Załącznik nr. 3

Faculty of Biomedical Engineering, Silesian University of Technology

Summary of Doctoral Dissertation

A data glove with reduced number of sensors for recognition of Polish Sign Language letters

Author:

Jakub Piskozub, M.Sc.

Promoter:

Prof. Paweł Strumillo, D.Sc.

Lodz, 2023

1. Introduction

1.1. Purpose of the work

The purpose of this dissertation is to build a sensory glove equipped with a set of sensors designed to study the possibility of its use in rehabilitation and diagnostic applications, as well as recognition of the letters of the Polish Sign Language (PSLA) alphabet. With it, it was decided to study the possibility of reducing the number of sensors in the glove for PSLA letter recognition. The work aims to analyze existing solutions, develop a new, more efficient and ergonomic model of the data glove, and apply advanced machine and deep learning techniques to improve gesture recognition performance.

1.2. Thesis and motivation for undertaking the research

The obvious choices among assistive devices for people with disabilities are technologies that can support or replace a damaged or malfunctioning organ. In the case of the deaf, these are usually hearing aids, cochlear implants or vocoders. However, there are cases in which the use of such a system is impossible, insufficient or unacceptable to the disabled person. With such situations in mind, a number of technologies have been developed to improve the quality of life and safety of their users. These include devices that help visually impaired and blind people with spatial orientation and access to visual information [1]. Technology over the past two decades has also made a significant contribution to improving the quality of life for deafblind people [2]. Thanks to smartphones, such people can sufficiently communicate with others who do not know sign languages, and with the spread of video calls, connect with a sign language interpreter [3]. However, it is not implausible that there is no access to a telephone or Internet connection necessary to make a call to an interpreter on an important matter, or that the sign language speaker does not know any other language. In such a situation, a system based on a data glove that translates sign language into speech can help. The device, has its genesis in cybernetics and has existed in the wider consciousness since it appeared as a game controller. Devices that track hand movements, such as the data glove developed by Sandin et al. [4] makes it feasible to develop such a system. The glove allows direct measurement of finger movements and is designed to not interfere with natural hand movements. This feature is crucial when tracking different gestures and sign languages. As the potential of human-computer interaction continues to be discovered, devices such as the data glove will play a key role in bridging the gap between the physical and virtual worlds, paving the way for more advanced and intuitive control systems.

Sign language is a key communication tool for the deaf and hard of hearing. Data gloves have the potential to facilitate communication between these people and the environment, but their size and the number of sensors may be a barrier to their wider dissemination. The motivation for this research topic is the desire to improve the quality of life for sign language speakers, as well as to accelerate the development of technologies to support such communication.

The most important of the data glove's many uses as a sign language interpreter may be its use in emergency situations. The glove can prove invaluable to emergency services, enabling them to communicate more effectively with deaf or hard-of-hearing people who require assistance. By providing real-time sign language interpretation, lives can be saved through effective emergency response. However, an obstacle to their widespread use is their high cost and unreliability, directly related to their complex design, as cited by their potential users. Reducing the number of sensors used for effective gesture recognition in sign language communication will contribute to solving both of these problems. Thus, the thesis of the following paper can be formulated as follows:

It is possible to effectively recognize the letters of the Polish Sign Language alphabet using a sensory glove with a reduced number of sensors, which can simplify the glove's design and improve its ergonomics.

2. Theoretical and technical fundamentals

One part of PSL is the sign alphabet, also known as the finger alphabet. It is mainly used to articulate proper names, geographical names, technical names, first and last names and other names that do not have their own sign in sign language In practice, the finger alphabet facilitates communication between deaf people and between deaf and hearing people who are not fluent in sign language. PSLA is also used as an educational tool for sign language learners. It helps them memorize gestures and letters and achieve fluency in communication. It is worth noting that different countries have different versions of the finger alphabet that correspond to the letters of the language. The PSL finger alphabet consists of 36 characters: 20 dynamic and 16 static. An interesting feature of PSLA is that diacritical marks such as "Ł" and "Ó" are dynamic equivalents of Latin alphabet characters. The full set of 36 PSLA letters is shown in Figure 2.1.2.

