



# Dynamic Gesture Recognition Using Surface EMG Signals Based on Multi-Stream Residual Network

Zhiwen Yang<sup>1,2</sup>, Du Jiang<sup>1,3,4\*</sup>, Ying Sun<sup>1,3,4\*</sup>, Bo Tao<sup>1,3,4\*</sup>, Xiliang Tong<sup>2,4</sup>, Guozhang Jiang<sup>2,4</sup>, Manman Xu<sup>1,2,3</sup>, Juntong Yun<sup>2,4</sup>, Ying Liu<sup>2,4</sup>, Baojia Chen<sup>5</sup> and Jianyi Kong<sup>2,3,4</sup>

<sup>1</sup>Key Laboratory of Metallurgical Equipment and Control Technology of Ministry of Education, Wuhan University of Science and Technology, Wuhan, China, <sup>2</sup>Research Center for Biomimetic Robot and Intelligent Measurement and Control, Wuhan University of Science and Technology, Wuhan, China, <sup>3</sup>Hubei Key Laboratory of Mechanical Transmission and Manufacturing Engineering, Wuhan University of Science and Technology, Wuhan, China, <sup>4</sup>Institute of Precision Manufacturing, Wuhan University of Science and Technology, Wuhan, China, <sup>5</sup>Hubei Key Laboratory of Hydroelectric Machinery Design and Maintenance, Three Gorges University, Yichang, China

## OPEN ACCESS

### Edited by:

Tinggui Chen,  
Zhejiang Gongshang University, China

### Reviewed by:

Jianjun Yang,  
University of North Georgia,  
United States  
Dongxu Gao,  
University of Portsmouth,  
United Kingdom

### \*Correspondence:

Du Jiang  
jiangdu@wust.edu.cn  
Ying Sun  
sunying65@wust.edu.cn  
Bo Tao  
taoboq@wust.edu.cn

### Specialty section:

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

**Received:** 18 September 2021

**Accepted:** 04 October 2021

**Published:** 22 October 2021

### Citation:

Yang Z, Jiang D, Sun Y, Tao B, Tong X, Jiang G, Xu M, Yun J, Liu Y, Chen B and Kong J (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

Gesture recognition technology is widely used in the flexible and precise control of manipulators in the assisted medical field. Our MResLSTM algorithm can effectively perform dynamic gesture recognition. The result of surface EMG signal decoding is applied to the controller, which can improve the fluency of artificial hand control. Much current gesture recognition research using sEMG has focused on static gestures. In addition, the accuracy of recognition depends on the extraction and selection of features. However, Static gesture research cannot meet the requirements of natural human-computer interaction and dexterous control of manipulators. Therefore, a multi-stream residual network (MResLSTM) is proposed for dynamic hand movement recognition. This study aims to improve the accuracy and stability of dynamic gesture recognition. Simultaneously, it can also advance the research on the smooth control of the Manipulator. We combine the residual model and the convolutional short-term memory model into a unified framework. The architecture extracts spatiotemporal features from two aspects: global and deep, and combines feature fusion to retain essential information. The strategy of pointwise group convolution and channel shuffle is used to reduce the number of network calculations. A dataset is constructed containing six dynamic gestures for model training. The experimental results show that on the same recognition model, the gesture recognition effect of fusion of sEMG signal and acceleration signal is better than that of only using sEMG signal. The proposed approach obtains competitive performance on our dataset with the recognition accuracies of 93.52%, achieving state-of-the-art performance with 89.65% precision on the Ninapro DB1 dataset. Our bionic calculation method is applied to the controller, which can realize the continuity of human-computer interaction and the flexibility of manipulator control.

**Keywords:** dynamic gesture recognition, sEMG, MResLSTM, signal fusion, deep neural network

## INTRODUCTION

The deep neural network is an intelligent heuristic algorithm used to solve complex real-world problems (He and Jiang, 2020). For example, deep learning is used for data mining to analyze user needs (Chen et al., 2021). The main purpose of the research on dynamic gesture recognition is to promote the development of dynamic human-computer interaction. The dynamic gesture recognition model is applied to the controller of the manipulator, which can improve the continuity and flexibility of the manipulator control. The surface electromyography signal (sEMG) contains a lot of information and can be used for gesture recognition and force prediction (Ma et al., 2020; Atzori et al., 2016; Sadikoglu et al., 2017; Baldacchino et al., 2018). Therefore, it is convenient and feasible to use it as an information interaction medium for human-computer interaction (Sun et al., 2020a; Hu et al., 2019; Jiang et al., 2019a; Shahzad et al., 2019). In biomedical signals, sEMG signals are widely accepted and decoded due to their neural basis and ease of use, so gesture recognition based on sEMG has become a research hotspot in manipulators and human-computer interaction (Xiao et al., 2021; Ahn et al., 2020; Gowtham et al., 2020). Many studies have found that sEMG-based deep learning approaches have great potential in gesture recognition. The gesture recognition model is applied to the controller of the Manipulator to control its actions (Rodríguez-Tapia et al., 2020). The control flow is shown in **Figure 1**.

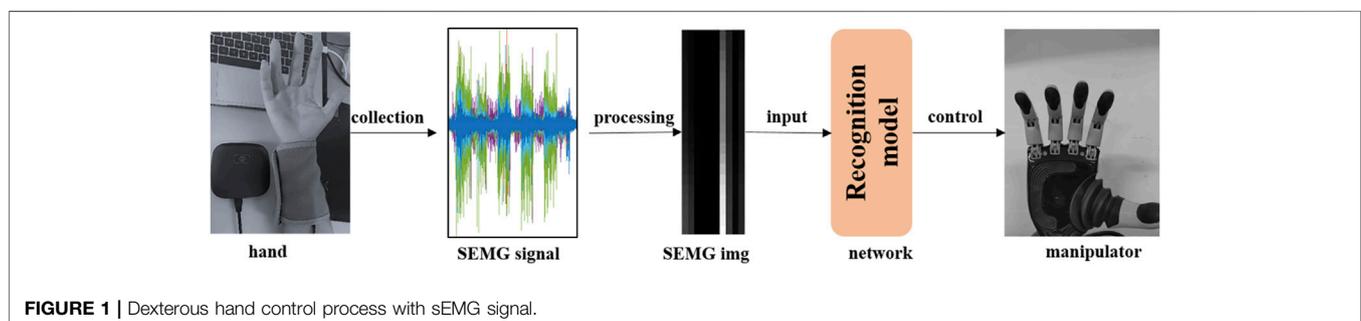
Surface sEMG signals represent a promising method to decode the movement intentions of amputees and control multifunctional dexterous hands in a non-invasive manner. The focus of sEMG signal research was to develop pattern recognition and classification techniques for detecting different hand movements. Therefore, many technologies, including fuzzy systems, neural networks, fuzzy support vector machines (SVM), hidden Markov models (HMM), and principal component analysis (PCA), have shown the high accuracy of hand motion recognition (Mendes Junior et al., 2020; Sun et al., 2020b; Cheng et al., 2021; Liao et al., 2021). Secondly, it is mainly from designing better features to improve the accuracy of the gesture recognition network. Then the process of feature extraction and selection is complicated. Third, different feature combinations have other recognition effects on the same model (Duan et al., 2021; Yu et al., 2019; Jiang et al., 2021a). However, deep learning can automatically

learn the characteristics of sEMG and avoid the disadvantages of manually extracting the features. Unlike vision-based gesture recognition methods, sEMG-based gesture recognition is not affected by the surrounding environment, such as background lighting and occlusion (Jiang et al., 2019b; Tian et al., 2020; Mujahid et al., 2021). However, different arm positions, electrode displacements, signal non-stationarity, and force changes greatly affect the accuracy and robustness of the sEMG-based recognition model. Finally, only relying on sEMG for gesture recognition cannot fully characterize the features of gestures in motion, making it difficult for the recognition model to converge during training. Therefore, Signal fusion technology is adopted to improve the accuracy and robustness of the network (Xu Zhang et al., 2011; Sun et al., 2018; Tan et al., 2020).

Dynamic gestures are a set of continuous motion gestures to represent a specific meaning, generally including hand movements and arm movements. In the paper, deep learning methods are used to analyze dynamic hand movements. The residual model and variant ConvLSTM model combined into a multi-stream network. For a multi-stream network, each stream independently learns representative features by ResNet. Then, it fuses the features learned from all streams into a unified feature map. Simultaneously, a dual-stream classifier fused with sEMG and ACC signals is used to recognize various dynamic actions to improve the accuracy of behavioral action recognition. The proposed MResLSTM can directly input the preprocessed EMG signal into the network for dynamic gesture recognition. The contributions of this paper are as follows:

- 1) Surface EMG signals and ACC signals are collected to construct datasets containing six different dynamic gestures.
- 2) Embedding the SE unit into the Residual module can effectively solve channel dependence. At the same time, the strategy of pointwise group convolution and channel shuffle is adopted to reduce the calculation amount of the model.
- 3) The proposed MResLSTM achieves state-of-the-art results in terms of dynamic hand movement recognition.

