

Article

Hand Gesture Recognition Using Compact CNN via Surface Electromyography Signals

Lin Chen ^{1,2,†}, Jianting Fu ^{1,2,†} , Yuheng Wu ^{1,3}, Haochen Li ^{1,2} and Bin Zheng ^{1,*}

¹ Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing 400700, China; chenlin18@cigit.ac.cn (L.C.); jiantingfu@cigit.ac.cn (J.F.); wuyuheng21@163.com (Y.W.); lihaochen@csu.ac.cn (H.L.)

² School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing 100049, China

³ School of Mechatronical Engineering, Changchun University of Science and Technology, Changchun 130022, China

* Correspondence: zhengbin@cigit.ac.cn

† These authors contributed equally.

Received: 25 December 2019; Accepted: 22 January 2020; Published: 26 January 2020



Abstract: By training the deep neural network model, the hidden features in Surface Electromyography (sEMG) signals can be extracted. The motion intention of the human can be predicted by analysis of sEMG. However, the models recently proposed by researchers often have a large number of parameters. Therefore, we designed a compact Convolution Neural Network (CNN) model, which not only improves the classification accuracy but also reduces the number of parameters in the model. Our proposed model was validated on the Ninapro DB5 Dataset and the Myo Dataset. The classification accuracy of gesture recognition achieved good results.

Keywords: surface electromyography (sEMG); convolution neural networks (CNNs); hand gesture recognition

1. Introduction

In recent years, Surface Electromyography (sEMG) signals have been widely used in artificial limb control, medical devices, human-computer interaction, and other fields. With the development of artificial intelligence and robotics technology, the intention of human hand movements can be obtained by using an artificial intelligence algorithm to analyze the sEMG signals collected from the residual limb. Robotics and artificial intelligence can be leveraged to better help the disabled people to independently complete some basic interactions in their daily life. The sEMG signals, which are non-stationary, represent the sum of subcutaneous athletic action potentials generated through muscular contraction [1]. Also, it is one of the main physical signals of an intelligent algorithm to identify motion intention.

Distinguishing sEMG signals collected from different gestures is the core part of the related applications using sEMG signals as intermediate media. At present, the literature on gesture recognition or artificial limb control by sEMG signals primarily focuses on the time and frequency domain feature extraction of sEMG signals, which aims to distinguish sEMG signals by feature recognition [1–3]. After years of exploration by researchers, some effective feature combinations have been proposed in both the time domain and frequency domain [4–6], and some fruitful results have been achieved with their respective datasets. Choosing feature extraction is particularly important in that different gestures can be distinguished by traditional methods. However, it is difficult to improve the performance of gesture recognition based on sEMG by traditional methods. Nevertheless, the process of designing

and selecting features can be complicated and the combinations of features are diverse, leading to increasing of workload and dissatisfied results [7].

Using deep neural networks to distinguish sEMG signals has been proposed by researchers. Wu et al. [7] proposed LCNN and CNN_LSTM models (These models can be thought of as autoencoders for automatic feature extraction.), which do not require the process of traditional feature extraction. In recent years, deep learning has achieved great success in the field of image recognition. An important idea was put forward in [8,9] that the signals of a channel can form a graph, after the short-time Fourier transform or wavelet transform of sEMG signals. It was a good idea to convert the sEMG signal into an image and inspired us with the transform of the sEMG signal. Researchers such as Côté-Allard et al. [8], who regarded the original sEMG signals as an image, constructed the ConvNet model to further improve the classification accuracy of sEMG signals. However, the LCNN and CNN_LSTM models proposed by Wu et al. [7], and the ConvNet model used by Côté-Allard et al. [8], contain a large number of parameters.

For deep learning algorithms, the final test accuracy is directly affected by the size of the training data, but one participant cannot be expected to generate tens of thousands of examples in one experiment during the data collection process. However, a large amount of data can be obtained by aggregating the records of multiple participants, so that the model can be pre-trained to reduce the amount of data required by new participants. On the other hand, designing a compact network structure to reduce the number of parameters can also reduce the demand for data size.

In order to reduce the number of model parameters and improve the accuracy of model classification, we present a new compact deep convolutional neural network model for gesture recognition, called as EMGNet. It was validated on the Myo Dataset that the average recognition accuracy of EMGNet can achieve 98.81%. The NinaPro DB5 dataset has often been used to test classical machine learning methods. The accuracy of the EMGNet on these datasets was higher than that of the traditional machine learning methods. Figure 1 shows the overall flow chart of sEMG signal acquisition and identification.