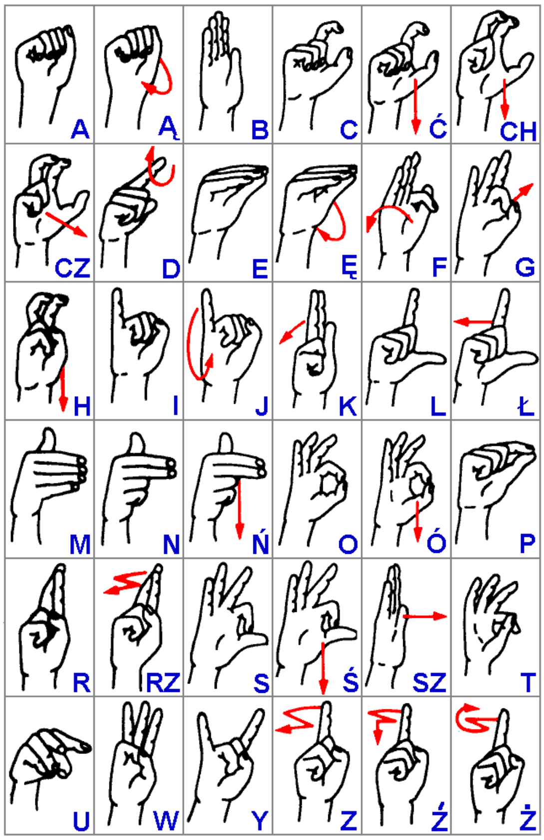


Fig 2.1.2 Alphabet of Polish Sign Language source: http://alfabet-migowy.blogspot.com/2013/04/alfabet-migowy.html

And while hand movements make up the majority in PSL grammar, they do not make up its entirety. Just as important as hand positioning during communication are facial expressions and body language. Thus, without an additional device that tracks and recognizes, in conjunction with glove operation, these other two aspects of PSL, it is impossible to build a complete interpreter. It should therefore be emphasized that the glove alone will never be a stand-alone sign language interpreter.

2.1. Data gloves: related works

W [5] Simoes Dias et al. present the development and analysis of a pattern recognition system for the Brazilian Sign Language (Libras) alphabet, based on a sensor glove. The developed glove consists of five flex sensors, an accelerometer, a gyroscope and two tactile sensors.

Lee et al. present the design and implementation of a portable solution for interpreting American Sign Language words by analyzing finger and hand movement patterns based on movement data from 5 inertial sensors [6]. A recurrent neural network model with an LSTM layer was tuned to achieve the best performance in classifying 27 word-based AJM gestures, with an average accuracy exceeding 99%.

W [7] Pezzouli et al. present the process of translating dynamic characters using a specially designed data glove, called Talking Hands. Various classifiers were tested, in particular random forests and neural networks. The study shows that good results in sign language recognition can be achieved without using an external camera, acquiring all the data from the glove.

The answer to the problem of segmenting a data stream containing gestures is presented in [8]. The paper presents a real-time sign language recognition system that uses a sensory glove based on 10 piezoresistive sensors sprayed directly onto the glove material. The developed gesture recognition system was designed to segment and recognize 17 dynamic gestures in real time, regardless of the number of times the gesture is repeated per measurement.

A study conducted by Saggio et al. [9] describes the use of a glove equipped with a set of 10 piezoresistive sensors placed in fabric pockets and one inertial sensor. The resulting data were classified using a nonparametric model combining a k-NN classifier with a DTW algorithm. The model showed good classification accuracy (96.6%).

W [10] Plawiak et al. described an attempt to analyze 22 hand gestures from generally accepted body language for identification. The experiment used data from a specialized glove (with ten sensors), which was preprocessed and analyzed using machine learning algorithms. The results confirmed that effective and fast recognition of hand body language is possible, and the best classifier achieved a sensitivity of 98.32%.

Dziubich et al. authors [11], conducted a study of the classification performance of PSLA letters using a glove consisting of 5 flex sensors. In their experiments, which included static and dynamic gesture classification tests, the ANN classifier performed worse than SVM. The results for both classifiers can be considered good, with results no worse than 87% and

82%, respectively. Classification of dynamic gestures showed that the performance of DTW and Hidden Markov Models (HMM) methods was much worse, with results of 48% and 53%, respectively. However, the authors did not specify which gestures or letters their study ultimately addressed. All that is known is that 25 different static and dynamic gestures were used.