The rest of this paper is organized as follows: *Related Work* discusses the related work, followed by the MResLSTM designed in *Method* and the optimization of the model. *Experiment* shows the experimental results and analysis, and *Conclusion* concludes the paper with a summary and future research directions.



## RELATED WORK

Surface EMG signals is a non-invasive technique for measuring the electrical activity of muscle groups on the skin surface, which makes it a simple and straightforward method that allows the user to actively control the prosthesis (Takaiwa et al., 2011; Gregory and Ren, 2019; Wu et al., 2017). The basic principle of the human-machine interface based on surface EMG signals is to convert sEMG into controllable signals through algorithms such as machine learning. With the precision, portability, and signal processing algorithm performance of the acquisition system, the high reliability of the man-machine interface and the robustness of the prosthetic hand control have become a reality. Recently, many researchers have paid more attention to deep learning in the field of EMG pattern recognition. It can automatically learn features of different abstract levels from many input samples, thereby avoiding cumbersome feature extraction and optimization processes and realizing end-to-end EMG gesture recognition (Weng et al., 2021; Su et al., 2021; Tsinganos et al., 2019; Chaiyaroj et al., 2019).

Atzori et al. (2016) proposed a LeNet-based convolutional neural network model AtzoriNet for end-to-end EMG gesture recognition. He et al. (2018) combined a Long short-term memory network and multilayer perceptrons and conducted experiments on the NinaPro DB1 dataset. When classifying the 52 hand movements of 27 subjects, the accuracy rate reached about 75%. Hu et al. (2019) proposed a CNN model based on the attention mechanism and tested it on the NinaProDB1, NinaProDB2, BioPatRec subdatabase, CapgMyo subdatabase, and csl-hdemg database. Its accuracy rates are 87.0, 82.2, 94.1, 99.7 and 94.5% respectively. Geng et al. (2016) proposed GengNet for gesture recognition based on transient EMG signals. They applied a pre-training strategy to make the EMG gesture recognition performance of the network surpassed the method of extracting signal features and inputting traditional classifier models for gesture recognition. Wu et al. (2018) proposed LSTM-CNN for the dynamic recognition of gestures. Mendes Junior et al., 2020 investigated multiple classification techniques for six hand gestures acquired from 13 participants using eight channels sEMG armband with a sampling rate of 2 kHz. Their best result, with an average accuracy of 94% was obtained from 40 features with the large margin nearest neighbor (LMNN) technique. Côté-Allard et al. (2020) presented an analysis of the features learned using deep learning to classify 11 hand gestures using sEMG. The LSTM model is used to extract timing information in signals. The CNN model can perform secondary feature extraction and signal classification (Peng et al., 2020).

As mentioned above, it is obvious that deep learning methods can overcome the limitation of feature engineering for better feature quality. Many studies have shown that the accuracy of using DNN to classify surface EMG signals is generally higher. However, EMG signal recognition based on deep learning models is expected to improve accuracy and feature extraction complexity (Jiang et al., 2019c; He et al., 2019; Sri-iesaranusorn et al., 2021).

## METHOD

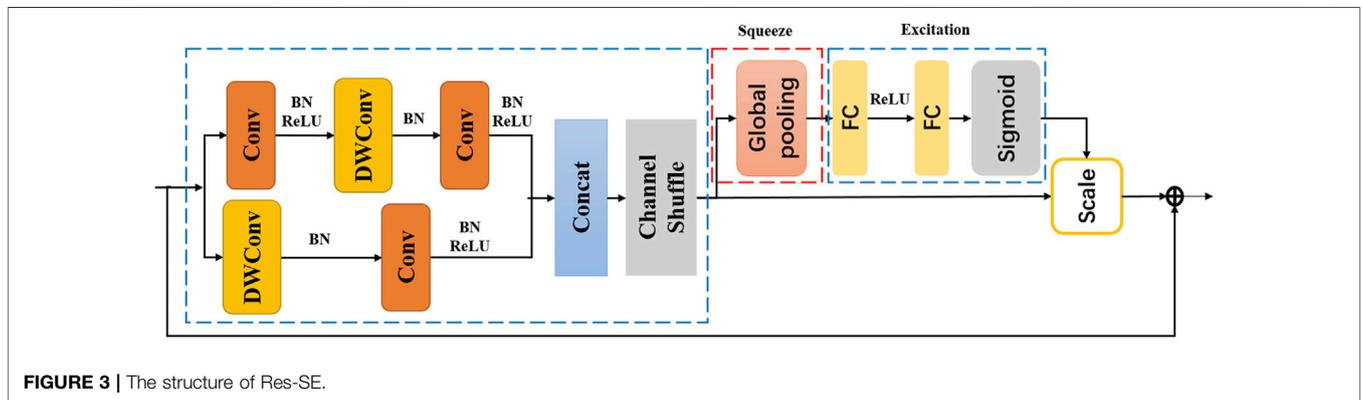
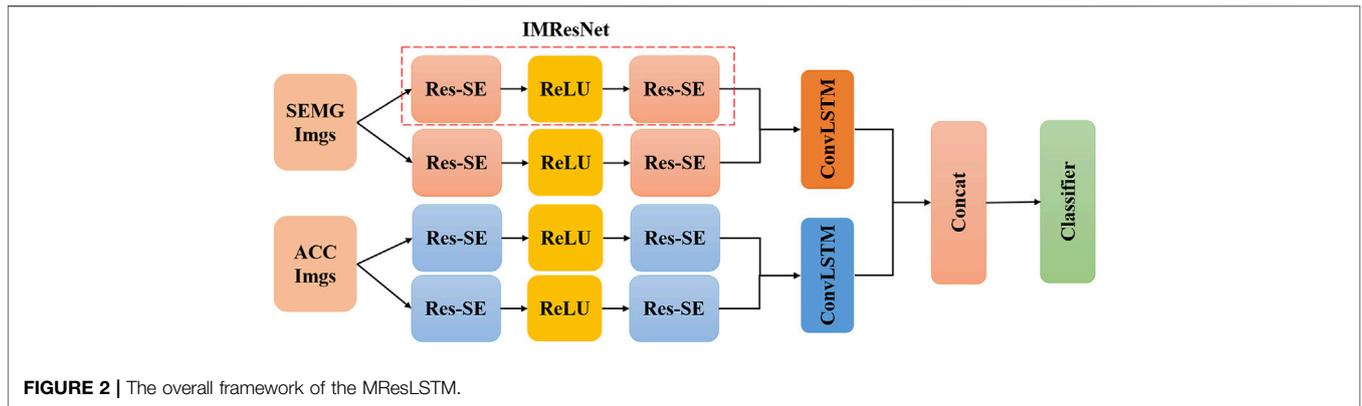
The advantage of dynamic gesture research lies in the ability to apply the trained model to the control of dexterous hands. Dynamic gestures are a set of continuous motion gestures to represent a specific meaning. The dynamic hand movement is regarded as a dynamic transfer action in which one gesture posture is converted to another (Zhang and Li, 2019; Zhang et al., 2021; Liu et al., 2021). In this paper, we formulate the sEMG-based gesture recognition problem as a DNN based image classification problem. In the context of dynamic gesture recognition, the EMG signal has a strong timing. Instantaneous sEMG images and simple classifiers may not fully capture the time information between multiple frames, so a time window is used to sample the sEMG signal, and the sEMG signal is converted to an sEMG image within the time window. In this paper, the MResLSTM is proposed for dynamic gesture recognition, and its overall framework is shown in **Figure 2**.

The model includes two stages: feature extraction and feature fusion. First, the original sEMG image is decomposed into  $n$  patches of equal size. Then these patches are input into a multi-stream network, and each stream independently learns representative features by IMResNet. During the fusion stage, it fuses the features learned from all streams into a unified feature map. The convolutional long short-term memory extracts spatiotemporal feature information from local, global and deep aspects, and combines feature fusion to alleviate the loss of feature information. Finally, the feature map is input to the classifier for classification. To prevent over-fitting, the ReLU nonlinear function is applied after each fully connected layer, batch normalization is performed, and a 50% dropout layer is added after the fully connected layer. Many studies have found that the recognition effect of information fusion technology is better than that of single information. Therefore, this paper proposes a novel dynamic gesture recognition scheme based on the information fusion of sEMG and ACC signals. The original signal is directly converted into images for training the recognition network after preprocessing.

### IMResNet

The IMResNet module is shown in **Figure 2** and consists of two Re-SE units. We embed the SE module into the residual network to form a Re-SE module, the structure of which is shown in **Figure 3**. The channel relationship of the image constructed by a convolutional neural network through convolution is local. Many researchers hope that the correlation of channels can be explicitly constructed to enhance the feature maps obtained by convolution. Squeeze-Excitation module (SE) is adopted to solve the above issue. This SE module enables the network to increase its sensitivity to signal characteristics to use these feature information in subsequent conversions. The SE module is composed of Squeeze and Excitation, as shown in **Figure 3**.

