

Research Article

Open Access



CAM-MR-MS based gesture recognition method using sEMG

Lina Tong¹, Yunbo Li¹, Yixia Liang¹, Chen Wang²

¹China University of Mining and Technology Beijing, Beijing 100083, China.

²State Key Laboratory of Multimodal Artificial Intelligence Systems Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China.

Correspondence to: Yunbo Li, China University of Mining and Technology Beijing, No. 11 Ding, Xueyuan Road, Haidian District, Beijing 100083, China. E-mail: lyblyw2000@163.com

How to cite this article: Tong, L.; Li, Y.; Liang, Y.; Wang, C. CAM-MR-MS based gesture recognition method using sEMG. *Intell. Robot.* 2025, 5, 292-312. <https://dx.doi.org/10.20517/ir.2025.15>

Received: 6 Nov 2024 **First Decision:** 18 Feb 2025 **Revised:** 2 Mar 2025 **Accepted:** 19 Mar 2025 **Published:** 8 Apr 2025

Academic Editor: Chunjie Chen **Copy Editor:** Pei-Yun Wang **Production Editor:** Pei-Yun Wang

Abstract

With the continuous concern for the disabled and the elderly, intelligent prosthetics and service robots have been widely applied. This paper provides a method for gesture recognition using forearm surface electromyography (sEMG), including an adaptive channel selection method to simplify the sEMG measurement. Based on the forearm muscle groups corresponding to different movements, surface skin areas are divided, and the Myo bracelet is used to collect sEMG signals from these areas. A method combined with channel attention module, multi-channel relationship feature extraction module and multi-scale skip connection module is built to adaptively select the signals from certain skin areas and recognize the seven gestures during experiment. The comparative experimental results indicate that this method can adaptively extract the optimal channel combination and show effective recognition results. It improved the practicability for the sEMG-based gesture recognition.

Keywords: Surface skin area, surface electromyography (sEMG), channel selection, gesture recognition, optimal channel combination

1. INTRODUCTION

With the increase in population aging and chronic diseases, the proportion of the disabled population is rising. Physical disabilities not only affect individuals' physical and mental health but also elevate their living costs. These patients with upper limb disabilities face significant challenges in daily life and work, resulting



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.



in lower social participation^[1]. As an assistive device, intelligent prosthetics can utilize the surface electromyography (sEMG) signals from the amputee's stump to control peripheral mechanical devices through advanced gesture recognition algorithms, thereby restoring some lost functions and significantly enhancing patients' self-care abilities. In order to utilize prosthetics for daily tasks such as fist clenching and finger extension, the system must first discern the operator's intended hand movements and execute corresponding gesture actions in accordance with the patient's intentions.

Currently, sEMG signals have found wider applications in the fields of gesture recognition and detection of other human movement intentions^[2]. They are bioelectric signals generated by the synergistic contraction of muscle tissue during human movement, typically arising 30 to 150 milliseconds before limb movement and reflecting changes caused by muscle activity^[3]. They can provide stable and abundant activity information, such as muscle contraction levels and muscle coordination, which can be used to identify different gesture movements and further control prosthetics^[4].

Research has indicated that as the number of sEMG channels increases, the accuracy of gesture recognition classification improves with the increase in the number of sEMG channels. However, once the number of channels surpasses a certain threshold, it can compromise the computational efficiency of the system^[5]. In such cases, it becomes crucial to screen the sEMG channels and extract appropriate combinations of channels. Qu *et al.* used the Relief-F method to calculate the weight of each sEMG feature value, averaged the feature weights of each channel as the weight of that channel, and ranked them^[6]. Experimental results indicate that the prediction accuracy for different gestures using the combination of the four highest-weighted channels is 98.75%, less than 1% lower than the prediction accuracy of more channels. Lele *et al.* proposed an algorithm based on the gradient boosting decision tree (GBDT) to select the optimal channel combination^[7]. The average recognition accuracy of the optimal channel combination for eight hand movements with four channels is 95.83%, which is basically the same as the recognition accuracy with six channels. This suggests that the proposed method can reduce the number of electrodes while ensuring high recognition accuracy. Wang *et al.* employed a genetic algorithm to reduce the number of sEMG channels from 16 to 11^[8]. This combination of 11 channels, which represents only 70% of the total number of channels, achieved 97% of the optimal classification performance, thereby validating the feasibility of using a genetic algorithm for channel reduction. Huang *et al.* employed a sequential forward selection (SFS) method for electrode channel selection, which significantly reduced the number of electrode channels while only decreasing the classification accuracy by 1.2% compared to using other electrode channels^[9]. Liu *et al.* used the Markov random field (MRF) method for optimal channel selection and classified the features of the selected channels using a K-nearest neighbor (KNN) classifier^[10]. This method can effectively reduce redundant information from different channels and obtain high classification accuracy similar to using all sEMG channels, but it also has a higher computational cost. Niu *et al.* built PCS-EMGNet based on convolutional neural network (CNN), which can automatically select channels and features^[11]. Compared with models with a similar number of parameters, the average accuracy improved by 9.96%. Compared with previous gesture recognition models, the parameter quantity of this model was reduced by an average of 80%, demonstrating that this network can reduce the model's parameter quantity while maintaining good accuracy.

Currently, there are many research methods for gesture recognition. For example, Zhang *et al.* used five features: mean absolute value (MAV), waveform length (WL), root mean square (RMS), slope sign change (SSC), and Hjorth parameters^[12]. With 12 subjects, the average accuracy for five gestures reached 97.8%. Lee *et al.* extracted six time-domain features such as RMS and VAR from three channels and used methods such as artificial neural network (ANN), support vector machines (SVM), random forest (RF), and logistic

regression (LR)^[13]. The average accuracy for ten gestures was 99.4%, 87.6%, 83.1%, and 53.9%, respectively. Geng *et al.* used a ConvNet architecture composed of four convolutional layers and two fully connected layers to recognize gestures from instantaneous sEMG signal images^[14]. On three sEMG signal benchmark databases, the recognition accuracy for 52 gestures was 77.8%, for eight gestures was 99.5%, and for 27 gestures was 89.3%. Zhang *et al.* decomposed the original sEMG images into equal-sized patch streams and then applied a multi-stream CNN fusion network for gesture recognition from the feature images, achieving an 85% recognition accuracy^[15]. Hu *et al.* proposed a hybrid CNN-RNN network structure based on the attention mechanism, achieving an average gesture recognition accuracy of 84.80% based on the NinaproDB1 dataset^[16]. Wang *et al.* improved the LSTM-CNN network by adding the convolutional block attention module (CBAM), increasing the recognition accuracy by 5.3%^[17].

Despite significant progress in research on sEMG channel selection and gesture recognition, several issues remain. For channel selection, most current studies are based on machine learning, which involves complex calculations. Additionally, the results of channel selection heavily depend on the position of the sEMG electrodes. Moreover, in some studies, although researchers conducted channel selection, the number of selected channels remained relatively high. For gesture recognition, current methods do not recognize different gesture movements by examining the relationships between sEMG channels.

Addressing the above issues, our research found commonalities in force generation patterns among different subjects. Therefore, based on the relationship between different muscles and movements, we divided eight surface skin areas and used channel attention module (CAM) to assign weights to the channels, achieving channel selection and extracting the optimal channel combination, which solved the problem of sEMG electrode placement. Furthermore, based on CNN and the CAM, this paper designed a multi-channel relationship extraction module and a multi-scale skip connection module, constructing the CMCS-ResNet classification model, which exhibited good classification performance on our dataset.

This paper is organized as follows: Section 2 introduces how we delineate surface skin areas and presents the relevant content of collecting sEMG signals. Section 3 describes the data processing and feature extraction of sEMG signals. Section 4 introduces the model constructed in this paper and the common patterns of muscle force generation among subjects that we have discovered. Section 5 presents the experimental results and validates the effectiveness of our proposed channel selection method. Finally, a summary of this article will be given in Section 6.

2. COLLECTION OF SEMG SIGNALS FROM SKIN AREAS

In daily life, the execution of different hand gestures requires corresponding muscle extension and contraction^[18]. This study designed seven different hand gestures and divided the surface skin areas based on the muscle regions required to complete these gestures. Ultimately, sEMG signals were collected from the surface skin areas of ten subjects.

2.1. Design of hand gesture actions

The hand gesture recognition technology presented in this paper is applicable for controlling prosthetics to perform common hand gesture actions in people's daily lives. There are a total of seven hand gesture actions: peace sign (V-sign), index finger extension, pronation of the palm, supination of the palm, thumbs-up, fist clenching, and palm opening [Figure 1].

