

# Article Study on Noise Reduction and Data Generation for sEMG Spectrogram Based User Recognition

Jae-Myeong Kim, Min-Gu Kim 🕩 and Sung-Bum Pan \*🕩

IT Research Institute, Chosun University, Gwang-Ju 61452, Korea; kjm11087@gmail.com (J.-M.K.); happy9433@chosun.ac.kr (M.-G.K.)

\* Correspondence: sbpan@chosun.ac.kr

Abstract: With the spread of the modern media industry, harmful genre contents are indiscriminately disseminated to teenagers. The password identification method used to block sensational and violent genre content has become a problem that teenagers can easily steal. Therefore, a user identification method with less risk of theft and hacking is required. The surface EMG (sEMG) signal, which is an electrical signal generated inside the body and has individual features, is being studied as a next-generation user identification method. sEMG involves measuring an individual's unique muscular strength activated over time as digital signals, thus giving it the advantage of generating different signal patterns. However, it is difficult to constantly and repeatedly acquire each motion signal and the number of repetitions for each motion is insufficient, thus there is a limit to improving user identification accuracy. In this paper, we propose a user identification system that solves the problem of insufficient data by applying the matching pursuit that enables signal generation to the sEMG signal from which the resting signal has been removed and improves classification accuracy by extracting STFT-based time-frequency features. As a result of the experiment, the user identification accuracy of the sEMG spectrogram with the resting state signal removed was 85.4%. In addition, when the training data were increased through data generation, the accuracy was improved, showing a user identification accuracy of 96.1%. Improved user recognition accuracy was confirmed when the training data of the sEMG signal from which the resting signal was removed were increased and multidimensional features including time-frequency were used.

Keywords: biometrics; sEMG; user recognition; spectrogram; CNN

# 1. Introduction

As the modern media industry expands due to the development of mobile devices such as smartphones and tablet PCs, media can be conveniently accessed without any place or time constraints [1]. Through such mobile devices, videos of harmful genre content such as crime and pornography are indiscriminately distributed and, in particular, youth with insufficient discriminant ability can easily access them, causing adverse social impacts such as imitation crimes [2]. Existing hazardous genre content approaches use a method of entering an ID and password permitted to access or an authentication method through personal information such as a phone number. However, the disadvantage is that anyone can access when the personal information is leaked. Therefore, as shown in Figure 1, it is necessary to strengthen user awareness to prove personal identity in order to block harmful genre content from youth with the most serious impact.

Existing methods that involve inputting passwords designated by users or using specific devices are accompanied by problems of password forgets, device misplacement, or theft [3]. Accordingly, biometric-information-based user identification technology using unique information or behavioral characteristics of users is garnering popularity. Biometric-information-based user identification technology is a technique of extracting unique features of an individual and converting them into information to identify different features for each person instead of using conventional passwords [4].



Citation: Kim, J.-M.; Kim, M.-G.; Pan, S.-B. Study on Noise Reduction and Data Generation for sEMG Spectrogram Based User Recognition. *Appl. Sci.* 2022, *12*, 7276. https:// doi.org/10.3390/app12147276

Academic Editor: Alexander E. Hramov

Received: 15 June 2022 Accepted: 18 July 2022 Published: 20 July 2022

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



**Copyright:** © 2022 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/).



Figure 1. Block harmful genre content through user recognition.

Biosignals are electrical signals generated inside the body and contain features that are unique to an individual. Representative biosignals include electromyogram (EMG), electrocardiogram (ECG), and electroencephalogram (EEG) [5]. Among such biosignals, various studies are being conducted to improve the performance of user identification technology using ECG signals that contain individual attributes based on the electrophysiological factors of the heart, the location and the size of the heart, and physical conditions. However, ECG signals cannot be changed when they are exposed externally because of problems including hacking. Furthermore, in ECG signals, the heart rate and waveform can change because of an individual's physical activity, the measurement time, or psychological effects. EMG signals, which can mitigate these drawbacks of ECG signals, are biosignals that are uniquely generated for each individual and generated by the muscle from behavioral characteristics, such as activities based on the development degree of each muscle and the muscle development degree between individuals. Additionally, as shown in Figure 2, surface electromyography (sEMG) signals can be measured on the skin surface of the desired muscle region to be acquired, rendering it easier to acquire signals compared with electrocardiography and electroencephalography. Moreover, sEMG signals can generate different signal patterns based on the muscles analyzed [6].



Signal acquisition using muscles of hands and arms

Signal acquisition using leg muscles

Figure 2. sEMG signal acquisition using various muscles.

