

Received January 11, 2021, accepted January 29, 2021, date of publication February 2, 2021, date of current version February 9, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3056412

Transfer Learning-Based Muscle Activity Decoding Scheme by Low-frequency sEMG for Wearable Low-cost Application

YURONG LI¹, WENXUAN ZHANG¹, QIAN ZHANG¹, AND NAN ZHENG¹

College of Electrical Engineering and Automation, Fuzhou University, Fuzhou 350108, China

Fujian Key Laboratory of Medical Instrumentation and Pharmaceutical Technology, Fuzhou University, Fuzhou 350108, China

Corresponding author: Yurong Li (liyurong@fzu.edu.cn)

This work was supported in part by the National Nature Science Foundation of China under Grant 61773124, and in part by Fujian Province Nature Science Foundation of China under Grant 2019J01544.

ABSTRACT The surface electromyogram (sEMG) contains a wealth of motion information, which can reflect user's muscle motion intentions. The decoding based on sEMG has been widely used to provide a safe and effective human-computer interaction (HCI) method for neural prosthesis and exoskeleton robot control. The motor intention decoding based on low sampling frequency sEMG may promote the application of wearable low-cost EMG sensors in HCI. Therefore, a motor intention decoding scheme suitable for low frequency EMG signal is proposed in this paper, that is, transfer learning based on Alexnet. Moreover, the effects of different feature extraction methods and data augmentation with Gaussian white noise are fully analyzed. The proposed algorithm is evaluated with the NinaPro database 5. The highest accuracy can reach $70.4\% \pm 4.36\%$ in 53 gestures identification of 10 subjects. Some classical machine learning algorithms such as support vector machine (SVM), linear discriminant analysis (LDA) and K Nearest Neighbor (KNN) are chosen to make comparison, where the SVM with Gaussian kernel function reaches to the maximum accuracy of $67.98\% \pm 4.56\%$. Two-way variance results show significant differences between each other. The experiment results show that the transfer learning is effective for decoding low-frequency sEMG for a large number of gestures.

INDEX TERMS EMG, hand gesture recognition, low-frequency sEMG, machine learning, motor intention decoding.

I. INTRODUCTION

The cerebral cortex generates a movement command called motor intention as soon as the human body performs a certain action. During this process, the α motor neuron receives a nerve impulse transmitted from the brain, which then causes the motor neuron to issue a pulse sequence of a fixed frequency, and forms an action potential sequence in the endplate region. At the same time, the potential is transmitted to the muscle fibers connected to the endplate, and the local voltage formed by the electrochemical form is conducted in two directions on the muscle fibers to form a current. At the moment, this potential change can be detected through the electromyography electrode attached to the skin

The associate editor coordinating the review of this manuscript and approving it for publication was Razi Iqbal¹.

surface, i.e., the surface electromyography (sEMG). It contains motion information related to human intentions in neural signals facilitating to decode human motion intentions. So it is often used for control signals of human-machine interfaces (HMI). Furthermore, as the core technology of HMI, pattern recognition based on sEMG has been widely used in research and clinical applications of autonomous prostheses and exoskeleton [1].

The raw myoelectric signal in each channel is non-stationary, non-linear, stochastic and unpredictable [2]. Thus, the sequence of myoelectric signals which is often transformed into a set of distinctive features is fed into the classifiers for decoding. Currently, the popular feature extraction methods include time domain features, frequency domain features, and time-frequency domain features [3]. These features are related to the strength of muscle movement, the

contribution of each muscle and the firing of motor unit action potentials, realizing by calculating the mean absolute value of the signal, root mean square and Willison amplitude and so on. Hence, is very necessary to reduce the cost of information processing and the complexity of the classifier. During this process, by using some classical algorithms such as support vector machine (SVM) [4], linear discriminant analysis (LDA) [5], K Nearest Neighbor (KNN) [6] and random forests [7] etc., the satisfactory results can be obtained in sEMG pattern recognition [3], [8]. Although these methods are greatly affected by the type of features, there is no feature that can fully stand for pattern of motion. Fortunately, deep learning has been applied to many fields, including computer vision and speech recognition.

As a network with unique ability to learn features from big amounts of raw data, deep learning may be able to mine the muscle information hidden in sEMG at a deeper level to achieve better decoding [9]. Zhai *et al.* [10] developed an auto regressive CNN model to classify sEMG patterns using data from the Ninapro database 2 and database 3 and found that deep networks outperformed the SVM. Xia *et al.* [11] proposed CNN-Recurrent neural network (CNN-RNN) architecture to address the variation in signals overtime. Their results verified that the hybrid CNN-RNN architecture outperformed CNN and support vector regression (SVR). A user-adaptive multilayer CNN algorithm was developed to decode six hand movements in [12], which show that CNN outperformed SVM in both the non-adaptation and adaptation scheme.

However, the researches on the above-mentioned feature extraction and classification strategies are mainly focused on sEMG with high sampling frequency. Yet, with the advancement of wireless communication and embedded technologies, the wearable sensors provide more possibilities for gesture information acquisition. Gesture recognition method based on wearable devices generally uses data gloves and sensors fixed on subject's arm to obtain the sEMG data for analysis and identification, which has better robustness and accuracy [13]. The more convenient sEMG acquisition is very convenient to the application of HCI methods based on sEMG pattern recognition. Specially, Myo armband developed by Thalmic Labs is a typical non-invasive sEMG acquisition method, which is convenient to collect and easy to wear, and provides a more convenient interface for HCI based on human gesture perception. However, the wearable sensor has a lower sampling frequency compared to the myoelectric acquisition device that has been widely used in clinical practice (1000 Hz vs 200 Hz).

Although the pattern recognition methods based on high-frequency sEMG are mature enough, whether these methods are suitable for sEMG below the Nyquist rate (1000 Hz) remains a challenging problem for scholars. It finds from many studies that the classification performance decreases as the sampling rate drops. Phinyomark *et al.* [14] classified the sEMG obtained at 200 Hz and 1000 Hz sampling frequencies using SVM classifier, and the classification

performance dropped significantly when dropping the sampling rate from 1000 Hz to 200 Hz. Li *et al.* [15] also showed that the accuracy decreases significantly when the sampling frequency is lower than 400 Hz. Compared with low-frequency sEMG, high-frequency sEMG contains more motion-related information, but the data processing and calculation burden also increase correspondingly. Therefore, some research focused on low-frequency sEMG decoding.

On one hand, some literatures research the impact of different feature extraction methods on decoding low-frequency sEMG. Zainal Arif *et al.* [16] collected sEMG from the Myo armband with low frequency sampling through extracting five time-domain features of mean absolute value (MAV), variance of EMG (VAR), Willison amplitude (WAMP), waveform length (WL) and zero crossing (ZC). The results verified that MAV and WL are the two features that can best distinguish different actions. Phinyomark *et al.* [14] proposed a novel feature and feature set suitable for low-frequency sEMG decoding. They decoded 12, 17 and 23 gestures, and the optimal results were 89.7%, 83.6%, and 78.9%, respectively. Mendez *et al.* [17] recorded data using Myo armband and they obtained a mean classification accuracy of $91.67\% \pm 6.8\%$ for 9 gestures after extracting 6 time-domain features. Three classic classifiers, including LDA, SVM and KNN, are employed by Chen *et al.* [18] to compare the CNNFeat with 25 traditional features. They focused on analyzing the performance of EMG feature extracted by CNN.

Some literatures effort to decode low-frequency sEMG using classical methods. Pizzolato *et al.* [19] realized classification of 41 gestures collected by Myo armband using SVM and random forests, and obtained the best average accuracy of 69.04%. Amirabdollahian and Walters [20] realized decoding of 4 gestures collected by the Myo armband using SVM with different kernel functions, and achieved a mean accuracy of 94.9%. Combining KNN with dynamic time planning algorithm, Benalcazar *et al.* [21] performed real-time recognition of 5 gestures using myoelectric signals collected by Myo armband, and reached an accuracy of 89.5%. The ideal results can be obtained via the above classical algorithms. In addition, the research on low-frequency sEMG decoding has also extended to the field of deep learning. Wei *et al.* [22] exploited a generative flow model (GFM), which is a recent flourishing branch of deep learning used with a SoftMax classifier for hand-gesture classification, achieved $63.86 \pm 5.12\%$ accuracy in classifying 53 different hand gestures from the NinaPro database 5. Hu *et al.* [23] presented the TL augmented ConvNet and achieved 68.98% accuracy on the NinaPro database 5. Rehman *et al.* [24] collected sEMG with the aid of Myo armband to construct CNN, and to decode six kinds of motion (plus rest), achieved mean accuracies of $97.60\% \pm 1.99\%$ and $98.12\% \pm 1.07\%$ for 7 hand movements with CNN and SSAE-f methods, respectively. Compared with the results obtained by LDA and stacked auto encoders, the accuracy and robustness of CNN are optimal, indicating that the data-driven feature extraction

method overcomes the problem of feature reconstruction and selection in sEMG control. The studies indicate that deep learning technology is a promising EMG control technology, and Myo armband is also promising for myoelectrical control schemes [25].

