



Improved Multi-Stream Convolutional Block Attention Module for sEMG-Based Gesture Recognition

Shudi Wang^{1,2}, Li Huang^{3,4*}, Du Jiang^{1,5*}, Ying Sun^{1,2,5*}, Guozhang Jiang^{1,2,5*}, Jun Li^{1,2}, Cejing Zou^{1,2}, Hanwen Fan^{1,2}, Yuanmin Xie⁵, Hegen Xiong⁵ and Baojia Chen^{6*}

¹Key Laboratory of Metallurgical Equipment and Control Technology of Ministry of Education, Wuhan University of Science and Technology, Wuhan, China, ²Research Center for Biomimetic Robot and Intelligent Measurement and Control, Wuhan University of Science and Technology, Wuhan, China, ³College of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan, China, ⁴Hubei Province Key Laboratory of Intelligent Information Processing and Real-time Industrial System, Wuhan University of Science and Technology, Wuhan, China, ⁵Hubei Key Laboratory of Mechanical Transmission and Manufacturing Engineering, Wuhan University of Science and Technology, Wuhan, China, ⁶Hubei Key Laboratory of Hydroelectric Machinery Design and Maintenance, China Three Gorges University, Yichang, China

OPEN ACCESS

Edited by:

Zhihua Cui,

Taiyuan University of Science and Technology, China

Reviewed by:

Nadica Miljković,

University of Belgrade, Serbia

Abdelalim Sadiq,

Ibn Tofail University, Morocco

*Correspondence:

Li Huang

huangli82@wust.edu.cn

Du Jiang

jiangdu@wust.edu.cn

Ying Sun

sunying65@wust.edu.cn

Guozhang Jiang

whjgz@wust.edu.cn

Baojia Chen

cbjia@163.com

Specialty section:

This article was submitted to Bionics and Biomimetics, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: 31 March 2022

Accepted: 17 May 2022

Published: 07 June 2022

Citation:

Wang S, Huang L, Jiang D, Sun Y, Jiang G, Li J, Zou C, Fan H, Xie Y, Xiong H and Chen B (2022) Improved Multi-Stream Convolutional Block Attention Module for sEMG-Based Gesture Recognition. *Front. Bioeng. Biotechnol.* 10:909023. doi: 10.3389/fbioe.2022.909023

As a key technology for the non-invasive human-machine interface that has received much attention in the industry and academia, surface EMG (sEMG) signals display great potential and advantages in the field of human-machine collaboration. Currently, gesture recognition based on sEMG signals suffers from inadequate feature extraction, difficulty in distinguishing similar gestures, and low accuracy of multi-gesture recognition. To solve these problems a new sEMG gesture recognition network called Multi-stream Convolutional Block Attention Module-Gate Recurrent Unit (MCBAM-GRU) is proposed, which is based on sEMG signals. The network is a multi-stream attention network formed by embedding a GRU module based on CBAM. Fusing sEMG and ACC signals further improves the accuracy of gesture action recognition. The experimental results show that the proposed method obtains excellent performance on dataset collected in this paper with the recognition accuracies of 94.1%, achieving advanced performance with accuracy of 89.7% on the Ninapro DB1 dataset. The system has high accuracy in classifying 52 kinds of different gestures, and the delay is less than 300 ms, showing excellent performance in terms of real-time human-computer interaction and flexibility of manipulator control.

Keywords: sEMG signals, gesture recognition, attention mechanisms, neural networks, multi-stream

INTRODUCTION

When the number of patients with physical disabilities is on the rise (Chen et al., 2021a; Chen et al., 2022a), dexterous hand devices are not fully satisfying the needs of patients. So it is extremely important to design a dexterous hand control system to help patients with forearm disabilities restore some of their limb functions and improve their quality of life (Ma et al., 2020; Xie et al., 2020). As a physiological signal closely related to human movement, the EMG signal can intuitively reflect the user's intention; and the EMG signal-based dexterous hand control system (Li et al., 2020; Yun et al., 2022; Xu et al., 2022) has received widespread attention. Because it is simple, safe to operate, and not susceptible to environmental influences such as light or environmental sound changes (Li et al., 2019a). There are two main approaches to obtain EMG signals: one uses needle electrodes to invade

the body and obtain physiological signals directly. The other is to analyze the user's movement status by placing electrodes to detect changes on the skin surface currently. Compared with the use of needle electrodes, sEMG has the advantages of non-invasive, painless measurement, easy acceptance by the subject, and simple operation. So it has been widely used in practice. For patients with hand disabilities, some forearm muscles remain at the residual limb, and their central nervous system is not damaged and can function normally.

However, due to the characteristics of very weak and noisy EMG signals, the effective recognition of EMG signals still needs further improvement (Jiang et al., 2019a; He et al., 2019; Tan et al., 2020; Yu et al., 2020). At present, the gesture recognition methods of EMG signals are mainly divided into traditional machine learning based and deep learning based (Lenz et al., 2015; He et al., 2016; Cheng et al., 2021; Duan et al., 2021); the traditional method consists of three parts, pre-processing (such as denoising and filtering), feature extraction, and classifier model classification. However, manual extraction of features and then classification is tedious, and the accuracy is not very satisfactory.

Deep learning is a method that requires massive data for experimentation. By pre-processing the initial signal and expanding the experimental data, researchers continuously optimize and improve various parameters in the deep learning model, repeatedly train the model using Convolutional Neural Network (CNN) (Weng et al., 2021; Tao et al., 2022), and continuously test to get the optimal experimental results, thus improving the recognition accuracy. At present, the deep learning model has made some progress in sEMG gesture recognition, but the accuracy is not high while ensuring real-time. The model's ability to fit multi-gesture sparse EMG signal data and extract features need to be further improved. The existing CNN-based EMG gesture recognition research does not make full use of the timing information of sEMG signal data, and difficult to apply in bionic hand control systems.

To solve the above problems, a multi-stream fusion network (Zhang et al., 2022) of one-dimensional convolutional neural network (CNN) + GRU is proposed, which embeds the attention mechanism (CBAM) in the CNN. CNN + GRU is used for processing to extract the hidden correlation characteristics between the sEMG sequence signals, and embeds an attention module to learn synergy of different sEMG feature channels, and the spatiotemporal features. At the same time, ACC signals are added for recognition to further improve the accuracy. Based on this network, a dexterous hand control system is established to classify the collected sEMG signals and control the bionic hand to do matching movements according to the user's intention, which can assist people with hand disabilities to live normally (Li et al., 2022; Tao et al., 2022). The contributions of this paper are as follows.

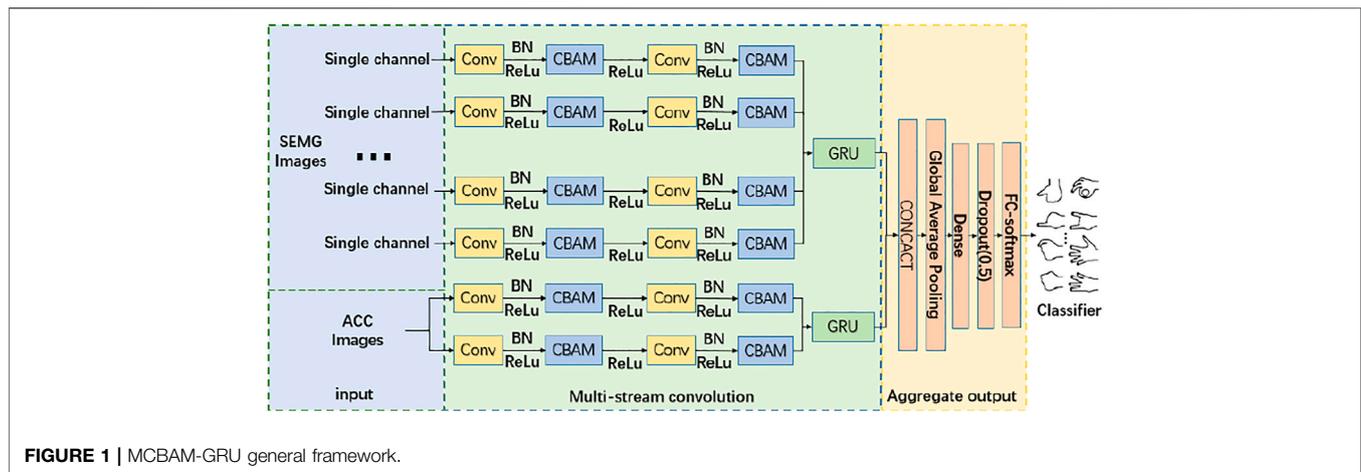
- 1) Acceleration signals and sEMG signals were collected to construct a dataset containing 52 different hand gestures.
- 2) Embedding CBAM units in a 1D convolutional network selectively emphasize informative features and suppress useless features on channels and spaces, enhancing the effective extraction of feature information from sparse channels while preventing overfitting.

- 3) A multi-stream fusion network based on CNN + GRU is designed to ensure accuracy while reducing the calculation time, making it more suitable for application in bionic hand control systems.

