

Article

Development of Sign Language Motion Recognition System for Hearing-Impaired People Using Electromyography Signal

Shigeyuki Tateno *, Hongbin Liu and Junhong Ou

Graduate School of Information, Production and Systems, Waseda University, Kitakyushu 808-0135, Japan; liuhongbin@toki.waseda.jp (H.L.); ohjunhong@ruri.waseda.jp (J.O.)

* Correspondence: tateno@waseda.jp

Received: 22 September 2020; Accepted: 12 October 2020; Published: 14 October 2020



Abstract: Sign languages are developed around the world for hearing-impaired people to communicate with others who understand them. Different grammar and alphabets limit the usage of sign languages between different sign language users. Furthermore, training is required for hearing-intact people to communicate with them. Therefore, in this paper, a real-time motion recognition system based on an electromyography signal is proposed for recognizing actual American Sign Language (ASL) hand motions for helping hearing-impaired people communicate with others and training normal people to understand the sign languages. A bilinear model is applied to deal with the on electromyography (EMG) data for decreasing the individual difference among different people. A long short-term memory neural network is used in this paper as the classifier. Twenty sign language motions in the ASL library are selected for recognition in order to increase the practicability of the system. The results indicate that this system can recognize these twenty motions with high accuracy among twenty participants. Therefore, this system has the potential to be widely applied to help hearing-impaired people for daily communication and normal people to understand the sign languages.

Keywords: motion recognition; electromyography; long short-term memory neural network; bilinear model; sign language

1. Introduction

According to the World Health Organization, 466 million people are suffering from hearing loss around the world in 2020 [1]. Sign language is an essential tool for them to communicate with others. Recently, studies of deafness have adopted more complex sociocultural perspectives, raising issues of community identity, formation and maintenance, and language ideology [2]. As a meaning to construct individual communication, sign languages do not share the same standard worldwide. Instead, cultural difference along with other factors create a huge difference among the sign languages [3–6]. The physical component of sign languages usually consists of movement of forearm and hand motion. The formation of sentences is also different in terms of grammar, vocabulary, and alphabets among various sign languages [7]. Obstacles stand between hearing-intact people and hearing-impaired people when communicating; therefore, special training is required to understand the sign languages.

Different approaches are applied to develop a sign language recognition system in which camera sensors and sensor-integrated gloves are commonly used. In 2014, C. H. Chuan et al. used the leap motion sensor to capture the hand movements of the user to recognize different hand gestures of the American Sign Language (ASL) [8,9]. H. E. Hayek et al. proposed a sign translator system by using a hand glove in 2015 [10]. However, in some situations, limitations of these sensors exist. The camera

sensor requires a lighting environment and has a limited detection range [11,12], and the glove is expensive and uneasy to be worn [13,14].

In order to solve the problems, researchers choose a kind of sensor based on electromyography (EMG) signal. The EMG signal collects bioelectrical signals of muscles during muscle extension or contraction from the forearm, which can avoid limitations of the camera sensors and the glove sensors. The EMG signal from the forearm can be utilized to control an artificial arm [15,16]. Moreover, functional states of muscle movements can be reflected by the EMG signal [17–21]. Savur and Sahin used a surface EMG signal to recognize the ASL letters alphabet to allow users to spell words and sentences with an accuracy of 60% [22]. Lionel and other researchers used convolutional neural networks (CNNs) on the EMG data of the forearm to recognize 20 Italian gestures [23]. Seongjoo and others used sensor fusion technology and group-dependent neural network models to recognize Korean sign language [24]. After the collection of EMG data, measures are applied to process raw EMG data. Features of the EMG data can be obtained by applying a model on each channel or modeling all channels as a whole [25]. In both the time domain and frequency domain, features are extracted to analyze the EMG data [26,27]. Recently, some studies are focusing on transforming time-series EMG data into images for utilizing image recognition techniques to avoid information loss during the feature extracting process [28,29]. Various classifiers are applied to identify different movements of muscles [30–32].

In essence, the sign language recognition system is to distinguish the time continuous gestures of forearms. Jaramillo-Yáñez and others made a systematic review on this subject [33]. Various devices of different sampling rate ranging from 100 Hz to 1200 Hz were used to collect raw data [34–37]. Approximately 95% of the signal power was below 400–500 Hz, which required the sampling rate to reach 1000 Hz in order to gather all the information according to the Nyquist sampling theory [38–40]. However, studies employing a low sampling rate device still obtained a decent accuracy with different approaches applied [41–43]. Some of other studies only carried the experiment on single participant and developed systems which had high accuracy [44–46]. However, as an electrophysiological signal, the EMG signal has individual differences, which will greatly reduce the accuracy of the system based on a single participant [47]. The same movements of the same muscles in different people can generate different EMG signals. Several methods were proposed to tackle this problem [48–51]. Some researchers developed a bilinear model to overcome individual differences. In 2000, J. B. Tenenbaum et al. proposed the definition and the algorithm of a bilinear model to deal with the problem of face recognition [52]. In 2013, Matsubara et al. applied the bilinear model for the first time in the recognition of five types of hand gestures (four types of motion gestures and one type of static relaxation) to control a robot hand [53]. Later, in 2014, Wang Tao et al. used a bilinear model to perform a single-finger compression experiment under different contractions [54].

Moreover, the review [33] showed that almost all the studies considered improving recognition accuracy and a few studies considered implementing actual real-time application on portable devices or embedded systems. In our research, a sign language motion recognition system based on electromyography (EMG) signal is proposed to realize a real-time application. Conventional EMG data processing methods are utilized to extract ten features from raw EMG data, and then the constructed features are input into a bilinear model. Conventional feature processing includes the feature dimensionality reduction and normalization. By extracting the features that have more contributions like the principal component analysis, as well as normalization of feature amplitude or time scale, individual differences can be eliminated within a certain range [55]. In addition, using a small number of training samples of test users to participate in the learning and training of the classifier can improve the recognition results of non-specific persons to a certain extent, like the transfer learning method [56]. By inputting the features into the bilinear model, the relationship between users and motions is considered. Thus, an EMG control interactive system with the ability to recognize non-specific human motions with high accuracy is constructed. By using this system, the sign language

motions which are commonly used in the daily life of hearing-impaired people can be recognized in real time and the meaning of these motions can be output to make normal people understand.

2. Mechanism and Algorithm

In this proposed system, the EMG signal of muscle is obtained to classify into different motion categories. A long short-term memory (LSTM) neural network is utilized as a classifier since the LSTM has good performance in time-series data classification. The memory units in the LSTM can be of great help in maintaining the useful information and discarding the interferences from the previous input data to affect the current state positively.

Firstly, the armband is worn on the user's forearm. During the process of making hand motions, the surface EMG signal is recorded to analyze the information of the muscles, such as muscle contraction, extension, and relaxed. Secondly, several widely used features are calculated to obtain the characteristics of the EMG data in both the time domain and the frequency domain. The feature selection needs to be applied to reduce the computation cost of the system. Therefore, the permutation feature importance algorithm [57] is conducted, and then several useful features will be selected. Thirdly, the parameters of a bilinear model will be adjusted, then the selected feature values are decomposed by the bilinear model for extracting the motion-dependent factors to decrease the individual difference of the EMG data since it differs obviously among different people. Fourthly, the obtained motion-dependent factors are input into the LSTM for recognition. Finally, the motion label of the EMG data is used to output the corresponding meaning of the hand motion. The flow chart of the whole system is shown in Figure 1.

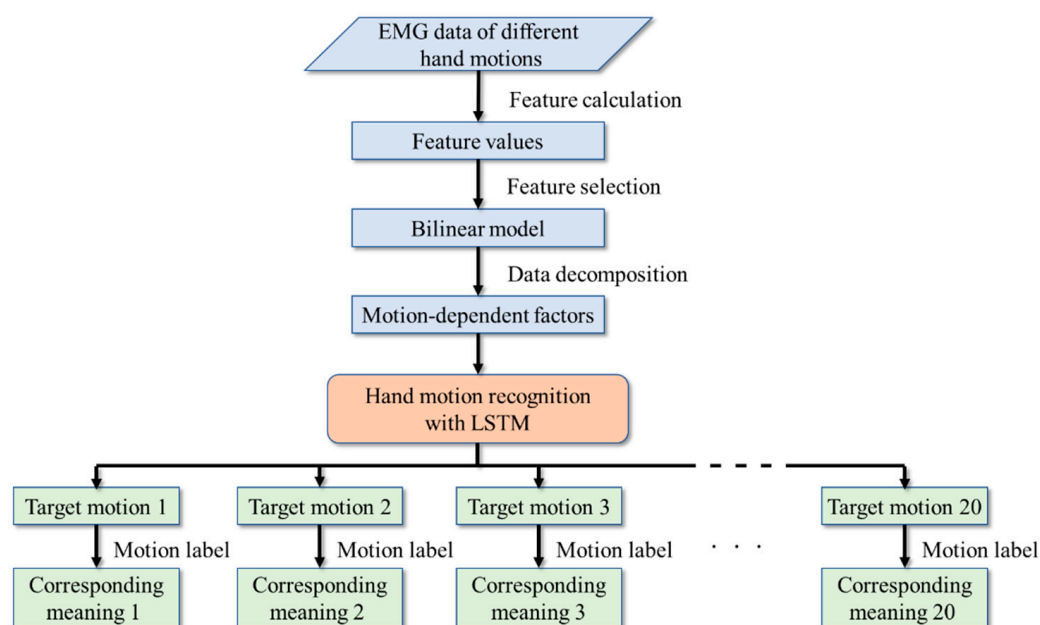


Figure 1. The flow chart of the system.

2.1. EMG Data Collection

The EMG sensor applied in this system is a Myo armband which is manufactured by Thalmic Labs. Compared with other EMG sensors such as electrodes made by Delsys and Otto Block company, the Myo armband transmits EMG data through Bluetooth, which can ensure the quality of EMG signals by reducing noise caused by cables and make it easy to wear. The Myo armband shown in Figure 2 has eight EMG sensors whose sampling rate is 200 Hz.