However, deep network requires a large number of training samples, and it takes a lot of time to collect the sEMG even if the number of people is small. Therefore, the sEMG often has an insufficient sample size. The lack of training samples when using deep learning can easily cause overfitting, which restricts its application in the decoding of EMG signals to a certain extent. More recently, to make use of deep learning methods to decode the motion intent, researchers proposed an effective approach, i.e., transfer learning, to overcome the deficiency of lacking sample size. Cote-Allard *et al.* [25] performed a real-time study with transfer learning based on CNN. They collected sEMG using an eight-channel Myo armband and controlled a 6-DoF robotic arm. Their proposed CNN classified 7 hand movements, achieved offline accuracy of 97.8% for 7 wrist movements, which was slightly better than the baseline CNN.

Pre-processing, net architecture and the optimal hyper parameter selection seem to be fundamental for the analysis of sEMG data with convolutional neural networks, since they can strongly change the final classification accuracy, and time to converge [26]. Some manually tuned hyper-parameters play an important role in the performance of deep learning networks, such as the improvement of recognition accuracy [26], [27]. However, the number of considered movements in the research is typically quite far from the dexterity of a human hand or from what a patient would need in daily life. In addition to this, different from the classification of a few actions, the overlap area of the sEMG in the spatial distribution between different actions will increase as the number of actions increases, making it difficult to distinguish movements [28]. Hence, whether the system is suitable for a large-scale gesture recognition library needs further exploration.

As a consequence, there are several open points in the literatures regarding hand gesture recognition with low sampling devices that can be investigated in more detail. First, most of the studies decode a small number of gestures compared to the number required for activities of daily living [17], [20], [24], [25], thus affecting the possibility to use the results for applications of targeting rehabilitation or assistive robotics. In addition, prosthesis with a large number of degrees can allow achieving more natural and dexterous control [26], [29]–[31]. Second, in the literatures, the decoding of low-frequency sEMG is mostly implemented using classical machine learning, but deep learning is a more promising method, it seems reasonable to investigate its abilities in sEMG as well [26]. Third, what features are beneficial to the proposed low-frequency sEMG decoding framework? In order to clarify the above mentioned open points of low-frequency sEMG decoding, in this paper, an Alexnet-based transfer learning scheme is proposed for decoding the

low-frequency sEMG of multi-gestures. Compared to the existing state of the art, the proposed method has the following advantages:

- 1) The number of classification actions is large for more natural and dexterous control, and the sampling frequency of sEMG is low for wearable low-cost application. The benchmark datasets Ninapro DB5 which records sEMG by the Myo armband is selected to verify the scheme. The number of the movement for classification is larger than 50 and the sEMG sampling frequency is low as 200 Hz.
- 2) Transfer learning method is used to deal with the overfitting problem caused by the lack of sEMG samples while the number of classification actions is large.
- 3) The suitable feature, data enhancement and network optimization are thoroughly discussed within the classification framework to improve the accuracy of classification.

II. MATERIALS AND METHODS

A. DATABASE

The NinaPro DB5 database is used in our study. It is publicly available on Zenodo (<https://zenodo.org/>) which has previously been used for research studies on hand movement recognition and decoding [19]. The DB5 contains sEMG data recordings from 10 intact subjects. The total number of analyzed movements equals to 52, corresponding to the Ninapro exercises A, B, and C plus rest. Each movement is repeated 6 times. DB5 is recorded with double Myo armband which samples 8 sEMG sensors at a 200 Hz frequency. The Thalmic Myo already presents a notch filter at 50Hz and the amplitude of the signal is limited –128 to 127 of the arbitrary Myo unit. We validate DB5 (Double Myo) by verifying that they allow the recognition of hand movements with a transfer learning network.

B. DATA SEGMENTATION

Perform motion segmentation based on the public data set: according to the label which is obtained after relabeling [5], 52 gestures plus rest are separated. Each movement is repeated 6 times of each subject, i.e., there are 6 groups for each movement. The sEMG of a channel is shown in Fig. 1 during the first movement (curved index finger) of the first subject.

It is divided into resting potential and action potential. The sEMG in the dotted rectangular frame are the divided motion signals. In order to avoid the same data used for training and testing, four non-adjacent groups (Groups 1, 3, 4, and 6) are selected for training, and the remaining two groups (Groups 2 and 5) are used for testing.

C. FEATURE SELECTION

As an intuitive form of information expression, images can be used as input samples for deep learning networks. sEMG used for deep learning can be implemented by generating EMG images. There are two common definitions

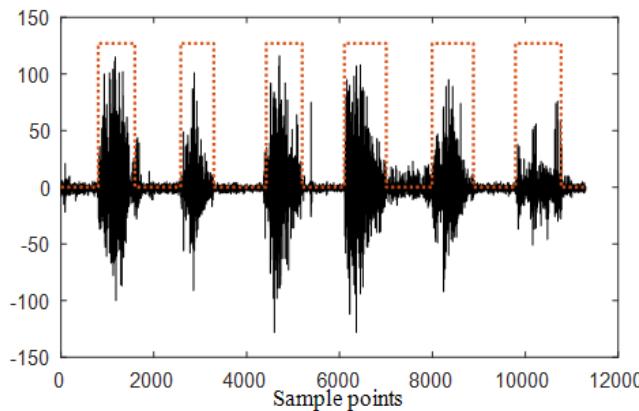


FIGURE 1. One channel sEMG signal of curved index finger of the first subject, and the action potential sEMG signals are inside the dotted rectangular frame.

of EMG images. One is generated directly from the raw sEMG without any additional signals or feature extractors. This method considers that the collected sEMG depicts the potential distribution in space. The heat map is corresponding to the EMG images, and the pixels in the EMG image is determined by the distribution of the electrode array in its acquisition equipment, including the number of electrodes and the distance between the electrodes. Another definition of EMG images is originated in the field of medicine, and is used for gesture recognition by Rojas-Martínez *et al.* [32], [33]. It is a two-dimensional average intensity map of sEMG over time, where each pixel is the root mean square of the sEMG for a channel over a specific time window.

The second definition method is adopted, the input data correspond to time windows of 250 ms, and the overlapping windows 125ms, spanning all the electrode measurements available (16 for the DB5). This choice aims to allow control in real time. The features within a time window are extracted to generate an EMG image. Different from the definition of Rojas *et al.*, in addition to the RMS feature, we also select other time domain features with discrimination [34]: mean absolute value (MAV), difference absolute standard deviation value (DASDV), root mean square (RMS), the mean value of the square root (MSR), integrated EMG (IEMG), waveform length (WL), simple square integral (SSI), variance of EMG (VAR), the maximum fractal length (MFL), zero crossing (ZC), Willison amplitude (WAMP). The extracted data is used to train the classifier, and the features are selected corresponding to the best classification performance.

The data normalization can be scaled according to a certain ratio without changing the original label of the signal, thereby mapping the original data to a new set of data located in a specific interval, and avoiding the phenomenon that the small numbers are overwhelmed in the large numbers. EMG is a kind of fluctuation random signal. Besides, the value of the same action varies on different channels, and the value range obtained after extracting features is different.

In order to avoid channels with small data changes being weakened by channels with large numerical and obtain

reliable results, the data are normalized in this paper to $[-1, 1]$ in the preprocessing [35]. Other than it is also beneficial to improve model convergence speed and calculation accuracy. The normalization method is shown in Equation (1):

$$y = \frac{(y_{max} - y_{min}) \times (x - x_{min})}{(x_{max} - x_{min})} + y_{min} \quad (1)$$

where x is the raw data; x_{max}, x_{min} represent the maximum and minimum values in the raw data respectively; y_{max}, y_{min} are the maximum and minimum values of the normalized range, and the values in this paper are 1 and -1; y is the normalized data.