Currently, to the best of the author's knowledge, the only research on the PSLA letter recognition glove has been conducted by Korzeniewska et al. described in the article [12]. The article describes the process of developing and testing a prototype sign language translation glove. The prototype allows users to enter letters using a mobile device or computer, replacing a keyboard, although the letter identification process is not fully accurate. The authors also focus on the challenges associated with the durability of thin conductive layers and the longevity of textronic sensors such as Velostat, proposing their use for finger bend sensors. Analysis of the electrical properties and selection of substrates for thin-film sensors enabled the creation of a glove that has been used to classify more than 500 character samples with a success rate of 86.5%.

3. Data glove design

The sensors were attached to the surface of the glove using Velcro, which was glued only to the tip of the sensor. At the same time, to ensure the stability of the sensor's movement, 3D printed guides were used, which were also attached to the glove using Velcro. This solution allowed the sensor to move freely, allowing it to bend along with the finger, while keeping it stable, eliminating unwanted movement. What's more, it made it possible to adjust the position of the sensor to the user's hand.

The use of FCL connectors made it possible to securely connect the sensors and quickly replace them when needed. Unlike traditional insulated cables, this version of the glove used a conductive thread sewn into the material. This made it possible to create connections that did not significantly affect sensor movement and were imperceptible to the user. Unfortunately, there was a problem of temporary short circuits between the various signal lines formed by this thread.

Attempts to solve this problem using commercially available isolation formulations that would not restrict the flexibility of the threads failed. However, it was mitigated by reorganizing the routing of the individual lines so that they were as far apart as possible without unnecessary slack, allowing the sensor to move freely.

3.1. Measurement and control unit

In each stage of the project, flex sensors were used in a circuit with a second resistor, creating a voltage divider that was sent to the input of the ADC. ADS1115 converters were used, which are 16-bit, four-channel analog-to-digital converters from Texas Instruments, operating at 860 samples per second. These converters communicated with the glove control unit via the I2C protocol.

Piezoresistive sensor designations use the convention: P[number of the finger on which the sensor is located]_[number of the row in which the sensor is located], for example, P2_1 is the sensor on the second finger in the first row. The block diagram of the latest version, along with an illustrative layout and labeling of various components, is shown in Figure 3.3.1.