2.2. Division of surface skin areas

The forearm of the human body contains multiple small muscle groups, each with distinct functions that

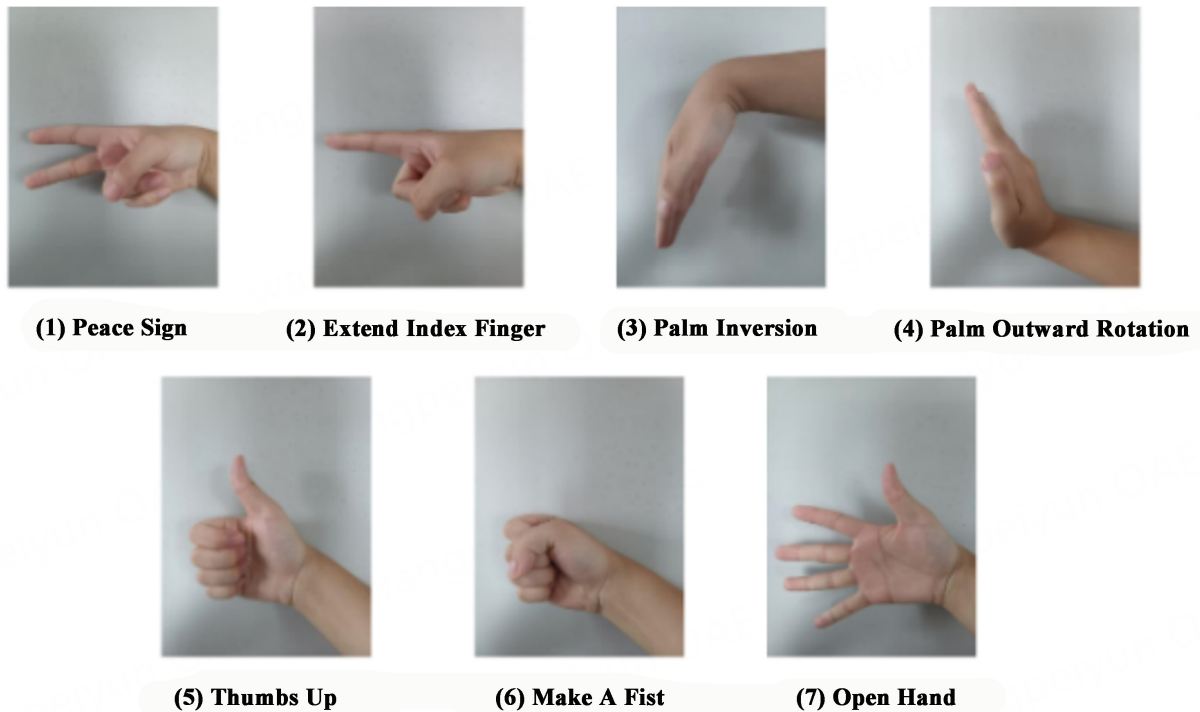


Figure 1. The seven hand gesture actions.

work together to perform various hand gesture actions. In the same hand gesture, different muscle groups contribute differently to the action, meaning that their electromyographic signals exhibit significant differences. This makes it possible to recognize different hand gesture actions based on electromyographic signals.

Figure 2 is a cross-sectional view of the forearm, depicting muscles such as the extensor carpi radialis brevis (ECRB), extensor carpi radialis longus (ECRL), brachioradialis (BR), flexor carpi ulnaris (FCU), palmaris longus (PL), flexor carpi radialis (FCR), flexor digitorum superficialis (FDS), extensor digitorum (ED), extensor digiti minimi (EDM), extensor carpi ulnaris (ECU), flexor digitorum profundus (FDP), pronator teres (PT), supinator (S), radius (R), and ulna (U)^[19]. The correspondence between the muscles and joint movements is shown in **Table 1**.

The seven hand gesture actions designed in this paper mainly involve the finger joints and wrist joints, corresponding to the muscles listed in the aforementioned table: ECRB, ECRL, BR, FCU, PL, FCR, FDS, ED, EDM, ECU, and FDP.

Based on the locations of the aforementioned muscles, this paper has divided the forearm into eight surface skin areas, each corresponding to different muscles and designated for placing sEMG electrodes. As shown in **Figure 3**, surface skin areas 1 to 8 correspond to the superficial muscles: ED and EDM, ECRB and BR, BR, FCR, PL, FCU, FDP, and FDP and ECU, respectively. An EMG electrode is placed in the center of each surface skin area to collect sEMG signals.

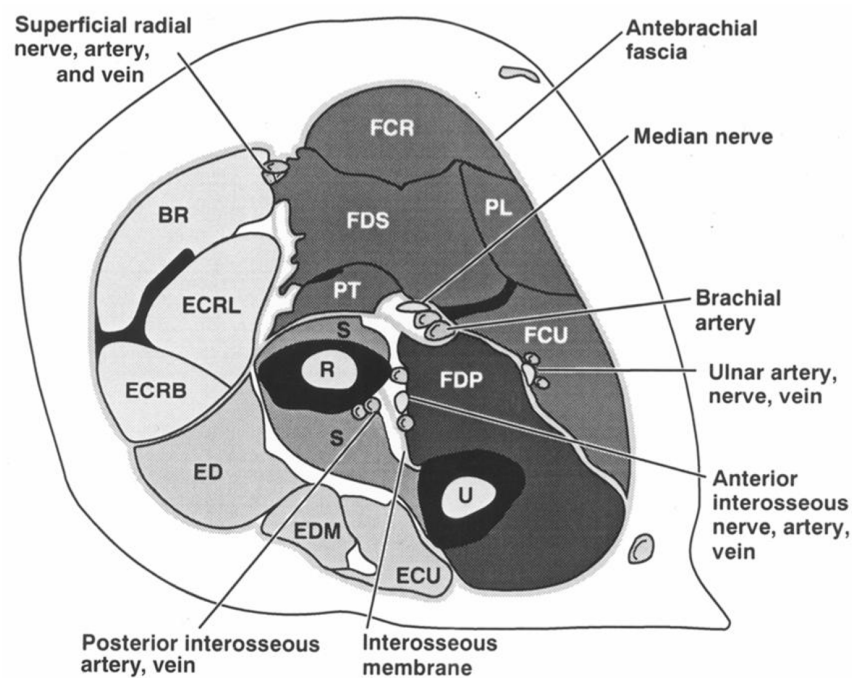
2.3. Acquisition of sEMG signals

In this paper, the Myo armband produced by Thalmic Labs in Canada is used to collect sEMG signals from

Table 1. Muscles and joints

Muscles	Joints
ECRB	Wrist
ECRL	Wrist
BR	Wrist
FCU	Wrist
PL	Wrist
FCR	Wrist
FDS	Wrist and finger
ED	Wrist and finger
EDM	Wrist and finger
ECU	Wrist
FDP	Finger

ECRB: Extensor carpi radialis brevis; ECRL: extensor carpi radialis longus; BR: brachioradialis; FCU: flexor carpi ulnaris; PL: palmaris longus; FCR: flexor carpi radialis; FDS: flexor digitorum superficialis; ED: extensor digitorum; EDM: extensor digiti minimi; ECU: extensor carpi ulnaris; FDP: flexor digitorum profundus.

**Figure 2.** Cross-sectional view of the forearm^[19].

the forearm. As shown in Figure 4, the Myo armband has eight electromyographic sensors with a sampling frequency of 200 Hz, which can conveniently collect sEMG signals for different hand gestures. It has the advantages of easy wearing and low hardware cost^[20]. Even non-professionals can apply it with ease, significantly broadening the scope of its application, such as in community rehabilitation and home care. However, since different electrodes correspond to different skin surface areas each time they are worn, algorithms are required to adaptively identify sEMG signals from effective regions. This is also one of the key objectives of our research.

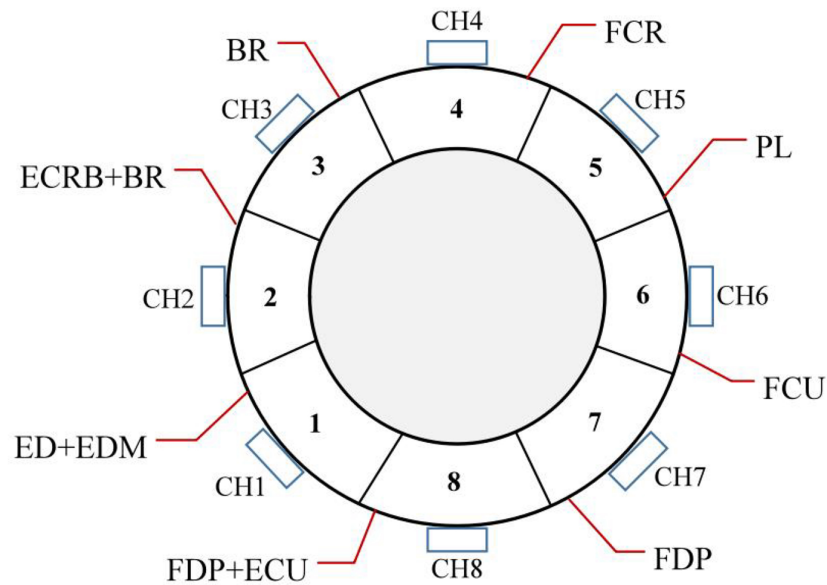


Figure 3. Surface skin areas and positions of electrodes.



Figure 4. Myo armband.

The collection of sEMG signals follows the following procedure:

First, before the collection begins, explain the experimental process to the subject and teach them how to perform the corresponding hand gesture actions.

Second, clean the skin surface with 75% alcohol to remove static electricity and contaminants.

Third, the subject wears the Myo armband for testing to ensure that it can normally collect data and that the subject understands how to perform the corresponding hand gesture actions.

Last, proceed with the formal data collection. For the seven hand gesture actions, the armband is worn on the subject's dominant forearm. Each action is collected ten times, with each lasting for three seconds. After

completing one action, the subject relaxes their arm muscles for six seconds. The rest interval between each action is two minutes.

Ten healthy subjects (Age: 24.7 ± 1.1) participated in the experiment; their specific information is shown in Table 2.

Figure 5 shows the raw sEMG signal of one action performed by one of the subjects.

3. DATA PROCESSING AND FEATURE EXTRACTION

During the process of collecting sEMG signals using the Myo armband, the signals may be affected by noise^[21]. Therefore, preprocessing operations such as filtering are required for the raw electromyographic data. After filtering, activity segment detection technology is used to extract the sEMG signals of the activity segments from the original continuous signals, and these activity segment data are segmented using a sliding window to increase the data volume.