Taking the extraction of RMS feature as an example, EMG images used the Jet coloring scheme are shown in Fig. 2. The raw sEMG contains 16 channels, so each subfigure is a color image with a size of 4×4 . The ‘imresize’ function in MATLAB can adjust the size of the input image on the premise of automatically maintaining the image aspect ratio. Hence the ‘imresize’ function is needed to convert the image into an RGB image of $227 \times 227 \times 3$ size before it is input into the network, which is in line with the input format of the network.

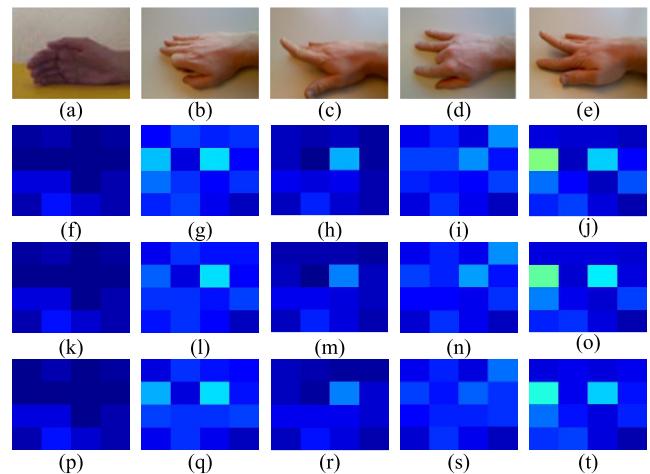


FIGURE 2. sEMG images generated by Jet coloring scheme, each column represents resting, index finger flexion, index finger extension, middle finger flexion, middle finger extension, respectively.

In Fig. 2, the 500ms steady-state sEMG images for 5 movements (i.e., resting, forefinger flexion and extension, middle finger flexion and extension) are selected for visualization. Each column in the picture (a)-(t) of Fig. 2 corresponds to three different stages of the same gesture. It is found that there are obvious distinguishing image features between the images generated by different gestures, which also provides a basis for our use of transfer learning to achieve classification. Yet, specific classification performance needs to rely on classifiers to complete.

D. DATA AUGMENTATION

As for sEMG, images represent muscle activation, which can be used for deep learning. However, the pixels in the sEMG

images indicate the muscle activation of the corresponding channel. After the affine transformation such as rotation and cropping, the label of the examples may be changed. Thus, most of these techniques are unsuitable and cannot be applied directly for sEMG. As such, specific data augmentation techniques must be employed.

For sEMG, the effects of electrode displacement, muscle fatigue, and the actual environment are several factors that need to be considered in the pattern recognition. Data can be realized by simulating the above factors. Since the data used in this paper are collected by two Myo armbands next to each other, during the subject sitting on a comfortable and adjustable seat with the arm naturally on the table. This process effectively reduces the influences of electrode displacement. At the same time, in order to avoid the effect of muscle fatigue on sEMG, there is sufficient rest time between motions. Therefore, the impact of different noises is only considered on the pattern recognition in the environment, and Gaussian white noise is added to achieve data enhancement confirming the original label unchanged.

In addition, sliding window augmentation is implemented which the resulting sequence of examples can be interpreted as the same data point, but shifted through time. For the data with length of L , it was divided into N analysis windows with length of W , and the length of overlapped data is $s=W-D$, then the number of sliding windows obtained is N :

$$N = \frac{L - s}{W - s} \quad (2)$$

A major advantage of this technique is that it does not create any synthetic examples in the dataset compared to the affine transformation employed with images [25]. Also, no new mislabeling occurs with designed sliding window well. Anyway, both sliding window and Gaussian white noise enhancement methods are used.

E. TRANSFER LEARNING ARCHITECTURE

The basic CNN model selected in this paper is the well-known Alexnet network model. Compared with other deep learning models, the convolutional neural network (CNN) can be applied to massive image database due to its strong learning ability and presentation ability. Moreover, CNN includes some beneficial skills, such as ReLU activation function, local response normalization operation, data augmentation and random inactivation to prevent overfitting, which ensure the model performance [13].

Although there are more complex networks with better classification effects, Alexnet with a simpler network structure is selected in order to shorten the training process and to reduce computational loss, which is conducive to the evaluation of methods such as feature selection, network parameters and data enhancement methods. Alexnet consists of 5 convolutional layers, 2 fully connected layers, and 1 classification layer. The network requires input images of size 227-by-227-by-3, but the images have different sizes, so the training images are resized automatically.

The deeper the network, the more training samples are required. Large-scale databases with millions of images such as ImageNet [36] or MNIST [36] are often used to train deep learning networks. However, due to the constraint of financial and time, getting a sufficiently large data set is difficult, so in the field of CNNs, very few researchers train a CNN network from scratch. Training networks with small data sets is prone to overfitting, so the general operation is to train a model on a large data set, and then use the model as an initialization parameter for similar tasks, and fine-tune the pre-trained model through its own training set to get better results, i.e., transfer learning. In this paper, Alexnet-based transfer learning is used. The training set of sEMG images is used to fine-tune the pre-trained Alexnet CNN and then the new input is classified. Fig. 3 describes the structure scheme and the corresponding network parameters of the transfer learning network used in this paper in detail.

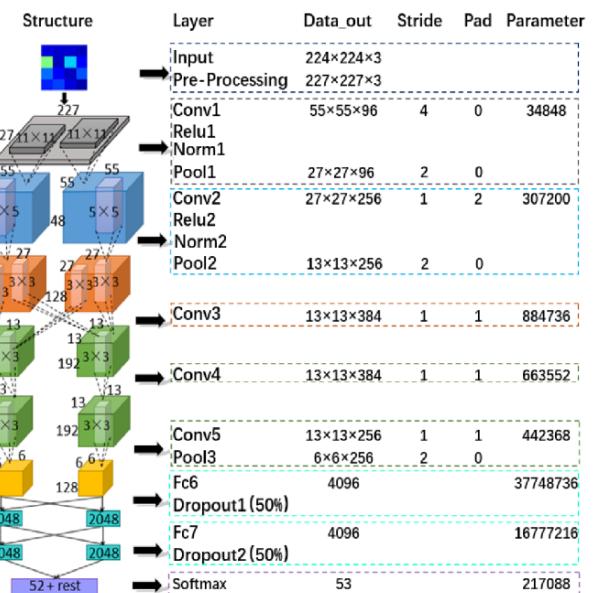


FIGURE 3. Structure scheme of Alexnet-based transfer learning network.

In terms of training time with gradient descent, non-saturating nonlinearity is more advantageous than saturating nonlinear activation functions. Neurons with this nonlinearity is called Rectified Linear Units (ReLUs), which successfully solved the gradient dispersion problem of Sigmoid in relatively deeper network.

$$\text{ReLU}(x) = \max(0, x) \quad (3)$$

ReLUs have the desirable property that they do not require input normalization to prevent them from saturating [38]. Local Response Normalization (LRN) simulated the lateral inhibition mechanism of the biological nervous system, created a competitive mechanism for the activity of local neurons, and improved the model generalization ability. After ReLU activation is completed in the first and second convolutional layers, LRD is applied as shown

in Equation (4):

$$b_{x,y}^i = \frac{a_{x,y}^i}{(k + \alpha \sum_{j=\max(0,i-n/2)}^{\min(N-1,i+n/2)} (a_{x,y}^j)^2)^\beta} \quad (4)$$

where $a_{x,y}^i$ is the activity of a neuron computed by applying kernel i at position (x, y) and $b_{x,y}^i$ is the response-normalized activity. The constants $k=2$, $n=5$, $\alpha = 10^{-4}$, and $\beta = 0.75$ are hyper-parameters whose values are determined using the validation set. The sum runs over n “adjacent” kernel maps at the same spatial position, and N is the total number of kernels in the layer.

Pooling layers in CNNs summarize the outputs of neighboring groups of neurons in the same kernel map. Max-pooling layers follows both response-normalization layers as well as the fifth convolutional layer.

Dropout layer consists of setting to zero the output of each hidden neuron with probability 0.5, which reduces complex co-adaptations of neurons. “Dropped out” in this way do not contribute to the forward pass and do not participate in back propagation [38].

The overall flow chart of transfer learning and related trials are shown in Fig. 4 (a)-(g). First, summon the Alexnet network. Second, fine-tune the last three layers of the pre-trained network, including two fully connected layers and a softmax classification layer, and extract all layers except the last three layers from the pre-trained network.