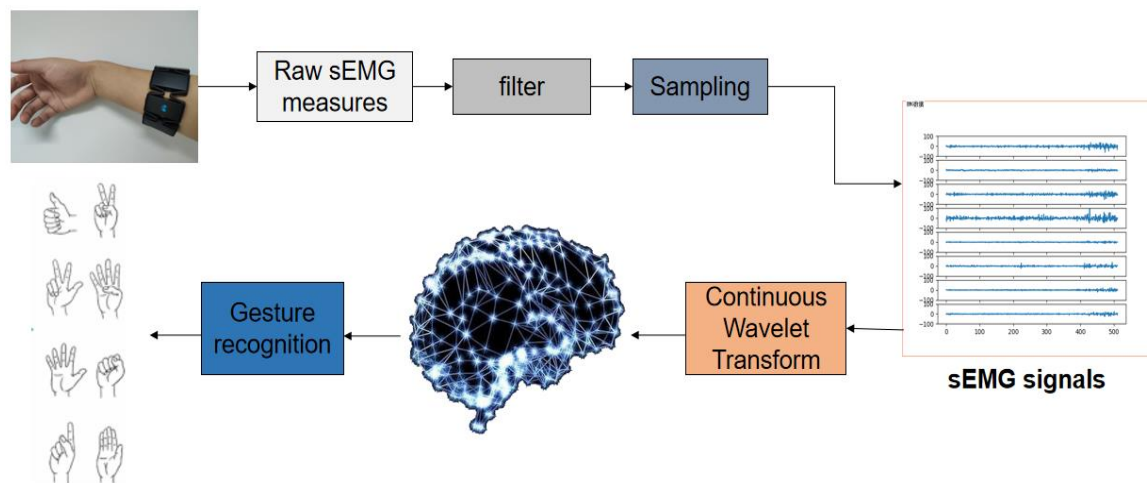


Figure 1. Surface Electromyography(sEMG) signals collection and classification process. Myo armband was used to collect the original signal, and then the collected signal was filtered and sampled to get sEMG signals. Continuous wavelet transform was selected to obtain the signal spectrum, and the neural network model was applied to classify the spectrum to achieve gesture recognition.

The rest of this paper is organized as follows. The related work of gesture recognition through deep learning is outlined in Section 2. The data processing and network architecture are described in detail in Section 3. The proposed network model is compared with the current excellent deep network framework and the classical machine learning classification method in Section 4. Finally, we present the conclusion in Section 5.

2. Related Work

People will produce different signals when completing the same action even lots of precise control electrodes are used to sense them [10]. Therefore, it is difficult to recognize sEMG signals. Since the AlexNet network proposed by Krizhevsky et al. [11] won the ImageNet challenge in 2012, deep learning has achieved great success in image classification, speech recognition, and other fields. Images can be accurately classified by training the neural network model to learn the characteristics of images. Nowadays, exploring network architecture has become part of deep learning research.

Currently, some researchers have successfully applied deep learning to sEMG signal classification and explored several effective network frameworks [8,12–16]. Using CNN to classify sEMG signals, literature [12,17,18] took the raw signals as input space. The spectrograms of raw sEMG signals were obtained by Short-Time Fourier Transform (STFT) and fed into the convolutional network (ConvNets) [13,19]. Literature [8] used the ConvNets to classify the characterizations of the sEMG signals extracted by short-time Fourier transform-based spectrogram and Continuous Wavelet Transform (CWT). Since sEMG signals correspond to the timing signal, we proposed to classify the sEMG signal by combining Long Short-Term Memory (LSTM) and CNN from our previous work [7]. The temporal information in the signal is retained and the ability of CNN to extract features is utilized. We take advantage of the complementarity of CNNs and LSTMs by combing them into one unified architecture. Meanwhile, we analyze the effect of adding CNN before the LSTM. We propose LCNN and CNN-LSTM models, which can directly input pre-processed EMG signals into the network [7]. In practical work, we verified that the performance of the LCNN model is better than CNN-LSTM. Figure 2 simply depicts the architecture of the LCNN model. We use PReLU [20] as the non-linear activation function. ADAM [21] is utilized for the optimization of the model.

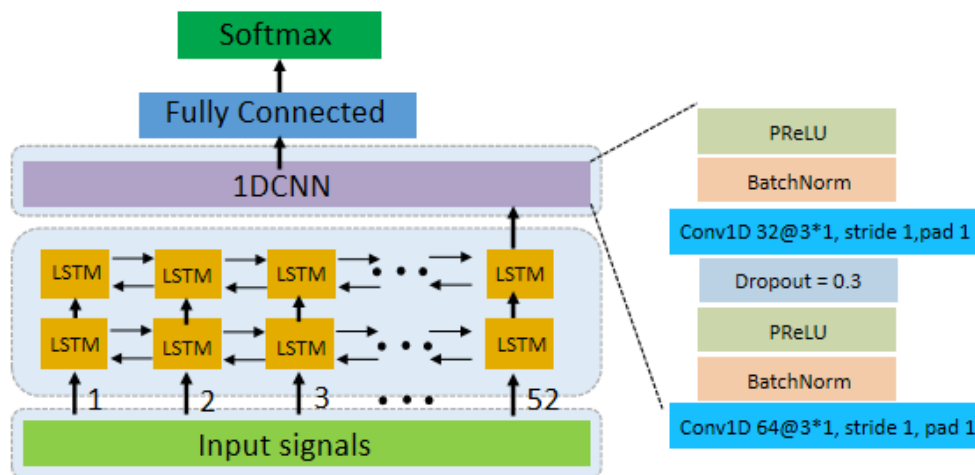


Figure 2. LCNN Architecture diagram, the LCNN consists of 2 LSTM layers, 2 one-dimensional convolution layers and 1 output layer. We use 2 LSTM layers, and each LSTM layer has 52 cells, and every cell has 64 hidden layers.

However, the ConvNets model (Shown in Figure 3) in [8] was too complicated, and the LSTM model was introduced in [7], which led to expensive computation in gesture recognition. Therefore, a new network model was proposed in this paper, and it was proved by experiment that this model not only improves the accuracy of recognition, but also reduces the complexity of the network model.

