sustainability

MDPI

Article

Multi-Session Surface Electromyogram Signal Database for Personal Identification

Jin-Su Kim, Cheol-Ho Song, EunSang Bak * and Sung-Bum Pan *

IT Research Institute, Chosun University, 309 Pilmun-daero, Gwang-Ju 61452, Korea; babotn5@gmail.com (J.-S.K.); songch0636@gmail.com (C.-H.S.)
* Correspondence: bakeunsang@gmail.com (E.B.); sbpan@chosun.ac.kr (S.-B.P.); Tel.: +82-62-230-6947 (S.-B.P.)

Abstract: Surface electromyogram (sEMG) refers to a biosignal acquired from the skin surface during the contraction of skeletal muscles, and a different signal waveform is generated, depending on the motion performed. Therefore, in contrast to generic personal identification, which uses only a piece of registered information, the sEMG changes the registered information in a personal identification method. The sEMG database (DB) for conventional personal identification has shortcomings, such as a few subjects and the inability to verify sEMG signal variability. In order to solve the problems of DBs, this paper describes a method for constructing a multi-session sEMG DB for many subjects. Data were obtained in two channels when each of the 200 subjects performed 12 motions. There were three sessions, and each motion was repeated 10 times in time intervals of a day or longer between each session. Furthermore, to verify the effectiveness of the constructed sEMG DB, we conducted a personal identification experiment. According to the experimental results, the accuracy for five subjects was 74.19%, demonstrating the applicability of the constructed multi-session sEMG DB.

Keywords: multi-session data; benchmarking data; electromyogram; personal identification

check for updates

Citation: Kim, J.-S.; Song, C.-H.; Bak, E.; Pan, S.-B. Multi-Session Surface Electromyogram Signal Database for Personal Identification. *Sustainability* **2022**, *14*, 5739. https://doi.org/10.3390/su14095739

Academic Editor: Ilsun You

Received: 25 March 2022
Accepted: 6 May 2022
Published: 9 May 2022

1. Introduction

As information security has become critical in modern society, numerous personal identification methods have been proposed that use information from the body. Personal identification methods using fingerprints, the face, and the iris are widely applied to the products used in everyday life, such as cellphones, because of benefits, such as low rejection and ease of use. However, fingerprints, the face, and the iris are displayed externally and can be collected without the user's consent. It can be forged by using photographs and lenses because of the nature of non-lively information. Furthermore, security problems such as malicious manipulation and personal information leakage may occur because the information cannot be changed in the event of registered information leakage [1–3]. Therefore, the relevant studies must be carried out to compensate for these disadvantages when biosignals such as electromyogram (EMG) and electrocardiogram (ECG) [4] are used.

The EMG is a biosignal generated while skeletal muscles contract, and a different signal waveform is generated, depending on the motion performed. Such a characteristic can solve the problem; the conventional personal identification methods cannot change the registered information. EMG signals display intricate details that reflect not only the physiological information of muscle tissues but also neuromuscular control information [5,6]. Furthermore, signals are conveniently acquired because they can be measured by attaching sensors to the arm and leg skin. The methods for obtaining EMG signals include invasive and non-invasive approaches. In the invasive approach, signals are sensed by inserting needle electrodes into the muscles. However, the invasive approach is unsuitable for personal identification, because of pain and discomfort. In the non-invasive method, signals are extracted from the user's skin surface, and it is suitable for personal identification because of the relative convenience in obtaining EMG signals. In particular, an EMG signal using the non-invasive approach is called a surface electromyogram (sEMG).

Personal identification using the conventional sEMG is performed with benchmarking databases (DBs) or sEMG DBs directly measured. Most benchmarking sEMG DBs are collected in a single session, not in a multi-session; thus, it ignores the variability of sEMG signals. The directly measured sEMG data are not disclosed in general, and the number of subjects is small for personal identification. Therefore, this paper constructed a multi-session benchmarking sEMG DB with a large number of subjects to solve this problem. The constructed DB was named "CS_sEMG DB", and sEMG data were obtained in a multi-session from 200 subjects in time intervals of a day or longer between each session. Furthermore, we conducted personal identification experiments to verify the DB quality and data effectiveness. According to the experimental results, the accuracy for five subjects was 74.19%, demonstrating the applicability of the constructed multi-session sEMG DB.

This paper is organized as follows. Section 2 analyzes the sEMG DBs used for personal identification in previous studies. In Section 3, the proposed multi-session sEMG DB construction method is described. Section 4 discusses the experimental results obtained to verify the effectiveness of the multi-session sEMG DB. Finally, in Section 5, conclusions are made.

2. Related Work

Experiments for conventional sEMG have been conducted by using benchmarking DBs or DBs acquired by the researchers. Benchmarking sEMG DBs are typically associated with motion recognition and obtained from a few subjects in a single session. A DB obtained by the researchers consists of sEMG signals acquired when subjects perform a few motions.

2.1. Benchmarking sEMG DBs

Benchmarking sEMG DBs include, firstly, an sEMG developed by Dr. R. Khushaba. The sEMG (Dr. R. Khushaba) DB was divided into six sets. Set 1 consisted of data acquired in two channels when eight subjects performed 10 hand movements. Sets 2 and 3 consisted of data obtained in eight channels when eight subjects performed 15 hand movements each. Set 4 consisted of sEMG signal data records of subjects whose one hand was amputated, and Set 5 comprised data recorded for the effect of the arm direction. Set 6 consisted of data recorded for the effect of the limb positions. For Sets 1, 2, 3, and 6, a bandpass filter (BPF) and a notch filter (NF) were applied at the preprocessing stage, and sEMGs were obtained from the extensor carpi ulnaris, extensor digitorum muscle, and extensor carpi radialis longus [7–13].

