

# A Hybrid Model Based on ResNet and GCN for sEMG-Based Gesture Recognition

Xianjing Xu, Haiyan Jiang<sup>✉</sup>

**Abstract:** The surface electromyography (sEMG) is one of the basic processing techniques to the gesture recognition because of its inherent advantages of easy collection and non-invasion. However, limited by feature extraction and classifier selection, the adaptability and accuracy of the conventional machine learning still need to promote with the increase of the input dimension and the number of output classifications. Moreover, due to the different characteristics of sEMG data and image data, the conventional convolutional neural network (CNN) have yet to fit sEMG signals. In this paper, a novel hybrid model combining CNN with the graph convolutional network (GCN) was constructed to improve the performance of the gesture recognition. Based on the characteristics of sEMG signal, GCN was introduced into the model through a joint voting network to extract the muscle synergy feature of the sEMG signal. Such strategy optimizes the structure and convolution kernel parameters of the residual network (ResNet) with the classification accuracy on the NinaPro DB1 up to 90.07%. The experimental results and comparisons confirm the superiority of the proposed hybrid model for gesture recognition from the sEMG signals.

**Keywords:** deep learning; graph convolutional network (GCN); gesture recognition; residual network (ResNet); surface electromyographic (sEMG) signals

## 1 Introduction

The surface electromyographic (sEMG) signals are a kind of biological electrical signals recorded on the skin surface, containing movement intentions of human body. The sEMG signals have inherent advantages in the human-machine interaction (HMI) like accessibility and non-invasion. Therefore, they can achieve the aim of HMI according to human autonomous consciousness.

A sEMG-based approach which is also known as the muscle-computer interface (MCI),

is widely used in prosthetic control [1], robot control [2], sign language recognition [3], and HMI [4]. For MCI, one key point is to precisely recognize gestures including the analyses of different categories such as hand posture, object grasping, and hand movements [5]. Such sEMG-based gesture recognition method has been widely used in many fields like the computer games, the assisted exoskeleton robots [6], and the medical rehabilitation robot [7], etc.

The recent researches about the sEMG-based gesture recognition mainly focus on the machine learning and the deep learning. The machine learning approach [8] classifies gestures by extracting sEMG features and constructing classifiers, such as the support vector machine (SVM), the naive Bayes classifiers, the random forest, and the neural network. Tavakoli et al. used high-dimensional spatial features and SVM to classify five gestures with the final classification accuracy of 95%-100% [9]. Jiang et al. com-

---

Manuscript received Oct. 19, 2022; revised Feb. 23, 2023; accepted Mar. 14, 2023. The associate editor coordinating the review of this manuscript was Dr. Hanshu Cai. This work was supported by the Development of Sleep Disordered Breathing Detection and Auxiliary Regulation System Project (No. 2019I1009).

Xianjing Xu and Haiyan Jiang are with Key Laboratory of Medical Instrumentation & Pharmaceutical Technology, College of Electrical and Automation, Fuzhou University, Fuzhou 350108, China.

✉ Corresponding author. Email: [jianghaiyan@fzu.edu.cn](mailto:jianghaiyan@fzu.edu.cn)  
DOI: [10.15918/j.jbit1004-0579.2022.116](https://doi.org/10.15918/j.jbit1004-0579.2022.116)

bined electromyography (EMG) and force myography (FMG) for the gesture recognition of American Sign Language digits with a classification accuracy of 91.6% [10]. Waris et al. extracted the sEMG signal features to classify gestures by using artificial neural network [11]. Above works prove that the artificial neural network shows the better performance than the classical K-nearest neighbor (KNN) and SVM.

A key limitation of machine learning methods is the feature extraction which is a cumbersome process. Therefore, it is a crucial issue to select an optimal classifier for gesture classification according to the selected feature. A number of features are most commonly used in sEMG-based gesture recognition, including time-domain features, frequency-domain features, and time-frequency features [7,12,13]. However, these features are susceptible to the muscle fatigue and the electrode displacement, which reduces the robustness and accuracy of the gesture recognition. Therefore, it is expected to get the optimal features from the principle of neural control movement to improve the robustness and accuracy [14]. The principle indicates that the muscle synergy, the smallest controlled unit of the central nervous system [15], presents the enhanced robustness under the influence of different factors. Luo et al. applied the muscle synergy method to recognize five gestures with 96% classification accuracy [16]. Zhang et al. designed a multi-degree-of-freedom parallel control framework based on the muscle synergy and achieved  $96.79\% \pm 2.46\%$  classification accuracy in 18 finger movements [17]. Masoumdoost et al. extracted the muscle synergy feature of six wrist movements and obtained  $99.78\% \pm 0.45\%$  classification accuracy in multilayer perceptron (MLP) [18]. These studies show that muscle synergy features have a promising application in sEMG-based gesture recognition. However, the convolutional neural network (CNN) is difficult to extract this feature because CNN is designed for Euclidean data, while muscle synergy feature is a graph data.

Recently, some literature reported that the graph convolutional network (GCN) has great potential in the field of action recognition [19]. The work in [20] proposed a recurrent neural network method with global context-aware capability by adding an attention mechanism (AT) to the long short-term memory (LSTM) network, reaching 77.1% classification accuracy on the benchmark dataset NTU-RGBD; Shi et al. designed a topology network, two-stream adaptive graph convolutional neural network (2S-AGCn), that can adaptively learn different graph convolution layers and skeleton data, significantly improving the accuracy of recognition [21]. Therefore, it would be a promising tactic to apply the GCN to improve the performance of the gestures classification.

