

Article

Two-Step Biometrics Using Electromyogram Signal Based on Convolutional Neural Network-Long Short-Term Memory Networks

Jin-Su Kim ¹, Min-Gu Kim ¹  and Sung-Bum Pan ^{1,2,*} 

¹ IT Research Institute, Chosun University, 309 Pilmun-daero, Dong-gu, Gwangju 61452, Korea; babotn5@gmail.com (J.-S.K.); happy9433@chosun.ac.kr (M.-G.K.)

² Interdisciplinary Program in IT-Bio Convergence System, Chosun University, 309 Pilmun-daero, Dong-gu, Gwangju 61452, Korea

* Correspondence: sbpan@chosun.ac.kr

Abstract: Electromyogram (EMG) signals cannot be forged and have the advantage of being able to change the registered data as they are characterized by the waveform, which varies depending on the gesture. In this paper, a two-step biometrics method was proposed using EMG signals based on a convolutional neural network–long short-term memory (CNN-LSTM) network. After preprocessing of the EMG signals, the time domain features and LSTM network were used to examine whether the gesture matched, and single biometrics was performed if the gesture matched. In single biometrics, EMG signals were converted into a two-dimensional spectrogram, and training and classification were performed through the CNN-LSTM network. Data fusion of the gesture recognition and single biometrics was performed in the form of an AND. The experiment used Ninapro EMG signal data as the proposed two-step biometrics method, and the results showed 83.91% gesture recognition performance and 99.17% single biometrics performance. In addition, the false acceptance rate (FAR) was observed to have been reduced by 64.7% through data fusion.

Keywords: biometrics; gesture recognition; user identification; electromyogram signal



Citation: Kim, J.-S.; Kim, M.-G.; Pan, S.-B. Two-Step Biometrics Using Electromyogram Signal Based on Convolutional Neural Network-Long Short-Term Memory Networks. *Appl. Sci.* **2021**, *11*, 6824. <https://doi.org/10.3390/app11156824>

Academic Editor: Juan A. Gómez-Pulido

Received: 9 June 2021

Accepted: 21 July 2021

Published: 25 July 2021

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



Copyright: © 2021 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

Biometrics refers to the verification of identity using an individual's behavioral and physical features. Currently, widely used biometrics include face, fingerprint, and iris recognition as physical features. However, faces, fingerprints, and irises are bodily information without liveness that can be forged using 3D printers, lenses, and silicon, raising security concerns [1]. To compensate for this issue, biometrics studies using biosignals, which are a behavioral feature, are being conducted. Biosignals are signals that measure the flow of microcurrents generated through biological activities, which cannot be replicated outside the body [2]. In addition, biosignals are signals with liveness, thereby having strong security advantages as they cannot be forged and falsified [3]. Examples of biosignals include electromyogram (EMG), electrocardiogram (ECG), and electroencephalogram (EEG) signals. Among them, the EMG signal represents a signal obtained in the form of a voltage value by measuring a microcurrent generated whenever there is a muscle contraction. There are two methods of measuring EMG signals: a needle electrode method and a surface electrode method. The needle electrode method measures an action potential occurring at a point in the muscle by inserting a needle electrode into a muscle. The surface electrode method measures an action potential by placing electrodes to the skin surface. Measurement of the EMG signal using the surface electrode method wherein an electrode is placed to the skin surface is shown in Figure 1 [1]. In terms of the electrode placement position, the signal is measured by placing the electrode at a muscle which change during experimentation or a muscle position that is frequently used in existing studies.

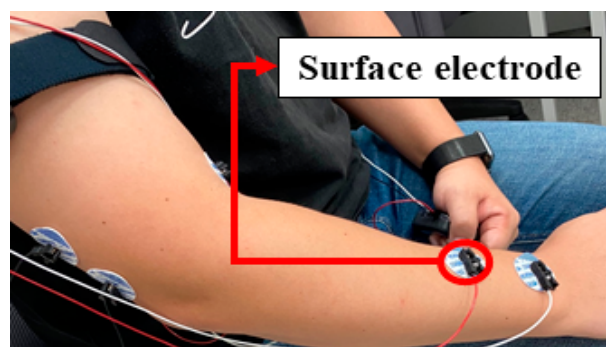


Figure 1. Electromyogram (EMG) signal measurement using the surface electrode method.

Biometrics studies using EMG signals are lacking compared to studies using ECG and EEG signals. However, the EEG signal gets distorted as it passes through the skull, which is a disadvantage while measuring the signal, and the ECG signal is a periodic signal, which is a drawback by itself in that the waveform of the signal cannot be changed if the registered data is hacked. By comparison, the EMG signal can be measured more easily than the EEG signal, and its registration data can be changed distinctly, unlike the ECG signal, using a feature representing a different waveform of a signal for each gesture. EMG signals in existing works have been used as gesture recognition or single biometrics.

In this paper, we proposed a method for using EMG signal-based gesture recognition and single biometrics together. The proposed method is a two-step biometrics method that uses CNN-LSTM network-based EMG signals. The time domain characteristics of the EMG signal and the LSTM network were used to determine whether the user's gesture matches. Single biometrics was performed using the spectrogram and CNN-LSTM network if it matched the registered gesture. Data fusion of gesture recognition and single biometrics was performed in the form of AND in which the false acceptance rate (FAR) was lowered. As a result of the experiment, it was confirmed that the FAR decreased by 64.7%. The remainder of this paper is structured as follows. Existing gesture recognition and biometrics studies using bioinformation are mentioned in Section 2. A two-step biometrics method using EMG signals based on the proposed CNN-LSTM network is described in Section 3. The results of the two-step biometrics experiment using our proposed method are explained in Sections 4 and 5 is the conclusion.

2. Related Works

Most existing works using EMG signals have been conducted on gesture recognition and biometrics from hand muscles measured while performing hand gestures [4–18]. These works include studies using EMG signals of the lower body when a subject is walking [14] and those using EMG signals of the mouth muscles when talking [15,18]. In addition, biometrics studies using multimodal information are being conducted to compensate for the shortcomings of existing biometrics [3,19–21].