Secondly, there is UCI's sEMG for basic-hand-movement dataset. It refers to data obtained from five subjects between the ages of 20 and 22 to recognize six hand movements. The data were obtained from the extensor carpi radialis and flexor carpi ulnaris. The BPF and NF were conducted as a preprocessing, and they consisted of two sets. Set 1 comprised data obtained when the five subjects repeated the hand movements 30 times in a single-session, and Set 2 comprised data obtained in a multi-session by having one subject repeat each movement 100 times [14,15].

Thirdly, there is UCI's EMG dataset in the lower-limb dataset. It refers to data obtained by observing the behavior of muscles around the knee, where 11 normal subjects and 11 patients performed three movements: sitting down, standing up, and walking. The sEMG signals were recorded from the rectus femoris muscle, biceps femoris, vastus medialis, and semitendinosus [16].

Fourthly, there is Ninapro DB2. It consisted of three sets for 49 hand movements performed by 40 subjects. Set 1 consisted of data acquired while spreading or rotating the fingers and wrist. Set 2 consisted of data acquired while grabbing or holding an object. Set 3 contained data obtained while bending a single finger or multiple fingers. Each movement lasted 5 s, with a 3 s rest between movements. The movements were repeated six times, and the data were obtained in 12 channels [17,18].

Lastly, there is Ninapro DB5. DB5 consisted of three sets with 52 hand movements performed by 10 subjects. Set 1 was obtained for basic finger movements, and Sets 2 and 3 were constructed similar to DB2. Each motion lasted 5 s, with a 3 s rest between movements. The movements were repeated six times, and the data were obtained in 16 channels [19,20].

2.2. Measured sEMG DBs

Li [21] obtained sEMGs for the movement of unlocking the screen of a smartphone. While 10 subjects were comfortably sitting in chairs, signals were obtained from each subject's right hand's flexor digitorum superficialis. Each screen-unlocking pattern was repeated 20 times, with the subjects resting for one minute before performing the pattern. The sEMG data were filtered by using a 5 Hz high-pass filter (HPF) and 60 Hz NF. Lu [22] used a Myo armband to obtain data in eight channels. A total of 21 subjects performed the hand-opening movement 30 times each. The repeated process involved a 2 s resting before movement, a 1.5 s movement, and a 2 s resting after movement.

Said [23] used sEMG to generate passwords by combining hand gestures. Fifty-six healthy subjects (16–62 ages) selected three gestures out of four gestures (fist the hand, open the fingers, wave-out, and wave-in), and sEMG was measured in random order, using the Myo armband. Each subject repeated the pattern of his/her choice 20 times.

Fan [24] measured sEMG while holding a smartphone on the desk. The gesture of holding and viewing a smartphone was not defined in advance, and 40 subjects measured sEMG in eight channels, using Myo armband. The disadvantage associated with sensor position is solved by using data augmentation technology that rolls the channel of the Myo armband. Yamaba [25] measured sEMG for a user-recognition method, using a "pass-gesture". The sEMG was measured from 11 subjects in a single channel, using DL-3100 and DL-141. Nine gestures in which the sEMG was clearly generated compared to other gestures were used. Among them, four gestures were selected to compose a "pass-gesture". Each gesture was measured a total of 30 times.

Jiang [26] constructed "high density sEMG (HD-sEMG) Recordings" for neural interface research. During two sessions, sEMG was measured from 20 subjects in 256 channels in the forearm muscle. The data consist of five sets. Set 1 consists of sEMG signals measured while performing 34 commonly used hand gestures. Sets 2–5 consist of sEMG signals measured by moving individual fingers. Raurale [27] obtained data by using an eight-channel Myo armband for sEMG-based personal identification; sEMG was obtained from five subjects while sitting. The movements performed were hand open, wrist flexion, wrist pronation, wrist ulnar flexion, wrist supination, hand close, wrist extension, and wrist radial flexion. Each movement was maintained for 10 s and repeated 20 times. Furthermore, the data were constructed in a multi-session to reflect signal variability. Jiang [28] used a 64-channel electrode array sensor to acquire HD-sEMG. Twenty-two subjects performed eight movements with an electrode attached to the back of the right hand, and they rested for 10 s before performing the next movement. A BPF of 10–900 Hz was performed for the obtained data, and multi-session-based data were constructed, with an average interval of 9 days between sessions.

Table 1 lists the existing sEMG DBs and the target DB. Every benchmarking DB was constructed in a single session, except for the multi-session-based DB for only one subject. Therefore, a problem arises that the variability of signals is neglected in personal identification that uses a single-session-based sEMG. On the other hand, the sEMG acquired by the researchers has other problems: the data are not accessible, or the number of subjects is small. To solve these problems, we constructed a multi-session benchmarking sEMG DB, which is now publicly accessible for personal identification. Subsequently, we conducted a personal identification experiment to verify the effectiveness of the data.

Table 1. Comparison of sEMG databases (DBs).

