | 48 | Typing movements | 2 | ✓ | (✓) | ✗ | ✗ | ✗ | ✓ | 27 | 8 | ~7.500 |
| **Achenbach et al. [5]** | **2023** | **Data glove (6 IMUs, 5 Flex)** | **VL2** | **95,55%** | **91,91%** | **Acc.** | **20** | **ASL letters/numbers** | **1** | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | **27** | **20** | **1.620** |
| Abhishek et al. [1] | 2016 | Data glove (8 Touch) | Template Matching | 92,00% | - | Acc. | 8 | ASL letters/numbers | 1 | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | 36 | - | 1.080 |
| Sekar et al. [188] | 2016 | Data glove (8 contact, 9 Flex) | Template Matching | 94,50% | - | Acc. | 17 | SL signs | 2 | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | 36 | 1 | 3.600 |
| **Achenbach et al. [5]** | **2023** | **Data glove (6 IMUs, 5 Flex)** | **VL2** | **93,28%** | **87,50%** | **Acc.** | **20** | **ASL handshapes** | **1** | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | **56** | **20** | **3.360** |

Table A.6: Overview of body-worn sensing-based recognition approaches for finger-alphabet-like gestures

| Source | Year | Hardware | Classifier | Within-User | Cross-User | Metric | Features | Gesture Type | Hands | HF | Ori. | Bew. | Loc. | Dyn. | Mobile | Gestures | TN | Samples |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cote-Allard et al. [53] | 2019 | sEMG (8 electrodes) | CNN | 97,88% | - | Acc. | 448 | Hand gestures | 1 | ✓ | ✓ | ✗ | ✗ | ✗ | ✓ | 7 | 36 | 1.960 |
| Jiang et al. [108] | 2017 | sEMG (4 electrodes), IMU | LDA | 92,60% | - | Acc. | 4 | Finger gestures | 1 | ✓ | ✓ | ✗ | ✗ | ✓ | ✓ | 8 | 10 | 1.280 |
| Simão et al. [197] | 2018 | sEMG (16 channels) | GAN | 95,01% | - | Acc. | 16 | Hand gestures | 1 | ✓ | ✓ | ✗ | ✗ | ✓ | ✓ | 8 | - | 880 |
| Antuvan and Masia [13] | 2019 | sEMG (6 electrodes) | LDA | 87,86% | - | Acc. | 6 | Hand gestures | 1 | ✓ | ✓ | ✗ | ✗ | ✗ | ✓ | 8 | 12 | 690 |
| Shukor et al. [195] | 2015 | Data glove (1 Accel., 10 inclination) | Template Matching | 88,88% | - | Acc. | 13 | BIM | 1 | ✓ | ✓ | ✗ | ✗ | ✗ | (✓) | 9 | 4 | 360 |
| Lu et al. [132] | 2016 | Data glove (18 IMUs) | ELM | 89,59% | - | Acc. | 54 | Numbers | 1 | ✓ | ✓ | ✗ | ✗ | ✗ | ✓ | 10 | - | - |
| Lu et al. [132] | 2016 | Data glove (18 IMUs) | SVM | 83,59% | - | Acc. | 54 | Numbers | 1 | ✓ | ✓ | ✗ | ✗ | ✗ | ✓ | 10 | - | - |
| Fang et al. [68] | 2017 | Data glove (18 IMUs) | ELM | 89,59% | - | Acc. | 54 | Numbers | 1 | ✓ | ✓ | ✗ | ✗ | ✗ | ✓ | 10 | 2 | 20 |
| Abreu et al. [2] | 2016 | sEMG (8 electrodes), IMU | SVM | 98,60% | - | Acc. | 8 | Libras letters | 1 | ✓ | ✓ | ✗ | ✗ | ✗ | (✓) | 20 | 1 | 13.600 |
| Mummadi et al. [152] | 2017 | Data glove (5 IMUs) | RF | 92,00% | - | Acc. | 10 | LSF letters | 1 | ✓ | ✓ | ✗ | ✗ | ✗ | ✓ | 21 | 57 | 1.250.000 |
| Abualola et al. [3] | 2016 | Data glove (6 IMUs) | LDA | 85,00% | - | Acc. | 5 | ASL letters | 1 | ✓ | ✓ | ✗ | ✗ | ✗ | ✓ | 24 | 9 | 216 |
| Yuan et al. [239] | 2021 | Data glove (10 Flex, 1 Gyro.), 2 IMUs | DFFN | 99,93% | - | Prec. | - | ASL letters | 1 | ✓ | ✓ | ✗ | ✗ | ✓ | ✓ | 24 | 6 | 5.178 |
| **Achenbach et al. [7]** | **2022** | **Data glove (8 IMUs)** | **SVM** | **99,50%** | **82,40%** | **Acc.** | **19** | **Finger gestures** | **1** | ✓ | ✓ | ✗ | ✗ | ✗ | ✓ | **25** | **9** | **2.250** |
| Bukhari et al. [33] | 2015 | Data glove (9 Flex, 1 IMU, 9 Touch) | PCA | 92,00% | - | Acc. | 17 | ASL letters | 1 | ✓ | ✓ | ✓ | ✗ | (✓) | (✓) | 26 | - | 520 |
| Jani et al. [107] | 2018 | Data glove (3 Flex, 1 IMU) | DTW + NMA | 96,50% | - | Acc. | 8 | ASL letters | 1 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 26 | 8 | 2.080 |
| Islam et al. [106] | 2022 | Data glove (1 IR, 1 Gyro.) | Template Matching | 97,72% | - | Acc. | - | IS letters/numbers | 1 | ✓ | ✓ | (✓) | ✗ | ✗ | ✓ | 46 | >12 | 1.700 |

Table A.7: Overview of body-worn sensing-based recognition approaches for gesture-sign-like gestures and simple signs

| Source | Year | Hardware | Classifier | Within-User | Cross-User | Metric | Features | Gesture Type | Hands | HF | Ori. | Bew. | Loc. | Dyn. | Mobile | Gestures | TN | Samples |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Melo [141] | 2018 | Data glove (7 IMUs) | HMM | 89,00% | - | Acc. | 15 | Hand gestures | 1 | ✓ | ✓ | ✓ | ✗ | ✗ | ✓ | 2 | - | 400 |
| Wen et al. [229] | 2020 | Data glove | CNN | 99,20% | - | Acc. | 10 | Hand gestures | 2 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 4 | - | 1.600 |
| Kau et al. [113] | 2015 | Data glove (10 Flex, 2 IMU) | Template Matching | 94,56% | - | Acc. | 20 | TSL | 2 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 5 | 5 | 1.250 |
| Smith et al. [199] | 2016 | iEMG (6 electrodes) | LR | 99,70% | - | TSR | 6 | Hand gestures | 1 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 6 | 8 | 2.592 |
| Melo [141] | 2018 | Data glove (7 IMUs) | HMM | 76,00% | - | Acc. | 15 | Hand gestures | 1 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 7 | - | 60 |
| Chu et al. [141] | 2021 | Data glove (5 Flex, 1 IMU) | SVM | 96,90% | 82,50% | Acc. | 28 | JSL | 1 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 7 | 3 | 2.100 |
| Dong et al. [60] | 2021 | Data glove (10 Flex, 2 IMU) | DGDL-GR | - | 98,69% | Acc. | 16 | ASL, CSL | 2 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 10 | 20 | 1.127 |
| Saggio et al. [179] | 2020 | Data glove (10 Flex, 6 IMUs) | CNN | 98,00% | - | Acc. | 22 | LIS | 2 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 10 | 7 | 7.000 |
| Calado et al. [35] | 2021 | Data glove (10 Flex, 6 IMUs) | ANN | - | 95,07% | Acc. | 288 | LIS | 2 | ✓ | ✓ | ✓ | ✗ | (✓) | ✓ | 10 | 17 | 10.596 |
| Calado et al. [35] | 2021 | Data glove (10 Flex, 3 IMUs) | ANN | - | 93,91% | Acc. | 191 | LIS | 2 | ✓ | ✓ | ✓ | ✗ | (✓) | ✓ | 10 | 17 | 10.596 |
| Calado et al. [36] | 2022 | Data glove (10 Flex, 6 IMUs) | Template Matching | 99,77% | 91,56% | Acc. | 46 | LIS | 2 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 10 | 5 | 5.000 |
| Wen et al. [229] | 2020 | Data glove | CNN | 95,23% | - | Acc. | 10 | Hand gestures | 2 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 11 | - | 2.200 |
| Fang et al. [68] | 2017 | Data glove (18 IMUs) | ELM | 82,50% | - | Acc. | 54 | Hand gestures | 1 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 16 | 2 | 20 |
| Dutta et al. [64] | 2022 | Data glove (6 Flex, 3 Force, 1 IMU) | SVM | 92,00% | - | Acc. | 15 | ARAT | 1 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 19 | 20 | 380 |
| Paudyal et al. [167] | 2016 | sEMG (16 electrodes), 2 IMUs | Template Matching | 97,72% | - | Acc. | 34 | ASL | 2 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 20 | 10 | 200 |
| Plawiak et al. [172] | 2016 | Data glove (3 Accels., 5 Flex, 2 Gyros.) | SVM | 98,32% | - | Acc. | 10 | Hand gestures | 1 | ✓ | ✓ | ✓ | ✗ | ✓ | (✓) | 22 | 10 | 2.200 |
| Simão et al. [197] | 2018 | Data glove (22 Flex), motion tracker | GAN | 90,20% | - | Acc. | 15 | Hand gestures | 1 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 24 | 8 | 2.400 |
| Müezzinoğlu and Karaköse [148] | 2021 | Data glove (10 Flex, 2 IMUs) | SVM | 98,02% | - | Acc. | 22 | Hand gestures | 2 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 25 | 20 | 49.000 |
| Rocha Silva et al. [175] | 2017 | Data glove (5 Flex, 9 IMUs) | ANN | 95,80% | - | Acc. | 53 | SL letters | 1 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 26 | 1 | 2.600 |
| Simoes Dias et al. [198] | 2022 | Data glove (2 contact, 5 Flex, 1 IMU) | RF | 96,15% | 85,18% | Acc. | 68 | Libras | 1 | ✓ | ✓ | ✓ | ✗ | ✓ | (✓) | 26 | 5 | 1.950 |
| Pezzuoli et al. [170] | 2020 | Data glove (10 Flex, 2 IMUs) | SVM | - | 96,70% | Acc. | 96 | Hand gestures | 1 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 27 | 5 | 810 |
| Pezzuoli et al. [170] | 2020 | Data glove (10 Flex, 2 IMUs) | RF | 99,70% | - | Acc. | 96 | Hand gestures | 1 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 27 | 5 | 810 |
| Lee and Lee [123] | 2018 | Data glove (5 Flex, 1 IMU, 2 Touch) | SVM | - | 98,20% | Acc. | 10 | ASL | 1 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 28 | 12 | 3.360 |
| Luzhnica et al. [134] | 2016 | Data glove (10 Flex, 5 Force, 7 IMUs) | LDA + LR | 98,00% | - | Acc. | 724 – 1284 | Hand gestures | 1 | ✓ | ✓ | ✓ | ✗ | ✓ | (✓) | 31 | 18 | 2.790 – 5.580 |
| Saleh et al. [181] | 2020 | Data glove (5 Flex, 1 IMU) | - | 90,00% | - | Acc. | 16 | ArSL | 2 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 35 | 2 | - |
| Haq et al. [92] | 2018 | Data glove (5 Flex, 1 IMU) | Template Matching | 91,00% | - | Acc. | 8 | SIBI | 1 | ✓ | ✓ | ✓ | ✗ | (✓) | ✓ | 36 | - | 360 |
| Shaheen and Mehmood [191] | 2018 | Data glove (10 Flex, 2 IMUs) | Template Matching | 93,01% | - | Acc. | 32 | PSL | 2 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 36 | 10 | 1.050 |
| Tubaiz et al. [219] | 2015 | Data glove (1 Accel., 5 Flex) | MkNN | 98,89% | - | Acc. | - | ArSL | 2 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 40 | 1 | 400 |
| Wu et al. [236] | 2015 | sEMG (4 electrodes), IMU | SVM | 95,14% | ~40,00% | Acc. | 30 | ASL | 1 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 40 | 2 | 12.000 |
| Galka et al. [73] | 2016 | Data glove (7 Accels.) | Parallel HMM | 99,75% | - | Sens. | 21 | SL signs | 1 | ✓ | ✓ | ✓ | ✗ | ✓ | (✓) | 40 | 5 | 2.000 |
| Yuan et al. [239] | 2021 | Data glove (10 Flex, 1 Gyro.), 2 IMUs | DFFN | 96,07% | - | F1 | - | CSL | 2 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 72 | 6 | 34.452 |
| Wu et al. [235] | 2016 | sEMG (4 electrodes), IMU | SVM | 96,16% | 85,24% | Acc. | 40 | ASL | 1 | ✓ | ✓ | ✓ | ✗ | ✓ | ✓ | 80 | 2 | 24.000 |

Table A.8: Overview of body-worn sensing-based recognition approaches for extensive signs

| Source | Year | Hardware | Classifier | Within-User | Cross-User | Metric | Features | Gesture Type | Hands | HF | Ori. | Bew. | Loc. | Dyn. | Mobile | Gestures | TN | Samples |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Oz and Leu [161] | 2011 | Data glove (15 Flex), position tracker | ANN | 90,00% | - | Acc. | 151 | ASL | 1 | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | 50 | 1 | 300 |
| Oz and Leu [161] | 2011 | Data glove (15 Flex), position tracker | ANN | 92,00% | - | Acc. | 151 | ASL | 1 | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | 50 | < 7 | 300 |
| Rozado et al. [177] | 2011 | Data glove, position tracker | HTM | 91,00% | - | Acc. | 11 | Auslan | 2 | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | 95 | 1 | 2.565 |
| Pezzuoli et al. [170] | 2020 | Data glove, position tracker | ANN | 97,40% | - | Acc. | 124 | Auslan | 2 | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | 95 | 1 | 2.565 |
| Mohandes and Deriche [145] | 2013 | Data glove, position tracker | LDA + MD | 98,10% | - | Acc. | 20 | ArSL | 2 | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | 100 | 1 | 2.000 |

## A.3 EVALUATION OF HANDSHAPE RECOGNITION

Table A.9: Evaluated parameter ranges for (pre-) grid search hyperparameter optimization with number of possible combinations

| **Parameter** | **Pre-Gridsearch Range** | (#) | **Gridsearch Range** | (#) |
|---|---|---|---|---|
| | | **SVM** | | |
| $C$ | $2^{-5}, 2^{-3}, \ldots, 2^{15}$ | (231) | $2^{0}, 2^{1}, \ldots, 2^{4}$ | (35) |
| $\gamma$ | $2^{-15}, 2^{-14}, \ldots, 2^{5}$ | | $2^{-8}, 2^{-7}, \ldots, 2^{-2}$ | |
| | | **RF** | | |
| *criterion* | gini, entropy | (220) | gini, entropy | (60) |
| *max_features* | $1, 2, \ldots, 10$ | | $1, 2, \ldots, 6$ | |
| *n_estimators* | $2^{0}, 2^{1}, \ldots, 2^{10}$ | | $2^{6}, 2^{7}, \ldots, 2^{10}$ | |
| | | **LR** | | |
| *penalty* | elasticnet | (440) | elasticnet | (40) |
| *solver* | newton-cg, lbfgs, sag, saga | | saga | |
| $C$ | $2^{-5}, 2^{-3}, \ldots, 2^{15}$ | | $2^{-5}, 2^{-4}, \ldots, 2^{4}$ | |
| $l_1$*_ratio* | $0.1, 0.2, \ldots, 1$ | | $0.2, 0.3, \ldots, 0.5$ | |
| *penalty* | none, $l_1$, $l_2$ | (132) | $l_2$ | (40) |
| *solver* | newton-cg, lbfgs, sag, saga | | newton-cg, lbfgs, sag, saga | |
| $C$ | $2^{-5}, 2^{-3}, \ldots, 2^{15}$ | | $2^{-5}, 2^{-4}, \ldots, 2^{4}$ | |
| | | **FNN** | | |
| *solver* | lbfgs | (36) | - | (0) |
| *hidden_layer_sizes* | $(2^2, 2^2), (2^3, 2^3), \ldots, (2^{10}, 2^{10})$ | | - | |
| *alpha* | $e^{-4}, e^{-3}, e^{-2}, e^{-1}$ | | - | |

Table A.10: Most frequent hyperparameter combinations depending on the investigated data preprocessing method for the 27 hand-shape dataset and the 56 hand-shape dataset

| **Model** | **Parameter** | **Data preprocessing** | | | |
|---|---|---|---|---|---|
| | | None | Out. | Feat. | Both |
| **27 Hand Shapes** | | | | | |
| SVM | $C$ | $2^4$ | $2^4$ | $2^4$ | $2^4$ |
| | $\gamma$ | $2^{-5}$ | $2^{-6}$ | $2^{-5}$ | $2^{-6}$ |
| RF | *criterion* | entropy | gini | entropy | entropy |
| | *max_features* | 3 | 2 | 2 | 2 |
| | *n_estimators* | $2^{10}$ | $2^{10}$ | $2^{10}$ | $2^{10}$ |
| LR | *penalty* | $l_2$ | $l_2$ | $l_2$ | $l_2$ |
| | *solver* | saga | lbfgs | saga | saga |
| | $C$ | $2^1$ | $2^1$ | $2^1$ | $2^1$ |
| **56 Hand Shapes** | | | | | |
| SVM | $C$ | $2^4$ | $2^4$ | $2^4$ | $2^4$ |
| | $\gamma$ | $2^{-5}$ | $2^{-6}$ | $2^{-5}$ | $2^{-6}$ |
| RF | *criterion* | entropy | gini | entropy | entropy |
| | *max_features* | 3 | 2 | 2 | 2 |
| | *n_estimators* | $2^{10}$ | $2^{10}$ | $2^{10}$ | $2^{10}$ |
| LR | *penalty* | $l_2$ | $l_2$ | $l_2$ | $l_2$ |
| | *solver* | saga | lbfgs | saga | saga |
| | $C$ | $2^1$ | $2^1$ | $2^1$ | $2^1$ |

Outlier Detection (Out.), Feature Selection (Feat.)