2.1. Gesture Recognition Using EMG Signals

EMG signal data can be measured easily on hand gestures; therefore, studies using hand gestures are actively being conducted [4–7]. Features of the EMG signal are extracted from the time and frequency domains. Zhang [4] used the time domain features of EMG signals and conducted a study to classify ten hand gestures using eight LSTM layers by extracting the features of the EMG signal in the time domain. Jaramillo [5] studied the time–frequency domain features of EMG signals. To classify five hand gestures, features were extracted using seventeen feature extraction methods such as root mean square (RMS) and zero crossing (ZC), and classification was conducted using the k-nearest neighbor (KNN). Oh [6] classified three hand motions and three hand gestures that express intention. Data were converted into a continuous wavelet transform (CWT), and biometrics were performed using CNN for time–frequency analysis in EMG signals. Qi [7] extracted features

from the time–frequency domain to classify nine hand gestures such as open palm and hand flexion, and performed normalization to match the range of each feature. Subsequently, the data dimension was reduced using principal component analysis (PCA), and each gesture was classified using a general regression neural network (GRNN). Chen [8] used sliding windows to partition the data and preprocess EMG signals via CWT conversion. Seven hand gestures were classified using CNN-based EMGNet, which consists of five convolutional layers. Asif [9] used a 5–500 Hz band pass filter (BPF) to denoise EMG signals and conducted a study to classify ten hand gestures using a 1D CNN consisting of three convolutional layers.

Zhang [4], who conducted the gesture recognition study, conducted the experiment using EMG signals from similar gestures, such as ‘hi gesture’ and ‘raising hand gesture’. However, other related works have conducted gesture recognition experiments without considering EMG signals when performing similar hand gestures. In addition, the number of gestures used in the experiment in all related works is less than ten, which leads to the problem of fewer gesture classes being classified.

2.2. Biometrics Using EMG Signal

Shreyas [10] studied biometrics using EMG signals of the hand muscles. Individuals were classified using euclidean distance (ED), KNN, and support vector machine (SVM) after feature extraction from EMG signals using PCA, unsupervised discriminant projection (UDP), and class dependent feature analysis (CFA). Shin [11] studied biometrics using two-channel EMG signals for three hand gestures. Noise was removed from the EMG signal by filtering with a 0.5 Hz high pass filter (HPF), a 200 Hz low pass filter (LPF), and a 60 Hz notch filter (NF). Subsequently, the time–frequency features were extracted to classify individuals through artificial neural networks (ANNs), SVMs, and KNNs. Lu [12] studied biometrics using EMG signals for open hand gestures in twenty-one people. EMG signals were converted into CWT for each channel to classify individuals through CNN. Shioji [13] studied gesture recognition and biometrics using three gestures. Noise was removed from the signal using a 20 Hz HPF and a 60 Hz NF; feature extraction and classification were performed using CNN. Lee [14] studied biometrics using muscles other than that of the hand. EMG signals of a subject while walking were used, and noise was removed with a 20–450 Hz BPF and a 60 Hz NF. Time–frequency features were extracted from EMG signals and individuals were classified using linear discriminant analysis (LDA). Morikawa [15] studied biometrics using EMG signals in the mouth muscles. EMG signals were collected while a subject pronounced five Japanese vowels, and noise was removed with a 5–500 Hz BPF and a 60 Hz NF; feature extraction and classification were performed using CNN. Lu [16] studied biometrics using EMG data from open hand gesture. After data conversion using discrete wavelet transform (DWT) and CWT, twenty-one people were classified as CNN-based siamese networks. Li [17] conducted biometrics studies using EMG signals when unlocking patterns. Noise in EMG signals was eliminated with a 5 Hz HPF and a 60 Hz NF, utilizing eleven time domain features and a one-class support vector machine (OCSVM) and local outlier factor (LOF). Khan [18] studied biometrics using speech EMG. Khan used empirical mode decomposition (EMD) to eliminate noise in signals and set areas of interest and utilize time-frequency domain features.

In most related works conducted on single biometrics, biometrics was performed using an EMG signal of only one gesture, indicating that the behavior was limited (except [11,13,15]). In addition, Shreyas [11], Shioji [13], and Morikawa [15] conducted biometrics experiments using up to eight people as subjects. Therefore, there is a need for biometrics experiments that classify more subjects.

2.3. Biometrics Using Multimodal Information

Biometrics using two or more types of bioinformation can improve biometrics performance by using data fusion of different pieces of information, such as feature level fusion and score level fusion [22]. Belgacem [3] conducted a study using feature level fusion,

extracted features from ECG signals using autocorrelation (AC) and fast fourier transformation (FFT), and extracted features from EMG signals using FFT. Features extracted from each biosignal were combined and classified using optimum patch forest (OPF). Liu [19] conducted a study using the score level fusion method. Liu extracted features from the face, fingerprint, and iris with a gabor wavelet and an eigenvector and classified each type of bioinformation to calculate the matching score. The final individual was then classified by combining the classified scores. Kumar [20] conducted a study using palm prints and hand geometry. Each bioinformation type was classified after features were extracted using morphological features and discrete cosine transform (DCT) coefficients, and a combination of each score was used. Rahman [21] conducted a study using EEG signals and faces. Time–frequency features were extracted from EEG signals and classified with backpropagation (BP), and facial features were extracted with a histogram oriented gradient and classified with KNN to calculate the score.

Biometrics studies using multimodal information based on data fusion have the advantage of improving biometrics performance through data fusion. However, bioinformation must be measured using a dedicated measuring equipment. Therefore, there are more equipment needed to measure multimodal information. In addition, there are inconveniences to real world use.

A summary of existing gesture recognition and biometrics studies using bioinformation, including data such as number of channels and feature extraction methods, is shown in Table 1. Most related works using EMG signals were conducted using EMG signals from the hand muscles, and noise was removed using NFs and BPFs via preprocessing. The method of extracting features from data has changed from the handcrafted method to the deep learning method, and studies that fuse data with other bioinformation are being conducted to compensate for the shortcomings of single biometrics. However, most of the related works performed gesture recognition using less than ten gestures, and all biometrics were performed using one gesture. In addition, similar gestures were not considered. EMG signals can be used to fuse gesture recognition information and biometrics information, and an advantage is that FAR can be improved by the data fusion of biometrics information in an AND form when performing biometrics [22]. In this paper, an experiment was performed using a public EMG database (DB), and EMG signal-based two-step biometrics were performed using gesture recognition information to improve FAR. The experiment was conducted using seventeen gestures of twenty subjects, including similar gestures. Two-step biometrics is a single biometrics method performed when the registered gesture and the input gesture match, which preprocesses the EMG signal and then extracts and classifies the features using a deep learning method.