Data Type	DB Names	Number of Channels	Number of Subjects	Session	Number of Motions
Bench marking	sEMG (Dr. Rami Khushaba): Sets 1–3 [7–13]	2–8	8	Single	10–15
	UCI's sEMG for basic hand movements dataset [14,15]	2	5	Single	6
		2	1	Multiple	6
	UCI's EMG dataset in lower-limb dataset [16]	4	22	Single	3
	Ninapro DB2 [17,18]	12	40	Single	49
	Ninapro DB5 [19,20]	16	10	Single	52
Measured	Li et al. [21]	1	10	Single	2
	Lu et al. [22]	8	21	Single	1
	Said et al. [23]	8	56	Single	4
	Fan et al. [24]	8	40	Single	1
	Yamaba et al. [25]	1	11	Single	9
	Jiang et al. [26]: Set 1	256	20	Multiple	34
	Raurale et al. [27]	8	5	Multiple	8
	Jiang et al. [28]	64	22	Multiple	8
The target DB	CS_sEMG DB	2	200	Multiple	12

2.3. Personal Identification Methods

Personal identification using sEMG is a process that distinguishes the differences between individuals caused by muscle development, activity, and habits when performing a particular gesture. Features for recognizing individuals are categorized into two parts. One part is the handcraft features in time and frequency domains, and the other is the features generated by a neural network. Li [21] studied the personal identification method that uses handcraft features. In the study, the measured sEMG was filtered to remove noises, and 11 handcraft features, such as mean absolute value (MAV), variance (VAR), waveform length (WL), zero crossing (ZC), etc., were extracted. A one-class support vector machine (OCSVM) and local outlier factor (LOF) were used as classifiers to recognize users with an accuracy of 98.2%. Yamaba [25] reorganized the training data by excluding the data that deviate from the mean by employing correlation coefficients and cross-correlation functions. The selected EMG signal was divided into 10 segments, and 11 handcraft features, such as sum, mean, skewness, standard deviation (SD), etc., were extracted from each segment and classified by a support vector machine (SVM) and dynamic time warping (DTW). Jiang [28] employed WL, frequency median (FMD), and spatial synchronization (SS) as features to recognize users and make each feature have equal weight, using an energy constraint. Twenty-two subjects were classified by using K-nearest neighbor (KNN), resulting in 85.8% accuracy.

Lu [22] conducted a user recognition by using neural networks. The measured sEMG was transformed into a continuous wavelet transform (CWT) to use time–frequency characteristics. He designed a convolutional neural network (CNN) composed of 4 convolutional layers and pooling layers, and it recognized users with an accuracy of 99.20%. Fan [24] did not preprocess a signal in order to avoid information loss. A Siamese network consisting of three convolutional layers with different number of filters and one fully connected layer structure was used, and the first convolutional layer learns inter-channel features. The output difference of the two subnetworks was calculated by using the Euclidean distance (ED), and an accuracy of 92.06% was achieved. Raurale [27] extracts features in sEMG by using band power (BP), root absolute sum square (RSS), and kernel fisher discrimination (KFD). Multilayer perceptron (MLP) and radial basis function (RBF) were used as classifiers, and the performance was 92.08%. S. H. Shin [29] used EMG signals measured from the fist gestures. Five handcraft features were extracted in the time domain, including SD, mean, and ZC. User recognition was carried out by an artificial neural network (ANN). Table 2 is used to summarize the personal recognition studies using the existing sEMG.

Table 2. Existing personal identification methods using sEMG.

Feature Type	Authors	Features	Classification
Handcraft	Li et al. [21]	MAV, VAR, RMS, etc.	OCSVM, LOF
	Yamaba et al. [25]	Skewness, SD, etc.	SVM, DTW
	Jiang et al. [28]	WL, FMD, SS	KNN
Neural Network	Lu et al. [22]	CWT, CNN	CNN
	Fan et al. [24]	Siamese CNN	Siamese CNN
Handcraft + Neural Network	Raurale et al. [27]	BP, RSS, KFD	MLP, RBF
	Shin et al. [29]	SD, ZC, etc.	ANN

3. Multi-Session sEMG DB Construction for Personal Identification

This section explains the method for building a multi-session benchmarking sEMG DB. In the constructed DB, sEMG was obtained in a multi-session from 200 subjects in intervals of a day or longer. In addition, a personal identification experiment was conducted to examine the DB quality and the effectiveness of the data.

3.1. sEMG DB Construction Method

Before proceeding with the data-acquisition process, we explained to the subjects every step of the process and the gestures required. In addition, we asked them at the end of the process about their caffeine intake, drinking habits, health condition, sleeping time, and any discomforts during the process (Figure A1 in Appendix A). Using Biopac MP160, we obtained sEMG from the right arms of 200 subjects (98 males and 102 females), whose average age was 24.69 ages (19–70 ages). The data were obtained in 2 channels by attaching Ag/AgCl sensors to the palmaris longus and extensor digitorum muscles of each subject, as shown in Figure 1. For the sensor positions, we selected the points where significant muscle changes occurred during hand gestures.

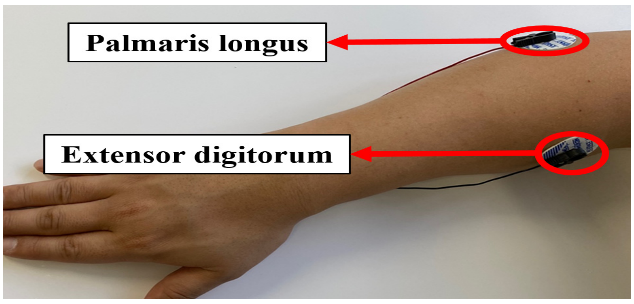

Figure 1. The sEMG sensor positions.

Figure 2 shows a data sequence for constructing the sEMG DB, which results in a file in the DB. The subjects performed each movement in a relaxed state while sitting in chairs, and each movement was repeated 10 times to obtain the data. When performing a gesture, the gesture was maintained for at least one second, proceeding with the "start (relaxing state)–hand gesture–end (relaxing state)" steps. The signals were obtained at a sampling rate of 2000 Hz, with a 16-bit ADC resolution. The subjects were selected among students, researchers at Chosun University, and people not associated with Chosun University.