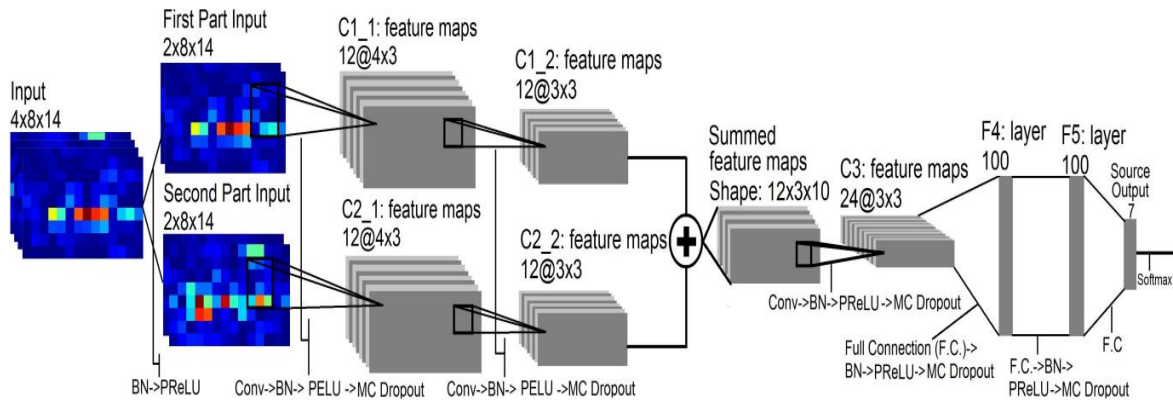


Figure 3. Schematic diagram of ConvNet architecture. In this figure, Conv refer to Convolution and F.C. to Fully Connected layers.

3. sEMG Signals Recognition Algorithm

3.1. sEMG Signals Feature Extraction

Designing a sEMG signal feature is one of the main tasks of the algorithm. It can not only easily distinguish the sEMG signals produced by different movements, but also maintain a small variance between the sEMG signals produced by the same movements. Four feature sets—Time domain characteristics (TD) [5], Enhanced TD [4], Nina Pro Features [6,22], and SampEn Pipeline [23]—were selected as features of the classical machine learning algorithm compared to the proposed method. The TD feature set includes four features: Mean Absolute Value (MAV), Zero Crossing (ZC), Slope Sign Changes (SSC), and Waveform Length (WL) [5]. The Enhanced TD includes features such as Root Mean Square (RMS) and Autoregressive Coefficient (AC) [3]. Nina Pro Features includes Root Mean Square, TD features, etc. [6,22]. SampEn Pipeline includes features such as Sample Entropy (SampEn) [24], Root Mean Square, and Waveform Length [23].

Since the sEMG signals is a non-stationary signal [25], the analysis of these signals using the Fourier transform is limited. One technique to solve this problem is the Short Time Fourier Transform (STFT), which separates the signal into smaller segments by applying a sliding window and separately calculates the Fourier transform for each of the segments. The spectrogram of the signal can be calculated from the square of the signal after STFT transformation. When the signal $x(t)$ and the window function $w(t)$ are given, the spectrogram is calculated as follows:

$$\text{spectrogram}(x(t), w(t)) = |STFT_x(t, f)|^2 \quad (1)$$

$$STFT_x(t, f) = \int_{-\infty}^{+\infty} [x(u)w(u-t)]e^{-j2\pi fu} du \quad (2)$$

where f represents frequency. The wavelet transform(WT) is similar to the STFT, but it overcomes the shortcomings of the window in the STFT which does not change with frequency. The WT adapts to the change in frequency in the signal by adjusting the width of the window. When the frequency in the signal climbs high, the WT increases the resolution by narrowing the time window. In addition, WT, which is an ideal signal analysis tool, has the ability to obtain the amplitude and frequency of mutations in a signal.

$$X(a, b) = \frac{1}{\sqrt{b}} \int_{-\infty}^{\infty} x(t)\varphi\left(\frac{t-a}{b}\right) dt \quad (3)$$

$$\int_{-\infty}^{+\infty} \frac{|\varphi(\omega)|^2}{\omega} d\omega < \infty \quad (4)$$

where the Fourier Transform $\varphi(\omega)$ of $\varphi(t)$ must satisfy Equation (4). $\varphi(t)$, also known as the mother wavelet function, is a signal with a limited duration, varying frequency and a mean of zero. The scaling factor b controls the scaling of the wavelet function. The translation amount a controls the translation of the wavelet function.

After the continuous wavelet transform of the sEMG signal, the corresponding spectrum information will be obtained, which is similar to the image in scale and also contains the frequency-domain information of the timing sequence data. The datasets tested in this article were all collected by Myo armband, as shown in Figure 4. Myo is an 8-channel, dry-electrodes, low-sampling rate (200 Hz), low-cost consumer grade sEMG armband, which is convenient to wear and easy to use [7]. The data of one channel was separated by applying sliding windows of 52 samples (260 ms). The mother wavelet of continuous wavelet transform adopts the Mexican Hat wavelet function. The CWTs were calculated with 32 scales obtaining a 32×52 matrix. The scale (32×52) sampled down 0.5 in spectrum information is taken as the input of EMGNet model. Thus the input of the EMGNet model has 8 channels, each of which consists of a matrix of size 15×25 . Figure 5b is the spectrum of the signal shown in Figure 5a after wavelet transformation.



Figure 4. The 5 hand/wrist gestures and Myo armband. In this figure, the left is a schematic diagram of five gestures, and the right is the Myo armband.

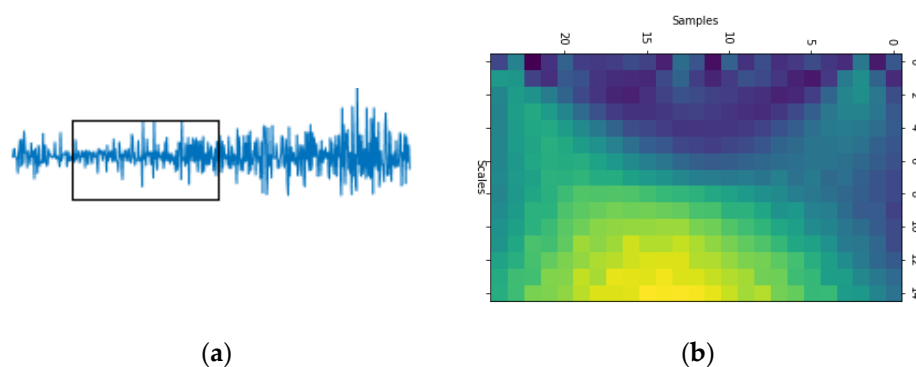


Figure 5. Part (a) shows the waveform of sEMG signals and Part (b) the spectrum of the sEMG signals shown in (a) after wavelet transformation.