Table 1. Analysis of related works using bioinformation.

Authors	Information	Identification	DB	Channel	Feature	Classification
Zhang, X. et al. [4]	EMG	Gesture	Acquired	8	MAV, VAR, WL, LSTM	LSTM
Jaramillo, A.G. et al. [5]	EMG	Gesture	Acquired	8	RMS, ZC, etc.	KNN
Oh, D.C. et al. [6]	EMG	Gesture	Acquired	8	CWT, CNN	CNN
Qi, J. et al. [7]	EMG	Gesture	Acquired	16	RMS, WL, etc.	GRNN
Chen, L. et al. [8]	EMG	Gesture	Publicly DB	8	CNN	CNN
Asif, A.R. et al. [9]	EMG	Gesture	Acquired	6	CNN	CNN
Shreyas, V. et al. [10]	EMG	Biometrics	Acquired	8	PCA, UDP, CFA	ED, KNN, SVM
Shin, S.H. et al. [11]	EMG	Biometrics	Acquired	2	VAR, Mean Length, ZC, MF	ANN, SVM, KNN
Lu, L. et al. [12]	EMG	Biometrics	Acquired	8	CWT, CNN	CNN

Table 1. Cont.

Authors	Information	Identification	DB	Channel	Feature	Classification
Shioji, R. et al. [13]	EMG	Gesture, Biometrics	Acquired	8	CNN	CNN
Lee, M. et al. [14]	EMG	Biometrics	Acquired	11	RMS, MAV, etc.	LDA
Morikawa, S. et al. [15]	EMG	Biometrics	Acquired	3	CNN	CNN
Lu, L. et al. [16]	EMG	Biometrics	Acquired	8	DWT, CWT, RMS, etc.	CNN, ExtraTrees
Li, Q. et al. [17]	EMG	Biometrics	Acquired	1	MAV, VAR, RMS, etc.	OCSVM, LOF
Khan, M.U. et al. [18]	EMG	Biometrics	Acquired	2	ZC, MF, etc.	SVM
Belgacem, N. et al. [3]	ECG, EMG	Biometrics	Acquired	1	AC, FFT	OPF
Liu, L. et al. [19]	Face, Iris, Finger Print	Biometrics	Acquired Publicly DB	-	Gabor Wavelet, Eigen Vector	PUBDNN
Kumar, A. et al. [20]	Palm Print, Hand Geometry	Biometrics	Acquired Publicly DB	-	Morphological Information, DCT	EN
Rahman, M.W. et al. [21]	EEG, Face	Biometrics	Publicly DB	15	SV, MAV, etc.	BP, KNN
Proposed	EMG	Gesture, Biometrics	Publicly DB	12	LSTM, CNN, MAV, VAR, etc.	LSTM, CNN

3. Materials and Methods

The two-step biometrics method using EMG signals based on the CNN-LSTM network proposed is shown in Figure 2. Noise in the signal was removed, and the EMG signal was then aligned using RMS to perform two-step biometrics using the EMG signal. After preprocessing the EMG signal, gesture recognition was performed using time domain features and the LSTM network to examine whether it matched the registered gesture, and single biometrics was performed if the gesture matched. For single biometrics, a one-dimensional EMG signal was converted into a spectrogram in the form of a two-dimensional image, and features were extracted and trained using the designed CNN-LSTM network. When the EMG signal for the two-step biometrics test was input, the above-mentioned preprocessing, gesture recognition, and spectrogram conversion were performed. Subsequently, the individual was classified using the information trained on the designed network.

3.1. EMG Signal Preprocessing

The EMG signal is in the form of a voltage value obtained by measuring microcurrent generated when a muscle contracts, and it contains considerable gesture information in a low-frequency band (500 Hz or below). The measured EMG signal contains various types of noise. The types of noise included in the EMG signal are baseline wander, power line noise, and external environmental noise [23,24]. In this study, NF was used in the 60 Hz band and BPF was used in the 5–500 Hz band to remove noise. Figure 3 represents changes in EMG signals in the time-frequency domains before and after preprocessing, with blue signals representing EMG signals before preprocessing and red signals representing EMG signals after preprocessing. It can be seen in Figure 3 (in red box) that NF application removed the 60 Hz band and BPF application passed the 5–500 Hz band. The EMG signal from which noise had been removed was divided according to purpose after aligning the EMG signal using RMS prior to performing gesture recognition.

