

Article

Data Glove with Bending Sensor and Inertial Sensor Based on Weighted DTW Fusion for Sign Language Recognition

Chenghong Lu , Shingo Amino and Lei Jing * 

Graduate School of Computer Science and Engineering, University of Aizu, Tsuruga, Ikki-machi, Aizuwakamatsu City 965-8580, Japan

* Correspondence: leijing@u-aizu.ac.jp

Abstract: There are numerous communication barriers between people with and without hearing impairments. Writing and sign language are the most common modes of communication. However, written communication takes a long time. Furthermore, because sign language is difficult to learn, few people understand it. It is difficult to communicate between hearing-impaired people and hearing people because of these issues. In this research, we built the Sign-Glove system to recognize sign language, a device that combines a bend sensor and WonderSense (an inertial sensor node). The bending sensor was used to recognize the hand shape, and WonderSense was used to recognize the hand motion. The system collects a more comprehensive sign language feature. Following that, we built a weighted DTW fusion multi-sensor. This algorithm helps us to combine the shape and movement of the hand to recognize sign language. The weight assignment takes into account the feature contributions of the sensors to further improve the recognition rate. In addition, a set of interfaces was created to display the meaning of sign language words. The experiment chose twenty sign language words that are essential for hearing-impaired people in critical situations. The accuracy and recognition rate of the system were also assessed.

Keywords: data glove; wearable device; sign language recognition; ubiquitous computing



Citation: Lu, C.; Amino, S.; Jing, L. Data Glove with Bending Sensor and Inertial Sensor Based on Weighted DTW Fusion for Sign Language Recognition. *Electronics* **2023**, *12*, 613. <https://doi.org/10.3390/electronics12030613>

Academic Editor: Paolo Visconti

Received: 11 December 2022

Revised: 23 January 2023

Accepted: 24 January 2023

Published: 26 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In Japan, there are about 341,000 hearing impaired people [1]. The general way to communicate between a hearing person and a hearing-impaired person is communication by writing or sign language. However, communication by writing takes a lot of time. Furthermore, sign language, which hearing impaired people use, is not always familiar to hearing people or those who acquired a hearing impairment. Each of the two approaches has problems that hinder smooth communication in society.

Sign language recognition has always been a research problem that has received a lot of attention. There has been a large number of studies on sign language recognition in recent years [2–6].

Sign language recognition systems can be divided into non-wearable and wearable approaches. Non-wearable generally include vision-based [7,8], while non-wearable use WiFi signal-based [9,10] methods. Another approach is to recognize sign language with wearable sensor-based data gloves [11,12].

Due to the development of deep learning methods in visual sign language recognition, the recognition rate has been improved. However, deep learning is driven by data, and the quality of data collection greatly affects the results. Insufficient video frames and occlusions will also reduce the recognition accuracy. Gerges et al. [13] established dynamic hand recognition based on MediaPipe's Landmarks and compared the recognition accuracy of three deep learning methods: gated recurrent unit (GRU), long short term memory (LSTM), and bi-directional LSTM (BiLSTM). Dataset collection requires complete characters, no occlusions, and a fixed duration. It is difficult to achieve these requirements in actual use.

Chang et al. [14] conducted research on recognizing sign language by detecting the places of nails and wrists using pictures of the hand. It recognizes sign language by a skeleton of the hand and the distribution of skin color from pictures taken of the hand. However, the systems that hearing-impaired people need to use in their daily lives can detect not only the hand shape part of sign language, but also the dynamic part of hand movement in sign language. In other vision-based methods, one approach uses color gloves and Kinect data stored by Microsoft. Shibata et al. [15] used color gloves for recognizing sign language. The color glove has every color at every finger and wrist. Furthermore, it operates by moving the distance and area of glove colors. However, in the detection step, if the background or the user's clothing are the same color as the part of the glove, it cannot be recognized in this way. Kinect can detect hand motions and hand places. Muaaz et al. [16] developed a system that can recognize American sign language with Kinect. This system has a high average recognition rate of 80%. Furthermore, this system can make easy sentences by recognizing sign language words. However, this system is also limited by the camera, and we can only use this system in limited positions without occlusion. In daily life, it can be a large barrier for the hearing impaired to use this system.

Vision-based sign language recognition is limited by the nature of camera view observation and is not good at capturing complex two-handed interaction movements because of occlusion. It is also susceptible to the influence of the environment in-between the camera and the object. The method of wearable sensors and data gloves forces users to accept some burdens. However, data gloves can collect data steadily in complex environments, without the problem of line of sight obstruction, noisy backgrounds, and inadequate light. It can even be used outdoors, and in low visibility. The camera method is subject to a variety of environmental constraints. Therefore, we plan to use wearable devices to capture the complex motion of the fingers.

In recent years, wearable sensor-based data gloves have been developed with continuous improvements in processing information technology and the miniaturization and high functionality of equipment. Wearable sensor-based data gloves have been able to operate a large amount of information and handle more complex processing.

Common wearable sensor data gloves for sign language recognition include flexible sensors [17], inertial measurement units (IMUs) [18], surface electromyography (sEMG) devices [19,20], and touch sensors [21]. EMG data have large individual variation. When using the bilinear model for classification, a new subject needs to perform at least one motion. Furthermore, the recognition rate will drop significantly without using a bilinear model.

The information directly related to the hand in sign language includes 21 degrees of freedom of the joints on the hand, and the spatial displacement and orientation of the hand. It is difficult to obtain the appropriate characteristics of complicated information through a single type of sensor. Korzeniewska et al. [17] chose Velostat to make bending sensors to collect data to identify Polish sign language and obtained a letter recognition rate of 86.5%. However, sign language generally uses words as the unit of recognition. Youngmin Na et al. [18] installed an accelerometer on the index finger to recognize static letter gestures in the Korean sign language alphabet, but sign language contains a lot of dynamic gestures, and only static gesture recognition was not enough. Jakub et al. [22] collected IMU sensor data installed on the palm and fingertips and used parallel hidden Markov model (HMM) approaches for sign language recognition. The finger shape data could be obtained by combining the IMU data on the fingertips and the IMU data on the palm. For collecting hand shape features, multiple inertial sensors are more expensive than multiple bending sensors.