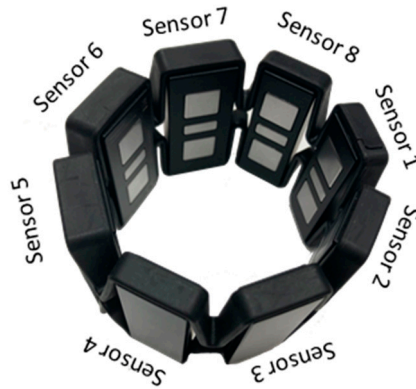


Figure 2. The Myo armband.

An example of the EMG data in eight channels is shown in Figure 3.

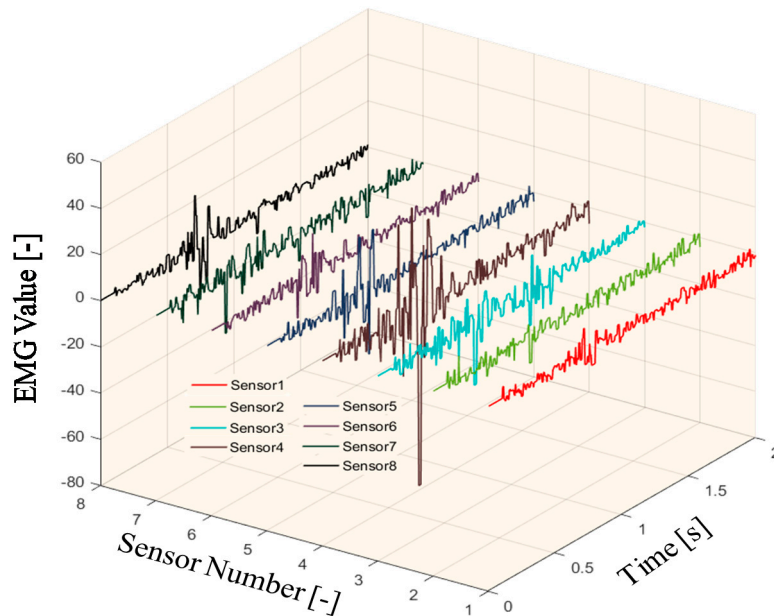


Figure 3. The electromyography (EMG) data in eight channels.

2.2. EMG Data Processing

Since the EMG data provided by the armband is the time-series data that describes how the muscle state varies during the process of performing hand motions, conventional data processing methods are selecting suitable features in both time and frequency domains to calculate the feature values as the input of classifiers.

The feature extraction is an important method to obtain useful characteristics of the EMG data and remove redundant or interfering information. Sometimes, the feature dimensionality reduction and the permutation feature importance are needed to select more important information from the feature values. Here, some commonly used features are listed as shown in Equations (1)–(10).

The first part is about the time-domain features which are calculated based on the raw time series EMG data. As the most basic and commonly used feature in statistical analysis, the mean absolute value (MAV) is calculated as:

$$MAV_j = \frac{1}{N} \sum_{i=1}^N |(EMG_i)_j|, j = 1, 2, \dots, C, \quad (1)$$

where j is the channel number of the EMG data, $(EMG_i)_j$ is the single value of the EMG data in channel j , N is the amount of the EMG data in channel j .

The second one is the standard deviation (STD), which describes the value variation of the EMG data:

$$STD_j = \sqrt{\frac{1}{N} \sum_{i=1}^N ((EMG_i)_j - \mu)^2}, \quad j = 1, 2, \dots, C, \quad (2)$$

where μ is the average value of the EMG data in channel j .

The third one is the root mean square (RMS). In the EMG analysis area, it is modeled as an amplitude modulated Gaussian random process, which relates to constant force and non-fatiguing contraction [58]. The RMS is calculated as follows:

$$RMS_j = \sqrt{\frac{1}{N} \sum_{i=1}^N [(EMG_i)_j]^2}, \quad j = 1, 2, \dots, C. \quad (3)$$

The fourth one is the log detector (LOG), which provides the estimations of muscle contraction force [59], as shown in Equation (4).

$$LOG_j = e^{\frac{1}{N} \sum_{i=1}^N \log|(EMG_i)_j|}, \quad j = 1, 2, \dots, C. \quad (4)$$

The final one is the average amplitude change (AAC), which is a measurement of the complexity of the EMG data and represents the average of the data difference over the time segment [59]. It can be calculated as:

$$AAC_j = \frac{1}{N-1} \sum_{i=1}^{N-1} |(EMG_{i+1})_j - (EMG_i)_j|, \quad j = 1, 2, \dots, C. \quad (5)$$

The second part is the frequency-domain features that represent the generated power of the working muscle during muscle movements and can be used to detect muscle fatigue. In order to obtain the frequency-domain features, the power spectrogram of the EMG data is firstly calculated which is based on the Welch's method. The data used in this research is obtained from a Myo band with a sampling rate of 200 Hz. The result is shown in Figure 4.

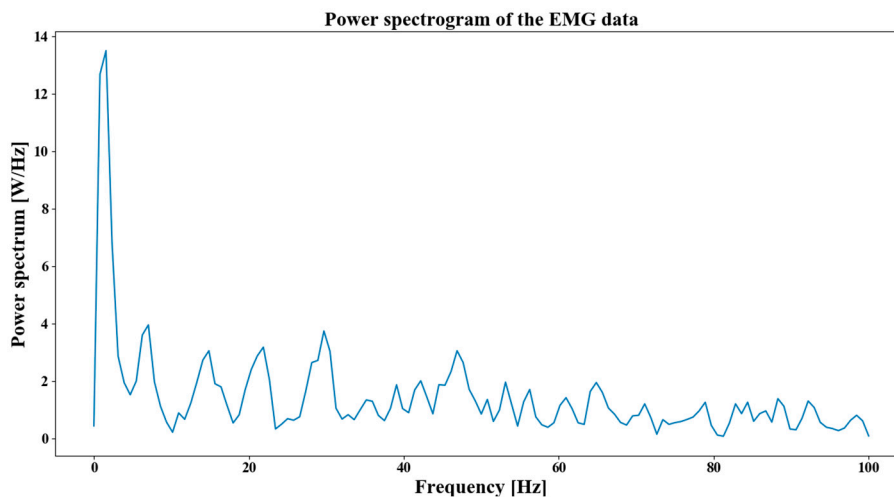


Figure 4. The power spectrogram of EMG data.

Then, the first frequency-domain feature is the mean frequency (MNF) which is the sum of the product of the EMG power spectrum and the frequency divided by the sum of spectrum intensity, as shown in Equation (6):

$$MNF_j = \frac{\sum_{i=1}^N f_i P_i}{\sum_{i=1}^N P_i}, \quad j = 1, 2, \dots, C. \quad (6)$$

where f_i is the frequency of the spectrum at the i -th frequency bin after the Fourier transform, P_i is the i -th spectrum value and N is the length of the frequency bin.

The second commonly used frequency-domain feature is the median frequency (MDF), which divides the spectrums into two regions with equal amplitude:

$$\sum_{i=1}^{MDF} P_{ij} = \sum_{i=MDF}^N P_{ij} = \frac{1}{2} \sum_{i=1}^N P_i, j = 1, 2, \dots, C. \quad (7)$$

Similar to the MNF, the mean power (MNP) is calculated as follows:

$$MNP_j = \sum_{i=1}^N f_i P_i / \sum_{i=1}^N f_i, j = 1, 2, \dots, C. \quad (8)$$

The next one is the power spectrum ratio (PSR), which measures the ratio between the maximum value and the whole energy of the EMG power spectrum:

$$PSR_j = \frac{P_0}{P} = \frac{\sum_{f_0-\varepsilon}^{f_0+\varepsilon} P_i}{\sum_{f_1}^{f_2} P_i}, j = 1, 2, \dots, C. \quad (9)$$

The last frequency-domain feature is the peak frequency (PKF), which is the frequency value where the maximum power value appears:

$$PKF_j = f_j (\max(P_i)), j = 1, 2, \dots, C. \quad (10)$$

After the feature calculation, the calculated feature values are firstly selected by a permutation feature importance process, and later input into a bilinear model to extract the motion-dependent factors for classification.

2.3. Bilinear Model Algorithm

According to the definition of the bilinear model, the EMG signal Y can be decomposed into the user-related factors Z and the motion-related factors X with a weight matrix W to describe the factor interactions, as shown in Figure 5 [52–54].

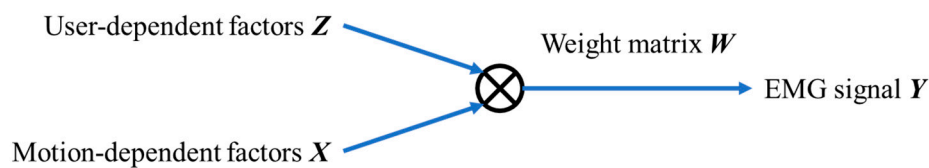


Figure 5. The composition of the EMG signal in the bilinear model.

For a single EMG signal value y , it can be represented in the following form:

$$y = z^T W_c x, \quad (11)$$

where $z \in R^I$ represents the user-related factor and $x \in R^J$ represents the motion-related factor. $W_c \in R^{I \times J}$ is the parameter matrix of the bilinear model which describes the factor interactions between z and x .