After completing the aforementioned data preprocessing, time-domain features and time-frequency domain features are extracted from the data in each time window to prepare for subsequent hand gesture recognition.

3.1. sEMG filtering and denoising

To avoid the influence of power supply, electronic devices, and environmental noise on sEMG signals during the signal acquisition process, we adopt 50 Hz power frequency filtering, wavelet transform denoising, and a 5–80 Hz band-pass filter to filter and denoise the sEMG signals. A comparison before and after filtering is shown in Figure 6.

3.2. Signal segmentation

The sEMG signals we collected contain multiple repeated actions, and we need to extract the starting time points of each action to facilitate further data processing. Commonly used activity segment detection methods include wavelet transform^[22] and short-time Fourier transform^[23]. These methods are overly complex in computation, while some methods, such as wavelet transform, require prior knowledge of sEMG signals, resulting in a cumbersome process for extracting active segments of sEMG signals and making them unsuitable for practical applications in sEMG signal control. In contrast, the method of average standard deviation^[24] involves applying a one-dimensional moving window to the signal and determining whether the beginning and end of a signal are within a window by calculating the threshold of the average standard deviation within that window.

In this paper, we choose the mean standard deviation method to detect and extract the activity segments of the sEMG signals, which is given as follows:

$$A_1[S, ch] = \frac{1}{L} \sum_{s=L}^S \left(X[S, ch] - \frac{1}{L} \sum_{s=L}^S X[S, ch] \right)^2 \quad (1)$$

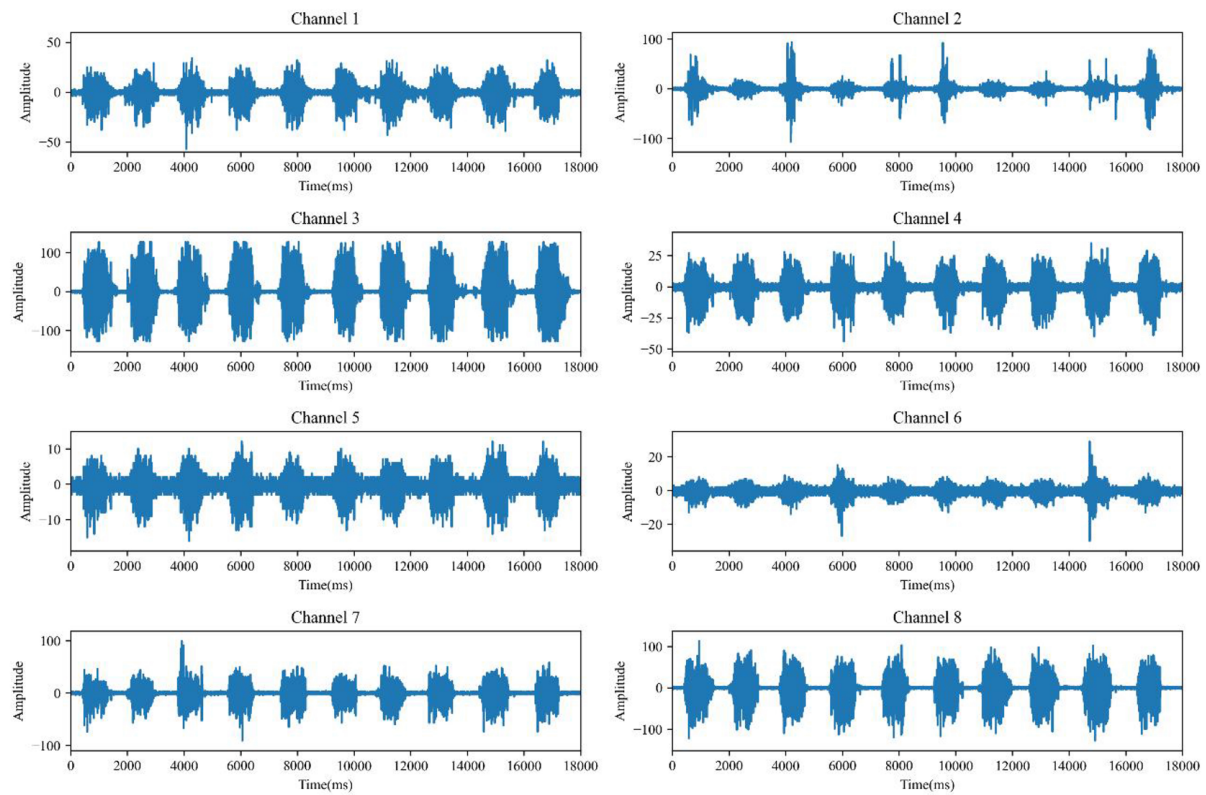
$$A_2[t] = \frac{1}{N} \sum_{ch=1}^N A_1[t, ch] \quad (2)$$

where S represents the time step of the sEMG signal, ch indicates the channel of the sEMG signal, L stands for the length of the sliding window, X signifies the data of the sEMG signal in a sample, and N points to the number of channels. The result of signal segmentation is shown in Figure 7.

Table 2. Information of subjects

Subjects	Gender	Height (m)	Weight (kg)	BMI
No.1	Male	1.71	63	21.55
No.2	Male	1.85	75	21.91
No.3	Male	1.90	83	22.99
No.4	Male	1.73	60	20.05
No.5	Male	1.80	54	16.67
No.6	Female	1.63	60	22.58
No.7	Female	1.66	62	22.50
No.8	Female	1.58	52	20.83
No.9	Female	1.70	56	19.38
No.10	Male	1.83	80	23.89
Average		1.74 ± 0.10	64.5 ± 10.39	21.23 ± 2.00

BMI: Body mass index.

**Figure 5.** Eight channels of sEMG signals. sEMG: Electromyography.

3.3. Data segmentation using sliding window

Sliding window segmentation is a commonly used data augmentation method that divides the sEMG signal into individual windows, which helps improve the accuracy of hand gesture recognition^[22]. Figure 8 illustrates the principle of the sliding window segmentation method.

The sliding window segmentation method includes overlapping window segmentation and non-overlapping window segmentation. Compared to non-overlapping window segmentation, overlapping window

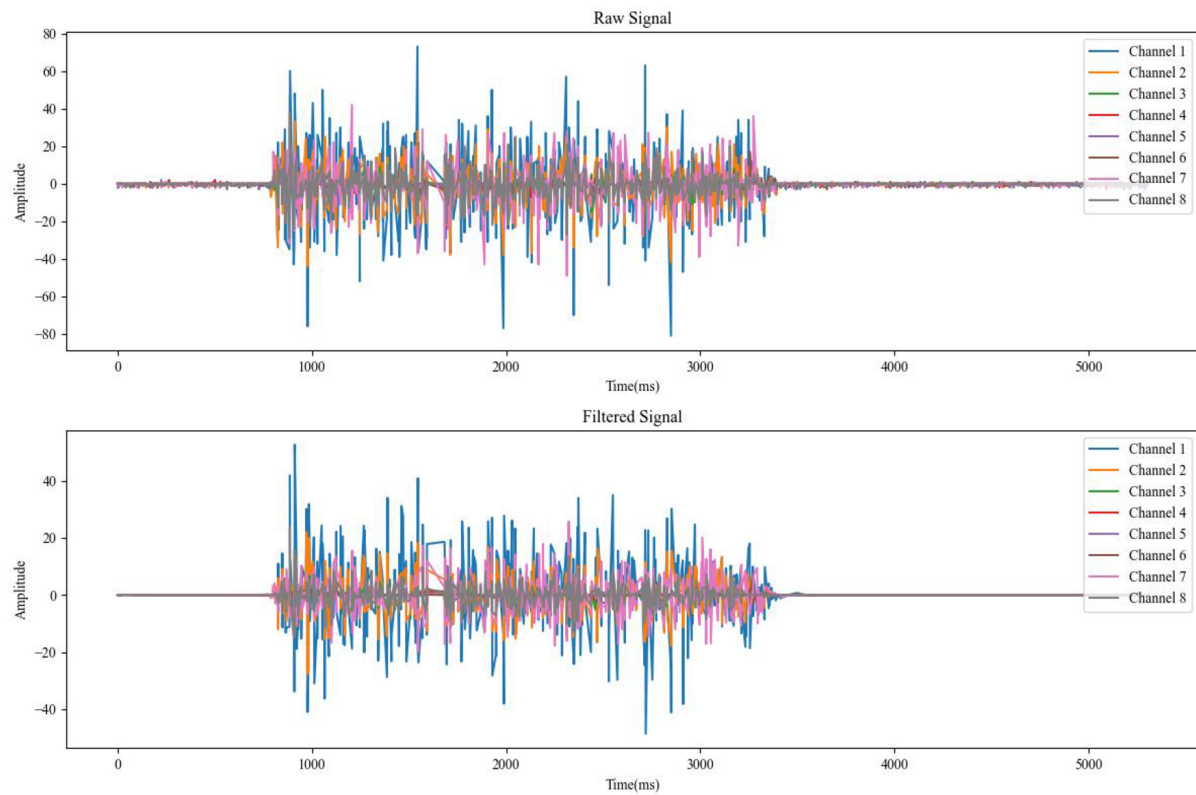


Figure 6. Raw signal and filtered signal.