3.2. EMGNet Architecture

A new network model called EMGNet is proposed in this paper, and shown in Figure 6. It consists of four convolutional layers and a max pooling layer without using the full connection layer as the final output. ConvNet architecture (Shown in Figure 3) contains 67179 learnable parameters used in the CWT+TL method proposed in Literature [8]. As shown in Table 1, the model proposed in this paper contains fewer parameters than the models used in current advanced methods. In view of the

better performance of the EMGNet model in the actual measurement, this section mainly introduces the EMGNet model.

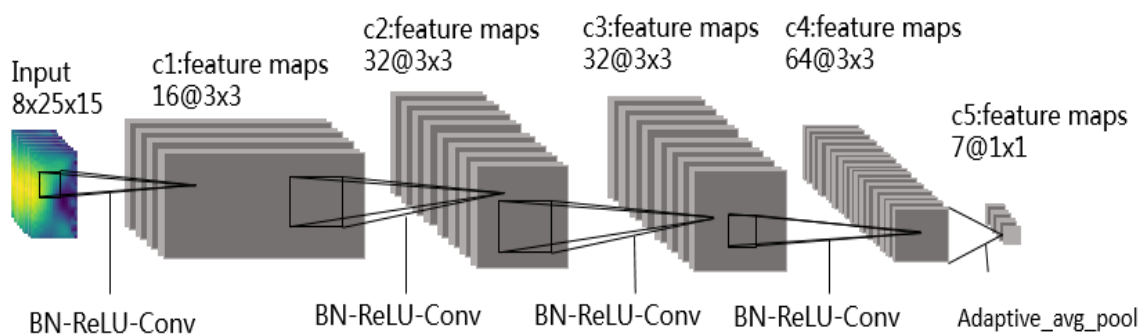


Figure 6. EMGNet Architecture contains four convolutional layers and a max pooling layer without using the full connection layer as the final output. In this figure, Conv refer to Convolution and avg_pool to a max pooling layer.

Table 1. Number of parameters used by various models.

Model	CNN_LSTM	LCNN	ConvNet	EMGNet
Number of parameters	56645	154353	67179	34311

The cost function Loss can be computed as follows:

$$Loss = -\sum_{i=1}^n y_i \log(y_i') \quad (5)$$

where y_i is the truth value of the i th category, n is the number of categories, and y_i' is the i th category predicted value of the output. Because we adopted One-Hot Encoding, the true value of the one category is 1 while the others are 0.

The advanced optimization method is used for the backpropagation of the EMGNet, and our final target is to minimize the cost function Loss. In the field of image recognition, the size of convolution kernels commonly used by researchers is 3×3 and 5×5 [11,26,27]. We found that the experimental results obtained by setting the size of the convolution kernel to 3×3 are better. Therefore, the size of all convolution kernels in the EMGNet model is set to 3×3 . Meanwhile, in order to reduce the parameters of the model as much as possible, the feature map of each layer of the EMGNet model is also set smaller. When the feature map of the network model increases, the step length in the convolution process is set to 2, so as to achieve the effect of halving. In other cases, the step length and the padding are set to 1, so that the feature remains unchanged. In order to further reduce the number of network parameters, the output of the model does not use the full connection layer, but carries out adaptive mean sampling first, and then uses the convolution layer for classification.

4. Experiment and Results

We evaluated our mode on two publicly available hand gesture recognition datasets composed of Myo Dataset [8,9] and NinaPro DB5 Dataset [6]. First, we compared the method proposed with the methods of classical machine learning and three other methods (CNN-LSTM, LCNN, CWT+TL) on the Myo Dataset [8,9]. Then, it was compared with the methods of classical machine learning and the methods (CNN-LSTM, LCNN) proposed earlier on the NinaPro DB5 Dataset [6].

4.1. Evaluation Dataset

Containing two different sub-datasets, this gesture dataset (Myo Dataset [8,9]) was collected using the Myo armband. The first Dataset was used as the pre-training dataset in [8] and the second includes the part of training and testing. The former is mainly used to establish, verify, and optimize the classification model, which consists of 19 subjects. The latter, used only for final training and validation, consists of 17 participants. The second Dataset contains a training section and two test sections in the literature [8,9], which is an unreasonable arrangement and leads to a decrease in the amount of training data. In order to facilitate the comparative experiment, this article uses the same settings. The Myo Dataset contains 7 types of gestures, and there are significant differences between gestures. It provides sufficient amount of data, and its gestures are shown in Figure 7.



Figure 7. The 7 hand/wrist gestures in the Myo Dataset. In the Myo dataset, seven gestures are included: Neutral, Hand Close, Wrist Extension, Ulnar Deviation, Hand Open, Wrist Flexion, and Radial Deviation.

This dataset (NinaPro DB5 [6]) is based on benchmark sEMG-based gesture recognition algorithms [6] containing data from 10 able-bodied participants divided into three exercise sets—Exercise A, B, and C contain 12, 17, and 23 different movements (including neutral) respectively. It is recorded by two Myo armbands, and only one of them is used in this work. One of the characteristics is that there are some similar gestures and the training data is not enough. Therefore, the model is prone to overfitting in the training process. Figure 8 shows the gesture categories in the dataset of Exercise A.