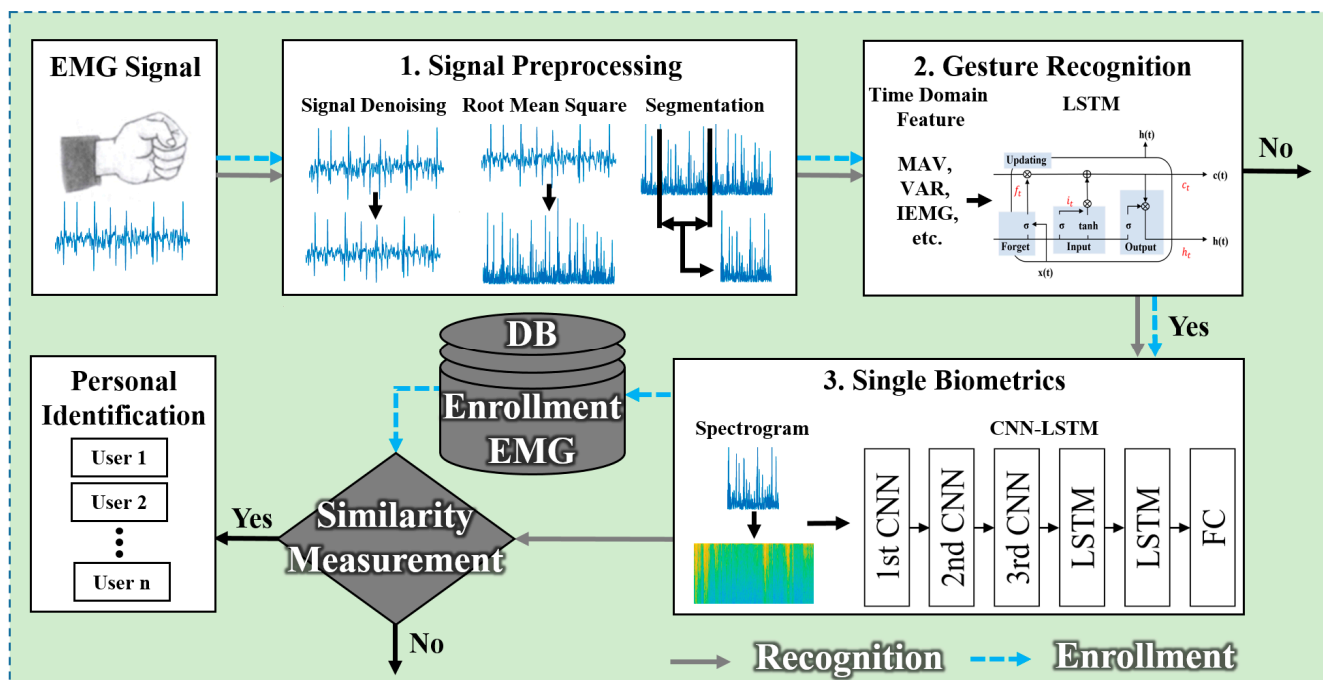


Figure 2. Convolutional neural network-long short-term memory (CNN-LSTM) network-based two-step biometrics flowchart.

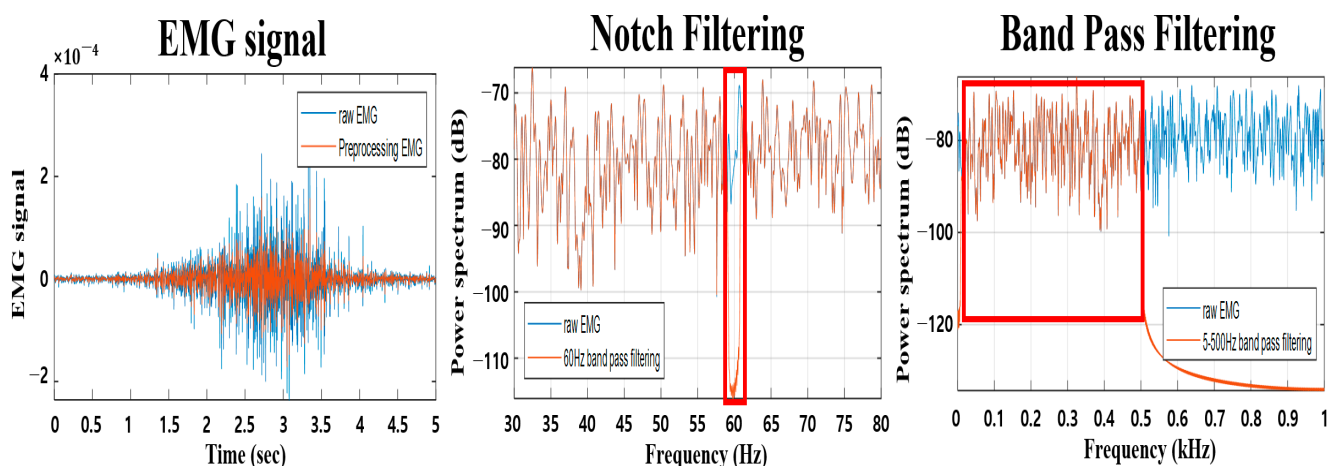


Figure 3. Comparison before and after EMG signal preprocessing in time-frequency domain of a notch filter (NF) and a band pass filter (BPF).

3.2. Gesture Recognition Method Using EMG Signals

For the two-step biometrics methods proposed herein, gesture recognition was performed as shown in Figure 4. Training and classification were performed with an LSTM network after extracting five time domain features from the preprocessed EMG signal, and the time domain features used in the experiment were mean absolute value (MAV), variance (VAR), integrated EMG (IEMG), MAV slope (MAVSLP), and waveform length (WL). MAV can detect muscle contraction, which is a feature widely used in control applications using EMG signals. VAR uses the power of the EMG signal as a feature, representing the power information of the EMG signal. IEMG is a feature used as an index to detect muscle activity, representing the force applied to a muscle. MAVSLP is an improved feature of MAV that can be calculated using the difference between adjacent MAVs, and WL is a feature representing the difference in the length of the waveform of the EMG signal. The

five time domain features can be calculated using the formula shown in Figure 4, where x_i denotes the EMG signal and N denotes the signal length [25,26].

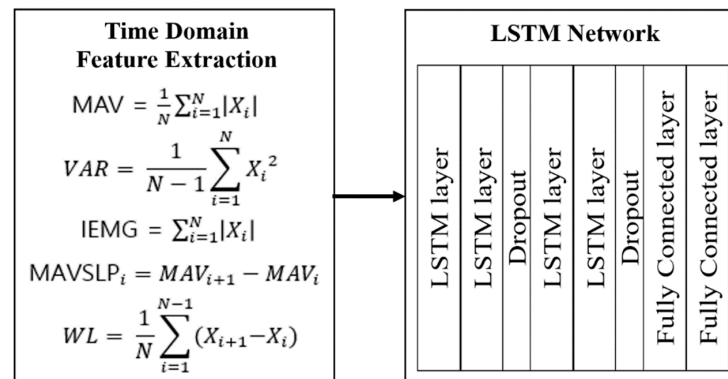


Figure 4. Flowchart of gesture recognition using time domain features and the LSTM network.

Training and classification were performed using an LSTM network after the features were extracted from the time domain of the EMG signal. The LSTM network is characterized by maintaining sequence information in the time domain, thereby an important role in improving classification accuracy [27]. In addition, the LSTM network is a structure that compensates for the disadvantage of the vanishing gradient problem of RNN. Four gates (forget, input, output, updating) can be used to selectively transfer or delete information from sequential inputs to maintain dependencies in the sequence, as shown in Figure 5.

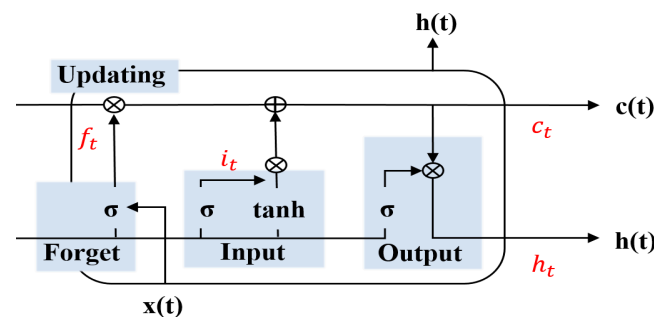


Figure 5. Basic LSTM network structure.