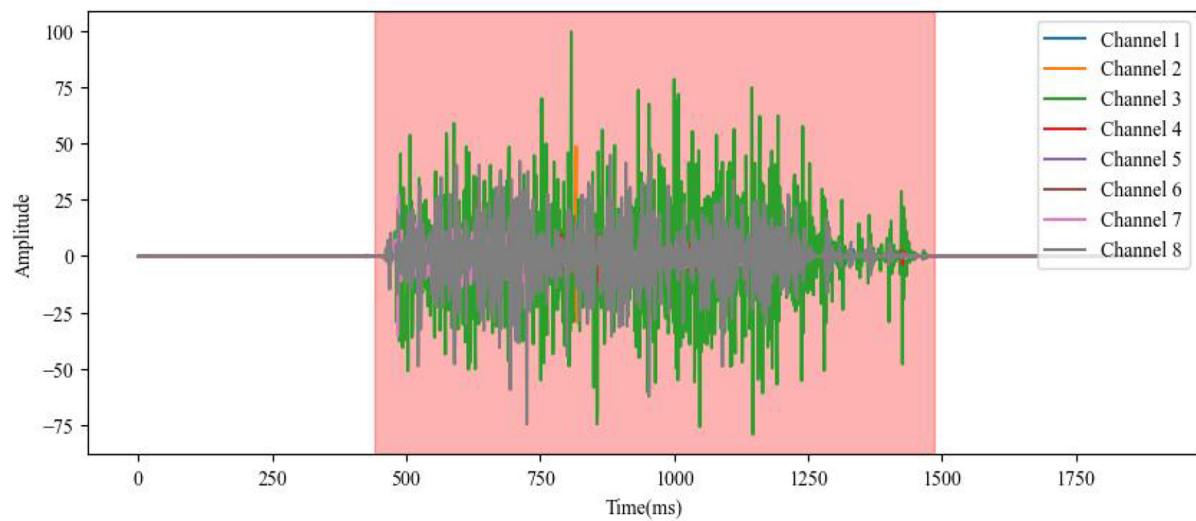


Figure 7. Interception of muscle activity time periods.

segmentation can expand the size of the sample data and include sEMG data that is more representative of the gesture features. Shorter windows may result in unstable signal features, affecting the accuracy of hand gesture recognition; longer windows may lead to insufficient data sample size. In this paper, a window width of 300 ms and a step size of 150 ms are selected.

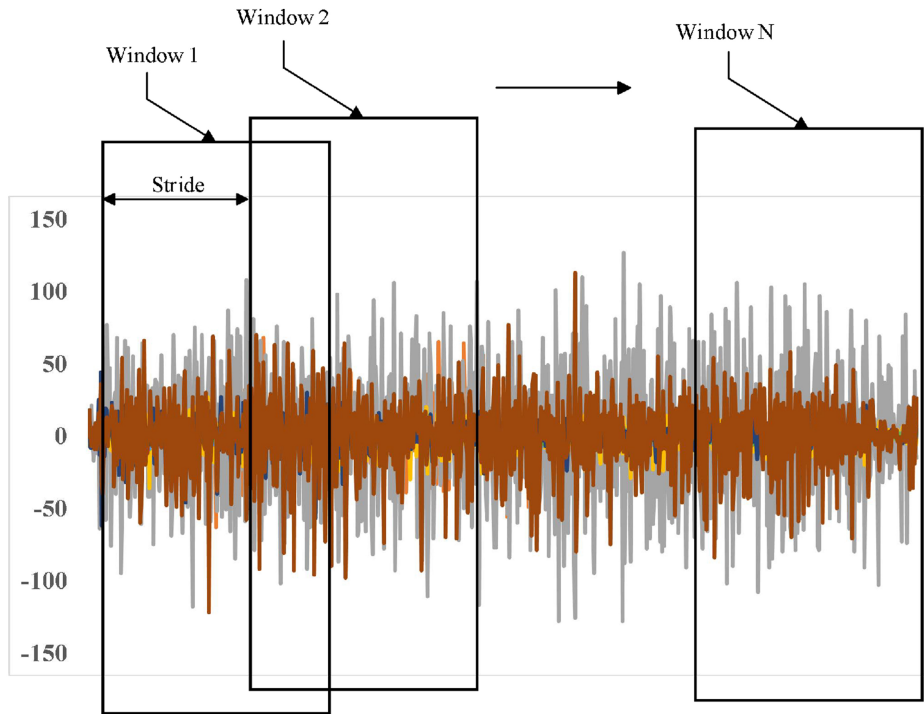


Figure 8. Sliding window segmentation.

3.4. Feature extraction

The feature extraction of sEMG signals can be classified into three types: time-domain features^[23], frequency-domain features^[24], and time-frequency domain features^[25]. Time-domain features can provide intuitive information about the signal within the time range, while time-frequency domain features can reveal the complex characteristics of the signal in terms of frequency. The combined use of both can more comprehensively reflect the essential characteristics of the signal, improving the accuracy of signal analysis and recognition^[25]. In this paper, eight time-domain features and eight time-frequency domain features are extracted for each channel of the sEMG signal.

MAV is the average of the absolute values of the sEMG amplitudes within a time window, and can be written as:

$$MAV = \frac{1}{n} \sum_{i=1}^n |s_i| \quad (3)$$

Variance (VAR) is the average of the squared deviations of each time sample, and can be written as:

$$VAR = \frac{1}{n-1} \sum_{i=1}^n s_i^2 \quad (4)$$

RMS is the square root of the average of the squared signal amplitudes, and can be written as:

$$RMS = \sqrt{\frac{1}{n} \sum_{i=1}^n s_i^2} \quad (5)$$

Difference absolute standard deviation value (DASDV) is the standard deviation of the lengths of the sEMG signals, and can be written as:

$$DASDV = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n-1} (s_{i+1} - s_i)^2} \quad (6)$$

Average amplitude change (AAC) is the average length of the sEMG waveforms within a time window, and can be written as:

$$ACC = \frac{1}{n} \sum_{i=1}^{n-1} |s_{i+1} - s_i| \quad (7)$$

Absolute value of the 3rd temporal moment (TM3) can be written as:

$$TM3 = \left| \frac{1}{n} \sum_{i=1}^n s_i^3 \right| \quad (8)$$

Zero crossing (ZC) represents the number of times the sEMG waveform crosses the zero point within a time window, and can be written as:

$$ZC = \sum_{i=1}^{n-1} f(-s_i \cdot s_{i+1}) \quad (9)$$

SSC represents the number of changes in the slope of the sEMG waveform within a time window, and can be written as:

$$SSC = \sum_{i=2}^{n-1} [[f(s_i - s_{i-1}) \cdot (s_i - s_{i+1})]] \quad (10)$$

The three-level wavelet packet transform (WPT) is applied to the sEMG signal to decompose it into eight equal-width frequency bands. The decomposition process is shown in [Figure 9](#).

The square mean of the coefficients obtained from each subspace after the WPT represents the energy value of each frequency band subspace, known as the energy of frequency band (EFB), which is calculated by

$$EFB_i = \frac{1}{N_i} \sum_{j=1}^{N_i} C_{ij}^2 \quad (11)$$

where i represents the different frequency band subspaces, ranging from 1 to 8, C_{ij} indicates the value of the j -th coefficient in the i -th frequency band. A total of eight features are extracted.

The dataset used in this paper is sourced from ten subjects, each performing seven gesture actions, totaling 14,180 samples. After feature extraction, the shape of the dataset is (14180, 8, 16).

4. CHANNEL SELECTION AND GESTURE RECOGNITION

This study has developed a model for sEMG signal channel selection and gesture recognition classification. Based on the feature data extracted from ten subjects, the model was trained and the weights of different channels for each subject were obtained and plotted as a heatmap. From the heatmap, common patterns in muscle activation among the ten subjects were identified, making it possible to screen out the optimal channel combinations.

4.1. Design of channel selection method

This paper extracted the weights of eight muscle regions from ten subjects and plotted them as heatmaps [[Figure 10](#)].

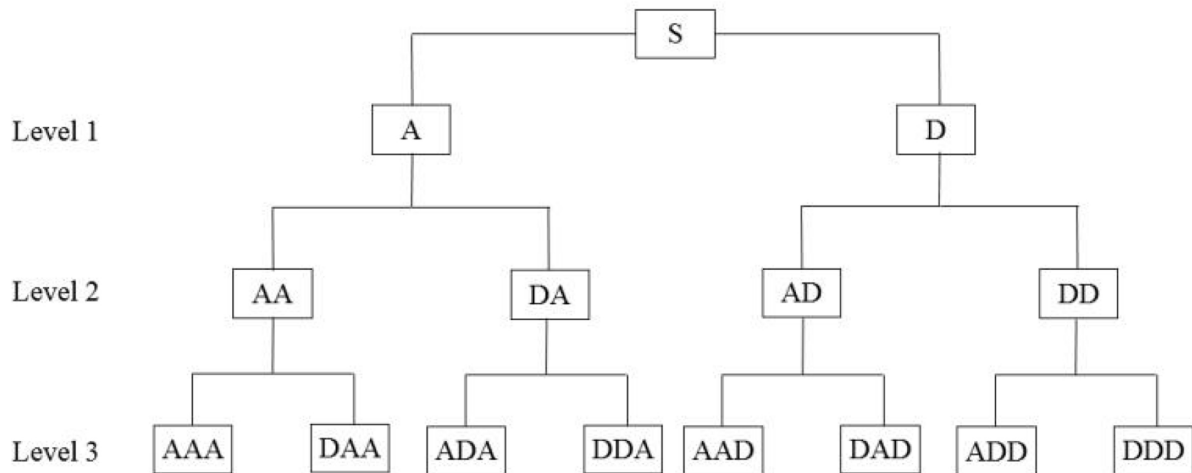


Figure 9. Decomposition process of WPT. WPT: Wavelet packet transform.