The Squeeze compresses the global information to each channel for description through global pooling, effectively



solving channel dependence. The output formula of the  $n$ th channel after global pooling is as follows:

$$z_n = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W I_n(i, j) \quad n = 1, 2, \dots, N, \quad (1)$$

Where  $I_n$  is the  $n$ th channel of the characteristic image;  $H$  and  $W$  are the height and width of the image, respectively;  $N$  is the number of channels of the picture. Global average pooling can make full use of the correlation of the channel, effectively shield the distribution information in the space, and make the calculation of the output characteristic information more accurate. After squeezing, the Excitation is used to capture the dependence of the channel fully. The Excitation is implemented with 2 fully connected layers. The full connection can use the correlation between channels to train the accurate image scale. The first fully connected layer compresses all channels  $C$  into  $C/k$  channels ( $k$  is the compression ratio). The second fully connected layer is restored to the original  $N$  channel. The purpose is to reduce the amount of calculation.

The dynamic gesture recognition has real-time requirements, so it is necessary to carry out a lightweight design to reduce network calculation. This paper adopted group convolution and channel shuffle, which greatly reduces the computational complexity of the model while maintaining accuracy. Group

convolution minimizes the amount of calculation of the network, but it causes the feature information between different groups to not be exchanged. The core design concept of ShuffleNet is to rearrange different channels to solve the drawbacks caused by grouped convolution (Zhang et al., 2019; Li et al., 2020). The channel reorganization of the feature map after the group convolution ensures that the information can flow between different groups. The IMResNet can directly input the processed EMG image and automatically extract the features of the image.

### Variants ConvLSTM

The surface EMG signal of dynamic gestures has a strong timing, so a timing network must be used to extract the timing characteristics of the signal. In this article, we improve the LSTM network structure. The LSTM unit has three thresholds: input gate  $i_t$ , forget gate  $f_t$ , and output gate  $o_t$ . The subscript  $t$  represents the time. In addition, use  $c_t$  to represent the cell state of the LSTM at time  $t$ . The LSTM network can process time-series data, but if the time series data is an image, adding a convolution operation based on LSTM will be more effective for image feature extraction. The ConvLSTM is a variant of LSTM (Peng et al.,). It not only can extract time-series features but also can describe spatial features. The structure of the LSTM cell and ConvLSTM cell is shown in Figure 4. The main change is that the weight calculation of  $W$  has become a

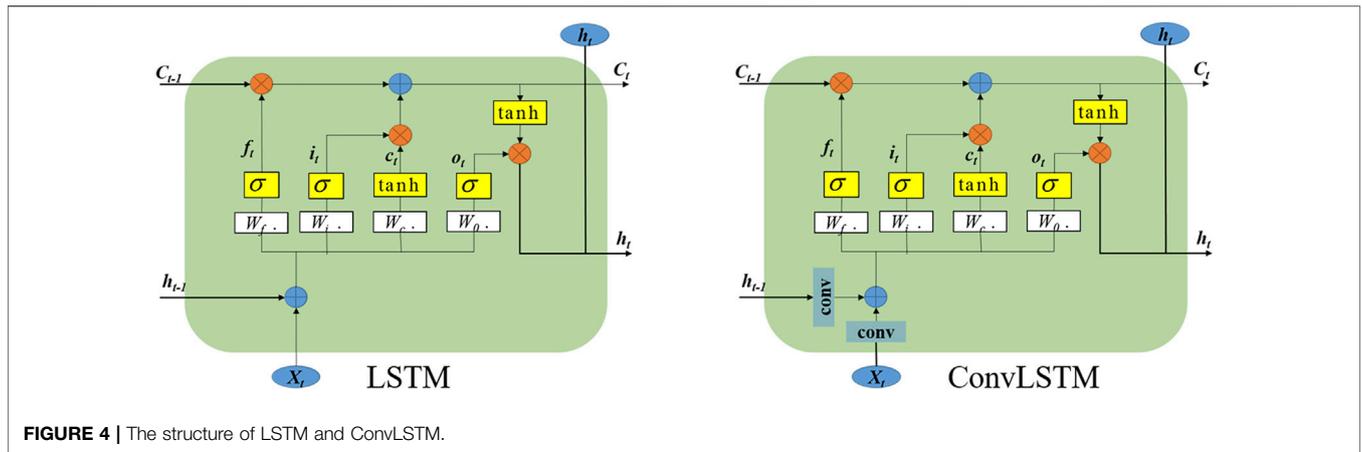


FIGURE 4 | The structure of LSTM and ConvLSTM.

convolution operation so that the characteristics of the image can be extracted.

$$\begin{aligned}
 f_t &= \sigma(W_{fh} \cdot h_{t-1} + W_{fx} \cdot x_t + b_f) \\
 i_t &= \sigma(W_{ih} \cdot h_{t-1} + W_{ix} \cdot x_t + b_i) \\
 C_t &= f_t \odot C_{t-1} + i_t \odot \tanh(W_{ch} \cdot h_{t-1} + W_{cx} \cdot x_t + b_c) \\
 o_t &= \sigma(W_{oh} \cdot h_{t-1} + W_{ox} \cdot x_t + b_o) \\
 h_t &= o_t \odot \tanh(C_t).
 \end{aligned} \tag{2}$$

Eq. 2 is the calculation formula of the LSTM unit. Where  $x_t$  is the input,  $C_t$  is the cell state,  $h_t$  is the hidden state. “ $\odot$ ” represents the Hadamard product.

$$\begin{aligned}
 f_t &= \sigma(W_{fh} * H_{t-1} + W_{fx} * X_t + b_f) \\
 i_t &= \sigma(W_{ih} * H_{t-1} + W_{ix} * X_t + b_i) \\
 C_t &= f_t \odot C_{t-1} + i_t \odot \tanh(W_{ch} * H_{t-1} + W_{cx} * X_t + b_c) \\
 o_t &= \sigma(W_{oh} * h_{t-1} + W_{ox} * x_t + b_o) \\
 h_t &= o_t \odot \tanh(C_t).
 \end{aligned} \tag{3}$$

Eq. 3 is the calculation formula of the ConvLSTM unit. Where  $X_t$  is the input,  $C_t$  is the cell state, and  $H_t$  is the hidden state. “ $*$ ” represents the convolutional operations, and “ $\odot$ ” means the Hadamard product. The ConvLSTM has a large number of parameters due to the convolution operation. In addition, the convolution in ConvLSTM has no spatial attention effect. The convolution of the three gates hardly affects the Spatio-temporal feature fusion. Therefore, reducing the convolution operation in the three gates can obtain better accuracy, fewer parameters and lower computational cost. This variant of ConvLSTM is improved on the basis of ConvSTLM, as shown in Figure 5.

The Variant ConvLSTM only retains the convolution at the input state in the ConvLSTM structure. The rest of the convolution operations are replaced by global average pooling and fully connected operations. The working principle of VConvLSTM can be expressed by:

$$\begin{aligned}
 f_t &= \sigma(W_{fh}GP(H_{t-1}) + W_{fx}GP(X_t) + b_f) \\
 i_t &= \sigma(W_{ih}GP(H_{t-1}) + W_{ix}GP(X_t) + b_i) \\
 C_t &= f_t \odot C_{t-1} + i_t \odot \tanh(W_{ch} * H_{t-1} + W_{cx} * X_t + b_c) \\
 o_t &= \sigma(W_{oh}GP(H_{t-1}) + W_{ox}GP(X_t) + b_o) \\
 h_t &= o_t \odot \tanh(C_t).
 \end{aligned} \tag{4}$$

Eq. 4 is the calculation formula of the Variant ConvLSTM unit. Where  $X_t$  is the input,  $C_t$  is the cell state, and  $H_t$  is the hidden state. “ $*$ ” represents the fully connected operations, and “ $\odot$ ” represents the Hadamard product. GP stands for global average pooling.

### Dataset Acquisition

The acquisition of sEMG and acceleration signals is the basis for realizing human hand movement recognition. In this article, a 16-channel sEMG instrument is used for signal acquisition. When the signal is collected, the installation of the equipment is shown in Figure 6.