The Alexnet model used stochastic gradient descent for training with a batch size of 128, momentum of 0.9, and weight decay of 0.0005. The iterative rule for weight w was:

$$v_{i+1} := 0.9 \cdot v_i - 0.0005 \cdot \epsilon \cdot w_i - \epsilon \cdot \langle \frac{\partial L}{\partial w} |_{w_i} \rangle D_i \quad (5)$$

$$w_{i+1} := w_i + v_{i+1} \quad (6)$$

where i is the iteration index, v is the momentum variable, ϵ is the learning rate, and $\langle \frac{\partial L}{\partial w} |_{w_i} \rangle D_i$ is the average over the i th batch D_i of the derivative of the objective with respect to w , evaluated at w_i .

During the training process, many factors that affect network performance need to be considered, such as the selection of learning rate and optimization methods for network. In order to determine the network parameters that are conducive to classification, the sEMG images extracted from the RMS features are borrowed to select the optimal network parameters. Hyper-parameters are identified via random search and manual hyper-parameter tuning [39] on a validation set composed of even numbered subjects selected from DB5. The optimization algorithm can minimize the loss function and make the predicted value as close to the real value as possible. The effects of three optimization algorithms such as ‘Sgdm’, ‘Rmsprop’ and ‘Adam’ are tested. The general framework of the algorithm includes calculating the gradient of the objective function g_t , calculating the first-order momentum function m_t or second-order momentum function v_t , calculating the descent gradient η_t and updating system parameters θ_{t+1} according to the descent gradient.

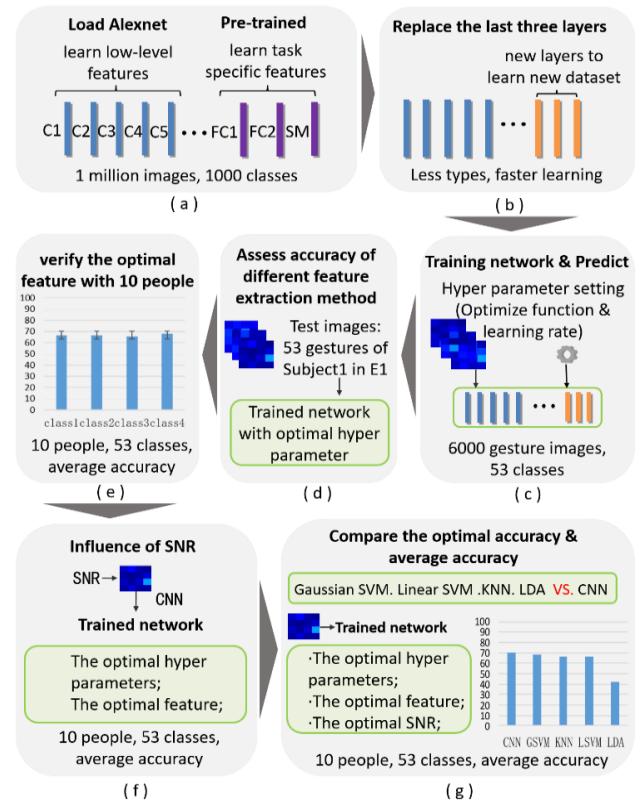


FIGURE 4. The complete flow chart of Alexnet-based transfer learning network. (a) The original Alexnet network structure; (b) The fine-tuned CNN network; (c) Adjust learning rate and optimize function during network training, so that the network can predict under the optimal hyper-parameters; (d) Assess accuracy of different feature extraction methods using the data of E1 object 1; (e) Confirm MRD is the best feature by comparing the average accuracy of 10 subjects; (f) Discuss the influence of different SNR on classification accuracy to find the optimal SNR; (g) Compare the CNN network based on transfer learning with the traditional classical machine learning method by taking the average accuracy of 10 subjects.

Assumed that the gradient of the objective function about the current parameter is:

$$g_t = \nabla J(\theta_t) \quad (7)$$

- 1) Sgdm: The combination of stochastic gradient descent method and first-order momentum method, and its update rules are as followed:

$$m_t = \beta \cdot m_{t-1} + (1 - \beta) \cdot g_t \quad (8)$$

$$\eta_t = \alpha \cdot m_t \quad (9)$$

$$\theta_{t+1} = \theta_t - \eta_t = \theta_t - \alpha \cdot m_t \quad (10)$$

where β is the momentum factor of the first-order momentum function, which is equal to 0.9 in this paper. The first-order momentum optimization algorithm can be used to solve the problem of large updating amplitude of the minimum block stochastic gradient descent optimization algorithm, and can accelerate the convergence speed of the network at the same time.

- 2) Rmsprop: The second-order momentum calculation method, and its update rules are as followed:

$$v_t = \beta \cdot v_{t-1} + (1 - \beta) \cdot g_t^2 \quad (11)$$

$$\eta_t = \alpha \cdot \frac{g_t}{\sqrt{v_t} + \epsilon} \quad (12)$$

$$\theta_{t+1} = \theta_t - \eta_t = \theta_t - \alpha \cdot \frac{g_t}{\sqrt{v_t} + \epsilon} \quad (13)$$

where β is the momentum factor of the second-order momentum function, which is chosen as 0.9 in this paper. This algorithm can reduce the variation range of parameters, decrease the fluctuation of loss function and accelerate the convergence speed of the function during update.

- 3) Adam: The combination of Sgdm and Rmsprop in order to train networks with better performance, and its update rules are as followed:

$$m_t = \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t \quad (14)$$

$$v_t = \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t^2 \quad (15)$$

$$\eta_t = \alpha \cdot \frac{m_t}{\sqrt{v_t} + \epsilon} \quad (16)$$

$$\theta_{t+1} = \theta_t - \eta_t = \theta_t - \alpha \cdot \frac{m_t}{\sqrt{v_t} + \epsilon} \quad (17)$$

where β_1 and β_2 are momentum factors, corresponding to the values in Sgdm and Rmsprop respectively, the values of the two variables are 0.9 and 0.99 respectively; ϵ is a smooth term with a small value 10^{-8} to prevent the denominator from being zero; α is learning rate, which is fine-tuned between 0.01 and 0.00001.

Further, learning rate can directly affect the gradient and parameter update during network training. Only by setting learning rate in an appropriate range can the network show better performance. The fixed and variable learning rate are examined. Using a large learning rate may result in oscillation because of the sparse training set. Therefore, under the condition that the initial state of fixed learning rate is 0.001, 0.0001 and 0.00001 respectively, three optimization algorithms are used to train the network. However, the classifier can be better trained and the accuracy can be improved by decreasing the learning rate step by step with a certain number of iterations. The Alexnet-based transfer learning network is optimized by the Deep Learning Toolbox of the MATLAB R2018 (MathWorks Company, US).

F. CLASSICAL CLASSIFICATION FOR COMPARISON

The used classifiers in this paper are well known, having previously been applied on sEMG in general and thoroughly described on the Ninapro data. They include: SVM with Gaussian kernel, SVM with linear kernel, KNN and LDA. The classification is performed on all the movements included in the database, including rest periods. The same data segmentation and feature extraction methods are implemented. The classical classifiers are trained by the Classification Learner Toolbox of MATLAB R2018 (MathWorks Company, US).

G. EVALUATION METHOD

The classification accuracy is used to evaluate various classification methods, i.e., the number of correctly classified samples divided by the total number of samples, as shown in Equation (2):

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \times 100\% \quad (18)$$

where TP represents the number of true positives which means testing pattern x classified as belonging to pattern x; FP represents the number of false positives which means another patterns classified as pattern x; FN is the number of false negatives which means pattern x classified as belonging to any other patterns; and TN is the number of true negatives which means any other patterns that are not classified as belonging to pattern x.

Further, a two-way analysis of variance (ANOVA) is used to perform statistic analysis. The p values less than 0.05 are considered significant. The ANOVA is implemented by Statistics Toolbox using Matlab programming language on the MATLAB R2018 platform (MathWorks Company).

III. ANALYSIS OF EXPERIMENTAL RESULTS

A. HYPER-PARAMETERS

Several network architectures, pre-processing parameters and hyper-parameters are tested on validation set. When the learning rate is fixed as 0.001, 0.0001 and 0.00001, respectively, three optimization algorithms are used to train the network. The results combining different network training schemes and learning rates are shown in Fig. 5.