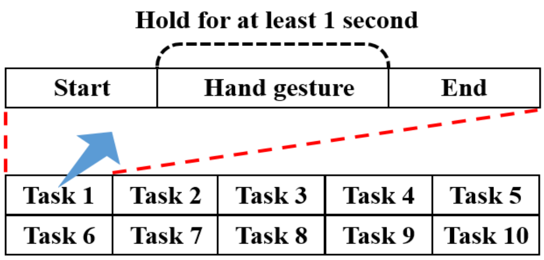

Figure 2. A data sequence of sEMG data acquisition for each gesture.

When obtaining sEMG, the subjects performed 12 gestures (Figure 3). The gestures were divided into static and dynamic gestures. For the hand gestures, we selected movements commonly used in everyday life. A static gesture is a motion of performing a single movement, whereas a dynamic gesture is a continuous movement. The sEMG signals were obtained during such gestures. CS_sEMG DB was obtained in 3 sessions to reflect the variability of sEMG signals in intervals of a day or longer between sessions. From each subject, data were obtained in 3 sessions for a gesture; 30 (10 repetitions $\times$ 3 sessions) sEMG signals were obtained. Because each subject performed 12 gestures, a total of 360 (12 gestures $\times$ 10 repetitions $\times$ 3 sessions) sEMG signals were obtained per person.

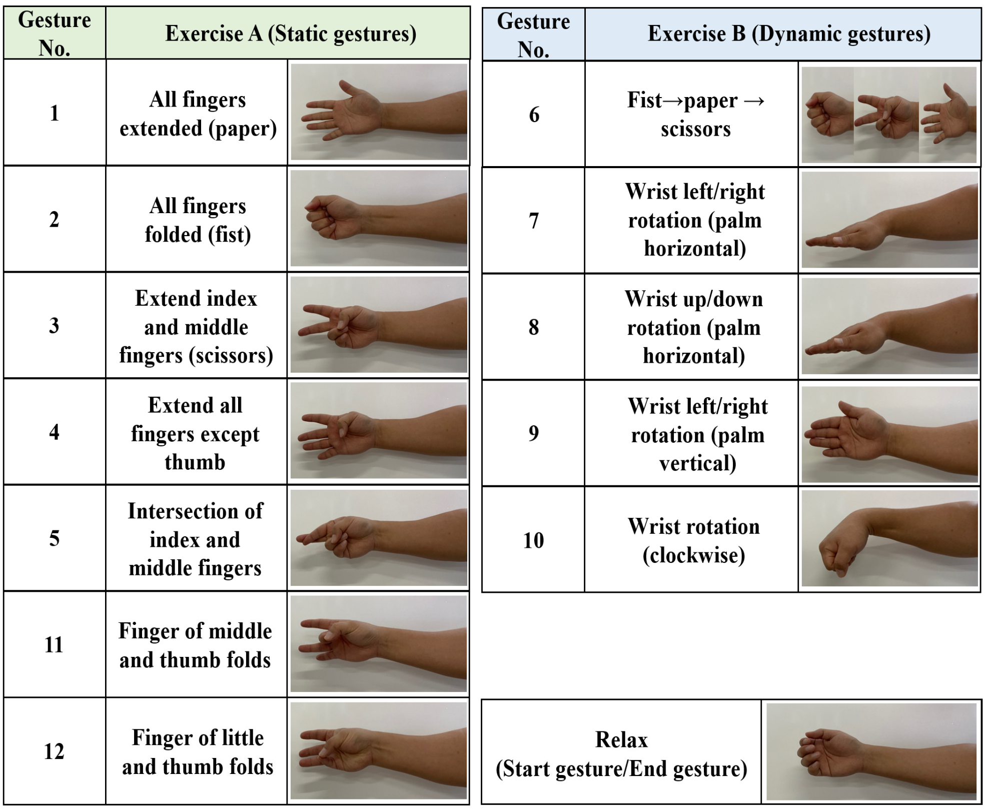

Figure 3. Hand gestures of CS_sEMG DB.

3.2. sEMG Signal Verification and Segmentation

To eliminate the effects of the incorrect measurements by acquirers and subjects, we excluded the subjects who produced one or more incorrect sEMG signals. The sEMG signals of 116 subjects, 58% of 200 subjects, were acquired correctly. Hence, 84 subjects were excluded for the following reasons (red box in Figure 4): "Device loses Bluetooth connection", "Incompletion of muscle contraction/relaxation", "Insufficient number of repetitions", and "Muscle not activated". Table 3 lists the labels of the excluded subjects; the data of the excluded subjects were not used in the experiments.

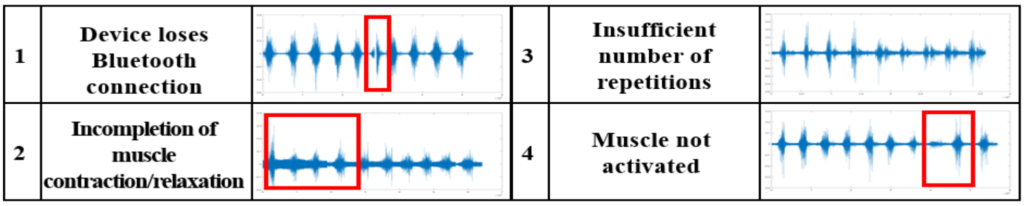

Figure 4. CS_sEMG signal-error types.