In the field of sEMG-based gesture recognition, it is demonstrated that the method based on the deep learning has the more advantages than the machine learning in improving classification accuracy, universality, and generalizability [22,23]. For instance, Atzori et al. [24] used a four-layer simple convolution network structure to classify 52 gestures on the NinaPro DBI [25] with a 2%-5% improvement in the classification accuracy compared to the classical KNN and SVM. Chen et al. [26] proposed a model based on long short-term memory and CNN to recognize gestures, which achieved a classification accuracy of 75.12% on the NinaPro DBI. Wei et al. [27] split each segments of the data samples into patches of images which are processed with a parallel multi-stream CNN architecture. The classification accuracy reached 88.2% on the NinaPro DBI, but the method is complex.

Here, according to the characteristics of the sEMG signal and the principle of muscle synergy, a novel hybrid model is proposed by combining residual network (ResNet) and GCN. The main contributions of this work are as follows.

- 1) A model combining ResNet and GCN is successfully used for the sEMG-based gesture recognition.

2) The proposed method obtains the augmented features by combining ResNet and GCN. Specifically, ResNet is used to extract the sEMG features, GCN is used to extract the muscle synergy features, and finally the sEMG features and the muscle synergy features are combined to obtain the augmented features.

3) We constructed 3D sEMG images by analyzing the intrinsic mode functions (IMF) of the sEMG signals from the variational mode decomposition (VMD) decomposition. Meanwhile, we evaluated the proposed method on the NinaPro DB1 and compared its performance with existing deep learning methods.

The organization of the paper is as follows: Section 2 presents the proposed methodology. Section 3 describes the dataset and data preprocessing. Section 4 is the experimental results and

discussion. Section 5 summarizes our work and discusses future work.

## 2 Methodology

### 2.1 The Hybrid Model Based on ResNet and GCN

The overall architecture of the hybrid model based on ResNet and GCN is shown in Fig. 1. The multichannel sEMG signals are reconstructed into  $L \times M \times C$  size sEMG images. Firstly, ResNet and GCN are used to extract sEMG signals feature  $\mathbf{X}$  and muscle synergy feature  $\mathbf{F}_q$ , respectively.  $\mathbf{X}$  and  $\mathbf{F}_q$  are then concatenated to the augmented feature  $\mathbf{X}_e$ , which is then fed into the global average pooling (GAP). Finally, the NN is used for gesture recognition to get the final classification result.

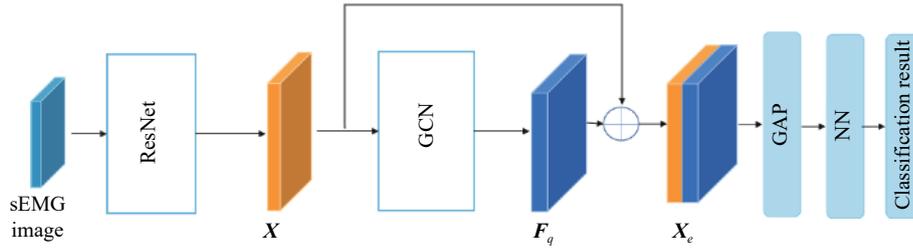


Fig. 1 Architecture of the model based on ResNet and GCN

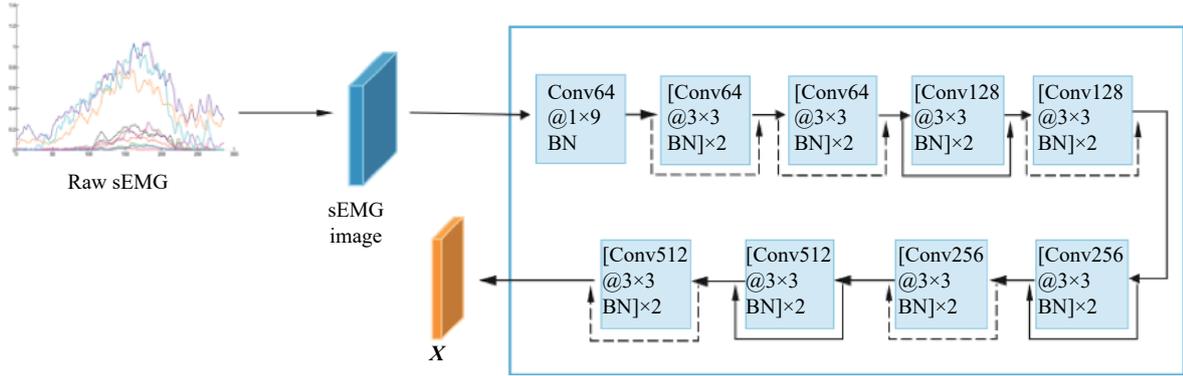
The NN in Fig. 1 has 4 hidden layers, which are 256, 256, 128, and 128 neurons respectively. In each hidden layer, the dropout layer is used to prevent overfitting during network training, and the loss rate of the dropout layer is set to 0.2. In the output layer, the softmax function is used as the activation function, and the L2 regularization parameter is set to 0.001. Furthermore, gradient descent optimization is performed using stochastic gradient descent (SGD) optimizer, where the learning rate and decay values of the optimizer are set to  $10^{-3}$  and  $10^{-4}$ , and epoch and batch size are set to 100 and 32, respectively.