Data gloves from a single type of sensor either collect much missing hand information or have a high price for implementation. Thus, multi-sensor fusion is a better solution. The use of wearable sensors and data gloves is moving toward practical applications as advanced MEMS technology sensors are being miniaturized. It also breaks down the spatial limitations of the hand, making multi-sensor data collection possible. Furthermore, among the multiple combinations, inertial sensors to collect hand motion and bending sensors

to collect hand shape are the common approaches [23–26]. Faisal et al. [23], using the K-nearest neighbor (KNN), classified 14 static and 3 dynamic gestures for sign language recognition. Faisal et al. [24] collected data from 25 subjects for 24 static and 16 dynamic American sign language gestures for validating the system. Boon Giin Lee et al. [25] used the support vector machine (SVM) to classify American sign language.

The combination of inertial sensors and bending sensors helps us to obtain hand shape and motion information at low cost. However, how to rationalize multiple sensor data for sign language recognition is still a difficult problem. The execution length of the actions of sign language varies greatly due to people's habits or usage scenarios. The dynamic time warping (DTW) algorithm is a solution to compare the similarity between time series data of different lengths. However, the current research on the application of the DTW algorithm to sign language recognition is insufficient. Chu et al. [26] studied DTW for sign language recognition on seven Japanese sign language datasets, and validation was performed using the leave one out (LOO) approach, with recognition rates of 82.5%. First of all, seven recognition actions were insufficient. On the other hand, the variation between different sensors was significant in providing useful information for sign language recognition. Thus, it is necessary to propose weighted DTW. As shown in Table 1, we compared the studies of various sensors. Portability in the table refers to whether good results can be obtained without any data from new users.

Table 1. Comparison of related research (KNN: K-nearest neighbor).

Research	Sensor	Accuracy	Subject	Kinds	Portability	Algorithm	Dynamic Motion
Muaaz et al. [16]	Kinect	95.6%	5	10	○	DTW	○
Tateno et al. [19]	EMG	97.7%	20	20	×	LSTM	○
Lee et al. [21]	Touch	92%	-	36	○	Tree	×
Faisal et al. [23]	Inertial and Flex	64%	35	3	○	KNN	○
Chu et al. [26]	Inertial and FlexForce	82.5%	3	7	○	DTW	○
Ours	Inertial and Flex	85.21%	8	20	○	weighted DTW	○

In this study, inertial sensors and bending sensors were deployed simultaneously in the hand space to collect hand shape and motion features. This method is a practical and promising solution to combine these two parts of features to recognize sign language. Thus, in this research, the Sign-Glove system was implemented, as shown in Figure 1. The development of such systems will give us a future where we wear sensors such as accessories that make it easier to communicate between a hearing person and a hearing-impaired person. When we developed the system to recognize sign language on portable devices with recent technology, for the recognition algorithm of sign language, we extended DTW to use on time series of multiple sensors. DTW is a general method for measuring the similarity between two temporal sequences. However, for data from multiple sensors, different sensors provide different recognition contributions. Thus, we proposed weighted DTW, an algorithm that improves the recognition rate by setting weighted values to raise the effect of key sensors. The contribution of the paper is as follows:

We developed a low-cost Sign-Glove system combining bending sensors and IMU for supporting communication in the context of hearing impairment. To determine hand shape and hand motion, the device supports the simultaneous collection of bending sensors and inertial sensors. We built the weighted DTW algorithm to implement multi-sensor fusion for sign language recognition. The algorithm does not limit the input data length. In addition, weights were assigned based on the contribution of sign language recognition. The contribution was analyzed based on the differences in the features of the sensors and the measurement locations of the hands. Assigning weights enhanced the influence of key sensors and reduced the errors caused by noisy data, effectively improving the recognition rate.

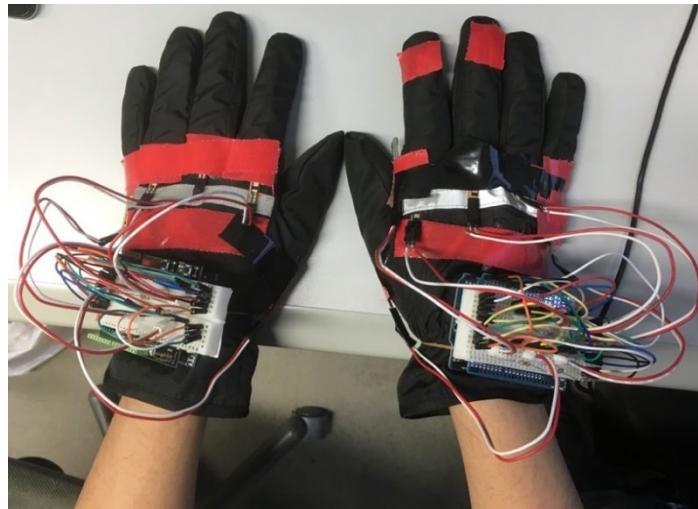


Figure 1. The Sign-Glove on hands.

The organization of this paper is as follows. The application model and sign language dataset are presented in Section 2. Our system design, implementation, and algorithms in this study are given in Section 3. The experimental setup and results evaluation are in Section 4. The discussion of the system is presented in Section 5. Finally, conclusions are given in Section 6.

2. Application Model and Sign Languages Datasets

2.1. Application Model

The system presented in this research shows the meaning of a sign language word on a PC for supporting communication between a healthy person and a hearing-impaired person. This system supposed that a user uses a pair of Sign-Gloves and shows the meaning of a sign language word on the PC.

A user wears Sign-Gloves on his/her hands. Furthermore, the user moves the motion of a sign language word. Then, the PC shows the mean of the sign language. A Sign-Glove is a glove-shaped device with a WonderSense device and a bending sensor. WonderSense is the device developed in this laboratory. This model supposed that the user wants to communicate a sign language word motion to another person. We explain the process of this system in Figure 2. A user moves the motion of the sign language word that he/she wants to communicate to another person. Sign-Glove measures the acceleration of hand motion and hand shape at this time. the WonderSense device of the Sign-Glove transmits the measured hand acceleration to WonderBox with Bluetooth low energy. WonderBox is the receiver device of WonderSense. WonderBox sends measured hand acceleration data to a PC with a serial connection. At the same time, the bending sensors of Sign-Glove measure the hand shape. Furthermore, an Arduino board sends measured data with a serial connection. An Arduino is an AVR Micon board and is used for taking data from the bending sensors and sending data to the PC. After the finished sign language gesture, the data values sent by sensors are computed to recognize a sign language word motion. This sign language word motion is converted into a message associated with the sign language word motion in the PC. Finally, the PC displays the message requested by the user. At this time, if the message is a serious one, the PC makes a sound.