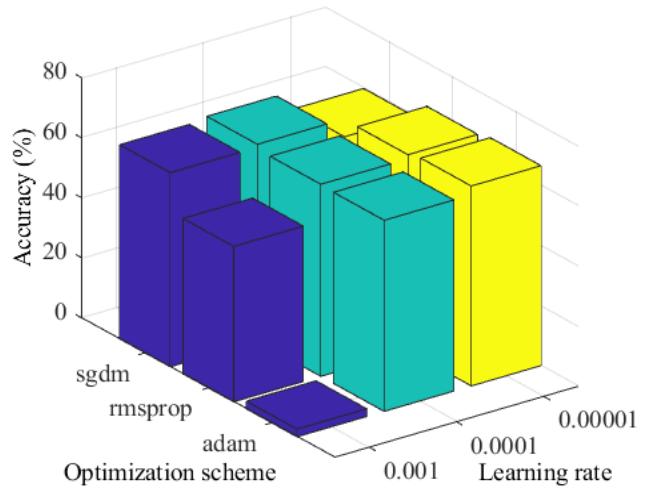


FIGURE 5. Classification accuracy of three optimization methods by using three different learning rates.

As shown in Fig. 5, using the ‘Sgdm’ optimization method, the classification effect is the best with learning rate 0.0001. The classification effect is the best with learning rate 0.00001 as the ‘Rmsprop’ and ‘Adam’ are used. Among these, the highest classification accuracy occurs as the fixed learning rate equals to 0.00001 for the ‘Adam’. However,

a fixed learning rate may cause the parameter vector not stable to a deeper loss function. Reducing the learning rate at an appropriate time can make the parameter vector reach to the best position and improve classification performance. Therefore, a variable learning rate method is further employed.

In fact, in the variable learning rate, the step-by-step attenuation method is applied more widely [40], because it involves fewer parameters and is easier to understand. In order to determine the number of generations of training that begin to decay, the relationship between the number of iterations and the loss function is plotted, as seen in Fig. 6 which shows the situation of subject 2 when the initial learning rate is 0.0001, 0.00001, and 0.000001 at the ‘Sgdm’, ‘Rmsprop’, and ‘Adam’ training modes. When the number of iterations reaches to the fifth generation, the loss function begins to slow down, hence the variable learning rate method is used that changes every five generations to train the classifier.

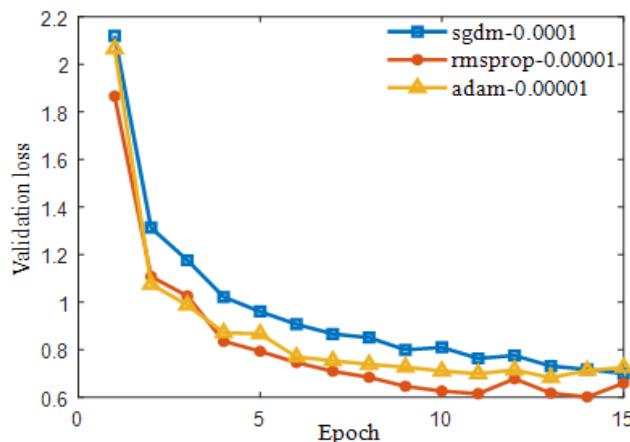


FIGURE 6. Relationship between the number of iterations and the loss of the validation set with different optimization methods.

The classification results when using the fixed learning rate and variable learning rate training scheme are shown in Fig. 7 under the optimal training scheme and initial learning rate mode.

As we can see from Fig. 7, the classification performance is improved using a variable learning rate for each optimization algorithm. In particular, the best results are obtained with ‘Adam’ and a variable learning rate. After several trials, the parameters were set in TABLE 1.

B. EFFECT OF DIFFERENT FEATURE EXTRACTION METHODS ON SEMG DECODING

The 53 gestures of subject 1 in E1 are classified with different feature extraction methods by CNN, and the corresponding results are shown in TABLE 2. CNN is trained with the optimal network parameters.

When the single feature MAV, RMS, and DASDV are extracted, CNN has a superior classification accuracy in TABLE 2. The three features are combined to form the feature set MAV+RMS+DASDV, and the optimal decoding accuracy is obtained, which equals to 78.70%. Consequently,

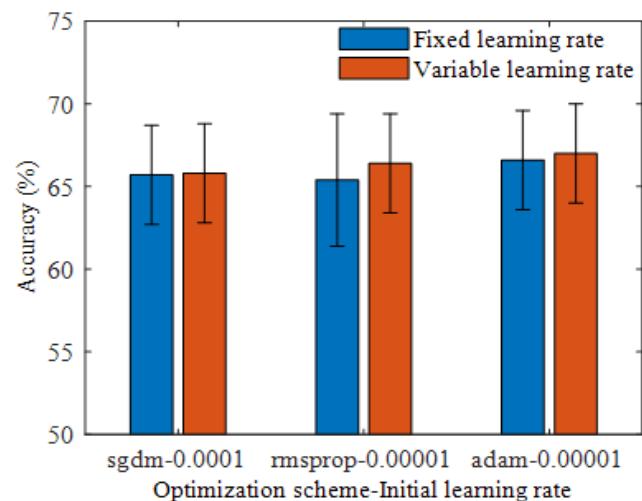


FIGURE 7. Classification accuracy of transfer learning algorithm with different parameter combinations.

TABLE 1. The optimal hyper-parameters setting of transfer learning network.

Parameter type	Parametric variable	Value
Training method	‘adam’	0.99 (momentum)
Variable learning rate	Learning rate drop factor	0.5
	Learning rate drop cycle	5
	Initial learning rate	0.00001
Iteration	Batch size	64
	Maximum iterative number	15
Full-connected layer	Weight attenuation factor	50
	Bias learning factor	50

TABLE 2. Classification results of 53 gestures of subject 1 in different feature extraction methods of transfer learning network.

Feature	Accuracy(%)	Feature	Accuracy(%)
MAV	76.42	WAMP	7.97
RMS	74.80	SSI	73.33
DASDV	77.72	VAR	59.68
MSR	74.76	MFL	7.25
IEMG	23.58	TD	73.15
WL	25.69	MAV+RMS+DASDV	78.70

we refer to the feature set of MAV + RMS + DASDV as MRD feature for short in the following analysis. The four features MAV, RMS, DASDV, and MRD with superior classification effects are further extracted to classify 53 kinds of gestures of 10 subjects. The classification results are shown in TABLE 3.

TABLE 3. Classification results of 53 gestures of 10 subjects in different feature extraction methods of transfer learning network.

Subject \ Feature	MAV	RMS	DASDV	MRD
1	67.419%	61.911%	65.926%	67.120%
2	69.925%	71.387%	68.918%	70.997%
3	74.645%	74.282%	71.662%	74.682%
4	63.032%	64.548%	60.677%	64.774%
5	65.937%	67.541%	71.536%	70.256%
6	65.336%	64.074%	63.135%	63.487%
7	61.822%	62.597%	57.484%	61.233%
8	61.723%	62.611%	60.474%	63.828%
9	71.453%	71.798%	69.693%	72.523%
10	67.128%	67.693%	69.010%	70.333%
Average±Std.	66.84%± 4.01%	66.84%± 4.19%	65.85%± 4.83%	67.92%± 4.23%
p value	0.0978	0.1149	0.0037	--

The maximum accuracy is 67.92% after decoding all 53 actions with MRD features. Statistical analysis of the results obtained from the other three feature extraction methods and the results obtained after extracting the MRD features. We find that there is no significant difference with the cases of single feature MAV and RMS, except for the significant difference between the results of MRD and DASDV. However, considering that the best decoding result is a feature set consisting of three single features, the subsequent research on the effect of different noises on sEMG decoding in the case of extracting the MRD feature set is studied.

C. INFLUENCE OF DIFFERENT NOISES ON SEMG DECODING

In order to increase the robustness of the system, based on the MRD extraction, using overlapping window and different noise combination data augmentation methods, the results for all subjects under different signal-to-noise ratios (SNRs) are shown in TABLE 4.

Seven types of white Gaussian noise are added into the raw sEMG. The results after adding different SNR and the raw sEMG are statistically analyzed. The results show that when the SNR is greater than 5, there is a significant difference of increasing noise, and the classification accuracy reaches to the maximum as the SNR equals to 35. Therefore, both a sliding window and a Gaussian white noise with a SNR of 35 are selected for data enhancement.