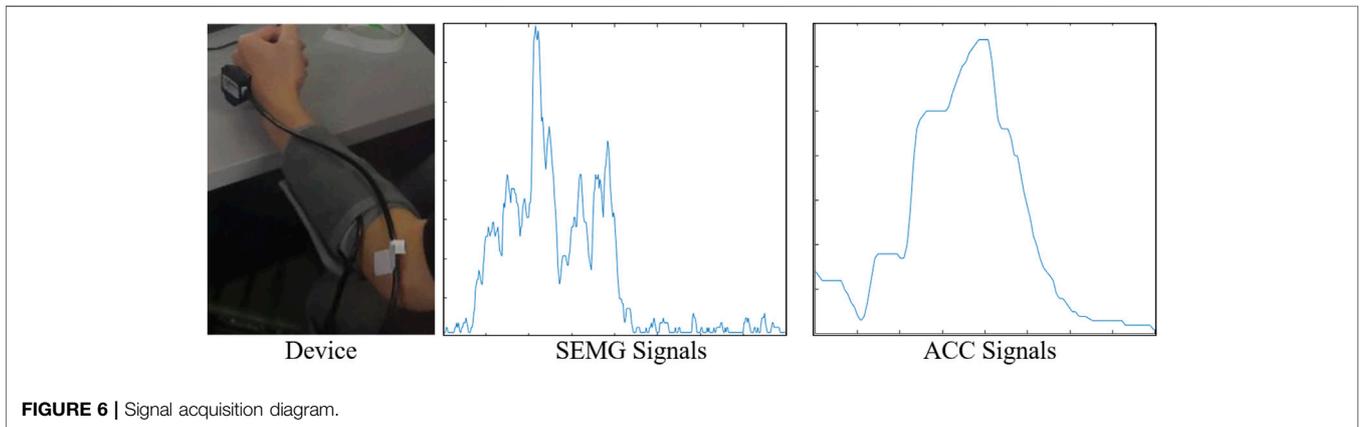
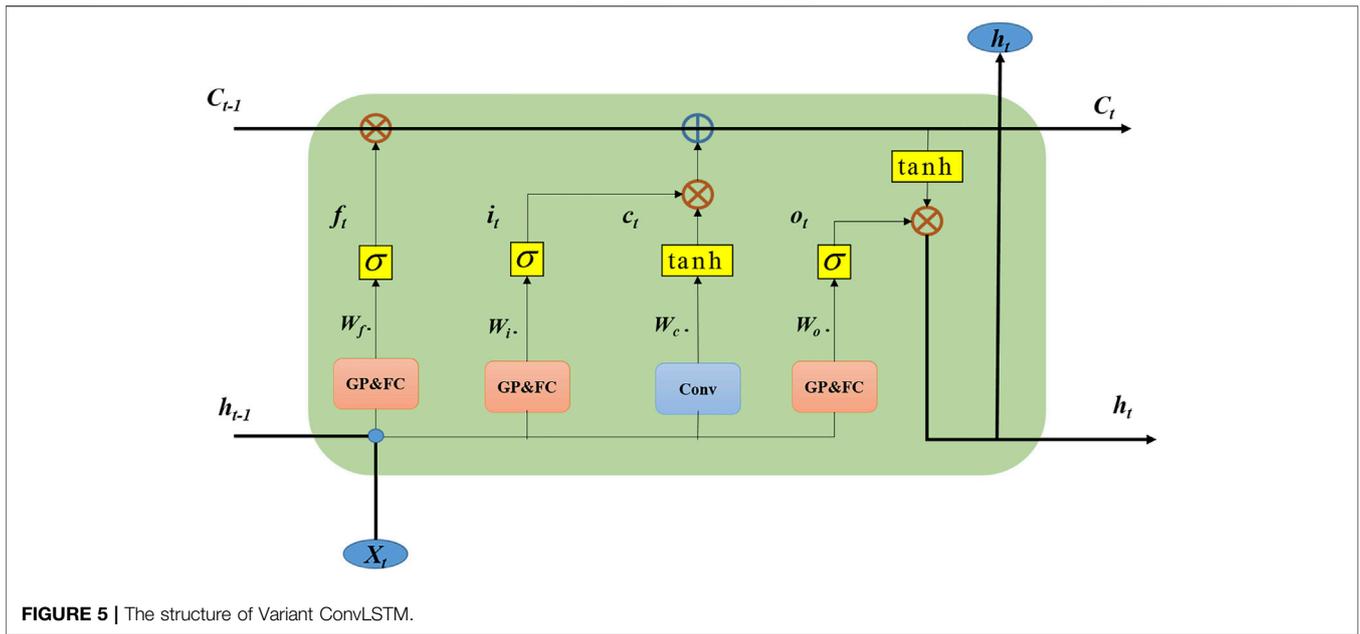
The ages of the experimenters were distributed among ten persons between 20 and 30 years old. The details of the subjects are summarized in Table 1. The electromyography cuff is worn on the left hand, and the acceleration sensor is close to the back of the hand. During the collection process, the forearm should be kept as level as possible.

The sampling frequency is set to 1000 Hz, the motion cycle of different gesture actions is set to 10 s, and a set of experiments are collected 20 times. During the experiment, taking into account the fatigue of the negative muscles, take a five-minute rest after each collection and proceed to the next set of experiments. In each experiment, the repeated method is to rest for 10 s, keep the action for 10 s, repeat twenty times, and collect for three consecutive days, using the same collection method every day. This method can be used to obtain temporal and spatial differences in myoelectric signals of the same individual. The complete paradigm is illustrated in Figure 7.

The six gestures involving the entire hand movement are shown in Figure 8, including two-finger left turn (TFTR), two-finger right turn (TFTR), flat palm flip (FPTL), flat palm left turn (FPTL), flat palm right turn (FPTR) and flat palm fist (FPMF).

### 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 500 sets for each gesture, and each test set contains 60 sets.



**TABLE 1 |** Demographic information the subjects.

Subjects	Hand	Status	Age	Sex
S0	Left	Healthy	25	M
S1	Left	Healthy	25	M
S2	Left	Healthy	26	M
S3	Left	Healthy	24	W
S4	Left	Healthy	26	M
S5	Left	Healthy	26	M
S6	Left	Healthy	23	M
S7	Left	Healthy	28	M
S8	Left	Healthy	26	M
S9	Left	Healthy	25	M

implemented by PyTorch 1.7.0 + cu110 on NVIDIA GTX 1080Ti GPU.

### Pretreatment

The process of sEMG signals collection is continuous, and the sEMG includes active segment signals and inactive segment information. To improve the accuracy and speed of the recognition model, it is necessary to eliminate non-active segment information. Research shows that the threshold method can efficiently extract active segments. The active segment detection formula is as follows:

$$S(n) = \sum_{c=1}^c (SEMG_c(n) - SEMG_c^{mean})^2 \geq TH, n = 1, 2, 3, \dots, n, \quad (5)$$

Experimental environment hardware: Intel(R) Core(TM) i5-10210U CPU@1.60 GHz; memory: 8.00 GB; system type: 64-bit operating system, x64-based processor. All experiments are

Where  $c$  is the number of acquisition channels of sEMG;  $N$  is the number of sampling points;  $SEMG_c(n)$  is the value of the  $n$ th sampling point of the  $c$  channel;  $SEMG_c^{mean}$  is the average value

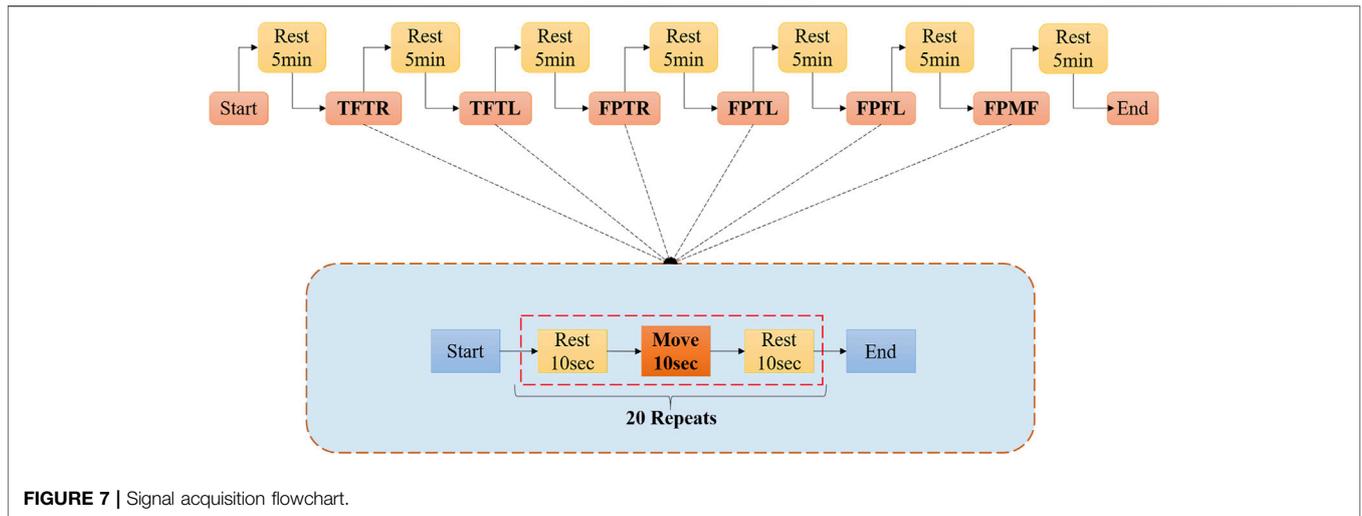


FIGURE 7 | Signal acquisition flowchart.