Suppose that the EMG signal is $Y \in R^C$ where $c \in 1 \sim C$ is the channel serial number of the EMG signal, the subject serial number is $u \in 1 \sim U$ the motion serial number is $m \in 1 \sim M$ and the data serial number for one motion is $n \in 1 \sim N$. Therefore, the problem of fitting the bilinear model can be described as searching for suitable variables $\{z^{uT}, W_c, x_n^m\}$ for all u, c, m , and n , to minimize the

difference between the constructed EMG signal which is calculated by Equation (11) and the original EMG signal Y . Therefore, the objective function of fitting the bilinear model is as follows:

$$E = \sum_{u=1}^U \sum_{n=1}^N \sum_{m=1}^M \sum_{c=1}^C \left\| y_{cn}^{um} - z^{uT} W_c x_n^m \right\|_2. \quad (12)$$

Conventionally, the definition of the EMG signal is a multiple-dimensional matrix, in which each dimension describes different information such as users, channels, motions, and so on, respectively. However, in the bilinear model, the matrices are always two-dimensional in order to utilize some standard matrix processing algorithms. Therefore, the multiple-dimensional matrices are expanded into the stacked two-dimensional matrices where the data is arranged in a specific order. By using this concept, the obtained dataset of EMG signal Y can be presented in the following form:

$$Y = \begin{bmatrix} y_{11}^{11} \cdots y_{1N}^{11} & y_{11}^{12} & \cdots & y_{1N}^{1M} \\ & \vdots & & \\ y_{C1}^{11} \cdots y_{CN}^{11} & y_{C1}^{12} & \cdots & y_{CN}^{1M} \\ y_{11}^{21} \cdots y_{1N}^{21} & y_{11}^{22} & \cdots & y_{1N}^{2M} \\ & \vdots & & \\ y_{C1}^{U1} \cdots y_{CN}^{U1} & y_{C1}^{U2} & \cdots & y_{CN}^{UM} \end{bmatrix} \in R^{CU \times MN}, \quad (13)$$

where U is the number of users, C is the number of channels, M is the number of motions, and N is the amount of data in one motion. Similarly, the definitions of the user-related matrix Z , the motion-related matrix X , and the weight matrix W are shown from Equations (14) to (16):

$$Z = \begin{bmatrix} z^1 & z^2 & \cdots & z^U \end{bmatrix} \in R^{I \times U}, \quad (14)$$

$$X = \begin{bmatrix} x_1^1 & x_2^1 & \cdots & x_N^1 & \cdots & x_N^M \end{bmatrix} \in R^{J \times MN}, \quad (15)$$

$$W = \begin{bmatrix} W_1 & W_2 & \cdots & W_C \end{bmatrix} \in R^{IC \times J}. \quad (16)$$

Basically, the calculation methods of the stacked matrices are the same as normal matrices. However, there are still some differences. When the channel number of the EMG signal is larger than one, the data of the same user and motion but different channels should be considered as a whole part for the calculation, especially for the matrix transpose. Therefore, the stacked transpose (ST) is defined as follows: for a $MC \times N$ stacked matrix, its ST can be defined as a $NC \times M$ matrix, as shown in Figure 6.

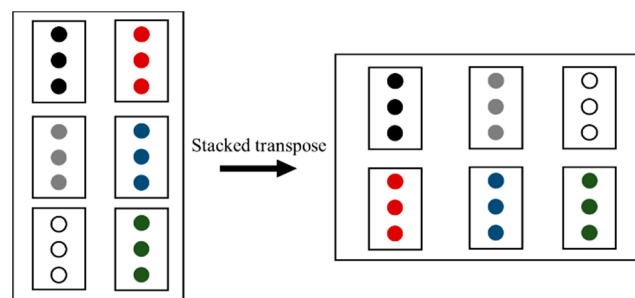


Figure 6. The Schematic diagram of the stacked transpose.

With all these definitions, two equivalent equations of the variables introduced above can be obtained as shown from Equations (17) and (18):

$$Y = [W^{ST} Z]^{ST} X, \quad (17)$$

$$\mathbf{Y}^{\text{ST}} = [\mathbf{W}\mathbf{X}]^{\text{ST}}\mathbf{Z}. \quad (18)$$

For determining the optimal matrices \mathbf{Z} , \mathbf{X} , and \mathbf{W} , the iterative procedure is as follows. Firstly, the singular value decomposition (SVD) of the EMG signal \mathbf{Y} is calculated as $\mathbf{Y} \xrightarrow{\text{SVD}} \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\text{T}}$ where \mathbf{U} and \mathbf{V}^{T} are unitary matrices and $\mathbf{\Sigma}$ is a diagonal matrix whose diagonal elements are the singular values of the \mathbf{Y} . Then, \mathbf{X} is initialized as the first J rows of the \mathbf{V}^{T} . Next, from the initialized \mathbf{X} and the EMG data \mathbf{Y} , \mathbf{Z} is updated as the first I rows of the \mathbf{V}^{T} , where $[\mathbf{Y}\mathbf{X}^{\text{T}}]^{\text{ST}} \xrightarrow{\text{SVD}} \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\text{T}}$. Finally, using the derived \mathbf{Z} and the EMG data \mathbf{Y} , updating the \mathbf{X} as the first J rows of the \mathbf{V}^{T} , where $[\mathbf{Y}^{\text{ST}}\mathbf{Z}^{\text{T}}]^{\text{VT}} \xrightarrow{\text{SVD}} \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\text{T}}$. Except for the initialization of the \mathbf{X} , the updating procedure of the \mathbf{Z} and the \mathbf{X} is a one-time iteration of the bilinear model algorithm. This algorithm converges within 10 iterations. Once the algorithm converges, the obtained matrix \mathbf{X} of which each component x_n^m corresponds to the motion label m is used for training a classifier.

For testing the performance of the bilinear model, a new subject needs to perform at least one motion m to extract his/her user-related matrix \mathbf{z}_{new} with Equation (19):

$$\mathbf{z}_{\text{new}} = \left\{ \left[[\mathbf{W}\mathbf{X}_m]^{\text{ST}} \right]^+ * \mathbf{Y}_{\text{new}_m}^{\text{ST}}, \right\} \quad (19)$$

where $\{\cdot\}^+$ means the pseudo-inverse matrix of matrix $\{\cdot\}$, $\mathbf{Y}_{\text{new}_m}$ is EMG data of the new subject when performing the motion m , and \mathbf{W} and \mathbf{X}_m are previous weight matrix and the corresponding motion matrix from the obtained matrix \mathbf{X} , respectively. With the new user-related matrix \mathbf{z}_{new} the new weight matrix $\mathbf{W}_{\text{new}_m}$ is derived by Equation (20):

$$\mathbf{W}_{\text{new}_m} = \left\{ \left[\mathbf{Y}_{\text{new}_m}^{\text{ST}} \{\mathbf{z}_{\text{new}}\}^+ \right]^{\text{ST}} \right\}^+ \times \{\mathbf{X}_m\}^+. \quad (20)$$

Finally, with the derived \mathbf{z}_{new} and $\mathbf{W}_{\text{new}_m}$ a new motion-related matrix $\mathbf{X}_{\text{new}_m}$ for motion m is extracted by Equation (21):

$$\mathbf{X}_{\text{new}_m} = \left\{ \left[\mathbf{W}_{\text{new}_m}^{\text{ST}} \mathbf{z}_{\text{new}} \right]^{\text{ST}} \right\}^+ \times \mathbf{Y}_{\text{new}_m}. \quad (21)$$

The new motion-related matrix $\mathbf{X}_{\text{new}_m}$ of which each component $x_{\text{new}_n}^m$ corresponds to the motion label m is obtained for testing the classifier.

2.4. Hand Motion Classify

In this paper, an LSTM is chosen as the hand motion classifier. In order to introduce the LSTM, firstly, the concept of a recurrent neural network (RNN) needs to be clarified. An RNN is a kind of deep neural network whose current node weight is not only decided by the current input but also affected by the previous input [60]. Therefore, the RNN is widely applied to deal with the data which can be considered to be related in time slices; that is, the state generated at the current time point is affected by the previous time point, and it will affect the output state at the subsequent time point.

However, for a long time-series data, since the structure of the RNN is essentially a recursive nested structure, which will cause the problem of “gradient explosion” or “gradient vanish” so that all the previous information is lost [60]. Therefore, the LSTM is introduced to add mainly three gates to control the information circulation: how much previous information dropping, how much current information inputting, and how much current information outputting. With these gates controlling the information circulation, the performance of the LSTM is much better than the simple RNN in a long time-series data classification task. Therefore, in this paper, the LSTM is selected as the classifier. The structure of the LSTM is shown in Figure 7.

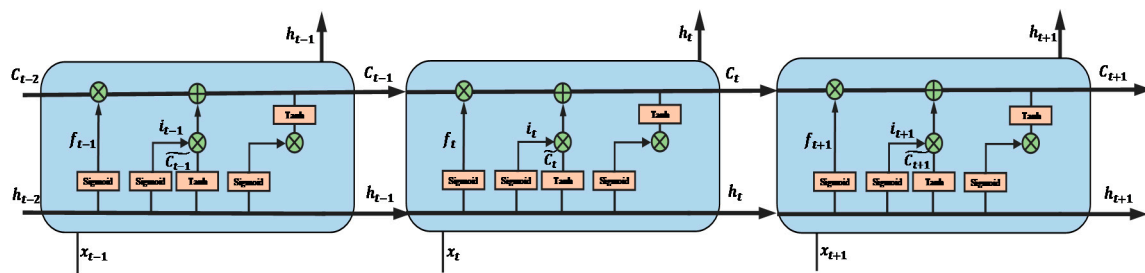


Figure 7. The structure of the long short-term memory (LSTM).

3. Experiment

The experiment consisted of two steps. The first step was to carry the experiment on a single participant. The second step was to experiment with multiple participants.

For the first step, the first thing to confirm was that the EMG signal could be used to recognize different hand motions for one participant. According to the ASL library, 20 hand motions of different meanings that are used for actual communication by hearing-impaired people were selected for recognition. The 20 hand motions are shown in Table 1. Some examples are shown in Figure 8.

Table 1. The 20 hand motions.

Motion No.	Motion Name	Motion No.	Motion Name
M1	How are you?	M11	Where is the store?
M2	Nice to meet you.	M12	How can I get food?
M3	See you later.	M13	How much does it cost?
M4	That's what I mean.	M14	Yes, thank you.
M5	I don't understand.	M15	I am sorry.
M6	What is your name?	M16	Where is the hospital?
M7	Where are you from?	M17	I don't feel good.
M8	What happens?	M18	Please help me.
M9	What is wrong?	M19	Please write it.
M10	Please call 911.	M20	I love you.