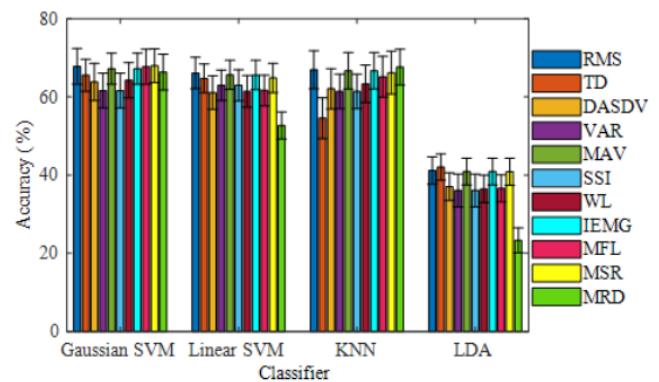
D. INFLUENCE OF DIFFERENT DECODING METHODS ON THE RESULTS OF SEMG DECODING

The classical machine learning methods and different feature extractions are considered to classify 53 gestures on the same

TABLE 4. Influence of different SNR on sEMG decoding.

Feature	SNR	Accuracy (%)	Std (%)	Error (%)	p value
MAV+ RMS+ DASDV	0	68.30	4.11	0	--
	0.5	69.05	4.26	1.13	0.0125
	5	68.60	4.57	0.68	0.2396
	15	69.98	4.41	2.06	0
	25	70.20	3.99	2.29	0
	35	70.40	4.36	2.48	0
	45	69.99	4.22	2.07	0
	55	69.55	4.15	1.63	0.0003

data set, and the results are shown in Fig. 8. Using the same classifier, the classification accuracy is significantly different when different features are extracted. When using Gaussian SVM, Linear SVM, KNN, and LDA for classification, the classification accuracy is the highest when extracting features MSR ($67.98\% \pm 4.56\%$), RMS ($65.93\% \pm 4.08\%$), MRD ($66.89\% \pm 4.87\%$), and TD ($42.06\% \pm 3.59\%$), respectively, but the classification performance is not much different in the case of the same time domain features. However, the optimal feature combination and the number of features in the optimal combination of different classifiers are not always identical. Some feature types were consistently selected more than other features for the different algorithms [3], while some feature combinations show difference in the classification accuracy of different classifiers [8]. Such a result indicates that the classification accuracy of gesture with different classifiers for a single feature are different, the key is whether the extracted features have good separability using the corresponding classifiers.

**FIGURE 8.** Classification accuracy of 53 gestures of 10 subjects in different feature extraction methods of classical machine learning.

The highest classification accuracy occurs in the SVM classifier with Gaussian kernel function within the four classifiers. When the feature is extracted as MSR, it reaches to $67.98\% \pm 4.56\%$. Among the four classifiers, LDA has the

worst classification result. The difference of single feature recognition is the reason why the classification accuracy of MRD features obtained in linear SVM and LDA classifiers is obviously lower than other two classifiers.

Aiming at the multi-gesture classification problem, the Alexnet-based transfer learning method is proposed, and the impact of data enhancement technology on the classification results is analyzed. The effectiveness of algorithm is verified on the Ninapro DB5 dataset. In order to compare the performance difference between the algorithm in this paper and the classical methods, Fig. 9 shows the histograms of the highest decoding results of the five classifiers. The data represent the average and standard deviation obtained by the 10 subjects.

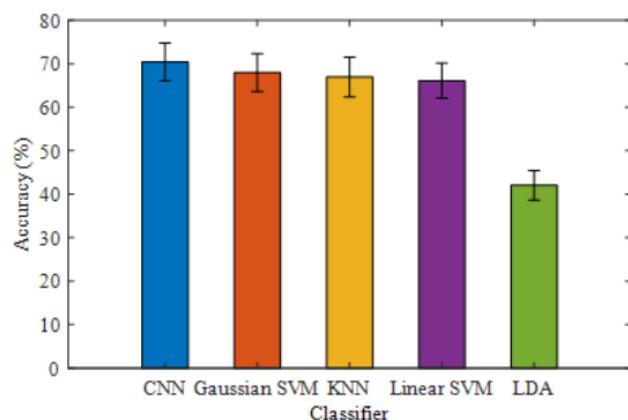


FIGURE 9. Comparison of 5 classifiers classifying 53 gestures of 10 subjects.

The average classification accuracy obtained using the CNN is $70.4\% \pm 4.36\%$, which is higher than that obtained by GFM [22] and ConvNet [23] used NinaPro DB5 for classification. The average classification accuracy obtained using all the classical techniques is $60.71 \pm 11.69\%$. Even if the LDA classifier with the worst result is removed, the average value obtained by the classical machine learning method is $66.93 \pm 4.44\%$, which is still less than the case of CNNs. The best classical classification technique (Gaussian SVM with MSR) obtains an average classification accuracy of $67.98\% \pm 4.56\%$. As represented in TABLE 4, Fig. 8, and Fig. 9, the classification accuracy obtained with CNN using transfer learning based on Alexnet is better than both the average and the highest results obtained from classical classification algorithms. Two-way analysis of variance for different classifiers shows that for different subjects, there is a significant difference between our proposed method and other methods ($p < 0.05$), indicating that the method of transfer learning is also applicable to lower frequency sEMG sampling, and achieves better classification results than traditional machine learning algorithms.

IV. CONCLUSION

Wearable sEMG sensors have been widely used in muscle fatigue detection [40], hand function evaluation [41],

hand motion prediction [42] and prosthetic control [43]. Yet, according to the characteristics of the sEMG, the energy of the sEMG is in the range of 0 - 500 Hz [44]. According to the Shannon sampling theorem, sampling frequency is greater than twice the highest frequency in the frequency spectrum to ensure that the signal is not distorted. Therefore, the sampling frequency of the sEMG sensor is usually in the range of 1000 – 2000 Hz [45], [46]. The Myo armband, as one of the most popular wearable myoelectric acquisition devices, has a sampling frequency of 200Hz. With recent advancements in wearable sensors, scholars tend to focus their research on whether the decoding methods of high frequency EMG signals are suitable for low frequency EMG signals. Some studies have collected EMG data directly from a wearable EMG armband at a low sampling rate, and evaluated the performance of EMG feature extraction methods and classification algorithms [13], [47], [48].

As a kind of machine learning method with strong ability to learn features from big amounts of raw data, deep learning has shown its advantages in the fields of language recognition, image recognition and has been widely used [26]. Although scholars still prefer to utilize classical machine learning methods for decoding low-frequency sEMG, deep learning also gains more and more attention for motion recognition of EMG signals [24], [25]. We find that for deep learning methods, its network performance is susceptible to network parameters and optimization methods, and both effective feature extraction methods and data enhancement methods also can improve the classification performance of the network. However, deep learning requires large samples, and the more complex the network is, the more samples require. The lack of samples is likely to cause overfitting. This may be the main reason that deep learning is rarely used for EMG signal decoding. However, fewer samples can be used to train a network with better performance using transfer learning. Therefore, transfer learning can be used to decode the multi-class problem of low-frequency sEMG.

In this paper, Alexnet-based transfer learning method has been used to construct a network to solve the multi-classification problem of low-frequency sEMG. As mentioned before, 53 hand movements with an average chance level lower than 2% were decoded. Compared to 12, 41 actions [13], [19], our movement has a lower chance of being correctly classified. Moreover, the different time-domain features are extracted, including MAV, DASDV, RMS, MSR, IEMG, WL, WAMP, SSI, VAR, MFL, and TD. It finds that the combination of MAV, RMS, and DASDV features improves the decoding results. Therefore, we extracted the feature set MRD from the sEMG to improve the decoding performance of the decoding framework in this paper. As described in the literature [26], the results are meaningful compared only with sEMG classification problems targeting a similar number of classes. Therefore, Gaussian kernel SVM, Linear kernel SVM, KNN and LDA were used, and the same classification strategy as the CNN was implemented to decode 53 gestures of low-frequency sEMG. The same

training scheme is used to train traditional classification methods and transfer learning algorithms, and the results are compared after classifying the same actions. The best result reaches to 67.98% with Gaussian kernel function SVM classifier, when the extracted features are MSR. Unfortunately, it is still lower than the average accuracy of 70.40% obtained by transfer learning. Two-way variance results show significant differences between each other. However, in actual exercises, myoelectric signals are affected by various factors such as the environment, the position of the limb, and the electrode displacement, etc. [25], [49]. The data used in this paper have been obtained in an ideal experimental environment without considering other factors. Further, the multi-class task decoding problem of low-frequency sEMG should be explored in a real environment. Transfer learning makes use of new task data to fine-tune a deep learning network pre-trained with a large amount of data, however, the currently net architecture that is chosen for transfer learning network is extremely simple. Nevertheless, it should be noted that although CNN performs well, more complex net architectures do exist and can be trained on sEMG data in the future work.