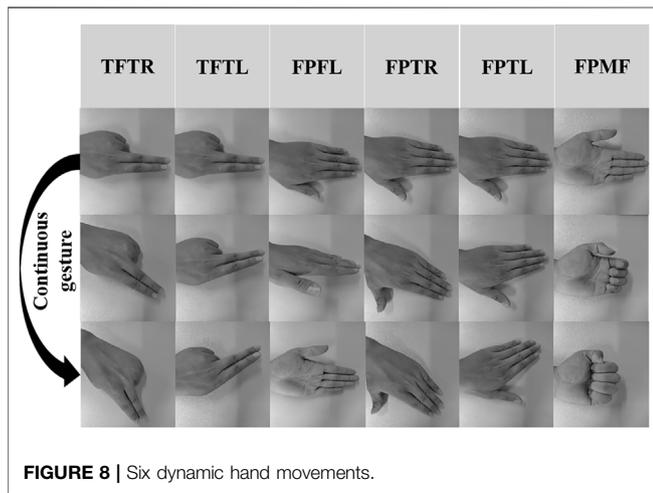


FIGURE 8 | Six dynamic hand movements.

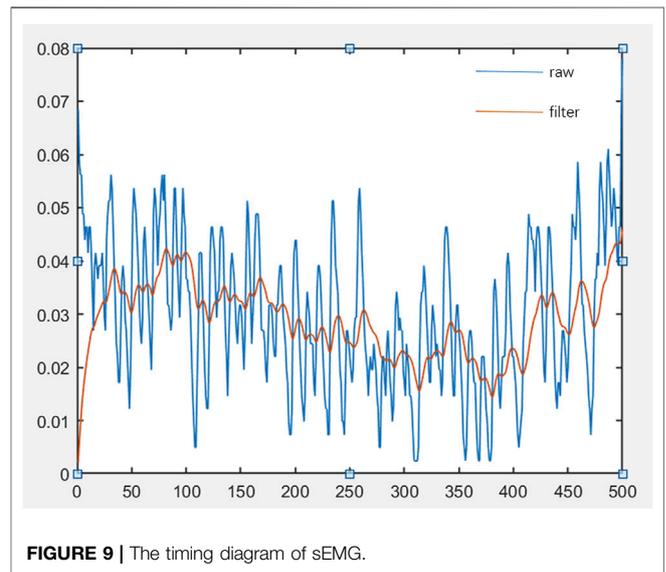


FIGURE 9 | The timing diagram of sEMG.

of the sEMG when the  $c$  channel is relaxed;  $TH$  is the set threshold. In this article,  $TH$  is 15% of the peak energy of each channel.

The raw sEMG contains a lot of noise, and the signal needs to be filtered and noise-reduced. The frequency of the power frequency noise in the environment is mainly concentrated at 50 Hz or the corresponding integer multiple of the frequency. A 20-order comb filter is used to filter it. Wavelet transform can highlight the signal characteristics in the time domain and frequency domain. Wavelet transform is to shift the basic wavelet function and then perform inner product with the signal that needs noise reduction at different scales. The wavelet transform is to shift the basic wavelet function, and then at different scales, the inner product with the signal that needs noise reduction, namely:

$$WT_x(\alpha, \tau) = \frac{1}{\sqrt{\alpha}} \int_{-\infty}^{+\infty} x(t) \varphi^*\left(\frac{t-\tau}{\alpha}\right) dt, \quad (6)$$

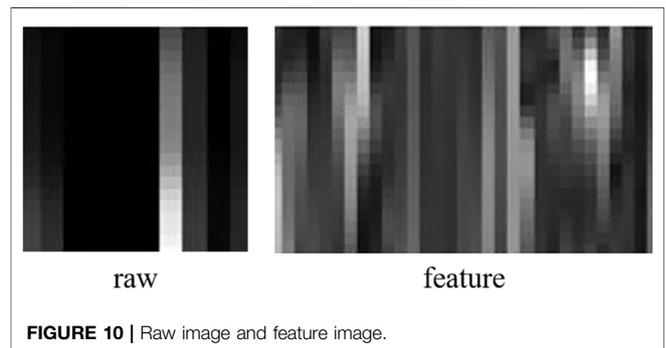


FIGURE 10 | Raw image and feature image.

Where  $\alpha > 0$ , is the scale factor, and its function is to expand and contract the basic wavelet  $\varphi(t)$  function and  $\tau$  represents the displacement. In this paper,  $\text{coif5}$  is used as the wavelet basis function, and the unbiased likelihood estimation threshold is

**TABLE 2 |** results of different networks.

Network model	Accuracy (%)
Single-stream network	73.21
Multi-stream network	84.35

used for threshold processing and the hard threshold function to process noise signals. The effect after sEMG treatment is shown in **Figure 9**.

The proposed recognition network compares the recognition effect of the original EMG image and the multi-EMG feature image as the input source. The raw image and feature image are shown in **Figure 10**. The dynamic recursive feature selection algorithm is used to calculate the correlation between each EMG feature and the target using mutual information. The EMG feature that is least relevant to the target is eliminated, and the optimal feature is selected.

This paper selects four characteristics: average absolute value (MAV), signal high and low-frequency ratio (FR), median frequency (MDF), and power spectrum average power (MNP) to construct a featured image. The calculation formulas for the four characteristics are as follows:

$$\begin{aligned}
 MAV &= \frac{1}{K} \sum_{i=1}^K |x_i|, FR = \frac{\sum_{i=LHC}^{LHC} P_i}{\sum_{i=HLC}^{HLC} P_i}, \sum_{i=1}^{MDF} P_i = \sum_{i=MDF}^M P_i \\
 &= \frac{1}{2} \sum_{i=1}^M P_i, MNP = \frac{1}{M} \sum_{i=1}^M P_i,
 \end{aligned}
 \tag{7}$$

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

**TABLE 3 |** results of information fusion.

Information fusion	Accuracy (%)
ACC	90.71
----	84.35

**TABLE 4 |** results of the different input sources.

Input source	Accuracy (%)
Raw img	90.71
Feature img	93.52

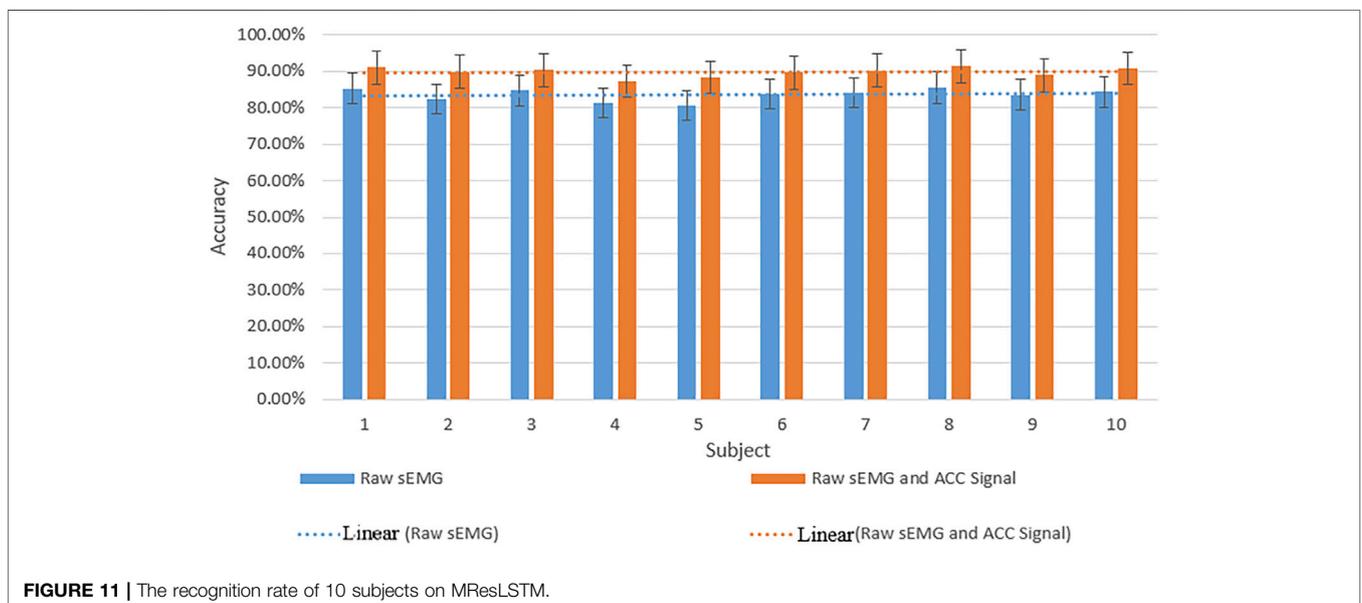
respectively; HLC and HHC are the lower and upper cut-off frequencies of the high-frequency band, respectively.