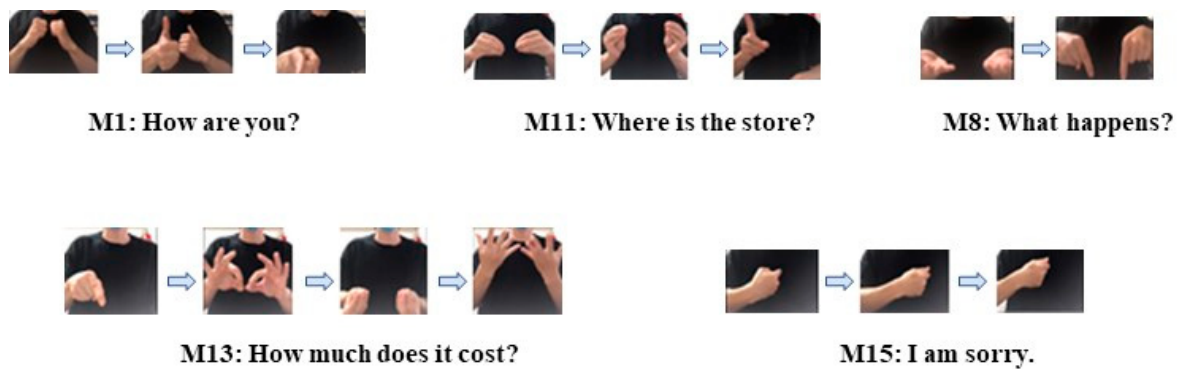


Figure 8. Sign language motions.

The participant was asked to wear the armband on his right hand since the selected 20 motions include the right-hand motions and the participant was right hand dominated. The experiment environment setting is shown in Figure 9.

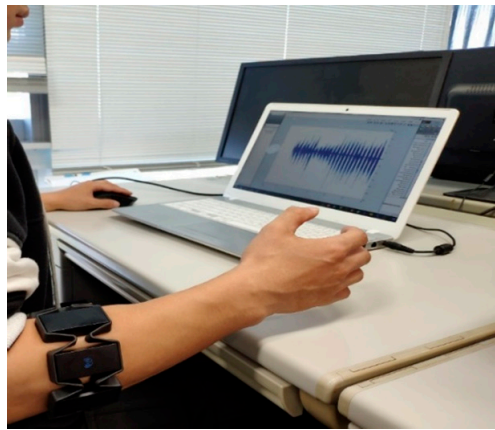


Figure 9. The environment settings of EMG data obtaining.

At the beginning of the experiment, the participant was asked to watch the videos of the sign language motions recorded in advance and practiced the motions in order to finish the motion within three seconds. What is more, before the experiment day, the participant was asked not to conduct heavy activities that require the right hand in dominate to avoid the muscle fatigue of his forearm.

For the second step, in order to apply the system to a wide range of users, the experiments on multiple participants were conducted. Twenty healthy male participants whose ages are between 23 and 25 and whose dominant hand is the right hand are recruited. For each participant, each hand motion is performed within three seconds, and each experiment was conducted under the same condition and environment as the participant one.

4. Results

4.1. Single-Person Experiment

In this experiment, the participant was required to perform each motion within three seconds which was corresponding to 600 pieces of EMG data. For each motion, the repeating cycle was set as 10, and moreover this cycle was repeated five times, changing the sensor position to obtain enough EMG data as much as possible in order to perform the permutation importance method. The obtained EMG data of one participant is shown in Figure 10.

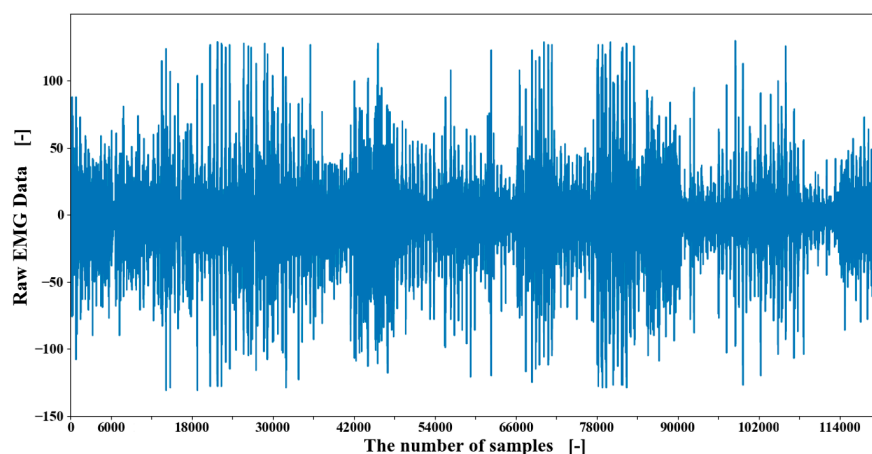


Figure 10. The obtained raw EMG data of 20 motions of participant one (10 times repeating for each motion).

The time-domain features and the frequency-domain features introduced in Section 2.2 were calculated with the window size of 600. Therefore, for each motion, 20 feature stacks were obtained.

For example, the calculated RMS feature values from one channel of the EMG data of participant one are shown in Figure 11. Each motion was performed 50 times, and there were 20 motions in total.

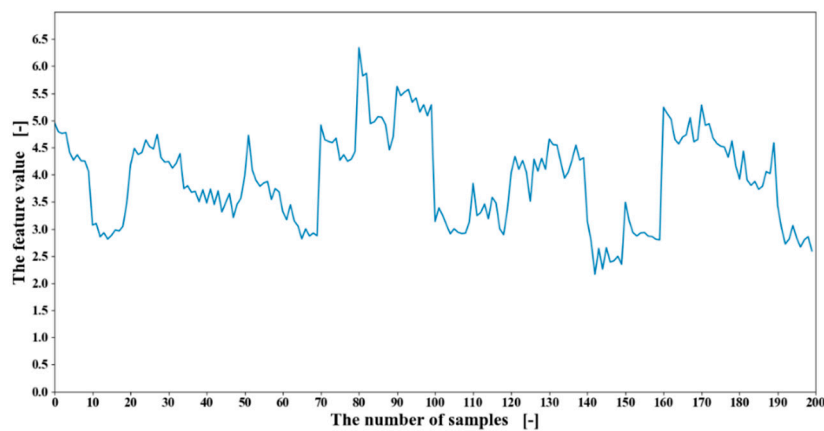


Figure 11. The root mean square (RMS) feature value of one channel from obtained EMG data of participant one.

First, all the 10 features were used for training and testing of the LSTM. In the LSTM, after using different learning rates ranging from 0.00001 to 0.001, the learning rate is set to 0.0001 for higher accuracy possible. The iteration time is set to 500 which can make the validation loss converge on the training dataset. The training dataset and the testing dataset are with a ratio of 0.8 and 0.2, respectively. Each motion is performed totally 50 times. The results of the LSTM are shown in Table 2.

Table 2. The classification result of the LSTM on the single participant.

		Predicted Label																			
		M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20
Actual Label	M1	47	0	2	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
	M2	0	49	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
	M3	1	0	47	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0
	M4	0	0	0	48	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0
	M5	0	0	1	0	49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	M6	0	0	0	0	0	49	0	0	0	1	0	0	0	0	0	0	0	0	0	0
	M7	0	0	0	0	0	0	49	0	0	0	1	0	0	0	0	0	0	0	0	0
	M8	0	0	0	1	0	0	0	47	0	0	0	0	1	0	0	1	0	0	0	0
	M9	0	0	0	0	0	0	0	0	50	0	0	0	0	0	0	0	0	0	0	0
	M10	0	0	0	0	0	0	0	0	0	50	0	0	0	0	0	0	0	0	0	0
	M11	0	0	0	0	0	0	1	0	0	0	49	0	0	0	0	0	0	0	0	0
	M12	0	0	0	0	0	0	0	0	0	0	0	49	0	0	0	0	0	0	1	0
	M13	0	0	0	0	0	1	0	0	0	0	0	0	49	0	0	0	0	0	0	0
	M14	0	0	0	0	0	0	0	0	0	0	0	2	0	48	0	0	0	0	0	0
	M15	0	0	0	0	1	0	0	0	0	0	0	0	0	0	49	0	0	0	0	0
	M16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	50	0	0	0	0
	M17	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	49	0	0	0
	M18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	50	0	0
	M19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	50	0
	M20	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	48

The result of permutation feature importance is shown in Table 3. From Table 3, the priority of the features based on the feature importance is that, for time-domain features, RMS > LOG > AAC >> STD > MAV, and for frequency-domain features, PSR > MNP > MDF >> MNF > PKF. As the result, six features, RMS, LOG, AAC, PSR, MNP, and MDF were selected. After the feature datasets were calculated and selected by evaluating the feature importance, the obtained six feature values were input into the classifier. The average of accuracy with the six features slightly decreased by 0.7%, meanwhile, the computation time decreased by 30%.

Table 3. The importance of selected ten features.

	Time-Domain Features					Frequency-Domain Features				
	MAV	STD	RMS	LOG	AAC	MNF	MDF	MNP	PSR	PKF
Accuracy	98.54%	97.38%	77.34%	87.68%	89.74%	97.94%	88.48%	86.52%	81.72%	96.30%
Importance	1.32%	2.48%	22.52%	12.18%	10.12%	2.92%	11.42%	14.34%	19.14%	3.56%

4.2. Multi-Person Experiment

After the single person experiment, multi-person experiment was performed. In this experiment, the participants were required to perform each motion within three seconds, which corresponded to 600 pieces of EMG data. For each motion, the repeating cycle is set as 10. To verify the effectiveness of the bilinear model, data, with and without, the bilinear model process were input to the LSTM to test the accuracy.

4.2.1. Classification without the Bilinear Model

With the feature calculation method mentioned in the previous section, the feature stacks from 19 participants were used for the training, and the features from one participant were used for the testing. This procedure repeats for 20 times as twenty-fold cross-validation with each time changing the testing participant until all the participants were tested once.

The calculated features of 20 motions were directly input into the LSTM for classification. The classification result of 20 participants is shown in Table 4. The learning rate was set to 0.0001, and the iteration time was set to 500.

Table 4. The total classification results of 20 participants.