Table 3. Removed subject labels with the corresponding error type.

Error Type	Subject No.
Device loses Bluetooth connection	2, 3, 5, 8, 10, 21, 28, 39, 44, 46, 47, 49, 57, 76, 77, 86, 89, 90, 92, 95, 100, 102, 105, 107, 108, 109, 111, 117, 120, 124, 127, 143, 144, 153, 165, 167, 172, 178, 184, 199
Incompletion of muscle contraction/relaxation	14, 20, 30, 33, 34, 45, 60, 67, 72, 81, 93, 114, 132, 140, 146, 151, 154, 159, 168, 173, 176, 182, 186, 191, 193
Insufficient number of repetitions	6, 16, 37, 69, 78, 116, 152, 196
Muscle not activated	17, 24, 29, 53, 65, 73, 121, 158, 161, 162, 181

The signals for each gesture were obtained continuously and represented as a file for 10 repetitions. The correctly acquired sEMG signals of the subjects were validated with the naked eye and then segmented manually after validation. Figure 5 shows the result obtained by the segmented sEMG signals for Gesture 1 of Subject 1. The constructed DB is publicly available at the IT Research Institute of Chosun University (http://www.chosun.ac.kr/riit, accessed on 21 February 2022), and the unfiltered raw sEMG data are provided in the form of both text files and MATLAB files. Data are available for non-commercial purposes, and only EMG data of Gesture Numbers 1 to 3 performed by 100 subjects (including subjects excluded by error type) are disclosed. The complete data will be available after signing an MOU with the IT Research Institute of Chosun University in public. The text file contains sEMG signals without notions of repetitions, whereas the MATLAB file contains sEMG signals with notions of repetitions. The provided DB file structure consists of 5 columns:

- Column 1: 2 channel sEMG data;
- Column 2: sampling information;
- Column 3: the subject number who performed the gesture;
- Column 4: the gesture number that the subject performed;
- Column 5: the corresponding session number among the 3 sessions.

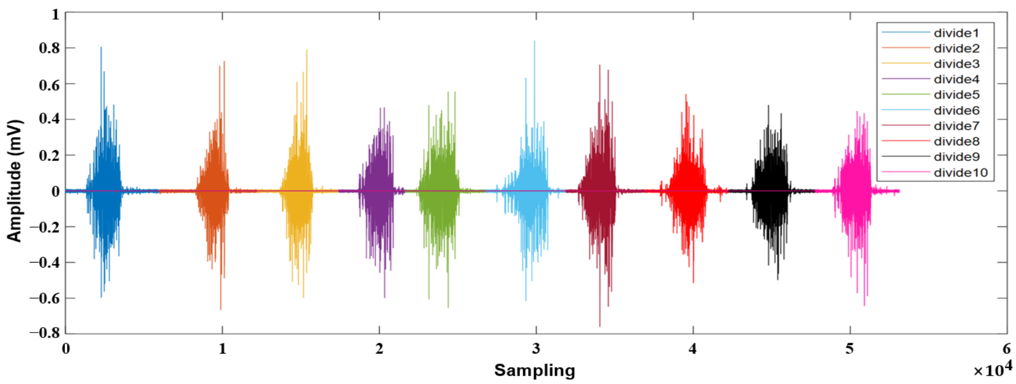

Figure 5. A sEMG signal example (Gesture 1) and its segmented result.

3.3. Personal Identification Method in the Experiment

Personal identification experiments were performed by using the single-session and multi-session sEMGs of 20 subjects to verify the effectiveness of the constructed CS_sEMG DB. Figure 6 shows the procedure of the experiment. In order to remove the baseline wander, power line noise, and external environmental noise of the measured sEMG, the notch filter (NF) in the 60 Hz band and bandpass filter (BPF) in the 5–500 Hz band were used as a preprocess. The preprocessed sEMG was converted into a spectrogram for simultaneous analysis in the time–frequency domain. The spectrogram was generated by using a fixed-length window along the time axis of a given signal. The generated spectrogram was resized into a 256 × 256 image and inputted to the designed CNN–LSTM. The designed CNN–LSTM uses the model consisting of 3 convolutional layers and 2 long short-term memory (LSTM) layers in the previous study [4]. The parameters of the CNN–LSTM are shown in Table 4, and the network learning was performed by using batch size 50, epoch 100, and sgdm optimizer.

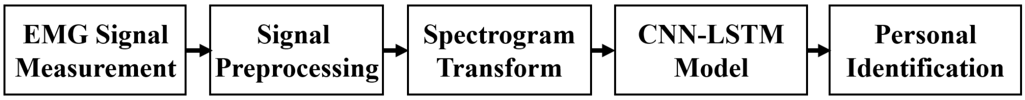

Figure 6. Convolutional neural network–long short-term memory (CNN–LSTM)-based personal identification procedure using CS_sEMG DB.

Table 4. The CNN–LSTM model structure used.