However, the databases used in existing user identification studies involving sEMG signals do not contain a sufficient number of subjects or repetitions for each motion. Because sEMG signals identify users using signals generated by muscles based on behavioral

characteristics, they can be applied to user identification only when each motion is acquired using a sufficient number of repetitions. If the number of repetitions is insufficient, each motion may not be recognized, thereby hindering the application of the signals to user identification. In addition, most existing sEMG-based user identification studies applied temporal domain feature extraction methods to one-dimensional (1D) sEMG signals and applied signals to identification algorithms. However, sEMG signals are continuous signals with features that change constantly over time; furthermore, it is difficult to derive a clear cycle from them as it is difficult to repeat the same motion while maintaining a constant muscle strength based on time change. In other words, analyzing 1D sEMG signals acquired without a constant muscle strength and time in the temporal domain results in degraded user identification performance [7].

Hence, a preprocessing method was used in this study to address the problems of irregular signals and insufficient data to apply 1D sEMG signals acquired based on behavioral characteristics to the user identification system; furthermore, a user identification method using multidimensional feature sEMG spectrograms, including time-frequency information, is presented. First, irregular resting state signals and noise contained in sEMG signals were removed and sEMG signals were generated using matching pursuit, which enables data generation using a small amount of available data. Subsequently, the preprocessed 1D sEMG signals were applied to short-time Fourier transform (STFT), which is a method of extracting multidimensional features including time-frequency information. After adjusting the sEMG signals to a resolution that can be efficiently analyzed, they were transformed into sEMG spectrograms, and a convolutional neural network (CNN) was used to identify users in the final step. The experiment's results revealed a 78.7% user identification accuracy based on the transformation of the raw sEMG signals into sEMG spectrograms for 40 subjects. By contrast, when the resting state signals were removed and transformed into sEMG spectrograms, as proposed herein, the user identification performance was 85.4%, which was a 6.7% improvement compared with using raw signals. Furthermore, when data were augmented by generating sEMG signals after removing the resting state signals, the performance increased by 10% to 96.1%, compared with before increasing the data amount.

This paper is organized as follows: The research trend involving sEMG signals is presented in Section 2. The proposed sEMG spectrogram-based user identification of this study is described in Section 3. The experimental methods employed for the proposed user identification and results analysis are presented in Section 4. Finally, the conclusions and future research directions are presented in Section 5.

## 2. Related Works

Techniques applied to existing user identification systems using sEMG signals are analyzed in this section. Figure 3 illustrates the user identification system using sEMG signals. First, sEMG signals are acquired independently or by using open databases to construct data. Subsequently, data preprocessing and normalization are conducted to remove noise in the data signals. The preprocessed signals are passed subjected to feature extraction and then the user identification performance is evaluated using a classifier.

For user identification, sEMG data are first constructed. sEMG data consist of muscle and motion data to be used. As summarized in Table 1, there are Ninapro DB2 and sEMG Basic Hand Movement Upatras databases for hand or wrist movement. Ninapro DB2 was composed by acquiring movements of fingers and wrists and holding movements from 40 subjects, and sEMG Basic Hand Movement Upatras is a database composed by acquiring daily hand holding movements from five subjects. The database for finger movement is Rami Khushaba's sEMG, and it is a database constructed by acquiring data from eight subjects by performing finger movements. Finally, the EMG dataset in lower limb dataset using the leg muscles was constructed by acquiring the sEMG signals generated when performing leg movements from 22 subjects [8–11].





Database Name	Channel	Individuals	Motion	Remarks		
sEMG Basic Hand Movement Upatras	2ch	5 subjects	6 hand motions	<ul> <li>Duration for each motion: 6 s.</li> <li>Repetitions for each motion: 30 times.</li> </ul>		
EMG (Dr. Rami Khushaba)	2ch	8 subjects	10 finger motions	<ul> <li>Duration for each motion: 5 s.</li> <li>Repetitions for each motion: 6 times.</li> </ul>		
Ninapro DB2	2 12ch 40 subjects 49 hand and wrist motions		<ul> <li>Duration for each motion: 5 s.</li> <li>Repetitions for each motion: 6 times.</li> </ul>			
Benchmark Datasets for Bilateral Lower-Limb	7ch	10 subjects	3 leg motions	• Repetitions for each motion: 3 times.		

However, the problem regarding insufficient data in sEMG DBs (an insufficient number of acquired subjects or motion repetitions) has been continuously problematic. Although multiple DBs can be synthesized to address this issue, limitations exist when combining multiple DBs into one, because each sEMG DB employs different motions, muscle channels, number of motion repetitions, and motion durations as initial acquisition conditions. Furthermore, when increasing the amount of data by segregating the sEMG signals of one motion cycle into predefined windows, it is difficult to identify the same motion signal because the frequency component changes over time even for the same motion [12].

The constructed sEMG DB must undergo preprocessing as it contains noise generated from various environments. The noise that should be removed includes power line noise generated from the measuring device, modulated waves from 60 Hz band, white noise from broadband, noise occurring from differences in performance and functions of disposable electrodes, noise caused by physiological interference, and noise caused by characteristics of muscle tissues. Preprocessing should be conducted to remove such noises. First, they were identified through frequency analysis and then removed by using various filters, such as high-pass, low-pass, band-pass, Butterworth [13], and notch filters [14].