		Predicted Label																			
Actual Label		M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20
	M1	101	2	5	1	3	7	7	4	10	3	5	4	8	1	11	6	7	6	4	5
	M2	6	107	3	3	6	7	2	4	6	6	3	5	3	6	6	4	4	8	5	6
	M3	6	11	117	3	7	4	4	5	3	8	4	3	2	4	1	4	4	3	4	3
	M4	6	4	7	109	4	4	1	7	3	4	3	5	5	2	4	1	9	7	5	10
	M5	7	5	3	2	116	2	5	8	4	4	1	2	4	3	3	1	5	8	9	8
	M6	4	2	9	1	4	126	8	2	3	3	5	3	5	3	4	3	2	6	3	4
	M7	6	5	2	4	5	7	113	7	6	1	5	4	4	1	6	5	7	5	4	3
	M8	5	1	7	4	3	8	1	120	4	1	3	3	2	6	1	6	5	15	2	3
	M9	4	5	4	3	3	0	9	2	112	4	3	4	2	2	3	6	8	8	12	6
	M10	5	5	7	8	7	4	7	4	4	104	4	4	3	3	3	10	2	8	4	4
	M11	5	5	2	5	7	6	11	3	4	7	103	1	2	5	2	3	5	8	7	9
	M12	6	4	5	4	4	8	3	4	3	6	4	107	3	5	7	5	5	6	6	5
	M13	3	3	5	4	5	5	7	7	9	5	3	4	98	4	1	9	6	6	7	9
	M14	5	4	5	6	3	2	8	6	2	3	4	3	3	122	1	2	3	6	6	6
	M15	5	2	5	3	7	2	11	2	6	5	3	1	5	6	115	5	3	7	3	4
	M16	4	3	8	4	8	3	4	1	4	2	5	3	4	9	5	116	1	2	7	7
	M17	5	6	4	10	3	5	6	3	4	5	4	2	6	2	3	8	108	9	6	1
	M18	4	4	3	4	3	5	4	10	2	3	4	3	6	4	8	3	6	117	3	4
	M19	4	2	4	5	4	5	8	4	10	2	2	4	7	6	3	2	4	4	115	5
	M20	9	4	3	5	8	1	7	2	7	7	2	4	8	3	5	2	6	7	6	104

The accuracy of 20 motions is shown in Table 5.

Table 5. The classification accuracy of 20 hand motions on 20 participants.

Motion Name	Accuracy	Motion Name	Accuracy
M1: How are you?	50.5%	M11: Where is the store?	51.5%
M2: Nice to meet you.	53.5%	M12: How can I get food?	53.5%
M3: See you later.	58.5%	M13: How much does it cost?	49.0%
M4: That's what I mean.	54.5%	M14: Yes, thank you.	61.0%
M5: I don't understand.	58.0%	M15: I am sorry.	57.5%
M6: What is your name?	63.0%	M16: Where is the hospital?	58.0%
M7: Where are you from?	56.5%	M17: I don't feel good.	54.4%
M8: What happens?	60.0%	M18: Please help me.	58.5%
M9: What is wrong?	56.0%	M19: Please write it.	57.5%
M10: Please call 911.	52.0%	M20: I love you.	52.0%

The accuracy of each participant is shown in Table 6.

Table 6. The classification accuracy of 20 participants.

Name	Accuracy	Name	Accuracy
Participant 1	60.5%	Participant 11	49.5%
Participant 2	50.5%	Participant 12	62.0%
Participant 3	71.5%	Participant 13	48.0%
Participant 4	61.0%	Participant 14	61.0%
Participant 5	49.0%	Participant 15	61.5%
Participant 6	64.0%	Participant 16	59.5%
Participant 7	54.0%	Participant 17	50.0%
Participant 8	49.5%	Participant 18	54.5%
Participant 9	57.5%	Participant 19	51.5%
Participant 10	46.5%	Participant 20	48.0%

4.2.2. Classification with the Bilinear Model

From Table 5, almost all the motions cannot be correctly classified when the training data is not obtained from the test participant. What is more, from Table 6, the average classification accuracy of 20 motions among 20 participants drops sharply to 55.70%. As a result, it can be considered that the individual difference among different participants has strong influences on the EMG data so that the classification results are far from ideal. In order to solve the problem caused by the individual difference of the EMG data, as mentioned in Section 2.3, the bilinear model algorithm was applied. In this research, the user number U was 19 which was corresponding to the number of participants whose data was used for the training. The channel number C was set to 8 since there were 8 EMG sensors in the armband. The motion number M and the amount of data in each motion N were set at 20 and 10, respectively. The selection of the parameter I and J has a big influence on the final decomposition results since the values of the I and the J are the size of z^u , which contains the user factors, and the size of x_n^m , which contains the motion factors, respectively.

However, no theory or formula can describe how to decide the best I and J values. The choices of the I and the J were based on the prior experience, which means that the I and the J were adjusted according to the final classification accuracy. In this research, the range of the I was from 1 to 10, and the range of the J was from 10 to 200. Therefore, the results of different pairs of the I and the J were compared to select the most suitable one. The results of how the accuracy varies with different pairs of the I and the J are shown in Figure 12.

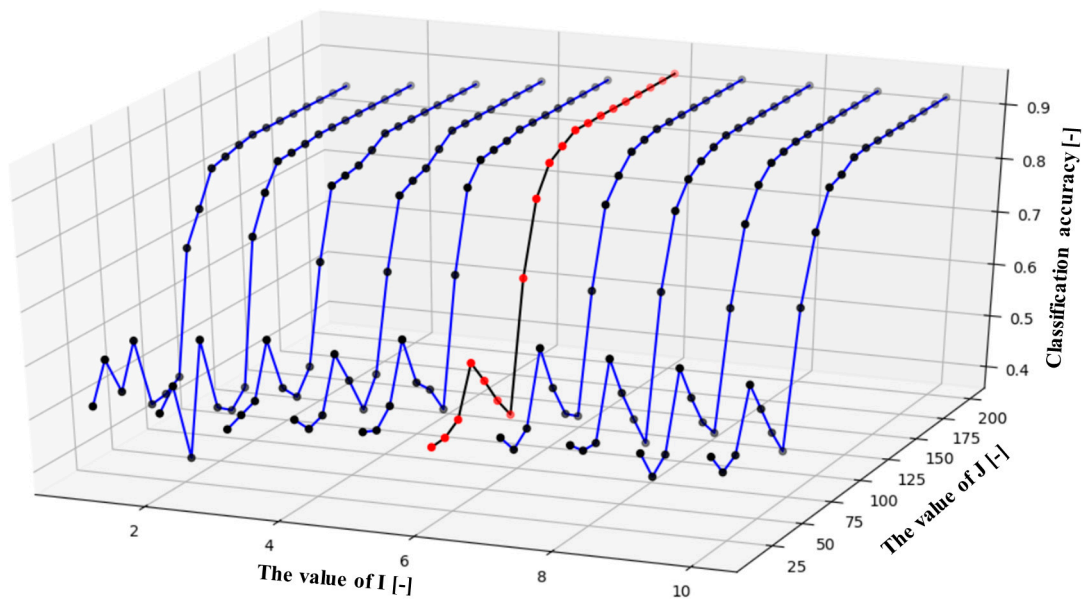


Figure 12. The influence of different I and J on the classification accuracy.

As shown in Figure 12, different pairs of the I and the J are compared to find the optimal choices. In Figure 12, the black curve with red points is the curve where the highest accuracy 90.5% occurs, of which the values of the I and the J are 6 and 120, respectively. Roughly, except for this line, the area of Figure 12 can be divided into three parts, which shows different situations of how the I and the J vary to influence the classification accuracy.

The first part is the lower part where the J is less than 80, no matter what value the I is, the classification accuracy was lower than 60%. It can be considered that when the J is lower than 80, the obtained motion matrix X does not contain enough motion factors for classification. This part is called the “lack of fitting” part.

The second part is the left top corner where the value of the I is less than 6 and the J is more than 120. In this case, accuracy was between 80% and 90%. It can be considered that the user factors cannot be completely separated from the EMG data; in other words, the obtained motion matrix X still had some user factors, which cause the accuracy was less than 90%. This part is called the “less suitable fitting” part.

The third part is the right top corner where the value of the I is more than 6 and the J is more than 120. In this case, the accuracy was almost the same which was from 87.5% to 90%. It can be concluded that almost all the user factors and the motion factors were well separated so that increasing the values of the I and the J had little influence on the accuracy, which is called as the “algorithm converged” part.

Therefore, in this research, the value of the I and the J were set at 6 and 120, respectively. The extracted motion matrix factor values are shown in Figure 13. The motion matrix factor values were input into the LSTM for classification. The twenty-fold cross-validation with each time changing the testing participant until all the participants were tested. The learning rate was set to 0.0001, and the iteration time was set to 500. With the bilinear model applied, the accuracy of 20 participants, the results of the participant one, the results of 20 participants, and the accuracy of 20 motions are shown from Tables 7–9.

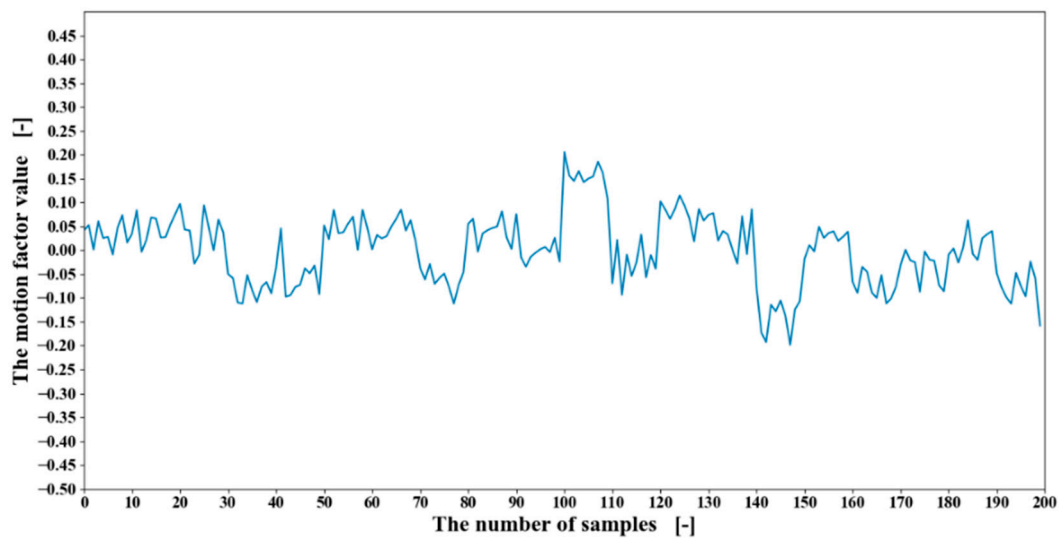


Figure 13. The extracted motion matrix factor values of participant one by the bilinear model.