### Experimental Results and Analysis

The calculation amount of a multi-stream network is larger than that of a single network. Therefore, it is necessary to construct a comparative experiment between a multi-stream network and a single network. In the comparison experiment, the input of both recognition models is all the original EMG images. At the same time, no ACC information fusion is added. In addition, the input matrix format of a single network model is different, and the input data format needs to be fine-tuned.

The experimental results are shown in **Table 2**. It can be seen from **Table 2** that the gesture recognition effect of the multi-stream network is better than that of the single network. The multi-stream network can extract more key features and prevent the gradient from disappearing.

Information fusion increases the workload of data collection, improves the complexity of the network, and reduces the identification efficiency of the network. Therefore, to verify the



**FIGURE 11 |** The recognition rate of 10 subjects on MResLSTM.

effectiveness of the fusion acceleration signal, a corresponding comparative experiment was carried out. In the experiments, the acceleration (ACC) signal is input into the network as an independent branch, the raw sEMG image is the input source of the network, and other conditions remain unchanged.

The comparison results are shown in **Table 3**. Only using SEMG for dynamic gesture recognition, its recognition effect is not as good as information fusion on the same model. The characteristic signals of a variety of signals are not entirely the same. Combining them may produce complementary information. These complementary features can improve the

recognition accuracy of the network. However, sometimes information fusion can also lead to information redundancy.

To judge the effectiveness of feature extraction, the feature image and the original EMG image are used as the input source of the network to conduct a comparative experiment. During the experiment, both networks added ACC signals. The difference is the input source of the network.

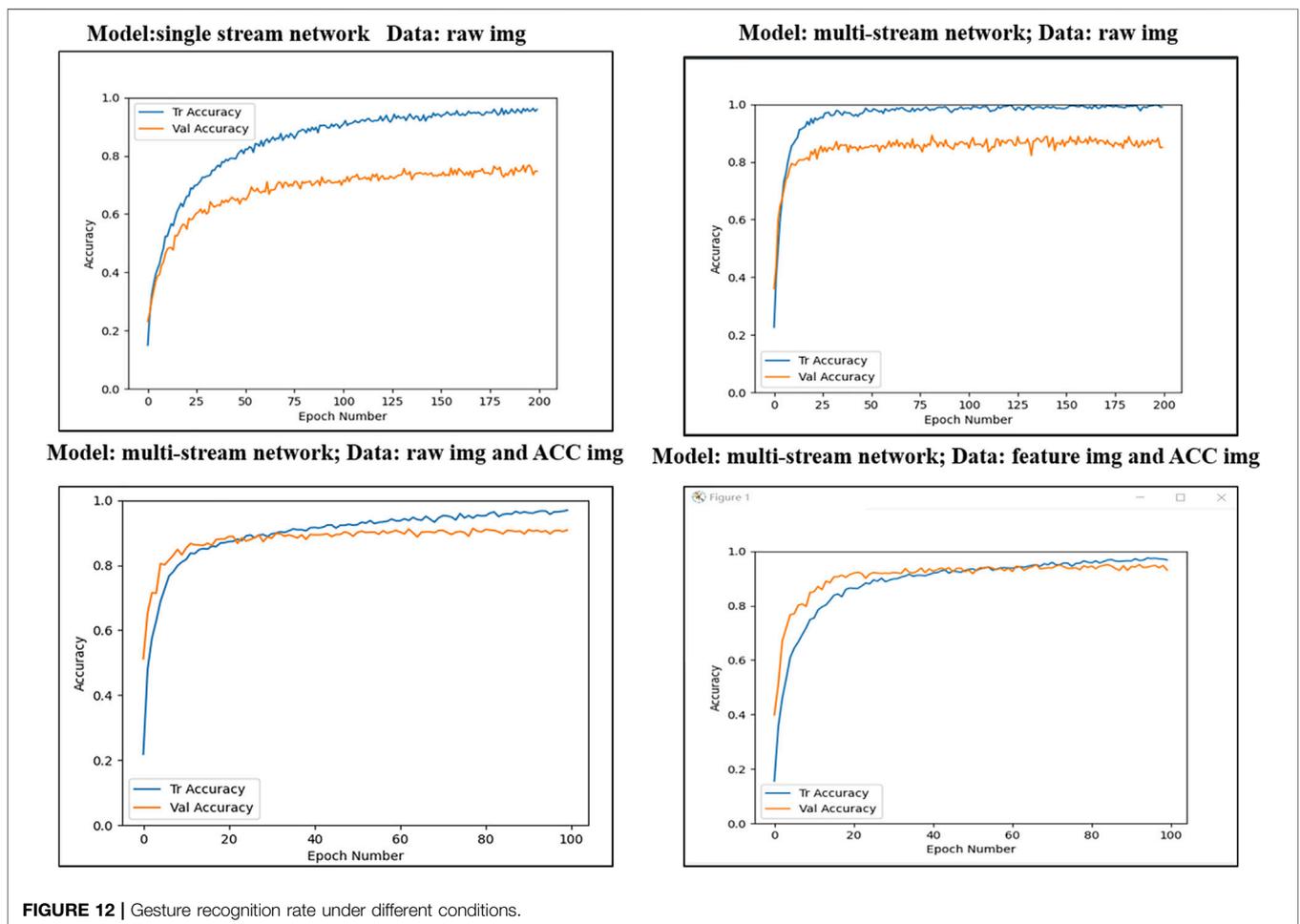
The experimental results are shown in **Table 4**. The recognition effect of the input feature image is better than the original EMG image. The featured image effectively retains the critical information, which significantly improves the recognition accuracy of the multi-stream network.

The average recognition rate of the proposed MResLSTM is 93.52%. However, it can be seen from **Figure 11** that the recognition effect of the model is affected by individual differences. Experimental results show that the recognition rate difference between subjects is about 8%. The reason may be that the position of the acquisition instrument has changed or that the hand movement is fast or slow during the signal acquisition process.

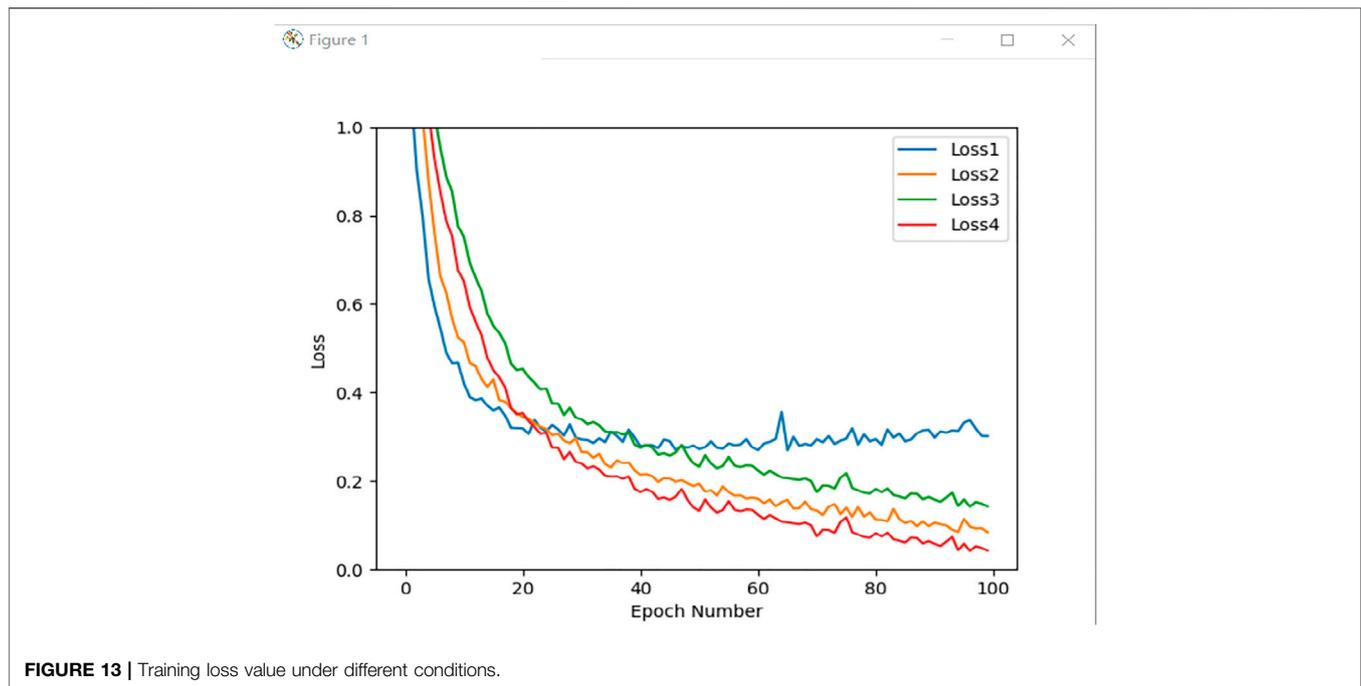
To show the advantages of our model, more comparisons with other neural networks should be added, so it is necessary to conduct an experiment on the public dataset Ninapro DB1. The

**TABLE 5** | Comparison results of different approaches on NinaPro DB1.

Algorithms	Accuracy (%)
Atzori_Net <sup>[Atzori et al.2016]</sup>	66.73
Geng_Net <sup>[Geng et al.2016]</sup>	77.80
Gene_ELM <sup>[Cene et al.2019]</sup>	75.11
Yu_CNN <sup>[Yu et al.2019]</sup>	79.50
GoogleNet	81.27
Wei_MSCNN <sup>[Wei et al.2019]</sup>	85.00
MResLSTM(our)	89.31



**FIGURE 12** | Gesture recognition rate under different conditions.



**FIGURE 13** | Training loss value under different conditions.

NinaPro DB1 dataset contains 52 different gestures of 27 healthy subjects, different from the gestures contained in the data set used in this article. It is necessary to fine-tune the model's classifier to enable it to perform 52 classifications. The experimental results are shown in **Table 5**. Experimental results show that our proposed multi-stream network is better than other algorithms.