Exercise A								
1	Index flexion		5	Ring flexion		9	Thumb adduction	
2	Index extension		6	Ring extension		10	Thumb abduction	
3	Middle flexion		7	Little finger flexion		11	Thumb flexion	
4	Middle extension		8	Little finger extension		12	Thumb extension	

Figure 8. The gesture categories in the Exercise A dataset.

4.2. Method of Training

Adam [11] is used as the optimization method of network model training in this work. The length of data collected by each gesture in the two datasets was the same. After the same segmentation method, the data amount of each gesture was the same, and the samples are balanced. In the Myo dataset, each person has 2280 samples for each gesture, with a total of 19 participants, while in the Nina Pro dataset, each person has 1140 samples for each gesture, with a total of 10 participants.

At the same time, after the samples are segmented, we used the shuffle algorithm to shuffle the samples of each gesture, and then took 60% as a training sample set, 10% as the verification sample set, and the last 30% of each gesture as the test sample set.

The amount of data fed into the network by each training batch was 128 samples, and a total of 50 rounds of iterative training were conducted. We initially set the learning rate at 0.01 and shrank the learning rate by 10 times at epoch 20 and 40. In order to prevent over-fitting of the network, L2 regularization is used in this paper. The parameter settings in the training process are shown in Table 2.

Table 2. Training policy parameter setting (lr represents the learning rate).

Dataset	Initial lr	lr in 20 Epoch	lr in 40 Epoch	L2 Regularization
Myo Dataset	0.01	0.001	0.0001	1e-2
DB5	0.01	0.001	1e-3	1e-3

4.3. Myo Dataset Classification Results

The loss and accuracy curves during training and testing on the Myo Dataset using the EMGNet model are shown in Figure 9.

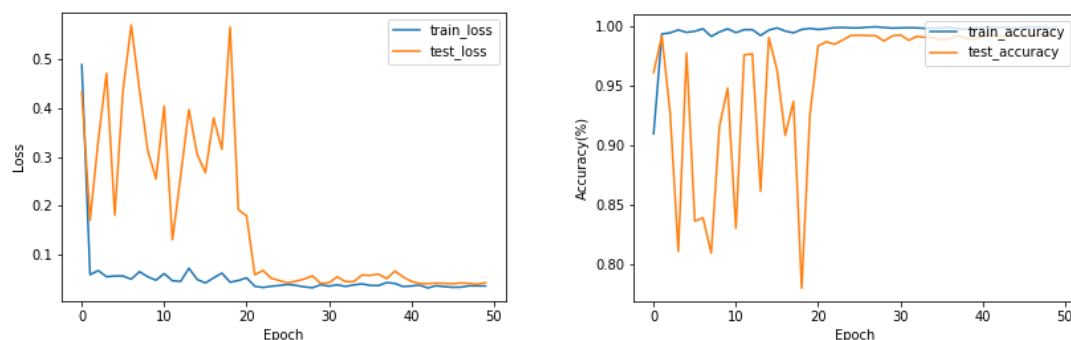


Figure 9. The loss and accuracy curves during training and testing on the Myo Dataset. when the training reaches convergence, there is no phenomenon where the accuracy of the training set is high and the accuracy of the test set is low, which is over-fitting.

As shown in Figure 9, the EMGNet network successfully completed the classification task in the Myo Dataset, and the over-fitting phenomenon did not appear in the training and testing. We tested the accuracy of our model and compared it with the current three most advanced methods. Table 3 shows the accuracy of each method. According to the results shown in Table 3, the accuracy of our proposed model is better than the current advanced methods.

Table 3. Classification accuracy on the Myo Dataset. (We used the same samples to test these methods).

Model	TD+LDA	TD+SVM	Enhanced TD+LDA	Enhanced TD+SVM	NinaPro+LDA	NinaPro+SVM	SampEn+LDA	SampEn+SVM	LCNN	CNN-LSTM	CWT+TL	CWT+EMGNet
Accuracy (%)	97.54	96.61	98.16	96.40	97.54	94.73	98.02	95.95	97.88	96.77	98.31	98.81
STD (%)	2.60	5.50	2.30	4.70	2.60	6.70	2.10	5.70	2.20	2.80	2.16	1.51

¹ The STD represents the standard deviation in accuracy for the 20 runs over the 17 participants.

4.4. NinaPro DB5 Dataset Classification Results

Figure 10 shows that the loss and accuracy curves during training and testing on exercise A of the NinaPro DB5 Dataset. During training of the EMGNet model, the phenomenon of overfitting appears. Reducing the number of layers in the EMGNet does not solve this problem.

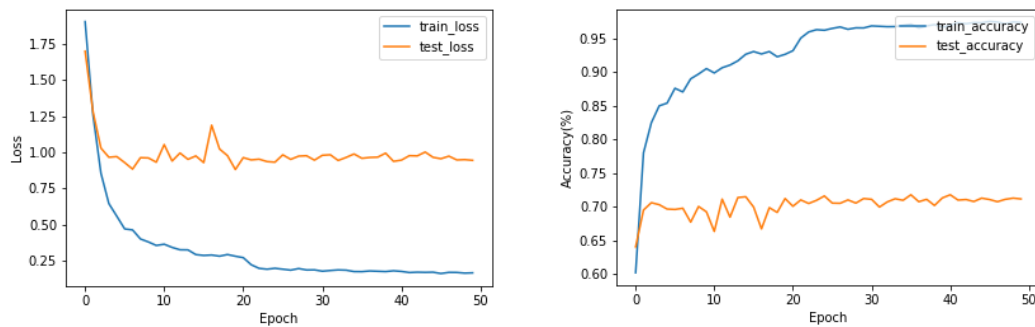


Figure 10. The loss and accuracy curves during training and testing on the exercise A of the NinaPro DB5 Dataset. During training and testing on the exercise A of the NinaPro DB5 Dataset, the phenomenon of overfitting appears.