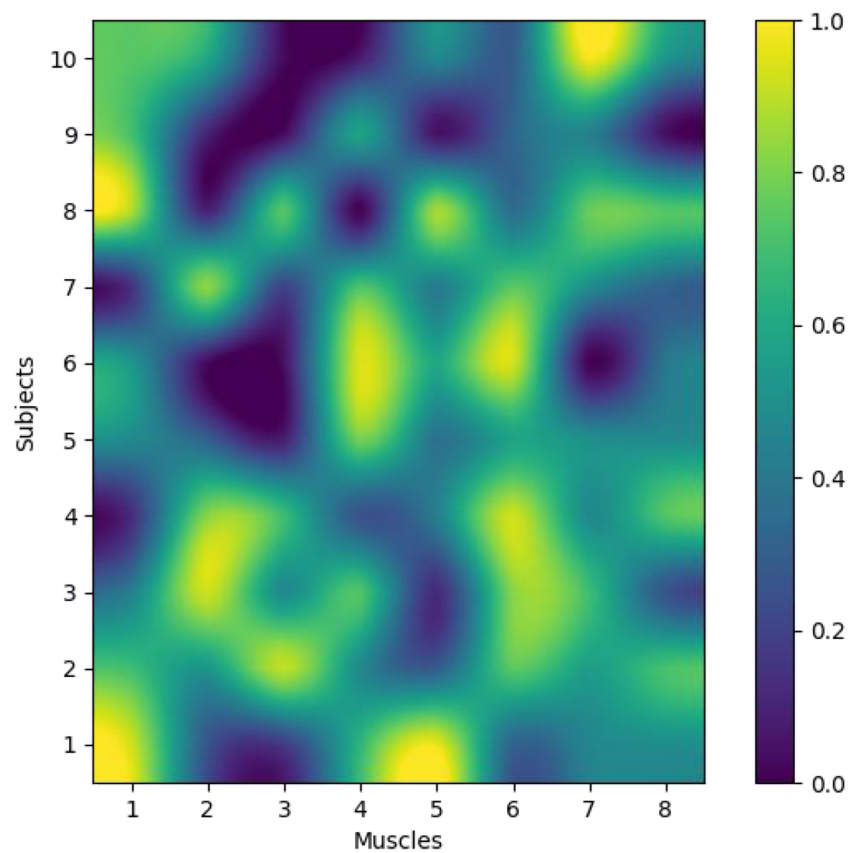


Figure 10. Heatmap of sEMG channel weights. sEMG: Electromyography.

From the heatmaps, a common pattern can be observed across the seven types of gestures performed by the ten subjects: muscle regions 1, 4, 5, 6, and 7 exhibit larger weights, indicating that these regions contribute significantly to the classification of gesture recognition. This demonstrates that the method we used is capable of assigning weights to different sEMG channels, making it possible to screen out important

channels and select the optimal channel combination.

To reduce redundant information during gesture recognition, enhance the operational efficiency of the model, and automatically select important sEMG channels, it is essential to perform channel selection on sEMG signals^[26]. With advancements in deep learning technologies, the CAM^[27] has emerged as a highly promising solution in sEMG channel selection. Unlike other attention mechanisms, CAM, as a specialized form of attention, excels in automatically assigning precise weights to various channels during model training. This capability allows it to accurately depict the individual contributions of each sEMG channel to gesture recognition, thereby outperforming other methods in this domain.

The structure of the CAM algorithm used in this paper is shown in [Figure 11](#). Its main components include max pooling, average pooling, a multi-layer perceptron (MLP), a sigmoid function, and a ReLU function. Assuming the input data has a shape of (Batchsize, C, H, W), the input data is first sent to both max pooling and average pooling to obtain two feature maps with shapes of (Batchsize, C, 1, 1). These feature maps are then fed into the MLP to produce two more feature maps with shapes of (Batchsize, C, 1, 1). After summing these feature maps, they pass through the sigmoid function to obtain the final channel weights.

Using the CAM to calculate the weight of 8-channel sEMG signals. The process is as follows:

First, input the subject's data into the model for training.

Second, extract the channel weights from the CAM layer of the trained network.

Third, sort the channels based on their weights.

Last, select a certain number of channels.

4.2. Design of gesture recognition classification model

ResNet has demonstrated excellent performance in gesture recognition^[28]. Based on this model, this paper introduces a multi-channel relationship feature extraction module (MR) to capture the relationship features among EMG channels at different scales, and a multi-scale skip connection module (MS) to further perceive and extract features at various scales. In addition to assigning weights to different sEMG channels, CAM can also assign weights to the channels of feature maps, reducing redundancy among channels. The overall structure of the model in this paper is shown in [Figure 12](#).

The structure of the MR module is shown in [Figure 13](#). This module employs parallel convolutional kernels of different sizes to extract features of the relationships between sEMG channels from both global and local perspectives.

The structure of the MS module is shown in [Figure 14](#). It employs multiple parallel convolutional and pooling branches to perceive and extract features at different scales. During the process of layer-by-layer convolution, only the high-level features extracted from the output of the last convolutional layer are ultimately retained, while the low-level features contained in the feature maps extracted by other convolutional layers are not preserved. Therefore, this paper introduces skip connections to retain the feature maps output by each convolutional kernel. After fusing the feature maps output by each convolutional kernel, the channel attention mechanism is used to remove redundant information in the feature maps.

5. EXPERIMENT

5.1. Verification experiments for channel selection methods

To further determine the number of sEMG channels and extract the optimal channel combination, we

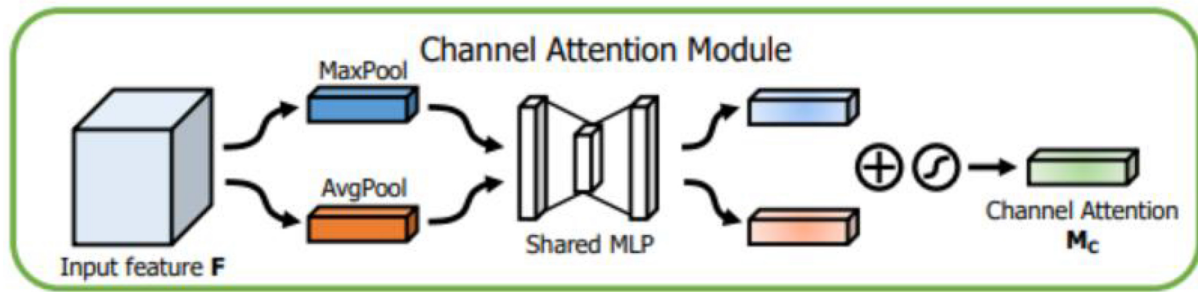


Figure 11. CAM. CAM: Channel attention module.

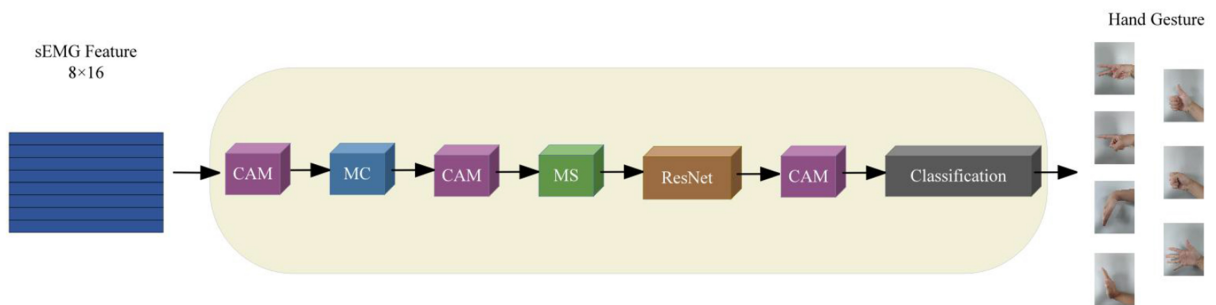


Figure 12. Model structure.

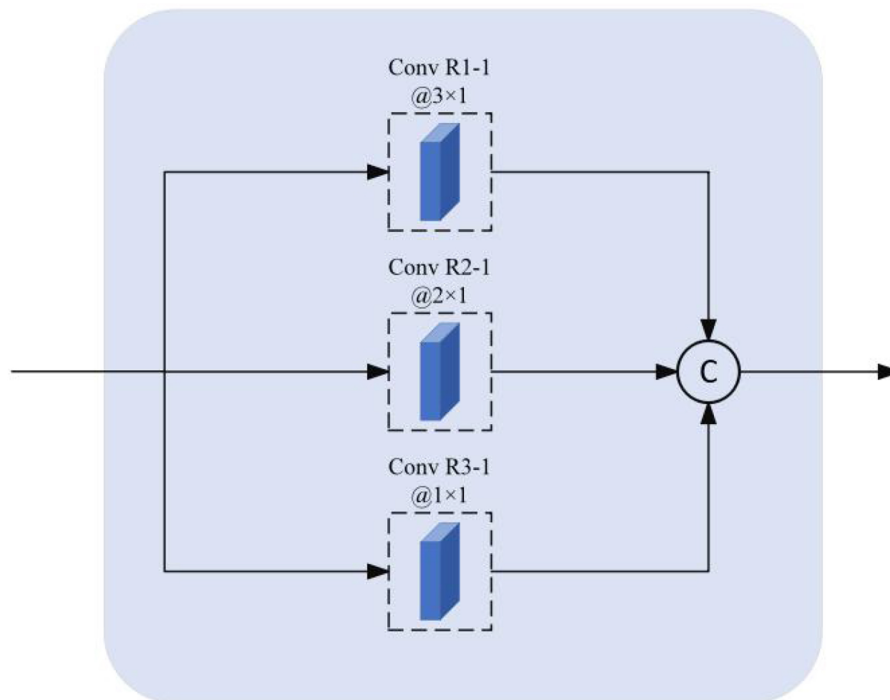


Figure 13. MR module. MR: Multi-channel relationship feature extraction.