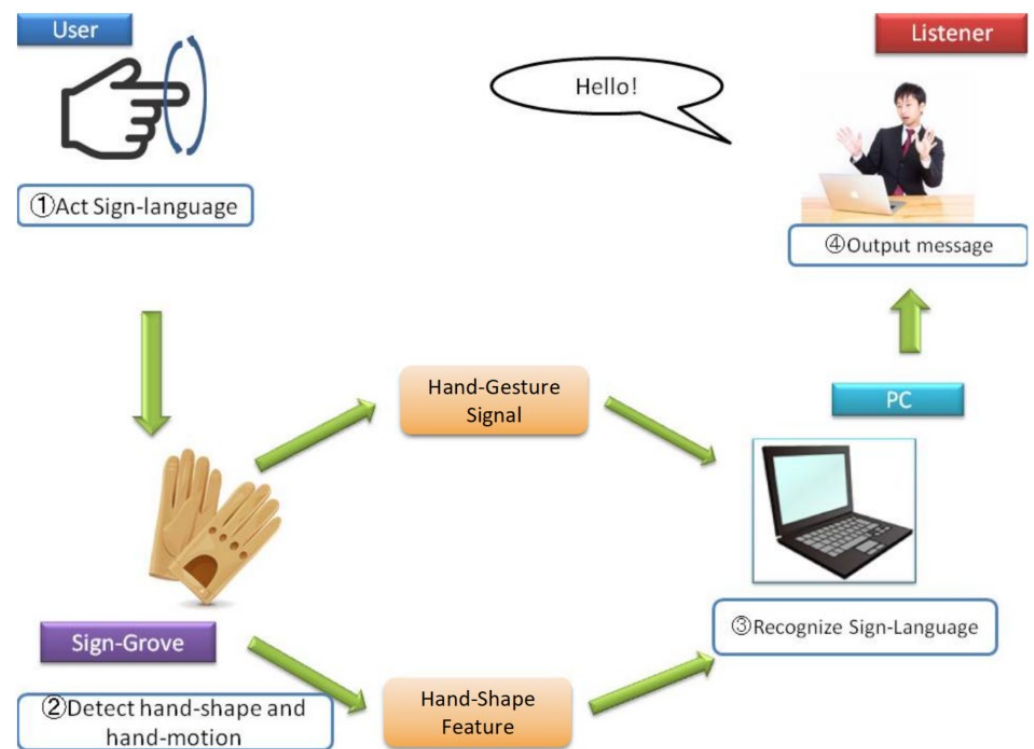


Figure 2. Usage of Sign-Glove for sign language recognition.

2.2. Sign Languages Datasets

2.2.1. Characteristics of Sign Language Data

Sign language consists of two main components in the hand part, namely, the shape of the hand and the overall movement of the hand. Static sign language is defined as a special case of dynamic sign language, which specifically means that the shape of the hand and the hand motion remain unchanged for a period of time. Figure 3 shows the hand shape parts and the hand motion parts of sign language.

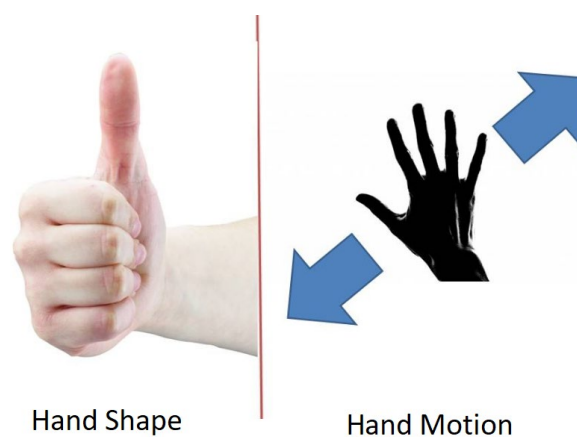


Figure 3. Hand shape and hand motion factors for sign languages.

2.2.2. Sign Language Dataset Definition

The key point to recognizing sign language is to recognize the hand shape and the hand motion at the same time. Missing one of them will significantly reduce the recognition rate, such as “please” and “good”, “sick” and “obstacle”, “down” and “I see”, as shown in Figure 4, because the hand motion of these sign languages is the same but the shape of the hands is different. If we detect only the hand motion of these sign language words,

the result is that this sign language is completely the same. In contrast, Sign-Glove used in this research can detect hand shape. Thus, we can increase the recognition rate of sign language words. Furthermore, for the same reason, we can also achieve the correct result of recognizing sign language words, which is the same hand shape and different hand motions.



Figure 4. Selection of sign language vocabularies with both hand shape and hand motion factors.

3. Methods

The system architecture is shown in Figure 5. The data glove collecting the physical features, and the communication structure are shown in Figure 6. We explain the design of the system in Section 4.1. We explain the recognition algorithm in Section 4.2.

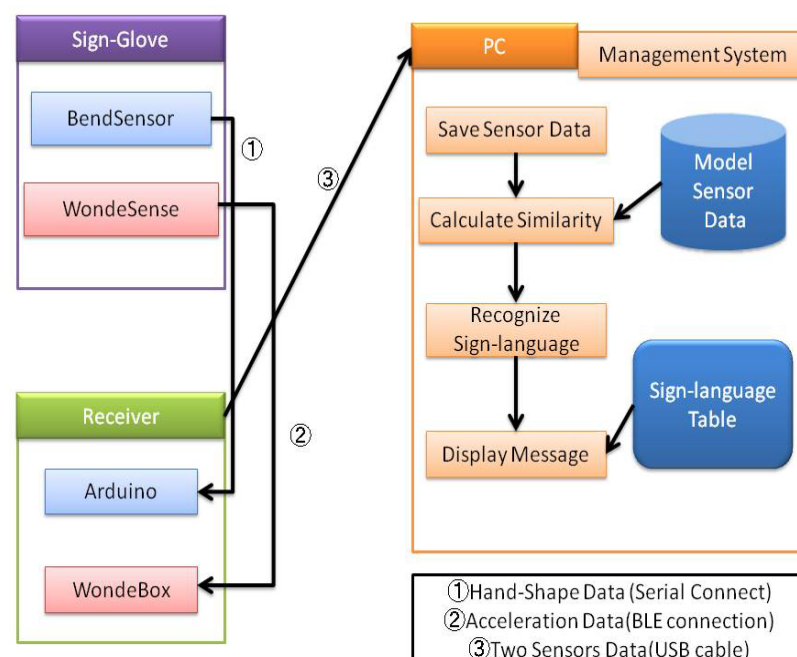


Figure 5. Architecture of the system.