REFERENCES

- [1] K. Li, J. Zhang, L. Wang, M. Zhang, J. Li, and S. Bao, "A review of the key technologies for sEMG-based human-robot interaction systems," *Biomed. Signal Process. Control*, vol. 62, Sep. 2020, Art. no. 102074, doi: [10.1016/j.bspc.2020.102074](https://doi.org/10.1016/j.bspc.2020.102074).
- [2] W. Geng, Y. Du, W. Jin, W. Wei, Y. Hu, and J. Li, "Gesture recognition by instantaneous surface EMG images," *Sci. Rep.*, vol. 6, no. 1, p. 36571, Dec. 2016, doi: [10.1038/srep36571](https://doi.org/10.1038/srep36571).
- [3] J. Camargo and A. Young, "Feature selection and non-linear classifiers: Effects on simultaneous motion recognition in upper limb," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 4, pp. 743–750, Apr. 2019, doi: [10.1109/TNSRE.2019.2903986](https://doi.org/10.1109/TNSRE.2019.2903986).
- [4] A. M. Alex, "An introduction to support vector machines and other Kernel-based learning methods," *Kybernetes*, vol. 30, no. 1, pp. 103–115, 2001, doi: [10.1108/k.2001.30.1.103.6](https://doi.org/10.1108/k.2001.30.1.103.6).
- [5] K. Englehart and B. Hudgins, "A robust, real-time control scheme for multifunction myoelectric control," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 7, pp. 848–854, Jul. 2003, doi: [10.1109/TBME.2003.813539](https://doi.org/10.1109/TBME.2003.813539).
- [6] F. A. Omari, J. Hui, C. Mei, and G. Liu, "Pattern recognition of eight hand motions using feature extraction of forearm EMG signal," *Proc. Nat. Acad. Sci., India A, Phys. Sci.*, vol. 84, no. 3, pp. 473–480, Sep. 2014, doi: [10.1007/s40010-014-0148-2](https://doi.org/10.1007/s40010-014-0148-2).
- [7] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, pp. 5–32, Oct. 2001, doi: [10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324).
- [8] M. Yu, G. Li, D. Jiang, G. Jiang, B. Tao, and D. Chen, "Hand medical monitoring system based on machine learning and optimal EMG feature set," *Pers. Ubiquitous Comput.*, pp. 1–17, Aug. 2019, doi: [10.1007/s00779-019-01285-2](https://doi.org/10.1007/s00779-019-01285-2).
- [9] B. Rim, N.-J. Sung, S. Min, and M. Hong, "Deep learning in physiological signal data: A survey," *Sensors*, vol. 20, no. 4, p. 969, Feb. 2020, doi: [10.3390/s20040969](https://doi.org/10.3390/s20040969).
- [10] X. Zhai, B. Jelfs, R. H. M. Chan, and C. Tin, "Self-recalibrating surface EMG pattern recognition for neuroprosthetic control based on convolutional neural network," *Frontiers Neurosci.*, vol. 11, p. 379, Jul. 2017, doi: [10.3389/fnins.2017.00379](https://doi.org/10.3389/fnins.2017.00379).
- [11] P. Xia, J. Hu, and Y. Peng, "EMG-based estimation of limb movement using deep learning with recurrent convolutional neural networks," *Artif. Organs*, vol. 42, no. 5, pp. E67–E77, May 2018, doi: [10.1111/aor.13004](https://doi.org/10.1111/aor.13004).
- [12] K.-H. Park and S.-W. Lee, "Movement intention decoding based on deep learning for multiuser myoelectric interfaces," in *Proc. 4th Int. Winter Conf. Brain-Comput. Interface (BCI)*, Yongpyong, South Korea, Feb. 2016, pp. 1–2.
- [13] D. Jiang, G. Li, Y. Sun, J. Kong, and B. Tao, "Gesture recognition based on skeletonization algorithm and CNN with ASL database," *Multimedia Tools Appl.*, vol. 78, no. 21, pp. 29953–29970, Nov. 2019, doi: [10.1007/s11042-018-6748-0](https://doi.org/10.1007/s11042-018-6748-0).
- [14] A. Phinayomark, R. N. Khushaba, and E. Scheme, "Feature extraction and selection for myoelectric control based on wearable EMG sensors," *Sensors*, vol. 18, no. 5, p. 1615, May 2018, doi: [10.3390/s18051615](https://doi.org/10.3390/s18051615).
- [15] G. Li, Y. Li, L. Yu, and Y. Geng, "Conditioning and sampling issues of EMG signals in motion recognition of multifunctional myoelectric prostheses," *Ann. Biomed. Eng.*, vol. 39, no. 6, pp. 1779–1787, Jun. 2011, doi: [10.1007/s10439-011-0265-x](https://doi.org/10.1007/s10439-011-0265-x).
- [16] Z. Arief, I. A. Sulistijono, and R. A. Ardiansyah, "Comparison of five time series EMG features extractions using Myo armband," in *Proc. Int. Electron. Symp. (IES)*, Surabaya, Indonesia, Sep. 2015, pp. 11–14.
- [17] I. Mendez, B. W. Hansen, C. M. Grabow, E. J. L. Smedsgaard, N. B. Skogberg, X. J. Uth, A. Bruhn, B. Geng, and E. N. Kamavuako, "Evaluation of the Myo armband for the classification of hand motions," in *Proc. Int. Conf. Rehabil. Robot. (ICORR)*, London, U.K., Jul. 2017, pp. 1211–1214.
- [18] H. Chen, Y. Zhang, G. Li, Y. Fang, and H. Liu, "Surface electromyography feature extraction via convolutional neural network," *Int. J. Mach. Learn. Cybern.*, vol. 11, no. 1, pp. 185–196, 2020, doi: [10.1007/s13042-019-00966-x](https://doi.org/10.1007/s13042-019-00966-x).
- [19] S. Pizzolato, L. Tagliapietra, M. Cognolato, M. Reggiani, H. Müller, and M. Atzori, "Comparison of six electromyography acquisition setups on hand movement classification tasks," *PLoS ONE*, vol. 12, no. 10, Oct. 2017, Art. no. e0186132, doi: [10.1371/journal.pone.0186132](https://doi.org/10.1371/journal.pone.0186132).
- [20] F. Amirabdollahian and M. L. Walters, "Application of support vector machines in detecting hand grasp gestures using a commercially off the shelf wireless myoelectric armband," in *Proc. Int. Conf. Rehabil. Robot. (ICORR)*, London, U.K., Jul. 2017, pp. 111–115.
- [21] M. E. Benalcazar, C. Motoche, J. A. Zea, A. G. Jaramillo, C. E. Anchundia, P. Zambrano, M. Segura, F. B. Palacios, and M. Perez, "Real-time hand gesture recognition using the Myo armband and muscle activity detection," in *Proc. IEEE 2nd Ecuador Tech. Chapters Meeting (ETCM)*, Salinas, Ecuador, Oct. 2017, pp. 1–6.
- [22] W. Wei, Q. Dai, Y. Wong, Y. Hu, M. Kankanhalli, and W. Geng, "Surface-electromyography-based gesture recognition by multi-view deep learning," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 10, pp. 2964–2973, Oct. 2019, doi: [10.1109/TBME.2019.289922](https://doi.org/10.1109/TBME.2019.289922).
- [23] Y. Hu, Y. Wong, W. Wei, Y. Du, M. Kankanhalli, and W. Geng, "A novel attention-based hybrid CNN-RNN architecture for sEMG-based gesture recognition," *PLoS ONE*, vol. 13, no. 10, Oct. 2018, Art. no. e0206049, doi: [10.1371/journal.pone.0206049](https://doi.org/10.1371/journal.pone.0206049).
- [24] M. Z. U. Rehman, A. Waris, S. Gilani, M. Jochumsen, I. Niazi, M. Jamil, D. Farina, and E. Kamavuako, "Multiday EMG-based classification of hand motions with deep learning techniques," *Sensors*, vol. 18, no. 8, p. 2497, Aug. 2018, doi: [10.3390/s18082497](https://doi.org/10.3390/s18082497).
- [25] U. Cote-Allard, C. L. Fall, A. Drouin, A. Campeau-Lecours, C. Gosselin, K. Glette, F. Laviolette, and B. Gosselin, "Deep learning for electromyographic hand gesture signal classification using transfer learning," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 4, pp. 760–771, Apr. 2019, doi: [10.1109/TNSRE.2019.2896269](https://doi.org/10.1109/TNSRE.2019.2896269).
- [26] M. Atzori, M. Cognolato, and H. Müller, "Deep learning with convolutional neural networks applied to electromyography data: A resource for the classification of movements for prosthetic hands," *Frontiers Neuro-robot.*, vol. 10, no. 9, pp. 1–10, Sep. 2016, doi: [10.3389/fnbot.2016.00009](https://doi.org/10.3389/fnbot.2016.00009).
- [27] L. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep Learning*, vol. 1. Cambridge, U.K.: MIT Press, 2016.
- [28] M. Atzori, A. Gijsberts, S. Heynen, A.-G.-M. Hager, O. Deriaz, P. van der Smagt, C. Castellini, B. Caputo, and H. Müller, "Building the ninapro database: A resource for the biorobotics community," in *Proc. 4th IEEE RAS EMBS Int. Conf. Biomed. Robot. Biomechatronics (BioRob)*, Rome, Italy, Jun. 2012, pp. 1258–1265.
- [29] A. Fougnier, O. Stavdahl, P. J. Kyberd, Y. G. Losier, and P. A. Parker, "Control of upper limb prostheses: Terminology and proportional myoelectric control—A review," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 20, no. 5, pp. 663–677, Sep. 2012, doi: [10.1109/TNSRE.2012.2196711](https://doi.org/10.1109/TNSRE.2012.2196711).
- [30] D. Farina, N. Jiang, H. Rehbaum, A. Holobar, B. Graimann, H. Dietl, and O. C. Aszmann, "The extraction of neural information from the surface EMG for the control of upper-limb prostheses: Emerging avenues and challenges," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 4, pp. 797–809, Jul. 2014, doi: [10.1109/TNSRE.2014.2305111](https://doi.org/10.1109/TNSRE.2014.2305111).