Layer	Input Size	Output Size	Number of Filter (Hidden Unit)
Input	-	256 × 256 × 3	-
Conv1	256 × 256 × 3	256 × 256 × 20	20
Pool1	256 × 256 × 20	128 × 128 × 20	-
Conv2	128 × 128 × 20	128 × 128 × 20	20
Pool2	128 × 128 × 20	64 × 64 × 20	-
Conv3	64 × 64 × 20	64 × 64 × 20	20
Pool3	64 × 64 × 20	32 × 32 × 20	-
LSTM1	20,480	1024	1024
LSTM2	1024	1024	1024
FC1	1024	512	-
FC2	512	Class (5, 10, 15, 20)	-

4. Experimental Results and Discussion

The personal identification experiments for verifying the effectiveness of the constructed CS_sEMG DB were conducted by using both single-session and multi-session signals. In the single-session-based personal identification experiment, we divided the sEMG signals into 70% for training and 30% for testing, yielding 84 training data and 36 test data for each subject. Table 5 lists the experimental results for the single-session sEMG. Based on the experimental results, 5, 10, 15, and 20 subjects produced an average accuracy of 93.22%, 90.47%, 87.09%, and 84.65%, respectively. Given that the same subjects participated, a difference in the personal identification accuracy between sessions occurred because of the variability of the sEMG signals. Such variations could minorly occur as a result of the changes of the electrodes' positions. When sEMG is acquired in each session, the electrodes could slide, because the adhesion between the skin and the electrodes could change, owing to sweat and other contaminants.

Table 5. Personal identification results for the single-session sEMG signals.

sEMG Session No.	Number of Subjects	Accuracy (%)
1	5	92.05
	10	91.11
	15	84.44
	20	83.72
2	5	94.92
	10	91.9
	15	89.84
	20	85.32
3	5	92.7
	10	88.41
	15	86.98
	20	84.91
Average	5	93.22
	10	90.47
	15	87.09
	20	84.65

During the multi-session-based personal identification experiment, two sessions were used for the model training, and the remaining one session was used for testing. For each subject, 240 training data and 120 test data were used. Table 6 lists the experimental results for the multi-session sEMG. Overall, 5, 10, 15, and 20 subjects produced an average accuracy of 74.19%, 55.73%, 45.50%, and 41.63%, respectively. It showed that a small group of people could be identified by using the multi-session sEMG. The experimental results were obtained by applying the previous method [4] without any modifications or improvements. Thus, when the multi-session sEMG DB was used, the personal identification accuracy of 20 subjects was relatively low. The performance can be improved by applying advanced feature-extraction and classification techniques.

Table 6. Personal identification results for the multi-session sEMG signals.

sEMG Session No.		Number of Subjects	Accuracy (%)
Training	Testing		
1, 2	3	5	69.91
		10	55.19
		15	51.46
		20	46.88
1, 3	2	5	81.14
		10	59.28
		15	45.14
		20	40.48

Table 6. *Cont.*

sEMG Session No.		Number of Subjects	Accuracy (%)
Training	Testing		
2, 3	1	5	71.52
		10	52.71
		15	39.91
		20	37.53
Average		5	74.19
		10	55.73
		15	45.5
		20	41.63

The existing sEMG-based personal identification method uses two sessions for learning and one for testing. The learning and testing sessions are mutually exclusive. As a result of the experiment, the performance of the personal identification method using multi-session data was similar to that of the previous study [4]. In particular, in the case of recognizing five subjects, every method using handcraft features achieved over 71% accuracy. Therefore, it is confirmed that the CU_sEMG DB is valid as a dataset for personal identification. Figure 7 summarizes the results.

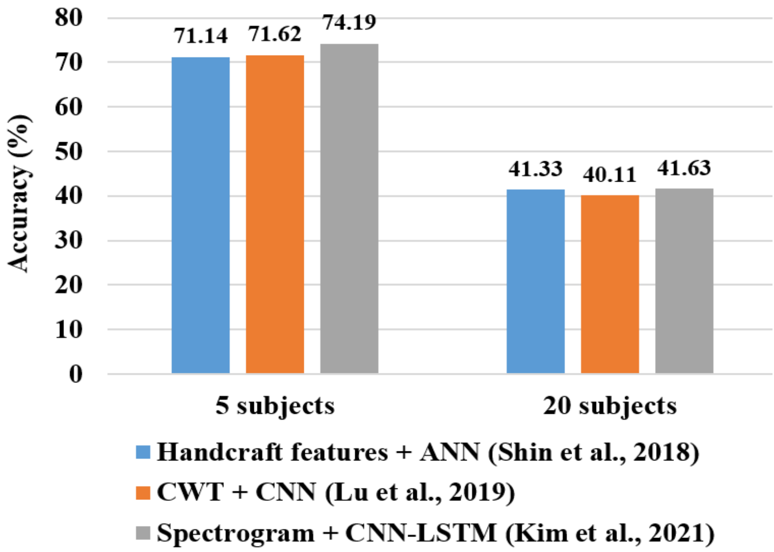

Figure 7. Personal identification experiment results using the existing methods [4,22,29].

5. Conclusions

In the literature, many personal identification methods based on sEMG have been performed by using benchmarking DBs or DBs acquired by the researchers. However, they have disadvantages, as the data reliability is vulnerable, owing to the small number of subjects, and the sEMG signal variability cannot be verified and replicated. In this study, we constructed and disclosed a large-capacity multi-session sEMG DB. Furthermore, to verify the effectiveness of the constructed CS_sEMG DB, we conducted personal identification experiments by using a previously developed method. The DB was constructed in two channels when 200 subjects performed 12 gestures. Each gesture was repeated 10 times in time intervals of a day or longer between three sessions. The constructed CS_sEMG DB was validated for quality, whereby the incorrectly acquired sEMG was excluded by the investigators, and the signals from the repeating gestures were segmented manually. The experiments for verifying the effectiveness of the DB were performed by using both single-session and multi-session sEMG DBs. As a result of the experiment using the existing methods, personal identification using multi-session sEMG achieved over 71% accuracy for five subjects. Consequently, CU_sEMG DB, the sEMG database created by the IT Research Institute of Chosun University, proved to be valid as a dataset for personal identification. Therefore, we concluded that the multi-session sEMG data could be used for personal identification. The CS_sEMG DB developed in this paper will contribute to the sEMG-based personal identification studies. In the future, we will continuously conduct personal identification experiments by using the CS_sEMG DB and research the advanced feature extraction and classification techniques for a better performance.