Table 7. The classification accuracy of 20 participants with the bilinear model.

		Predicted Label																			
		M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20
Actual Label	M1	196	0	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0	1	0	0
	M2	0	197	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0
	M3	0	1	198	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	M4	0	0	1	194	1	0	0	1	0	0	0	1	0	0	0	0	2	0	0	0
	M5	0	0	0	0	194	0	0	0	0	0	0	2	0	0	0	0	0	2	0	2
	M6	0	1	0	0	0	197	0	0	0	1	0	0	0	0	0	1	0	0	0	0
	M7	0	0	0	0	0	1	195	0	0	1	0	1	0	0	0	1	0	0	0	1
	M8	0	0	0	0	0	1	0	199	0	0	0	0	0	0	0	0	0	0	0	0
	M9	0	0	0	0	0	0	0	0	199	0	0	0	0	0	0	0	1	0	0	0
	M10	0	1	1	0	1	3	1	0	1	188	0	0	1	1	0	0	0	0	2	0
	M11	0	0	0	0	0	0	0	0	0	0	196	0	0	1	0	1	0	0	1	1
	M12	0	2	0	0	0	0	0	1	0	0	0	196	0	0	0	0	0	1	0	0
	M13	0	0	1	0	0	0	0	0	1	0	0	0	198	0	0	0	0	0	0	0
	M14	0	0	0	0	0	0	0	0	1	0	0	0	1	198	0	0	0	0	0	0
	M15	0	0	1	0	1	0	0	0	0	0	0	0	1	0	193	0	0	2	0	2
	M16	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	198	0	0	0	1
	M17	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	194	1	1	1
	M18	0	0	1	0	1	1	0	0	0	0	0	0	0	0	2	0	1	194	0	0
	M19	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	1	196	0
	M20	0	0	0	0	1	1	0	0	0	0	0	1	0	1	4	1	0	0	1	190

Table 8. The classification accuracy of 20 motions with the bilinear model.

Motion Name	Accuracy	Motion Name	Accuracy
M1: How are you?	98.0%	M11: Where is the store?	98.0%
M2: Nice to meet you.	98.5%	M12: How can I get food?	98.0%
M3: See you later.	99.0%	M13: How much does it cost?	99.0%
M4: That's what I mean.	97.0%	M14: Yes, thank you.	99.0%
M5: I don't understand.	97.0%	M15: I am sorry.	96.5%
M6: What is your name?	98.5%	M16: Where is the hospital?	99.0%
M7: Where are you from?	97.5%	M17: I don't feel good.	97.0%
M8: What happens?	99.5%	M18: Please help me.	97.0%
M9: What is wrong?	99.5%	M19: Please write it.	98.0%
M10: Please call 911.	94.0%	M20: I love you.	95.0%

Table 9. The classification accuracy of 20 participants with the bilinear model.

Name	Accuracy	Name	Accuracy
Participant 1	98.5%	Participant 11	95.0%
Participant 2	98.0%	Participant 12	100.0%
Participant 3	99.0%	Participant 13	95.0%
Participant 4	100.0%	Participant 14	98.0%
Participant 5	96.5%	Participant 15	98.5%
Participant 6	98.5%	Participant 16	94.5%
Participant 7	99.0%	Participant 17	95.5%
Participant 8	97.5%	Participant 18	98.0%
Participant 9	98.5%	Participant 19	99.0%
Participant 10	98.0%	Participant 20	98.0%

5. Discussion

To demonstrate the effectiveness of the bilinear model, the comparison of RMS feature values and motion factors from the other two participants were conducted as shown in Figures 14 and 15. In Figure 14, graph (a) and graph (b) are of the same EMG data channel and the same feature, but from two participants. As shown in Figure 14, the RMS feature values of the two participants differ greatly. In contrast, after the introduction of the bilinear model, as shown in Figure 15, almost all the interferences of user factors are removed so that the values of the motion factors are much more similar. Although there are some differences between the motion factor values, it can be considered that the motion factors extracted from 19 participants are a little limited to represent the others' motions.

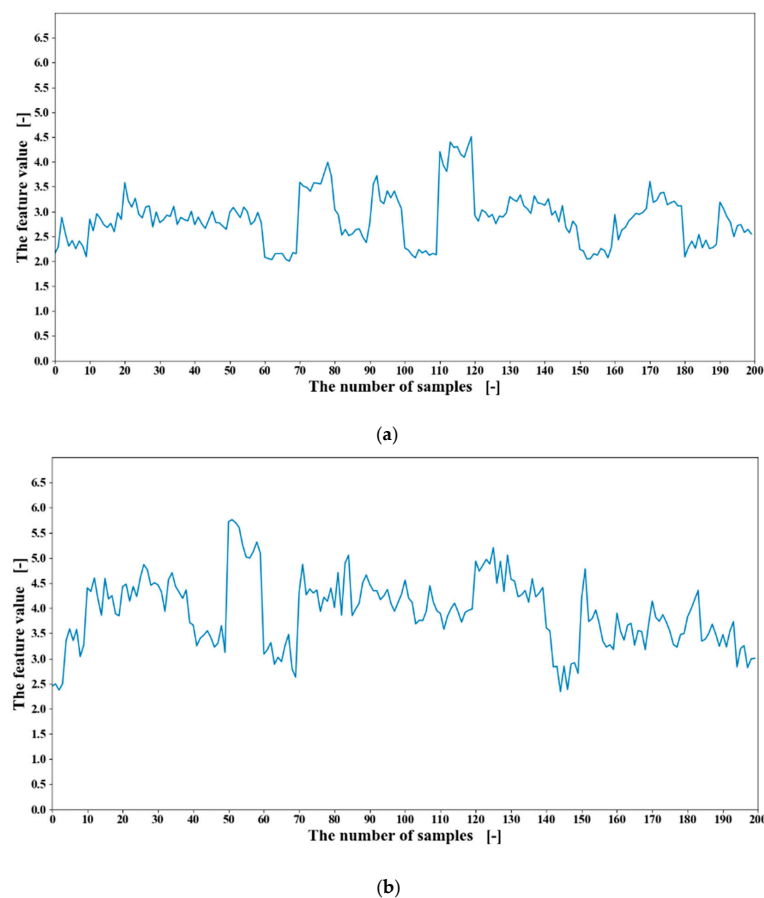
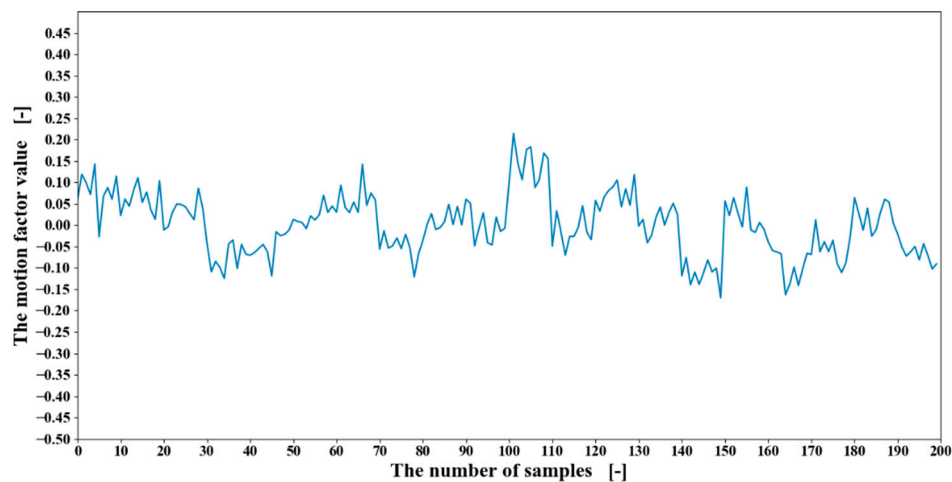
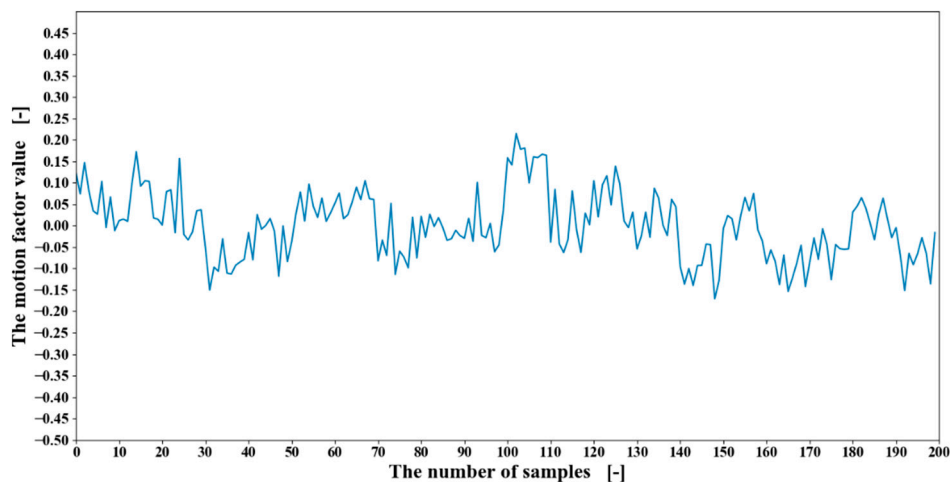


Figure 14. The RMS feature values of 20 motions from different participants:(a) participant four; (b) participant five.



(a)



(b)

Figure 15. The motion matrix factor values of 20 motions from different participants; (a) participant four; (b) participant five.

In Table 6, participant three has the highest accuracy of 71.5%, and participant ten has the lowest accuracy of 46.5%. The accuracy difference between the two participants is 25%, which can indicate the existence of the individual difference. In Table 9, participant four and twelve have the highest accuracy of 100%, and participant sixteen has the lowest accuracy of 94.5%. The accuracy difference among the participants has decreased to 5.5%. Moreover, the average accuracy has increased to 99.7%. Compared with the results of Tables 6 and 9, by applying the bilinear model algorithm, the influence of the individual difference has been largely decreased, which shows that the bilinear model is very effective in the decomposition of user factors and motion factors.