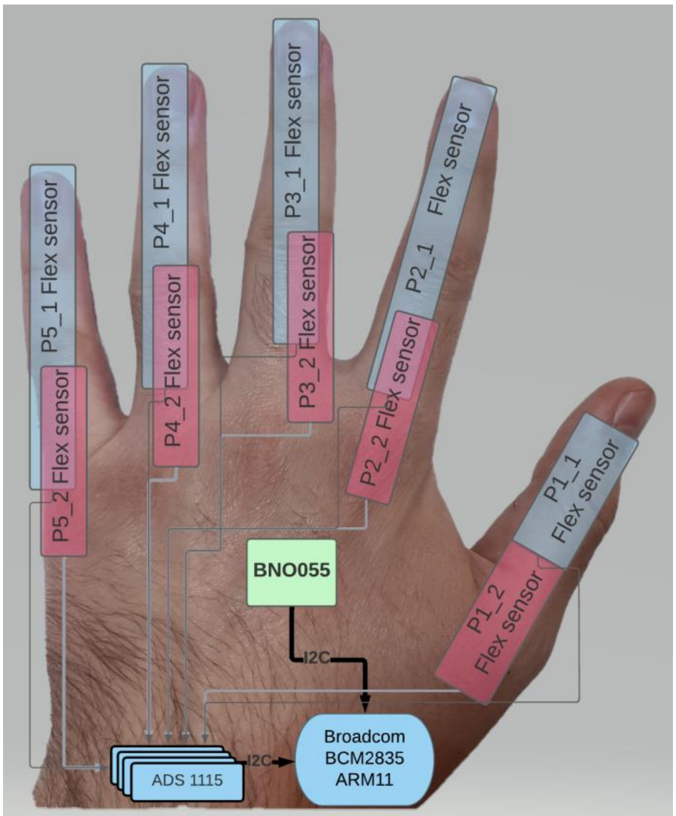


Figure 3.3.1: Block diagram of the device with illustrative component layout

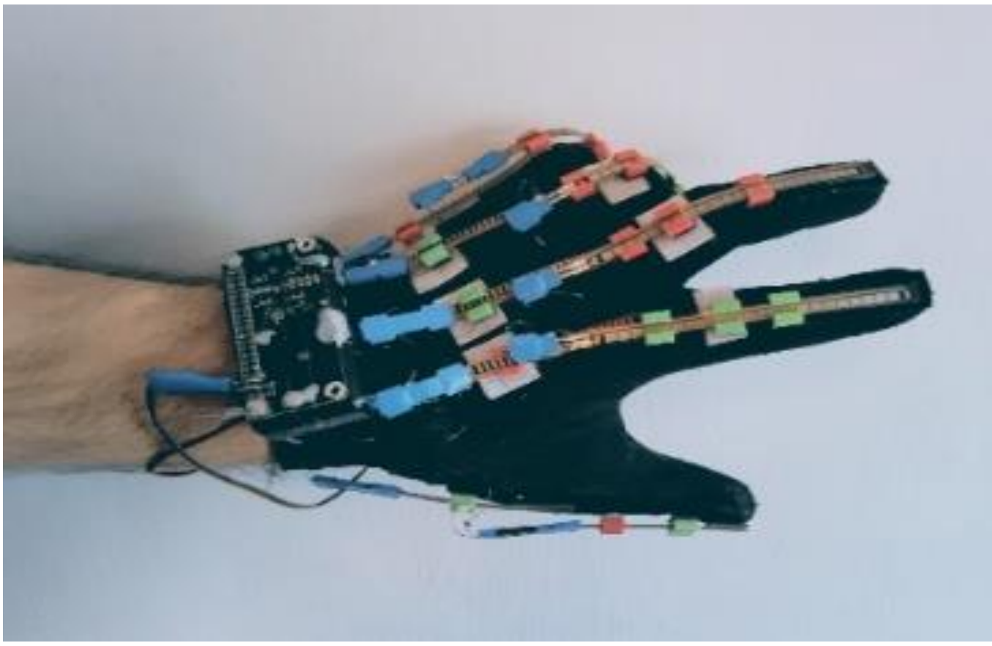


Figure 3.3.2: Photo of the current version of the device

4. Research methodology

The most important stage of the research was the work on PAM recognition based on data from the research set recorded from 16 subjects. The effectiveness of the glove in acquiring data to recognize dynamic hand movements that correspond to PAM letters was evaluated. For this purpose, deep learning techniques were applied, which are discussed later in this section. Tests of selected neural network training data configurations were also conducted to determine the most efficient and effective method for classifying the data.

4.1. Measurement method

The study collected measurements while participants performed gestures corresponding to all PSLA letters, both dynamically and statically. To facilitate data collection, the time taken to perform each gesture was assumed to be 3 seconds, which was determined by empirical observations. This allowed us to collect 75 data samples from each of the 10 piezoresistive sensors and 6 accelerometer axes for each letter of the alphabet. Each test participant repeated each gesture 10 times. As a result of this stage of testing, a total of about 5,800 signals were collected for PSLA gestures, which translates into about 437,000 individual samples. The data glove tests were approved by the Institutional Ethics Committee of the Technical University of Lodz (No. 1/2021, approval date: 28 X 2021).

4.2. Use of machine learning and deep ugo for PSLA letter recognition

The architecture of the network used for data classification was chosen in studies of the effectiveness of different network configurations, including the use of both LSTM and GRU layers, different layer sizes and activation functions. The neural network model built using the sequential API from the Keras library consists of multiple layers and activation functions. The input of the network is given training data.

The first stage of data processing is batch normalization. Then the data passes through a one-dimensional convolution layer (Conv1D) with 64 filters, a kernel of size 3 and a sigmoid activation function. After another batch normalization, the data is routed to gated recursive units (GRUs), which use the "selu" activation function and L2 regularization with a coefficient equal to 0.03. The GRU layer is also equipped with an abandonment mechanism, with DP and RDP values for regular and recursive abandonment, respectively. This process is repeated three times. In the final step, after normalizing the data batch, the data is routed to the Dense layer (Dense) with 36 neurons and a softmax activation function, which is used to generate classification results.