### 2.2 The ResNet Architecture

Conventional ResNet solves the problem of model degradation caused by CNN deepening, which is suitable for image classification. However, since the sEMG data of each channel corresponds to

the time series signal generated by a single muscle, the classification accuracy of conventional ResNet is not high when it is used for gesture recognition. In order to solve this problem and make ResNet better learn the features of sEMG signals, we adjust the convolution kernel size of each module of ResNet to solve the degradation problem of deep CNN model while preserving the ResNet architecture.

As shown in Fig. 2, the ResNet architecture is composed of 1 initial convolution layer, 8 residual blocks, and 1 full connection layer. The initial convolution kernel is adjusted to  $1 \times 9$ , the step size is set to 1, and there is no maximum pooling layer to avoid information loss. The convolution kernel size in each residual block is set to  $3 \times 3$ .  $\mathbf{X} \in \mathbf{R}^{H \times W \times V}$  obtained from the improved ResNet.



Note: Conv and BN denote the convolution layer and batch normalization, respectively. The number following the layer name denotes the number of filters, and the numbers after the ampersand (@) denote the convolution kernel size.

Fig. 2 Architecture and parameters of the improved ResNet

### 2.3 The GCN

The CNN achieved impressive performance in a wide variety of fields. Despite the merit, CNN fails to properly address problems with non-Euclidean data. However, the muscle synergy is a unique kind of graph relationship. The most basic characteristic of muscle synergy is that each muscle has its feature, and the synergy between each two muscles is different and uncertain. To overcome this problem, this paper introduced the GCN by the voting network to automatically extract muscle synergy features. GCN operates directly on non-Euclidean data can learn

node feature and structure feature information end-to-end with strong applicability.

The GCN in this paper is shown in Fig. 3, the network starting unit is a joint voting network. The feature matrix of each muscle is obtained by voting the features of sEMG signal. The voting weight matrix function is expressed as

$$\mathbf{W} = \Phi(\psi(\mathbf{X})) \quad (1)$$

where  $\psi(\cdot)$  is a transform function implemented by a  $1 \times 1$  convolution,  $\Phi$  is the spatial softmax normalization, and  $\mathbf{W} \in \mathbf{R}^{H \times L \times N}$  is the voting weight, where the  $k$ th channel  $\mathbf{W}_k \in \mathbf{R}^{H \times W}$  represents the voting matrix of the  $k$ th muscle.

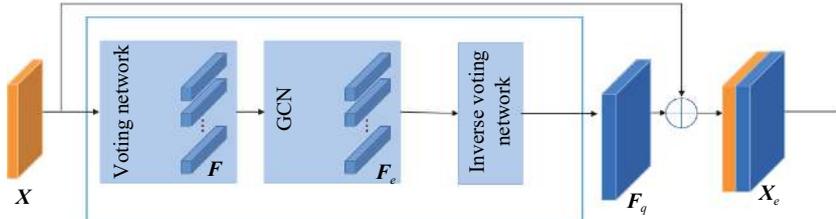


Fig. 3 Structures of the proposed GCN

The  $k$ -th muscle feature, denoted as  $f_k$ , is calculated as follows

$$f_k = \sum_i \mathbf{W}_{ki} \phi(\mathbf{X}_i) \quad (k = 1, 2, \dots, N) \quad (2)$$

where  $\mathbf{X}_i$  refers to the sEMG feature of the  $i$ th channel,  $\phi(\cdot)$  refers to the transform function implemented by a  $1 \times 1$  convolution layer, and  $\mathbf{W}_{ki}$ , an element of  $\mathbf{W}_k$ , is the voting weight matrix for  $\mathbf{X}_i$ . Thus, we have the feature  $\mathbf{F}$  of all  $N$  muscles as below:

$$\mathbf{F} = [f_1^T \cdots f_N^T] \quad (3)$$

To intelligently extract the muscle synergy features, the connection graph  $G$  between muscles is assumed as fully connected. Then,  $\mathbf{A} \in \mathbf{R}^{N \times N}$  is defined as a matrix representing  $G$  with elements 1 except the diagonal element 0,  $\mathbf{I}$  is the identity matrix, and  $\mathbf{D}$  is the degree matrix of  $\mathbf{A}$ . Following GCN defined in [28], we perform graph reasoning over feature  $\mathbf{F}$  with matrix multiplication, resulting the evolved mus-

cle synergy features  $F_e$

$$F_e = \sigma \left( D^{-\frac{1}{2}} A_e D^{-\frac{1}{2}} F W_e \right) \quad (4)$$

where  $A_e = A + I$ ,  $W_e$  is a trainable transformation matrix, and  $\sigma(\cdot)$  is a transform function implemented by a  $1 \times 1$  convolution with BN and ReLU. The evolved muscle synergy features can be used to augment the sEMG features. Specifically, the evolved muscle synergy features are mapped back to the muscle features and then combined with  $X$  to calculate the augmented features.

$$C_{ik} = W_{ik} F_e^k \quad (5)$$

$$F_q = \rho \left( \frac{1}{N} \sum_k C_{ik} \right) \quad (6)$$

where  $F_e^k$  is the muscle synergy feature evolved from the  $k$ th muscle,  $W_{ik}$  is the same as the voting weight  $W_{ik}$ ,  $\rho(\cdot)$  is a transform function implemented by a  $1 \times 1$  convolution with BN and ReLU, and then the final muscle synergy feature  $F_q$  is calculated using the mean value of  $C_{ik}$ . Finally, we concatenate  $X$  and  $F_q$  to obtain the augmented feature  $X_e$

$$X_e = \tau ([X, F_q]) \quad (7)$$

where  $\tau(\cdot)$  is a transform function implemented by a  $1 \times 1$  convolution with BN and ReLU.