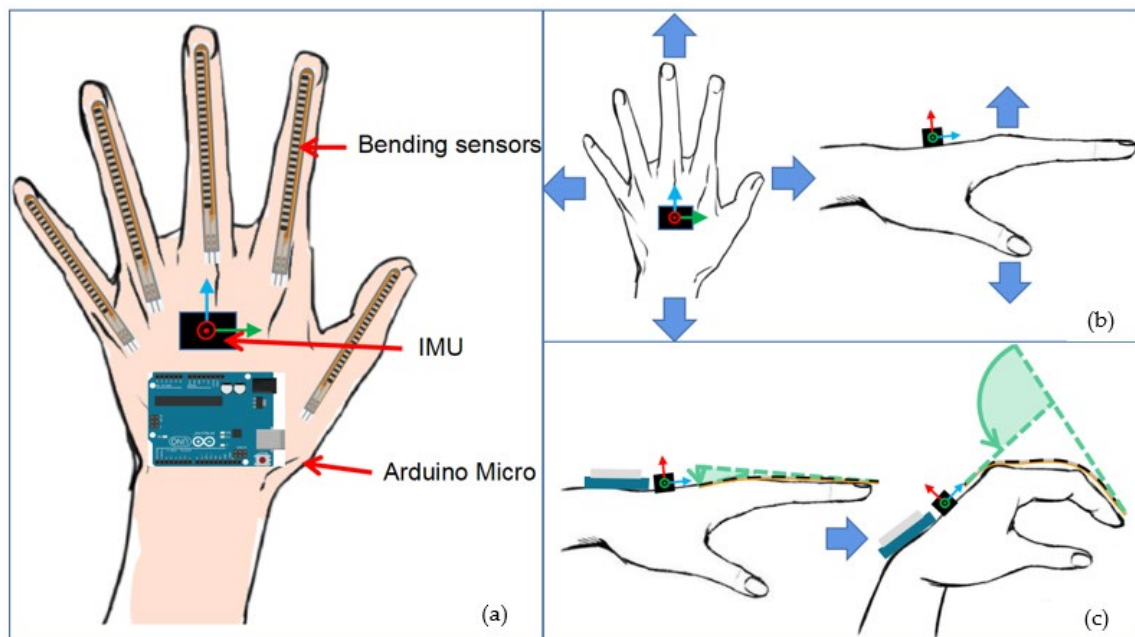


Figure 6. (a) System structure diagram; (b) IMU collecting the hand motion data; (c) bending sensor collecting the hand shape data.

3.1. System Design

In this section, we explain how a sign language word is recognized by the system constructed in this research and the way to detect it. First of all, we explain the operation of this system. This research system recognizes a sign language word based on bending condition data of fingers detected by a bending sensor and acceleration data of hands detected by WonderSense.

First, the Sign-Glove worn on a user's hands takes bending fingers condition data and acceleration data of hands from the bending sensor and the 3-axis acceleration sensor. The bending fingers condition data are sent from the Arduino to a PC with a serial connection via the USB cable. At the same time, the acceleration data of hands detected by WonderSense are sent to WonderBox, which is a data receiver with a BLE connection. Furthermore, WonderBox sends the data to a PC with a serial connection. Until the end of a sign language word motion, Sign-Glove continues to receive the values. Acceleration and bending finger state data are stored in the computer. After that, the model data are read.

The model data are defined with the 3-axis acceleration data and bending data of the fingers for all sign language words. For the recognition process, we use the weighted DTW algorithm to calculate the similarity between the sign language data to be recognized and the model data. The highest similarity of the model data is selected. Finally, the meaning of the sign language word is extracted from the sign language table and displayed on the screen.

As shown in Figure 6, the data glove collects the physical features of the hand. The IMU collects the motion features of the hand, as shown in Figure 6b. The bending sensor collects the shape features of the hand, as shown in Figure 6c. Both sensors are stitched to the cloth glove at the corresponding locations for fixation. The bending sensor is fixed in a special way. When the finger is bending, the skin is stretched. However, the length of the bend sensor is fixed, so we fix the top of the bend sensor to the fingertip position of the glove, and the middle of the sensor is restricted by the wire without shifting from left to right. Furthermore, the back end of the sensor is not fixed but is free to stretch, only allowing the sensor and the back of the hand to move as far as possible to fit.

3.2. Implementation

3.2.1. Hardware

We explain the construction of the hardware in this research. Sign-Glove is the device that takes the acceleration of the hand gesture and hand shape. Sign-Glove has two kinds of sensors. One of the two sensors is the bending sensor. Figure 7a shows the bending sensor. The bending sensor changes its resistance by the bending condition. The bending sensor has a polymer ink on one side of the sensor, which has about 30 K Ω of resistance when straight and increases when it is bent. The Arduino board detects the change as a voltage change value and sends it to the PC with a USB cable. The Arduino is an AVR Micon board. We use it as a bending sensor receiver. The other sensor is WonderSense. Figure 7b shows WonderSense.

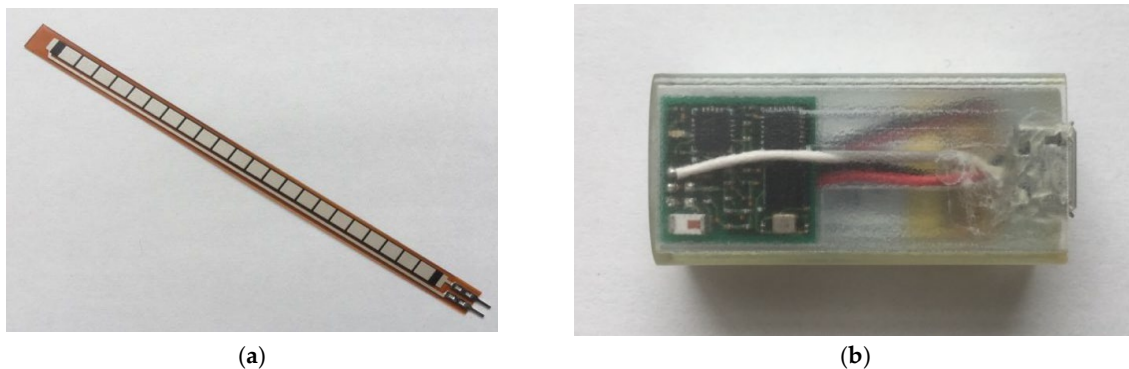


Figure 7. Sensors and data collection devices: (a) bending sensor; (b) WonderSense.