After that, features are extracted from the preprocessed EMG signal. Feature extraction can be mainly classified into time domain, frequency domain, and time–frequency domain. The time and frequency domain functions include the mean absolute value (MAV) for detecting muscle activity, the slope sign change (SSC) expressing the frequency domain characteristics of the EMG signal calculated in the time domain, and the root mean square (RMS) related to continuous force and contraction. In addition, there are waveform lengths (WL) indicating the waveform length with respect to time segment, variance (VAR) indicating the characteristic of force and the integrated EMG (IEMG) used as an onset detection index, and zero crossing (ZC), indicating the number of times the amplitude of the EMG signal crosses zero. Table 2 summarizes the formulas used for each feature extraction [15].

Feature Name	Formula				
MAV	$\frac{1}{N}\sum_{n=1}^{N} x_{n} $				
SC	$\sum_{n=2}^{N-1} [f[(x_n - x_{n-1}) \times (x_n - x_{n+1})]],$ $f(x) = \begin{cases} 1, ifx \ge threshold\\ 0, otherwise \end{cases}$				
RMS	$\sqrt{rac{1}{N}}\sum\limits_{n=1}^{N}x_n^2$				
WL	$\sum_{n=1}^{N-1}  x_{n+1} - x_n $				
VAR	$rac{1}{N-1}\sum\limits_{n=1}^N x_n^2$				
IEMG	$\sum_{n=1}^{N}  x_n $				
ZC	$\sum_{n=1}^{N-1} [sgn(x_n \times x_{n+1}) \cap  x_n - x_{n+1}  \ge threshold],$ $sgn(x) = \begin{cases} 1, if x \ge threshold\\ 0, otherwise \end{cases}$				

Table 2. Formulas for temporal domain features.

sEMG signals are continuous signals with features that change over time; therefore, it is difficult to locate a clear periodicity in them as it is difficult to repeat the same motion at a constant time and intensity. One of the most widely used feature extraction methods for the frequency domain is fast Fourier transform (FFT). The FFT transforms signals of a temporal domain into a frequency domain. However, the FFT is restricted in evaluating the frequency component for the desired time point because of temporal limitations [16].

STFT is used as a representative feature extraction method in the time–frequency domain [17]. The STFT is a method that compensates for the temporal limitations of the FFT; it selects the desired window length, partitions the time into short segments, and Fourier transforms each of the partitioned segments. Accordingly, because the STFT analyzes the frequency components based on the temporal domain, it has been proven to be more efficient as an analysis method than using temporal and frequency features as it analyzes time–frequency multidimensional features by applying sEMG signals, whose features change over time [7].

Methods used for sEMG signal classification include machine learning and deep learning approaches. Machine learning is a technique that trains a machine to perform a task on behalf of humans to achieve the desired result. Machine learning can be primarily categorized into supervised learning and unsupervised learning. In supervised learning, a model is trained with correct answers to obtain correct predictions; one of the most widely used supervised learning methods is the support vector machine [18]. Unsupervised learning is a method of obtaining similar patterns through clustering without using correct answers. One of the most widely used unsupervised learning methods is the k-nearest neighbor (KNN) algorithm. In addition to the KNN, other classifiers used include decision tree learning, random forest, principal component analysis [19], and linear discriminant analysis [20]. Deep learning, a subset of machine learning, is a technique that uses artificial neural networks as its basis; it attempts to solve the problem of weak training by deeply configuring neural networks. The most widely used deep learning networks include CNNs and long short-term memory (LSTM). LSTM is an algorithm that solves the long-term dependence problem of recurrent neural networks and is often used for data with temporal characteristics. The CNN is the most widely used deep learning method and is useful for obtaining patterns in images [21].

#### 3. Proposed sEMG Spectrogram-Based User Identification

This section presents a preprocessing method for removing resting state signals and noise to solve the irregularity problem of sEMG signals, as well as a user identification system that applies a time–frequency multidimensional analysis method to 1D sEMG signals. Figure 4 illustrates the overall flowchart of the proposed sEMG spectrogram-based user identification.



Figure 4. Flowchart of proposed sEMG spectrogram-based user identification.

Among the open databases, Ninapro DB2 is used to compose the sEMG signal dataset. Subsequently, sEMG signals are partitioned into one motion cycle signals, and a preprocessing process that includes removing noise and irregular resting state signals generated in the signal acquisition process is performed. Noise occurring in the sEMG signals is removed using filters such as band-stop and band-pass filters. Additionally, the sEMG signals of one motion cycle are partitioned into non-overlapping frames; the energy and spectrum center for each frame are calculated to set a threshold value and the irregular resting state signals are removed by extracting only signals containing activity information. The final step of the preprocessing increases the amount of sEMG data through matching pursuit. Subsequently, the preprocessed 1D sEMG signals are transformed into two-dimensional (2D) sEMG spectrogram images by applying a time–frequency multidimensional feature extraction method. Finally, the user identification performance is verified using the CNN, the most widely used technique for image classification.