RELATED WORK

With the deepening of sEMG detection technology and the rapid development of computer technology, sEMG controlled human-machine interaction (Sun et al., 2020a; Chen et al., 2022c) systems can analyze the sEMG generated during the user's movement to obtain the human body's intention, and eventually control peripheral devices by transmitting movement commands (Pinto and Gupta, 2016; Zhang et al., 2019; Li et al., 2021; Liu et al., 2022a). Early prostheses were generally single-degree-of-freedom robotic arms with grasping capability only. Reiter (Scott, 1989) first used sEMG signals for prosthetic control in 1948, that's the earliest. Carrozza (Carozza et al., 2005) implemented a single-degree-of-freedom sEMG prosthetic hand control (Andreas et al., 2017; Sun et al., 2020b; Liu et al., 2022b) using a finite state machine. Many researchers have worked on the problem of the multiclassification of sEMG. Traditional machine learning algorithms (Yu et al., 2019; Chen et al., 2022b) need to extract time-domain, frequency-domain, or time-frequency-domain features from sEMG data and select appropriate classifiers, such as extreme learning machines (Shi et al., 2013; Sun et al., 2022b), random forests (Atzori et al., 2014a; Asif et al., 2017), linear discriminant analysis, and support vector machines (Chen and Zhang, 2019), to accomplish the gesture recognition task. However, traditional machine learning tends to decrease the recognition accuracy significantly with the increase of gesture movements. From machine learning to deep learning (Lenz et al., 2015; Qi et al., 2020; Huang et al., 2021; Tao et al., 2022), the accuracy of the multi-classification of EMG signals has been improving. Deep learning performs better in the multi-classification problem of sEMG gestures because of its strong data fitting and feature extraction ability (Li et al., 2019c; Sun et al., 2022a). In the method of gesture recognition using deep learning (Han et al., 2018; Jiang et al., 2019b; Liao et al., 2021), there are several major types of mainstream network algorithms: 1) convolutional neural networks (Tao et al., 2022; Sun et al., 2022c); 2) recurrent neural networks (Liu et al., 2022b); 3) network combining multi-class models (Zhao et al., 2022); 4) some novel network (Huang et al., 2019; Chen et al., 2021b; Sun et al., 2021; Liu et al., 2022c; Liu et al., 2022d; Wu et al., 2022; Zhao et al., 2022).

In terms of EMG gesture recognition, CNN-based deep learning algorithms (Shi et al., 2022) have proven to be highly advantageous by many researchers (Huang et al., 2020; Jiang et al., 2021a; Yang et al., 2021)). Manfredo et al. (Manfredo et al., 2016) found that the average result of recognition accuracy of classical classification methods is easily surpassed by a simple CNN structure on the Nina Pro database. P. Tsiganos (Tsiganos et al., 2018) et al. used CNN networks for gesture recognition and improved the accuracy by 3%. Yu Hu (Hu Y. et al., 2018) et al. proposed an attention-based hybrid CNN-RNN



(Recurrent Neural Network) model to process Nina Pro DB1, Nina Pro DB2, and Bio Pat Rec-26MOV compared to the normal hybrid CNN-RNN model databases, the accuracy was improved by 2.0%, 7.4% and 1.6%, respectively. Yuru Chen et al. use MYO hand ring to acquires upper limb EMG signals for data preprocessing, classification, and identification followed by real-time control of upper limb mechanical devices. Migratory learning, long and short-term memory networks, and recurrent neural networks were applied to EMG signal gesture recognition (Naik et al., 2014; Côté-Allard et al., 2017; Ding et al., 2018; Quivira et al., 2018; Xie et al., 2018; Chen et al., 2020), and Tsinganos et al. (Zardoshti-Kermani et al., 1995) researchers treated EMG signal-based gesture recognition as a sequence classification problem and introduced temporal convolutional networks for gesture recognition, with an improvement in recognition accuracy of about 5%. In general, machine learning methods for sEMG gesture recognition require low training data set size and short training time, but the requirements for researchers are relatively high; while deep learning-based methods have low or basically no requirements for feature set selection and certain requirements for sample data volume, because insufficient data will lead to poor recognition accuracy. We will combine attentional mechanisms and long short-term memory networks. Consider the sEMG signal as an image classification problem and time series classification as the basis for network design (Li et al., 2019b; Luo et al., 2020), establish a new network architecture for gesture recognition, further explore the optimization of network models, improve the recognition accuracy, and solve the problems of relatively long computation time, high hardware requirements in use, and unsuitable for the application of bionic. The problem of relatively long computation time, high hardware requirements in use, and unsuitable for the application of the actual control process of bionic hands (Xiao et al., 2021; Liu X. et al., 2022).

METHODS AND MATERIALS

sEMG signal is a signal with temporal order (Luo et al., 2020), and RNN has shown excellent prediction capability in dealing with

time series problems. GRU (Chung et al., 2014) is a further development based on RNN that can solve the problems of long-term learning dependence and long-term preservation of RNN and avoid gradient disappearance. The speedy, lightweight Conv1D combined with sequential-sensitive GRU to build the model can balance accuracy and speed, with fewer training parameters and faster convergence and iteration, which is beneficial to real-time recognition. Conv1D serves as a pre-processing step for GRU to shorten the identified sequences and extract local information before GRU processes the timing-related information.

MCBAM-GRU Net Architecture

A multistream CNN is proposed, which fusing attention mechanism and long short-term memory network, called MCBAM-GRU network, whose general framework is shown in **Figure 1**. After supervised training, end-to-end gesture recognition of surface EMG signals can be performed. The overall model can be divided into three stages, the data input, the multistream convolution and the aggregated output.

First, the data input phase divides the EMG signal by acquisition channel dimension to obtain a single-channel EMG map. The multistream convolutional network (Sun et al., 2020c; Tran and Lin, 2021) has six independent network branches. The features of the EMG signals of the six channels are extracted separately for each stream. Meanwhile, to improve the recognition effect, the sEMG and acceleration are fused by information fusion technique (Sun et al., 2020a; Liao et al., 2020; Tian et al., 2020; Hao et al., 2021a; Jiang et al., 2021b; Bai et al., 2022; Huang et al., 2022), and the signal surface EMG and ACC signals are used as the input of each stream of independent CNN, and each stream learns features independently by MCBAM. Since the single-channel sEMG map is essentially one-dimensional time-series data, a one-dimensional convolution kernel is used in the batch normalized convolution layer to learn the hidden correlation between sEMG sequence signals.

The aggregation output stage aggregates all the outputs of the multi-stream convolution stage and obtains the final recognition results. The first layer is the aggregation layer, which aggregates the outputs of the multi-stream convolution using a cascade

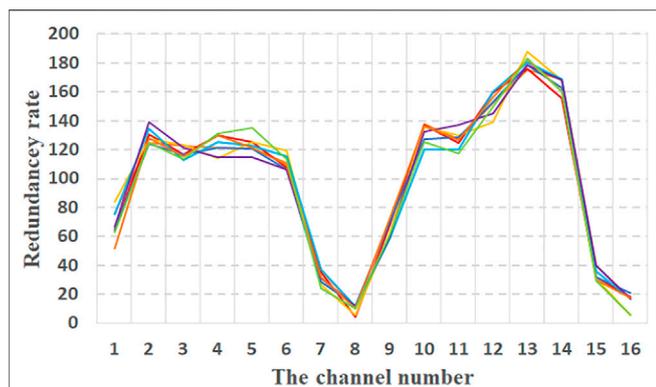


FIGURE 2 | Weighted channel redundancy.

stitching of feature channel dimensions. The second layer is a global mean pooling layer, which averages all pixel values of the feature map to obtain a value. The third layer is a full connected layer with dropout, which reduces the dimensionality of the output vector and adds dropout to prevent overfitting. The final layer uses a softmax-activated fully-connected layer to obtain the classification results. This layer first obtains a label vector g of the length of the number of gesture categories through the fully-connected layer and subsequently uses a softmax function to predict the category to which the label vector g belongs.

$$p(m|g) = \frac{\exp(g(m))}{\sum_{j=1}^M \exp(g(j))} \quad (1)$$

Where M represents the total number of gesture categories; $p(m|g)$ represents the probability that the vector g belongs to the Class m gesture, and the final category with the highest probability is the classification result m_0 .

$$m_0 = \arg \max(p(1|g), p(2|g), \dots, p(M|g)) \quad (2)$$

Redundant Channel Removal

For the region of EMG signal distribution studied in the experiment, the sEMG signals generated by different hand movements of muscles some distributed in forearm muscles have a certain regularity, that is, some muscles in do not produce useful signals or redundant information channels, these regions not only interfere with the classification, and will increase the amount of data and increase the computational burden, thus affecting the classification speed.

Therefore, a method of removing redundant channels is necessary (Sun et al., 2020b), which takes the arm without movement as a benchmark and takes the variance of the signal values of different actions to represent the degree of signal redundancy.

By using the above variance calculation values (Zhou et al., 2010) to grade the redundancy of 16 channels in each action. Simple sequencing coding methods are used to assign positive correlation weights to different levels. In order to ensure the

stability of the method, the cross-validation method (Scheme and Englehart, 2011) is used to verify. The collected data is divided into eight parts. One of them is taken out in turn and tested in the remaining data. Figure 2 shows the redundancy rate of 16 channels obtained after eight tests.

According to the weighting result, with high redundancy rate ten common redundant channels 2, 3, 4, 5, 6, 10, 11, 12, 13 and 14, channels are removed.

One Dimension Convolutional Block Attention Module

Attentional mechanisms (Hao et al., 2021b) has aroused the interest of many researchers because it have fewer parameters, faster speed, and better results in important areas such as machine translation (Hao et al., 2021a), speech recognition (Bahdanau et al., 2016; Rogowski et al., 2020), image recognition (Wang et al., 2017), and gradually started to be applied in sEMG gesture recognition (Hu J. et al., 2018; Jiang et al., 2019c; Liu et al., 2021) Previous CNN gesture recognition models often do not give enough attention to the characteristics of the EMG signal and do not make full use of the temporal information (Atzori et al., 2014b; Atzori et al., 2016; Geng et al., 2016; Ketykó et al., 2019). Therefore, this paper introduces a convolutional attention mechanism into the EMG gesture recognition method and designs a one-dimensional convolutional attention module based on time and feature channels to make it more applicable to EMG gesture recognition. A one-dimensional CBAM is added after the ReLU nonlinear function, which can redistribute the weights of EMG signals in different time frames and adaptively adjust the weights of each feature map to focus more on the effective features and suppress the useless features to some extent. The CBAM1D shown in Figure 3 is divided into the feature channel attention module and temporal attention module.

The specific main process of the feature channel attention module shows as follows: an input feature $F \in \mathbb{R}^{T \times C}$ along the time dimension to perform feature compression, a real number is used to represent the time dimension information. Temporal information is aggregated through the average pooling and maximum pooling. Generated two different feature vectors are $F_b^C \in \mathbb{R}^{1 \times C}$ and $F_n^C \in \mathbb{R}^{1 \times C}$. The formula is shown below.

$$F_b^C \in \frac{1}{T} \sum_{i=1}^T F^C(i) \quad (3)$$

$$F_n^C = \max(F^C(i)) \quad i = 1, 2, \dots, T$$

Among them: $F_b^C = [F_b^1, F_b^2, F_b^3, \dots, F_b^C]$; $F_n^C = [F_n^1, F_n^2, F_n^3, \dots, F_n^C]$.