WonderSense collects acceleration data using a 9-axis inertial sensor module MPU9250. WonderBox is the data receiver of WonderSense. The core chip of the WonderBox is the PCA10040 for Bluetooth data reception. WonderBox sends data to a PC with a USB cable. A Sign-Glove device is a pair of gloves. Sign-Glove is constructed by ten bending sensors, two Arduino boards, two WonderSense sensors, and one WonderBox device. To facilitate synchronization, we sampled both the inertial and bending sensor data sampling rate to 50 Hz.

3.2.2. Software

In this research, the environment of the PC for recognizing sign language was structured by the Java language written by the integrated development environment of Eclipse in the OS of the Mac OS X 64-bit system. The DB system is MySQL of the XAMMP application. Furthermore, we used WonderTerminal to manage the data of WonderBox taken by WonderSense. WonderTerminal is software designed for controlling WonderBox and WonderSense.

3.2.3. User Interface

Figure 8 shows the user interface of our system. We explain here every function of our user interface. ① is the start button. When we click this button, recognition of sign language starts. Furthermore, when we click again, recognition of sign language stops, and ⑨ shows a recognized sign language word. ② is the save button. When the save button is clicked, sign language data that the system now contains is saved to the database. The destination database in which to save is decided by database selectors. Furthermore, what kinds of sign language words to save is decided by a sign language selector. ⑦ comprises the extract buttons. The extract AccData button is the button that reads the acceleration data of a database. The extract BendData button is the button that reads the bending data of a database. At this time, the user interface decides which database to read from and what kind of sign language word by ③. In addition, the user interface decides to read how many

times data are in the data base by ⑤, ID selector. ⑥ is the Insert ModelData button. The Insert Model Data button has a function that inserts the sign language word ModelData to DataBase. Finally, ⑧ is the research button.

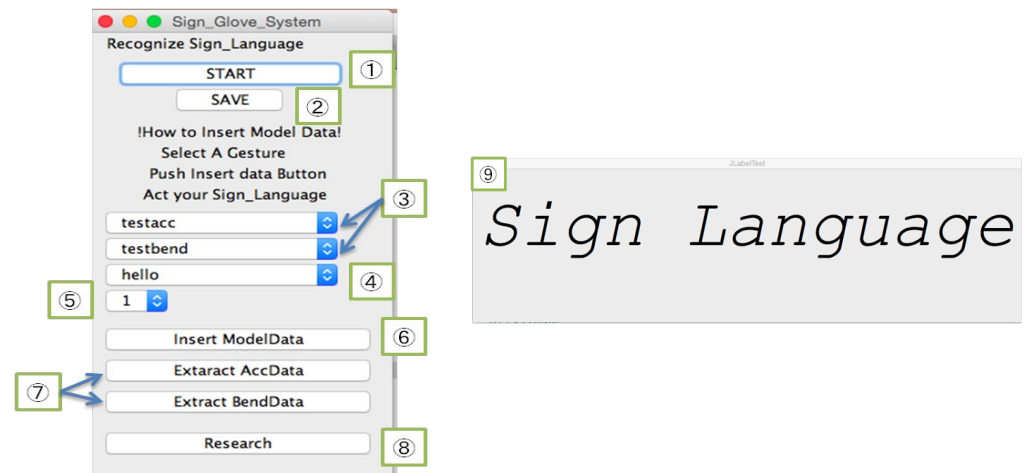


Figure 8. System user interface.

The Research button is the button that calculates the similarity between sign language data that contain current and model data.

3.2.4. Data Format

We explain here data taken from WonderSense. As above, the data taken from WonderSense are acceleration data. The acceleration data taken from WonderSense are sent to WonderTerminal on a PC through WonderBox. WonderTerminal has a function that builds a server. The server of WonderTerminal sends the acceleration data to our Sign system. The acceleration data format is String and the frequency is 50 Hz. We use the acceleration data for recognition. Our research system can save the acceleration data from WonderSense into DataBase.

We explain data here taken from bending sensors. As above, the data taken from bending sensors are the resistance values of finger bending. The resistance data taken from every bending sensor are sent to our research system in a PC through an Arduino. In this system, we can save the data from bending sensors into a database.

3.3. Recognition Method

3.3.1. Dynamic Time Warping

The dynamic time warping (DTW) algorithm is used for measuring waveform similarity. The DTW algorithm calculates the similarity of time-series data using Euclidean distance. The feature of the DTW algorithm is that the length of sample data does not become a problem for the calculation. The duration of sign language varies while expressing the same word according to habit, proficiency, and other factors. Even in this situation, the DTW algorithm can calculate similarity. Next, we explain how to calculate the DTW algorithm for a single sensor.

1. To calculate the similarity of a sequence $X = x_1, x_2, \dots, x_M$ $M \in \mathbf{N}$ and sequence $Y = y_1, y_2, \dots, y_N$ $N \in \mathbf{N}$, make similarity arrays $D(i, j)$ of size $M \times N$.

$$D(i, j) (i = 1, 2, \dots, M) (j = 1, 2, \dots, N) \quad (1)$$

2. Assign 0 to $D(0, 0)$ and ∞ to the others.

$$\begin{aligned} D(0, 0) &= 0 \\ D(i, j) &= \infty (i = 1, 2, \dots, M) (j = 1, 2, \dots, N) \end{aligned} \quad (2)$$

3. Calculate the similarity of two time-series data, $D(M, N)$, with calculation (3) for $i = 1, 2, \dots, M$ and $j = 1, 2, \dots, N$. $f(x_i, y_j)$ is the cost function.

$$D(i, j) = f(x_i, y_j) + \min(D_{i,j-1}, D_{i-1,j}, D_{i-1,j-1})$$

$$f(x_i, y_j) = \sqrt{(x_i^2 - y_j^2)} \quad (3)$$

4. The DTW distance we need is the result of calculation in all combinations $D(M, N)$.

$$D(M, N) \quad (4)$$

3.3.2. Weighted DTW

DTW can calculate model data and sensor data similarity for a single sensor. Furthermore, by assigning weights, the weighted DTW can effectively fuse data from multiple sensors. The model data are the ideal data generated by analyzing the average value of the standard action and the waveform trend of each sensor for multiple executions.