Through the comparison of various recognition algorithms in **Table 5**, it can be seen that the recognition rate of the MResLSTM on the public dataset is 89.31%, which is 4 percentage points higher than MSCNN. It is not difficult to see from the comparative experimental results that with the further development of deep learning in EMG gesture recognition in recent years, the advantages of deep convolutional neural networks in the research of EMG pattern recognition have become more and more apparent. Among them, the average gesture recognition rate based on multi-stream CNN proposed by Wei reached 85.00%. The network is divided into a multi-stream decomposition stage and a fusion stage. In the multi-stream decomposition stage, each stream independently learns representative features through CNN. Then in the fusion stage, it merges the features learned from all streams into a unified feature map and then inputs it into the fusion network to recognize gestures. The experimental results show that the multi-stream network can make up for the single input data information and retain richer features.

When the following four experiments are performed, the data batch size is 128, and Epoch is 200. The four experiments are as follows: Experiment 1: the recognition model is a single network structure, and the input source is the original EMG image. Experiment 2: the recognition model is a multi-stream network structure, and the input source is the raw EMG image. Experiment 3: The recognition model is a multi-stream

network structure, and the input source is the original EMG and ACC signal image. Experiment 4: The recognition model is a multi-stream network structure, and the input sources are characteristic EMG images and ACC signal images. The training accuracy and verification accuracy during network training are shown in **Figure 12**.

Comparing Experiment 1 and Experiment 2, it can be seen that the multi-stream network converges faster during training, and the network is more robust. Secondly, the multi-stream network can retain more key features and improve the recognition accuracy of the network. Comparing Experiment 2 and Experiment 3, we can find that Signals fusion can effectively compensate for the shortcomings of single information, making the learned features richer. Comparing Experiment 3 and Experiment 4, we can see that the overall recognition rate of the original EMG image as the input of the network model is low. This is because only limited abstract features can be extracted from the original EMG image through convolution operation.

**Figure 13** is the training loss graph of four different experiments. Loss1 represents the loss function of Experiment 1, and Loss2 indicates the loss function of Experiment 2. Loss3 means the training loss of Experiment 3, and Loss4 represents the data input is the loss of Experiment 4.

It is not difficult to see from **Figure 13** that the network is challenging to converge when a single network is trained with the original sEMG as the input source. This is because a single network has limited features extracted from the sEMG, which is prone to overfitting. The multi-stream network can retain more effective information, making the accuracy and stability of gesture recognition better. Multi-stream networks have better generalization capabilities.

## CONCLUSION

The motivation of research on dynamic gesture recognition based on sEMG signals is that it can promote the flexible control of manipulators. In this paper, the MResLSTM is proposed for dynamic gestures recognition. The problem of gesture recognition research based on EMG signal is that the amount of data is relatively small and easy to overfit. A multi-stream network structure can retain more crucial information to solve the issue. The strategy of pointwise convolution and channel shuffle is adopted to achieve the real-time requirements of the recognition model. This article uses feature correlation to select key features. The recognition rate of the MResLSTM on the feature image is 93.52%, and the accuracy on the original EMG image is 90.71%. Experimental results show that decent feature images can improve the recognition accuracy of the network. The comparative experiment results on the dataset Ninapro DB1 show our proposed model outperforms the state-of-the-art methods.

SEMG signals are one of the most widely used biological signals to predict the movement intention of the upper limbs. Converting sEMG signals into effective control signals often requires a lot of computing power and complicated processes. The high variability of sEMG and the lack of existing data limit the application of gesture recognition technology (Li et al., 2021; Aranceta-Garza and Conway, 2019). In future work, high-density sEMG (Chen et al., 2020) and multiple information fusion will be the direction of dynamic gesture recognition research. Secondly, the influence of the speed and cycle of hand actions on the model will be a meaningful direction.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES

- Ahn, B., Ko, S. Y., and Yang, G.-H. (2020). Compliance Control of Slave Manipulator Using EMG Signal for Telemanipulation. *Appl. Sci.* 10 (4), 1431. doi:10.3390/app10041431
- Aranceta-Garza, A., and Conway, B. A. (2019). Differentiating Variations in Thumb Position from Recordings of the Surface Electromyogram in Adults Performing Static Grips, a Proof of Concept Study. *Front. Bioeng. Biotechnol.* 7, 123. doi:10.1016/j.procs.2017.11.25910.3389/fbioe.2019.00123
- 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, 9–10. doi:10.3389/fnbot.2016.00009
- Baldacchino, T., Jacobs, W. R., Anderson, S. R., Worden, K., and Rowson, J. (2018). Simultaneous Force Regression and Movement Classification of Fingers via Surface EMG within a Unified Bayesian Framework. *Front. Bioeng. Biotechnol.* 6, 13. doi:10.3389/fbioe.2018.00013
- 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
- Ceolini, E., Frenkel, C., Shrestha, S. B., Taverni, G., Khacef, L., Payvand, M., et al. (2020). Hand-Gesture Recognition Based on EMG and Event-Based Camera

## ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

## AUTHOR CONTRIBUTIONS

ZY and DJ provided research ideas and plans; YS, BT and XT wrote programs and conducted experiments. GJ and MX analyzed and explained the simulation results; BC and JK improved the algorithm. JY and YL 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 of the National Natural Science Foundation of China (Grant Nos.52075530, 51575407, 51975324, 51505349, 61733011, 41906177); the Grants of Hubei Provincial Department of Education (D201911105); 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 & Maintenance in Three Gorges University (2020KJX02, 2021KJX13).

Sensor Fusion: A Benchmark in Neuromorphic Computing. *Front. Neurosci.* 14, 637. doi:10.3389/fnins.2020.00637