Table 4 shows the accuracy of three subsets of DB5 Dataset. Time domain characteristics (TD) [5], Enhanced TD [4], Nina Pro Features [6,22] and SampEn Pipeline [23] were selected as the classification features of the classical machine learning algorithm (LDA and SVM). The LCNN and CNN_LSTM we proposed previously did not perform feature extraction on the data and directly processed the sEMG signal. The proposed method uses continuous wavelet transform (CWT) to process the data as the input of EMGNet. From the experimental results, we can conclude that the classification accuracy of our proposed EMGNet model is higher than that of the classical machine learning algorithm.

With the increase of the categories of gestures, the classification algorithms decline at different degrees. The degree of decline of EMGNet is lower than that of the classical machine learning algorithms (see Figure 11 and Table 4).

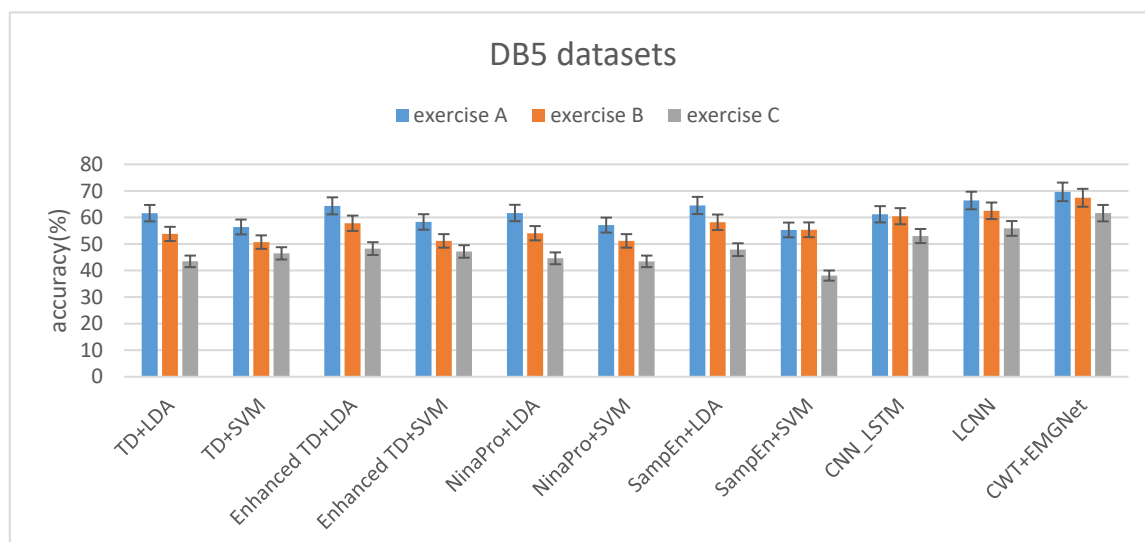


Figure 11. The average accuracy of three subsets on the DB5 Dataset.

Table 4. Classification accuracy on the DB5 Dataset. (We use the same samples to test these methods).

Model	TD+LDA	TD+SVM	Enhanced TD+LDA	Enhanced TD+SVM	NinaPro+LDA	NinaPro+SVM	SampEn+LDA	SampEn+SVM	CNN_LSTM	LCNN	CWT+EMGNet
EA(%)	61.61	56.40	64.36	58.32	61.66	57.12	64.53	55.28	61.21	66.38	69.62
EB(%)	53.80	50.70	57.80	51.17	54.05	51.18	58.20	55.34	60.46	62.51	67.42
EC(%)	43.48	46.49	48.26	47.18	44.61	43.46	47.89	38.12	52.99	55.87	61.63
$\frac{EA - EC}{EA}$	29.4%	17.5%	25.0%	19.1%	29.5%	23.9%	25.8%	31.0%	13.4%	15.8%	11.5%

¹ EA, EB and EC represent the accuracy of exercise A, exercise B and exercise C respectively.

The classification accuracy of both the classical machine learning method and EMGNet is not as high as that tested on the Myo Dataset (see Table 4). Two reasons are as follows:

- (1) The DB5 dataset has a relatively small amount of data per gesture and a relatively large number of gesture categories. For example, there are at most 7 gestures in the Myo dataset, while the smallest exercise A in DB5 dataset has 12 gestures.
- (2) There are a large number of similar gestures in the DB5 dataset. Figure 12 shows the sEMG signal spectrum of two processed samples in exercise A. The two samples shown in Figure 12c,d belong to different gesture categories, and it can be seen from Figure 12 that the channels with large fluctuations of the two gestures are almost the same. When the sample number of each gesture is insufficient, the model can easily misidentify it as the same gesture.