The two feature vectors are sent to a shared network to generate two new feature vectors, which are combined using element-by-element summation to generate the feature channel attention map $M_C \in \mathbb{R}^{1 \times C}$. The shared network consists of a two-layer multilayer perceptron. $\mathbb{R}^{G_r \times 1}$ is the activation size of the first layer, which can reduce the number of parameters, where r is the compression rate; The second layer output size is $\mathbb{R}^{C \times 1}$. The feature channel attention module can be expressed as shown in Eq. 4.

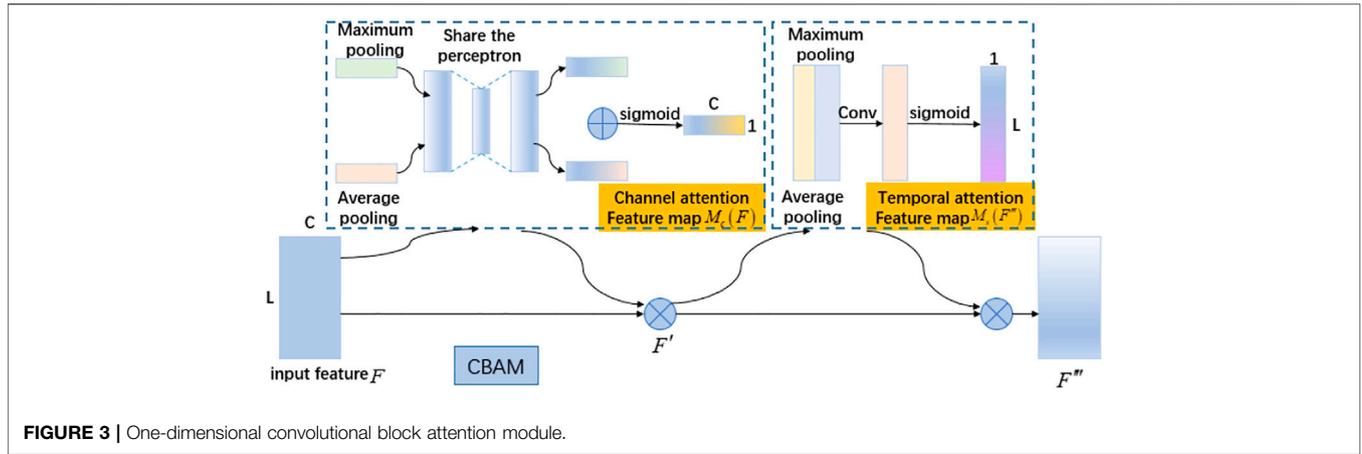


FIGURE 3 | One-dimensional convolutional block attention module.

$$M_C(F) = \sigma_g(W_1(W_0(F_b^C)) + W_1(W_0(F_n^C))) \quad (4)$$

where: σ_g is the sigmoid activation function; $W_0 \in \mathbb{R}^{C \times C}$; $W_1 \in \mathbb{R}^{C \times C}$.

The spatial attention module uses the time-series relationship of the myoelectric map to generate the temporal attention map $M_T \in \mathbb{R}^{T \times 1}$. Feature channel attention focuses on the effective feature maps, while temporal attention focuses more on the effective time frames in the decision window, which further complements the feature channel attention. The specific process of the temporal attention module is: feature compression along the feature channel dimension, using a real number to represent the feature channel dimension information of the feature. The average pooling and maximum pooling are used to aggregate the feature channel information to obtain two feature vectors, $F_b^T \in \mathbb{R}^{T \times 1}$ and $F_n^T \in \mathbb{R}^{T \times 1}$.

$$F_b^T = \frac{1}{C} \sum_{j=1}^C F^t(j) \quad (5)$$

$$F_n^T = \max(F^t(j)) \quad j = 1, 2, \dots, C$$

Among them: $F_b^T = [F_b^1, F_b^2, F_b^3, \dots, F_b^T]$; $F_n^T = [F_n^1, F_n^2, F_n^3, \dots, F_n^T]$

Two feature vectors are cascaded and stitched by feature channel axes to generate a one-dimensional temporal attention map M_T through a standard convolutional layer. Temporal attention module can be represented by Eq. 6.

$$M_T(F) = \sigma_g(f^{3 \times 1}([F_b^T; F_n^T])) \quad (6)$$

where: σ_g is the sigmoid activation function; $f^{3 \times 1}$ represents the convolution kernel is a 3*1 one-dimensional convolution.

Combining the time and feature channel attention modules, the total process can be summarized as follows

$$\begin{aligned} F' &= M_C(F) \otimes F \\ F'' &= M_T(F') \otimes F' \end{aligned} \quad (7)$$

where: \otimes represents the element-by-element multiplication; F'' is the final output of the 1D CBAM module, the input features after remapping.

TABLE 1 | The correspondence between 16 electrodes and forearm muscle.

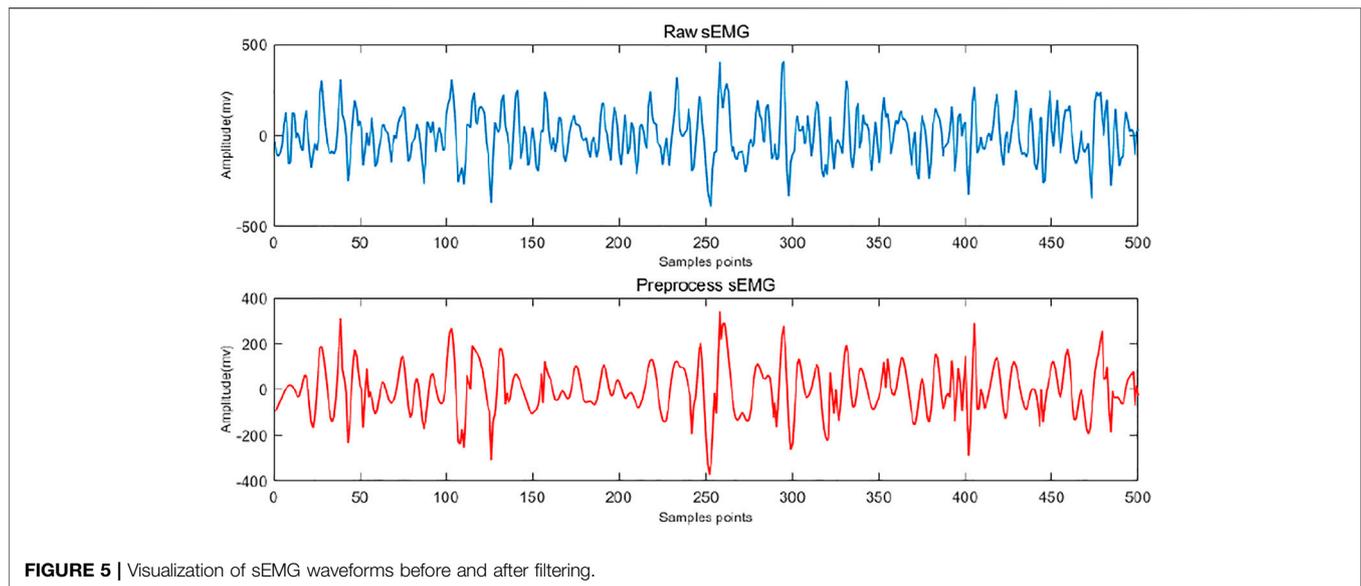
Electrode Number	Strong Related Muscles	Weak Related Muscles
Electrodes 1 and 9	Finger deep flexors	
Electrodes 2 and 10	Ulnar carpal flexor	Finger deep flexors
Electrodes 3 and 11	Superficial finger flexors	Palmaris Longus
Electrodes 4 and 12	Radial wrist flexor	Palmaris Longus
Electrodes 5 and 13	Brachioradialis	
Electrodes 6 and 14	Radial wrist extensors	
Electrodes 7 and 15	Finger extensor muscle	
Electrodes 8 and 16	Ulnar carpal extensor	Little finger extensors

EMG Signal Acquisition

The experiments use the ELONXI EMG instrument produced by Hangzhou Jiaopu Technology Company as the measurement sensor. ELONXI electromyograph has a 16-channel sEMG signal sensor. The sampling frequency is 1,000 Hz. This electromyograph has a high sampling frequency and is easy to wear, simple to use and low cost. This device is convenient for data acquisition and uses with dexterous hand mechanical devices. Before the experiment, the surface of each electrode of the cuff and the skin surface of the volunteer were gently wiped with alcohol cotton. Volunteers need to wait a few minutes for the skin surface to dry naturally and put the cuff on the forearm. The 16 electrodes of the cuff were numbered and evenly distributed on the surface of the arm. The correspondence between 16 electrodes and forearm muscles is shown in Table 1.

This experiment was approved by the Research Ethics Committee of Wuhan University of Science and Technology of China. Before the experiment, the relevant content of the experiment and the risks have been informed in detail to the 10 healthy subjects. And then they have signed informed consent forms. The experimental environment was quiet and free of absolute noise, Subjects seated in a chair where their hands can comfortably be placed on the table. They were asked to perform the corresponding right-handed movements according to the cues on the computer screen.

Volunteers should try to maintain the same speed and force, with the muscle relaxation state as the initial state and each target movement as the end state of one movement. To ensure that there



Where x_i and P_i represents the peak value of the i th point of sEMG in the time sequence; P_i represents the power value of the i th point of sEMG on the spectrum; N represents the number of signal sampling points; M is the signal bandwidth. LLC and LHC are the lower and upper cut-off frequencies of the low-frequency band, respectively; HLC and HHC are the lower and upper cut-off frequencies of the high-frequency band.