Table A.11: Number of recorded handshapes and handshapes used for handshape classification with original and transformed labels

| Original | (#) | Transformed | (#) | Used? |
|---|---|---|---|---|
| Curved 1 | (174) | Curved 1 | (408) | yes |
| Curved V | (174) | Curved V | (348) | yes |
| Curved L | (174) | Curved L | (348) | yes |
| Flat N | (234) | Flat N | (234) | yes |
| Q | (234) | Q | (234) | yes |
| K | (234) | K | (234) | yes |
| C | (174) | C | (174) | yes |
| S | (174) | S | (174) | yes |
| D | (174) | D | (174) | yes |
| Open B | (174) | Open B | (174) | yes |
| 7 | (174) | 7 | (174) | yes |
| Horns | (174) | Horns | (174) | yes |
| Curved H | (174) | Curved H | (174) | yes |
| W | (174) | W | (174) | yes |
| Flat 4 | (174) | Flat 4 | (174) | yes |
| T | (174) | T | (174) | yes |
| Flat ILY | (174) | Flat ILY | (174) | yes |
| U | (174) | U | (174) | yes |
| Curved 4 | (174) | Curved 4 | (174) | yes |
| O | (174) | O | (174) | yes |
| I | (174) | I | (174) | yes |
| Closed E | (174) | Closed E | (174) | yes |
| 3 | (174) | 3 | (174) | yes |
| R | (174) | R | (174) | yes |
| Flat M | (174) | Flat M | (174) | yes |
| F | (174) | F | (174) | yes |
| L | (174) | L | (174) | yes |
| Baby O | (174) | Baby O | (174) | yes |
| Spread Flat 5 | (174) | Spread Flat 5 | (174) | yes |
| Open 8 | (174) | Open 8 | (174) | yes |
| Open F | (174) | Open F | (174) | yes |
| Flat B | (174) | Flat B | (174) | yes |
| N | (174) | N | (174) | no |
| Open H | (174) | Open H | (174) | yes |
| Goody Goody | (174) | Goody Goody | (174) | yes |

Continued on next page

Continued from previous page

| **Original** | (#) | **Transformed** | (#) | **Used?** |
|---|---|---|---|---|
| Closed B | (174) | Closed B | (174) | yes |
| Stacked 5 | (174) | Stacked 5 | (174) | yes |
| Flat 1 | (174) | Flat 1 | (174) | yes |
| E | (174) | E | (174) | yes |
| V | (174) | V | (174) | yes |
| A | (174) | A | (174) | yes |
| Beak | (174) | Beak | (174) | yes |
| 1 | (174) | 1 | (174) | yes |
| Flat Horns | (174) | Flat Horns | (174) | yes |
| Y | (174) | Y | (174) | yes |
| 4 | (174) | 4 | (174) | yes |
| ILY | (174) | ILY | (174) | yes |
| M | (174) | M | (174) | no |
| 8 | (174) | 8 | (174) | yes |
| Flat V | (174) | Flat V | (174) | yes |
| Curved 5 | (174) | Curved 5 | (174) | yes |
| Spread E | (174) | Spread E | (174) | yes |
| Open E | (174) | Open E | (174) | yes |
| 5 | (174) | 5 | (174) | yes |
| Spread Open E | (174) | Spread Open E | (174) | yes |
| Open A | (114) | Open A | (114) | no |
| B | (114) | B | (114) | no |
| Flat H | (114) | Flat H | (114) | yes |
| Open Horns | (114) | Open Horns | (114) | no |
| J | (60) | J | (60) | no |
| Z | (60) | Z | (60) | no |
| X | (234) | Curved 1 | (0) | no |
| Bent V | (174) | Curved V | (0) | no |
| Bent L | (174) | Curved L | (0) | no |
| Total | 10,908 | | 10,908 | 10,098 |

Table A.12: Classification report for the final handshape classifier. Reported are Precision, Recall, $F_1$-score, and Support for each handshape as well as aggregated average values

| Label | Precision | Recall | $F_1$-Score | Support |
|---|---|---|---|---|
| 1 | 0.86 | 0.84 | 0.85 | 174 |
| Curved 1 | 0.87 | 0.91 | 0.89 | 408 |
| Flat 1 | 0.81 | 0.73 | 0.77 | 174 |
| 3 | 0.89 | 0.92 | 0.91 | 174 |
| 4 | 0.76 | 0.80 | 0.78 | 174 |
| Curved 4 | 0.75 | 0.67 | 0.71 | 174 |
| Flat 4 | 0.89 | 0.92 | 0.91 | 174 |
| 5 | 0.95 | 0.95 | 0.95 | 174 |
| Curved 5 | 0.84 | 0.74 | 0.78 | 174 |
| Spread Flat 5 | 0.89 | 0.86 | 0.87 | 174 |
| Stacked 5 | 0.91 | 0.93 | 0.92 | 174 |
| 7 | 0.96 | 0.97 | 0.96 | 174 |
| 8 | 0.94 | 0.91 | 0.92 | 174 |
| Open 8 | 0.89 | 0.89 | 0.89 | 174 |
| A | 0.91 | 0.88 | 0.89 | 174 |
| Closed B | 0.91 | 0.95 | 0.93 | 174 |
| Flat B | 0.98 | 0.98 | 0.98 | 174 |
| Open B | 0.91 | 0.98 | 0.94 | 174 |
| Beak | 0.88 | 0.97 | 0.92 | 174 |
| C | 0.90 | 0.86 | 0.88 | 174 |
| D | 0.85 | 0.88 | 0.86 | 174 |
| E | 0.71 | 0.78 | 0.74 | 174 |
| Closed E | 0.78 | 0.76 | 0.77 | 174 |
| Open E | 0.79 | 0.81 | 0.80 | 174 |
| Spread Open E | 0.77 | 0.73 | 0.75 | 174 |
| Spread E | 0.68 | 0.67 | 0.68 | 174 |
| F | 0.92 | 0.91 | 0.92 | 174 |
| Open F | 0.92 | 0.92 | 0.92 | 174 |
| Goody Goody | 0.97 | 0.96 | 0.96 | 174 |
| Curved H | 0.75 | 0.49 | 0.60 | 174 |
| Flat H | 0.98 | 0.82 | 0.90 | 114 |

Continued on next page

Continued from previous page

| Label | Precision | Recall | $F_1$-Score | Support |
|---|---|---|---|---|
| Open H | 0.95 | 0.93 | 0.94 | 174 |
| Horns | 0.89 | 0.93 | 0.91 | 174 |
| Flat Horns | 0.88 | 0.91 | 0.90 | 174 |
| I | 0.98 | 0.98 | 0.98 | 174 |
| ILY | 0.87 | 0.91 | 0.89 | 174 |
| Flat ILY | 0.91 | 0.90 | 0.90 | 174 |
| K | 0.87 | 0.92 | 0.90 | 234 |
| L | 0.89 | 0.96 | 0.93 | 174 |
| Curved L | 0.92 | 0.89 | 0.91 | 348 |
| Flat M | 0.94 | 0.93 | 0.93 | 174 |
| Flat N | 0.66 | 0.79 | 0.72 | 174 |
| O | 0.78 | 0.82 | 0.80 | 174 |
| Baby O | 0.91 | 0.80 | 0.85 | 174 |
| Q | 0.88 | 0.88 | 0.88 | 234 |
| R | 0.92 | 0.90 | 0.91 | 174 |
| S | 0.73 | 0.78 | 0.76 | 174 |
| T | 0.88 | 0.84 | 0.86 | 174 |
| U | 0.88 | 0.88 | 0.88 | 174 |
| V | 0.85 | 0.93 | 0.89 | 174 |
| Curved V | 0.74 | 0.82 | 0.78 | 348 |
| Flat V | 0.70 | 0.60 | 0.65 | 174 |
| W | 0.98 | 0.97 | 0.97 | 174 |
| Y | 0.99 | 1 | 0.99 | 174 |
| accuracy | 0.87 | 0.87 | 0.87 | 10,038 |
| macro avg | 0.87 | 0.86 | 0.86 | 10,038 |
| weighted avg | 0.87 | 0.87 | 0.86 | 10,038 |

## A.4 EVALUATION

In the following, an overview of all signs recorded as part of the evaluation of the *ParaSignRec* is provided (see Section 6.2.1). In total, 210 different signs were recorded 1,877 times, divided across three lists, which are shown in Table A.13 to Table A.19. Following the lists, Table A.22 provides an overview of which signs were recorded how often by which participant.

The signs are presented within the tables sorted by their name. The tables include the following columns:

- **Sign**: Name of the sign, derived from the sign ID (e.g. "all_together" as ALL TOGETHER)
- **BST**: Battison Sign Type (see Section 2.1.2), i.e.
  - **Type 0**: One-handed sign, corresponds to type ∅/X
  - **Type 1**: Two-handed, symmetrical sign
  - **Type 2**: Two-handed, asymmetrical sign with the same handshape
  - **Type 3**: Two-handed, asymmetrical sign with different handshape
- **Comp.**: The phonological complexity of the recorded sign on a scale from 0 (easy) to 6 (difficult), taken from *ASL-LEX*. The maximum phonological complexity was limited to 4 during recording.
- **HANDSHAPE**: Handshape of the dominant (d) and non-dominant (nd) hand at the beginning and end of the execution.
- **LOCATION**: Location of the dominant (d) and non-dominant (nd) hand at the beginning and end of the execution.
- **MOVEMENT**: Movement path of the dominant (d) and non-dominant (nd) hand.
- **$FC_x$**: Change in finger flexion of the dominant (d) and non-dominant (nd) hand, specified for each of the four fingers:
  **i**: index finger (*index*), **m**: middle finger, **r**: ring finger, **p**: little finger (*pinky*).
- **$SC_x$**: Change in finger spreading of the dominant (d) and non-dominant (nd) hand, specified for the following two finger pairings:
  **im**: index finger / middle finger, **rp**: ring finger / little finger.
- **UR**: Rotation of the dominant (d) and non-dominant forearm (ulnar).
- **ORIENTATION**: Orientation of the dominant (d) and non-dominant (nd) hand at the beginning and end of the execution.

**Note:** The class * is equivalent to all available classes of the parameter concerned and is used when the parameter is not relevant to the meaning of the sign. This may occur, for example, for the non-dominant handshape of one-handed signs or in a few cases of FC and UR, where it could not be determined unambiguously whether the corresponding movement is present or not.

Table A.13: List 1 of the recorded signs with the parameters examined in this work (Part 1)

| Sign | BST | Complex. | Handshape(End) | | Handshape(Start) | |
|---|---|---|---|---|---|---|
| | | | d | nd | d | nd |
| Acquire | Type 1 | 2 | *8*, *Open 8* | *8*, *Open 8* | *8*, *Open 8* | *8*, *Open 8* |
| Adopt | Type 1 | 1 | *A*, *S* | *A*, *S* | *5*, *Spread Flat 5* | *5*, *Spread Flat 5* |
| Airplane | Type 0 | 2 | *Ily*, *Curved L* | * | *Ily*, *Curved L* | * |
| Badge | Type 1 | 1 | *Curved L* | *Curved L* | *Curved L* | *Curved L* |
| Brag | Type 1 | 2 | *A*, *T* | *A*, *T* | *A*, *T* | *A*, *T* |
| Bury 3 | Type 1 | 1 | *5*, *Open B* | *5*, *Open B* | *5*, *Open B* | *5*, *Open B* |
| Carry | Type 1 | 1 | *5*, *Open B* | *5*, *Open B* | *5*, *Open B* | *5*, *Open B* |
| Cheese grater | Type 3 | 3 | *5*, *Curved 5*, *C*, *Open E*, *Spread Open E* | *O*, *S* | *5*, *Curved 5*, *C*, *Open E*, *Spread Open E* | *O*, *S* |
| Come | Type 0 | 2 | *Flat B*, *Open B* | * | *Closed B*, *Open B* | * |
| Correct | Type 3 | 4 | *1*, *Curved 1*, *D* | *Closed B* | *1*, *Curved 1*, *D* | *Closed B* |
| Credit card | Type 3 | 2 | *A*, *S*, *T* | *Closed B* | *A*, *S*, *T* | *Closed B* |
| Cry 3 | Type 1 | 2 | *4*, *Curved 4* | *4*, *Curved 4* | *4*, *Curved 4* | *4*, *Curved 4* |
| Dictionary | Type 3 | 4 | *D* | *Closed B* | *D* | *Closed B* |
| Dig up | Type 1 | 3 | *A*, *S* | *A*, *S* | *A*, *S* | *A*, *S* |
| Doctor 2 | Type 3 | 2 | *D*, *Horns* | *1*, *5*, *A*, *Closed B*, *C*, *D*, *Horns*, *O*, *S* | *D*, *Horns* | *1*, *5*, *A*, *Closed B*, *C*, *D*, *Horns*, *O*, *S* |
| Drag | Type 1 | 1 | *A*, *Open E*, *O*, *S* | *A*, *Open E*, *O*, *S* | *5*, *Curved 5*, *Spread Flat 5* | *5*, *Curved 5*, *Spread Flat 5* |
| Drawer | Type 1 | 1 | *Curved 5*, *E*, *Closed E*, *Spread Open E*, *Spread E*, *S* | *Curved 5*, *E*, *Closed E*, *Spread Open E*, *Spread E*, *S* | *4*, *Curved 4*, *5*, *Curved 5*, *S* | *4*, *Curved 4*, *5*, *Curved 5*, *S* |
| Each other | Type 1 | 3 | *A* | *A* | *A* | *A* |
| Edit 1 | Type 1 | 1 | *Curved 1* | *Curved 1* | *Curved 1* | *Curved 1* |
| Evaluate | Type 1 | 1 | *E*, *Closed E*, *Spread E*, *O* | *E*, *Closed E*, *Spread E*, *O* | *E*, *Closed E*, *Spread E*, *O* | *E*, *Closed E*, *Spread E*, *O* |
| Experiment | Type 1 | 2 | *E*, *Spread E* | *E*, *Spread E* | *E*, *Spread E* | *E*, *Spread E* |
| Explanation | Type 1 | 1 | *F* | *F* | *F* | *F* |
| Fade | Type 1 | 1 | *A*, *T* | *A*, *T* | *Beak* | *Beak* |
| Frustrate | Type 0 | 1 | *4*, *Closed B* | * | *4*, *Closed B* | * |
| Get off | Type 3 | 4 | *Curved H*, *Curved V* | *C*, *Spread Open E* | *Curved H*, *Curved V* | *C*, *Spread Open E* |
| Gloves 4 | Type 2 | 1 | *4*, *5*, *Spread Flat 5*, *Closed B*, *Open B* | *4*, *5*, *Spread Flat 5*, *Closed B*, *Open B* | *5*, *Spread Flat 5* | *5*, *Spread Flat 5* |
| Grieve | Type 1 | 2 | *A*, *O*, *S* | *A*, *O*, *S* | *5* | *5* |
| Hairdresser | Type 1 | 3 | *V* | *V* | *V* | *V* |
| Handsome 3 | Type 0 | 1 | *A* | * | *Open E*, *Spread Open E* | * |
| Hunt | Type 1 | 2 | *3*, *Open H*, *L* | *3*, *Open H*, *L* | *3*, *Open H*, *L* | *3*, *Open H*, *L* |
| Hurdle 1 | Type 3 | 4 | *A*, *T* | *1*, *D* | *K* | *1*, *D* |
| Inspiring | Type 0 | 1 | *5*, *Spread Flat 5* | * | *Spread Open E*, *O* | * |
| Japan | Type 1 | 1 | *Curved L*, *Baby O*, *Q*, *T* | *Curved L*, *Baby O*, *Q*, *T* | *L*, *Curved L*, *Baby O*, *Q* | *L*, *Curved L*, *Baby O*, *Q* |
| Lick envelope 2 | Type 1 | 2 | *Flat H* | *Flat H* | *Flat H* | *Flat H* |
| Look at | Type 0 | 2 | *V*, *Curved V*, *Flat V* | * | *V*, *Curved V*, *Flat V* | * |
| Match | Type 1 | 1 | *Curved 5*, *Open E*, *Spread Open E* | *Curved 5*, *Open E*, *Spread Open E* | *Curved 5*, *Open E*, *Spread Open E* | *Curved 5*, *Open E*, *Spread Open E* |
| Measure 2 | Type 1 | 2 | *Y* | *Y* | *Y* | *Y* |
| Meditate 3 | Type 1 | 3 | *Stacked 5*, *W* | *Stacked 5*, *W* | *Stacked 5*, *W* | *Stacked 5*, *W* |

Continued on next page

Continued from previous page

| Sign | BST | Complex. | Handshape(End) | | Handshape(Start) | |
|---|---|---|---|---|---|---|
| | | | d | nd | d | nd |
| Mischievous | Type 1 | 3 | *A*, *Open E*, *Spread Open E*, *Curved V* | *A*, *Open E*, *Spread Open E*, *Curved V* | *3*, *Open H* | *3*, *Open H* |
| Mix 3 | Type 1 | 2 | *1*, *Curved 1*, *D* | *1*, *Curved 1*, *D* | *1*, *Curved 1*, *D* | *1*, *Curved 1*, *D* |
| Monkey | Type 1 | 1 | *5*, *Curved 5*, *Flat B*, *Spread Open E* | *5*, *Curved 5*, *Flat B*, *Spread Open E* | *5*, *Curved 5*, *Spread Open E* | *5*, *Curved 5*, *Spread Open E* |
| Museum | Type 1 | 2 | *4*, *Flat M*, *W* | *4*, *Flat M*, *W* | *4*, *Flat M*, *W* | *4*, *Flat M*, *W* |
| Next to | Type 2 | 2 | *Flat B*, *C* | *Flat B*, *C* | *Flat B*, *C* | *Flat B*, *C* |
| Niece | Type 0 | 2 | *Flat M*, *Flat N*, *Flat V* | * | *Flat M*, *Flat N*, *Flat V* | * |
| Octopus | Type 3 | 3 | *Stacked 5*, *Beak* | *O* | *5*, *Stacked 5* | *O* |
| Paper | Type 2 | 2 | *5*, *Open B* | *5*, *Open B* | *5*, *Open B* | *5*, *Open B* |
| Problem 1 | Type 1 | 2 | *Curved V* | *Curved V* | *Curved V* | *Curved V* |
| Ready | Type 1 | 3 | *R*, *U* | *R*, *U* | *R*, *U* | *R*, *U* |
| Scarcely | Type 1 | 2 | *Open 8* | *Open 8* | *Open 8* | *Open 8* |
| Scout | Type 0 | 2 | *W* | * | *W* | * |
| Sentence | Type 1 | 1 | *F* | *F* | *F* | *F* |
| Shark 1 | Type 0 | 1 | *4*, *Closed B* | * | *4*, *Closed B* | * |
| Shooting | Type 1 | 3 | *L* | *L* | *L* | *L* |
| Shoulder | Type 0 | 1 | *Flat B*, *Open B* | * | *Flat B*, *Open B* | * |
| Since | Type 1 | 2 | *1*, *Curved 1*, *Flat 1*, *D* | *1*, *Curved 1*, *Flat 1*, *D* | *1*, *Curved 1*, *Flat 1*, *D* | *1*, *Curved 1*, *Flat 1*, *D* |
| Spend | Type 1 | 2 | *Curved L*, *Baby O* | *Curved L*, *Baby O* | *Baby O* | *Baby O* |
| Spread 3 | Type 2 | 3 | *5*, *Closed B*, *Open B* | *5*, *Closed B*, *Open B* | *5*, *Closed B*, *Open B* | *5*, *Closed B*, *Open B* |
| Starbucks | Type 1 | 2 | *Curved 5*, *Horns*, *Ily* | *Curved 5*, *Horns*, *Ily* | *Horns* | *Horns* |
| Suspenders 1 | Type 1 | 2 | *Curved H*, *Flat N*, *U*, *V*, *Curved V* | *Curved H*, *Flat N*, *U*, *V*, *Curved V* | *Curved H*, *U*, *V*, *Curved V* | *Curved H*, *U*, *V*, *Curved V* |
| Sweater | Type 1 | 2 | *5*, *Curved 5*, *Spread Flat 5*, *C* | *5*, *Curved 5*, *Spread Flat 5*, *C* | *5*, *Curved 5*, *Spread Flat 5*, *C* | *5*, *Curved 5*, *Spread Flat 5*, *C* |
| Thin | Type 1 | 2 | *I*, *Y* | *I*, *Y* | *I*, *Y* | *I*, *Y* |
| Transgender | Type 0 | 2 | *Curved 5*, *Beak* | * | *4*, *5*, *Curved 5* | * |
| Visit | Type 1 | 4 | *K* | *K* | *K* | *K* |
| Wave | Type 1 | 3 | *4*, *5*, *Closed B*, *Open B* | *4*, *5*, *Closed B*, *Open B* | *4*, *5*, *Closed B*, *Open B* | *4*, *5*, *Closed B*, *Open B* |
| Weird | Type 0 | 2 | *W* | * | *W* | * |
| Wheelchair | Type 1 | 2 | *1*, *Curved 1*, *D* | *1*, *Curved 1*, *D* | *1*, *Curved 1*, *D* | *1*, *Curved 1*, *D* |
| Whisper | Type 0 | 1 | *4*, *Flat 4* | * | *4*, *Flat 4* | * |