Author Contributions: Conceptualization, J.-S.K., C.-H.S. and E.B.; Methodology, J.-S.K. and E.B.; Software, J.-S.K. and C.-H.S.; Validation, J.-S.K., E.B. 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., C.-H.S., E.B. 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 the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (No. 2017R1A6A1A03015496); and the NRF grant, funded by the Korean government (MSIT) (No. NRF-2021R1A2C1014033).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data supporting the findings presented in this article are available at http://www.chosun.ac.kr/riit (accessed on 21 February 2022).

Acknowledgments: This paper presents the results of a study on the "HPC Support Project", supported by the "Ministry of Science and ICT" and NIPA.

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

Abbreviations

sEMG	surface electromyogram
DB	database
EMG	electromyogram
ECG	electrocardiogram
BPF	bandpass filter
NF	notch filter
HD-sEMG	high-density surface electromyogram
HPF	high-pass filter
LSTM	long short-term memory
CNN–LSTM	convolutional neural network–long short-term memory

Appendix A

<Questionnaire on Biosignal Construction Experiment>

This questionnaire is used by the IT Research Institute to access your conditions accurately to help construct a biosignal DB and personal identification research. The contents of this questionnaire are used as basic data for research and product testing and will not be used for any other purpose.

Name: ___________, Sex/Age: ____/_____

Date of Experiment: ________ (Year/Month/Day)

Start: ____:____:____, Finish: ____:____:____

Height: ______cm, Weight: ______kg

◆ Please mark the illnesses you currently have among those listed below, and if they are not on the list, please write them down.

Illness you currently have	Hospital where you are receiving treatment	Are you taking medications?
□ Diabetes		□ Yes □ No
□ Hypertension		□ Yes □ No
□ Heart arrhythmia		□ Yes □ No
□ Lung disease		□ Yes □ No
□ Thyroid disease		□ Yes □ No
□ Depression		□ Yes □ No

◆ Please write down any medicines (including nutritional supplements) that you have taken in the last two weeks or are taking currently.

____

◆ On average, how much do you consume the following beverages?

- Do you drink coffee? □ Yes □ No

If yes, how much do you do drink? ____ times per week, Coffee type: _______ Amount consumed: ____

When was the last time you drank coffee? (Please specify the time and amount.) ______________

- Do you drink alcohol? □ Yes □ No

If yes, how much do you drink? ____ times per week, Type of drink: _______ Drinking capacity: _____

When was the last time you drank alcohol? (Please specify the time and amount.) ______________

- Have you ever smoked, or do you smoke? □ Yes □ No

On average, how much did/do you smoke per day? ____pack/day or ____cigarettes/day

How old were you when you started smoking? ________ years old

If you have quit, how many years did you quit? ________

When was the last time you smoked? (Please specify the time and amount.) ______________

◆ What is your body condition today?

- What is your current body condition?

① Very tired ② a little tired ③ nothing special ④ feel refreshing and invigorating

- When is your bedtime?

- How many hours do you sleep?

< Questionnaire After Experiment>

- What was your tension/drowsiness level during the experiment?

① Heavily tensed ② a little tensed ③ normal ④ a little drowsy ⑤ heavily drowsy

- Was it cold or hot during the experiment?

① Very cold ② a little cold ③ average ④ a little hot ⑤ very hot

- What was your body condition before the experiment?

① Very tired ② a little tired ③ nothing special ④ refreshing and invigorating

- What was your body condition during the experiment?

① Very tired ② a little tired ③ nothing special ④ refreshing and invigorating

- How do you feel now after the experiment?

① Very tired ② a little tired ③ nothing special ④ refreshing and invigorating

- Did you experience any physical discomfort during the experiment?

① Very uncomfortable ② a little uncomfortable ③ normal ④ a little comfortable ⑤ very comfortable

- If you experienced any discomfort during the experiment or have suggestions for improvements, please specify.

____

Thank you for participating in this study. I hereby consent to the use of the contents of this questionnaire as reference data for medical treatment and research activities.

Name__________ (Signature)__________

Figure A1. Post-test questionnaires for investigating a subject's behaviors.