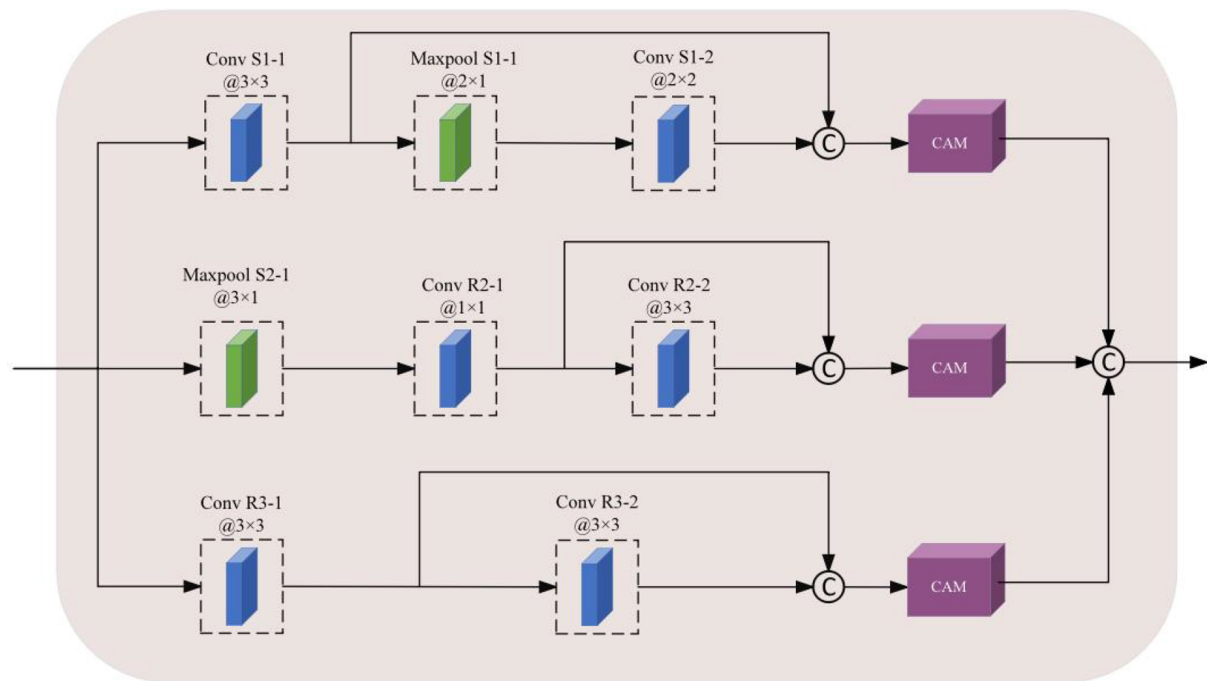


Figure 14. MS module. MS: Multi-scale skip connection.

investigated the relationship between the number of channels, classification accuracy, and model runtime. Based on this relationship, the number of sEMG channels was determined. Finally, two subjects based on body mass index (BMI) were selected to verify the effectiveness of the optimal channel selection method. In this paper, a 5-fold cross-validation method is employed to partition the dataset and train the model.

This paper calculates the average recognition accuracy and prediction time for different numbers of sEMG channels across ten subjects [Figure 15].

We can observe that as the number of channels increases, the average recognition accuracy of the model tends to increase. When the number of channels exceeds 3, the slope of accuracy improvement decreases significantly. When the number of channels exceeds 5, the growth in accuracy becomes very slow. This indicates that our proposed method prioritizes retaining more important channels during channel selection, resulting in a rapid initial increase in accuracy followed by a slower change later on, which aligns with our design expectations.

When using the 8-channel signals from the entire skin area, the recognition accuracy reaches its peak at 93.22%. However, this approach requires a relatively large number of model parameters and substantial computational resources, making it unsuitable for all applications. For instance, when computational time is limited, one can opt for fewer channels by selecting the channels with higher weights for recognition. From Figure 15, it can be seen that the model runtime is 60.90 ms with three channels, which is a reduction of 107.03% compared to the 8-channel configuration. When the number of channels is 4, the runtime increases by 19.8% compared to three channels, but the prediction accuracy only increases by 4.47%. Therefore, we choose three sEMG channels as the optimal channel combination example to verify the effectiveness of this optimal channel combination. Here, one subject with normal BMI and one with underweight BMI from the ten subjects were selected, to compare the gesture recognition accuracy of the optimal channel combination

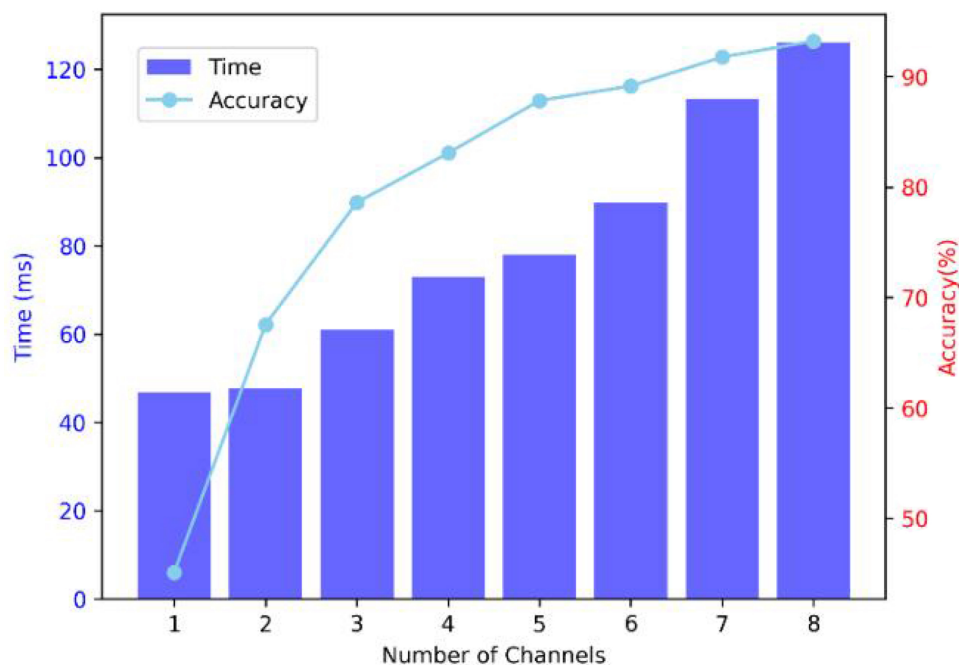


Figure 15. Relationship between accuracy, prediction time and the number of sEMG channels. sEMG: Electromyography.

for each subject with the accuracy of the remaining 55 channel combinations (Since there are 56 combinations of 3 channels from 8). Figures 16 and 17 show the classification accuracy of the models for underweight and normal BMI subjects, respectively, under different combinations, with the orange line representing the accuracy of the optimal channel combination. It can be seen that the accuracy of the optimal channel combination selected in this paper is superior to that of the other 55 channel combinations. This proves the effectiveness of the adaptive optimal channel selection method.

5.2. Comparison experiment of gesture recognition models

To verify the performance advantages of the model designed in this paper, we compared the gesture recognition classification performance of the models with the most references in the field, such as SE-CNN^[29], VMSCNN^[30], and SqueezeNet^[31], on self-built datasets and the Ninapro DB5 public dataset, under both low-channel and full-channel conditions. The experimental results are shown in Tables 3 and 4 respectively. It can be seen that on both datasets, the CAM-MR-MS model has shown better classification performance than the other models.

5.3. Ablation experiment

To verify the effectiveness of each module in the CAM-MR-MS model, ablation experiments were conducted on different datasets. Tables 5 and 6 show the experimental results on the self-built dataset and the Ninapro DB5 dataset, respectively. It can be seen that adding each module to the model helps improve the accuracy of gesture recognition classification. This verifies the effectiveness of each module.

6. CONCLUSIONS

The acquisition of sEMG has always been challenging, and it is difficult for non-professionals to operate, which greatly limits the practical application of sEMG-based gesture recognition and hinders the implementation of community rehabilitation and elderly care services at home. Wristband-based sEMG acquisition methods are easy to use and can be operated by non-professionals. However, due to variations

Table 3. Model comparison experiments on self-built datasets with different channel numbers

Model	Accuracy (3 Channels)	Accuracy (8 Channels)
CAM-MR-MS	78.61%	94.14%
SE-CNN	73.35%	86.87%
VMSCNN	72.57%	87.21%
SqueezeNet	70.18%	82.69%

CAM: Channel attention module; MR: multi-channel relationship feature extraction module; MS: multi-scale skip connection module.

Table 4. Model comparison experiments on Ninapro DB5 datasets with different channel numbers

Model	Accuracy (3 Channels)	Accuracy (8 Channels)
CAM-MR-MS	70.82%	87.35%
SE-CNN	68.10%	84.63%
VMSCNN	68.84%	83.55%
SqueezeNet	67.48%	84.10%

CAM: Channel attention module; MR: multi-channel relationship feature extraction module; MS: multi-scale skip connection module.