Table A.14: List 1 of the recorded signs with the parameters examined in this work (Part 2)

| Sign | Orientation(Start) | | Location(Start) | | Movement | | $SC_{im}$ | | $SC_{rp}$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | d | nd | d | nd | d | nd | d | nd | d | nd |
| Acquire | *DownBack*, *FrontDown*, *FrontIn*, *InDown*, *UpIn* | *DownBack*, *FrontDown*, *FrontIn*, *InDown*, *UpIn* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Adopt | *DownBack*, *FrontDown* | *DownBack*, *FrontDown* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✓ | ✓ | ✓ | ✓ |
| Airplane | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Neutral* | *None* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Badge | *InDown*, *UpFront*, *UpIn* | *InDown*, *UpFront*, *UpIn* | *Shoulder* | *Shoulder* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Brag | *InDown*, *InFront* | *InDown*, *InFront* | *TorsoBottom* | *TorsoBottom* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Bury 3 | *FrontIn*, *InBack* | *FrontIn*, *InBack* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Carry | *FrontIn*, *FrontUp*, *OutFront*, *OutUp*, *UpBack*, *UpIn*, *UpOut* | *FrontIn*, *FrontUp*, *OutFront*, *OutUp*, *UpBack*, *UpIn*, *UpOut* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Cheese grater | *FrontIn*, *InBack* | *DownOut*, *InBack*, *InDown* | *NonDominant* | *Neutral* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Come | *FrontUp*, *OutUp*, *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Neutral* | *None* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Correct | *FrontIn*, *InBack*, *InDown*, *UpIn* | *FrontIn*, *FrontUp*, *InBack*, *OutUp* | *NonDominant* | *Neutral* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Credit card | *FrontIn*, *FrontUp*, *InBack*, *UpBack*, *UpIn* | *FrontIn*, *FrontUp*, *InUp* | *NonDominant* | *Neutral* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Cry 3 | *FrontIn*, *InBack*, *InDown*, *UpBack* | *FrontIn*, *InBack*, *InDown*, *UpBack* | *Head* | *Head* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Dictionary | *FrontDown*, *FrontIn*, *UpFront*, *UpIn* | *FrontIn*, *FrontUp*, *InUp* | *NonDominant* | *Neutral* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Dig up | *DownBack*, *DownFront*, *DownIn*, *DownOut*, *FrontIn*, *FrontUp* | *DownBack*, *DownFront*, *DownIn*, *DownOut*, *FrontIn*, *FrontUp* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Doctor 2 | *FrontDown*, *InDown*, *UpIn* | *FrontUp* | *NonDominant* | *Neutral* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Drag | *DownBack*, *FrontDown*, *InDown*, *OutDown*, *UpIn* | *DownBack*, *FrontDown*, *InDown*, *OutDown*, *UpIn* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✓ | ✓ | ✓ | ✓ |
| Drawer | *FrontUp*, *UpBack* | *FrontUp*, *UpBack* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✓ | ✓ | ✓ | ✓ |
| Each other | *FrontDown*, *FrontIn*, *InBack*, *InDown*, *InFront*, *UpFront*, *UpIn* | *FrontDown*, *FrontIn*, *InBack*, *InDown*, *InFront*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Edit 1 | *FrontDown*, *FrontIn*, *UpFront*, *UpIn* | *FrontDown*, *FrontIn*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Evaluate | *FrontDown*, *InDown*, *UpFront*, *UpIn* | *FrontDown*, *InDown*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Experiment | *FrontDown*, *UpFront* | *FrontDown*, *UpFront* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Explanation | *FrontDown*, *FrontIn*, *InDown*, *UpFront*, *UpIn* | *FrontDown*, *FrontIn*, *InDown*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Fade | *FrontIn*, *FrontUp*, *InBack*, *InUp* | *FrontIn*, *FrontUp*, *InBack*, *InUp* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Frustrate | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Get off | *FrontDown*, *InDown*, *UpFront*, *UpIn* | *FrontIn*, *InBack* | *NonDominant* | *Neutral* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Gloves 4 | *FrontDown* | *DownBack*, *FrontDown*, *FrontIn* | *NonDominant* | *Neutral* | *Straight* | *None* | ✓ | ✗ | ✓ | ✗ |
| Grieve | *DownBack*, *DownIn*, *FrontDown*, *FrontIn*, *InDown*, *InFront*, *InUp*, *UpIn* | *DownBack*, *DownIn*, *FrontDown*, *FrontIn*, *InDown*, *InFront*, *InUp*, *UpIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✓ | ✓ | ✓ | ✓ |
| Hairdresser | *InDown*, *OutUp*, *UpFront*, *UpIn* | *InDown*, *OutUp*, *UpFront*, *UpIn* | *Head* | *Head* | *Curved* | *Curved* | ✓ | ✓ | ✗ | ✗ |
| Handsome 3 | *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | *None* | *None* | ✗ | ✗ | ✗ | ✗ |
| Hunt | *FrontUp*, *InBack*, *UpFront*, *UpIn* | *FrontUp*, *InBack*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Hurdle 1 | *FrontDown*, *FrontIn*, *UpFront* | *DownOut*, *InBack*, *InDown*, *UpIn* | *NonDominant* | *Neutral* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Inspiring | *FrontUp*, *InBack*, *InUp*, *UpOut* | *DownBack*, *DownFront*, *DownIn* | *TorsoTop* | *None* | *Straight* | *None* | ✓ | ✗ | ✓ | ✗ |
| Japan | *FrontIn*, *FrontUp*, *UpBack*, *UpIn* | *FrontIn*, *FrontUp*, *UpBack*, *UpIn* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Lick envelope 2 | *FrontIn*, *InBack*, *InDown*, *UpBack*, *UpIn* | *FrontIn*, *InBack*, *InDown*, *UpBack*, *UpIn* | *Head* | *Head* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Look at | *FrontDown*, *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Match | *FrontIn*, *FrontUp*, *InBack*, *UpBack* | *FrontIn*, *FrontUp*, *InBack*, *UpBack* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Measure 2 | *InBack*, *InDown*, *InFront*, *UpFront*, *UpIn* | *InBack*, *InDown*, *InFront*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Meditate 3 | *InBack*, *UpBack*, *UpIn* | *InBack*, *UpBack*, *UpIn* | *Head* | *Head* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |

Continued on next page

Continued from previous page

| Sign | Orientation(Start) | | Location(Start) | | Movement | | $SC_{im}$ | | $SC_{rp}$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | d | nd | d | nd | d | nd | d | nd | d | nd |
| Mischievous | *InDown*, *UpFront*, *UpIn* | *InDown*, *UpFront*, *UpIn* | *Head* | *Head* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Mix 3 | *DownBack*, *DownOut*, *FrontDown*, *FrontIn*, *InBack*, *InDown* | *DownBack*, *DownOut*, *FrontDown*, *FrontIn*, *InBack*, *InDown* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Monkey | *DownIn*, *InBack*, *InUp* | *DownIn*, *InBack*, *InUp* | *TorsoBottom* | *TorsoBottom* | *None* | *None* | ✗ | ✗ | ✗ | ✗ |
| Museum | *FrontDown*, *FrontIn*, *UpFront* | *FrontDown*, *FrontIn*, *UpFront* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Next to | *FrontIn*, *OutFront* | *InBack* | *NonDominant* | *Neutral* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Niece | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Octopus | *DownBack*, *DownOut*, *FrontDown*, *FrontIn*, *InBack*, *InDown*, *UpFront*, *UpIn* | *DownOut*, *FrontDown*, *InBack*, *InDown* | *NonDominant* | *Neutral* | *Straight* | *None* | ✓ | ✗ | ✓ | ✗ |
| Paper | *FrontDown*, *FrontIn*, *InDown*, *UpFront*, *UpIn* | *DownFront*, *FrontIn*, *FrontOut*, *FrontUp*, *InBack*, *InUp* | *NonDominant* | *Neutral* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Problem 1 | *InBack*, *InDown*, *UpFront*, *UpIn* | *InBack*, *InDown*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | *None* | *None* | ✗ | ✗ | ✗ | ✗ |
| Ready | *FrontDown*, *InDown*, *InFront*, *UpFront*, *UpIn* | *FrontDown*, *InDown*, *InFront*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Scarcely | *FrontIn*, *FrontUp*, *InBack*, *UpBack*, *UpIn* | *FrontIn*, *FrontUp*, *InBack*, *UpBack*, *UpIn* | *Head* | *Head* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Scout | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Sentence | *FrontDown*, *FrontIn* | *FrontDown*, *FrontIn* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Shark 1 | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Shooting | *FrontIn*, *InBack*, *OutFront* | *FrontIn*, *InBack*, *OutFront* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Shoulder | *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Shoulder* | *None* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Since | *InBack*, *UpBack*, *UpIn* | *InBack*, *UpBack*, *UpIn* | *Shoulder* | *Shoulder* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Spend | *DownFront*, *DownIn*, *FrontIn*, *FrontUp*, *InBack* | *DownFront*, *DownIn*, *FrontIn*, *FrontUp*, *InBack* | *TorsoBottom* | *TorsoBottom* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Spread 3 | *FrontDown*, *FrontIn*, *InDown* | *FrontIn*, *FrontUp* | *NonDominant* | *Neutral* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Starbucks | *UpFront*, *UpIn* | *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Suspenders 1 | *FrontDown*, *UpFront*, *UpIn* | *FrontDown*, *UpFront*, *UpIn* | *TorsoTop* | *TorsoTop* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Sweater | *InDown*, *InFront*, *UpIn* | *InDown*, *InFront*, *UpIn* | *TorsoTop* | *TorsoTop* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Thin | *FrontIn*, *FrontUp*, *InBack* | *FrontIn*, *FrontUp*, *InBack* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Transgender | *InDown*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *TorsoTop* | *None* | *Curved* | *None* | ✓ | ✗ | ✓ | ✗ |
| Visit | *FrontIn*, *FrontUp*, *UpBack*, *UpIn* | *FrontIn*, *FrontUp*, *UpBack*, *UpIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Wave | *DownBack*, *FrontDown*, *FrontIn*, *InDown* | *DownBack*, *FrontDown*, *FrontIn*, *InDown* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Weird | *InDown*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Wheelchair | *FrontDown*, *FrontIn*, *InBack* | *FrontDown*, *FrontIn*, *InBack* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Whisper | *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | *None* | *None* | ✗ | ✗ | ✗ | ✗ |

Table A.15: List 1 of the recorded signs with the parameters examined in this work (Part 3)

| Sign | Orientation(End) | | Location(End) | | $FC_i$ | | $FC_m$ | | $FC_p$ | | $FC_r$ | | UR | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | d | nd | d | nd | d | nd | d | nd | d | nd | d | nd | d | nd |
| Acquire | *FrontDown*, *InDown*, *UpFront*, *UpIn* | *FrontDown*, *InDown*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | * | * | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Adopt | *FrontDown*, *FrontIn*, *UpFront*, *UpIn* | *FrontDown*, *FrontIn*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ |
| Airplane | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Neutral* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Badge | *InDown*, *InFront*, *UpFront*, *UpIn* | *InDown*, *InFront*, *UpFront*, *UpIn* | *Shoulder* | *Shoulder* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Brag | *DownOut*, *FrontDown*, *InDown* | *DownOut*, *FrontDown*, *InDown* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Bury 3 | *DownIn*, *FrontIn* | *DownIn*, *FrontIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Carry | *FrontIn*, *FrontUp*, *InBack*, *InUp* | *FrontIn*, *FrontUp*, *InBack*, *InUp* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Cheese grater | *DownBack*, *DownIn*, *FrontDown*, *FrontIn*, *InBack* | *DownOut*, *InBack*, *InDown* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Come | *InDown*, *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Neutral* | *None* | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✗ | ✗ |
| Correct | *FrontIn*, *InBack*, *InDown* | *FrontIn*, *FrontUp*, *InBack*, *OutUp* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Credit card | *FrontIn*, *FrontUp*, *InBack*, *UpBack*, *UpIn* | *FrontIn*, *FrontUp*, *InUp* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Cry 3 | *DownBack*, *FrontIn*, *InBack*, *InDown* | *DownBack*, *FrontIn*, *InBack*, *InDown* | *Head* | *Head* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Dictionary | *FrontDown*, *FrontIn*, *UpIn* | *FrontIn*, *FrontUp*, *InUp* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Dig up | *FrontIn*, *InBack*, *InDown*, *OutFront*, *OutUp*, *UpFront*, *UpIn* | *FrontIn*, *InBack*, *InDown*, *OutFront*, *OutUp*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Doctor 2 | *FrontDown*, *FrontIn*, *InDown*, *UpIn* | *FrontUp* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Drag | *DownBack*, *FrontDown*, *FrontIn*, *InDown*, *UpFront*, *UpIn* | *DownBack*, *FrontDown*, *FrontIn*, *InDown*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ |
| Drawer | *FrontIn*, *FrontUp* | *FrontIn*, *FrontUp* | *Neutral* | *Neutral* | * | * | * | * | * | * | * | * | ✗ | ✗ |
| Each other | *FrontDown*, *FrontIn*, *InBack*, *InDown*, *InFront*, *UpFront*, *UpIn* | *FrontDown*, *FrontIn*, *InBack*, *InDown*, *InFront*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Edit 1 | *DownIn*, *FrontDown*, *FrontIn* | *DownIn*, *FrontDown*, *FrontIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Evaluate | *FrontDown*, *InDown*, *UpFront*, *UpIn* | *FrontDown*, *InDown*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Experiment | *FrontDown*, *UpFront*, *UpIn* | *FrontDown*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Explanation | *FrontDown*, *FrontIn*, *InDown*, *UpFront*, *UpIn* | *FrontDown*, *FrontIn*, *InDown*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Fade | *FrontIn*, *FrontUp*, *OutFront*, *OutUp* | *FrontIn*, *FrontUp*, *OutFront*, *OutUp* | *Neutral* | *Neutral* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ |
| Frustrate | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Get off | *DownBack*, *FrontDown* | *FrontIn*, *InBack* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Gloves 4 | *FrontDown*, *InDown* | *DownBack*, *FrontDown*, *FrontIn* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Grieve | *DownIn*, *InBack*, *InDown*, *InUp*, *UpIn* | *DownIn*, *InBack*, *InDown*, *InUp*, *UpIn* | *Neutral* | *Neutral* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Hairdresser | *InDown*, *OutUp*, *UpFront*, *UpIn* | *InDown*, *OutUp*, *UpFront*, *UpIn* | *Head* | *Head* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Handsome 3 | *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✗ | ✗ |
| Hunt | *FrontIn*, *InBack*, *OutFront*, *UpFront*, *UpIn* | *FrontIn*, *InBack*, *OutFront*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Hurdle 1 | *FrontDown*, *FrontIn*, *UpFront* | *DownOut*, *InBack*, *InDown*, *UpIn* | *NonDominant* | *Neutral* | ✓ | ✗ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Inspiring | *FrontUp*, *UpBack*, *UpOut* | *DownBack*, *DownFront*, *DownIn* | *TorsoTop* | *None* | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✗ | ✗ |
| Japan | *FrontIn*, *FrontUp*, *OutFront*, *UpBack* | *FrontIn*, *FrontUp*, *OutFront*, *UpBack* | *Neutral* | *Neutral* | * | * | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Lick envelope 2 | *FrontIn*, *InBack*, *InDown*, *UpIn* | *FrontIn*, *InBack*, *InDown*, *UpIn* | *Head* | *Head* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Look at | *FrontDown*, *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Match | *FrontIn*, *FrontUp*, *InBack*, *UpBack* | *FrontIn*, *FrontUp*, *InBack*, *UpBack* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Measure 2 | *InBack*, *InDown*, *InFront*, *UpFront*, *UpIn* | *InBack*, *InDown*, *InFront*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Meditate 3 | *InBack*, *UpBack*, *UpIn* | *InBack*, *UpBack*, *UpIn* | *Head* | *Head* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ |

Continued on next page

Continued from previous page

| Sign | Orientation(End) | | Location(End) | | $FC_i$ | | $FC_m$ | | $FC_p$ | | $FC_r$ | | UR | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | d | nd | d | nd | d | nd | d | nd | d | nd | d | nd | d | nd |
| Mischievous | *InDown*, *UpFront*, *UpIn* | *InDown*, *UpFront*, *UpIn* | *Head* | *Head* | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Mix 3 | *DownBack*, *DownOut*, *FrontDown*, *InBack*, *InDown* | *DownBack*, *DownOut*, *FrontDown*, *InBack*, *InDown* | *Neutral* | *Neutral* | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Monkey | *DownIn*, *InBack*, *InUp* | *DownIn*, *InBack*, *InUp* | *TorsoBottom* | *TorsoBottom* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ |
| Museum | *FrontDown*, *OutDown*, *UpFront* | *FrontDown*, *OutDown*, *UpFront* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Next to | *FrontIn*, *OutFront* | *InBack* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Niece | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Octopus | *DownBack*, *FrontDown*, *FrontIn*, *InDown*, *UpFront* | *DownOut*, *FrontDown*, *InBack*, *InDown* | *NonDominant* | *Neutral* | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✗ | ✗ |
| Paper | *FrontIn*, *InBack*, *InDown*, *UpFront*, *UpIn* | *DownFront*, *FrontIn*, *FrontOut*, *FrontUp*, *InBack*, *InUp* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Problem 1 | *DownOut*, *FrontIn*, *InBack*, *InDown* | *DownOut*, *FrontIn*, *InBack*, *InDown* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✓ | ✓ |
| Ready | *FrontDown*, *UpFront*, *UpIn* | *FrontDown*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Scarcely | *UpBack*, *UpIn* | *UpBack*, *UpIn* | *Head* | *Head* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Scout | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Sentence | *FrontDown*, *FrontIn*, *UpFront* | *FrontDown*, *FrontIn*, *UpFront* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Shark 1 | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Shooting | *UpBack*, *UpFront*, *UpIn* | *UpBack*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Shoulder | *InDown*, *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Shoulder* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Since | *FrontIn*, *FrontUp*, *InBack*, *UpBack* | *FrontIn*, *FrontUp*, *InBack*, *UpBack* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Spend | *FrontUp*, *UpBack*, *UpIn* | *FrontUp*, *UpBack*, *UpIn* | *Neutral* | *Neutral* | * | * | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Spread 3 | *FrontDown*, *FrontIn*, *InBack*, *InDown* | *FrontIn*, *FrontUp* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Starbucks | *InDown*, *UpFront*, *UpIn* | *InDown*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✓ | ✓ | ✗ | ✗ | ✓ | ✓ | ✗ | ✗ |
| Suspenders 1 | *DownBack*, *DownIn*, *InBack*, *InDown* | *DownBack*, *DownIn*, *InBack*, *InDown* | *TorsoBottom* | *TorsoBottom* | * | * | * | * | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Sweater | *DownIn*, *FrontIn*, *InBack*, *InUp* | *DownIn*, *FrontIn*, *InBack*, *InUp* | *TorsoBottom* | *TorsoBottom* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✓ | ✓ |
| Thin | *DownFront*, *FrontIn*, *FrontUp*, *UpBack*, *UpIn* | *DownFront*, *FrontIn*, *FrontUp*, *UpBack*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Transgender | *FrontUp*, *InUp*, *OutUp*, *UpBack*, *UpIn*, *UpOut* | *DownBack*, *DownFront*, *DownIn* | *TorsoTop* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✓ | ✗ |
| Visit | *FrontIn*, *FrontUp*, *InBack*, *UpBack*, *UpIn* | *FrontIn*, *FrontUp*, *InBack*, *UpBack*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Wave | *DownBack*, *DownIn*, *FrontDown*, *InDown*, *OutDown* | *DownBack*, *DownIn*, *FrontDown*, *InDown*, *OutDown* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Weird | *InDown*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | ✓ | ✗ | ✓ | ✗ | ✗ | ✗ | ✓ | ✗ | ✗ | ✗ |
| Wheelchair | *DownBack*, *FrontIn*, *InBack* | *DownBack*, *FrontIn*, *InBack* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | * | * |
| Whisper | *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✗ | ✗ |