The contribution of each sensor to sign language recognition is different. In this research, we used both bending sensors and inertial sensors. Furthermore, two inertial sensors measure the movement of two hands, and ten bending sensors measure the bending of ten fingers. On the one hand, the types of sensors are different, so the effectiveness of information is different. On the other hand, even with the same sensor, for sign language recognition, the thumb, index finger, and middle finger of the right hand provide more critical information in many cases, while the other fingers most of the time make little contribution to distinguishing sign language. Due to a large number of static states, the waveform has less effective information and is more affected by noise. Thus, setting the same weight is unreasonable. We set different weights between 10 bending sensors, different weights between 2 inertial sensors, and different weights between 2 types of sensors. The weights calculation process is as follows:

1. Combine 10 bending sensors added weight β_i of 10 fingers.

$$DTW(B) = \sum_{i=1}^{10} \beta_i B_i \left(\sum_{i=1}^{10} \beta_i = 1 \right) \quad (5)$$

2. Combine 2 WonderSense added weight γ_j of both hands.

$$DTW(WS) = \sum_{j=1}^2 \gamma_j WS_j \left(\sum_{j=1}^2 \gamma_j = 1 \right) \quad (6)$$

3. Combine (1) data and (2) data added weight α .

$$DTW = \alpha DTW(WS) + (1 - \alpha) DTW(B) \quad (7)$$

4. Experiment and Evaluation

In the experiment, we evaluated the performance of the sign language data glove. In this section, we first describe the experimental setup. Next, we show how the experiments compared the recognition performance of hand shape, hand motion, and combined data of both. After that, we verify the recognition performance of our weighted DTW.

4.1. Experimental Setting

We recruited 8 volunteers and collected data on 20 sign language words. The average age of subjects was 22. Each person repeated each sign language three times, and we collected a total of $8 \times 20 \times 3$ data points. Model data was the average value of multiple executions of the standard action. Table 2 shows the weight parameters in the experiment. Next, we introduced the usage of Sign-Glove.

Table 2. Setting of the weight parameters in the experiment.

Weight	Value	Weight	Value
α	0.05	β_6	0.0002
β_1	0.2448	β_7	0.0002
β_2	0.3772	β_8	0.0002
β_3	0.3776	β_9	0.0002
β_4	0.0002	β_{10}	0.0002
β_5	0.0002	γ	0.5

Usage of the Sign-Glove

In this section, we explain how to wear the Sign-Glove and the starting position of recognizing a sign language word, as shown in Figure 4, with this system. First, the Sign-Glove is worn on the hand. The fingers should be inserted into the Sign-Glove because the Sign-Glove has a bending sensor in each finger part. Figure 6a shows the correct wearing of Sign-Gloves.

Figure 9a,b show a pose and hand position when we start to recognize a sign language word. The basic position is sitting in a chair and putting the hands on the knees. Recognizing a sign language word must begin from this basic position. In this research, we started the system during recognition. Furthermore, when a sign language word motion was finished, we stopped the system. Then, the hands were returned to the basic position.

**Figure 9.** The basic pose and hand position during the usage of the system: (a) side view; (b) front view.

4.2. Experiment Results

4.2.1. Comparison between the Hand Shape, Hand Motion, and Combination Methods

As shown in Figure 10, the recognition ratio of twenty kinds of sign languages in this experiment was obtained. Experiments were performed to calculate the recognition rate of sign language for three kinds of feature data: combined hand motion data and hand shape data, data only based on hand motion, and data only based on hand shape. The motion data of the hand originates from the inertial sensor, which is shown as a red rectangle on the graph. The shape of the hand originates from the bending sensor, which is shown as a blue rectangle on the graph.

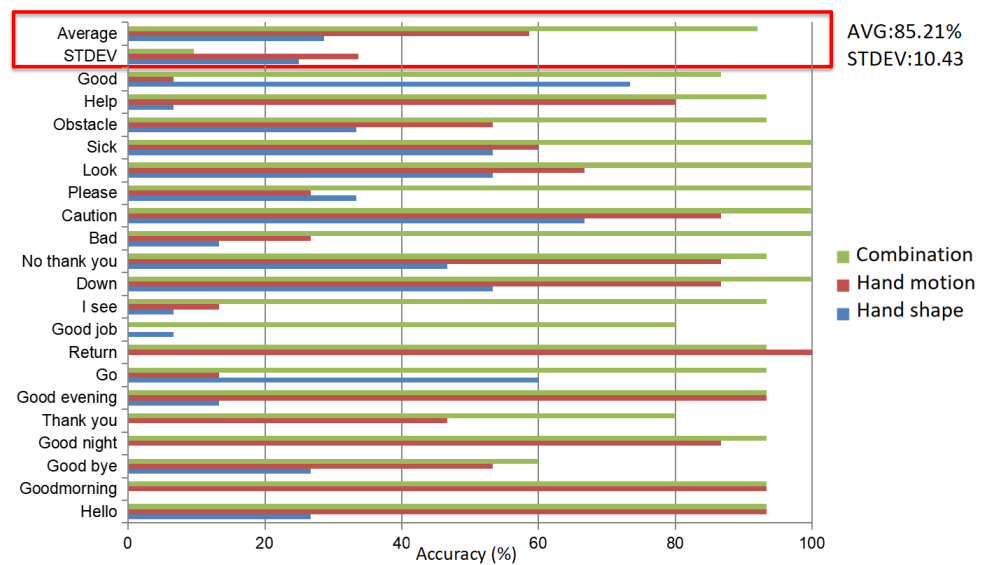


Figure 10. Comparison between the hand shape, hand motion, and combination methods. (The red boxes on the right give the values of AVG and STDEV for combination method.).

The combined sign language recognition rate was the highest, with an average recognition rate of 85.21% and a standard deviation of 10.43. The next highest recognition rate was using only the hand motion features, and the lowest recognition rate was using only the hand shape features for sign language recognition. Depending on the features of different sign languages, the contribution of hand motion data and hand shape data were different. In this dataset, hand motion features contributed more to the recognition rate.

The hand motion and hand shape parts of the data are complementary most of the time. Except for the word “return”, where the combined features are not as good as just the hand motion part of the data, the hand shape part of the data plays a reverse role.

4.2.2. Comparison between Using Our Proposed Weighted DTW or Not

We obtained result of cases in which we used different weights or the same weight as data fusion, as shown in Figure 11.