Table 5. Ablation experiments on self-built datasets with different channel numbers

Model	Accuracy (3 Channels)	Accuracy (8 Channels)
ResNet18	58.26%	63.36%
ResNet18 + CAM	64.95%	72.57%
ResNet18 + MR	65.13%	68.16%
ResNet18 + MS	67.06%	66.95%
ResNet18 + MS + MR	73.31%	81.90%
ResNet18 + MR + CAM	70.47%	81.49%
ResNet18 + MS + CAM	70.63%	77.93%
CAM-MR-MS	78.61%	94.14%

CAM: Channel attention module; MR: multi-channel relationship feature extraction module; MS: multi-scale skip connection module.

Table 6. Ablation experiments on Ninapro DB5 datasets with different channel numbers

Model	Accuracy (3 Channels)	Accuracy (8 Channels)
ResNet18	55.35%	61.46%
ResNet18 + CAM	64.80%	70.80%
ResNet18 + MR	63.45%	67.58%
ResNet18 + MS	64.28%	71.29%
ResNet18 + MS + MR	69.67%	81.76%
ResNet18 + MR + CAM	68.13%	80.39%
ResNet18 + MS + CAM	67.66%	82.17%
CAM-MR-MS	75.94%	87.35%

CAM: Channel attention module; MR: multi-channel relationship feature extraction module; MS: multi-scale skip connection module.

in wearing positions each time, the electrodes may not align properly with the skin surface. This paper proposes an adaptive sEMG channel selection method for gesture recognition, allowing users to pay less attention to the correspondence between electrode plates and skin areas. We initially extract the weights of eight sEMG channels from all subjects, plot heatmaps of muscle regions, and verify the feasibility of our

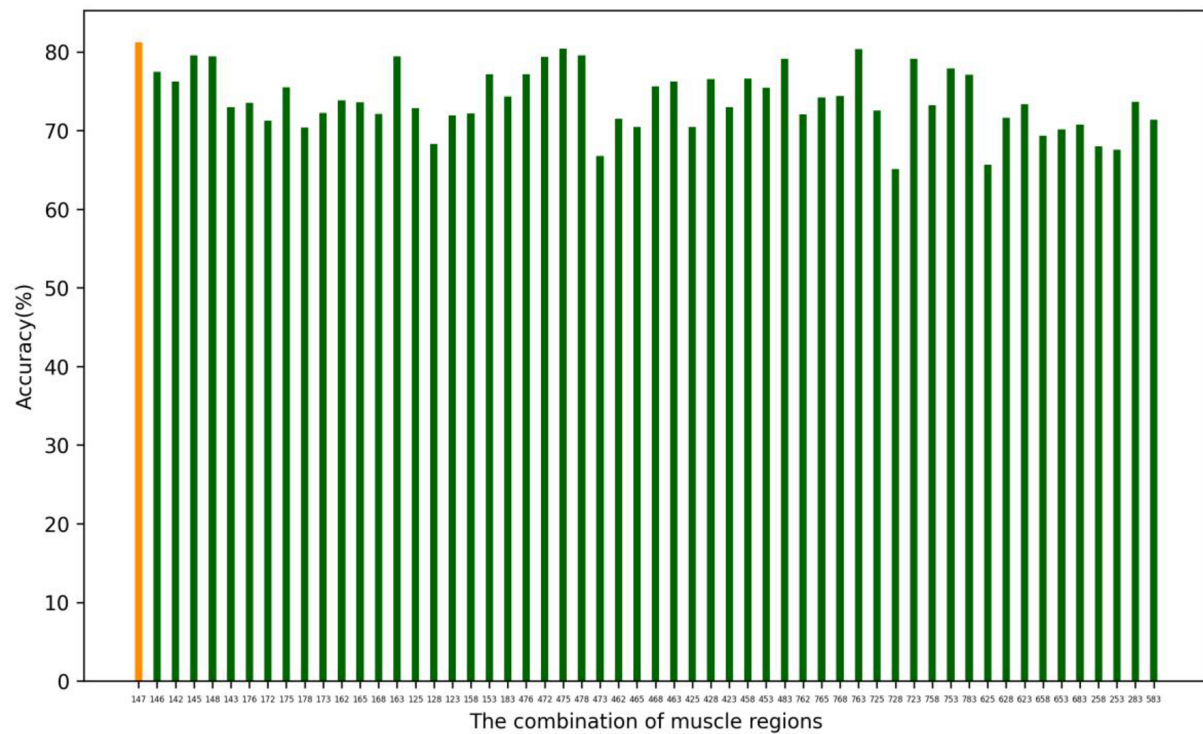


Figure 16. Accuracy of 56 channel combinations for subject with underweight BMI. BMI: Body mass index.

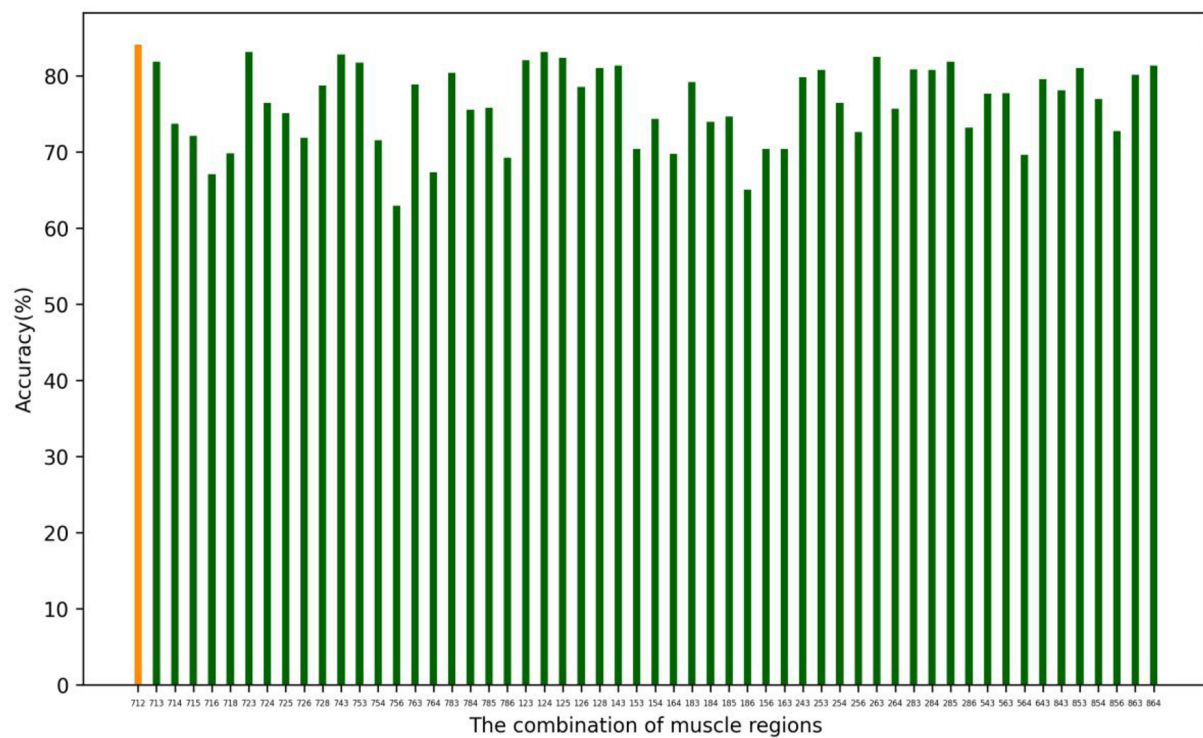


Figure 17. Accuracy of 56 channel combinations for subject with normal BMI. BMI: Body mass index.

channel selection method based on the channel attention mechanism. Subsequently, we plot the relationship between the number of channels and both the average prediction accuracy and average runtime, finding that both the accuracy and runtime of the model tend to increase as the number of channels increases. It is noteworthy that when the number of channels exceeds 3, the slope of prediction accuracy decreases significantly; when the number exceeds 5, the growth in prediction accuracy of the model slows down considerably. This verifies the effectiveness of our channel selection method. Considering both prediction accuracy and model runtime, we select three channels as the optimal channel combination. We then choose two subjects, one with normal BMI and one with underweight BMI, and compare the prediction accuracy of their optimal channel combination with that of the remaining 55 combinations. The results show that the prediction accuracy of the optimal channel combination we extracted is the highest, superior to the other 55 combinations, validating our method's ability to select the best channel combination. Older adults experience muscle loss^[32], and amputee patients may have altered muscle structures^[33]. For these two types of subjects, it is challenging to select muscle regions for gesture recognition. Therefore, the Myo armband can be used to directly collect sEMG signals from around the forearm. By combining our method, we can select skin areas that contribute to gesture recognition tasks, providing a foundation for personalized modeling.

In future work, more daily gesture actions should be considered to enrich the gesture functions of prosthetics in daily life. Additionally, it is necessary to collect sEMG signals from a larger number of subjects, while expanding the age distribution range of the subjects and ensuring a balanced gender ratio. Furthermore, electromyographic signal data from amputee patients should be collected to compare and analyze the differences in gesture recognition performance between healthy subjects and amputee patients. In terms of prosthetics, based on the current offline system, an online system should be designed and developed, with further deployment and optimization of the model to enable the online system to recognize users' intended gesture actions in real time. In terms of the classification network, the structure of the classification network should be optimized to further improve the classification accuracy and reduce resource consumption. For channel selection methods, more deep learning and machine learning techniques should be employed to select channels, with the aim of finding methods that offer superior performance.

DECLARATIONS

Acknowledgments

We express our sincere gratitude to all the subjects who generously participated in our experiment.