Table A.16: List 2 of the recorded signs with the parameters examined in this work (Part 1)

| Sign | BST | Complex. | Handshape(End) | | Handshape(Start) | |
|---|---|---|---|---|---|---|
| | | | d | nd | d | nd |
| Accept | Type 1 | 1 | *4*, *Curved 5*, *Stacked 5*, *O* | *4*, *Curved 5*, *Stacked 5*, *O* | *4*, *5*, *Curved 5*, *Spread Flat 5*, *Stacked 5* | *4*, *5*, *Curved 5*, *Spread Flat 5*, *Stacked 5* |
| Admire | Type 1 | 3 | *V*, *Curved V*, *Flat V* | *V*, *Curved V*, *Flat V* | *Flat H*, *Flat N*, *Curved V*, *Flat V* | *Flat H*, *Flat N*, *Curved V*, *Flat V* |
| All day | Type 2 | 2 | *Open B* | *Open B* | *5*, *Open B* | *5*, *Open B* |
| All together | Type 1 | 2 | *Curved 5*, *Stacked 5*, *Beak*, *C*, *O* | *Curved 5*, *Stacked 5*, *Beak*, *C*, *O* | *4*, *Curved 5*, *Spread Flat 5*, *Stacked 5* | *4*, *Curved 5*, *Spread Flat 5*, *Stacked 5* |
| All way | Type 0 | 2 | *4*, *5*, *Open B* | * | *4*, *5*, *Open B* | * |
| Angel | Type 1 | 1 | *Flat B*, *Open B* | *Flat B*, *Open B* | *Flat B*, *Open B* | *Flat B*, *Open B* |
| Bag 2 | Type 3 | 2 | *4*, *Flat 4*, *Flat B* | *1*, *4*, *Flat 4*, *5*, *A*, *Closed B*, *Flat B*, *C*, *O*, *S* | *4*, *Flat 4*, *Flat B* | *1*, *4*, *Flat 4*, *5*, *A*, *Closed B*, *Flat B*, *C*, *O*, *S* |
| Band aid 2 | Type 3 | 4 | *Curved H*, *U*, *Curved V* | *Closed B* | *Curved H*, *U*, *Curved V* | *Closed B* |
| Barrette | Type 0 | 2 | *Flat H*, *Flat N* | * | *Flat H*, *Flat N* | * |
| Bib | Type 1 | 2 | *1*, *Curved 1*, *Flat 1* | *1*, *Curved 1*, *Flat 1* | *1*, *Curved 1*, *Flat 1* | *1*, *Curved 1*, *Flat 1* |
| Boiling mad | Type 3 | 2 | *Spread Flat 5* | *Closed B* | *Spread Flat 5* | *Closed B* |
| Bottle | Type 3 | 2 | *A*, *C*, *O*, *S* | *Closed B*, *Open B* | *5*, *Curved 5*, *C*, *O*, *S* | *Closed B*, *Open B* |
| Bowling 2 | Type 0 | 2 | *Curved 5*, *C* | * | *Curved 5*, *C* | * |
| Bracelet | Type 3 | 1 | *Open 8*, *F*, *Open F*, *Q* | *Closed B* | *Open 8*, *F*, *Open F*, *Q* | *Closed B* |
| Building | Type 1 | 2 | *Flat B*, *Open B*, *Flat H*, *Open H*, *Flat M*, *Flat N*, *U* | *Flat B*, *Open B*, *Flat H*, *Open H*, *Flat M*, *Flat N*, *U* | *Flat B*, *Open B*, *Flat H*, *Open H*, *Flat M*, *Flat N*, *U* | *Flat B*, *Open B*, *Flat H*, *Open H*, *Flat M*, *Flat N*, *U* |
| Camping | Type 1 | 3 | *A*, *Horns*, *Flat Horns*, *T*, *Y* | *A*, *Horns*, *Flat Horns*, *T*, *Y* | *Horns*, *Ily* | *Horns*, *Ily* |
| Cloud 1 | Type 1 | 2 | *Curved 5*, *Spread Open E* | *Curved 5*, *Spread Open E* | *Curved 5*, *Spread Open E* | *Curved 5*, *Spread Open E* |
| Cowboy | Type 1 | 3 | *3* | *3* | *3* | *3* |
| Crash plane | Type 3 | 3 | *Ily* | *5*, *Open B* | *Ily* | *5*, *Open B* |
| Diaper | Type 1 | 2 | *Curved H*, *Flat N*, *U*, *Flat V* | *Curved H*, *Flat N*, *U*, *Flat V* | *3*, *Flat N*, *U*, *V*, *Curved V* | *3*, *Flat N*, *U*, *V*, *Curved V* |
| Dress up | Type 1 | 2 | *4*, *5*, *Open B* | *4*, *5*, *Open B* | *4*, *5*, *Open B* | *4*, *5*, *Open B* |
| Duty | Type 3 | 3 | *D*, *Horns* | *4*, *Closed B* | *D*, *Horns* | *4*, *Closed B* |
| Easter | Type 1 | 1 | *E*, *Closed E*, *Spread E* | *E*, *Closed E*, *Spread E* | *E*, *Closed E*, *Spread E* | *E*, *Closed E*, *Spread E* |
| Elementary school | Type 3 | 3 | *E*, *Closed E* | *Closed B* | *E*, *Closed E* | *Closed B* |
| Erupt 2 | Type 1 | 2 | *5*, *Spread Flat 5* | *5*, *Spread Flat 5* | *Open E* | *Open E* |
| Express | Type 1 | 2 | *5*, *Spread Flat 5* | *5*, *Spread Flat 5* | *A*, *O*, *S* | *A*, *O*, *S* |
| Feedback | Type 1 | 1 | *5*, *Spread Flat 5*, *F* | *5*, *Spread Flat 5*, *F* | *F* | *F* |
| Frown 2 | Type 1 | 2 | *Curved 1*, *Flat 1* | *Curved 1*, *Flat 1* | *Curved 1*, *Flat 1* | *Curved 1*, *Flat 1* |
| Glue 2 | Type 1 | 2 | *8*, *F* | *8*, *F* | *8* | *8* |
| Goody goody shoe opposite | Type 0 | 3 | *Goody Goody* | * | *Goody Goody* | * |
| Goosebumps | Type 3 | 1 | *Curved 4* | *1*, *Curved 4*, *5*, *A*, *Closed B*, *C*, *O*, *S* | *Curved 4* | *1*, *Curved 4*, *5*, *A*, *Closed B*, *C*, *O*, *S* |
| Half hour | Type 2 | 2 | *4*, *Flat 4*, *Closed B* | *4*, *Flat 4*, *Closed B* | *4*, *Closed B* | *4*, *Closed B* |
| Hang clothes | Type 1 | 1 | *F* | *F* | *F* | *F* |
| Hippo 2 | Type 1 | 2 | *Horns*, *Flat Horns* | *Horns*, *Flat Horns* | *Horns*, *Flat Horns* | *Horns*, *Flat Horns* |
| Hurricane | Type 1 | 2 | *Flat H*, *Open H* | *Flat H*, *Open H* | *Open H* | *Open H* |
| Incredible | Type 1 | 2 | *L* | *L* | *L* | *L* |
| Invisible | Type 1 | 2 | *A*, *S* | *A*, *S* | *8*, *Open 8* | *8*, *Open 8* |
| Jewelry | Type 1 | 2 | *Curved 5*, *Stacked 5*, *Spread E* | *Curved 5*, *Stacked 5*, *Spread E* | *4* | *4* |

Continued on next page

Continued from previous page

| Sign | BST | Complex. | Handshape(End) | | Handshape(Start) | |
|---|---|---|---|---|---|---|
| | | | d | nd | d | nd |
| Load | Type 1 | 1 | *Curved 4*, *Curved 5* | *Curved 4*, *Curved 5* | *Curved 4*, *Curved 5* | *Curved 4*, *Curved 5* |
| Maid | Type 0 | 2 | *Horns* | * | *Horns* | * |
| Mexican 2 | Type 0 | 2 | *Stacked 5*, *Flat M* | * | *Stacked 5*, *Flat M* | * |
| Mime | Type 1 | 2 | *4*, *Flat M* | *4*, *Flat M* | *4*, *Flat M* | *4*, *Flat M* |
| Mustache | Type 1 | 1 | *Curved 1*, *Baby O*, *Q* | *Curved 1*, *Baby O*, *Q* | *Curved 1*, *Curved L*, *Q* | *Curved 1*, *Curved L*, *Q* |
| Navy | Type 1 | 1 | *Beak* | *Beak* | *Stacked 5*, *Beak* | *Stacked 5*, *Beak* |
| Necklace 2 | Type 1 | 1 | *Curved 1*, *Flat 1* | *Curved 1*, *Flat 1* | *Curved 1*, *Flat 1* | *Curved 1*, *Flat 1* |
| Network | Type 1 | 3 | *Open 8* | *Open 8* | *Open 8* | *Open 8* |
| Neutral | Type 0 | 2 | *Flat H*, *Flat N*, *Flat V* | * | *Flat H*, *Flat N*, *Flat V* | * |
| Pants | Type 1 | 3 | *5*, *Spread Flat 5*, *Closed B*, *Open B* | *5*, *Spread Flat 5*, *Closed B*, *Open B* | *5*, *Spread Flat 5*, *Closed B*, *Open B* | *5*, *Spread Flat 5*, *Closed B*, *Open B* |
| Penguin | Type 1 | 1 | *5*, *Spread Flat 5*, *Closed B*, *Open B* | *5*, *Spread Flat 5*, *Closed B*, *Open B* | *5*, *Spread Flat 5*, *Closed B*, *Open B* | *5*, *Spread Flat 5*, *Closed B*, *Open B* |
| Place | Type 1 | 4 | *K* | *K* | *K* | *K* |
| Play | Type 1 | 2 | *Y* | *Y* | *Y* | *Y* |
| Porcupine | Type 3 | 3 | *Curved 4*, *Curved 5*, *Spread Open E*, *Spread E* | *Curved 5*, *C* | *Curved 4*, *Curved 5*, *Spread Open E*, *Spread E* | *Curved 5*, *C* |
| Put aside | Type 1 | 1 | *5*, *Open B* | *5*, *Open B* | *5*, *Open B* | *5*, *Open B* |
| React | Type 1 | 4 | *R*, *U* | *R*, *U* | *R*, *U* | *R*, *U* |
| Record | Type 1 | 3 | *Open 8* | *Open 8* | *Open 8* | *Open 8* |
| Reduce | Type 2 | 1 | *Flat B* | *Flat B* | *Flat B* | *Flat B* |
| Right 1 | Type 1 | 1 | *1*, *D* | *1*, *D* | *1*, *D* | *1*, *D* |
| Rocking chair 2 | Type 1 | 2 | *V*, *Curved V*, *Flat V* | *V*, *Curved V*, *Flat V* | *3*, *V*, *Curved V*, *Flat V* | *3*, *V*, *Curved V*, *Flat V* |
| Rollerskating | Type 1 | 3 | *Curved H*, *Curved V* | *Curved H*, *Curved V* | *Curved H*, *Curved V* | *Curved H*, *Curved V* |
| Sauce | Type 3 | 4 | *A* | *Closed B* | *A* | *Closed B* |
| Schedule | Type 3 | 4 | *4*, *5*, *Spread Flat 5* | *5*, *Open B* | *4*, *5*, *Spread Flat 5* | *5*, *Open B* |
| Seven | Type 0 | 2 | *7* | * | *7* | * |
| Sharpen 1 | Type 2 | 1 | *A*, *Spread Open E*, *O*, *Baby O*, *S* | *A*, *Spread Open E*, *O*, *Baby O*, *S* | *A*, *Spread Open E*, *O*, *Baby O*, *S* | *A*, *Spread Open E*, *O*, *Baby O*, *S* |
| Sled 2 | Type 1 | 2 | *Curved 1* | *Curved 1* | *Curved 1* | *Curved 1* |
| Smooth | Type 1 | 2 | *4*, *Stacked 5*, *O* | *4*, *Stacked 5*, *O* | *4*, *Stacked 5* | *4*, *Stacked 5* |
| Soup | Type 3 | 4 | *Curved H*, *Flat N*, *U*, *Curved V* | *Closed B*, *Open B* | *Curved H*, *Flat N*, *U*, *Curved V* | *Closed B*, *Open B* |
| Spaghetti | Type 1 | 3 | *I* | *I* | *I* | *I* |
| Sweep | Type 2 | 2 | *5*, *Closed B*, *Flat B*, *Open B* | *5*, *Closed B*, *Flat B*, *Open B* | *5*, *Closed B*, *Flat B*, *Open B* | *5*, *Closed B*, *Flat B*, *Open B* |
| Truth | Type 3 | 3 | *U* | *Closed B* | *U* | *Closed B* |
| Underwear | Type 1 | 1 | *L*, *Curved L*, *Baby O*, *Q* | *L*, *Curved L*, *Baby O*, *Q* | *L*, *Curved L* | *L*, *Curved L* |
| Vain | Type 1 | 3 | *Flat N*, *V*, *Flat V* | *Flat N*, *V*, *Flat V* | *V*, *Flat V* | *V*, *Flat V* |
| Vase | Type 1 | 3 | *L*, *Curved L* | *L*, *Curved L* | *L*, *Curved L* | *L*, *Curved L* |
| Videogame | Type 1 | 2 | *A*, *T* | *A*, *T* | *A* | *A* |
| Voice | Type 0 | 3 | *V*, *Flat V* | * | *Flat H*, *V*, *Curved V*, *Flat V* | * |
| Win | Type 2 | 2 | *A*, *O*, *S* | *A*, *O*, *S* | *5*, *Curved 5*, *O*, *S* | *5*, *Curved 5*, *O*, *S* |
| World | Type 1 | 3 | *W* | *W* | *W* | *W* |
| Zebra | Type 1 | 1 | *4*, *Flat 4* | *4*, *Flat 4* | *4*, *Flat 4* | *4*, *Flat 4* |

Table A.17: List 2 of the recorded signs with the parameters examined in this work (Part 2)

| Sign | Orientation(Start) | | Location(Start) | | Movement | | $SC_{im}$ | | $SC_{rp}$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | d | nd | d | nd | d | nd | d | nd | d | nd |
| Accept | *FrontDown*, *FrontIn*, *InDown*, *UpIn* | *FrontDown*, *FrontIn*, *InDown*, *UpIn* | *TorsoTop* | *TorsoTop* | *Straight* | *Straight* | ✓ | ✓ | ✓ | ✓ |
| Admire | *DownBack*, *FrontDown*, *FrontIn*, *FrontUp*, *InBack*, *InDown*, *UpIn* | *DownBack*, *FrontDown*, *FrontIn*, *FrontUp*, *InBack*, *InDown*, *UpIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| All day | *FrontIn*, *OutBack*, *OutUp*, *UpIn* | *InDown* | *Neutral* | *Neutral* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| All together | *FrontDown*, *FrontUp*, *InDown*, *InFront*, *InUp*, *UpIn* | *FrontDown*, *FrontUp*, *InDown*, *InFront*, *InUp*, *UpIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✓ | ✓ | ✓ | ✓ |
| All way | *BackIn*, *BackUp*, *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Neutral* | *None* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Angel | *InDown*, *UpBack*, *UpIn* | *InDown*, *UpBack*, *UpIn* | *Shoulder* | *Shoulder* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Bag 2 | *UpIn* | *InDown* | *NonDominant* | *Neutral* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Band aid 2 | *DownOut*, *InBack*, *InDown* | *DownOut*, *FrontDown*, *InBack*, *InDown* | *NonDominant* | *Neutral* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Barrette | *InUp*, *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | *None* | *None* | ✗ | ✗ | ✗ | ✗ |
| Bib | *InBack*, *InDown*, *UpBack* | *InBack*, *InDown*, *UpBack* | *TorsoTop* | *TorsoTop* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Boiling mad | *FrontUp*, *InBack*, *InUp*, *UpOut* | *InBack*, *InDown* | *TorsoBottom* | *Neutral* | *None* | *None* | ✗ | ✗ | ✗ | ✗ |
| Bottle | *FrontIn*, *InBack* | *FrontUp*, *InBack*, *InUp* | *NonDominant* | *Neutral* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Bowling 2 | *BackDown*, *DownFront*, *DownIn*, *OutDown* | *DownBack*, *DownFront*, *DownIn* | *Neutral* | *None* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Bracelet | *FrontDown*, *InDown* | *DownBack*, *DownOut*, *FrontDown*, *InDown* | *NonDominant* | *Neutral* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Building | *InDown*, *UpIn* | *InDown*, *UpIn* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Camping | *InDown*, *UpIn* | *InDown*, *UpIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Cloud 1 | *InFront*, *UpFront* | *InFront*, *UpFront* | *Head* | *Head* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Cowboy | *FrontIn*, *FrontUp*, *InUp* | *FrontIn*, *FrontUp*, *InUp* | *TorsoBottom* | *TorsoBottom* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Crash plane | *FrontDown*, *InDown*, *UpFront*, *UpIn* | *DownBack*, *DownOut*, *FrontDown*, *InDown* | *NonDominant* | *Neutral* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Diaper | *DownBack*, *DownOut*, *FrontDown*, *InBack*, *InDown* | *DownBack*, *DownOut*, *FrontDown*, *InBack*, *InDown* | *TorsoBottom* | *TorsoBottom* | *None* | *None* | ✗ | ✗ | ✗ | ✗ |
| Dress up | *FrontDown*, *FrontIn*, *InBack*, *InDown*, *UpIn* | *FrontDown*, *FrontIn*, *InBack*, *InDown*, *UpIn* | *TorsoTop* | *TorsoTop* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Duty | *FrontDown*, *FrontIn*, *InBack*, *InDown*, *UpFront*, *UpIn* | *DownBack*, *FrontDown*, *InBack*, *InDown* | *NonDominant* | *Neutral* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Easter | *InDown*, *UpFront*, *UpIn* | *InDown*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | *None* | *None* | ✗ | ✗ | ✗ | ✗ |
| Elementary school | *FrontDown*, *UpFront* | *InDown* | *NonDominant* | *Neutral* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Erupt 2 | *FrontDown*, *FrontIn*, *InDown* | *FrontDown*, *FrontIn*, *InDown* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✓ | ✓ | ✓ | ✓ |
| Express | *InBack*, *InUp*, *UpBack* | *InBack*, *InUp*, *UpBack* | *TorsoTop* | *TorsoTop* | *Curved* | *Curved* | ✓ | ✓ | ✓ | ✓ |
| Feedback | *UpFront*, *UpIn* | *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Frown 2 | *InBack*, *InDown*, *UpIn* | *InBack*, *InDown*, *UpIn* | *Head* | *Head* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Glue 2 | *FrontIn*, *FrontUp* | *FrontIn*, *FrontUp* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Goody goody shoe opposite | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Neutral* | *None* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Goosebumps | *FrontDown*, *FrontIn*, *InDown*, *UpFront*, *UpIn* | *DownOut*, *FrontDown*, *InDown* | *NonDominant* | *Neutral* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Half hour | *UpFront*, *UpIn* | *InDown*, *UpIn* | *NonDominant* | *Neutral* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Hang clothes | *DownBack*, *DownOut*, *FrontDown*, *InDown* | *DownBack*, *DownOut*, *FrontDown*, *InDown* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Hippo 2 | *FrontDown*, *FrontIn*, *FrontUp*, *InBack*, *InDown*, *UpBack*, *UpFront*, *UpIn* | *FrontDown*, *FrontIn*, *FrontUp*, *InBack*, *InDown*, *UpBack*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Hurricane | *InBack*, *InFront*, *UpBack*, *UpFront*, *UpIn* | *InBack*, *InFront*, *UpBack*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Incredible | *FrontIn*, *OutUp*, *UpFront*, *UpIn* | *FrontIn*, *OutUp*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Invisible | *FrontDown*, *FrontIn*, *InBack*, *InDown*, *InUp*, *UpFront*, *UpIn* | *FrontDown*, *FrontIn*, *InBack*, *InDown*, *InUp*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Jewelry | *InBack*, *InDown* | *InBack*, *InDown* | *TorsoTop* | *TorsoTop* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |

Continued on next page

Continued from previous page

| Sign | Orientation(Start) | | Location(Start) | | Movement | | $SC_{im}$ | | $SC_{rp}$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | d | nd | d | nd | d | nd | d | nd | d | nd |
| Load | *FrontUp*, *InBack*, *InUp*, *OutUp* | *FrontUp*, *InBack*, *InUp*, *OutUp* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Maid | *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Mexican 2 | *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Mime | *FrontDown*, *FrontIn*, *InDown*, *UpIn* | *FrontDown*, *FrontIn*, *InDown*, *UpIn* | *TorsoTop* | *TorsoTop* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Mustache | *InDown*, *UpFront*, *UpIn* | *InDown*, *UpFront*, *UpIn* | *Head* | *Head* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Navy | *FrontDown*, *FrontIn*, *InBack*, *InDown* | *FrontDown*, *FrontIn*, *InBack*, *InDown* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Necklace 2 | *InBack*, *InDown*, *UpBack*, *UpIn* | *InBack*, *InDown*, *UpBack*, *UpIn* | *Shoulder* | *Shoulder* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Network | *UpFront*, *UpIn* | *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Neutral | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Neutral* | *None* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Pants | *FrontDown*, *FrontIn*, *InBack*, *InDown* | *FrontDown*, *FrontIn*, *InBack*, *InDown* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Penguin | *DownBack*, *DownIn*, *FrontDown*, *FrontIn*, *OutDown* | *DownBack*, *DownIn*, *FrontDown*, *FrontIn*, *OutDown* | *TorsoBottom* | *TorsoBottom* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Place | *FrontIn*, *FrontUp*, *InBack*, *UpIn* | *FrontIn*, *FrontUp*, *InBack*, *UpIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Play | *FrontDown*, *InDown*, *OutFront*, *UpIn* | *FrontDown*, *InDown*, *OutFront*, *UpIn* | *Neutral* | *Neutral* | *None* | *None* | ✗ | ✗ | ✗ | ✗ |
| Porcupine | *FrontIn*, *FrontUp*, *OutFront*, *UpFront* | *DownOut*, *FrontOut*, *InDown* | *NonDominant* | *Neutral* | *None* | *None* | ✗ | ✗ | ✗ | ✗ |
| Put aside | *FrontDown*, *FrontIn*, *FrontOut* | *FrontDown*, *FrontIn*, *FrontOut* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| React | *UpFront*, *UpIn* | *UpFront*, *UpIn* | *Head* | *Head* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Record | *FrontIn*, *FrontUp*, *InBack*, *InDown*, *InUp*, *UpBack* | *FrontIn*, *FrontUp*, *InBack*, *InDown*, *InUp*, *UpBack* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Reduce | *UpFront*, *UpIn* | *FrontUp*, *InUp* | *NonDominant* | *Neutral* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Right 1 | *FrontIn*, *InBack*, *UpBack*, *UpIn* | *FrontIn*, *InBack*, *UpBack*, *UpIn* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Rocking chair 2 | *FrontIn*, *UpFront*, *UpIn* | *FrontIn*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Rollerskating | *FrontIn*, *FrontUp*, *InBack* | *FrontIn*, *FrontUp*, *InBack* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Sauce | *FrontDown*, *InDown*, *InFront*, *UpFront*, *UpIn* | *FrontIn*, *FrontUp*, *InBack*, *InUp* | *NonDominant* | *Neutral* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Schedule | *InDown*, *UpIn* | *FrontIn* | *NonDominant* | *Neutral* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Seven | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Neutral* | *None* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Sharpen 1 | *FrontDown*, *FrontIn* | *FrontDown*, *FrontIn*, *InBack* | *NonDominant* | *Neutral* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Sled 2 | *BackDown*, *InBack*, *UpBack* | *BackDown*, *InBack*, *UpBack* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Smooth | *FrontIn*, *FrontUp*, *InBack*, *UpBack* | *FrontIn*, *FrontUp*, *InBack*, *UpBack* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Soup | *InBack*, *InUp*, *UpIn* | *FrontUp*, *InUp* | *NonDominant* | *Neutral* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Spaghetti | *FrontIn*, *InUp*, *UpBack*, *UpIn* | *FrontIn*, *InUp*, *UpBack*, *UpIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Sweep | *FrontIn*, *InBack*, *UpBack*, *UpIn* | *FrontIn*, *FrontUp* | *NonDominant* | *Neutral* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Truth | *FrontIn*, *FrontUp*, *InBack* | *FrontIn*, *FrontUp* | *NonDominant* | *Neutral* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Underwear | *DownBack*, *InBack* | *DownBack*, *InBack* | *TorsoBottom* | *TorsoBottom* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Vain | *FrontIn*, *FrontUp*, *UpBack*, *UpIn* | *FrontIn*, *FrontUp*, *UpBack*, *UpIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Vase | *FrontIn*, *InBack* | *FrontIn*, *InBack* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Videogame | *FrontDown*, *FrontIn* | *FrontDown*, *FrontIn* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Voice | *UpBack* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Win | *FrontDown*, *FrontIn*, *UpFront*, *UpIn* | *FrontDown*, *FrontIn*, *InBack* | *NonDominant* | *Neutral* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| World | *InBack*, *InDown*, *UpIn* | *InBack*, *InDown*, *UpIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Zebra | *InBack*, *InDown* | *InBack*, *InDown* | *TorsoTop* | *TorsoTop* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |

Table A.18: List 2 of the recorded signs with the parameters examined in this work (Part 3)

| Sign | Orientation(End) | | Location(End) | | $FC_i$ | | $FC_m$ | | $FC_p$ | | $FC_r$ | | UR | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | d | nd | d | nd | d | nd | d | nd | d | nd | d | nd | d | nd |
| Accept | *InBack*, *InDown*, *UpBack*, *UpIn* | *InBack*, *InDown*, *UpBack*, *UpIn* | *TorsoTop* | *TorsoTop* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ |
| Admire | *FrontDown*, *InDown*, *UpFront*, *UpIn* | *FrontDown*, *InDown*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | * | * | * | * | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| All day | *InDown* | *InDown* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| All together | *FrontDown*, *FrontIn*, *InBack*, *InDown*, *InUp*, *UpFront*, *UpIn* | *FrontDown*, *FrontIn*, *InBack*, *InDown*, *InUp*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| All way | *DownBack*, *FrontDown*, *FrontIn* | *DownBack*, *DownFront*, *DownIn* | *Neutral* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Angel | *FrontDown*, *UpFront* | *FrontDown*, *UpFront* | *Neutral* | *Neutral* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Bag 2 | *FrontUp*, *InUp* | *InDown* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✓ | ✗ |
| Band aid 2 | *FrontDown*, *InDown*, *UpFront*, *UpIn* | *DownOut*, *FrontDown*, *InBack*, *InDown* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Barrette | *InUp*, *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | ✓ | ✗ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Bib | *BackDown*, *BackOut*, *InBack*, *InDown*, *UpBack* | *BackDown*, *BackOut*, *InBack*, *InDown*, *UpBack* | *Shoulder* | *Shoulder* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Boiling mad | *FrontUp*, *InBack*, *InUp*, *UpOut* | *InBack*, *InDown* | *TorsoBottom* | *Neutral* | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✗ | ✗ |
| Bottle | *FrontIn*, *InBack* | *FrontUp*, *InBack*, *InUp* | *NonDominant* | *Neutral* | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✗ | ✗ |
| Bowling 2 | *FrontUp*, *UpBack* | *DownBack*, *DownFront*, *DownIn* | *Neutral* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Bracelet | *InDown* | *DownBack*, *DownOut*, *FrontDown*, *InDown* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Building | *InDown*, *InFront*, *UpIn* | *InDown*, *InFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Camping | *FrontDown*, *InDown*, *UpIn* | *FrontDown*, *InDown*, *UpIn* | *Neutral* | *Neutral* | ✓ | ✓ | ✗ | ✗ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ |
| Cloud 1 | *FrontOut*, *UpFront* | *FrontOut*, *UpFront* | *Head* | *Head* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Cowboy | *FrontIn*, *FrontUp* | *FrontIn*, *FrontUp* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Crash plane | *FrontDown*, *FrontIn*, *InBack*, *InDown* | *DownBack*, *DownOut*, *FrontDown*, *InDown* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Diaper | *DownBack*, *DownIn*, *DownOut*, *FrontDown*, *InDown* | *DownBack*, *DownIn*, *DownOut*, *FrontDown*, *InDown* | *TorsoBottom* | *TorsoBottom* | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Dress up | *DownBack*, *FrontDown*, *FrontIn*, *UpFront*, *UpIn* | *DownBack*, *FrontDown*, *FrontIn*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Duty | *FrontIn*, *InBack*, *InDown* | *DownBack*, *FrontDown*, *InBack*, *InDown* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Easter | *InDown*, *UpFront*, *UpIn* | *InDown*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Elementary school | *FrontDown*, *UpFront* | *InDown* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Erupt 2 | *FrontIn*, *UpIn* | *FrontIn*, *UpIn* | *Neutral* | *Neutral* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | * | * |
| Express | *FrontUp*, *InBack*, *InUp*, *UpBack* | *FrontUp*, *InBack*, *InUp*, *UpBack* | *Neutral* | *Neutral* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Feedback | *FrontDown*, *InDown*, *UpBack*, *UpFront*, *UpIn* | *FrontDown*, *InDown*, *UpBack*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Frown 2 | *InDown*, *UpFront*, *UpIn* | *InDown*, *UpFront*, *UpIn* | *Head* | *Head* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Glue 2 | *FrontIn*, *FrontUp* | *FrontIn*, *FrontUp* | *Neutral* | *Neutral* | ✗ | ✗ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Goody goody shoe opposite | *DownBack*, *FrontDown*, *InBack* | *DownBack*, *DownFront*, *DownIn* | *Neutral* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Goosebumps | *UpIn* | *DownOut*, *FrontDown*, *InDown* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Half hour | *FrontDown* | *InDown*, *UpIn* | *NonDominant* | *Neutral* | * | ✗ | * | ✗ | * | ✗ | * | ✗ | ✗ | ✗ |
| Hang clothes | *FrontDown*, *FrontIn*, *InDown* | *FrontDown*, *FrontIn*, *InDown* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Hippo 2 | *FrontDown*, *FrontIn*, *FrontUp*, *InBack*, *InDown*, *UpBack*, *UpIn* | *FrontDown*, *FrontIn*, *FrontUp*, *InBack*, *InDown*, *UpBack*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Hurricane | *InBack*, *InFront*, *UpBack*, *UpFront*, *UpIn* | *InBack*, *InFront*, *UpBack*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Incredible | *FrontIn*, *FrontUp*, *OutFront*, *OutUp* | *FrontIn*, *FrontUp*, *OutFront*, *OutUp* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Invisible | *FrontIn*, *FrontUp*, *UpFront*, *UpIn* | *FrontIn*, *FrontUp*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Jewelry | *BackDown*, *InBack*, *InDown*, *UpBack* | *BackDown*, *InBack*, *InDown*, *UpBack* | *Shoulder* | *Shoulder* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ |

Continued on next page

Continued from previous page

| Sign | Orientation(End) | | Location(End) | | $FC_i$ | | $FC_m$ | | $FC_p$ | | $FC_r$ | | UR | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | d | nd | d | nd | d | nd | d | nd | d | nd | d | nd | d | nd |
| Load | *FrontIn*, *FrontUp*, *InBack*, *InUp*, *UpIn* | *FrontIn*, *FrontUp*, *InBack*, *InUp*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Maid | *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Mexican 2 | *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Mime | *FrontDown*, *InDown*, *UpFront*, *UpIn* | *FrontDown*, *InDown*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Mustache | *InDown*, *UpFront*, *UpIn* | *InDown*, *UpFront*, *UpIn* | *Head* | *Head* | * | * | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Navy | *FrontDown*, *FrontIn*, *InBack*, *InDown* | *FrontDown*, *FrontIn*, *InBack*, *InDown* | *Neutral* | *Neutral* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ |
| Necklace 2 | *InBack*, *UpBack*, *UpIn* | *InBack*, *UpBack*, *UpIn* | *TorsoTop* | *TorsoTop* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Network | *FrontIn*, *UpFront*, *UpIn* | *FrontIn*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Neutral | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Neutral* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Pants | *DownIn*, *FrontIn*, *InBack*, *InUp*, *OutDown* | *DownIn*, *FrontIn*, *InBack*, *InUp*, *OutDown* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Penguin | *DownBack*, *DownIn*, *FrontDown*, *FrontIn*, *OutDown* | *DownBack*, *DownIn*, *FrontDown*, *FrontIn*, *OutDown* | *TorsoBottom* | *TorsoBottom* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Place | *FrontIn*, *FrontUp*, *InBack* | *FrontIn*, *FrontUp*, *InBack* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Play | *FrontDown*, *InDown*, *OutFront*, *UpFront*, *UpIn* | *FrontDown*, *InDown*, *OutFront*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Porcupine | *UpFront*, *UpIn* | *DownOut*, *FrontOut*, *InDown* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Put aside | *FrontDown*, *FrontIn*, *FrontOut*, *InBack*, *OutBack* | *FrontDown*, *FrontIn*, *FrontOut*, *InBack*, *OutBack* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| React | *FrontDown*, *InDown* | *FrontDown*, *InDown* | *Head* | *Head* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Record | *FrontIn*, *FrontUp*, *InBack*, *InDown*, *InUp* | *FrontIn*, *FrontUp*, *InBack*, *InDown*, *InUp* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Reduce | *InUp*, *UpIn* | *FrontUp*, *InUp* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Right 1 | *FrontIn*, *InBack*, *UpBack*, *UpIn* | *FrontIn*, *InBack*, *UpBack*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Rocking chair 2 | *FrontIn* | *FrontIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Rollerskating | *FrontIn*, *FrontUp*, *UpBack* | *FrontIn*, *FrontUp*, *UpBack* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Sauce | *FrontDown*, *InDown*, *InFront*, *UpFront*, *UpIn* | *FrontIn*, *FrontUp*, *InBack*, *InUp* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Schedule | *FrontIn*, *InBack* | *FrontIn* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✓ | ✗ |
| Seven | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Neutral* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Sharpen 1 | *FrontDown*, *FrontIn* | *FrontDown*, *FrontIn*, *InBack* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Sled 2 | *DownFront*, *FrontIn*, *FrontUp*, *InBack*, *InUp* | *DownFront*, *FrontIn*, *FrontUp*, *InBack*, *InUp* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Smooth | *FrontIn*, *FrontUp*, *InBack* | *FrontIn*, *FrontUp*, *InBack* | *Neutral* | *Neutral* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ |
| Soup | *FrontIn*, *InBack*, *InUp*, *UpIn* | *FrontUp*, *InUp* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Spaghetti | *FrontIn*, *InUp*, *UpBack*, *UpIn* | *FrontIn*, *InUp*, *UpBack*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Sweep | *FrontIn*, *FrontUp*, *InBack*, *InUp*, *UpBack*, *UpIn* | *FrontIn*, *FrontUp* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Truth | *FrontIn*, *FrontUp*, *InBack* | *FrontIn*, *FrontUp* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Underwear | *DownBack*, *InBack* | *DownBack*, *InBack* | *Neutral* | *Neutral* | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Vain | *BackDown*, *InBack*, *UpBack* | *BackDown*, *InBack*, *UpBack* | *Neutral* | *Neutral* | * | * | * | * | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Vase | *FrontIn*, *InBack*, *UpBack*, *UpIn* | *FrontIn*, *InBack*, *UpBack*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | * | * |
| Videogame | *FrontDown*, *FrontIn* | *FrontDown*, *FrontIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Voice | *FrontUp*, *InBack*, *UpBack* | *DownBack*, *DownFront*, *DownIn* | *Neutral* | *None* | ✓ | ✗ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Win | *FrontIn*, *UpBack*, *UpIn* | *FrontDown*, *FrontIn*, *InBack* | *NonDominant* | *Neutral* | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ |
| World | *InBack*, *InDown*, *UpIn* | *InBack*, *InDown*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Zebra | *InBack*, *InDown*, *UpBack* | *InBack*, *InDown*, *UpBack* | *TorsoTop* | *TorsoTop* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |

Table A.19: List 3 of the recorded signs with the parameters examined in this work (Part 1)

| Sign | BST | Complex. | Handshape(End) | | Handshape(Start) | |
|---|---|---|---|---|---|---|
| | | | d | nd | d | nd |
| Baby 2 | Type 0 | 1 | *5*, *Open B* | * | *5*, *Open B* | * |
| Basic | Type 2 | 2 | *5*, *Open B* | *5*, *Open B* | *5*, *Open B* | *5*, *Open B* |
| Belt | Type 1 | 2 | *Curved L*, *U* | *Curved L*, *U* | *Curved L*, *U* | *Curved L*, *U* |
| Belt 2 | Type 1 | 3 | *Curved H*, *Flat N*, *Curved V*, *Flat V* | *Curved H*, *Flat N*, *Curved V*, *Flat V* | *U*, *V*, *Curved V* | *U*, *V*, *Curved V* |
| Bra | Type 1 | 1 | *C*, *L*, *Curved L* | *C*, *L*, *Curved L* | *C*, *L*, *Curved L* | *C*, *L*, *Curved L* |
| Bread | Type 3 | 2 | *Flat B*, *Open E*, *Spread Open E* | *Closed B*, *Flat B*, *Open B* | *Flat B*, *Open E*, *Spread Open E* | *Closed B*, *Flat B*, *Open B* |
| Brown | Type 0 | 1 | *4*, *Closed B*, *Open B* | * | *4*, *Closed B*, *Open B* | * |
| Climb ladder | Type 1 | 1 | *A*, *O*, *S* | *A*, *O*, *S* | *Curved 5*, *O* | *Curved 5*, *O* |
| Cooperate | Type 1 | 2 | *Open 8* | *Open 8* | *Open 8* | *Open 8* |
| Crawl 2 | Type 1 | 2 | *Curved 4*, *Curved 5* | *Curved 4*, *Curved 5* | *Curved 4*, *Curved 5* | *Curved 4*, *Curved 5* |
| Crewcut | Type 1 | 2 | *Curved 1*, *Flat 1*, *Q* | *Curved 1*, *Flat 1*, *Q* | *Curved 1*, *Flat 1*, *Q* | *Curved 1*, *Flat 1*, *Q* |
| Cross heart | Type 0 | 3 | *1*, *Curved 1*, *Flat 1* | * | *1*, *Curved 1*, *Flat 1* | * |
| Develop 2 | Type 3 | 3 | *D* | *Closed B* | *D* | *Closed B* |
| Die | Type 1 | 2 | *5*, *Open B* | *5*, *Open B* | *5*, *Open B* | *5*, *Open B* |
| Dig | Type 3 | 3 | *Curved 4*, *Flat 4*, *C* | *Closed B* | *Curved 4*, *Flat 4*, *C* | *Closed B* |
| Dinner 2 | Type 0 | 2 | *D* | * | *D* | * |
| Dissect | Type 3 | 3 | *A* | *Closed B* | *A* | *Closed B* |
| Dresser | Type 1 | 3 | *Curved 5*, *A*, *O* | *Curved 5*, *A*, *O* | *Curved 5*, *Spread Open E* | *Curved 5*, *Spread Open E* |
| Emotion | Type 1 | 2 | *Curved 4*, *E*, *Closed E*, *Open E*, *Spread E*, *O* | *Curved 4*, *E*, *Closed E*, *Open E*, *Spread E*, *O* | *Curved 4*, *E*, *Closed E*, *Open E*, *Spread E*, *O* | *Curved 4*, *E*, *Closed E*, *Open E*, *Spread E*, *O* |
| Feed 2 | Type 0 | 4 | *4*, *5*, *Spread Flat 5*, *Closed B* | * | *4*, *Stacked 5* | * |
| Frisk | Type 1 | 3 | *5* | *5* | *5* | *5* |
| Ghost | Type 1 | 1 | *F* | *F* | *F* | *F* |
| Glue 1 | Type 3 | 3 | *Flat H*, *Open H*, *T* | *4*, *Curved 4*, *5*, *Spread Flat 5*, *Closed B*, *Open B* | *Flat H*, *Flat V* | *4*, *Curved 4*, *5*, *Spread Flat 5*, *Closed B*, *Open B* |
| Goody goody | Type 0 | 2 | *Goody Goody* | * | *Goody Goody* | * |
| Haircut | Type 1 | 3 | *U*, *V*, *Curved V*, *Flat V* | *U*, *V*, *Curved V*, *Flat V* | *V*, *Curved V* | *V*, *Curved V* |
| Imagine | Type 1 | 3 | *I* | *I* | *I* | *I* |
| Inspire | Type 1 | 2 | *5*, *Spread Flat 5* | *5*, *Spread Flat 5* | *Beak* | *Beak* |
| Interview | Type 1 | 2 | *I* | *I* | *I* | *I* |
| Introduce | Type 1 | 2 | *Flat B*, *Open B* | *Flat B*, *Open B* | *Flat B*, *Open B* | *Flat B*, *Open B* |
| Intuitive | Type 0 | 1 | *Flat B* | * | *Flat B* | * |
| Knitting 3 | Type 1 | 3 | *Horns* | *Horns* | *Horns* | *Horns* |
| Lift 1 | Type 1 | 2 | *5*, *Open B* | *5*, *Open B* | *5*, *Open B* | *5*, *Open B* |
| Load 2 | Type 1 | 1 | *5*, *Spread Flat 5*, *Open B* | *5*, *Spread Flat 5*, *Open B* | *A*, *Closed E*, *O*, *S* | *A*, *Closed E*, *O*, *S* |
| Magnet | Type 1 | 2 | *F* | *F* | *8*, *Open 8* | *8*, *Open 8* |
| Missing | Type 3 | 2 | *A*, *Beak*, *O*, *S* | *C* | *4*, *Spread Flat 5*, *Stacked 5* | *C* |
| Model 1 | Type 0 | 2 | *Open 8* | * | *Open 8* | * |
| Monday | Type 0 | 3 | *Closed E*, *Flat M* | * | *Closed E*, *Flat M* | * |
| Nephew | Type 0 | 2 | *Flat N*, *Flat V* | * | *Flat N*, *Flat V* | * |

Continued on next page

Continued from previous page

| Sign | BST | Complex. | Handshape(End) | | Handshape(Start) | |
|---|---|---|---|---|---|---|
| | | | d | nd | d | nd |
| Paranoid | Type 1 | 3 | *K* | *K* | *K* | *K* |
| Perceive | Type 1 | 2 | *A*, *Curved H*, *Curved L*, *O*, *S* | *A*, *Curved H*, *Curved L*, *O*, *S* | *3*, *Open H* | *3*, *Open H* |
| Perspective | Type 3 | 4 | *Curved V*, *Flat V* | *1*, *D* | *Curved V*, *Flat V* | *1*, *D* |
| Rain | Type 1 | 1 | *Curved 4*, *5*, *Curved 5*, *Open E*, *Spread Open E* | *Curved 4*, *5*, *Curved 5*, *Open E*, *Spread Open E* | *Curved 4*, *5*, *Curved 5*, *Open E*, *Spread Open E* | *Curved 4*, *5*, *Curved 5*, *Open E*, *Spread Open E* |
| Respect | Type 1 | 4 | *R*, *U* | *R*, *U* | *R*, *U* | *R*, *U* |
| Retire | Type 1 | 3 | *R* | *R* | *R* | *R* |
| Rob | Type 1 | 3 | *3*, *Open H*, *U* | *3*, *Open H*, *U* | *3*, *Open H*, *U* | *3*, *Open H*, *U* |
| Scoop | Type 3 | 4 | *A*, *O*, *S* | *C* | *A*, *O*, *S* | *C* |
| Skate | Type 1 | 2 | *Curved 1* | *Curved 1* | *Curved 1* | *Curved 1* |
| Smile | Type 1 | 2 | *1*, *Curved 1*, *Flat 1*, *Flat B* | *1*, *Curved 1*, *Flat 1*, *Flat B* | *1*, *Curved 1*, *Flat 1*, *Flat B* | *1*, *Curved 1*, *Flat 1*, *Flat B* |
| Snap | Type 0 | 3 | *A*, *S* | * | *Curved H*, *Curved V* | * |
| Sophisticated | Type 0 | 3 | *Horns* | * | *Horns* | * |
| Spend lots of money | Type 1 | 2 | *Curved 1*, *L*, *Curved L*, *Baby O* | *Curved 1*, *L*, *Curved L*, *Baby O* | *Curved 1*, *Baby O* | *Curved 1*, *Baby O* |
| Stay | Type 1 | 2 | *Y* | *Y* | *Y* | *Y* |
| Story 2 | Type 1 | 1 | *Beak*, *O* | *Beak*, *O* | *Beak*, *O* | *Beak*, *O* |
| Suit | Type 1 | 3 | *A* | *A* | *A* | *A* |
| Suspend | Type 1 | 1 | *Curved 1* | *Curved 1* | *Curved 1* | *Curved 1* |
| Suspenders 2 | Type 1 | 3 | *Curved V* | *Curved V* | *Open H*, *Curved V* | *Open H*, *Curved V* |
| Take off | Type 3 | 4 | *Open 8*, *Ily* | *Closed B* | *Open 8*, *Ily* | *Closed B* |
| Tornado | Type 1 | 1 | *1*, *Open 8*, *L*, *Curved L*, *Q* | *1*, *Open 8*, *L*, *Curved L*, *Q* | *1*, *Open 8*, *Curved L*, *Q* | *1*, *Open 8*, *Curved L*, *Q* |
| Underwear 2 | Type 0 | 1 | *E*, *Closed E*, *S* | * | *E*, *Closed E* | * |
| Unicorn | Type 0 | 1 | *1*, *D* | * | *1*, *D* | * |
| Wade 2 | Type 1 | 2 | *4*, *5*, *Spread Flat 5*, *Stacked 5* | *4*, *5*, *Spread Flat 5*, *Stacked 5* | *4*, *5*, *Spread Flat 5* | *4*, *5*, *Spread Flat 5* |
| Water | Type 0 | 2 | *W* | * | *W* | * |
| We | Type 0 | 2 | *1*, *Curved 1*, *D* | * | *1*, *Curved 1*, *D* | * |
| Weather | Type 1 | 4 | *5*, *W* | *5*, *W* | *5*, *W* | *5*, *W* |
| Weave | Type 1 | 1 | *4* | *4* | *4* | *4* |
| Why | Type 0 | 2 | *Open 8*, *Y* | * | *Open 8*, *Open B*, *Y* | * |

Table A.20: List 3 of the recorded signs with the parameters examined in this work (Part 2)

| Sign | Orientation(Start) | | Location(Start) | | Movement | | $SC_{im}$ | | $SC_{rp}$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | d | nd | d | nd | d | nd | d | nd | d | nd |
| Baby 2 | *FrontUp*, *InBack*, *InUp* | *DownBack*, *DownFront*, *DownIn* | *TorsoTop* | *None* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Basic | *InDown* | *InDown* | *NonDominant* | *Neutral* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Belt | *DownBack*, *DownIn*, *InBack* | *DownBack*, *DownIn*, *InBack* | *TorsoBottom* | *TorsoBottom* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Belt 2 | *FrontIn*, *OutFront* | *FrontIn*, *OutFront* | *TorsoBottom* | *TorsoBottom* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Bra | *InBack*, *InDown* | *InBack*, *InDown* | *TorsoTop* | *TorsoTop* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Bread | *InBack*, *InDown* | *InBack*, *InDown* | *NonDominant* | *Neutral* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Brown | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Climb ladder | *FrontDown*, *UpFront*, *UpIn* | *FrontDown*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Cooperate | *FrontIn*, *InBack*, *UpFront*, *UpIn* | *FrontIn*, *InBack*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Crawl 2 | *DownBack*, *FrontDown*, *FrontIn* | *DownBack*, *FrontDown*, *FrontIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Crewcut | *InDown*, *UpIn* | *InDown*, *UpIn* | *Head* | *Head* | *None* | *None* | ✗ | ✗ | ✗ | ✗ |
| Cross heart | *InBack*, *InDown*, *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *TorsoTop* | *None* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Develop 2 | *FrontDown*, *FrontIn*, *UpFront* | *FrontDown*, *FrontIn* | *NonDominant* | *Neutral* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Die | *FrontDown*, *FrontIn*, *FrontUp*, *InBack*, *InUp*, *UpFront* | *FrontDown*, *FrontIn*, *FrontUp*, *InBack*, *InUp*, *UpFront* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Dig | *FrontDown*, *FrontIn* | *DownIn*, *FrontUp*, *InBack*, *InUp* | *NonDominant* | *Neutral* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Dinner 2 | *InUp*, *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Dissect | *FrontDown*, *InDown* | *FrontIn*, *FrontUp* | *NonDominant* | *Neutral* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Dresser | *FrontIn*, *FrontUp*, *InBack* | *FrontIn*, *FrontUp*, *InBack* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✓ | ✓ | ✓ | ✓ |
| Emotion | *InBack*, *UpBack*, *UpIn* | *InBack*, *UpBack*, *UpIn* | *TorsoTop* | *TorsoTop* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Feed 2 | *DownBack*, *InBack*, *InUp* | *DownBack*, *DownFront*, *DownIn* | *Neutral* | *None* | *Curved* | *None* | ✓ | ✗ | ✓ | ✗ |
| Frisk | *FrontDown*, *FrontIn*, *OutFront* | *FrontDown*, *FrontIn*, *OutFront* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Ghost | *FrontOut*, *FrontUp*, *InBack*, *InDown*, *InUp* | *FrontOut*, *FrontUp*, *InBack*, *InDown*, *InUp* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Glue 1 | *FrontDown*, *FrontIn* | *DownFront*, *FrontIn*, *FrontUp* | *NonDominant* | *Neutral* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Goody goody | *UpFront* | *DownBack*, *DownFront*, *DownIn* | *Neutral* | *None* | *None* | *None* | ✗ | ✗ | ✗ | ✗ |
| Haircut | *UpBack*, *UpIn* | *UpBack*, *UpIn* | *Head* | *Head* | *Curved* | *Curved* | ✓ | ✓ | ✗ | ✗ |
| Imagine | *InDown*, *UpBack*, *UpIn* | *InDown*, *UpBack*, *UpIn* | *Head* | *Head* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Inspire | *FrontIn*, *InBack*, *UpBack*, *UpIn* | *FrontIn*, *InBack*, *UpBack*, *UpIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✓ | ✓ | ✓ | ✓ |
| Interview | *InDown*, *UpIn* | *InDown*, *UpIn* | *Head* | *Head* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Introduce | *FrontIn*, *FrontUp*, *OutFront*, *UpFront*, *UpIn* | *FrontIn*, *FrontUp*, *OutFront*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Intuitive | *DownBack*, *DownIn*, *FrontDown*, *FrontIn*, *InBack* | *DownBack*, *DownFront*, *DownIn* | *TorsoBottom* | *None* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Knitting 3 | *FrontDown*, *InDown*, *UpFront*, *UpIn* | *FrontDown*, *InDown*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Lift 1 | *FrontIn*, *FrontUp*, *OutUp*, *UpBack* | *FrontIn*, *FrontUp*, *OutUp*, *UpBack* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Load 2 | *FrontDown*, *FrontIn*, *InDown*, *UpFront*, *UpIn* | *FrontDown*, *FrontIn*, *InDown*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✓ | ✓ | ✓ | ✓ |
| Magnet | *UpFront*, *UpIn* | *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Missing | *FrontUp*, *UpBack* | *InBack*, *UpIn* | *NonDominant* | *Neutral* | *Straight* | *None* | ✓ | ✗ | ✓ | ✗ |
| Model 1 | *InDown*, *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Shoulder* | *None* | *None* | *None* | ✗ | ✗ | ✗ | ✗ |
| Monday | *FrontUp*, *OutUp*, *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Neutral* | *None* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Nephew | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | *None* | *None* | ✗ | ✗ | ✗ | ✗ |

Continued on next page

Continued from previous page

| Sign | Orientation(Start) | | Location(Start) | | Movement | | $SC_{im}$ | | $SC_{rp}$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | d | nd | d | nd | d | nd | d | nd | d | nd |
| Paranoid | *BackUp*, *UpBack*, *UpFront*, *UpIn* | *BackUp*, *UpBack*, *UpFront*, *UpIn* | *Head* | *Head* | *None* | *None* | ✗ | ✗ | ✗ | ✗ |
| Perceive | *InDown*, *UpIn* | *InDown*, *UpIn* | *Head* | *Head* | *None* | *None* | ✗ | ✗ | ✗ | ✗ |
| Perspective | *FrontDown*, *FrontIn*, *InDown*, *InFront*, *UpFront*, *UpIn* | *FrontDown*, *InDown*, *UpFront*, *UpIn* | *NonDominant* | *Neutral* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Rain | *InDown*, *UpFront*, *UpIn* | *InDown*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Respect | *FrontDown*, *InDown*, *UpIn* | *FrontDown*, *InDown*, *UpIn* | *Head* | *Head* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Retire | *InBack*, *InDown*, *UpIn* | *InBack*, *InDown*, *UpIn* | *Shoulder* | *Shoulder* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Rob | *FrontIn*, *UpFront*, *UpIn* | *FrontIn*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Scoop | *FrontDown* | *FrontDown*, *FrontIn*, *InBack*, *InDown* | *NonDominant* | *Neutral* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Skate | *FrontIn*, *FrontUp*, *InBack* | *FrontIn*, *FrontUp*, *InBack* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Smile | *InDown*, *UpIn* | *InDown*, *UpIn* | *Head* | *Head* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Snap | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Neutral* | *None* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Sophisticated | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Spend lots of money | *FrontIn*, *FrontUp* | *FrontIn*, *FrontUp* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Stay | *FrontDown*, *UpFront*, *UpIn* | *FrontDown*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Story 2 | *FrontIn*, *InBack*, *InDown*, *InUp*, *UpFront*, *UpIn* | *FrontIn*, *InBack*, *InDown*, *InUp*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Suit | *FrontIn*, *FrontUp*, *UpBack*, *UpIn* | *FrontIn*, *FrontUp*, *UpBack*, *UpIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Suspend | *InBack*, *InDown* | *InBack*, *InDown* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Suspenders 2 | *InFront*, *UpFront*, *UpIn* | *InFront*, *UpFront*, *UpIn* | *Shoulder* | *Shoulder* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Take off | *FrontDown*, *FrontIn*, *InDown*, *InFront*, *UpIn* | *FrontUp*, *InBack*, *InUp* | *NonDominant* | *Neutral* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Tornado | *FrontIn*, *FrontUp*, *InBack*, *InDown*, *InUp*, *UpBack*, *UpIn* | *FrontIn*, *FrontUp*, *InBack*, *InDown*, *InUp*, *UpBack*, *UpIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Underwear 2 | *UpBack* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | *None* | *None* | ✗ | ✗ | ✗ | ✗ |
| Unicorn | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| Wade 2 | *FrontDown*, *InBack*, *InDown*, *UpIn* | *FrontDown*, *InBack*, *InDown*, *UpIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Water | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |
| We | *BackDown*, *BackOut*, *InBack*, *InDown*, *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Shoulder* | *None* | *Curved* | *None* | ✗ | ✗ | ✗ | ✗ |
| Weather | *UpFront*, *UpIn* | *UpFront*, *UpIn* | *Neutral* | *Neutral* | *Curved* | *Curved* | ✗ | ✗ | ✗ | ✗ |
| Weave | *InDown*, *InFront* | *InDown*, *InFront* | *Neutral* | *Neutral* | *Straight* | *Straight* | ✗ | ✗ | ✗ | ✗ |
| Why | *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | *Straight* | *None* | ✗ | ✗ | ✗ | ✗ |