The LSTM network had two transfer parameters, $c(t)$ and $h(t)$, calculated by using four gates. The forget gate was calculated using Equation (1), and the input gate was calculated using Equation (2). Parameters $c(t)$ and $h(t)$ to be transferred to the next LSTM network were calculated using Equations (3) and (4), respectively. In the equation, σ denotes the sigmoid function, $*$ denotes the convolution, and W denotes the weight. In this paper, we performed gesture recognition using an LSTM network consisting of four LSTM layers and two fully connected layers. The layer placement and parameters of the network were selected through experiments, and the designed network deployed one dropout (0.3 probability) for two LSTM networks. An LSTM network uses a many-to-one structure that receives the time domain characteristics of a multi EMG channel as input and outputs the gesture class. Network learning is done by setting adam optimizer, hidden unit 512 and batch size 128, and epoch with 100.

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \quad (1)$$

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \quad (2)$$

$$c_t = f_t * c_{t-1} + i_t * \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (3)$$

$$h_t = \sigma(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \quad (4)$$

3.3. Single Biometrics Method Using EMG Signals

The single biometrics method among the two-step biometrics methods proposed herein used a spectrogram and the CNN-LSTM network. The time domain of the EMG signal represents information on muscle activity, and the frequency domain has the advantage of being immune to noise [1]. Therefore, the one-dimensional EMG signal was converted into a two-dimensional spectrogram to simultaneously analyze time–frequency domain information. The spectrogram was generated using a window function of a fixed length within the time domain of a given signal. In the generated spectrogram, the horizontal axis represented time, and the vertical axis represented frequency information.

In the case of biometrics using the existing EMG signal, a handcrafted method is performed wherein a subject extracts features by using a specific feature extraction algorithm. However, this method lacks versatility and has a problem in that it may not be able to extract features optimized for a signal. To solve this problem, deep learning methods such as CNN and LSTM networks, which have been actively studied recently, can automatically extract optimal features from data without extracting features via the handcrafted method [28].

A CNN-LSTM network was designed herein to perform single biometrics using EMG signals as shown in Figure 6. In this paper, a single biometrics was performed using a network of three convolutional layers, three pooling layers, two LSTM layers, and two fully connected layers. Batch normalization was performed between each convolutional layer and the pooling layer, and a dropout (0.3 probability) was placed after three convolutional layers. The layer placement and parameters of the network were selected through experiments, and a rectified linear unit (ReLU) was used as the activation function. Network learning was done by setting sgdm optimizer, batch size 50, and epoch with 100. The spectrogram input to the network was normalized to 256×256 , and the output data size, padding, stride, and number of filters used in the experiment are shown in Table 2. The stride and padding were set to 1 and 2, respectively, to ensure that the data size remained unchanged in the convolutional layer, and the stride and padding were set to 2 and 0, respectively, to reduce the data size in the pooling layer.

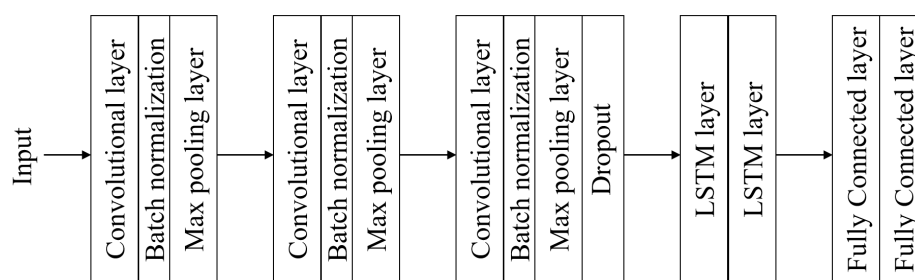


Figure 6. CNN-LSTM network structure for single biometrics.

Table 2. Designed CNN-LSTM network parameters.

Layer	Input Size	Output Size	Stride	Padding	Filter Count (Hidden Unit)
Input	-	$256 \times 256 \times 3$	-	-	-
Conv1	$256 \times 256 \times 3$	$256 \times 256 \times 20$	1	2	20
Pool1	$256 \times 256 \times 20$	$128 \times 128 \times 20$	2	-	-
Conv2	$128 \times 128 \times 20$	$128 \times 128 \times 20$	1	2	20
Pool2	$128 \times 128 \times 20$	$64 \times 64 \times 20$	2	-	-
Conv3	$64 \times 64 \times 20$	$64 \times 64 \times 20$	1	2	20
Pool3	$64 \times 64 \times 20$	$32 \times 32 \times 20$	2	-	-
LSTM1	20,480	1024	-	-	1024
LSTM2	1024	1024	-	-	1024
FC1	1024	512	-	-	-
FC2	512	20	-	-	-

4. Results and Discussion

The publicly available EMG signal dataset Ninapro DB2 was used to verify the performance of the two-step biometrics method using EMG signals based on the CNN-LSTM network proposed herein. The Ninapro DB2 comprises hand gestures performed by a subject, and hand gestures can be divided into three groups (groups 1–3). The data used in the experiment were EMG signals of seventeen gestures measured within group 1, and the experiment was conducted with twenty subjects. EMG signal measurements were collected at twelve channels using Delsys Trigno Wireless EMG equipment and measured at a 2000 Hz sampling rate. When measuring the EMG data, each movement lasted for five seconds, and a three seconds rest period was performed between movements. The subjects were asked to move the right hand repeatedly while measuring the EMG data and each movement was repeated six times [29].

Experimental performance was evaluated using accuracy, recall, and precision, and the above evaluation indicators can be calculated as true positive (TP), true negative (TN), false positive (FP), and false negative (FN). Accuracy refers to the performance of when the model's true label predicts true and false label predicts false, which is an evaluation index representing the most intuitive performance of the model. The method to obtain accuracy is shown in Equation (5). Recall refers to the ratio of the model's true label predicts true or false, which can be calculated through TP and FN as shown in Equation (6). Precision refers to the actual ratio of true classifications among all classifications by the trained model, which can be calculated through TP and FP as shown in Equation (7).

$$\text{accuracy} = (TP + TN) / (TP + TN + FP + FN) \quad (5)$$

$$\text{recall} = TP / (TP + FN) \quad (6)$$

$$\text{precision} = TP / (TP + FP) \quad (7)$$

The data used in the two-step biometrics experiment were recorded from each gesture repeated six times by twenty subjects. However, there is a problem that the amount of training and test data were insufficient when using an LSTM network for gesture recognition and a CNN-LSTM network for single biometrics. Therefore, three-fold cross-validation (each subject used data from four training sets and two test sets) was performed for gesture recognition (1:N comparison) and single biometrics (1:N comparison).