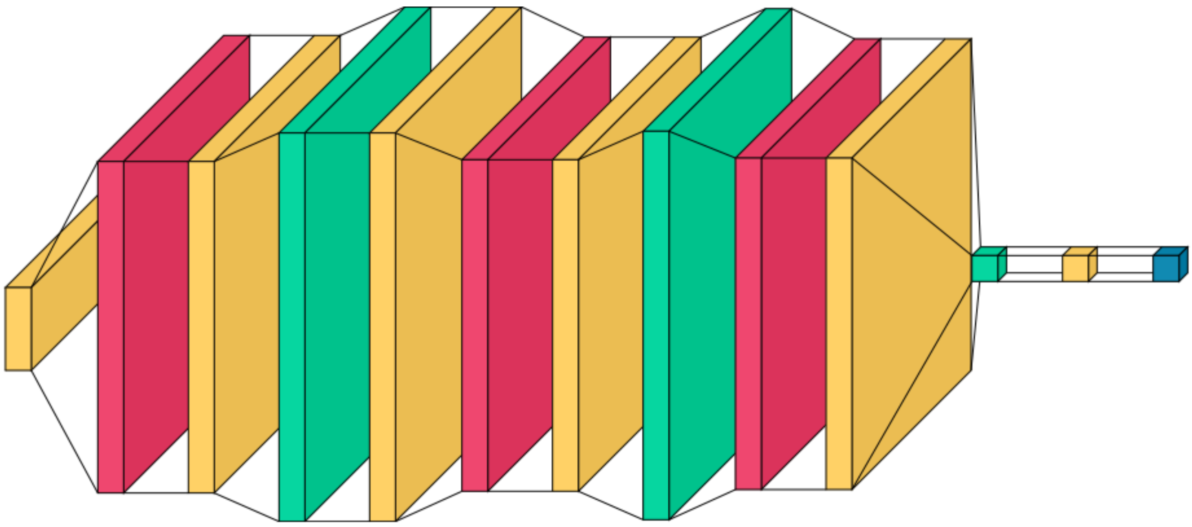


Figure 4.5.1 Neural network architecture (yellow - batch normalization, red - convolutional layer, green - GRU layer, blue - dense layer)

5. Results

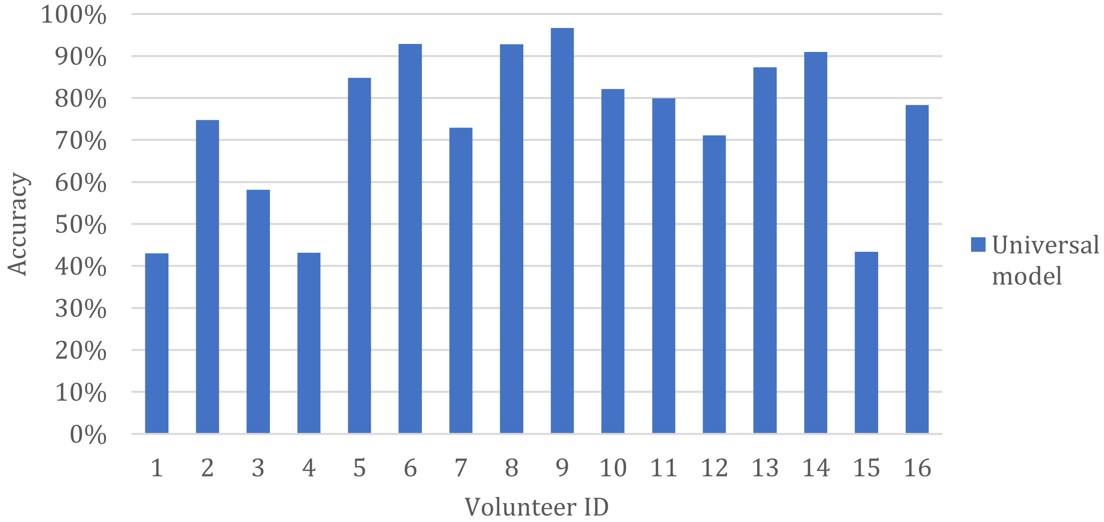
The results of the previous stages of work were described in two publications [13], [14].

The performance of the various classification algorithms was evaluated based on the relevance value expressed by the formula:

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions} \qquad (1)$$

5.1. Classification results of Polish Sign Language letters based on reduced feature vector

The network trained on a set consisting of data from all participants achieved an accuracy of 76%. The network reached its highest efficiency after about 50 epochs lasting about 40 minutes. The result was confirmed by 5-fold cross-validation.



Due to the lack of effective methods for assessing feature validity for recurrent neural networks, especially those designed for classifying time series data, calculations were carried out using the brute-force method. The effectiveness of neural network learning was tested on various data sets, including all possible combinations of three of the ten available sensors, as well as data from an inertial sensor. The results of the accuracy obtained, along with the sets of three sensors on which these results were based, are shown in Table 5.2.1.