Table A.21: List 3 of the recorded signs with the parameters examined in this work (Part 3)

| Sign | Orientation(End) | | Location(End) | | $FC_i$ | | $FC_m$ | | $FC_p$ | | $FC_r$ | | UR | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | d | nd | d | nd | d | nd | d | nd | d | nd | d | nd | d | nd |
| Baby 2 | *InBack*, *InUp* | *DownBack*, *DownFront*, *DownIn* | *TorsoTop* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Basic | *InDown* | *InDown* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Belt | *DownBack*, *FrontIn*, *InBack*, *InDown* | *DownBack*, *FrontIn*, *InBack*, *InDown* | *TorsoBottom* | *TorsoBottom* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Belt 2 | *FrontIn*, *InBack* | *FrontIn*, *InBack* | *TorsoBottom* | *TorsoBottom* | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Bra | *DownBack*, *DownOut*, *InBack*, *InDown* | *DownBack*, *DownOut*, *InBack*, *InDown* | *TorsoTop* | *TorsoTop* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Bread | *InBack*, *InDown* | *InBack*, *InDown* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Brown | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Climb ladder | *UpFront*, *UpIn* | *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ |
| Cooperate | *FrontIn*, *UpBack*, *UpFront*, *UpIn* | *FrontIn*, *UpBack*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Crawl 2 | *DownBack*, *FrontDown* | *DownBack*, *FrontDown* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Crewcut | *BackIn*, *BackUp*, *InFront*, *UpFront* | *BackIn*, *BackUp*, *InFront*, *UpFront* | *Head* | *Head* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Cross heart | *InBack*, *InDown*, *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *TorsoTop* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Develop 2 | *FrontDown*, *UpFront* | *FrontDown*, *FrontIn* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Die | *DownBack*, *FrontDown*, *FrontUp*, *InDown*, *OutUp* | *DownBack*, *FrontDown*, *FrontUp*, *InDown*, *OutUp* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✓ | ✓ |
| Dig | *InUp*, *UpBack*, *UpIn* | *DownIn*, *FrontUp*, *InBack*, *InUp* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | * | ✗ |
| Dinner 2 | *InUp*, *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Dissect | *FrontDown*, *InDown* | *FrontIn*, *FrontUp* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Dresser | *DownFront*, *FrontIn*, *FrontUp* | *DownFront*, *FrontIn*, *FrontUp* | *Neutral* | *Neutral* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ |
| Emotion | *InBack*, *UpBack*, *UpIn* | *InBack*, *UpBack*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Feed 2 | *FrontIn*, *FrontUp*, *InBack*, *InUp* | *DownBack*, *DownFront*, *DownIn* | *Neutral* | *None* | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✗ | ✗ |
| Frisk | *FrontDown*, *FrontIn* | *FrontDown*, *FrontIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Ghost | *FrontUp*, *InBack*, *InDown*, *InUp*, *UpFront*, *UpIn* | *FrontUp*, *InBack*, *InDown*, *InUp*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Glue 1 | *FrontDown*, *InBack*, *InDown*, *UpIn* | *DownFront*, *FrontIn*, *FrontUp* | *NonDominant* | *Neutral* | ✓ | ✗ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Goody goody | *UpFront* | *DownBack*, *DownFront*, *DownIn* | *Neutral* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Haircut | *BackDown*, *UpBack*, *UpIn* | *BackDown*, *UpBack*, *UpIn* | *Head* | *Head* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Imagine | *InDown*, *UpBack*, *UpFront*, *UpIn* | *InDown*, *UpBack*, *UpFront*, *UpIn* | *Head* | *Head* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Inspire | *FrontIn*, *UpBack*, *UpIn* | *FrontIn*, *UpBack*, *UpIn* | *Neutral* | *Neutral* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ |
| Interview | *InDown*, *UpIn* | *InDown*, *UpIn* | *Head* | *Head* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Introduce | *FrontIn*, *FrontUp*, *InBack*, *InUp* | *FrontIn*, *FrontUp*, *InBack*, *InUp* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Intuitive | *DownBack*, *DownIn*, *FrontDown*, *FrontIn*, *InBack* | *DownBack*, *DownFront*, *DownIn* | *TorsoBottom* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Knitting 3 | *FrontDown*, *InDown*, *UpFront*, *UpIn* | *FrontDown*, *InDown*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Lift 1 | *FrontUp*, *UpBack* | *FrontUp*, *UpBack* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Load 2 | *DownBack*, *DownOut*, *FrontDown*, *InDown*, *OutDown* | *DownBack*, *DownOut*, *FrontDown*, *InDown*, *OutDown* | *Neutral* | *Neutral* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ |
| Magnet | *InDown*, *UpFront*, *UpIn* | *InDown*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Missing | *FrontUp*, *InBack*, *InUp*, *UpOut* | *InBack*, *UpIn* | *NonDominant* | *Neutral* | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✗ | ✗ |
| Model 1 | *InDown*, *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Shoulder* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Monday | *FrontUp*, *OutUp*, *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Neutral* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Nephew | *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✓ | ✗ |

Continued on next page

Continued from previous page

| Sign | Orientation(End) | | Location(End) | | $FC_i$ | | $FC_m$ | | $FC_p$ | | $FC_r$ | | UR | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | d | nd | d | nd | d | nd | d | nd | d | nd | d | nd | d | nd |
| Paranoid | *UpIn* | *UpIn* | *Head* | *Head* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Perceive | *InDown*, *UpIn* | *InDown*, *UpIn* | *Head* | *Head* | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Perspective | *InDown*, *UpBack*, *UpIn* | *FrontDown*, *InDown*, *UpFront*, *UpIn* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✓ | ✗ |
| Rain | *InDown*, *InFront*, *UpIn* | *InDown*, *InFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Respect | *InDown*, *UpIn* | *InDown*, *UpIn* | *Head* | *Head* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Retire | *InBack*, *InDown*, *UpIn* | *InBack*, *InDown*, *UpIn* | *Shoulder* | *Shoulder* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Rob | *FrontDown*, *FrontIn* | *FrontDown*, *FrontIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Scoop | *InDown*, *UpFront*, *UpIn* | *FrontDown*, *FrontIn*, *InBack*, *InDown* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✓ | ✗ |
| Skate | *FrontIn*, *FrontUp*, *InBack*, *UpBack* | *FrontIn*, *FrontUp*, *InBack*, *UpBack* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Smile | *BackDown*, *InDown*, *UpBack*, *UpIn* | *BackDown*, *InDown*, *UpBack*, *UpIn* | *Head* | *Head* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Snap | *FrontOut*, *FrontUp*, *InUp*, *UpBack* | *DownBack*, *DownFront*, *DownIn* | *Neutral* | *None* | ✗ | ✗ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✓ | ✗ |
| Sophisticated | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Spend lots of money | *FrontUp*, *UpBack*, *UpIn* | *FrontUp*, *UpBack*, *UpIn* | *Neutral* | *Neutral* | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Stay | *DownBack*, *FrontDown* | *DownBack*, *FrontDown* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Story 2 | *FrontIn*, *InDown*, *UpBack*, *UpFront*, *UpIn* | *FrontIn*, *InDown*, *UpBack*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✓ | ✓ |
| Suit | *InBack*, *InDown* | *InBack*, *InDown* | *TorsoTop* | *TorsoTop* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | * | * |
| Suspend | *FrontIn*, *InBack*, *InDown*, *UpIn* | *FrontIn*, *InBack*, *InDown*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Suspenders 2 | *DownBack*, *InBack*, *InDown*, *InUp* | *DownBack*, *InBack*, *InDown*, *InUp* | *TorsoBottom* | *TorsoBottom* | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | ✓ |
| Take off | *FrontDown*, *UpFront* | *FrontUp*, *InBack*, *InUp* | *NonDominant* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Tornado | *FrontIn*, *InBack*, *InDown*, *InUp*, *UpBack*, *UpFront*, *UpIn* | *FrontIn*, *InBack*, *InDown*, *InUp*, *UpBack*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | * | * | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Underwear 2 | *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✓ | ✗ | ✗ | ✗ |
| Unicorn | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Wade 2 | *FrontDown*, *InBack*, *InDown* | *FrontDown*, *InBack*, *InDown* | *Neutral* | *Neutral* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ |
| Water | *UpFront*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| We | *BackDown*, *BackOut*, *InDown*, *InFront*, *UpBack* | *DownBack*, *DownFront*, *DownIn* | *Shoulder* | *None* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✓ | ✗ |
| Weather | *FrontDown*, *FrontIn*, *UpFront*, *UpIn* | *FrontDown*, *FrontIn*, *UpFront*, *UpIn* | *Neutral* | *Neutral* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Weave | *InDown* | *InDown* | *Neutral* | *Neutral* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ |
| Why | *UpBack*, *UpIn* | *DownBack*, *DownFront*, *DownIn* | *Head* | *None* | ✓ | ✗ | ✓ | ✗ | ✗ | ✗ | ✓ | ✗ | ✗ | ✗ |

Table A.22: Recorded signs by participants before exclusion of erroneous recordings

| Sign | P1 | P2 | P3.1 | P3.2 | Total |
|---|---|---|---|---|---|
| Accept | 3 | 3 | 3 | 0 | 9 |
| Acquire | 3 | 3 | 0 | 3 | 9 |
| Admire | 3 | 3 | 3 | 0 | 9 |
| Adopt | 3 | 3 | 0 | 3 | 9 |
| Airplane | 3 | 3 | 0 | 3 | 9 |
| All day | 3 | 3 | 2 | 0 | 8 |
| All together | 3 | 3 | 3 | 0 | 9 |
| All way | 3 | 3 | 3 | 0 | 9 |
| Angel | 3 | 3 | 3 | 0 | 9 |
| Baby 2 | 3 | 3 | 3 | 0 | 9 |
| Badge | 3 | 3 | 0 | 3 | 9 |
| Bag 2 | 3 | 3 | 3 | 0 | 9 |
| Band aid 2 | 3 | 3 | 3 | 0 | 9 |
| Barrette | 3 | 3 | 3 | 0 | 9 |
| Basic | 3 | 3 | 3 | 0 | 9 |
| Belt | 3 | 3 | 3 | 0 | 9 |
| Belt 2 | 3 | 3 | 3 | 0 | 9 |
| Bib | 3 | 3 | 3 | 0 | 9 |
| Boiling mad | 3 | 3 | 3 | 0 | 9 |
| Bottle | 3 | 3 | 3 | 0 | 9 |
| Bowling 2 | 3 | 3 | 3 | 0 | 9 |
| Bra | 3 | 3 | 3 | 0 | 9 |
| Bracelet | 3 | 3 | 3 | 0 | 9 |
| Brag | 3 | 3 | 0 | 3 | 9 |
| Bread | 3 | 3 | 3 | 0 | 9 |
| Brown | 3 | 3 | 3 | 0 | 9 |
| Building | 3 | 3 | 3 | 0 | 9 |
| Bury 3 | 3 | 3 | 0 | 3 | 9 |
| Camping | 3 | 3 | 3 | 0 | 9 |
| Carry | 3 | 3 | 0 | 3 | 9 |
| Cheese grater | 3 | 3 | 0 | 3 | 9 |
| Climb ladder | 3 | 3 | 3 | 0 | 9 |
| Cloud 1 | 3 | 3 | 3 | 0 | 9 |
| Come | 3 | 3 | 0 | 3 | 9 |
| Cooperate | 3 | 3 | 3 | 0 | 9 |
| Correct | 3 | 3 | 0 | 3 | 9 |

Continued on next page

Continued from previous page

| Sign | P1 | P2 | P3.1 | P3.2 | Total |
|---|---|---|---|---|---|
| Cowboy | 3 | 3 | 3 | 0 | 9 |
| Crash plane | 3 | 3 | 3 | 0 | 9 |
| Crawl 2 | 3 | 3 | 3 | 0 | 9 |
| Credit card | 3 | 3 | 0 | 3 | 9 |
| Crewcut | 3 | 3 | 3 | 0 | 9 |
| Cross heart | 3 | 3 | 3 | 0 | 9 |
| Cry 3 | 3 | 3 | 0 | 3 | 9 |
| Develop 2 | 3 | 3 | 3 | 0 | 9 |
| Diaper | 3 | 3 | 3 | 0 | 9 |
| Dictionary | 3 | 3 | 0 | 3 | 9 |
| Die | 3 | 3 | 3 | 0 | 9 |
| Dig | 3 | 3 | 3 | 0 | 9 |
| Dig up | 3 | 3 | 0 | 3 | 9 |
| Dinner 2 | 3 | 3 | 3 | 0 | 9 |
| Dissect | 3 | 3 | 3 | 0 | 9 |
| Doctor 2 | 3 | 0 | 0 | 3 | 6 |
| Drag | 3 | 3 | 0 | 3 | 9 |
| Drawer | 3 | 3 | 0 | 3 | 9 |
| Dress up | 3 | 3 | 3 | 0 | 9 |
| Dresser | 3 | 3 | 3 | 0 | 9 |
| Duty | 3 | 3 | 3 | 0 | 9 |
| Each other | 3 | 3 | 0 | 3 | 9 |
| Easter | 3 | 3 | 3 | 0 | 9 |
| Edit 1 | 3 | 3 | 0 | 3 | 9 |
| Elementary school | 3 | 3 | 3 | 0 | 9 |
| Emotion | 3 | 3 | 3 | 0 | 9 |
| Erupt 2 | 3 | 3 | 3 | 0 | 9 |
| Evaluate | 3 | 3 | 0 | 3 | 9 |
| Experiment | 3 | 3 | 0 | 3 | 9 |
| Explanation | 3 | 3 | 0 | 3 | 9 |
| Express | 3 | 3 | 3 | 0 | 9 |
| Fade | 3 | 3 | 0 | 3 | 9 |
| Feed 2 | 3 | 3 | 3 | 0 | 9 |
| Feedback | 3 | 3 | 3 | 0 | 9 |
| Frisk | 3 | 3 | 3 | 0 | 9 |
| Frown 2 | 3 | 3 | 3 | 0 | 9 |
| Frustrate | 3 | 3 | 0 | 3 | 9 |

Continued on next page

Continued from previous page

| Sign | P1 | P2 | P3.1 | P3.2 | Total |
|---|---|---|---|---|---|
| GET OFF | 3 | 3 | 0 | 3 | 9 |
| GHOST | 3 | 3 | 3 | 0 | 9 |
| GLOVES 4 | 3 | 3 | 0 | 3 | 9 |
| GLUE 1 | 3 | 3 | 3 | 0 | 9 |
| GLUE 2 | 3 | 3 | 3 | 0 | 9 |
| GOODY GOODY | 3 | 3 | 3 | 0 | 9 |
| GOODY GOODY SHOE OPPOSITE | 3 | 3 | 3 | 0 | 9 |
| GOOSEBUMPS | 3 | 3 | 3 | 0 | 9 |
| GRIEVE | 3 | 3 | 0 | 3 | 9 |
| HAIRCUT | 3 | 3 | 3 | 0 | 9 |
| HAIRDRESSER | 3 | 3 | 0 | 3 | 9 |
| HALF HOUR | 3 | 3 | 3 | 0 | 9 |
| HANDSOME 3 | 3 | 3 | 0 | 3 | 9 |
| HANG CLOTHES | 3 | 3 | 3 | 0 | 9 |
| HIPPO 2 | 3 | 3 | 3 | 0 | 9 |
| HUNT | 0 | 3 | 0 | 3 | 6 |
| HURDLE 1 | 3 | 3 | 0 | 3 | 9 |
| HURRICANE | 3 | 3 | 3 | 0 | 9 |
| IMAGINE | 3 | 3 | 3 | 0 | 9 |
| INCREDIBLE | 3 | 3 | 3 | 0 | 9 |
| INSPIRE | 3 | 3 | 3 | 0 | 9 |
| INSPIRING | 3 | 3 | 0 | 3 | 9 |
| INTERVIEW | 3 | 3 | 3 | 0 | 9 |
| INTRODUCE | 3 | 3 | 3 | 0 | 9 |
| INTUITIVE | 3 | 3 | 3 | 0 | 9 |
| INVISIBLE | 3 | 3 | 3 | 0 | 9 |
| JAPAN | 3 | 3 | 0 | 3 | 9 |
| JEWELRY | 3 | 3 | 3 | 0 | 9 |
| KNITTING 3 | 3 | 3 | 3 | 0 | 9 |
| LICK ENVELOPE 2 | 3 | 3 | 0 | 3 | 9 |
| LIFT 1 | 3 | 3 | 3 | 0 | 9 |
| LOAD | 3 | 3 | 3 | 0 | 9 |
| LOAD 2 | 3 | 3 | 3 | 0 | 9 |
| LOOK AT | 3 | 3 | 0 | 3 | 9 |
| MAGNET | 3 | 3 | 3 | 0 | 9 |
| MAID | 3 | 3 | 3 | 0 | 9 |
| MATCH | 3 | 3 | 0 | 3 | 9 |

Continued on next page

Continued from previous page

| Sign | P1 | P2 | P3.1 | P3.2 | Total |
|---|---|---|---|---|---|
| Measure 2 | 3 | 3 | 0 | 3 | 9 |
| Meditate 3 | 3 | 3 | 0 | 3 | 9 |
| Mexican 2 | 3 | 3 | 3 | 0 | 9 |
| Mime | 3 | 3 | 3 | 0 | 9 |
| Mischievous | 3 | 3 | 0 | 3 | 9 |
| Missing | 3 | 3 | 3 | 0 | 9 |
| Mix 3 | 3 | 3 | 0 | 3 | 9 |
| Model 1 | 3 | 3 | 3 | 0 | 9 |
| Monday | 3 | 3 | 3 | 0 | 9 |
| Monkey | 3 | 3 | 0 | 3 | 9 |
| Museum | 3 | 3 | 0 | 3 | 9 |
| Mustache | 3 | 3 | 3 | 0 | 9 |
| Navy | 3 | 3 | 3 | 0 | 9 |
| Necklace 2 | 3 | 3 | 3 | 0 | 9 |
| Nephew | 3 | 3 | 3 | 0 | 9 |
| Network | 3 | 3 | 3 | 0 | 9 |
| Neutral | 3 | 3 | 3 | 0 | 9 |
| Next to | 3 | 3 | 0 | 3 | 9 |
| Niece | 3 | 3 | 0 | 3 | 9 |
| Octopus | 3 | 3 | 0 | 3 | 9 |
| Pants | 3 | 3 | 3 | 0 | 9 |
| Paper | 3 | 3 | 0 | 3 | 9 |
| Paranoid | 3 | 3 | 3 | 0 | 9 |
| Penguin | 3 | 3 | 3 | 0 | 9 |
| Perceive | 3 | 3 | 3 | 0 | 9 |
| Perspective | 3 | 3 | 3 | 0 | 9 |
| Place | 3 | 3 | 3 | 0 | 9 |
| Play | 3 | 3 | 3 | 0 | 9 |
| Porcupine | 3 | 3 | 3 | 0 | 9 |
| Problem 1 | 3 | 3 | 0 | 3 | 9 |
| Put aside | 3 | 3 | 3 | 0 | 9 |
| Rain | 3 | 3 | 3 | 0 | 9 |
| React | 3 | 3 | 3 | 0 | 9 |
| Ready | 3 | 3 | 0 | 3 | 9 |
| Record | 3 | 3 | 3 | 0 | 9 |
| Reduce | 3 | 3 | 3 | 0 | 9 |
| Respect | 3 | 3 | 3 | 0 | 9 |