### 3 Data

#### 3.1 Dataset Description

NinaPro is a publicly available database consisting of ten data sets that provide benchmark electromyography data sources of the upper limbs to test the accuracy of machine learning algorithms. Among these databases, NinaPro DB1 consists of sEMG signals extracted from 27 subjects (20 males and 7 females, aged  $28 \pm 3.4$ -year-old) while they performed 52 different hand activities (tag No.0-52), as shown in Tab. 1, there are 53 labels and corresponding gesture categories. Fig. 4 shows 52 hand movements in the NinaPro DB1 dataset, the movements are grouped into three

exercises, namely exercise A, exercise B, and exercise C, respectively. The exercise A involves basic movements of the fingers, such as flexions and extensions. The exercise B consists of multiple finger flexion and extension movements as well as wrist movements. The exercise C consists of grasping household objects. Each activity is performed for a duration of 5 s with a rest period of 3 s and repeated over 10 trials and the sEMG sampling rate is 100 Hz.

Tab. 1 The label of gesture activity on the NinaPro DB1

Label	Gesture activity classification
0	Tranquillization
1 – 12	Basic movements of the fingers
13 – 29	Isometric, isotonic hand configurations and basic wrist movements
30 – 52	Grasping and functional movements



Fig. 4 The 52 hand movements in the NinaPro DB1: (a) 12 basic movements of the fingers; (b) 8 isometric and isotonic hand configurations and 9 basic movements of the wrist; (c) 23 grasping and functional movements

#### 3.2 Data Processing

When dealing with real life conditions, the accuracy of sEMG-based gesture recognition is often affected by undesirable noises and disturbances, such as poor contact, electrode displacement, and the subject’s skin condition. Therefore, several signal processing steps were performed before data analysis and classification. The DB1 dataset has been preliminarily preprocessed toward sEMG, including signal full-wave rectification, synchronization, and relabeling. Therefore, only the high-frequency noises of the acquisition

device need to be considered. In this work, the first-order Butterworth filter as in previous studies [29–31] was used to carry out low-pass filtering on the raw sEMG of each channel to remove the high-frequency noise. Fig. 5 shows the comparison between the denoised signals and the raw signals.

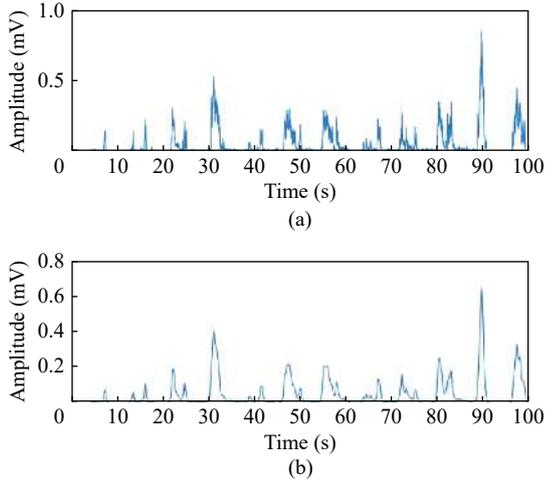
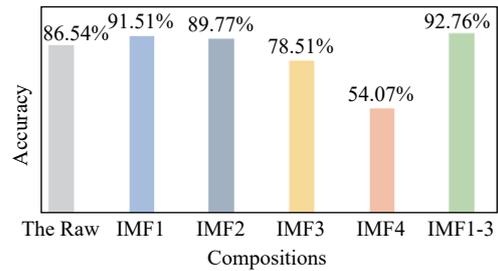


Fig. 5 Comparison of denoised and the raw signals: (a) the raw signals; (b) the denoised signals

In order to construct more recognizable sEMG images, the VMD method is used to decompose the raw signal into multiple IMF. Each IMF signal has its center frequency, and each mode is smoothed after demodulation. Finally, IMF components with narrow bands were obtained according to the frequency domain characteristics of the actual signal. The raw sEMG signal and the IMF1-IMF4 decomposed by the VMD method were respectively used for training, and the results are shown in Fig. 6. The accuracy of recognition of the raw sEMG image was 86.54%, and the accuracy of IMF1-IMF4 reached 91.51%, 89.77%, 78.51%, and 54.07%, respectively. The greater contribution of low-frequency domain to the result may be related to the greater correlation of amplitude and other time-domain features with gesture.