### 3.1. Noise Removal Including Resting State Signals

In Ninapro DB2, raw signals are acquired by repeating 40 hand and wrist motions six times each. To partition the sEMG signals into one motion cycle signals for each motion, the signals are partitioned using the label assigned to each motion in the data, as shown in Figure 5.



Figure 5. Partitioned one motion cycle signals.

After partitioning them into one motion cycles, noise contained in the sEMG signals is removed. The types of noises contained in the sEMG signals include power line noise caused by the measuring device, noise occurring from differences in performance and functions of disposable electrodes, and noise caused by physiological interference. In this study, a band-pass filter was used to pass the 10–500 Hz frequency band containing activity information without attenuating it, whereas the remainder of the frequency band was attenuated and removed. The power line noise, which was generated by the poor grounding of the measuring device or by high power cables around the device, was generally observed at 60 Hz. To remove such power line noise, a band-stop filter was employed to remove the 60 Hz frequency band.

Subsequently, the resting state signals in the sEMG signals were removed. The resting state signal defined in this study refers to a resting signal included before and after motion is performed during the motion execution time set as the signal acquisition condition. To remove the resting state signal, the sEMG signals of one motion cycle are partitioned into non-overlapping frames and the mean energy for each frame is calculated and set as the threshold. Based on the configured threshold, if the signal is greater than the threshold value, then it is regarded as a motion signal containing activity information and is extracted; otherwise, it is regarded as a resting state signal and is removed. Because the sEMG signals extracted in this process involve different durations, the size of all signals is adjusted to be the same through resampling. Figure 6 illustrates the result of removing the resting state signals from the sEMG signals partitioned into one motion cycle signals.



Figure 6. Removal of resting state signals.

#### 3.2. Data Increase Using Matching Pursuit

For the sEMG signals processed through noise and resting state signal removal, the amount of data was increased by generating sEMG signals. The generative adversarial

network (GAN) is the most widely used technique for data generation. GAN comprises a generator that generates data and a discriminator that assesses the generated data; during training, the generator and the discriminator compete against each other to improve the performance. The GAN requires a substantial amount of data in advance to generate data [13]. However, it is difficult to apply the GAN for generating sEMG signals as the sEMG database is acquired using a small number of motion repetitions, i.e., the amount of data obtained is insufficient. Hence, matching pursuit, which can generate signals using a small amount of data and enables quick data generation using a relatively simple formula compared with other data generation techniques, was employed in this study. The matching pursuit algorithm was first introduced by Mallat and Zhang [16]. The basic idea of matching pursuit is to first select atoms individually to identify the atom with the highest inner product using the current signal after expressing signals with approximation, subtracting an approximation that uses only that one atom from the signal, and repeating the process until the residual signal is decomposed. The approximate decomposition of the matching pursuit algorithm can be expressed as shown in Equation (1) below, where  $\mathscr{R}^{(m)}$ denotes the residual signal and, based on the number of repetitions (m), the index  $r_i$  is obtained and  $\alpha_r$  is derived.

$$\mathcal{S} = \sum_{i=1}^{m} \alpha_{r_i} \mathcal{O}_{r_i} + \mathcal{R}^{(m)} \tag{1}$$

Accordingly, a signal that is similar to the current signal can be generated by applying matching pursuit to the sEMG signals. Signals can be generated by changing the similarity based on the number of repetitions. The signals were generated to exhibit cross-correlation similarity between 90% and 99%. In this study, preprocessing was performed as follows. First, noise and resting state signals in the signals were removed using the band-pass and band-stop filters mentioned earlier. Subsequently, sEMG data signals were generated using matching pursuit. All the preprocessed one motion cycle EMG signals were convenient to visually check muscle activation and are combined into 12-channel signals in the temporal domain, as shown in Figure 7.



Figure 7. 12-channel synthesis using temporal domain.