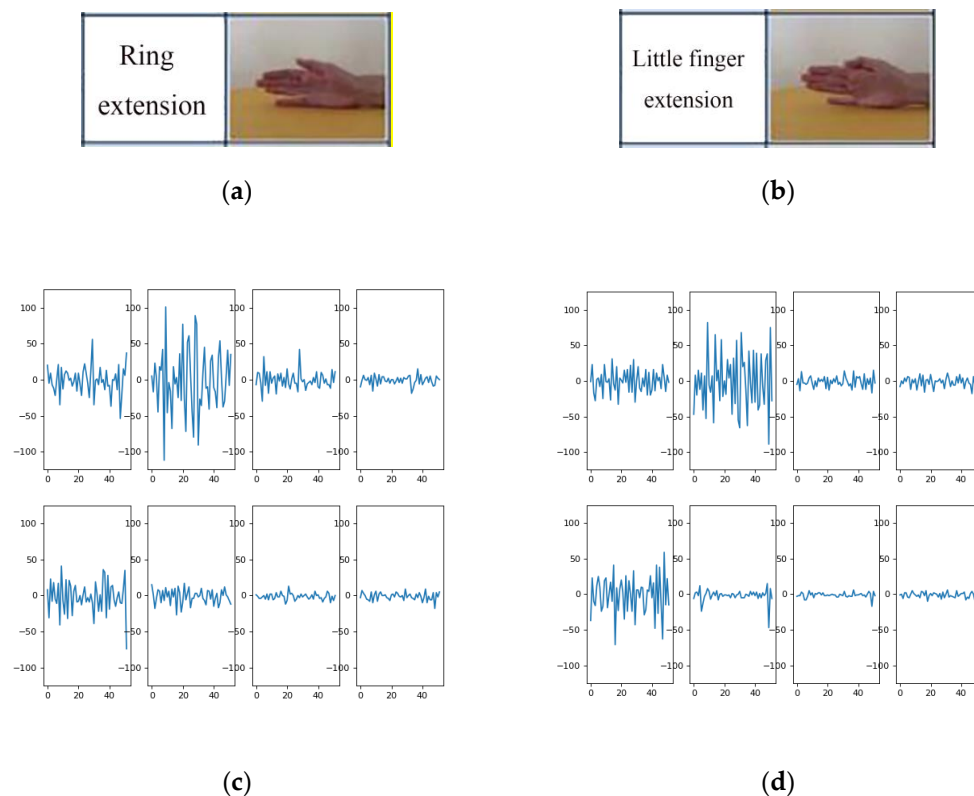


Figure 12. Parts (a) and (b) represent pictures of two different gestures in exercise A, Parts (c) and (d) show the spectrum diagrams of the sEMG signal generated by (a) and (b) respectively.

5. Conclusions

This paper presents a novel CNN architecture consisting of four convolutional layers and a max pooling layer with a compact structure and fewer parameters. The experimental results show that the proposed EMGNet not only reduces the complexity of the model but also improves the accuracy of the sEMG signal classification. It is highly competitive with the classical classifiers and deep learning frameworks currently used to classify sEMG signals, which also shows the potential of deep learning in classifying sEMG signals. (The elimination of transition time between successive gestures is explained in Appendix A).

Author Contributions: Writing—original draft, L.C.; Investigation, L.C. and Y.W.; Conceptualization and methodology, J.F.; Software, Y.W. and H.L.; Visualization, H.L.; Experiment design, Y.W.; Data curation, Y.W.; Writing—review and editing, B.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Priority research & design program of Chongqing technology innovation and application demonstration, Grant No. cstc2017zdcy-zdyfx0036.

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

Abbreviations

The following abbreviations are used in this manuscript:

EMG	Electromyography
SEMG	surface Electromyography
Adam	adaptive moment estimation
CNN	convolutional neural network
TD	time domain
MAV	mean absolute value
SSC	slop sign changes
ZC	zero crossing
WL	waveform length
WT	wavelet transform
RMS	root mean square
AR	autoregressive coefficient
STFT	short time Fourier transform
CWT	continuous wavelet transform
LSTM	Long Short-Term Memory

Appendix A

We assume that each gesture can be held for a certain time. As shown in the figure below, if three sliding windows among four consecutive sliding windows are judged as gesture A, then we judge this signal as gesture A, if not, we judge this signal as the previous gesture remains unchanged. We use the same method for other gestures. (The numbers 4 and 3 used above are only examples, not actual numbers.) Our proposed method can identify the signals generated by different gestures during the signal acquisition process.

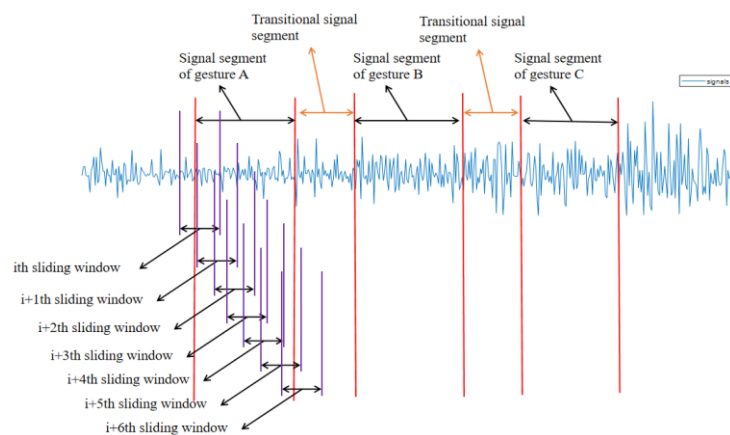


Figure A1. Schematic of continuous signal recognition.

References

- Oskoei, M.A.; Hu, H. Myoelectric control systems—A survey. *Biomed. Signal Process. Control* **2007**, *2*, 275–294. [[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 ECTI-CON2010: The 2010 ECTI International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, Chiang Mai, Thailand, 19–21 May 2010; pp. 856–860.