The results of the gesture recognition experiment using time domain features and LSTM network are as shown in Table 3. As a result of the experiment, the basic gestures of the seventeen hands showed an average of 83.91% accuracy, 91.26% recall, and 91.15% precision performance. In terms of the recognition performance by gesture, the 16th gesture was the best classified at 99.62%, and the 11th gesture was the worst classified. The 11th gesture was most often misclassified as the 9th gesture, indicating similar hand gestures.

Table 3. Gesture recognition experiment results using the time domain features and LSTM.

Fold	Accuracy (%)	Recall (%)	Precision (%)
1	80.54	89.83	88.62
2	89.17	93.96	94.6
3	82.01	90	90.24
Average	83.91	91.26	91.15

The results of comparative experiments with conventional gesture recognition studies using EMG signals are shown in Table 4. The EMG signal has a characteristic that shows different waveforms for each gesture when the same subject performs different gestures. Therefore, the performance of gesture recognition using the EMG degraded as the number of classification gesture classes increased. In this paper, we conducted more hand gesture classification experiments than other related works by utilizing seventeen gestures. In

addition, the data size used in the experiment was 60ea, which is smaller than in related works (except [4]).

Table 4. Gesture recognition comparison experiment.

Author	Gesture	Input Data Size (ea)	Accuracy (%)
Zhang, X. et al. [4]	10	24	77.78
Jaramillo, A.G. et al. [5]	5	136	86
Oh, D.C. et al. [6]	6	1600	94
Qi, J. et al. [7]	9	64	95.15
Shioji, R. et al. [13]	3	1024	94.5
Chen, L. et al. [8]	7	3000	98.81
Asif, A.R. et al. [9]	10	7200	92
Proposed	17	60	83.91

The results of the single biometrics experiment using spectrogram and CNN-LSTM network are as shown in Table 5. The experiments were conducted correctly using all seventeen gestures from Ninapro DB group 1. As a result of the experiment, twenty subjects showed an average of 99.17% accuracy, 99.76% recall, and 99.41% precision performance. Nine people (4th, 5th, 7th, 8th, 11th, 12th, 15th, 19th, and 20th) were classified 100%, and the 14th subject was classified 97.06%, showing the worst performance. The 14th subject was most often misclassified as the 11th subject, and the 9th, 10th, and 11th gestures were most often misclassified, due to the similarity of the hand gestures.

Table 5. Results of single biometrics experiments using the CNN-LSTM network.

Fold	Accuracy (%)	Recall (%)	Precision (%)
1	98.68	99.85	98.83
2	99.27	99.42	99.85
3	99.56	100	99.56
Average	99.17	99.76	99.41

The results of a comparative experiment with the conventional single biometrics study using EMG signals are shown in Table 6. The EMG signal has a characteristic that shows different waveforms for each gesture when the same subject performs different gestures. Therefore, biometrics experiments with a large number of gestures degrade performance. The single biometrics method using EMG signals based on the CNN-LSTM network performed herein was tested with seventeen gestures in Ninapro DB group 1. By comparison, most of the related studies conducted biometrics experiments using one gesture, and no study used more than six gestures. In addition, we showed 99.17% performance on twenty subjects, which was more subjects than in related works (except [12,16]).

Table 6. Single biometrics comparison experiment.

Author	Gestures	Subjects	Accuracy (%)
Shreyas, V. et al. [10]	1	14	99.7
Shin, S.H. et al. [11]	3	5	81.6
Lu, L. et al. [12]	1	21	99.2
Shioji, R. et al. [13]	3	8	94.6
Lee, M. et al. [14]	1	12	93
Morikawa, S. et al. [15]	5	6	47.66
Lu, L. et al. [16]	1	21	99.29
Li, Q. et al. [17]	1	10	98.2
Khan, M.U. et al. [18]	1	10	95.3
Proposed	17	20	99.17

Figure 7 shows the 9th, 10th, and 11th gestures as performed by the 11th and 14th subjects, which showed the highest misclassification rates in a single biometrics test. The Ninapro DB 9th, 10th, and 11th gestures are EMG signals for performing ‘wrist supination’ operations, measured by varying axis and rotational direction. Therefore, we find that classification performance is low due to the use of EMG signals for similar hand gestures.

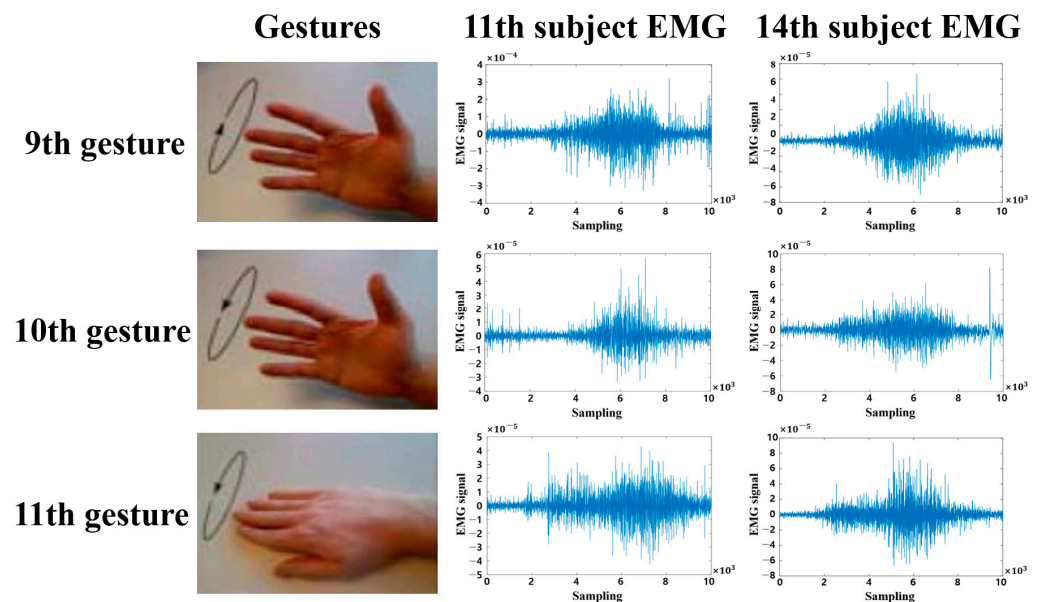


Figure 7. Ninapro DB (database) 11th and 14th subjects EMG signal.