#### 3.3. User Identification Using sEMG Spectrogram

In this study, multidimensional features were extracted from the preprocessed and normalized 1D sEMG signals by applying the STFT, a time–frequency feature extraction method. The STFT is a method that compensates for the disadvantages of the existing FFT, a frequency–domain feature extraction method; it enables the extraction of multidimensional features, including time and frequency, by analyzing the frequency components at a desired time point for signals that change over time [22]. The application of the STFT is expressed as shown in Equation (2).

$$\mathscr{X}(\mathscr{R},w) = \int_{-\infty}^{\infty} x(t)w(t-\mathscr{R})e^{-i\omega t}dt = s$$
<sup>(2)</sup>

The spectrogram transformation is performed based on the FFT length using the input signal  $\mathcal{X}(t)$  and the window function w(t), where  $\mathcal{R}$  denotes the window length, w the angular frequency, and s the spectrogram value. Hence, the frequency information over time can be included by applying Equation (1) to 1D sEMG signals, such that multidimensional features containing time–frequency information can be extracted. In the process, the temporal resolution was enhanced by applying a 50% overlap, because both the time and frequency resolutions cannot be improved simultaneously. Figure 8 illustrates the result of transforming the extracted multidimensional features into the sEMG spectrogram.



Figure 8. Transformation of sEMG signal into spectrogram.

The transformed 2D sEMG spectrogram images were derived by adjusting the window length  $\mathscr{R}$  to verify the time–frequency resolution based on the change in the window length  $\mathscr{R}$ , which is a parameter used in the STFT. The minimum value of the window length  $\mathscr{R}$  was set to 64, and it was increased by two-fold increments until the maximum length of 512 was attained. The change in the frequency resolution increases and the time resolution decreases based on the increases in the window length as illustrated by the spectrograms shown in Figure 9 [23].



Figure 9. sEMG spectrograms based on changes in STFT window length.

In the final user identification, a typical deep learning CNN based on a 2D image input was employed. The constructed CNN comprised three convolutional layers, two max pooling layers, two fully connected layers, and a ReLu activation function. The final classification was proceeded by the softmax function. Figure 10 illustrates the overall network structure. The filter size of the CNN's convolutional layers was set to  $3 \times 3$ , whereas those of the pooling layers and stride were set to  $2 \times 2$  and 2, respectively. The maximum number of iterations was set to 150, the initial weight of each layer was set to random, and the user was identified in the output layer by the softmax function in the final step [21].



Figure 10. CNN structure.

Furthermore, DenseNet201, which is designed with a deep neural network, and MobileNet-v2 network, which can be applied to limited environments without a high-performance computer, were used in addition to the CNN constructed in this study to compare the identification performance. DenseNet201 was designed with a densely connected CNN structure. This model improves the flow between layers by connecting the feature maps of the previous layer with the inputs of the subsequent layer. Such a structure is applied to supplement the problem of information loss that occurs as the information on input passes through multiple layers with the increase in the depth of the network. Therefore, DenseNet201 can improve the vanishing gradient, strengthen feature propagation, encourage feature reuse, and reduce the number of parameters [24].

MobileNet is a network designed based on studies regarding lightweight networks that can be used in a limited environment without requiring a high-performance computer. The early MobileNet network was designed to achieve fast training and accuracy improvement under low power consumption through its small size. The core idea of MobileNet is to perform convolution operations on filters corresponding to each channel after segregating the input data by channels using depthwise separable convolutions obtained from Xception. Hence, the number of filters is equivalent to the number of channels of the input data, and the feature map that has completed a convolution operation is again passed on to the convolution operation to be output as a final result of the 1-channel [25]. MobileNet-v2 was designed based on MobileNet-v1. In the second version, the number of required tasks and memory was reduced, while the same accuracy was maintained by segregating the entire convolution into two separate layers with different strides [26].

#### 4. Experimental Methods and Results

Ninapro DB2, a representative open database, was used to evaluate the reproducibility of the proposed sEMG spectrogram-based user identification performance. Table 3 summarizes the detailed composition of the database. A total of 40 subjects were used in the database, and Figure 11 illustrates the hand motions employed in this study. The hand motions included in Ninapro DB2 comprised movements of an entire arm or a hand. However, motions with large movements were not suitable for use as user identification passwords; hence, such motions were removed and the data were constructed using motions one to seven, which can be performed within the palm range. Each motion was repeated six times for 5 s with a 3 s break; the data were acquired at a sampling rate of 2000 Hz and 12 channels were used, including the forearm periphery, biceps, and triceps [9]. As the amount of data was insufficient because each motion was repeated only six times, the training and test data for each motion were initially composed of three data entries each when the data were partitioned by applying a ratio of 5:5. In other words, the total amount of data for one subject was 21 training data and 21 test data; hence, the total amount of data for 40 subjects was 840 training data and 840 test data.

Table 3. sEMG signal database composition.

Item	Data
Number of subjects	40
Number of motions	7
Number of channels	12
Number of repetitions	6
Sampling rate (Hz)	2000



Figure 11. Seven hand motions of Ninapro DB2.

Noise, including resting state signals, was removed from the acquired sEMG signals; subsequently, the amount of data was increased using matching pursuit. The sEMG signals generated through the predefined number of repetitions in matching pursuit can be confirmed by measuring the similarity with the actual signals through cross-correlation. Table 4 summarizes the cross-correlation similarity between the actual sEMG signals and the sEMG signals generated through matching pursuit. It was confirmed that the signals were generated with a similarity between 90% and 99%. Accordingly, when the initial training dataset was increased by ten times, the data of each motion were composed of 33 training data and 3 test data; the amount of training data was increased for the data of 40 subjects through such data generation to acquire a composition ratio of 9240 training data and 840 test data.