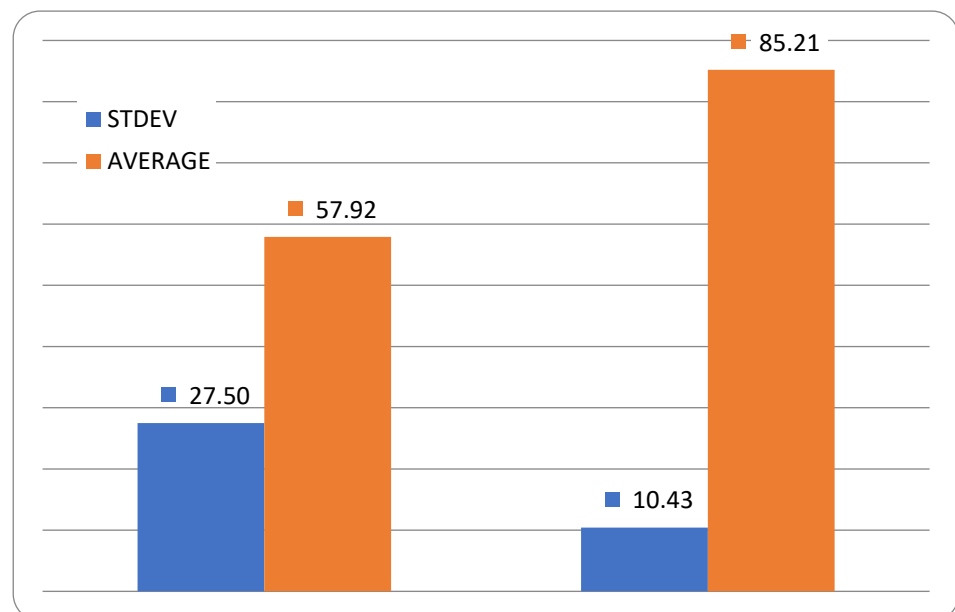


Figure 11. Comparison between use our proposed weighted DTW and the original DTW.

When we used our weight by data fusion algorithm, we could take the average recognition rate of 85.21%, and the standard deviation was about 10.43. When we did not use our weight by data fusion algorithm, the average recognition ratio was about 57.92%, and the standard deviation was 27.50. Thus, we understand that this data fusion algorithm was increasing the recognition ratio and decreasing the standard deviation. Thus, we can say that this algorithm is useful for recognizing sign language.

5. Discussion

We built a data glove based on bending sensors and inertial sensors to capture hand shape and motion features, and then it uses weighted DTW fusion features to recognize sign language. We experimentally verified that both hand shape and hand motion contribute to sign language recognition. Moreover, the two features are complementary, and a higher recognition rate can be obtained by fusing the two features to recognize sign language. Adjusting the weight values to fuse the features, we found that the quality of information provided by sensors with different placements is different. By adjusting the weights to focus on the sensors with large value changes during the execution of sign language, the recognition accuracy can be improved. We collected data for 20 dynamic sign language words from eight volunteers, and the recognition accuracy was 85.21%. The feasibility of the system was verified.

In comparison with similar systems, although there has been a large number of studies on sign language recognition, the defined sign language countries are different, and the numbers of participants in the experiments are different. The number of sign language word data contained in the dataset is different. We chose Chu's system, which is similar in structure to our system, and both use bending sensors and IMUs, and we also used Japanese sign language, for comparison. The results are shown in Table 3, which shows that the weighted DTW has a better recognition rate when the number of participants and the number of recognized sign language words are both greater.

Table 3. Our system compared with Chu's system.

Research (Years)	Subject	Number of Signs	Algorithm	Sensor	Cross-Recognition
This study	8	20	Weighted DTW	Bend and IMU	85.21%
Chu et al. [26] (2021)	3	7	DTW	Bend and IMU	82.5%

There are still many limitations of our system. The data glove prototype system uses a breadboard, so the system is rather bulky. For some palm-related sign language words, the system sometimes causes inaccurate movements. However, the semantic impact on sign language expression is minimal. It is still able to recognize sign language words in sign language communication. Regarding the impact of data collection, there will be data loss or disconnection problems during long time periods of data collection.

In addition to hand shape features and hand motion features, collecting other features in sign language has the potential to further improve recognition rates in the future, for example, the relationship between head and hand position, body posture, facial expressions, etc. In addition, the data features of some locations on the hand do not contribute much to recognition, offering the possibility of simplifying the device in the future.

6. Conclusions

In this research, we built a Sign-Glove system to recognize sign language. By analyzing the process of sign language, we noticed that sign language is composed of both hand motion and the hand shape in time. Therefore, we decided to use IMU to detect the hand motion part and the bend sensor to detect the hand shape part. Then, we combined this information and used the weighted DTW algorithm to fuse the features and recognize sign

language words. In the experiments, we verified the performance of the Sign-Glove system and obtained high recognition rates of sign language. Such a wearable glove system has the potential to greatly reduce the cost of communication for people with hearing impairment.

In the future, with further improvements, we will exchange the cables for wireless connections such as BLE and Xbee. In addition, word-by-word sign language recognition was achieved, but sign language is often used to construct meaning through continuous use. We will replace the breadboard connection with a printed circuit board (PCB) and flexible flat cable (FFC) connections to achieve more stable data collection over a long period of time in daily use. We hope to build a system capable of continuous sign language recognition in the future. A more concise system will provide more convenient and complete sign language expression.