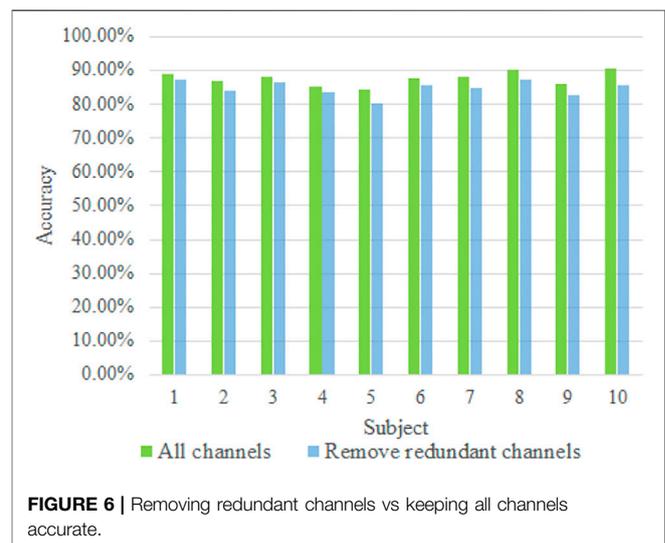
EXPERIMENT

The dataset is randomly divided into two groups: one is the training set, and the other is the test set. The training set contains 800 sets for each gesture, and each test set contains 100 sets. Experimental environment hardware: Intel(R) Core(TM) i7-9700K CPU@3.60 GHz; memory: 32.00 GB. All experiments are implemented by PyTorch 1.7.0 + cu110 on NVIDIA GTX 1080Ti GPU.

Experimental Results and Analysis

During the experiment, redundant channels were removed to reduce the amount of computational data and to speed up the computational rate. Therefore, it is necessary to construct a comparison experiment between the sEMG signals after all channels were selected and the redundant channels were removed. In the comparison experiment, the network model was fine-tuned, and to prove its validity, no ACC information was added for fusion while the inputs were the original sEMG images. The comparison validation yielded the results shown in **Figure 6**.

By comparing the recognition rates of ten groups, the experimental results show that when the redundant channels are removed, the recognition efficiency does not change significantly, and considering the recognition rate and data dimension, it is necessary to remove the redundant channels.



Using the step-by-step debugging function of the Pycharm development environment, we input the action segment sEMG signal feature data from any test set into the trained network, and count the time from receiving data to output signal category as the delay time to measure the network computation speed. In order to exclude the chance influence of a certain input data, this paper brings in 10 times of data and calculates the average of network delay time. To ensure that the experimental results are not affected by the particular sEMG signal characteristics of a particular test subject, three subjects from the Nina Pro-DB1 database (Atzori et al., 2012; Atzori et al., 2014a) are selected for individual training and testing in this paper. The experimental results are shown in **Table 2**.

From **Table 2**, it can be seen that removing redundant channels has lower latency and higher accuracy than keeping all channels.

TABLE 2 | Classification under removing redundant channels and keeping all channels.

Subjects	Channel selection	Delay (ms)	Testing Set Accuracy
Subject 1	All channels	82.5	85.3
	Removing redundant channels	64.3	82.2
Subject 2	All channels	76.1	78.1
	Removing redundant channels	61.9	76.6
Subject 3	All channels	80.8	87.0
	Removing redundant channels	63.8	84.2

TABLE 3 | Average gesture recognition accuracy ablation experiment results.

Method	Average Gesture Recognition Accuracy Rate/%
multi-stream convolution	68.3
+ batch normalization (BN)	80.2
+ BN + Gate Recurrent Unit (GRU)	82.1
+ BN+ 1D CBAM	83.6
+ BN+ 1D CBAM + GRU	86.0

Therefore, removing the redundant channels with less data is more suitable for the application of the bionic hand control process.

To better validate the effectiveness of the MCBAM-GRU sEMG gesture recognition network, the following ablation experiments were performed using a matchless normalized multistream convolutional network as the baseline model for the five experiments. Raw sEMG images were used as the input source of the network and no ACC information was input for fusion.

Experiment I: multi-stream convolution.

Experiment II: multi-stream convolution + batch normalization (BN).

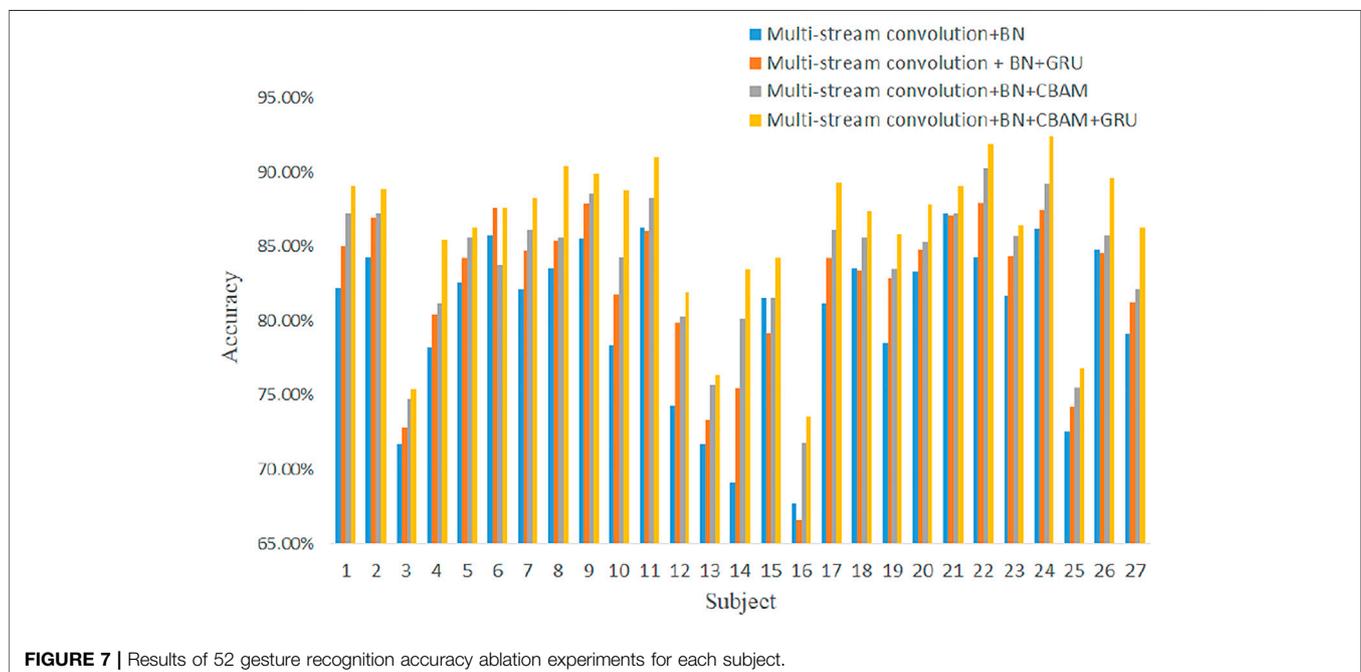
Experiment III: multi-stream convolution + batch normalization + Gate Recurrent Unit (GRU).

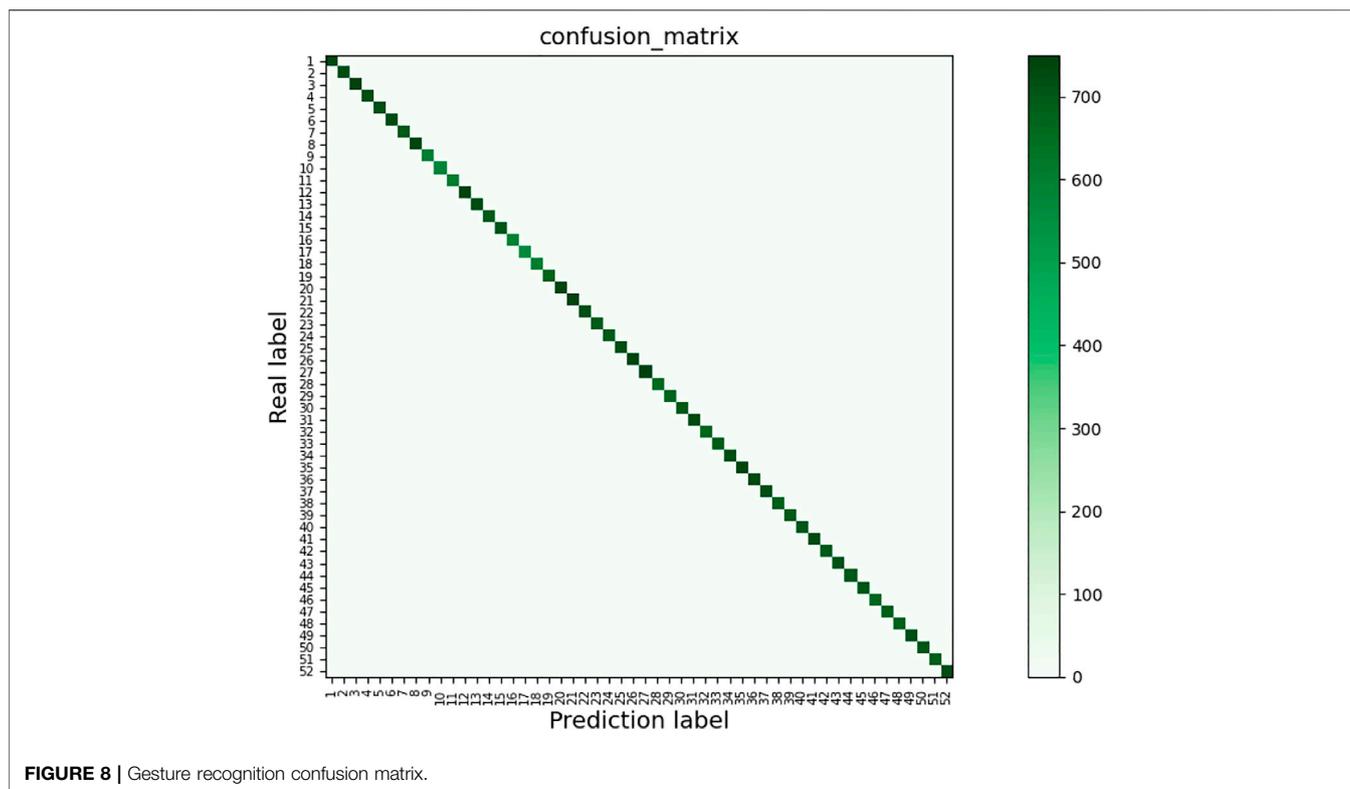
Experiment IV: multi-stream convolution + batch normalization + one-dimensional Convolutional Block Attention Module (CBAM).