Tab 5.2.1 The 5 most effective tercets of features for Polish Sign Language alphabet classification using a neural network

Sensor			Tractability
P1_2	P3_1	P5_1	85,58%
P1_2	P3_1	P4_1	84,92%
P1_2	P2_2	P5_1	84,85%
P1_2	P4_2	P5_1	84,51%
P1_2	P3_1	P5_2	84,51%

5.2. Comparison with related work

Comparison of the previously described concept with others is hampered by the lack of alternatives for reducing the size of gloves that could be used for sign language gesture recognition. W [15], where the authors reviewed papers on data gloves for hand motion recognition, only 4 articles can be found that dealt with a device that worked properly with less than 5 sensors. However, these articles dealt only with basic gestures or movements. They are listed in Table 5.3.1.

Tab 5.3.1 Articles describing devices with less than 5 sensors

Reference	Number and type of sensors	Max. effectiveness
Gupta et al. (2020)	2 IMU	Tracking forearm movement only
Devnath et al. (2019)	2 IMU	Tracking only two fingers
Zhang et al. (2019)	4 pressure sensors	Testing basic gestures
Huang et al. (2019)	2 IMU	Tracking forearm movement only

It is also necessary to compare the effectiveness of the classification algorithms described in this work. Several works on devices used to test and classify synthetically separated dynamic gestures or letters of the sign alphabet using neural networks are summarized along with the number of sensors used in this work in Table 5.3.2.

Tab 5.3.2 Comparison with results obtained in other studies

Reference	Number and type of sensors	Max. effectiveness
[5] Simoes Dias et al. (2022)	5 flex + IMU sensor + 2 contact sensors	96,15%
[6] Lee et al. (2020)	5 IMU sensor	99%
[7] Pezzuoli et al. (2021)	10 flex + IMU sensor	99,7%
[8] Bae et al. (2021)	10 flex + IMU sensor	94%
[9] Saggio et al. (2020)	10 flex + IMU sensor	96,6%
This work	3 flexs + IMU sensor	99%

So, as you can see, the concept described in this paper makes it possible to achieve equally good, if not better, results using a number of sensors even several times smaller than comparable solutions.

6. Discussion and conclusions

6.1. Summary of research results

Among the best results achieved for the set of all letters and for the set with expanded readings from one person, a set of tercets stands out: 1_2, 3_1, 5_1 (marked in blue), as indicated in Table 5.2.1. Among the experimental results for different combinations of the two sensors, a recurring set can also be identified for both data sets: P2_1, P3_2. Even more importantly, the classification performance based on data from only two piezoresistive sensors and a 6-axis inertial sensor yielded classification accuracy of 99% for extended data and 82% for data from all subjects together. This confirms another benefit of using an approach that uses data from only one person to train the network.

The conclusions of the study confirm the main thesis of this work, i.e. the possibility of effective recognition of gestures of the letters of the Polish Sign Language alphabet, using only three sensors. This is particularly important in the context of further development of such gloves as components of sign language translation systems. Ergonomic shortcomings and low performance often form the basis for criticism of this technology. To date, few research teams have studied and tested various sensors, assuming that at least five of them are needed. The results of the present study show that it is possible to significantly reduce the size of gloves used to recognize the letters of the Polish Sign Language alphabet.

6.2. Contribution and impact of the results obtained

To summarize the results obtained:

- It has been proven that it is possible to successfully recognize synthetically segmented Polish Sign Language letters using data from just three piezoresistive sensors with an accuracy of no worse than 90%.

The results of the present work can provide a valuable reference point for researchers and constructors to reduce the size of the glove, which can be a key component of sign language translation systems. The confirmation of the possibility of reducing the size of these gloves by reducing the number of sensors can be considered the most important achievement of the work. In the present study, a neural network was trained using only data from three finger flex sensors and data from a six-axis inertial sensor. The results of this research not only allow not to reduce the efficiency of gesture recognition, but can reduce production costs and improve the ergonomics of the device.

Equally interesting results came from an experiment in which, by extending data from a single person, greater gesture recognition performance was achieved than when classifiers were trained on data from many people. This result suggests that extensive measurements from a large number of people are not necessarily required to build a universal model. A short calibration performed on a single person appears to be sufficient, greatly simplifying the process of developing effective gesture recognition systems

The results of the present work open up the possibility of developing a universal model by using transfer learning. This method involves using previously trained neural network models and adapting them to new data, saving computational resources and computational time. Using data from multiple individuals, the model can be adapted to different styles and idiosyncrasies of hand movements, which can contribute to the development of a more versatile and universal gesture recognition system.