Authors' contributions

Made substantial contributions to conception and design of the study and performed data analysis and interpretation: Tong, L.; Li, Y.

Performed data acquisition and material support: Liang, Y.; Wang, C.

Availability of data and materials

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Financial support and sponsorship

This work was supported by the National Natural Science Foundation of China under Grant 62203441, the Beijing National Natural Science Foundation under Grant 4232053, and China University of Mining and Technology, Beijing, Innovation Training Program for undergraduates (No.202413024).

Conflicts of interest

Wang, C. is an Editorial Board Member of the journal *Intelligence & Robotics*. She is not involved in any steps of editorial processing, notably including reviewers' selection, manuscript handling and decision making. The other authors declare that there are no conflicts of interest.

Ethical approval and consent to participate

The Ethics Committees of the Chinese Academy of Sciences (approval number: IA11-2202-31) approved the protocols utilized in our investigation. Written informed consent was obtained from each participant in this study.

Consent for publication

Written informed consent was obtained from the participants.

Copyright

© The Author(s) 2025.

REFERENCES

1. Sabariego, C.; Fellinghauer, C.; Lee, L.; et al. Generating comprehensive functioning and disability data worldwide: development process, data analyses strategy and reliability of the WHO and World Bank Model Disability Survey. *Arch. Public. Health.* **2022**, *80*, 6. DOI PubMed PMC
2. Ersen, M.; Oztop, E.; Sarel, S. Cognition-enabled robot manipulation in human environments: requirements, recent work, and open problems. *IEEE. Robot. Automat. Mag.* **2017**, *24*, 108-22. DOI
3. Sun, Y.; Xu, C.; Li, G.; et al. Intelligent human computer interaction based on non redundant EMG signal. *Alex. Eng. J.* **2020**, *59*, 1149-57. DOI
4. Igual, C.; Pardo, L. A.; Hahne, J. M.; Igual, J. Myoelectric control for upper limb prostheses. *Electronics* **2019**, *8*, 1244. DOI
5. Mane, S. M.; Kambli, R. A.; Kazi, F. S.; Singh, N. M. Hand motion recognition from single channel surface EMG using wavelet & artificial neural network. *Procedia. Comput. Sci.* **2015**, *49*, 58-65. DOI
6. Qu, Y.; Shang, H.; Teng, S. Reduce sEMG channels for hand gesture recognition. In: *2020 3rd IEEE International Conference on Information Communication and Signal Processing (ICICSP)*, Shanghai, China. Sep 12-15, 2020. IEEE, 2020. pp. 215-20. DOI
7. Ma, L.; Zhao, X.; Li, Z.; Zhang, D.; Xu, Z. An optimal channel selection method for EMG signals based on gradient boosting decision tree. *Inf. Control.* **2020**, *49*, 114-21. DOI
8. Wang, Z.; Fang, Y.; Li, G.; Liu, H. Facilitate sEMG-Based human-machine interaction through channel optimization. *Int. J. Human. Robot.* **2019**, *16*, 1941001. DOI
9. Huang, H.; Zhou, P.; Li, G.; Kuiken, T. A. An analysis of EMG electrode configuration for targeted muscle reinnervation based neural machine interface. *IEEE. Trans. Neural. Syst. Rehabil. Eng.* **2008**, *16*, 37-45. DOI PubMed PMC
10. Liu, J.; Li, X.; Li, G.; Zhou, P. EMG feature assessment for myoelectric pattern recognition and channel selection: a study with incomplete spinal cord injury. *Med. Eng. Phys.* **2014**, *36*, 975-80. DOI PubMed PMC
11. Niu, Y.; Chen, W.; Zeng, H.; Gan, Z.; Xiong, B. Optimizing sEMG gesture recognition: leveraging channel selection and feature compression for improved accuracy and computational efficiency. *Appl. Sci.* **2024**, *14*, 3389. DOI
12. Zhang, Z.; Tang, Y.; Zhao, S.; Zhang, X. Real-time surface EMG pattern recognition for hand gestures based on support vector machine. In: *2019 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Dali, China. Dec 06-08, 2019. IEEE, 2019; pp. 1258-62. DOI
13. Lee, K. H.; Min, J. Y.; Byun, S. Electromyogram-based classification of hand and finger gestures using artificial neural networks. *Sensors* **2021**, *22*, 225. DOI PubMed PMC
14. Geng, W.; Du, Y.; Jin, W.; Wei, W.; Hu, Y.; Li, J. Gesture recognition by instantaneous surface EMG images. *Sci. Rep.* **2016**, *6*, 36571. DOI PubMed PMC
15. Zhang, J.; Ling, C.; Li, S. Human movements classification using multi-channel surface EMG signals and deep learning technique. In: *2019 International Conference on Cyberworlds (CW)*, Kyoto, Japan. Oct 02-04, 2019. IEEE, 2019; pp. 267-73. DOI
16. Hu, Y.; Wong, Y.; Wei, W.; Du, Y.; Kankanhalli, M.; Geng, W. A novel attention-based hybrid CNN-RNN architecture for sEMG-based gesture recognition. *PLoS. One.* **2018**, *13*, e0206049. DOI PubMed PMC
17. Wang, L.; Fu, J.; Zheng, B.; Zhao, H. Research on sEMG based gesture recognition using the Attention-based LSTM-CNN with stationary wavelet packet transform. In: *2022 4th International Conference on Advances in Computer Technology, Information Science and Communications (CTISIC)*, Suzhou, China. Apr 22-24, 2022. IEEE, 2022; p. 1-6. DOI
18. Alzahrani, M.; Almalki, M.; Althunayan, T.; Almohawis, A.; Almehaid, F.; Umedani, L. Functional anatomy of the hand: prevalence

- of the linburg–Comstock anomaly in a young saudi population. *J. Musculoskelet. Surg. Res.* **2018**, *2*, 21. DOI
19. Boles, C. A.; Kannam, S.; Cardwell, A. B. The forearm: anatomy of muscle compartments and nerves. *AJR. Am. J. Roentgenol.* **2000**, *174*, 151–9. DOI PubMed
 20. Liu, X.; Zhang, M.; Wang, J.; et al. Gesture recognition of continuous wavelet transform and deep convolution attention network. *Math. Biosci. Eng.* **2023**, *20*, 11139–54. DOI
 21. Tang, D.; Yu, Z.; He, Y.; et al. Strain-insensitive elastic surface electromyographic (sEMG) electrode for efficient recognition of exercise intensities. *Micromachines* **2020**, *11*, 239. DOI PubMed PMC
 22. Gijssberts, A.; Atzori, M.; Castellini, C.; Muller, H.; Caputo, B. Movement error rate for evaluation of machine learning methods for sEMG-based hand movement classification. *IEEE. Trans. Neural. Syst. Rehabil. Eng.* **2014**, *22*, 735–44. DOI PubMed
 23. Wang, M.; Wang, X.; Peng, C.; Zhang, S.; Fan, Z.; Liu, Z. Research on EMG segmentation algorithm and walking analysis based on signal envelope and integral electrical signal. *Photon. Netw. Commun.* **2019**, *37*, 195–203. DOI
 24. Khan, T. I.; Moznuzzaman, M.; Ide, S. Analysis of aging effect on lower limb muscle activity using short time Fourier transform and wavelet decomposition of electromyography signal. *AIP. Adv.* **2023**, *13*, 055011. DOI
 25. Kundu, A. S.; Mazumder, O.; Lenka, P. K.; Bhaumik, S. Hand gesture recognition based omnidirectional wheelchair control using IMU and EMG sensors. *J. Intell. Robot. Syst.* **2018**, *91*, 529–41. DOI
 26. Wang, Q.; Wu, B.; Zhu, P.; et al. ECA-Net: efficient channel attention for deep convolutional neural networks. In: *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, USA. Jun 13–19, 2020. IEEE, 2020. pp. 11534–42. DOI
 27. Woo, S.; Park, J.; Lee, J.; Kweon, I. S. CBAM: convolutional block attention module. *arXiv* **2018**, arXiv:1807.16521. Available online: <https://doi.org/10.48550/arXiv.1807.06521>. [accessed on 1 Apr 2025]
 28. Xu, X.; Jiang, H. A hybrid model based on ResNet and GCN for sEMG-based gesture recognition. *J. Beijing. Inst. Technol.* **2023**, *32*, 219–29. DOI
 29. Xu, Z.; Yu, J.; Xiang, W.; et al. A novel SE-CNN attention architecture for sEMG-based hand gesture recognition. *Comput. Model. Eng. Sci.* **2023**, *134*, 157–77. DOI
 30. Zhang, Y.; Yang, F.; Fan, Q.; Yang, A.; Li, X. Research on sEMG-based gesture recognition by dual-view deep learning. *IEEE. Access.* **2022**, *10*, 32928–37. DOI
 31. Minu, M. S.; Aroul, C. R.; Subashka, R. S. S. Optimal squeeze net with deep neural network-based arial image classification model in unmanned aerial vehicles. *TS.* **2022**, *39*, 275–81. DOI
 32. Sardeli, A. V.; Komatsu, T. R.; Mori, M. A.; Gáspari, A. F.; Chacon-Mikahil, M. P. T. Resistance training prevents muscle loss induced by caloric restriction in obese elderly individuals: a systematic review and meta-analysis. *Nutrients* **2018**, *10*, 423. DOI PubMed PMC
 33. Bramley, J. L.; Worsley, P. R.; Bader, D. L.; et al. Changes in tissue composition and load response after transtibial amputation indicate biomechanical adaptation. *Ann. Biomed. Eng.* **2021**, *49*, 3176–88. DOI PubMed PMC