The two-step biometrics method proposed herein performs single biometrics after gesture recognition, wherein data fusion of the two experimental results is performed in an AND form. This data fusion has an advantage in that the FAR can be lowered when it is performed in an AND form. Results of the comparison of FAR between two-step biometrics performing data fusion in an AND form and single biometrics is shown in Figure 8. As a result of the experiment, it was confirmed that the FAR of the two-step biometrics that was data fused in AND form was 0.0155%, which was 64.7% lower than that of single biometrics. As the FAR decreased, the probability of misclassifying an unregistered subject as a registered subject decreased [22]. Therefore, it was confirmed that the EMG based two-step biometrics method proposed in this paper was less likely to misclassify unregistered subjects as registered subjects. In addition, the existing biometrics problem in which registration data cannot be changed was solved by adding a gesture recognition process.

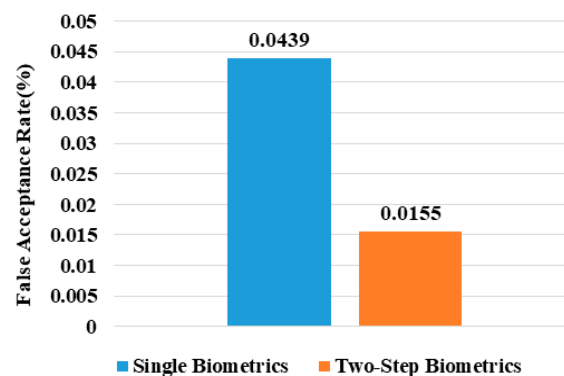


Figure 8. Comparison of false acceptance rate (FAR) between single biometrics and two-step biometrics.

5. Conclusions

A two-step biometrics method using EMG signals based on the CNN-LSTM network was proposed. Noise in the EMG signal was removed using NF and BPF, and preprocessing was performed with RMS. After preprocessing the EMG signal, the time domain features and LSTM network were used to examine whether the gesture matched, and single biometrics was performed if the gesture matched. In single biometrics, EMG signals were converted into a two-dimensional spectrogram, and training and classification were performed through the CNN-LSTM network. Data fusion of the gesture registration and single biometrics was performed in the form of an AND, which lowered FAR, thus proving that the proposed method was less likely to misclassify unregistered subjects as registered subjects. A publicly available EMG signal dataset (Ninapro DB) was used to evaluate the performance of the proposed two-step biometrics method using EMG signals. As a result of the experiment, seventeen hand gestures were classified with an accuracy of 83.91%, and twenty people were classified with an accuracy of 99.17%. In addition, we found that the fusion of gesture recognition and single biometrics reduced the FAR, which is important in biometrics. A two-step biometrics experiment will be conducted with a greater number of subjects and hand gestures in the future, and the hand gesture recognition performance using EMG signals will be improved. In addition, check the performance according to the spectrogram image size, and the intra- and inter-individual differences depending on each gesture of the EMG signal will be examined.

Author Contributions: Conceptualization, J.-S.K. and S.-B.P.; methodology, J.-S.K. and S.-B.P.; software, J.-S.K.; validation, J.-S.K., M.-G.K. and S.-B.P.; formal analysis, J.-S.K.; investigation, J.-S.K.; writing—original draft preparation, J.-S.K.; writing—review and editing, J.-S.K., M.-G.K. and S.-B.P.; supervision, S.-B.P.; project administration, S.-B.P.; funding acquisition, S.-B.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. 2017R1A6A1A03015496), the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2021R1A2C1014033), and the HPC Support Project, supported by the Ministry of Science and ICT and NIPA.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data supporting the findings of the article are available in the Ninapro DB2 at [29].

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

Abbreviations

EMG	Electromyogram
CNN-LSTM	convolutional neural network-long short-term memory
FAR	false acceptance rate
ECG	electrocardiogram
EEG	electroencephalogram
RMS	root mean square
ZC	zero crossing
KNN	K-nearest neighbor
CWT	continuous wavelet transform
PCA	principal component analysis
GRNN	general regression neural network
BPF	band pass filter
DB	database
ED	euclidean distance

SVM	support vector machine
UDP	unsupervised discriminant projection
CFA	class-dependent feature analysis
HPF	high pass filter
LPF	low pass filter
NF	notch filter
ANNs	artificial neural networks
LDA	linear discriminant analysis
DWT	discrete wavelet transform
OCSVM	one-class support vector machine
LOF	local outlier factor
EMD	empirical mode decomposition
AC	autocorrelation
FFT	fourier transform
OPF	optimum patch forest
DCT	discrete cosine transform
BP	backpropagation
MF	mean frequency
MAV	mean absolute value
VAR	variance
IEMG	integrated EMG
MAVSLP	MAV slope
WL	waveform length
ReLU	rectified linear unit
TP	true positive
TN	true negative
FP	false positive
FN	false negative