- Chaiyaroj, A., Sri-Iesaranusorn, P., Buekban, C., Dumnin, S., Thanawattano, C., and Surangsrirat, D. (2019). Deep Neural Network Approach for Hand, Wrist, Grasping and Functional Movements Classification Using Low-Cost Semg Sensors. in IEEE International Conference on Bioinformatics and Biomedicine (San Diego, CA), 1443–1448. doi:10.1109/BIBM47256.2019.8983049
- Chen, J., Bi, S., Zhang, G., and Cao, G. (2020). High-Density Surface EMG-Based Gesture Recognition Using a 3D Convolutional Neural Network. *Sensors* 20, 1201. doi:10.3390/s20041201
- Chen, T., Peng, L., Yang, J., and Cong, G. (2021). Analysis of User Needs on Downloading Behavior of English Vocabulary APPs Based on Data Mining for Online Comments. *Mathematics* 9 (12), 1341. doi:10.3390/math9121341
- Cheng, Y., Li, G., Yu, M., Jiang, D., Yun, J., Liu, Y., et al. (2021). Gesture Recognition Based on Surface Electromyography -feature Image. *Concurrency Computat Pract. Exper* 33 (6), e6051. doi:10.1002/cpe.6051
- Côté-Allard, U., Campbell, E., Phinyomark, A., Laviolette, F., Gosselin, B., and Scheme, E. (2020). Interpreting Deep Learning Features for Myoelectric Control: a Comparison with Handcrafted Features. *Front. Bioeng. Biotechnol.* 8, 158. doi:10.3389/fbioe.2020.00158
- 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. *Concurrency Computat 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. doi:10.1038/srep36571
- Gowtham, S., Krishna, K. M. A., Srinivas, T., Raj, R. G. P., and Joshuva, A. (2020). EMG-based Control of a 5 DOF Robotic Manipulator," 2020 International Conference on Wireless Communications Signal Processing and Networking, Chennai, India, 52–57. doi:10.1109/WISPNET48689.2020.9198439
- Gregory, U., and Ren, L. (2019). Intent Prediction of Multi-Axial Ankle Motion Using Limited Emg Signals. *Front. Bioeng. Biotechnol.* 7, 335. doi:10.3389/fbioe.2019.00335
- He, J., and Jiang, N. (2020). Biometric from Surface Electromyogram (sEMG): Feasibility of User Verification and Identification Based on Gesture Recognition. *Front. Bioeng. Biotechnol.* 8, 58. doi:10.3389/fbioe.2020.00058
- 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. *Cluster Comput.* 22 (Suppl. 5), 10935–10946. doi:10.1007/s10586-017-1237-1
- Hu, J., Sun, Y., Li, G., Jiang, G., and Tao, B. (2019). Probability Analysis for Grasp Planning Facing the Field of Medical Robotics. *Measurement* 141, 227–234. doi:10.1016/j.measurement.2019.03.010
- 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.* doi:10.1007/s12652-020-02843-w
- Jiang, D., Li, G., Sun, Y., Kong, J., Tao, B., and Chen, D. (2019b). 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. (2019a). 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 Generation Comput. Syst.* 123, 94–104. doi:10.1016/j.future.2021.04.019
- Jiang, D., Zhong, Z., Li, G., Sun, Y., Kong, J., Jiang, G., et al. (2019c). Gesture Recognition Based on Binocular Vision. *Cluster Comput.* 22 (Suppl. 6), 13261–13271. doi:10.1007/s10586-018-1844-5
- 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. (2019). Human Lesion Detection Method Based on Access 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. (2019). 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, W., Shi, P., and Yu, H. (2021). Gesture Recognition Using Surface Electromyography and Deep Learning for Prostheses Hand: State-Of-The-Art, Challenges, and Future. *Front. Neurosci.* 15. doi:10.3389/fnins.2021.621885
- Liao, S., Li, G., Wu, H., Jiang, D., Liu, Y., Yun, J., et al. (2021). Occlusion Gesture Recognition Based on Improved SSD. *Concurrency Computat Pract. Exper* 33 (6), e6063. doi:10.1002/cpe.6063
- 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. Intelligence Neurosci.* 2021, 1–12. doi:10.1155/2021/4828102
- 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
- Mendes Junior, J. J. A., Freitas, M. L. B., Siqueira, H. V., Lazzaretti, A. E., Pichorim, S. F., and Stevan, S. L. (2020). Feature Selection and Dimensionality Reduction: an Extensive Comparison in Hand Gesture Classification by Semg in Eight Channels Armband Approach. *Biomed. Signal Process. Control.* 59, 101920. doi:10.1016/j.bspc.2020.101920
- Mujahid, A., Awan, M. J., Yasin, A., Mohammed, M. A., Damaševičius, R., Maskeliūnas, R., et al. (2021). Real-Time Hand Gesture Recognition Based on Deep Learning YOLOv3 Model. *Appl. Sci.* 11, 4164. doi:10.3390/app11094164
- Peng, Y., Tao, H., Li, W., Yuan, H., and Li, T. (2020). Dynamic Gesture Recognition Based on Feature Fusion Network and Variant ConvLSTM. *IET image process* 14 (11), 2480–2486. doi:10.1049/iet-ipr.2019.1248
- Rodríguez-Tapia, B., Soto, I., Martínez, D. M., and Arballo, N. C. (2020). Myoelectric Interfaces and Related Applications: Current State of EMG Signal Processing—A Systematic Review. *IEEE Access* 8, 7792–7805. doi:10.1109/ACCESS.2019.2963881
- Sadikoglu, F., Kavalcioğlu, C., and Dagman, B. (2017). Electromyogram (EMG) Signal Detection, Classification of EMG Signals and Diagnosis of Neuropathy Muscle Disease. *Proced. Comput. Sci.* 120, 422–429. doi:10.1016/j.procs.2017.11.259
- Shahzad, W., Ayaz, Y., Khan, M. J., Naseer, N., and Khan, M. (2019). Enhanced Performance for Multi-Forearm Movement Decoding Using Hybrid IMU-sEMG Interface. *Front. Neurobot.* 13, 43. doi:10.3389/fnbot.2019.00043
- Sri-iesaranusorn, P., Chaiyaroj, A., Buekban, C., Duminin, S., Pongthornseri, R., Thanawattano, C., et al. (2021). Classification of 41 Hand and Wrist Movements via Surface Electromyogram Using Deep Neural Network. *Front. Bioeng. Biotechnol.* 9, 394. doi:10.3389/fbioe.2021.548357
- Su, Z., Liu, H., Qian, J., Zhang, Z., and Zhang, L. (2021). Hand Gesture Recognition Based on sEMG Signal and Convolutional Neural Network. *Int. J. Patt. Recogn. Artif. Intell.* 35, 2151012. doi:10.1142/S0218001421510125
- Sun, Y., Li, C., Li, G., Jiang, G., Jiang, D., Liu, H., et al. (2018). Gesture Recognition Based on Kinect and sEMG Signal Fusion. *Mobile Netw. Appl.* 23 (4), 797–805. doi:10.1007/s11036-018-1008-0
- 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
- Takaiwa, M., Noritsugu, T., Noritsugu, T., Ito, N., and Sasaki, D. (2011). Wrist Rehabilitation Device Using Pneumatic Parallel Manipulator Based on EMG Signal. *Ijat* 5 (4), 472–477. doi:10.20965/ijat.2011.p0472
- 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
- Tian, J., Cheng, W., Sun, Y., Li, G., Jiang, D., Jiang, G., et al. (2020). Gesture Recognition Based on Multilevel Multimodal Feature Fusion. *IFS* 38 (3), 2539–2550. doi:10.3233/JIFS-179541
- Tsinganos, P., Cornelis, B., Cornelis, J., Jansen, B., and Skodras, A. (2019). Improved Gesture Recognition Based on sEMG Signals and TCN. IEEE International Conference on Acoustics, Speech and Signal Processing, IEEE, Brighton, United Kingdom, 1169–1173. doi:10.1109/ICASSP.2019.8683239
- Wei, W., Wong, Y., Du, Y., Hu, Y., Kankanalli, M., and Geng, W. (2019). A Multi-Stream Convolutional Neural Network for sEMG-Based Gesture Recognition in Muscle-Computer Interface. *Pattern Recognition 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. *Concurrency Computat Pract. Exper* 33 (5), e5976. doi:10.1002/cpe.5976
- Wu, C., Zeng, H., Song, A., and Xu, B. (2017). Grip Force and 3D Push-Pull Force Estimation Based on sEMG and GRNN. *Front. Neurosci.* 11, 343. doi:10.3389/fnins.2017.00343
- Wu, Y., Zheng, B., and Zhao, Y. (2018). Dynamic Gesture Recognition Based on LSTM-CNN. *Chin. Automation Congress*, 2446–2450. doi:10.1109/CAC.2018.8623035
- 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. *Mechanism Machine Theor.* 159, 104265. doi:10.1016/j.mechmachtheory.2021.104265
- Xie, T., Leng, Y., Zhi, Y., Jiang, C., Tian, N., Luo, Z., et al. (2020). Increased Muscle Activity Accompanying with Decreased Complexity as Spasticity Appears: High-Density EMG-Based Case Studies on Stroke Patients. *Front. Bioeng. Biotechnol.* 8, 1338. doi:10.3389/fbioe.2020.589321

- Xu Zhang, X., Xiang Chen, X., Yun Li, Y., Lantz, V., Kongqiao Wang, K., and Jihai Yang, J. (2011). A Framework for Hand Gesture Recognition Based on Accelerometer and EMG Sensors. *IEEE Trans. Syst. Man. Cybern. A.* 41 (6), 1064–1076. doi:10.1109/TSMCA.2011.2116004
- Yu, B., Luo, Z., Wu, H., and Li, S. (2020). Hand Gesture Recognition Based on Attentive Feature Fusion. *Concurrency Computat Pract. Exper* 32 (22), e5910. doi:10.1002/cpe.5910
- 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
- Zhang, W., Shuai, L., and Kan, H. (2021). Real-time Gesture Recognition Based on Improved Artificial Neural Network and sEMG Signals. *IEEE International Conference on Mechatronics and Automation (ICMA)*, IEEE, Takamatsu, Japan, 981–986. doi:10.1109/ICMA52036.2021.9512756
- Zhang, X., and Li, X. (2019). Dynamic Gesture Recognition Based on MEMP Network. *Future Internet* 11, 91. doi:10.3390/fi11040091
- Zhang, X., Zhou, X., Lin, M., and Sun, J. (2018). Shufflenet: An Extremely Efficient Convolutional Neural Network for mobile Devices. *Proc. IEEE Conf. Comput. Vis. pattern recognition* 11, 6848–6856. doi:10.1109/CVPR.2018.00716
- Zhang, Z., Yang, K., Qian, J., and Zhang, L. (2019). Real-Time Surface EMG Pattern Recognition for Hand Gestures Based on an Artificial Neural Network. *Sensors* 19, 3170. doi:10.3390/s19143170
- 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 © 2021 Yang, Jiang, Sun, Tao, Tong, Jiang, Xu, Yun, Liu, Chen and Kong. 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.