Experiment V: multi-stream convolution + batch normalization + one-dimensional CBAM + Gate Recurrent Unit (GRU).

The results of the average gesture recognition accuracy ablation experiment are shown in **Table 3**. The results of the 52 gesture recognition accuracy ablation experiments for each subject are shown in **Figure 7**.

From the experimental results in **Figure 7**, it is clear that the use of batch normalization has the greatest impact on recognition accuracy. The reason is that batch normalization can reduce the effect of internal covariate bias, reduce the sensitivity of different choices of parameter initialization and learning rate on the impact of model performance, and also facilitate gradient descent, which helps the model converge quickly. Therefore, adding batch normalization can help the network model fit the data set better and achieve higher performance. Adding one-dimensional convolutional





attention to the multi-stream convolutional network can increase the average gesture recognition accuracy to 83.6% for 52 gestures for 27 subjects; adding GRU to the multi-stream convolutional attention mechanism network structure can increase the average gesture recognition accuracy to 86.0%. CBAM is introduced in the multi-stream network, which enables the network to learn saliency information in the image, making the important features in the image more salient and improving the expressiveness of the network without adding too many extra parameters and training time. The attention module generates temporal attention maps in the temporal dimension, giving more weight coefficients to the more important time frames in the time window and suppressing them with fewer weight coefficients on the contrary; in addition, it generates feature channel attention maps in the feature channel dimension, reinforcing the more effective feature maps and weakening the useless ones. The addition of the GRU module can further improve the network accuracy, enhance feature learning, and optimize network performance.

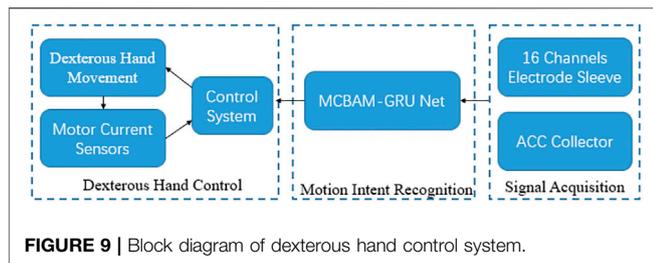
Figure 8 shows the confusion matrix of the MCBAM-GRU network for gesture recognition, containing the prediction results for 52 gestures for 27 individuals. In this experiment, the acceleration (ACC) signal is input into the network as an independent branch, the characteristic EMG signal is the input source of the network, and other conditions remain unchanged. The horizontal coordinates are the predicted gesture labels and the vertical coordinates are the actual gesture labels. From **Figure 8**, we can see that the

TABLE 4 | Comparison results of different approaches on NinaPro DB1.

Algorithms	Accuracy (%)
Random Forest (Atzori et al., 2014b)	75.2
Atzori_Net (Atzori et al., 2016)	66.6
Geng_Net (Geng et al., 2016)	77.8
CNN (Du et al., 2017)	79.5
ELM (Cene et al., 2019)	75.1
MSFusionNet (Wei et al., 2019)	85.0
MCBAM-GRU (ours)	89.7

recognition accuracy of most of the movements are relatively high and are concentrated on the diagonal, but there are also individual gestures with low recognition accuracy, such as gestures 9, 10, 11, 16, and 17, which are easily misidentified because they are relatively similar. 9, 10, and 11 are all thumb movement gestures and have very similar force points, so the recognition is poor; 16 and 17 are also thumb movement gestures, but the difference of only the direction of thumb movement is very similar, resulting in poor recognition.

To show the advantages of our model, MCBAM-GRU network recognition model is compared with other models that have been studied in recent years. The results on the NinaPro DB1 dataset are shown in **Table 4**. The NinaPro DB1 dataset contains 52 different gestures of 27 healthy subjects, the same gestures contained in the dataset used in this paper. So we can compare the experimental results horizontally to conclude that our proposed MCBAM-GRU network algorithm outperforms other algorithms.



From the comparison of the results in the above table, it is obvious that with the application of deep learning on sEMG gestures, many researchers did not achieve better recognition results, and most of them did not exceed 80% recognition rate, which is only comparable to the recognition results of traditional methods. Except for the results in this Section, the recognition rate achieved by Wei's proposed multi-stream convolutional neural network can reach 85.0%, while the multi-view fusion network proposed also can achieve a satisfactory recognition rate on the NinaPro DB1 dataset, which can reach 89.7%, 4.7% higher than the former.

Validation Experiments

A 6-channel sEMG gesture recognition based dexterous hand control system is designed as a validation experiment platform, and the block diagram of the dexterous hand control system is shown in **Figure 9**. The system mainly contains three parts: sEMG signal acquisition, action intention recognition and prosthetic hand control. The control process of the dexterous hand is as follows: sEMG signal and acceleration acquisition equipment collects sEMG data from the user's forearm in real time and sends it to the computer by wireless communication. The collected data is pre-processed and the redundant channels are removed as the input of the MCBAM-GRU network. The trained MCBAM-GRU network outputs the predicted current gesture, which is sent as the control signal of the dexterous hand to the dexterous hand console. Fifty-two bionic hand movements were preset in the console program, corresponding to 52 movements from the Nina-Pro-DB1 database (Atzori et al., 2016).

The 16-channel sEMG sensor cuff can adapt to different users' arm circumference and keep the electrode in good fit with the forearm; meanwhile, the sensor electrode is made of stainless steel, which is convenient to wear. The device connects the EMG collector to the computer via Bluetooth for fast data upload. The EMG cuff is worn on the right hand, and the accelerometer is placed close to the back of the hand to keep the forearm horizontal. The dexterous prosthesis used is an adult-sized anthropomorphic SR-RH8D (shown in **Figure 10**) developed by Seed Robotics in the United Kingdom. It is designed with underdrive technology, which allows the dexterous hand to be adaptive for precise control of objects with different or complex shapes.

The SR-RH8D has a total of 19 degrees of freedom, including the contralateral thumb and a full spherical wrist joint. In addition, it has eight compact and powerful internal actuators (seed_1-seed_8), all of which are contained within the unit and

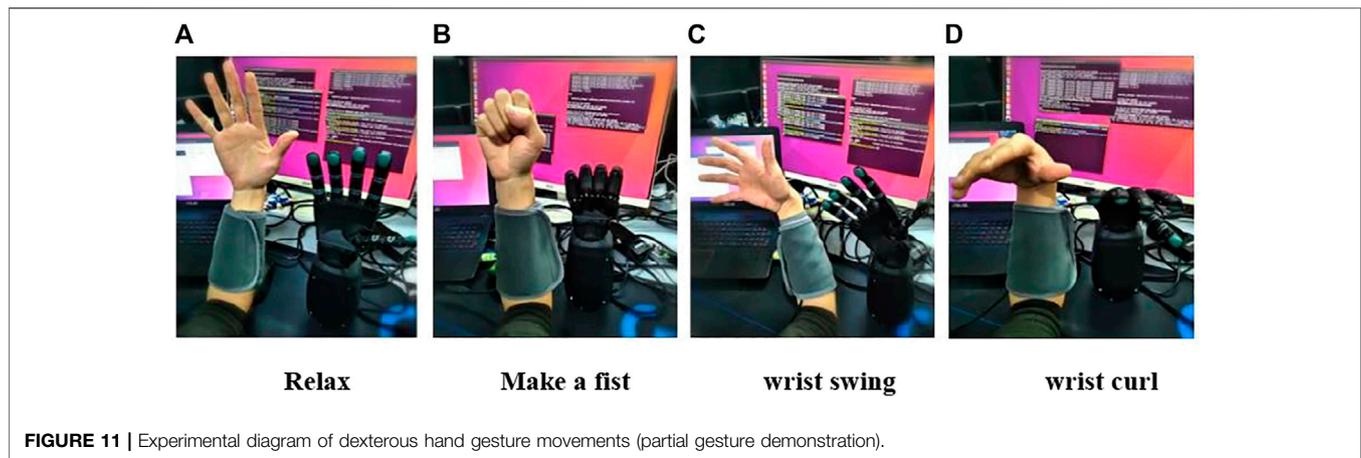


TABLE 5 | Servo and dexterous hand motion correspondence.

Servo Number	Function
Seed_1	Wrist rotation
Seed_2	Wrist swaying from side to side
Seed_3	Flexion of the wrist
Seed_4	Inward thumb
Seed_5	Thumb flexion
Seed_6	Index finger flexion
Seed_7	Flexion of the middle finger
Seed_8	Ring finger and pinky joint action

can be controlled independently by the user, and the relative relationship between the actuators and the dexterous hand motion is shown in **Table 5**.

The motor and control circuit of the dexterous hand are fully integrated into the dexterous hand, and which can be connected to a computer *via* a serial communication protocol. The trained multi-stream convolutional attention network is used to predict the gesture actions for the new input sEMG signals, and finally the predicted results are sent to the SR-RH8D dexterous hand as instructions to complete the corresponding gesture actions. The resting action was added to the experiment as a transition



between the previous gesture and the next gesture action. The experimental results of some gestures are shown in **Figure 11**.