In Table 5, motion six has the highest accuracy of 63.0%, which means that the motions are barely recognized. In Table 8, motion eight and nine have the highest accuracy of 99.5%, and motion ten has the lowest accuracy of 94.0%. The accuracy difference is 5.5%. It can be concluded that almost all the 20 motions are well classified. Misjudgments can be considered that they are mostly caused by the similarities among the 20 motions. For motion ten, the reason it has more misjudgments than the other motions is that it includes more sequential gestures than the others. Another reason can be that the user factors cannot be completely removed and the motion factors are just representative among the 19 participants, which means there are still some differences between the training motion

matrix and the testing motion matrix. If more motion data can be obtained from different people, the influence of the individual difference will be no longer a significant problem that has influence on the classification accuracy.

The Myo band has a sampling rate of 200 Hz, which in some cases, is not enough to obtain all the information of the EMG signal. However, in this research, with an accuracy of 99.7%, it is reasonable to address that this limitation of sampling rate has little impact on this system.

Table 10 shows the performance and characteristics of our proposed system and other studies using the Myo band as a sensor device [61]. Even though the performance metrics of these studies cannot be compared directly due to different experiment settings, they are helpful for qualitative comparisons. Among them, the studies of [43,62] had the high accuracy more than 99%; however, the number of recognized gestures was less than 10, and [62] needed about 1 s to perform the recognition. On the other hand, although the study of [42] needed only 3 ms, the accuracy is 85.1%. Compared with other studies, our proposed system can recognize 20 sign language motions with 97.7% accuracy among 20 participants, and also can perform a real-time recognition with a delay time of less than 50 ms on a PC platform with Intel Core-i7 3.2 GHz and no GPU. Therefore, our system has the potential to be widely applied, and may be implemented into smart phones to realize a real-time daily conversation system by hand gestures for hearing-impaired people.

Table 10. Comparison with other studies using a Myo band.

Study	RTP (ms)	Gestures	Duration (s)	Participants	Repetition	Classifier	Accuracy (%)
Savur [22]	NI	27	2	10	20	SVM	60.9
Hu [34]	NI	52	5	27	10	LCNN	87.0
Kerber [41]	500	5	NI	14	NI	SVM	95.0
Chung [42]	3	5	5	120	50	ANN	85.1
Raurale [43]	4.5/8.8	9	5	10	20	RBF	99.0
Zhang [61]	200	21	2	13	30	GRU	89.6
Nasri [62]	940	6	10	35	195	GRU	99.8
Ali [63]	NI	18	5	40	6	LSTM	89.5
He [64]	400	52	5	27	10	LSTM	75.5
Ours	50	20	3	20	10	BL + LSTM	97.7

Note: RTP represents real-time performance. NI means the corresponding term is not indicated in the paper clearly. LCNN is the combination of LSTM and CNN. ANN is an artificial neural network. RBF is a radial basis function neural network. GRU means gated recurrent units. BL means a bilinear model.

6. Conclusions

This paper presented a user-independent sign language motion recognition system based on the electromyography signal for both helping hearing-impaired people communicate with others more easily in their daily life and training normal people to understand the sign language motions. This proposed motion recognition system could recognize 20 meaningful and widely used ASL motions with high accuracy.

In this paper, the characteristic of the EMG signal was analyzed and utilized for motion recognition. The EMG signal itself is a strong indicator representing the muscle movements; however, it has obvious individual differences among multiple people. Therefore, in this research, the bilinear model algorithm was applied to obtain motion factors for classification. With the introduction of the bilinear model, the interferences of user factors were largely decreased and the motion factors were extracted for classification. Finally, the LSTM was used as the classifier of motions. Moreover, the permutation importance of the features was performed to select the most important features to reduce computation time-consuming. As a result, the LSTM with the bilinear model could realize real-time hand gesture recognition with very high accuracy among 20 participants.

Author Contributions: The contributions to this article include: conceptualization, S.T. and H.L.; methodology, S.T.; software, H.L.; validation, H.L.; formal analysis, H.L.; investigation, S.T.; resources, J.O.; data curation, H.L.; writing—original draft preparation, J.O.; writing—review and editing, S.T. and J.O.; visualization, H.L.; supervision, S.T.; project administration, S.T.; funding acquisition, S.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

References

- World Health Organization Website. Available online: <https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss> (accessed on 9 August 2020).
- Senghas, R.J.; Monaghan, L. SIGNS OF THEIR TIMES: Deaf Communities and the Culture of Language. *Annu. Rev. Anthropol.* **2002**, *31*, 69–97. [\[CrossRef\]](#)
- Galea, L.C.; Smeaton, A.F. Recognising Irish Sign Language Using Electromyography. In Proceedings of the 2019 International Conference on Content-Based Multimedia Indexing, Dublin, Ireland, 4–6 September 2019; pp. 1–4.
- Lucas, C. *The Sociolinguistics of Sign Languages*; Cambridge University Press: Cambridge, UK, 2001; ISBN 9780521794749.
- Efthimiou, E.; Fotinea, S.-E. An environment for deaf accessibility to education content. In Proceedings of the International Conference on ICT & Accessibility (GSRT, M3. 3, id 35), Hammamet, Tunisia, 12–14 April 2007; pp. 12–14.
- Steinberg, A.; Sullivan, V.; Loew, R. Cultural and linguistic barriers to mental health service access: The deaf consumer’s perspective. *Am. J. Psychiatry* **1998**, *155*, 982–984. [\[CrossRef\]](#)
- Meurant, L.; Sinté, A.; Herreweghe, M.V.; Vermeerbergen, M. Sign language research, uses and practices: A Belgian perspective. In *Sign Language Research, Uses and Practices*; Meurant, L., Sinté, A., van Herreweghe, M., Vermeerbergen, M., Eds.; Mouton De Gruyter: Berlin, Germany, 2013; Volume 1, pp. 1–14.
- Chuan, C.-H.; Regina, E.; Guardino, C. American Sign Language Recognition Using Leap Motion Sensor. In Proceedings of the 2014 13th International Conference on Machine Learning and Applications, Detroit, MI, USA, 3–6 December 2014; pp. 541–544.
- Smith, R.G.; Nolan, B. Emotional facial expressions in synthesised sign language avatars: A manual evaluation. *Univers. Access Inf. Soc.* **2016**, *15*, 567–576. [\[CrossRef\]](#)
- Hayek, H.E.; Nacouzi, J.; Mosbeh, P.O.B.Z. Sign to Letter Translator System using a Hand Glove. In Proceedings of the Third International Conference on e-Technologies and Networks for Development, Beirut, Lebanon, 29 April–1 May 2014; pp. 146–150.
- Savur, C.; Sahin, F. Real-Time American Sign Language Recognition System Using Surface EMG Signal. In Proceedings of the 2015 IEEE 14th International Conference on Machine Learning and Applications, Miami, FL, USA, 9–11 December 2015; pp. 497–502.
- Farulla, G.A.; Russo, L.O.; Pintor, C.; Pianu, D.; Micotti, G.; Salgarella, A.R.; Camboni, D.; Controzzi, M.; Cipriani, C.; Oddo, C.M.; et al. Real-Time Single Camera Hand Gesture Recognition System for Remote Deaf-Blind Communication. In Proceedings of the International Conference on Augmented and Virtual Reality, Lecce, Italy, 17–20 September 2014; Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer: Cham, Switzerland, 2014; pp. 35–52, ISBN 978-3-319-13968-5.
- Cyber Gloves Website. Available online: <http://www.cyberglovesystems.com/> (accessed on 9 August 2020).
- Lu, G.; Shark, L.-K.; Hall, G.; Zeshan, U. Immersive manipulation of virtual objects through glove-based hand gesture interaction. *Virtual Real.* **2012**, *16*, 243–252. [\[CrossRef\]](#)
- Raghavan, A.; Joseph, S. EMG analysis and control of artificial arm. *Int. J. Cybern. Inform.* **2016**, *5*, 317–327. [\[CrossRef\]](#)
- Saridis, G.N.; Gootee, T.P. EMG Pattern Analysis and Classification for a Prosthetic Arm. *IEEE Trans. Biomed. Eng.* **1982**, *BME-29*, 403–412. [\[CrossRef\]](#)