Author Contributions: Conceptualization, C.L. and S.A.; software, S.A.; validation, S.A. and C.L.; formal analysis, C.L.; investigation, C.L. and S.A.; resources, L.J.; data curation, S.A.; writing—original draft preparation, S.A. and C.L.; writing—review and editing, C.L.; visualization, C.L. and S.A.; supervision, L.J.; project administration, L.J.; funding acquisition, L.J. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by JSPS KAKENHI Grant Number 22K12114, the JKA Foundation, and NEDO Intensive Support for Young Promising Researchers Number JPNP20004.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Dataset used in this paper is available and can be accessed in the Google Drive repository: https://drive.google.com/drive/folders/1dpd2QMe18tIRI_uMXA6l9WGyb7hwHQyr?usp=sharing.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ministry of Health. Labour and Welfare Home Page, “2016 Survey on Difficulty in Life (Nationwide Fact-Finding Survey on Children with Disabilities at Home) Results”. Available online: https://www.mhlw.go.jp/toukei/list/dl/seikatsu_chousa_c_h28.pdf (accessed on 20 January 2023).
2. Rastgoo, R.; Kiani, K.; Escalera, S. Sign Language Recognition: A Deep Survey. *Expert Syst. Appl.* **2021**, *164*, 113794. [CrossRef]
3. Amin, M.S.; Rizvi, S.T.; Hossain, M.M. A Comparative Review on Applications of Different Sensors for Sign Language Recognition. *J. Imaging* **2022**, *8*, 98. [CrossRef] [PubMed]
4. Jiang, S.; Kang, P.; Song, X.; Lo, B.P.; Shull, P.B. Emerging Wearable Interfaces and Algorithms for Hand Gesture Recognition: A Survey. *IEEE Rev. Biomed. Eng.* **2021**, *15*, 85–102. [CrossRef] [PubMed]
5. Seçkin, A.Ç. Multi-Sensor Glove Design and Bio-Signal Data Collection. *Nat. Appl. Sci. J. Full Pap. 2nd Int. Congr. Updates Biomed. Eng.* **2021**, *3*, 87–93.
6. Seçkin, M.; Seçkin, A.Ç.; Genç, Ç. Biomedical Sensors and Applications of Wearable Technologies on Arm and Hand. *Biomed. Mater. Devices* **2022**, *1*, 1–13. [CrossRef]
7. Aloysius, N.; Geetha, M.K. Understanding vision-based continuous sign language recognition. *Multimed. Tools Appl.* **2020**, *79*, 22177–22209. [CrossRef]
8. Sharma, S.; Singh, S. Vision-Based Sign Language Recognition System: A Comprehensive Review. In Proceedings of the 2020 International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 26–28 February 2020; pp. 140–144.
9. Ma, Y.; Zhou, G.; Wang, S.; Zhao, H.; Jung, W. SignFi: Sign Language Recognition Using WiFi. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2018**, *2*, 23:1–23:21. [CrossRef]
10. He, W.; Wu, K.; Zou, Y.; Ming, Z. WiG: WiFi-Based Gesture Recognition System. In Proceedings of the 2015 24th International Conference on Computer Communication and Networks (ICCCN), Las Vegas, NV, USA, 3–6 August 2015; pp. 1–7.
11. Kudrinko, K.; Flavin, E.; Zhu, X.; Li, Q. Wearable Sensor-Based Sign Language Recognition: A Comprehensive Review. *IEEE Rev. Biomed. Eng.* **2020**, *14*, 82–97. [CrossRef] [PubMed]
12. Lokhande, P.M.; Prajapati, R.; Pansare, S. Data Gloves for Sign Language Recognition System. *Int. J. Comput. Appl.* **2015**, *975*, 8887.
13. Samaan, G.H.; Wadie, A.R.; Attia, A.K.; Asaad, A.M.; Kamel, A.E.; Slim, S.O.; Abdallah, M.S.; Cho, Y. MediaPipe’s Landmarks with RNN for Dynamic Sign Language Recognition. *Electronics* **2022**, *11*, 3228. [CrossRef]
14. Kohei, M.; Youngha, C.; Nobuhiko, M. *Recognition of Fingerspelling in Japanese Sign Language based on Nail Detection and Wrist Position*; ITE Technical Report; ITE: Singapore, 2013; pp. 199–202.
15. Shibata, H.; Hiromitsu, N.; Hiroshi, T.; Daisuke, K. Similarity Analysis of Motion Difference for Sign Language Recognition using Colored Gloves. *Forum Inf. Technol.* **2015**, *14*, 551–554.

16. Salagar, M.; Kulkarni, P.; Gondane, S. Implementation of Dynamic Time Warping for Gesture Recognition in Sign Language Using High Performance Computing. In Proceedings of the 2013 International Conference on Human Computer Interactions (ICHCI), Chennai, India, 23–24 August 2013; pp. 1–6.
17. Korzeniewska, E.; Kania, M.; Zawislak, R. Textronic Glove Translating Polish Sign Language. *Sensors* **2022**, *22*, 6788. [[CrossRef](#)] [[PubMed](#)]
18. Na, Y.; Yang, H.; Woo, J. Classification of the Korean Sign Language Alphabet Using an Accelerometer with a Support Vector Machine. *J. Sensors* **2021**, *2021*, 9304925:1–9304925:10. [[CrossRef](#)]
19. Tateno, S.; Liu, H.; Ou, J. Development of Sign Language Motion Recognition System for Hearing-Impaired People Using Electromyography Signal. *Sensors* **2020**, *20*, 5807. [[CrossRef](#)] [[PubMed](#)]
20. Khomami, S.A.; Shamekhi, S. Persian sign language recognition using IMU and surface EMG sensors. *Measurement* **2021**, *168*, 108471. [[CrossRef](#)]
21. Abhishek, K.S.; Qubeley, L.C.F.; Ho, D. Gloved-Based Hand Gesture Recognition Sign Language Translator Using Capacitive touch sensor. In Proceedings of the IEEE International Conference on Elrctron Devices and Solid-State Circuits (EDSSC), Hong Kong, China, 3–5 August 2016; pp. 334–337.
22. Gałka, J.; Masiar, M.; Zaborski, M.; Barczewska, K. Inertial Motion Sensing Glove for Sign Language Gesture Acquisition and Recognition. *IEEE Sens. J.* **2016**, *16*, 6310–6316. [[CrossRef](#)]
23. Faisal, M.A.; Abir, F.F.; Ahmed, M.U. Sensor Dataglove for Real-Time Static and Dynamic Hand Gesture Recognition. In Proceedings of the 2021 Joint 10th International Conference on Informatics, Electronics & Vision (ICIEV) and 2021 5th International Conference on Imaging, Vision & Pattern Recognition (icIVPR), Kitakyushu Virtual, Japan, 16–20 August 2021; pp. 1–7.
24. Faisal, M.A.; Abir, F.F.; Ahmed, M.U.; Ahad, M. Exploiting domain transformation and deep learning for hand gesture recognition using a low-cost dataglove. *Sci. Rep.* **2022**, *12*, 21446. [[CrossRef](#)] [[PubMed](#)]
25. Lee, B.; Lee, S.M. Smart Wearable Hand Device for Sign Language Interpretation System with Sensors Fusion. *IEEE Sens. J.* **2018**, *18*, 1224–1232. [[CrossRef](#)]
26. Chu, X.; Liu, J.; Shimamoto, S. A Sensor-Based Hand Gesture Recognition System for Japanese Sign Language. In Proceedings of the 2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech), Nara, Japan, 9–11 March 2021; pp. 311–312.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.