Table 4. Similarity of sEMG signals generated for one motion cycle.

Item	Number of Repetitive Generations									
	1	2	3	4	5	6	7	8	9	10
Cross-correlation similarity	0.908	0.919	0.927	0.934	0.942	0.955	0.96	0.973	0.985	0.99
	0.903	0.913	0.922	0.932	0.941	0.949	0.956	0.964	0.971	0.981
	0.907	0.918	0.928	0.937	0.944	0.956	0.963	0.971	0.979	0.989
	0.900	0.912	0.923	0.933	0.94	0.952	0.959	0.966	0.972	0.983
	0.904	0.916	0.924	0.934	0.943	0.953	0.961	0.969	0.974	0.986
	0.907	0.917	0.925	0.937	0.949	0.956	0.961	0.968	0.972	0.988
	0.909	0.919	0.927	0.936	0.945	0.955	0.964	0.971	0.977	0.989

To analyze the sEMG spectrogram-based user identification performance, the window length, a parameter of the STFT, was set to 64, 128, 256, and 512, separately, and the sEMG spectrogram was generated by changing the time–frequency resolution. As shown in Figure 12, the highest identification accuracy of 85.4% was observed for the signals with resting state signals removed when using the window length of 256. Therefore, the window



length parameter having the most suitable time-frequency resolution for the sEMG signal was set to 256.

As shown in Figure 13, when the sEMG spectrograms were generated using the window length of 256, the user identification accuracy for 40 subjects was 78.7% for the raw signals before resting state signal removal and 85.4% for the signals after resting state signal removal, indicating a 6.7% improvement after removing the static rest signal that did not contain muscle information. Furthermore, the amount of training data increased through matching pursuit after the resting state signals were removed to address the insufficient data problem of the sEMG database. The sEMG signals were generated with a similarity of 90–99% for the signals of each motion; the user identification performance was evaluated after augmenting the amount of data training data by 10 times from the existing insufficient data. When the number of insufficient training data was increased by 10 times, the sEMG spectrogram improved by more than 10%, showing 96.1% of user recognition performance, and it was confirmed that when the number of training data was increased by 10 times or more, it converged to a constant value.



Figure 13. Identification performance based on preprocessing.

Figure 12. Identification performance based on changes in window length.

For performance comparison, the user identification performance was compared using the DenseNet201 neural network, which is composed of a deep neural network, and the MobileNet-v2 neural network, which can be applied to a mobile environment by reducing the computational cost and model size. Figure 14 illustrates the user identification performance using the raw signals and the user identification performance after noise removal and data generation.



Figure 14. User identification performance comparison.

In this paper, the number of learning data is increased tenfold by removing the resting signal included in the sEMG signal and applying the matching pursuit. In addition, the time–frequency resolution of STFT suitable for sEMG signals was set to 256 through an experiment and converted into an sEMG spectrogram, and then the deep learning-based CNN was directly constructed to confirm the user identification accuracy. The batch size of CNN was set to 128, maxEpochs to 150, and filter size to  $3 \times 3$ . As a result of the experiment, when the method proposed in this paper was applied, the user recognition accuracy was 96.1%. Accuracy was improved by 22.4% compared with before increasing the training data, and accuracy by 17.4% was improved compared with when a one-dimensional signal was used as an input. When the user recognition accuracy was checked using MobileNet-v2 and DenseNet201, as well as the directly constructed CNN, it was confirmed that the user recognition accuracy was improved in the directly constructed CNN.

In addition, the identification accuracy was compared with the previous study using the same Ninapro DB2. Zhai [7] and Huang [17] conducted pattern identification by converting a one-dimensional sEMG signal into a spectrogram. Zhai compared and analyzed the accuracy according to the feature extraction method and confirmed the identification accuracy of 77% using the spectrogram-based SVM. It was demonstrated that sEMG signals can be efficiently analyzed when the spectrogram, a multidimensional feature extraction method with time–frequency information, is applied rather than the time–domain feature extraction method, which is a one-dimensional analysis method. In addition, Huang compared and analyzed the accuracy according to the classifier and confirmed the identification accuracy of 79.4% using the spectrogram-based CNN-LSTM. The deep learning CNN-LSTM network showed higher identification accuracy than the existing machine learning SVM. In this paper, it was confirmed that the user recognition accuracy was improved to 96.1% of identification accuracy when the learning data were increased by applying the matching pursuit that can generate a signal to the sEMG signal from which the resting signal was removed, and the STFT-based time–frequency feature was used.