17. Shi, J.; Dai, Z. Research on Gesture Recognition Method Based on EMG Signal and Design of Rehabilitation Training System. In Proceedings of the IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference, Chongqing, China, 12–14 October 2018; pp. 835–838.
18. Sathiyarayanan, M.; Rajan, S. Myo armband for physiotherapy healthcare: A case study using gesture recognition application. In Proceedings of the 2016 8th International Conference on Communication Systems and Networks (COMSNETS), Bangalore, India, 5–10 January 2016; pp. 1–6.
19. Sathiyarayanan, M.; Mulling, T. Map navigation using hand gesture recognition: A case study using myo connector on apple maps. *Procedia Comput. Sci.* **2015**, *58*, 50–57. [\[CrossRef\]](#)
20. Lu, Z.; Chen, X.; Li, Q.; Zhang, X.; Zhou, P. A hand gesture recognition framework and wearable gesture-based interaction prototype for mobile devices. *IEEE Trans. Hum. Mach. Syst.* **2014**, *44*, 293–299. [\[CrossRef\]](#)
21. Muhammad, Z.U.R.; Asim, W.; Syed, O.G.; Mads, J.; Imran, K.N.; Mohsin, J.; Dario, F.; Ernest, N.K. Multiday EMG-Based Classification of Hand Motions with Deep Learning Techniques. *Sensors* **2018**, *18*, 2497.
22. Savur, C.; Sahin, F. American Sign Language Recognition system by using surface EMG signal. In Proceedings of the 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Budapest, Hungary, 9–12 October 2016; pp. 2872–2877.
23. Pigou, L.; Dieleman, S.; Kindermans, P.; Schrauwen, B. *Sign Language Recognition Using Convolutional Neural Networks*; Springer: Cham, Switzerland, 2015; pp. 572–578. Available online: <https://biblio.ugent.be/publication/5796137> (accessed on 28 August 2020).
24. Shin, S.; Baek, Y.; Lee, J.; Eun, Y.; Son, S.H. Korean sign language recognition using EMG and IMU sensors based on group-dependent NN models. In Proceedings of the 2017 IEEE Symposium Series on Computational Intelligence (SSCI), Honolulu, HI, USA, 27 November–1 December 2017; pp. 1–7.
25. Hu, X.; Nenov, V. Multivariate AR modeling of electromyography for the classification of upper arm movements. *Clin. Neurophysiol.* **2004**, *115*, 1276–1287. [\[CrossRef\]](#)
26. Zivanovic, M. Time-Varying Multicomponent Signal Modeling for Analysis of Surface EMG Data. *IEEE Signal Process. Lett.* **2014**, *21*, 692–696. [\[CrossRef\]](#)
27. Wang, P.; Wang, Y.; Ru, F.; Wang, P. Develop a home-used EMG sensor system to identify pathological gait with less data via frequency analysis. *Rev. Sci. Instrum.* **2019**, *90*, 043113. [\[CrossRef\]](#)
28. Karlsson, S.; Gerdle, B. Mean frequency and signal amplitude of the surface emg of the quadriceps muscles increase with increasing torquea study using the continuous wavelet transform. *J. Electromyogr. Kinesiol.* **2001**, *11*, 131–140. [\[CrossRef\]](#)
29. Ismail, A.R.; Asfour, S.S. Continuous wavelet transform application to EMG signals during human gait, Conference Record of Thirty-Second Asilomar Conference on Signals. *Syst. Comput.* **1998**, *1*, 325–329.
30. Alkan, A.; Günay, M. Identification of EMG signals using discriminant analysis and SVM classifier. *Expert Syst. Appl.* **2012**, *39*, 44–47. [\[CrossRef\]](#)
31. Arvind, T.; Elizabeth, T.; Enrico, C.; Bastien, B.; Thierry, P.; Eleni, V. An Ensemble Analysis of Electromyographic Activity during Whole Body Pointing with the Use of Support Vector Machines (SVM Analysis of EMG Activity from Complex Movement). *PLoS ONE* **2011**, *6*, e20732.
32. Alberto, D.B.; Emanuele, G.; Giorgio, C.; Angelo, D.; Rinaldo, S.; Eugenio, G.; Loredana, Z. NLR, MLP, SVM, and LDA: A comparative analysis on EMG data from people with trans-radial amputation. *J. Neuroeng. Rehabil.* **2017**, *14*, 82.
33. Andrés, J.; Marco, B.; Elisa, M. Real-Time Hand Gesture Recognition Using Surface Electromyography and Machine Learning: A Systematic Literature Review. *Sensors* **2020**, *20*, 2467.
34. Hu, Y.; Wong, Y.; Wei, W.; Du, Y.; Kankanhalli, M.S.; Geng, W. A novel attention-based hybrid CNN-RNN architecture for sEMG-based gesture recognition. *PLoS ONE* **2018**, *13*, e0206049. [\[CrossRef\]](#)
35. Ameri, A.; Akhaee, M.A.; Scheme, E.; Englehart, K. Regression convolutional neural network for improved simultaneous EMG control. *J. Neural Eng.* **2019**, *16*, 036015. [\[CrossRef\]](#)
36. Mane, S.M.; Kambli, R.A.; Kazi, F.S.; Singh, N.M. Hand motion recognition from single channel surface EMG using wavelet & artificial neural network. *Procedia Comput. Sci.* **2015**, *49*, 58–65.
37. Tavakoli, M.; Benussi, C.; Lourenco, J.L. Single channel surface EMG control of advanced prosthetic hands: A simple, low cost and efficient approach. *Expert Syst. Appl.* **2017**, *79*, 322–332. [\[CrossRef\]](#)
38. Clancy, E.; Morin, E.; Merletti, R. Sampling, noise-reduction and amplitude estimation issues in surface electromyography. *J. Electromyogr. Kinesiol.* **2002**, *113*, 1–16. [\[CrossRef\]](#)

39. Li, G.; Li, Y.; Zhang, Z.; Geng, Y.; Zhou, R. Selection of sampling rate for EMG pattern recognition based prosthesis control. In Proceedings of the 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, Buenos Aires, Argentina, 31 August–4 September 2010; Volume 2010, pp. 5058–5061.
40. Winter, D.A. *Biomechanics and Motor Control of Human Movement*; John Wiley & Sons: Hoboken, NJ, USA, 2009.
41. Kerber, F.; Puhl, M.; Krüger, A. User-Independent Real-Time Hand Gesture Recognition Based on Surface Electromyography. In Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services, Vienna, Austria, 4–7 September 2017; p. 36.
42. Chung, E.A.; Benalcázar, M.E. Real-Time Hand Gesture Recognition Model Using Deep Learning Techniques and EMG Signals. In Proceedings of the 27th European Signal Processing Conference (EUSIPCO), Coruña, Spain, 2–6 September 2019; pp. 1–5.
43. Raurale, S.; McAllister, J.; del Rincon, J.M. EMG wrist-hand motion recognition system for real-time Embedded platform. In Proceedings of the 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 12–17 May 2019; pp. 1523–1527.
44. Das, A.K.; Laxmi, V.; Kumar, S. Hand Gesture Recognition and Classification Technique in Real-Time. In Proceedings of the 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN), Tamil Nadu, India, 30–31 March 2019; pp. 1–5.
45. Luo, X.Y.; Wu, X.Y.; Chen, L.; Hu, N.; Zhang, Y.; Zhao, Y.; Hu, L.T.; Yang, D.D.; Hou, W.S. Forearm Muscle Synergy Reducing Dimension of the Feature Matrix in Hand Gesture Recognition. In Proceedings of the 3rd International Conference on Advanced Robotics and Mechatronics (ICARM), Singapore, 18–20 July 2018; pp. 691–696.
46. Zanghieri, M.; Benatti, S.; Burrello, A.; Kartsch, V.; Conti, F.; Benini, L. Robust Real-Time Embedded EMG Recognition Framework Using Temporal Convolutional Networks on a Multicore IoT Processor. *IEEE Trans. Biomed. Circuits Syst.* **2019**, *14*, 244–256. [[CrossRef](#)]
47. Divya, B.; Delpha, J.; Badrinath, S. Public speaking words (Indian sign language) recognition using EMG. In Proceedings of the 2017 International Conference on Smart Technologies for Smart Nation (SmartTechCon), Bangalore, India, 17–19 August 2017; pp. 798–800.
48. Sheng, X.; Lv, B.; Guo, W.; Zhu, X. Common spatial-spectral analysis of EMG signals for multiday and multiuser myoelectric interface. *Biomed. Signal Process. Control* **2019**, *53*, 101572. [[CrossRef](#)]
49. Yang, C.; Xi, X.; Chen, S.; Miran, S.M.; Hua, X.; Luo, Z. SEMG-based multifeatures and predictive model for knee-joint-angle estimation. *AIP Adv.* **2019**, *9*, 095042. [[CrossRef](#)]
50. Zhang, L.; Shi, Y.; Wang, W.; Chu, Y.; Yuan, X. Real-time and user-independent feature classification of forearm using EMG signals. *J. Soc. Inf. Disp.* **2019**, *27*, 101–107. [[CrossRef](#)]
51. Khushaba, R.N. Correlation Analysis of Electromyogram Signals for Multiuser Myoelectric Interfaces. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2014**, *22*, 745–755. [[CrossRef](#)]
52. Tenenbaum, J.B.; Freeman, W.T. Separating style and content with bilinear models. *Neural Comput.* **2000**, *12*, 1247–1283. [[CrossRef](#)]
53. Matsubara, T.; Morimoto, J. Bilinear Modeling of EMG Signals to Extract User-Independent Features for Multiuser Myoelectric Interface. *IEEE Trans. Biomed. Eng.* **2013**, *60*, 2205–2213. [[CrossRef](#)]
54. Wang, T.; Hou, W. Analysis of the sEMG bilinear model for the control of hand prosthesis. *Chin. J. Sci. Instrum.* **2014**, *35*, 1907.
55. Frigo, C.; Crenna, P. Multichannel SEMG in clinical gait analysis: A review and state-of-art. *Clin. Biomech.* **2009**, *24*, 236–245. [[CrossRef](#)] [[PubMed](#)]
56. Liu, J.; Zhong, L.; Wickramasuriya, J. A real-time EMG pattern recognition system based on linear-nonlinear feature projection for a multifunction myoelectric hand. *IEEE Trans. Biomed. Eng.* **2006**, *53*, 657–675.
57. Huang, N.; Lu, G.; Xu, D. A Permutation Importance-Based Feature Selection Method for Short-Term Electricity Load Forecasting Using Random Forest. *Energies* **2016**, *9*, 767. [[CrossRef](#)]
58. Arjunan, S.P.; Kumar, D.K.; Naik, G.R. Fractal feature of sEMG from Flexor digitorum superficialis muscle correlated with levels of contraction during low-level finger flexions. In Proceedings of the 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, Buenos Aires, Argentina, 31 August–4 September 2010; pp. 4614–4617.
59. Collí, A.; Guillermo, J. Implementation of User-Independent Hand Gesture Recognition Classification Models Using IMU and EMG-Based Sensor Fusion Techniques. Master's Thesis, Western University, London, ON, Canada, 2019.

60. Li, X.; Wu, X. Constructing long short-term memory based deep recurrent neural networks for large vocabulary speech recognition. In Proceedings of the 2015 IEEE International Conference on Acoustics, Speech and Signal Processing, Brisbane, Australia, 19–24 April 2015; pp. 4520–4524.
61. Zhang, Z.; He, C.; Kuo, Y. A Novel Surface Electromyographic Signal-Based Hand Gesture Prediction Using a Recurrent Neural Network. *Sensors* **2020**, *20*, 3994. [[CrossRef](#)]
62. Nasri, N.; Orts-Escolano, S.; Gomez-Donoso, F.; Cazorla, M. Inferring Static Hand Poses from a Low-Cost Non-Intrusive sEMG Sensor. *Sensors* **2019**, *19*, 371. [[CrossRef](#)]
63. Ali, S. Gated Recurrent Neural Networks for EMG-based Hand Gesture Classification. A Comparative Study. In Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, USA, 18–21 July 2018; pp. 1094–1097.
64. He, Y.; Fukuda, O.; Bu, N.; Okumura, H.; Yamaguchi, N. Surface EMG Pattern Recognition Using Long Short-Term Memory Combined with Multilayer Perceptron. In Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, USA, 18–21 July 2018; pp. 5636–5639.

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).