CONCLUSION

Aiming at the insufficient effective feature extraction of sEMG timing information, the poor performance of gesture recognition speed and accuracy of sparse surface EMG signals, and the difficulty of application in dexterous hand control system, a multi-stream convolutional neural network MCBAM-GRU that integrates attention mechanism is proposed. It uses multi-stream convolution to excavate characteristics of the sparse channels of multiple acquisition channels. The one-dimensional convolutional attention module and the GRU module are added to learn important timing information and focus on more differentiated signal areas. Meanwhile the network adaptively learns the importance of different feature maps and strengthens the feature maps with stronger correlation, which ensure accuracy and greatly reduce the calculation time. The network can recognize 52 gestures using 6-channel surface electromyography and acceleration signals as model inputs. This multi-stream convolutional attention of this network can effectively prevent the overfitting problem of sEMG signals, while the use of GRU further improves the network accuracy. The prediction accuracy of the collected experimental data reached 94.1%, which is a satisfactory practicality. A dexterous hand control system is built to verify its feasibility. This network great potential for wider applications in future fields such as muscle fatigue and sensor electrode deflection. In subsequent studies, it is necessary to address differences in hand gestures, differences in arm size, and differences in the speed and strength of hand movements between individuals. There is a need for an in-depth study to obtain a pervasive multi-gesture recognition

REFERENCES

Andreas, P., Marcus, G., Kate, S., and Platt, R. (2017). Grasp Pose Detection in Point Clouds. *Int. J. Robotics Res.* 36 (13-14), 1455–1473.

algorithm, Further simplify the network model, and broaden its application in resource-constrained embedded platforms.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

SW and DJ provided research ideas and plans; YS, HF, and YX wrote programs and conducted experiments. GJ and HX analyzed and explained the simulation results; BC and LH improved the algorithm. JL and CZ co-authored the manuscript, and were responsible for collecting data; DJ and YS revised the manuscript for the corresponding author and approved the final submission.

FUNDING

This work was supported by grants from the National Natural Science Foundation of China (Grant Nos.52075530, 51575407, 51505349, 51975324, 61733011, and 41906177); the Grants of Hubei Provincial Department of Education (D20191105); the Grants of National Defense PreResearch Foundation of Wuhan University of Science and Technology (GF201705) and Open Fund of the Key Laboratory for Metallurgical Equipment and Control of Ministry of Education in Wuhan University of Science and Technology (2018B07,2019B13) and Open Fund of Hubei Key Laboratory of Hydroelectric Machinery Design and Maintenance in Three Gorges University(2020KJX02, 2021KJX13).

Asif, U., Bennamoun, M., and Sohel, F. A. (2017). RGB-D Object Recognition and Grasp Detection Using Hierarchical Cascaded Forests. *IEEE Trans. Robot.* 33 (3), 547–564. doi:10.1109/tro.2016.2638453

Atzori, M., Gijsberts, A., Castellini, C., Caputo, B., Hager, A. G., Elsig, S., et al. (2014a). Electromyography Data for Non-invasive Naturally-Controlled

- Robotic Hand Prostheses. *Sci. Data* 1 (1), 140053–140113. doi:10.1038/sdata.2014.53
- Atzori, M., Gijsberts, A., Kuzborskij, I., Elsig, S., Hager, A. G., Deriaz, O., Castellini, C., Müller, H., and Caputo, B. (2014b). Characterization of a Benchmark Database for Myoelectric Movement Classification. *IEEE Trans. Neural Syst. Rehabil. Eng.* 23 (1), 73–83. doi:10.1109/TNSRE.2014.2328495
- Atzori, M., Cognolato, M., and Müller, H. (2016). Deep Learning with Convolutional Neural Networks Applied to Electromyography Data: A Resource for the Classification of Movements for Prosthetic Hands. *Front. Neurobot.* 10, 9. doi:10.3389/fnbot.2016.00009
- Atzori, M., Gijsberts, A., Heynen, S., Hager, A. G. M., Deriaz, O., Van Der Smagt, P., and Müller, H. (2012). “Building the Ninapro Database: A Resource for the Biorobotics Community,” in 2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechanics (BioRob) (IEEE), 1258–1265.
- Bahdanau, D., Chorowski, J., Serdyuk, D., Brakel, P., and Bengio, Y. (2016). “March). End-To-End Attention-Based Large Vocabulary Speech Recognition,” in 2016 IEEE international conference on acoustics, speech and signal processing (ICASSP) (IEEE), 4945–4949.
- Bai, D., Sun, Y., Tao, B., Tong, X., Xu, M., Jiang, G., et al. (2022). Improved SSD Target Detection Method Based on Deep Feature Fusion. *Concurrency Comput. Pract. Exp.* 34 (4), e6614. doi:10.1002/CPE.6614
- Carozza, M. C., Cappiello, G., Stellin, G., Zaccone, F., Vecchi, F., Micera, S., et al. (2005). “On the Development of a Novel Adaptive Prosthetic Hand with Compliant Joints: Experimental Platform and EMG Control,” in 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems (IEEE), 1271–1276. doi:10.1109/iroso.2005.1545585
- Cene, V., Tosin, M., Machado, J., and Balbinot, A. (2019). Open Database for Accurate Upper-Limb Intent Detection Using Electromyography and Reliable Extreme Learning Machines. *Sensors* 19 (8), 1864. doi:10.3390/s19081864
- Chen, H., Zhang, Y., Li, G., Fang, Y., and Liu, H. (2020). “Surface Electromyography Feature Extraction via Convolutional Neural Network,” 185–196. doi:10.1007/s13042-019-00966-xInt. J. Mach. Learn. Cyber.111
- Chen, T., Jin, Y., Yang, J., and Cong, G. (2022b). Identifying Emergence Process of Group Panic Buying Behavior under the COVID-19 Pandemic. *J. Retail. Consumer Serv.* 67, 102970. doi:10.1016/j.jretconser.2022.102970
- Chen, T., Peng, L., Yang, J., Cong, G., and Li, G. (2021a). Evolutionary Game of Multi-Subjects in Live Streaming and Governance Strategies Based on Social Preference Theory during the COVID-19 Pandemic. *mathematics* 9 (21), 2743. doi:10.3390/math9212743
- Chen, T., Qiu, Y., Wang, B., and Yang, J. (2022c). Analysis of Effects on the Dual Circulation Promotion Policy for Cross-Border E-Commerce B2B Export Trade Based on System Dynamics during COVID-19. *Systems* 10 (1), 13. doi:10.3390/systems10010013
- Chen, T., Rong, J., Yang, J., and Cong, G. (2022a). Modeling Rumor Diffusion Process with the Consideration of Individual Heterogeneity: Take the Imported Food Safety Issue as an Example during the COVID-19 Pandemic. *Front. Public Health* 10. doi:10.3389/fpubh.2022.781691
- Chen, T., Yin, X., Yang, J., Cong, G., and Li, G. (2021b). Modeling Multi-Dimensional Public Opinion Process Based on Complex Network Dynamics Model in the Context of Derived Topics. *Axioms* 10 (4), 270. doi:10.3390/axioms10040270
- Chen, W., and Zhang, Z. (2019). “Hand Gesture Recognition Using sEMG Signals Based on Support Vector Machine,” in 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITTAIC) (IEEE), 230–234.
- Cheng, Y., Li, G., Li, J., Sun, Y., Jiang, G., Zeng, F., et al. (2020). Visualization of Activated Muscle Area Based on sEMG. *Ijs* 38 (3), 2623–2634. doi:10.3233/jifs-179549
- Cheng, Y., Li, G., Yu, M., Jiang, D., Yun, J., Liu, Y., et al. (2021). Gesture Recognition Based on Surface Electromyography -feature Image. *Concurr. Comput. Pract. Exper* 33 (6), e6051. doi:10.1002/cpe.6051
- Chung, J., Gulcehre, C., Cho, K., and Bengio, Y. (2014). *Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling*. arXiv preprint arXiv:1412.3555.
- Côté-Allard, U., Fall, C. L., Campeau-Lecours, A., Gosselin, C., Lavolette, F., and Gosselin, B. (2017). “Transfer Learning for sEMG Hand Gestures Recognition Using Convolutional Neural Networks,” in 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (IEEE), 1663–1668.
- Ding, Z., Yang, C., Tian, Z., Yi, C., Fu, Y., and Jiang, F. (2018). sEMG-Based Gesture Recognition with Convolution Neural Networks. *Sustainability* 10 (6), 1865. doi:10.3390/su10061865
- Du, Y., Jin, W., Wei, W., Hu, Y., and Geng, W. (2017). Surface EMG-Based Inter-session Gesture Recognition Enhanced by Deep Domain Adaptation. *Sensors* 17 (3), 458. doi:10.3390/s17030458
- Duan, H., Sun, Y., Cheng, W., Jiang, D., Yun, J., Liu, Y., et al. (2021). Gesture Recognition Based on Multi-modal Feature Weight. *Concurr. Comput. Pract. Exper* 33 (5), e5991. doi:10.1002/cpe.5991
- Geng, W., Du, Y., Jin, W., Wei, W., Hu, Y., and Li, J. (2016). Gesture Recognition by Instantaneous Surface EMG Images. *Sci. Rep.* 6 (1), 36571–36578. doi:10.1038/srep36571
- Han, J., Zhang, D., Cheng, G., Liu, N., and Xu, D. (2018). Advanced Deep-Learning Techniques for Salient and Category-specific Object Detection: a Survey. *IEEE Signal Process. Mag.* 35 (1), 84–100. doi:10.1109/msp.2017.2749125
- Hao, Z., Wang, Z., Bai, D., Tao, B., Tong, X., and Chen, B. (2021b). Intelligent Detection of Steel Defects Based on Improved Split Attention Networks. *Front. Bioeng. Biotechnol.* 9. doi:10.3389/fbioe.2021.810876
- Hao, Z., Wang, Z., Bai, D., and Zhou, S. (2021a). Towards the Steel Plate Defect Detection: Multidimensional Feature Information Extraction and Fusion. *Concurr. Comput. Pract. Exper* 33 (21), e6384. doi:10.1002/CPE.6384
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep Residual Learning for Image Recognition. IEEE Conference on Computer Vision and Pattern Recognition, 770–778.
- He, Y., Li, G., Liao, Y., Sun, Y., Kong, J., Jiang, G., et al. (2019). Gesture Recognition Based on an Improved Local Sparse Representation Classification Algorithm. *Clust. Comput.* 22 (5), 10935–10946. doi:10.1007/s10586-017-1237-1
- Hu, J., Shen, L., and Sun, G. (2018). Squeeze-and-excitation Networks. In Proceedings of the IEEE conference on computer vision and pattern recognition. (pp. 7132–7141).doi:10.1109/cvpr.2018.00745
- Hu, Y., Wong, Y., Wei, W., Du, Y., Kankanhalli, M., and Geng, W. (2018). A Novel Attention-Based Hybrid CNN-RNN Architecture for sEMG-Based Gesture Recognition. *PLoS one*, 13(10), e0206049.doi:10.1371/journal.pone.0206049
- Huang, L., Chen, C., Yun, J., Sun, Y., Tian, J., Hao, Z., et al. (2022). Multi-scale Feature Fusion Convolutional Neural Network for Indoor Small Target Detection. *Front. Neurobotics* 85 (16), 881021. doi:10.3389/fnbot.2022.881021
- Huang, L., Fu, Q., He, M., Jiang, D., and Hao, Z. (2021). Detection Algorithm of Safety Helmet Wearing Based on Deep Learning. *Concurr. Comput. Pract. Exper* 33 (13), e6234. doi:10.1002/CPE.6234
- Huang, L., Fu, Q., Li, G., Luo, B., Chen, D., and Yu, H. (2019). Improvement of Maximum Variance Weight Partitioning Particle Filter in Urban Computing and Intelligence. *IEEE Access* 7, 106527–106535. doi:10.1109/ACCESS.2019.2932144
- Huang, L., He, M., Tan, C., Jiang, D., Li, G., and Yu, H. (2020). Jointly Network Image Processing: Multi-task Image Semantic Segmentation of Indoor Scene Based on CNN. *IET image process* 14 (15), 3689–3697. doi:10.1049/iet-ipr.2020.0088
- Jiang, D., Li, G., Sun, Y., Hu, J., Yun, J., and Liu, Y. (2021a). Manipulator Grabbing Position Detection with Information Fusion of Color Image and Depth Image Using Deep Learning. *J. Ambient. Intell. Hum. Comput.* 12 (12), 10809–10822. doi:10.1007/s12652-020-02843-w
- Jiang, D., Li, G., Sun, Y., Kong, J., Tao, B., and Chen, D. (2019c). Grip Strength Forecast and Rehabilitative Guidance Based on Adaptive Neural Fuzzy Inference System Using sEMG. *Pers. Ubiquit Comput.* doi:10.1007/s00779-019-01268-3
- Jiang, D., Li, G., Sun, Y., Kong, J., and Tao, B. (2019b). Gesture Recognition Based on Skeletonization Algorithm and CNN with ASL Database. *Multimed. Tools Appl.* 78 (21), 29953–29970. doi:10.1007/s11042-018-6748-0
- Jiang, D., Li, G., Tan, C., Huang, L., Sun, Y., and Kong, J. (2021b). Semantic Segmentation for Multiscale Target Based on Object Recognition Using the Improved Faster-RCNN Model. *Future Gener. Comput. Syst.* 123, 94–104. doi:10.1016/j.future.2021.04.019
- Jiang, D., Zheng, Z., Li, G., Sun, Y., Kong, J., Jiang, G., et al. (2019a). Gesture Recognition Based on Binocular Vision. *Clust. Comput.* 22 (6), 13261–13271. doi:10.1007/s10586-018-1844-5
- Ketykó, I., Kovács, F., and Varga, K. Z. (2019). “Domain Adaptation for Semg-Based Gesture Recognition with Recurrent Neural Networks,” in 2019 International Joint Conference on Neural Networks(IJCNN) (IEEE), 1–7.