#### 5. Conclusions

In existing user identification studies using sEMG signals, experiments with sEMG signals were conducted after removing noise using only simple filters. sEMG signals are time series data acquired over time and are generated based on different activity degrees of each muscle when performing a motion; hence, they can be applied to user identification once a sufficient amount of data is acquired. However, most sEMG databases contain a minimal amount of data and signals are acquired irregularly as it is difficult to acquire constant signals based on the conditions involved when acquiring data by repeating motions. Furthermore, because sEMG signals, which are time series data, cannot be repeated while maintaining a constant muscle strength over time, it is difficult to obtain a clear periodicity; hence, user identification performance is degraded when analyzed as 1D features.

In this study, a preprocessing method was employed for solving problems of irregular signals and insufficient data for a user identification system using 1D sEMG signals obtained based on behavioral characteristics; furthermore, a user identification method using multidimensional feature sEMG spectrograms containing time-frequency information was proposed. After removing irregular resting state signals and noise included in the sEMG signals, the sEMG signal data were generated using matching pursuit, which enables signals to be generated using a small amount of data and quick data generation using a relatively simple formula compared with other data generation techniques. The similarity of the generated sEMG signals was verified using cross-correlation similarity, which yielded 90% to 99% similarity with the raw data. The preprocessed 1D sEMG signals were applied with STFT, a multidimensional feature extract method containing time-frequency information, and the resolution was changed to enable efficient analysis of the sEMG signals. Subsequently, after transforming the signals into sEMG spectrograms, a CNN model was used to perform final user identification. The proposed system comprised processes of sEMG data composition, sEMG data preprocessing and normalization, transformation of 1D sEMG signals into spectrograms, and final classification.

Based on experiments, the user identification accuracy obtained using the 1D sEMG signals was 59.3% before performing preprocessing and 66.7% after performing preprocessing, indicating only a slight performance increase. Furthermore, the user identification performance of 40 subjects using the proposed method was 78.4% before preprocessing and 85.4% after preprocessing when the sEMG signals were transformed into spectrograms by applying a window length of 256 in 12 channels; these results indicated 19.4% and 18.7% accuracy improvements, respectively, compared with the case of using 1D sEMG signals. When the insufficient amount of data was increased by 10 times by applying matching pursuit, the user identification performance was 96.1%, a 10% increase compared with before data augmentation. By conducting user identification using DenseNet201 and MobileNet-v2 networks in addition to the CNN employed in this study, it was demonstrated that the user identification performance improved when applying the method proposed. Accordingly, the possibility of performing user identification was verified based on the use of multidimensional feature sEMG spectrograms transformed through the STFT after the removal of noise and unnecessary resting state signals in the 1D sEMG signals, as well as training data augmentation. In the future, we plan to acquire sEMG signals directly from the wearable device environment, build a database, and conduct sEMG signal-based user identification research that can be applied in real life.

**Author Contributions:** Conceptualization, J.-M.K. and S.-B.P.; methodology, J.-M.K. and S.-B.P.; software, J.-M.K.; validation, J.-M.K., M.-G.K., and S.-B.P.; formal analysis, J.-M.K.; investigation, J.-M.K.; writing—original draft preparation, J.-M.K.; writing—review and editing, J.-M.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 study was supported by research fund from Chosun University (2020).

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 http://ninaweb.hevs.ch/ (accessed on 1 September 2020), reference number [9].

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

#### Abbreviations

Surface electromyography (sEMG), electromyogram (EMG), electrocardiogram (ECG), electroencephalogram (EEG), one-dimensional (1D), short-time Fourier transform (STFT), convolutional neural network (CNN), mean absolute value (MAV), slope sign change (SSC), root mean square (RMS), waveform length (WL), variance (VAR), integrated EMG (IEMG), zero crossing (ZC), Fourier transform (FFT), k-nearest neighbor (KNN), long short-term memory (LSTM), two-dimensional (2D), generative adversarial network (GAN).