References

- Kim, J.S.; Pan, S.B. A study on EMG-based biometrics. *J. Internet Serv. Inf. Secur.* **2017**, *7*, 19–31.
- Kim, J.S.; Kim, S.H.; Pan, S.B. Personal recognition method using coupling image of ECG signal. *Smart Media J.* **2019**, *8*, 62–69. [\[CrossRef\]](#)
- Belgacem, N.; Fournier, R.; Ali, A.N.; Reguig, F.B. A novel biometric authentication approach using ECG and EMG signals. *J. Med. Eng. Technol.* **2015**, *39*, 226–238. [\[CrossRef\]](#) [\[PubMed\]](#)
- Zhang, X.; Yang, Z.; Chen, T.; Chen, D.; Huang, M.C. Cooperative sensing and wearable computing for sequential hand gesture recognition. *IEEE Sens. J.* **2019**, *19*, 5575–5583. [\[CrossRef\]](#)
- Jaramillo, A.G.; Benalcazar, M.E. Real-time hand gesture recognition with EMG using machine learning. In Proceedings of the IEEE Ecuador Technical Chapters Meeting, Salinas, Ecuador, 16–20 October 2017.
- Oh, D.C.; Jo, Y.U. EMG-based hand gesture classification by scale average wavelet transform and CNN. In Proceedings of the International Conference on Control, Automation and Systems, Jeju, Korea, 15–18 October 2019.
- Qi, J.; Jiang, G.; Li, G. Surface EMG hand gesture recognition system based on PCA and GRNN. *Neural Comput. Appl.* **2020**, *32*, 6343–6351. [\[CrossRef\]](#)
- Chen, L.; Fu, J.; Wu, Y.; Li, H.; Zheng, B. Hand gesture recognition using compact CNN via surface electromyography signals. *Sensors* **2020**, *20*, 672. [\[CrossRef\]](#) [\[PubMed\]](#)
- Asif, A.R.; Waris, A.; Gilani, S.O.; Jamil, M.; Ashraf, H.; Shafique, M.; Niazi, K. Performance evaluation of convolutional neural network for hand gesture recognition using EMG. *Sensors* **2020**, *20*, 1642. [\[CrossRef\]](#) [\[PubMed\]](#)
- Shreyas, V.; Felix, J.X.; Benjamin, C.; Marios, S. Electromyograph and keystroke dynamics for spoof-resistant biometric authentication. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Boston, MA, UAS, 8–10 June 2015.
- Shin, S.H.; Jung, J.H.; Kim, Y.T. A study of an EMG-based authentication algorithm using an artificial neural network. In Proceedings of the IEEE Sensors, Glasgow, UK, 29 October–1 November 2017.
- Lu, L.; Mao, J.; Wang, W.; Ding, G.; Zhang, Z. An EMG-based personal identification method using continuous wavelet transform and convolutional neural networks. In Proceedings of the IEEE Biomedical Circuits and Systems Conference, Nara, Japan, 17–19 October 2019.
- Shioji, R.; Ito, S.; Ito, M.; Fukumi, M. Personal authentication and hand motion recognition based on wrist EMG analysis by a convolutional neural network. In Proceedings of the IEEE International Conference on Internet of Things and Intelligence System, Bali, Indonesia, 1–3 November 2018.

14. Lee, M.; Ryu, J.; Youn, I. Biometric personal identification based on gait analysis using surface EMG signals. In Proceedings of the IEEE International Conference on Computational Intelligence and Applications, Beijing, China, 8–11 September 2017.
15. Morikawa, S.; Ito, S.; Ito, M.; Fukumi, M. Personal authentication by lips EMG using dry electrode and CNN. In Proceedings of the IEEE International Conference on Internet of Things and Intelligence System, Bali, Indonesia, 1–3 November 2018.
16. Lu, L.; Mao, J.; Wang, W.; Ding, G.; Zhang, Z. A study of personal recognition method based on EMG signal. *IEEE Trans. Biomed. Circuits Syst.* **2020**, *14*, 681–691. [[CrossRef](#)] [[PubMed](#)]
17. Li, Q.; Dong, P.; Zheng, J. Enhancing the security of pattern unlock with surface EMG-based biometrics. *Appl. Sci.* **2020**, *10*, 541. [[CrossRef](#)]
18. Khan, M.U.; Choudry, Z.A.; Aziz, S.; Naqvi, S.Z.H.; Aymin, A.; Imtiaz, M.A. Biometric authentication based on EMG signals of speech. In Proceedings of the International Conference on Electrical, Communication, and Computer Engineering, Istanbul, Turkey, 12–13 June 2020.
19. Liu, L.; Gu, X.F.; Li, J.P.; Lin, J.; Shi, J.X.; Huang, Y.Y. Research on data fusion of multiple biometric features. In Proceedings of the International Conference on Apperceiving Computing and Intelligence Analysis, Chengdu, China, 23–25 October 2009.
20. Kumar, A.; Zhang, D. Improving biometric authentication performance from the user quality. *IEEE Trans. Instrum. Meas.* **2010**, *59*, 730–735. [[CrossRef](#)]
21. Rahman, M.W.; Gavrilova, M. Emerging EEG and kinect face fusion for biometric identification. In Proceedings of the IEEE Symposium Series on Computational Intelligence, Honolulu, HI, USA, 27 November–1 December 2017.
22. Faundez-Zanuy, M. Data fusion in biometrics. *IEEE Aerosp. Electron. Syst. Mag.* **2005**, *20*, 34–38. [[CrossRef](#)]
23. Choi, Y.J.; Yu, H.J. Human motion tracking based on sEMG signal processing. *J. Inst. Control Robot. Syst.* **2007**, *13*, 769–776.
24. Lee, C.H.; Kang, S.I.; Bae, S.H.; Kwon, J.W.; Lee, D.H. A study of a module of wrist direction recognition using EMG signals. *J. RWEAT* **2013**, *7*, 51–58.
25. Phinyomark, A.; Limsakul, C.; Phukpattaranont, P. A novel feature extraction for robust EMG pattern recognition. *J. Comput.* **2009**, *1*, 71–80.
26. Phinyomark, A.; Hirunviriya, S.; Limsakul, C.; Phukpattaranont, P. Evaluation of EMG feature extraction for hand movement recognition based on Euclidean distance and standard deviation. In Proceedings of the International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, Chiang Mai, Thailand, 19–21 May 2010.
27. Choi, G.H.; Bak, E.S.; Pan, S.B. User identification system using 2D resized spectrogram features of ECG. *IEEE Access* **2019**, *7*, 34862–34873. [[CrossRef](#)]
28. Taherisadr, M.; Asnani, P.; Salster, S.; Dehzangi, O. ECG-based driver inattention identification during naturalistic driving using mel-frequency spectrum 2-D transform and convolutional neural networks. *Smart Health* **2018**, *9*, 50–61. [[CrossRef](#)]
29. Atzori, M.; Gijssberts, A.; Castellini, C.; Caputo, B.; Hager, A.G.M.; Elsig, S.; Giatsidis, G.; Hassetto, F.; Muller, H. Electromyography data for non-invasive naturally-controlled robotic hand prostheses. *Sci. Data* **2014**, *1*, 1–13. [[CrossRef](#)] [[PubMed](#)]