7. Bibliography

[1] "Assistive listening devices / Deaf and Hard of Hearing." Accessed: Jul. 14, 2023 [Online]. Available: https://mn.gov/deaf-hard-of-hearing/assistive-technology/assistive-listening-devices/

[2] "Assistive Devices for People with Hearing or Speech Disorders | NIDCD." Accessed: Jul. 14, 2023 [Online]. Available: https://www.nidcd.nih.gov/health/assistive-devices-people-hearing-voice-speech-or-language-disorders#4

[3] "Migam Interpreter - Migam." Accessed: Jul. 14, 2023 [Online]. Available: https://migam.org/tlumacz-migam/

[4] D. J. Sandin and T. A. DeFanti, "Sayre Glove - US NEA R60-34-163 Final Project Report." pp. 1-30, 1977.

[5] T. Simoes Dias, J. J. A. M. Junior, and S. F. Pichorim, "An Instrumented Glove for Recognition of Brazilian Sign Language Alphabet," *IEEE Sens J*, vol. 22, no. 3, pp. 2518-2529, 2022, doi: 10.1109/JSEN.2021.3136790.

[6] B. G. Lee, T. W. Chong, and W. Y. Chung, "Sensor Fusion of Motion-Based Sign Language Interpretation with Deep Learning," *Sensors 2020, vol. 20, page 6256, vol. 20*, no. 21, p. 6256, Nov. 2020, doi: 10.3390/S20216256.

[7] F. Pezzuoli, D. Corona, and M. L. Corradini, "Recognition and Classification of Dynamic Hand Gestures by a Wearable Data-Glove,". *SN Comput Sci*, vol. 2, no. 1, 2021, doi: 10.1007/s42979-020-00396-5.

[8] M. Lee and J. Bae, "Deep Learning Based Real-Time Recognition of Dynamic Finger Gestures Using a Data Glove," *IEEE Access,* vol. 8, pp. 219923-219933, 2020, doi: 10.1109/ACCESS.2020.3039401.

[9] G. Saggio, P. Cavallo, M. Ricci, V. Errico, J. Zea, and M. E. Benalcázar, "Sign language recognition using wearable electronics: Implementing K-nearest neighbors with dynamic

time warping and convolutional neural network algorithms," *Sensors (Switzerland)*, vol. 20, no. 14, pp. 1-14, 2020, doi: 10.3390/s20143879.

[10] P. Plawiak, T. Sosnicki, M. Niedźwiecki, Z. Tabor, and K. Rzecki, "Hand Body Language Gesture Recognition Based on Signals From Specialized Glove and Machine Learning Algorithms," *IEEE Trans Industr Inform,* vol. 12, no. 3, pp. 1104-1113, 2016, doi: 10.1109/TII.2016.2550528.

[11] T. Dziubich and J. Szymanski, "Automatic classification of Polish sign language words | Automatic classification of Polish sign je{ogonek} sign language,". *Przeglad Elektrotechniczny*, vol. 90, no. 3, 2014, doi: 10.12915/pe.2014.03.43.

[12] E. Korzeniewska, M. Kania, and R. Zawiślak, "Textronic Glove Translating Polish Sign Language,". *Sensors*, vol. 22, no. 18, 2022, doi: 10.3390/s22186788.

[13] J. PISKOZUB, "Recognition of static hand gestures using a sensor glove,". *ELECTROTECHNICAL REVIEW*, vol. 1, no. 9, pp. 144-147, Sep. 2019, doi: 10.15199/48.2019.09.29.

[14] J. Piskozub and P. Strumillo, "Reducing the Number of Sensors in the Data Glove for Recognition of Static Hand Gestures," *Applied Sciences*, vol. 12, no. 15, p. 7388, Jul. 2022, doi: 10.3390/app12157388.

[15] Z. R. Saeed, Z. B. Zainol, B. B. Zaidan, and A. H. Alamoodi, "A Systematic Review on Systems-Based Sensory Gloves for Sign Language Pattern Recognition: An Update from 2017 to 2022," *IEEE Access*, pp. 1-1, 2022, doi: 10.1109/ACCESS.2022.3219430.