Continued on next page

Continued from previous page

| Sign | P1 | P2 | P3.1 | P3.2 | Total |
|---|---|---|---|---|---|
| RETIRE | 3 | 3 | 3 | 0 | 9 |
| RIGHT 1 | 3 | 3 | 3 | 0 | 9 |
| ROB | 3 | 3 | 3 | 0 | 9 |
| ROCKING CHAIR 2 | 3 | 3 | 3 | 0 | 9 |
| ROLLERSKATING | 3 | 3 | 3 | 0 | 9 |
| SAUCE | 3 | 3 | 3 | 0 | 9 |
| SCARCELY | 3 | 3 | 0 | 3 | 9 |
| SCHEDULE | 3 | 3 | 3 | 0 | 9 |
| SCOOP | 3 | 3 | 3 | 0 | 9 |
| SCOUT | 3 | 3 | 0 | 3 | 9 |
| SENTENCE | 3 | 3 | 0 | 3 | 9 |
| SEVEN | 3 | 3 | 3 | 0 | 9 |
| SHARK 1 | 3 | 3 | 0 | 3 | 9 |
| SHARPEN 1 | 3 | 3 | 3 | 0 | 9 |
| SHOOTING | 0 | 3 | 0 | 3 | 6 |
| SHOULDER | 3 | 3 | 0 | 3 | 9 |
| SINCE | 3 | 3 | 0 | 3 | 9 |
| SKATE | 3 | 3 | 3 | 0 | 9 |
| SLED 2 | 3 | 3 | 3 | 0 | 9 |
| SMILE | 3 | 3 | 3 | 0 | 9 |
| SMOOTH | 3 | 3 | 3 | 0 | 9 |
| SNAP | 3 | 3 | 3 | 0 | 9 |
| SOPHISTICATED | 3 | 3 | 3 | 0 | 9 |
| SOUP | 3 | 3 | 3 | 0 | 9 |
| SPAGHETTI | 3 | 3 | 3 | 0 | 9 |
| SPEND | 3 | 3 | 0 | 3 | 9 |
| SPEND LOTS OF MONEY | 3 | 3 | 3 | 0 | 9 |
| SPREAD 3 | 3 | 3 | 0 | 3 | 9 |
| STARBUCKS | 3 | 3 | 0 | 3 | 9 |
| STAY | 3 | 3 | 3 | 0 | 9 |
| STORY 2 | 3 | 3 | 3 | 0 | 9 |
| SUIT | 3 | 3 | 3 | 0 | 9 |
| SUSPEND | 3 | 0 | 3 | 0 | 6 |
| SUSPENDERS 1 | 3 | 3 | 0 | 3 | 9 |
| SUSPENDERS 2 | 3 | 3 | 3 | 0 | 9 |
| SWEATER | 3 | 3 | 0 | 3 | 9 |
| SWEEP | 3 | 3 | 3 | 0 | 9 |

Continued on next page

Continued from previous page

| Sign | P1 | P2 | P3.1 | P3.2 | Total |
|---|---|---|---|---|---|
| Take off | 3 | 3 | 3 | 0 | 9 |
| Thin | 3 | 3 | 0 | 3 | 9 |
| Tornado | 3 | 3 | 3 | 0 | 9 |
| Transgender | 3 | 3 | 0 | 3 | 9 |
| Truth | 3 | 3 | 3 | 0 | 9 |
| Underwear | 3 | 3 | 3 | 0 | 9 |
| Underwear 2 | 3 | 3 | 3 | 0 | 9 |
| Unicorn | 3 | 3 | 3 | 0 | 9 |
| Vain | 3 | 3 | 3 | 0 | 9 |
| Vase | 3 | 3 | 3 | 0 | 9 |
| Videogame | 3 | 3 | 3 | 0 | 9 |
| Visit | 3 | 3 | 0 | 3 | 9 |
| Voice | 3 | 3 | 3 | 0 | 9 |
| Wade 2 | 3 | 3 | 3 | 0 | 9 |
| Water | 3 | 3 | 3 | 0 | 9 |
| Wave | 3 | 3 | 0 | 3 | 9 |
| We | 3 | 3 | 3 | 0 | 9 |
| Weather | 3 | 3 | 3 | 0 | 9 |
| Weave | 3 | 3 | 3 | 0 | 9 |
| Weird | 3 | 3 | 0 | 3 | 9 |
| Wheelchair | 3 | 3 | 0 | 3 | 9 |
| Whisper | 3 | 3 | 3 | 0 | 9 |
| Why | 3 | 3 | 3 | 0 | 9 |
| Win | 3 | 3 | 3 | 0 | 9 |
| World | 3 | 3 | 3 | 0 | 9 |
| Zebra | 3 | 3 | 3 | 0 | 9 |
| Total | 624 | 624 | 431 | 198 | 1877 |

## A.5 LIST OF ABBREVIATIONS

| | |
|---|---|
| **Acc.** | Accuracy |
| **AMK** | Aggregative Meta-Classifier |
| **ANN** | Artificial Neural Network |
| **AR** | Augmented Reality |
| **ASCII** | American Standard Code for Information Interchange |
| **ASL** | American Sign Language |
| **ASLLRP** | American Sign Language Linguistic Research Project |
| **ASLLVD** | American Sign Language Lexicon Video Dataset |
| **ASLSJ** | ASL Sign Jotting |
| **Auslan** | Australian Sign Language |
| **Bew.** | Movement |
| **BFSG** | Accessibility Strengthening Act |
| **BLE** | Bluetooth Low Energy |
| **BPNN** | Back Propagation Neural Network |
| **BSL** | British Sign Language |
| **BST** | Battison Sign Type |
| **CNN** | Convolutional Neural Network |
| **DBSCAN** | Density-Based Spatial Clustering of Applications with Noise |
| **DGS** | German Sign Language |
| **DIP** | Distal Interphalangeal |
| **DoF** | Degrees of Freedom |
| **DSP** | DawnSignPress |
| **DST** | Dempster-Shafer Theory |
| **DT** | Decision Tree |
| **Dyn.** | Dynamic |
| **EIT** | Electrical Impedance Tomography |
| **EMF** | Electromagnetic Field |
| **EMG** | Electromyography |
| **F1** | F1-Score |
| **FC** | Change in flexion (Flexion Change) |
| **FinSL** | Finnish Sign Language |
| **FN** | False Negative |
| **FNN** | Feed-forward Neural Network |
| **FP** | False Positive |

| | |
|---|---|
| **FPU** | Floating Point Unit |
| **GA** | Genetic Algorithm |
| **HamNoSys** | Hamburg Sign Language Notation System |
| **HCI** | Human-Computer Interaction |
| **HF** | Handshape |
| **HMM** | Hidden Markov Model |
| **iEMG** | Intramuscular Electromyography |
| **IMU** | Inertial Measurement Unit |
| **IPA** | International Phonetic Alphabet |
| **ISL** | Israeli Sign Language |
| **KMK** | Composite Meta-Classifier |
| **kNN** | k-Nearest Neighbor |
| **LDA** | Linear Discriminant Analysis |
| **LiPo** | Lithium Polymer |
| **LOSOCV** | Leave-One-Subject-Out Cross-Validation |
| **LR** | Logistic Regression |
| **LSF** | French Sign Language |
| **MAD** | Mean Absolute Deviation |
| **MCP** | Metacarpophalangeal |
| **ML** | Machine Learning |
| **MMG** | Mechanomyography |
| **MQTT** | Message Queuing Telemetry Transport |
| **NB** | Naive Bayes |
| **NGT** | Dutch Sign Language |
| **Ori.** | Orientation |
| **OvA** | One-versus-All |
| **OvO** | One-versus-One |
| **ParaSignRec** | Parametric Framework for Sign Language Recognition |
| **PCA** | Principal Component Analysis |
| **PIP** | Proximal Interphalangeal |
| **PMHC** | Prosodic Model Handshape Coding |
| **Prec.** | Precision |
| **RBF** | Radial Basis Function |
| **RF** | Random Forest |
| **RIT** | Rochester Institute of Technology |
| **RMS** | Root Mean Square |
| **RNN** | Recurrent Neural Network |

| | |
|---|---|
| **RPS** | Rock-Paper-Scissors |
| **RSL** | Russian Sign Language |
| **SC** | Change in spread (Spread Change) |
| **SDK** | Software Development Kit |
| **sEMG** | Surface Electromyography |
| **Sens.** | Sensitivity |
| **SLAAASh** | Sign Language Acquisition Annotation Archiving and Sharing |
| **SLIPA** | Sign Language IPA |
| **SLPA** | Sign Language Phonetic Annotation |
| **SNR** | Signal-to-Noise Ratio |
| **STS** | Swedish Sign Language |
| **SVM** | Support Vector Machine |
| **TN** | True Negative |
| **TNR** | True Negative-Rate |
| **TP** | True Positive |
| **TPR** | True Positive-Rate |
| **TSFEL** | Time Series Feature Extraction Library |
| **TSL** | Taiwanese Sign Language |
| **TSR** | Task Success Rate |
| **UR** | Rotation of the forearm (Ulnar) |
| **VL2** | Voting (Level 2) Meta-Classifier |
| **VR** | Virtual Reality |
| **WiFi** | Wireless Fidelity |
| **WLASL** | Word-Level American Sign Language |
| **ÖGS** | Austrian Sign Language |

# B

## SUPERVISED STUDENT DEGREE THESES

Master's degree theses

[1] Philipp Kaul. "Extraction and Augmentation of Sensor Data from Video Material using MediaPipe." Master Thesis. TU Darmstadt, 2024.

[2] Nicolai Kellerer. "Classification of signs according to Battison based on their articulatory complexity." Master Thesis. TU Darmstadt, 2022.

[3] Marius Kempf. "Konzeption und Entwicklung eines Mechanomyografie-Controllers zur Erkennung von einfachen Fingergesten." Master Thesis. TU Darmstadt, 2021.

[4] Felix Klose. "Recognition of gesture movement, orientation and location using IMUs and Inverse Kinematics." Master Thesis. TU Darmstadt, 2021.

[5] Sarah Kuhlmann. "IMU data segmentation for pose recognition in sign language." Master Thesis. TU Darmstadt, 2023.

[6] Dennis Purdack. "Towards metaclassification of complex handshape recognition." Master Thesis. TU Darmstadt, 2022.

[7] Rebekka Schiller. "Live Classification of Battison's Sign Types." Master Thesis. TU Darmstadt, 2024.

[8] Alexandra Skogseide. "Automatic generation of a parametric sign lexicon with MediaPipe." Master Thesis. TU Darmstadt, 2023.

[9] Qilin Tan. "Towards arm posture detection using sensor fusion and extended Kalman filter." Master Thesis. TU Darmstadt, 2022.

[10] Martin Wende. "Entwicklung einer App zur Ermittlung des ökologischen Fußabdrucks im Kontext einer Mitfahrgelegenheit mit Hilfe von ML." Master Thesis. TU Darmstadt, 2020.

Bachelor's degree theses

[11] Elena Bock. "Classification of hand gestures using MediaPipe with qualitative feedback." Bachelor Thesis. TU Darmstadt, 2023.

[12] Tobias Hartmann. "Sensorbasierte Unterarmgestenerkennung zur Interaktion mit mobilen Anwendungen." Bachelor Thesis. TU Darmstadt, 2020.

[13] Thanh Huan Hoang. "Dancing with the gloves on: Using data gloves and arm controllers to teach the Macarena." Bachelor Thesis. TU Darmstadt, 2023.

[14] Benedikt Hock. "Towards sign parametrization by using MediaPipe." Bachelor Thesis. TU Darmstadt, 2022.

[15] Theo Kastner-Guhl. “Generation of a dynamic virtual construction site based on Building Information Modeling.” Bachelor Thesis. TU Darmstadt, 2020.

[16] Nico Kunz. “Recognition and classification of handshapes of American finger alphabet.” Bachelor Thesis. TU Darmstadt, 2022.

[17] Sebastian Laux. “Augmenting wearable sensor data for sign language recognition.” Bachelor Thesis. TU Darmstadt, 2022.

[18] Leon Petri. “Augmenting Finger Motion Data for Hand Gesture Recognition using Video-based Zero-Shot Domain Adaption.” Bachelor Thesis. TU Darmstadt, 2022.

[19] Minh Pham. “Live Classification of Battison’s Sign Types.” Bachelor Thesis. TU Darmstadt, 2024.

[20] Dennis Purdack. “Entwicklung und Evaluation zweier Ansätze zur Erkennung primitiver Handgesten anhand eines Leap Motion Controllers.” Bachelor Thesis. TU Darmstadt, 2021.

[21] Lea Schott. “Erfassung einfacher Handformen mittels Elektromyografie unter Berücksichtigung verschiedener Konfigurationen.” Bachelor Thesis. TU Darmstadt, 2021.

[22] Maximilian Segers. “Hand Orientation Recognition in Sign Language Videos: Data Fusion and Tool Development.” Bachelor Thesis. TU Darmstadt, 2025.

[23] Tobias Alexander Wach. “Konzeption und Implementierung eines Fingergesten-gesteuerten Spiels.” Bachelor Thesis. TU Darmstadt, 2020.

[24] Sebastian Wolf. “Comparing the suitability of Decision Tree classifier and Support Vector Machines for hand gesture recognition.” Bachelor Thesis. TU Darmstadt, 2021.

# C

# PUBLICATIONS OF THE AUTHOR

## PUBLICATIONS AS FIRST AUTHOR


[1] Philipp Achenbach, Yasmin Göksu, Timo Kullmann, Thomas Tregel, and Stefan Göbel. "Towards handshape identification for automatic gesture recognition using sign notation systems." en. In: *Proceedings of the 8th European Conference on Social Media: A Virtual Conference Hosted By University of Central Lancaster, UCLan, Cyprus 1-2 July 2021*. University of Central Lancashire, Cyprus: Academic Conferences International Limited, July 2021. ISBN: 978-1-914587-01-6. DOI: 10.34190/ESM.21.100.

[2] Philipp Achenbach, Sebastian Laux, Dennis Purdack, Philipp Niklas Müller, and Stefan Göbel. "Give Me a Sign: Using Data Gloves for Static Hand-Shape Recognition." en. In: *Sensors* 23.24 (Dec. 2023), p. 9847. ISSN: 1424-8220. DOI: 10.3390/s23249847. URL: https://www.mdpi.com/1424-8220/23/24/9847 (Last accessed on Jan. 6, 2024).

[3] Philipp Achenbach, Philipp Niklas Muller, Tobias Alexander Wach, Thomas Tregel, and Stefan Gobel. "Rock beats Scissor: SVM based gesture recognition with data gloves." en. In: *2021 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*. Kassel, Germany: IEEE, Mar. 2021, pp. 617–622. ISBN: 978-1-6654-0424-2. DOI: 10.1109/PerComWorkshops51409.2021.9430962. URL: https://ieeexplore.ieee.org/document/9430962/ (Last accessed on Jan. 3, 2024).

[4] Philipp Achenbach, Dennis Purdack, Sebastian Wolf, Philipp Niklas Müller, Thomas Tregel, and Stefan Göbel. "Paper Beats Rock: Elaborating the Best Machine Learning Classifier for Hand Gesture Recognition." en. In: *Serious Games*. Ed. by Heinrich Söbke, Pia Spangenberger, Philipp Müller, and Stefan Göbel. Vol. 13476. Series Title: Lecture Notes in Computer Science. Cham: Springer International Publishing, 2022, pp. 229–245. ISBN: 978-3-031-15324-2 978-3-031-15325-9. DOI: 10.1007/978-3-031-15325-9_17. URL: https://link.springer.com/10.1007/978-3-031-15325-9_17 (Last accessed on Jan. 3, 2024).


## PUBLICATIONS AS CO-AUTHOR


[5] Polona Caserman, Philipp Achenbach, and Stefan Gobel. "Analysis of Inverse Kinematics Solutions for Full-Body Reconstruction in Virtual Reality." en. In: *2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH)*. Kyoto, Japan: IEEE, Aug. 2019, pp. 1–8. ISBN: 978-1-7281-0300-6. DOI:

10.1109/SeGAH.2019.8882429. URL: https://ieeexplore.ieee.org/document/8882429/ (Last accessed on Jan. 6, 2024).

[6] Philipp Niklas Müller, Philipp Achenbach, André Mihca Kleebe, Jan Ulrich Schmitt, Ute Lehmann, Thomas Tregel, and Stefan Göbel. "Flex Your Muscles: EMG-Based Serious Game Controls." en. In: *Serious Games*. Ed. by Minhua Ma, Bobbie Fletcher, Stefan Göbel, Jannicke Baalsrud Hauge, and Tim Marsh. Vol. 12434. Series Title: Lecture Notes in Computer Science. Cham: Springer International Publishing, 2020, pp. 230–242. ISBN: 978-3-030-61813-1 978-3-030-61814-8. DOI: 10.1007/978-3-030-61814-8_18. URL: https://link.springer.com/10.1007/978-3-030-61814-8_18 (Last accessed on Jan. 6, 2024).

[7] Philipp Niklas Müller, Alexander Josef Müller, Philipp Achenbach, and Stefan Peter Göbel. *IMU-Based Fitness Activity Recognition Using CNNs for Time Series Classification*. preprint. Computer Science and Mathematics, Nov. 2023. DOI: 10.20944/preprints202311.1411.v1. URL: https://www.preprints.org/manuscript/202311.1411/v1 (Last accessed on Jan. 8, 2024).

[8] Philipp Niklas Müller, Felix Rauterberg, Philipp Achenbach, Thomas Tregel, and Stefan Göbel. "Physical Exercise Quality Assessment Using Wearable Sensors." en. In: *Serious Games*. Ed. by Bobbie Fletcher, Minhua Ma, Stefan Göbel, Jannicke Baalsrud Hauge, and Tim Marsh. Vol. 12945. Series Title: Lecture Notes in Computer Science. Cham: Springer International Publishing, 2021, pp. 229–243. ISBN: 978-3-030-88271-6 978-3-030-88272-3. DOI: 10.1007/978-3-030-88272-3_17. URL: https://link.springer.com/10.1007/978-3-030-88272-3_17 (Last accessed on Jan. 6, 2024).

[9] Elisabeth Rotter, Philipp Achenbach, Birgit Ziegler, and Stefan Göbel. "Finding Appropriate Serious Games in Vocational Education and Training: A Conceptual Approach." In: *European Conference on Games Based Learning* 16.1 (Sept. 2022), pp. 473–481. ISSN: 2049-100X, 2049-0992. DOI: 10.34190/ecgbl.16.1.577. URL: https://papers.academic-conferences.org/index.php/ecgbl/article/view/577 (Last accessed on Jan. 6, 2024).

[10] Thomas Tregel, Tim Dutz, Patrick Hock, Philipp Niklas Müller, Philipp Achenbach, and Stefan Göbel. "StreetConqAR: Augmented Reality Anchoring in Pervasive Games." en. In: *Serious Games*. Ed. by Minhua Ma, Bobbie Fletcher, Stefan Göbel, Jannicke Baalsrud Hauge, and Tim Marsh. Vol. 12434. Series Title: Lecture Notes in Computer Science. Cham: Springer International Publishing, 2020, pp. 3–16. ISBN: 978-3-030-61813-1 978-3-030-61814-8. DOI: 10.1007/978-3-030-61814-8_1. URL: http://link.springer.com/10.1007/978-3-030-61814-8_1 (Last accessed on Jan. 6, 2024).

# D

## DECLARATION PURSUANT TO THE DOCTORAL REGULATIONS

*§8 Abs. 1 lit. c PromO*

I hereby confirm that the electronic version of my dissertation matches the written version.

*§8 Abs. 1 lit. d PromO*

I hereby confirm that, at an earlier point in time, no attempt has been made to obtain a doctorate. In this case, further details must be provided regarding the time, the university, the dissertation topic, and the outcome of this attempt.

*§9 Abs. 1 PromO*

I hereby confirm that the present dissertation was written independently and using only the specified sources.

*§9 Abs. 2 PromO*

To date, the work has not been used for examination purposes.

*Frankfurt, January 27, 2026*

Philipp Achenbach