After the decomposition using the VMD method, several IMF components were chosen to reconstruct the signals. It can be seen that the accuracy of 3D sEMG images construction using

the IMF1-IMF3 is further improved to 92.76%. Therefore, we employed a 5-level VMD decomposition for each channel of the raw sEMG signal and constructed a 3D sEMG images. Fig. 7 shows the image construction of a one-dimensional sEMG signal, the processed signal is segmented into continuous stream of windows with a sliding length of 200 ms and an increment of 170 ms. Data segmentation can not only multiply the number of the samples, but also improve the real-time performance of the model [32], the number of samples per subject can reach more than 10000 after processing. Since the generation of sEMG signal is 30–150 ms earlier than human muscle movements, the setting of 200 ms sliding window can meet the real-time requirements of HMI technology.



Note: Gesture recognition accuracy of the sEMG image constructed by using different signal compositions

Fig. 6 The accuracy of the different IMF

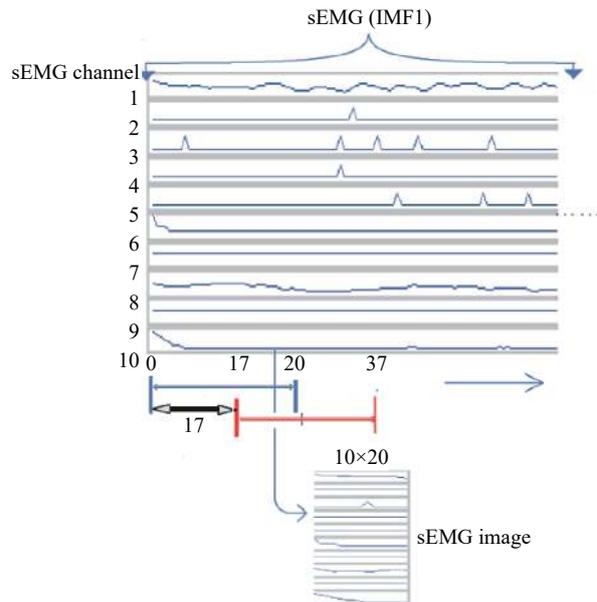


Fig. 7 Sliding windows of the sEMG data

## 4 Results and Discussion

### 4.1 Results

For each gesture, we followed the same validation scheme described in previous studies [24,29,30,33]. Specifically, the test set is composed of approximately 1/3 of the movement repetitions, and the training set is composed of the remaining repetitions. According to the above method, the final classification result is obtained through the cross-entropy loss function. The formula is as follows

$$L = -\frac{1}{n} \sum_{i=1}^n y_i \log y_i^* \quad (8)$$

where  $y_i^*$  is the label of the training sample,  $y_i$  is the output value of the model, and  $n$  is the number of samples.

The proposed approach is used to train and test on the DB1 dataset, and the classification accuracy is evaluated as follows

$$\text{Classification accuracy} = \frac{\text{Number of correct classifications}}{\text{Total number of test samples}} \times 100\% \quad (9)$$

Fig. 8 shows the accuracy and loss curves of the model in the training and test sets during the training of one subject of NinaPro DB1. It can be seen that the model can converge to the optimal value relatively quickly. In addition, the mixture matrix shows that the model has good gesture recognition performance.

Fig. 9 shows the recall rate and precision rate of the gesture recognition on one subject's test set. Most of the gestures achieved high recall rate and precision rate except for the gesture labeled 49, which, according to the mixing matrix in Fig. 8, were similar in power to the gestures labeled 50, 44, 39, and 34. The overall average recall rate and precision rate reached 93.01% and 92.84%, respectively. It can be seen that the proposed model has strong generalization ability and recognition ability.

Fig. 10 shows the evaluation results of 52

gestures performed by 27 subjects, with an average accuracy of 90.07%. The results show that the proposed model architecture is completely feasible.

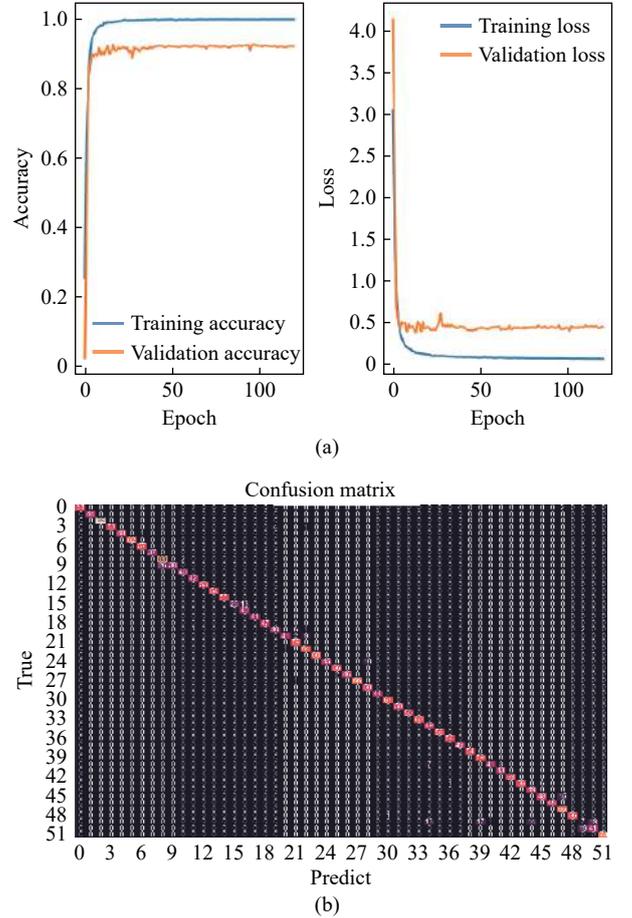


Fig. 8 Performance of the model on one subject's data: (a) the training curve of one subject; (b) the mixture matrix of one subject's test set