#### References

- 1. Kim, G.Y. SmartPhone Security technology in the open mobile environment. *Korea Inf. Process. Soc.* 2009, 19, 21–28.
- Baek, J.H.; Lee, D.K.; Hong, C.Y.; Ahn, B.T. Multimodal approach for blocking obscene and violent contents. J. Converg. Inf. Technol. 2017, 7, 113–121.
- 3. Xiao, Q. Technology review-biometrics-application, challenge, and computational intelligence solutions. *IEEE Comput. Intell. Mag.* **2007**, *2*, 5–25. [CrossRef]
- Wayman, J.L. Technical testing and evaluation of biometric identification devices. In *Biometrics*; Jain, A.K., Bolle, R., Pankanti, S., Eds.; Springer: Boston, MA, USA, 1996; pp. 345–368.
- Luis-Garcia, R.D.; Alberola-Lopez, C.; Aghzout, O.; Ruiz-Alzola, J. Biometric identification systems. *Signal Process.* 2003, *83*, 2539–2557. [CrossRef]
- 6. Scheme, E.; Englehart, K. Electromyogram pattern recognition for control of powered upper-limb prostheses: State of the art and challenges for clinical use. *J. Rehabil. Res. Dev.* 2011, *48*, 643–659. [CrossRef] [PubMed]
- Zhai, X.; Jelfs, B.; Chan, R.H.M.; Tin, C. Short latency hand movement classification based on surface EMG spectrogram with PCA. In Proceedings of the IEEE Engineering in Medicine and Biology Society, Orlando, FL, USA, 16–20 August 2016; pp. 327–330.
- 8. Khushaba, R.N.; Takruri, M.; Kodagod, S.; Dissanayake, G. Toward improved control of prosthetic fingers using surface electromyogram(EMG) signals. *Expert. Syst. Appl.* **2012**, *39*, 10731–10738. [CrossRef]
- Atzori, M.; Gijsberts, A.; Castellini, C.; Caputo, B.; Hager, A.M.; Elsig, S.; Giatsidis, G.; Bassetto, F.; Muller, H. Electromyography data for non-invasive naturally-controlled robotic hand prostheses. *Sci. Data* 2014, *1*, 1–13. [CrossRef] [PubMed]
- 10. Sapsanis, C.; Georgoulas, G.; Tzes, A. EMG based classification of basic hand movements based on time-frequency features. In Proceedings of the Mediterranean Conference on Control and Automation, Platanias, Greece, 25–28 June 2013; pp. 716–722.
- 11. Hu, B.; Rouse, E.; Hargrove, L. Benchmark datasets for bilateral lower-limb neuromechanical signals form wearable sensors during unassisted locomotion in able-bodied individuals. *Front. Robot. AI* **2018**, *5*, 14. [CrossRef] [PubMed]
- 12. Phinyomark, A.; Scheme, E. EMG pattern recognition in the era of big data and deep learning. *Big Data Cogn. Comput.* **2018**, *2*, 21. [CrossRef]
- 13. Selesnick, I.W.; Burrus, C.S. Generalized digital butterworth filter design. *IEEE Trans. Signal. Process.* **1998**, *46*, 1688–1694. [CrossRef]
- 14. Nehorai, A. A minimal parameter adaptive notch filter with constrained poles and zeros. *IEEE Trans. Acoust.* **1985**, *33*, 983–996. [CrossRef]
- 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; pp. 856–860.

- Jayasree, D.D. Classification of power quality disturbance signals using FFT, STFT, wavelet transforms and neural networks—a comparative analysis. In Proceedings of the Computational Intelligence and Multimedia Applications, Sivakasi, India, 13–15 December 2007; pp. 335–340.
- 17. Huang, D.; Chen, B. Surface EMG decoding for hand gestures based on spectrogram and CNN-LSTM. In Proceedings of the Cognitive Computing and Hybrid Intelligence, Xi'an, China, 21–22 September 2019; pp. 123–126.
- 18. Oskoei, M.A.; Hu, H. Support vector machine-based classification scheme for myoelectric control applied to upper limb. *IEEE Trans. Biomed. Eng.* **2008**, *55*, 1956–1965. [CrossRef] [PubMed]
- 19. Wold, S.; Esbensen, K.; Geladi, P. Principal component analysis. Chemometr. Intell. Lab. Syst. 1987, 2, 37–52. [CrossRef]
- Mika, S.; Ratsch, G.; Weston, J.; Scholkopf, B.; Mullers, K.R. Fisher discriminant analysis with kernels. In Proceedings of the IEEE Signal Processing Society Workshop, Madison, WI, USA, 25–25 August 1999; pp. 41–48.
- Bao, T.; Zaidi, A.; Xie, S.; Zhang, Z. Surface-EMG based wrist kinematics estimation using convolutional neural network. In Proceedings of the IEEE International Conference on Wearable and Implantable Body Sensor Networks, Chicago, IL, USA, 19–22 May 2019; pp. 1–4.
- Kim, J.M.; Choi, G.H.; Kim, J.S.; Pan, S.B. User recognition using electromyogram 2D spectrogram images based on CNN. J. KIIT 2021, 19, 107–117. [CrossRef]
- Zawawi, T.N.S.T.; Abdullah, A.R.; Shair, E.F.; Halim, I.; Rawaida, O. Electromyography signal analysis using spectrogram. In Proceedings of the IEEE Student Conference on Research and Developement, Putrajaya, Malaysia, 16–17 December 2013; pp. 319–324.
- Huang, G.; Liu, Z.; Maaten, V.D.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 2261–2269.
- 25. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861.
- Sandler, M.; Howard, A.G.; Zhu, M.; Zhmoginov, A.; Chen, L. MobileNetV2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 4510–4520.