- Lenz, I., Lee, H., and Saxena, A. (2015). Deep Learning for Detecting Robotic Grasps. *Int. J. Robotics Res.* 34 (4/5), 705–724. doi:10.1177/0278364914549607
- Li, C., Li, G., Jiang, G., Chen, D., and Liu, H. (2020). Surface EMG Data Aggregation Processing for Intelligent Prosthetic Action Recognition. *Neural Comput. Applic* 32 (22), 16795–16806. doi:10.1007/s00521-018-3909-z
- Li, G., Jiang, D., Zhou, Y., Jiang, G., Kong, J., and Manogaran, G. (2019a). Human Lesion Detection Method Based on Image Information and Brain Signal. *IEEE Access* 7, 11533–11542. doi:10.1109/access.2019.2891749
- Li, G., Li, J., Ju, Z., Sun, Y., and Kong, J. (2019b). A Novel Feature Extraction Method for Machine Learning Based on Surface Electromyography from Healthy Brain. *Neural Comput. Applic* 31 (12), 9013–9022. doi:10.1007/s00521-019-04147-3
- Li, G., Xiao, F., Zhang, X., Tao, B., and Jiang, G. (2022). An Inverse Kinematics Method for Robots after Geometric Parameters Compensation. *Mech. Mach. Theory* 174, 104903. doi:10.1016/j.mechmachtheory.2022.104903
- Li, G., Zhang, L., Sun, Y., and Kong, J. (2019c). Towards the SEMG Hand: Internet of Things Sensors and Haptic Feedback Application. *Multimed. Tools Appl.* 78 (21), 29765–29782. doi:10.1007/s11042-018-6293-x
- Li, T., Wang, F., Ru, C., Jiang, Y., and Li, J. (2021). Keypoint-Based Robotic Grasp Detection Scheme in Multi-Object Scenes. *Sensors* 21 (6), 2132. doi:10.3390/s21062132
- Liao, S., Li, G., Li, J., Jiang, D., Jiang, G., Sun, Y., et al. (2020). Multi-object Intergroup Gesture Recognition Combined with Fusion Feature and KNN Algorithm. *Ijs* 38 (3), 2725–2735. doi:10.3233/jifs-179558
- Liao, S., Li, G., Wu, H., Jiang, D., Liu, Y., Yun, J., et al. (2021). Occlusion Gesture Recognition Based on Improved SSD. *Concurrency Comput. Pract. Exp.* 33 (6), e6063. doi:10.1002/cpe.6063
- Liu, X., Jiang, D., Tao, B., Jiang, G., Sun, Y., Kong, J., et al. (2022). Genetic Algorithm-Based Trajectory Optimization for Digital Twin Robots. *Front. Bioeng. Biotechnol.* 9, 793782. doi:10.3389/fbioe.2021.793782
- Liu, Y., Jiang, D., Duan, H., Sun, Y., Li, G., Tao, B., et al. (2021). Dynamic Gesture Recognition Algorithm Based on 3D Convolutional Neural Network. *Comput. Intell. Neurosci.* 2021, 1–12. doi:10.1155/2021/4828102
- Liu, Y., Jiang, D., Tao, B., Qi, J., Jiang, G., Yun, J., et al. (2022b). Grasping Posture of Humanoid Manipulator Based on Target Shape Analysis and Force Closure. *Alexandria Eng. J.* 61 (5), 3959–3969. doi:10.1016/j.aej.2021.09.017
- Liu, Y., Jiang, D., Yun, J., Sun, Y., Li, C., Jiang, G., et al. (2022d). Self-tuning Control of Manipulator Positioning Based on Fuzzy PID and PSO Algorithm. *Front. Bioeng. Biotechnol.* 9, 817723. doi:10.3389/fbioe.2021.817723
- Liu, Y., Li, C., Jiang, D., Chen, B., Sun, N., Cao, Y., et al. (2022c). Wrist Angle Prediction under Different Loads Based on GA-ELM Neural Network and Surface Electromyography. *Concurrency Comput. Pract. Exp.* 34 (3), e6574. doi:10.1002/cpe.6574
- Liu, Y., Xiao, F., Tong, X., Xu, M., Jiang, G., Chen, B., et al. (2022a). Manipulator Trajectory Planning Based on Work Subspace Division. *Concurrency Comput. Pract. Exp.* 34 (5), e6710. doi:10.1002/cpe.6710
- Luo, B., Sun, Y., Li, G., Chen, D., and Ju, Z. (2020). Decomposition Algorithm for Depth Image of Human Health Posture Based on Brain Health. *Neural Comput. Applic* 32 (10), 6327–6342. doi:10.1007/s00521-019-04141-9
- Ma, R., Zhang, L., Li, G., Jiang, D., Xu, S., and Chen, D. (2020). Grasping Force Prediction Based on sEMG Signals. *Alexandria Eng. J.* 59 (3), 1135–1147. doi:10.1016/j.aej.2020.01.007
- Naik, G. R., Kumar, D. K., and Palaniswami, M. (2014). Signal Processing Evaluation of Myoelectric Sensor Placement in Low-Level Gestures: Sensitivity Analysis Using Independent Component Analysis. *Expert Syst.* 31 (1), 91–99. doi:10.1111/exsy.12008
- Pinto, L., and Gupta, A. (2016). “Supersizing Self-Supervision: Learning to Grasp from 50k Tries and 700 Robot Hours,” in 2016 IEEE International Conference on Robotics and Automation, Piscataway, USA (IEEE), 3406–3413.
- Qi, J., Jiang, G., Li, G., Sun, Y., and Tao, B. (2020). Surface EMG Hand Gesture Recognition System Based on PCA and GRNN. *Neural Comput. Applic* 32 (10), 6343–6351. doi:10.1007/s00521-019-04142-8
- Quivira, F., Koike-Akino, T., Wang, Y., and Erdogmus, D. (2018). “Translating sEMG Signals to Continuous Hand Poses Using Recurrent Neural Networks,” in 2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI) (IEEE), 166–169.
- Rogowski, A., Bieliszczuk, K., and Rapcewicz, J. (2020). Integration of Industrially-Oriented Human-Robot Speech Communication and Vision-Based Object Recognition. *Sensors (Basel)* 20 (24), e7287. doi:10.3390/s20247287
- Scheme, E., and Englehart, K. (2011). Electromyogram Pattern Recognition for Control of Powered Upper-Limb Prostheses: State of the Art and Challenges for Clinical Use. *J. Rehabil. Res. Dev.* 48 (6), 643–659. doi:10.1682/jrrd.2010.09.0177
- Scott, R. N. (1989). “Biomedical Engineering in Upper-Extremity Prosthetics,” in *Comprehensive Management of the Upper-Limb Amputee* (New York, NY: Springer), 173–189. doi:10.1007/978-1-4612-3530-9_16
- Shi, J., Cai, Y., Zhu, J., Zhong, J., and Wang, F. (2013). SEMG-based Hand Motion Recognition Using Cumulative Residual Entropy and Extreme Learning Machine. *Med. Biol. Eng. Comput.* 51 (4), 417–427. doi:10.1007/s11517-012-1010-9
- Shi, K., Huang, L., Jiang, D., Sun, Y., Tong, X., Xie, Y., et al. (2022). Path Planning Optimization of Intelligent Vehicle Based on Improved Genetic and Ant Colony Fusion Algorithm. *Front. Bioeng. Biotechnol.* 10, 905983. doi:10.3389/fbioe.2022.905983
- Sun, Y., Hu, J., Jiang, G., Bai, D., Liu, X., Cao, Y., et al. (2022b). Multi-objective Location and Mapping Based on Deep Learning and Visual Slam. *Front. Bioeng. Biotechnol.* 2022 (10), 903261. doi:10.3389/fbioe.2022.903261
- Sun, Y., Hu, J., Li, G., Jiang, G., Xiong, H., Tao, B., et al. (2020b). Gear Reducer Optimal Design Based on Computer Multimedia Simulation. *J. Supercomput.* 76 (6), 4132–4148. doi:10.1007/s11227-018-2255-3
- Sun, Y., Huang, P., Cao, Y., Jiang, G., Yuan, Z., Bai, D., et al. (2022a). Multi-objective Optimization Design of Ladle Refractory Lining Based on Genetic Algorithm. *Front. Bioeng. Biotechnol.* 10, 900655. doi:10.3389/fbioe.2022.900655
- Sun, Y., Weng, Y., Luo, B., Li, G., Tao, B., Jiang, D., et al. (2020a). Gesture Recognition Algorithm Based on Multi-scale Feature Fusion in RGB-D Images. *IET image process* 14 (15), 3662–3668. doi:10.1049/iet-ipr.2020.0148
- Sun, Y., Xu, C., Li, G., Xu, W., Kong, J., Jiang, D., et al. (2020b). Intelligent Human Computer Interaction Based on Non Redundant EMG Signal. *Alexandria Eng. J.* 59 (3), 1149–1157. doi:10.1016/j.aej.2020.01.015
- Sun, Y., Yang, Z., Tao, B., Jiang, G., Hao, Z., and Chen, B. (2021). Multiscale Generative Adversarial Network for Real-world Super-resolution. *Concurr. Comput. Pract. Exper* 33 (21), e6430. doi:10.1002/cpe.6430
- Sun, Y., Zhao, Z., Jiang, D., Tong, X., Tao, B., Jiang, G., et al. (2022c). Low-illumination Image Enhancement Algorithm Based on Improved Multi-Scale Retinex and ABC Algorithm Optimization. *Front. Bioeng. Biotechnol.* 10, doi:10.3389/fbioe.2022.865820
- Tan, C., Sun, Y., Li, G., Jiang, G., Chen, D., and Liu, H. (2020). Research on Gesture Recognition of Smart Data Fusion Features in the IoT. *Neural Comput. Applic* 32 (22), 16917–16929. doi:10.1007/s00521-019-04023-0
- Tao, B., Liu, Y., Huang, L., Chen, G., and Chen, B. (2022). 3D Reconstruction Based on Photoelastic Fringes. *Concurrency Comput. Pract. Exp.* 34 (1), e6481. doi:10.1002/cpe.6481
- Tao, B., Wang, Y., Qian, X., Qian, X., Tong, X., He, F., et al. (2022). Photoelastic Stress Field Recovery Using Deep Convolutional Neural Network. *Front. Bioeng. Biotechnol.* 10, doi:10.3389/fbioe.2022.818112
- Tian, J., Cheng, W., Sun, Y., Li, G., Jiang, D., Jiang, G., et al. (2020). Gesture Recognition Based on Multilevel Multimodal Feature Fusion. *Ijs* 38 (3), 2539–2550. doi:10.3233/jifs-179541
- Tran, L. V., and Lin, H.-Y. (2021). BiLUNetCP: A Deep Neural Network for Object Semantic Segmentation and 6D Pose Recognition. *IEEE Sensors J.* 21 (10), 11748–11757. doi:10.1109/jsen.2020.3035632
- Tsinganos, P., Cornelis, B., Cornelis, J., Jansen, B., and Skodras, A. (2018). “Deep Learning in EMG-Based Gesture Recognition,” in 5th International Conference on Physiological Computing Systems, Seville, Spain, 107–114.
- Wang, F., Jiang, M., Qian, C., Yang, S., Li, C., Zhang, H., and Tang, X. (2017). “Residual Attention Network for Image Classification,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 3156–3164. doi:10.1109/cvpr.2017.683
- Wei, W., Wong, Y., Du, Y., Hu, Y., Kankanhalli, M., and Geng, W. (2019). A Multi-Stream Convolutional Neural Network for sEMG-Based Gesture Recognition in Muscle-Computer Interface. *Pattern Recognit. Lett.* 119, 131–138. doi:10.1016/j.patrec.2017.12.005
- Weng, Y., Sun, Y., Jiang, D., Tao, B., Liu, Y., Yun, J., et al. (2021). Enhancement of Real-time Grasp Detection by Cascaded Deep Convolutional Neural Networks. *Concurr. Comput. Pract. Exper* 33 (5), e5976. doi:10.1002/cpe.5976