3. Phinyomark, A.; Phukpattaranont, P.; Limsakul, C. Feature reduction and selection for EMG signal classification. *Expert Syst. Appl.* **2012**, *39*, 7420–7431. [[CrossRef](#)]
4. Khushaba, R.N.; Kodagoda, S. Electromyogram (EMG) Feature Reduction Using Mutual Components Analysis for Multifunction Prosthetic Fingers Control. In Proceedings of the 2012 12th International Conference on Control Automation Robotics Vision (ICARCV), Guangzhou, China, 5–7 December 2012; pp. 1534–1539.
5. Englehart, K.; Hudgins, B. A robust, real-time control scheme for multifunction myoelectric control. *IEEE Trans. Biomed. Eng.* **2003**, *50*, 848–854. [[CrossRef](#)] [[PubMed](#)]
6. Pizzolato, S.; Tagliapietra, L.; Cognolato, M.; Reggiani, M.; Müller, H.; Atzori, M. Comparison of six electromyography acquisition setups on hand movement classification tasks. *PLoS ONE* **2017**, *12*, e0186132. [[CrossRef](#)] [[PubMed](#)]
7. Wu, Y.; Zheng, B.; Zhao, Y. Dynamic Gesture Recognition Based on LSTM-CNN. In Proceedings of the Chinese Automation Congress (CAC), Xi'an, China, 30 November—2 December 2018; pp. 2446–2450.
8. Côté-Allard, U.; Fall, C.L.; Drouin, A.; Campeau-Lecours, A.; Gosselin, C.; Glette, K.; Laviolette, F.; Gosselin, B. Deep Learning for Electromyographic Hand Gesture Signal Classification Using Transfer Learning. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2019**, *27*, 760–771. [[CrossRef](#)] [[PubMed](#)]
9. Cote-Allard, U.; Fall, C.L.; Campeau-Lecours, A.; Gosselin, C.; Laviolette, F.; Gosselin, B. Transfer learning for sEMG hand gestures recognition using convolutional neural networks. In Proceedings of the 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Banff, AB, Canada, 5–8 October 2017; pp. 1663–1668.
10. Castellini, C.; Fiorilla, A.E.; Sandini, G. Multi-subject/dailylife activity EMG-based control of mechanical hands. *Neuroeng. Rehabil.* **2009**, *6*, 41. [[CrossRef](#)] [[PubMed](#)]
11. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In *Proceedings of the Advances in Neural Information Processing Systems 25 (NIPS 2012)*, Lake Tahoe, NV, USA, 3–6 December 2012; Neural Information Processing Systems Foundation Inc.: La Jolla, CA, USA, 2012.
12. Zia ur Rehman, M.; Waris, A.; Gilani, S.O.; Jochumsen, M.; Niazi, I.K.; Jamil, M.; Farina, D.; Kamavuako, E.N. Multiday EMG-Based Classification of Hand Motions with Deep Learning Techniques. *Sensors* **2018**, *18*, 2497. [[CrossRef](#)] [[PubMed](#)]
13. Allard, U.C.; Nougrou, F.; Fall, C.L.; Giguere, P.; Gosselin, C.; LaViolette, F.; Gosselin, B. A convolutional neural network for robotic arm guidance using sEMG based frequency-features. In Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Daejeon, Korea, 9–14 October 2016; pp. 2464–2470.
14. Tommasi, T.; Orabona, F.; Castellini, C.; Caputo, B. Improving Control of Dexterous Hand Prostheses Using Adaptive Learning. *IEEE Trans. Robot.* **2012**, *29*, 207–219. [[CrossRef](#)]
15. Patricia, N.; Tommasi, T.; Caputo, B. Multi-source Adaptive Learning for Fast Control of Prosthetics Hand. In Proceedings of the 2014 22nd International Conference on Pattern Recognition, Stockholm, Sweden, 24–28 August 2014; pp. 2769–2774.
16. Orabona, F.; Castellini, C.; Caputo, B.; Fiorilla, A.E.; Sandini, G. Model adaptation with least-squares SVM for adaptive hand prosthetics. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), Kobe, Japan, 12–17 May 2009; pp. 2897–2903.
17. Atzori, M.; Cognolato, M.; Müller, H. Deep learning with convolutional neural networks applied to electromyography data: A resource for the classification of movements for prosthetic hands. *Front. Neurobotics* **2016**, *10*, 9. [[CrossRef](#)] [[PubMed](#)]
18. Geng, W.; Du, Y.; Jin, W.; Wei, W.; Hu, Y.; Li, J. Gesture recognition by instantaneous surface EMG images. *Sci. Rep.* **2016**, *6*, 36571. [[CrossRef](#)] [[PubMed](#)]
19. Du, Y.; Jin, W.; Wei, W.; Hu, Y.; Geng, W. Surface EMG-Based Inter-Session Gesture Recognition Enhanced by Deep Domain Adaptation. *Sensors* **2017**, *17*, 458. [[CrossRef](#)] [[PubMed](#)]
20. He, K.; Zhang, X.; Ren, S.; Sun, J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–3 December 2015; pp. 1026–1034.
21. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
22. Atzori, M.; Gijsberts, A.; Castellini, C. Electromyography data for non-invasive naturally-controlled robotic hand prostheses. *Sci. Data* **2014**, *1*, 1–3. [[CrossRef](#)] [[PubMed](#)]

23. Phinyomark, A.; Quaine, F.; Charbonnier, S.; Serviere, C.; Tarpin-Bernard, F.; Laurillau, Y. EMG feature evaluation for improving myoelectric pattern recognition robustness. *Expert Syst. Appl.* **2013**, *40*, 4832–4840. [[CrossRef](#)]
24. Zhang, X.; Zhou, P. Sample entropy analysis of surface EMG for improved muscle activity onset detection against spurious background spikes. *J. Electromyogr. Kinesiol.* **2012**, *22*, 901–907. [[CrossRef](#)] [[PubMed](#)]
25. 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](#)]
26. LeCun, Y.; Haffner, P.; Bottou, L.; Bengio, Y. Object Recognition with Gradient-Based Learning. In *Shape, Contour and Grouping in Computer Vision*; Forsyth, D.A., Mundy, J.L., di Gesù, V., Cipolla, R., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 1999; pp. 319–345.
27. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).