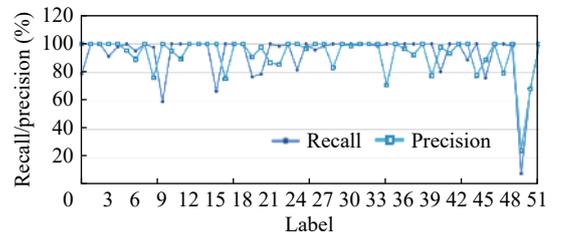


Fig. 9 The recall rate and precision rate of various gestures of one subject's test set

As shown in Tab. 2, the proposed model is also compared with existing networks, such as CNN, LS-SVM, CNN-RNN, Multi-view CNN, and EvCNN on NinaPro DB1 dataset. The sliding window size of each model in Tab. 2 is set to

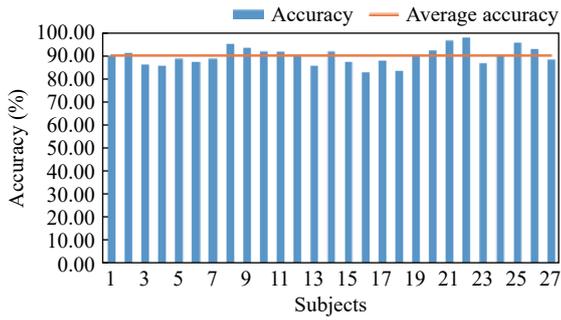


Fig. 10 Classification accuracy of each subject on the NinaPro DB1

200 ms. Compared with ResNet, the accuracy of the proposed model is improved by 7.82%, which proves the effectiveness of the method. Compared with RCNN, multi-stream ResNet, and multi-stream convolutional, the gesture classification accuracy of the proposed model is improved by 1.2%, 0.42%, and 0.37%, respectively. The experimental results show that the proposed model has certain advantages in sEMG gesture recognition, and achieves higher accuracy than the existing methods.

Tab. 2 Compared with the classification accuracy (%) of different types of networks

Classification accuracy of different types of networks (%)	NinaPro DB1
CNN by Atzori et al. [24]	66.60%
Benchmark classifier with handcrafted features by Atzori et al. [29]	75.32%
LDA by Nazemi et al. [34]	84.23%
LS-SVM by Nazemi et al. [34]	85.19%
Energy features by Karnam NK et al. [35]	88.80%
MsCNN by Wei et al. [30]	85.00%
EvCNN by Olsson AE et al. [36]	81.40%
CNN-RNN by Hu et al. [33]	87.00%
Multi-view CNN by Wei et al. [27]	88.20%
Multi-sEMG-features image by Cheng et al. [31]	82.54%
Multi-stream ResNet by Yang et al. [37]	<b>89.65%</b>
Multi-stream convolutional by Wang et al. [38]	<b>89.70%</b>
RCNN by Xu et al. [39]	<b>88.87%</b>
Dual-view by Zhang et al. [40]	86.72%
ResNet by Cao et al. [41]	82.25%
The proposed ResNet-GCN model	<b>90.07%</b>

## 4.2 Discussion

The performance of the proposed hybrid model based on ResNet and GCN is improved due to

the following reasons:

1) The sEMG signal is a strong nonlinear and non-stationary time series signal, and VMD can well process this kind of signal. The time-domain characteristics such as the amplitude of the IMFs in the low-frequency domain are strongly correlated with hand gestures. Therefore, the 3D sEMG images constructed by decomposing IMFs with VMD can well deal with the problem of gesture classification.

2) The GCN in the proposed model is used to extract the optimal muscle synergy features, which further improves the accuracy of gesture recognition. And the architecture of the ResNet also contributes to the performance of the model.

## 5 Conclusion

This paper proposed a novel approach based on the ResNet and the GCN to achieve better gesture recognition performance of sEMG-based HCI. Based on the characteristics of sEMG data, the convolution kernel parameters in the ResNet are improved for gesture recognition. According to the characteristics of muscle synergy, GCN was introduced into the model through a joint voting network to extract the muscle synergy feature of the sEMG signals. Moreover, VMD was also used to decompose the signal into multiple variable components to reconstruct more recognizable 3D sEMG images. The average classification accuracy of the hybrid model based on ResNet and GCN for the NinaPro DB1 dataset is superior to that of the existing research methods, reaching 90.07%. Clearly, the experimental results showed that the proposed model is effective due to the improved convolution kernel parameters of the ResNet, and the introduction of the GCN can further improve the accuracy of gesture recognition.

Our future work will focus on improving the robustness of our proposed hybrid model through other methods. For example, from the perspec-

tive that sEMG signals are sequential data, we can try to continue to improve the robustness of model recognition by introducing recurrent neural networks (RNNs).