- Woo, S., Park, J., Lee, J.-Y., and Kweon, I. S. (2018). "Cbam: Convolutional Block Attention Module," in Proceedings of the European conference on computer vision (ECCV), 3–19. doi:10.1007/978-3-030-01234-2_1
- Wu, X., Jiang, D., Yun, J., Liu, X., Sun, Y., Tao, B., et al. (2022). Attitude Stabilization Control of Autonomous Underwater Vehicle Based on Decoupling Algorithm and PSO-ADRC. *Front. Bioeng. Biotechnol.* 10, 843020. doi:10.3389/fbioe.2022.843020
- Xiao, F., Li, G., Jiang, D., Xie, Y., Yun, J., Liu, Y., et al. (2021). An Effective and Unified Method to Derive the Inverse Kinematics Formulas of General Six-DOF Manipulator with Simple Geometry. *Mech. Mach. Theory* 159, 104265. doi:10.1016/j.mechmachtheory.2021.104265
- Xie, B., Li, B., and Harland, A. (2018). "Movement and Gesture Recognition Using Deep Learning and Wearable-Sensor Technology," in Proceedings of the 2018 International Conference on Artificial Intelligence and Pattern Recognition, 26–31. doi:10.1145/3268866.3268890
- Xie, T., Leng, Y., Zhi, Y., Jiang, C., Tian, N., Luo, Z., and Song, R. (2020). Increased Muscle Activity Accompanying with Decreased Complexity as Spasticity Appears: High-Density EMG-Based Case Studies on Stroke Patients. *Front. Bioeng. Biotechnol.*, 1338. doi:10.3389/fbioe.2020.589321
- Yang, Z., Jiang, D., Sun, Y., Tao, B., Tong, X., Jiang, G., et al. (2021). Dynamic Gesture Recognition Using Surface EMG Signals Based on Multi-Stream Residual Network. *Front. Bioeng. Biotechnol.* 9, 779353. doi:10.3389/fbioe.2021.779353
- Yu, M., Li, G., Jiang, D., Jiang, G., Tao, B., and Chen, D. (2019). Hand Medical Monitoring System Based on Machine Learning and Optimal EMG Feature Set. *Pers. Ubiquit Comput.* doi:10.1007/s00779-019-01285-2
- Yu, M., Li, G., Jiang, D., Jiang, G., Zeng, F., Zhao, H., et al. (2020). Application of PSO-RBF Neural Network in Gesture Recognition of Continuous Surface EMG Signals. *Ijs* 38 (3), 2469–2480. doi:10.3233/jifs-179535
- Yun, J., Sun, Y., Li, C., Jiang, D., Tao, B., Li, G., et al. (2022). Self-Adjusting Force/Bit Blending Control Based on Quantitative Factor-Scale Factor Fuzzy-PID Bit Control. *Alexandria Eng. J.* 61 (6), 4389–4397. doi:10.1016/j.aej.2021.09.067
- Zardoshti-Kermani, M., Wheeler, B. C., Badie, K., and Hashemi, R. M. (1995). EMG Feature Evaluation for Movement Control of Upper Extremity Prostheses. *IEEE Trans. Rehab. Eng.* 3 (4), 324–333. doi:10.1109/86.481972
- Zhang, H., Lan, X., Bai, S., Zhou, X., Tian, Z., and Zheng, N. (2019). "ROI-based Robotic Grasp Detection for Object Overlapping Scenes," in 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Macau, China, 4768–4775.
- Zhang, X., Xiao, F., Tong, X., Yun, J., Liu, Y., Sun, Y., et al. (2022). Time Optimal Trajectory Planning Based on Improved Sparrow Search Algorithm. *Front. Bioeng. Biotechnol.* 10. doi:10.3389/fbioe.2022.852408
- Zhao, G., Jiang, D., Liu, X., Tong, X., Sun, Y., Tao, B., et al. (2022). A Tandem Robotic Arm Inverse Kinematic Solution Based on an Improved Particle Swarm Algorithm. *Front. Bioeng. Biotechnol.* doi:10.3389/fbioe.2022.832829
- Zhou, B., Ji, X. L., Zhang, R., Zhang, H., and He, A. (2010). Design of Surface EMG Acquisition Instrument Based on Embedded Technology [J]. *Mod. Electron. Tech.* 33 (6), 72–75.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Wang, Huang, Jiang, Sun, Jiang, Li, Zou, Fan, Xie, Xiong and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.