References

1. Kim, J.S.; Pan, S.B. A study on EMG-based biometrics. *J. Internet Serv. Inf. Secur.* **2017**, *7*, 19–31.
2. Caputo, D.; Verderame, L.; Ranieri, A.; Merlo, A.; Caviglione, L. Fine-hearing google home: Why silence will not protect your privacy. *J. Wirel. Mob. Netw. Ubiquitous Comput. Dependable Appl.* **2020**, *11*, 35–53.
3. Ranieri, A.; Caputo, D.; Verderame, L.; Merlo, A.; Caviglione, L. Deep adversarial learning on google home device. *J. Internet Serv. Inf. Secur.* **2021**, *11*, 33–43.
4. 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. [CrossRef]
5. Karuna, M.; Guntur, S.R. EMG signal analysis using intrinsic mode functions to discriminate upper limb movements. In Proceedings of the International Conference on Artificial Intelligence and Signal Processing, Amaravati, India, 10–12 January 2020.
6. Song, X.; Guo, S.; Gao, B.; Wang, Z. Motion recognition of the bilateral upper-limb rehabilitation using sEMG based on ensemble EMD. In Proceedings of the IEEE International Conference on Mechatronics and Automation, Tianjin, China, 3–6 August 2014.
7. Khushaba, R.N.; Kodagoda, S.; Takruri, M.; Dissanayake, G. Toward improved control of prosthetic fingers using surface electromyogram (EMG) signals. *Expert Syst. Appl.* **2012**, *39*, 10731–10738. [CrossRef]
8. Khushaba, R.N.; Kodagoda, S. Electromyogram (EMG) feature reduction using mutual components analysis for multifunction prosthetic fingers control. In Proceedings of the International Conference on Control Automation Robotics and Vision, Guangzhou, China, 5–7 December 2012.
9. Khushaba, R.N.; Kodagoda, S.; Liu, D.; Dissanayake, G. Muscle computer interfaces for driver distraction reduction. *Comput. Methods Programs Biomed.* **2013**, *110*, 137–149. [CrossRef] [PubMed]
10. AI-Timemy, A.H.; Khushaba, R.N.; Bugmann, G.; Escudero, J. Improving the performance against force variation of EMG controlled multifunctional upper-limb prostheses for transradial amputees. *IEEE Trans. Neural Syst. Rehabilitation Eng.* **2015**, *24*, 650–661. [CrossRef] [PubMed]

11. Khushaba, R.N.; AI-Timemy, A.; Kodagoda, S.; Nazarpour, K. Combined influence of forearm orientation and muscular contraction of EMG pattern recognition. *Expert Syst. Appl.* **2016**, *61*, 154–161. [CrossRef]
12. Khushaba, R.N.; Takruri, M.; Miro, J.V.; Kodagoda, S. Towards limb position invariant myoelectric pattern recognition using time-dependent spectral features. *Neural Netw.* **2014**, *55*, 42–58. [CrossRef] [PubMed]
13. EMG Datasets Repository. Available online: http://www.rami-khushaba.com/electromyogram-emg-repository.html (accessed on 21 February 2022).
14. Sapsanis, C.; Georgoulas, G.; Tzes, A.; Lymberopoulos, D. Improving EMG based classification of basic hand movements using EMD. In Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society, Osaka, Japan, 3–7 July 2013.
15. UCI Machine Learning Repository: sEMG for Basic Hand Movements Data Set. Available online: http://archive.ics.uci.edu/ml/datasets/sEMG+for+Basic+Hand+movements (accessed on 21 February 2022).
16. UCI Machine Learning Repository: EMG Dataset in Lower Limb Data Set. Available online: http://archive.ics.uci.edu/ml/datasets/emg+dataset+in+lower+limb (accessed on 21 February 2022).
17. Atzori, M.; Gijsberts, 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*, 140053. [CrossRef] [PubMed]
18. Ninaweb. Available online: http://ninapro.hevs.ch/node/17 (accessed on 21 February 2022).
19. Pizzolato, S.; Tagliapietra, L.; Cognolato, M.; Reggiani, M.; Muller, H.; Atzori, M. Comparison of six electromyography acquisition setups on hand movement classification tasks. *PLoS ONE* **2017**, *12*, e0186132. [CrossRef] [PubMed]
20. Ninaweb. Available online: http://ninapro.hevs.ch/DB5_DoubleMyo (accessed on 21 February 2022).
21. Li, Q.; Dong, P.; Zheng, J. Enhancing the security of pattern unlock with surface EMG-based biometrics. *Appl. Sci.* **2020**, *10*, 541. [CrossRef]
22. 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 Biomedical Circuits and Systems Conferences, Nara, Japan, 17–19 October 2019.
23. Said, S.; Karar, A.S.; Beyrouthy, T.; Alkork, S.; Nait-ali, A. Biometrics verification modality using multi-channel sEMG wearable bracelet. *Appl. Sci.* **2020**, *10*, 6960. [CrossRef]
24. Fan, B.; Liu, X.; Su, X.; Hui, P.; Niu, J. EmgAuth: An EMG-based smartphone unlocking system using siamese network. In Proceedings of the International Conference on Pervasive Computing and Communications, Austin, TX, USA, 23–27 March 2020.
25. Yamaba, H.; Usuzaki, S.; Takatsuka, K.; Aburada, K.; Katayama, T. On a user authentication method to realise an authentication system using s-EMG. *Int. J. Grid Util. Comput.* **2020**, *11*, 725–734. [CrossRef]
26. Jiang, X.; Liu, X.; Fan, J.; Ye, X.; Dai, C.; Clancy, E.A.; Akay, M.; Chen, W. Open access dataset, toolbox and benchmark processing results of high-density surface electromyogram recordings. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2021**, *29*, 1035–1046. [CrossRef] [PubMed]
27. Raurale, S.A.; McAllister, J.; Rincon, J.M.D. EMG biometric systems based on different wrist-hand movements. *IEEE Access* **2021**, *9*, 12256–12266. [CrossRef]
28. Jiang, X.; Xu, K.; Liu, X.; Dai, C.; Clifton, D.A.; Clancy, E.A.; Akay, M.; Chen, W. Cancelable HD-sEMG-based biometrics for cross-application discrepant personal identification. *IEEE J. Biomed. Health Inform.* **2021**, *25*, 1070–1079. [CrossRef] [PubMed]
29. Shin, S.; Jung, J.; Kang, M.; Kim, Y.T. A study on EMG signal acquisition modules and artificial neural networks for personal authentication. In Proceedings of the International Conference on Computational Science and Computational Intelligence, Las Vegas, NV, USA, 12–14 December 2018.