## References:

- [1] W. Guo, P. Yao, X. Sheng, D. Zhang, and X. Zhu, "An enhanced human-computer interface based on simultaneous sEMG and NIRS for prostheses control," in *2014 IEEE International Conference on Information and Automation (ICIA)*, pp. 204-207, 2014.
- [2] Y. Fan and Y. Yin, "Active and progressive exoskeleton rehabilitation using multisource information fusion from EMG and force-position EPP," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 12, pp. 3314-3321, 2013.
- [3] Y. Li, X. Chen, X. Zhang, K. Wang, and J. Z. Wang, "A sign-component-based framework for chinese sign language recognition using accelerometer and sEMG data," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 10, pp. 2695-2704, 2012.
- [4] J. Cheng, X. Chen, Z. Y. Lu, K. Q. Wang, and M. F. Shen, "Key-press gestures recognition and interaction based on sEMG signals," in *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction (ICMI-MLMI'10)*, Association for Computing Machinery, New York, NY, USA, Article 36, pp. 1-4, 2010.
- [5] W. Qi, H. Su, and A. Aliverti, "A smartphone-based adaptive recognition and real-time monitoring system for human activities," *IEEE Transactions on Human-Machine Systems*, vol. 50, no. 5, pp. 414-423, 2020.
- [6] P. Artemiadis, "EMG-based robot control interfaces: Past, present and future," *Advances in Robotics & Automation*, vol. 1, no. 2, pp. 1-3, 2012.
- [7] Q. C. Ding, A. B. Xiong, X. G. Zhao, and J. D. Han, "A review on researches and applications of sEMG-based motion intent recognition method," *Acta Automatica Sinica*, vol. 42, no. 1, pp. 13-25, 2016.
- [8] X. H. Hu, A. G. Song, and H. J. Li, "Dexterous robot hand control system based on surface electromyography image," *Control Theory and Applications*, vol. 35, no. 12, pp. 1707-1714, 2018.
- [9] M. Tavakoli, C. Benussi, P. A. Lopes, L. B. Osorio, and A. T. De Almeida, "Robust hand gesture recognition with a double channel surface EMG wearable armband and SVM classifier," *Biomedical Signal Processing and Control*, vol. 46, pp. 121-130, 2018.
- [10] S. Jiang, Q. Gao, H. Liu, and P B Shull, "A novel, co-located EMG-FMG-sensing wearable armband for hand gesture recognition," *Sensors and Actuators A Physical*, vol. 301, no. 1, pp. 111738, 2020.
- [11] A. Waris, K. I. Khan, J. Mohsin, E. Kevin, J. Winnie, and K. E. Nlandu, "Multiday evaluation of techniques for EMG-based classification of hand motions," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 4, pp. 1526-1534, 2019.
- [12] A. Phinyomark, E. Campbell, and E. J. Scheme, *Surface electromyography (EMG) signal processing, classification, and practical considerations*. Singapore: Springer, pp. 3-29, 2020. DOI: 10.1007/978-981-13-9097-5\_1.
- [13] J. Too, A. R. Abdullah, T. N. S. T. Zawawi, N. M. Saad, and H. Musa, "Classification of EMG signal based on time domain and frequency domain features," *International Journal of Human and Technology Interaction (IJHaTI)*, vol. 1, no. 1, pp. 25-30, 2017.
- [14] N. Zheng, Y. R. Li, W. X. Zhang, and J. X. Li, "User-independent EMG gesture recognition based on task-specific muscle synergy," *Chinese Journal of Scientific Instrument*, vol. 42, no. 9, pp. 253-261, 2021.
- [15] M. Ghislieri, V. Agostini, and M. Knafitz, "Muscle synergies extracted using principal activations: Improvement of robustness and interpretability," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 2, pp. 453-460, 2020.
- [16] X. Luo, Wu X, Chen L, Y. Zhao, L. Zhang, G. Li, and W. Hou, "Synergistic myoelectrical activities of forearm muscles improving robust recognition of multi-fingered gestures," *Sensors*, vol. 19, no. 3, pp. 610, 2019.
- [17] S. Zhang, X. Zhang, S. Cao, X. Gao, X. Chen, and P. Zhou, "Myoelectric pattern recognition based on muscle synergies for simultaneous control of dexterous finger movements," *IEEE Transactions on Human-Machine Systems*, vol. 47, no. 4, pp. 576-582, 2017.

- [18] A. Masoumdoost, R. Saadatyar, and H. R. Kobravi, "A muscle synergies-based movements detection approach for recognition of the wrist movements," *EURASIP Journal on Advances in Signal Processing*, vol. 2020, no. 1, pp. 1-19, 2020.
- [19] F. Y. Zhou, L. P. Jin, and J. Dong, "Review of convolutional neural network," *Chinese Journal of Computers*, vol. 40, no. 6, pp. 1229-1251, 2017.
- [20] J. Liu, G. Wang, P. Hu, L. Y. Duan, and A. C. Kot, "Global context-aware attention LSTM networks for 3D action recognition," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, pp. 3671-3680, 2017.
- [21] L. Shi, Y. Zhang, J. Cheng, and H. Lu, "Two-stream adaptive graph convolutional networks for skeleton-based action recognition," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, pp. 12018-12027, 2019.
- [22] M. Jafarzadeh, D. C. Hussey, and Y. Tadesse, "Deep learning approach to control of prosthetic hands with electromyography signals," in *2019 IEEE International Symposium on Measurement and Control in Robotics (ISMCR)*, pp. A1-4-1-A1-4-11, 2019.
- [23] E. A. Chung and M. E. Benalcázar, "Real-time hand gesture recognition model using deep learning techniques and EMG signals," in *2019 27th European Signal Processing Conference (EUSIPCO)*, pp. 1-5, 2019.
- [24] M. Atzori, M. Cognolato, and H. Muller, "Deep learning with convolutional neural networks applied to electromyography data: A resource for the classification of movements for prosthetic hands," *Frontiers in Neurobotics*, vol. 10, pp. 9, 2016. doi: 10.3389/fnbot.2016.00009.
- [25] M. Atzori and H. Muller, "The NinaPro database: A resource for sEMG naturally controlled robotic hand prosthetics," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Milan, Italy, pp. 7151-7154, 2015.
- [26] S. J. Chen and Z. Z. Luo, "Research on gesture EMG recognition based on long short-term memory and convolutional neural network," *Chinese Journal of Scientific Instrument*, vol. 42, no. 2, pp. 162-170, 2021.
- [27] W. Wei, Q. Dai, Y. Wong, Y. Hu, M. Kankanhalli, and W. Geng, "Surface electromyography-based gesture recognition by multi-view deep learning," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 10, pp. 2964-2973, 2019.
- [28] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," arXiv preprint arXiv: 1609.02907, 2016.
- [29] M. Atzori, A. Gijsberts, C. Castellini, B. Caputo, and H. Müller, "Electromyography data for non-invasive naturally controlled robotic hand prostheses," *Sci Data*, vol. 1, pp. 140053, 2014.
- [30] W. Wei, Y. Wong, Y. Du, Y. Hu, M. Kankanhalli, and W. Geng, "A multi-stream convolutional neural network for sEMG-based gesture recognition in muscle-computer interface," *Pattern Recognition Letters*, vol. 119, no. 1, pp. 131-138, 2019.
- [31] Y. Cheng, G. Li, M. Yu, D. Jiang, J. Yun, Y. Liu, Y. Liu, and D. Chen, "Gesture recognition based on surface electromyography-feature image," *Concurrency and Computation: Practice and Experience*, vol. 33, no. 6, pp. e6051, 2021.
- [32] K. Englehart and B. Hudgins, "A robust, real-time control scheme for multifunction myoelectric control," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 7, pp. 848-854, 2003.
- [33] Y. Hu, Y. Wong, W. Wei, Y. Du, M. Kankanhalli, and W. Geng, "A novel attention-based hybrid CNN-RNN architecture for sEMG-based gesture recognition," *PLoS One*, vol. 13, no. 10, pp. 1-18, 2018.
- [34] A. Nazemi and A. Maleki, "Artificial neural network classifier in comparison with LDA and LS-SVM classifiers to recognize 52 hand postures and movements," in *2014 4th International Conference on Computer and Knowledge Engineering (ICCKE)*, IEEE, pp. 18-22, 2014.
- [35] N. K. Karnam, A. C. Turlapaty, S. R. Dubey, and B. Gokaraju, "Classification of sEMG signals of hand gestures based on energy features," *Biomed Signal Process Control*, vol. 70, pp. 102948, 2021.
- [36] A. E. Olsson, A. Björkman, and C. Antfolk, "Automatic discovery of resource-restricted convolutional neural network topologies for myoelectric pattern recognition," *Computers in Biology and Medicine*, vol. 120, pp. 103723, 2020.
- [37] Z. Yang, D. Jiang, Y. Sun, B. Tao, X. Tong, G. Jiang, M. Xu, J. Yun, Y. Liu, B. Chen, and J. Kong, "Dynamic gesture recognition using surface EMG signals based on multi-stream residual network," *Front Bioeng Biotechnol*, vol. 9, pp. 779353, 2021.
- [38] S. Wang, L. Huang, D. Jiang, Y. Sun, G. Jiang, J.

- Li, C. Zou, H. Fan, Y. Xie, H. Xiong, and B. Chen, "Improved multi-stream convolutional block attention module for sEMG-based gesture recognition," *Front Bioeng Biotechnol*, vol. 10, pp. 909023, 2022.
- [39] P. Xu, F. Li, and H. Wang, "A novel concatenate feature fusion RCNN architecture for sEMG-based hand gesture recognition," *PLoS ONE*, vol. 17, no. 1, pp. 1-19, 2022.
- [40] Y. Zhang, F. Yang, Q. Fan, A. Yang, and X. Li, "Research on sEMG-based gesture recognition by dual-view deep learning," *IEEE Access*, vol. 10, pp. 32928-32937, 2022.
- [41] S. Cao and C. Xu, "Research on gesture recognition algorithm of surface EMG signal based on deep learning model," *Science and Technology Vision*, vol. 259, no. 1, pp. 118-120, 2019.



**Xianjing Xu** received his B. Sc. degree from Fujian Agriculture and Forestry University in 2016. He is currently a master student at Fuzhou University. His research interests include intelligent detection and signal processing.



**Haiyan Jiang** received her B. Sc. degree from Shandong University of Science and Technology in 1998, and received her M. Sc. degree and Ph. D. degree both from Fuzhou University in 2005 and 2013, respectively. She is currently an associate professor in College of Electrical Engineering and Automation at Fuzhou University. Her main research interests include biomedical signal detection and processing.