- [31] N. J. Jarque-Bou, A. Scano, M. Atzori, and H. Müller, "Kinematic synergies of hand grasps: A comprehensive study on a large publicly available dataset," *J. Neuroeng. Rehabil.*, vol. 16, no. 1, p. 63, Dec. 2019, doi: [10.1186/s12984-019-0536-6](https://doi.org/10.1186/s12984-019-0536-6).
- [32] M. Rojas-Martínez, M. A. Mañanas, and J. F. Alonso, "High-density surface EMG maps from upper-arm and forearm muscles," *J. Neuroeng. Rehabil.*, vol. 9, no. 1, p. 85, Dec. 2012, doi: [10.1186/1743-0003-9-85](https://doi.org/10.1186/1743-0003-9-85).
- [33] M. Rojas-Martínez, M. A. Mañanas, J. F. Alonso, and R. Merletti, "Identification of isometric contractions based on high density EMG maps," *J. Electromyogr. Kinesiol.*, vol. 23, no. 1, pp. 33–42, Feb. 2013, doi: [10.1016/j.jelekin.2012.06.009](https://doi.org/10.1016/j.jelekin.2012.06.009).
- [34] A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "Feature reduction and selection for EMG signal classification," *Expert Syst. Appl.*, vol. 39, no. 8, pp. 7420–7431, Jun. 2012, doi: [10.1016/j.eswa.2012.01.102](https://doi.org/10.1016/j.eswa.2012.01.102).
- [35] H. Cao, S. Sun, and K. Zhang, "Modified EMG-based handgrip force prediction using extreme learning machine," *Soft Comput. Fusion Found. Methodol. Appl.*, vol. 21, no. 2, pp. 491–500, Jan. 2017, doi: [10.1007/s00500-015-1800-8](https://doi.org/10.1007/s00500-015-1800-8).
- [36] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Miami Beach, FL, USA, Jun. 2009, pp. 248–255, doi: [10.1109/cvpr.2009.5206848](https://doi.org/10.1109/cvpr.2009.5206848).
- [37] L. Deng, "The MNIST database of handwritten digit images for machine learning research [best of the Web]," *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 141–142, Nov. 2012, doi: [10.1109/MSP.2012.2211477](https://doi.org/10.1109/MSP.2012.2211477).
- [38] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017, doi: [10.1145/3065386](https://doi.org/10.1145/3065386).
- [39] R. Ge, S. M. Kakade, R. Kidambi, and P. Netrapalli, "The step decay schedule: A near optimal, geometrically decaying learning rate procedure for least squares," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, vol. 32, Vancouver, BC, Canada, Dec. 2019, pp. 1–12.
- [40] M. Montoya, O. Henao, and J. Muñoz, "Muscle fatigue detection through wearable sensors: A comparative study using the Myo armband," in *Proc. XVIII Int. Conf. Hum. Comput. Interact.*, Cancún, Mexico, Sep. 2017, p. 30, doi: [10.1145/3123818.3123855](https://doi.org/10.1145/3123818.3123855).
- [41] V. Montoya-Leal, A. Orozco-Duque, J. P. Ugarte, M. A. Portela, J. C. Franco, and V. Z. Perez, "Assessment protocol of wrist flexion and extension to support processes in occupational health using Myo Armband," in *Proc. VII Latin Amer. Congr. Biomed. Eng. (CLAIB)*, vol. 60, Bucaramanga, Colombia, Oct. 2017, pp. 585–588.
- [42] H. J. Kim, Y. S. Lee, and D. Kim, "Arm motion estimation algorithm using MYO armband," in *Proc. 1st IEEE Int. Conf. Robot. Comput. (IRC)*, Taichung, Taiwan, Apr. 2017, pp. 376–381, doi: [10.1109/IRC.2017.32](https://doi.org/10.1109/IRC.2017.32).
- [43] C. Igual, J. Igual, J. M. Hahne, and L. C. Parra, "Adaptive auto-regressive proportional myoelectric control," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 2, pp. 314–322, Feb. 2019, doi: [10.1109/TNSRE.2019.2894464](https://doi.org/10.1109/TNSRE.2019.2894464).
- [44] J.-U. Chu, I. Moon, Y.-J. Lee, S.-K. Kim, and M.-S. Mun, "A supervised Feature-projection-based real-time EMG pattern recognition for multifunction myoelectric hand control," *IEEE/ASME Trans. Mechatronics*, vol. 12, no. 3, pp. 282–290, Jun. 2007, doi: [10.1109/TMECH.2007.897262](https://doi.org/10.1109/TMECH.2007.897262).
- [45] E. A. Clancy, E. L. Morin, and R. Merletti, "Sampling, noise-reduction and amplitude estimation issues in surface electromyography," *J. Electromyogr. Kinesiol.*, vol. 12, no. 1, pp. 1–16, 2002, doi: [10.1016/S1050-6411\(01\)00033-5](https://doi.org/10.1016/S1050-6411(01)00033-5).
- [46] J.-U. Chu, I. Moon, and M.-S. Mun, "A real-time EMG pattern recognition system based on linear-nonlinear feature projection for a multi-function myoelectric hand," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 11, pp. 2232–2239, Nov. 2006, doi: [10.1109/TBME.2006.883695](https://doi.org/10.1109/TBME.2006.883695).
- [47] A. Phinyomark and E. Scheme, "A feature extraction issue for myoelectric control based on wearable EMG sensors," in *Proc. IEEE Sensors Appl. Symp. (SAS)*, Seoul, South Korea, Mar. 2018, pp. 232–237.
- [48] K. Muzaffar, K. S. Jai, and T. Mukesh, "Review of EMG signal classification for diagnosis," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 5, no. 12, p. 6, 2015.
- [49] G. Kanitz, C. Cipriani, and B. B. Edin, "Classification of transient myoelectric signals for the control of multi-grasp hand prostheses," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 9, pp. 1756–1764, Sep. 2018, doi: [10.1109/TNSRE.2018.2861465](https://doi.org/10.1109/TNSRE.2018.2861465).



YURONG LI received the master's degree in industry automation and the Ph.D. degree in control theory and control engineering from Zhejiang University, in 1997 and 2001, respectively. She is currently a Professor with Fuzhou University. Since 2007, she has been a member of the Fujian Key Laboratory of Medical Instrumentation and Pharmaceutical Technology. Her research interests include biomedical instrument and intelligent information processing.



WENXUAN ZHANG received the B.E. degree in automation from the College of Electrical Engineering and Automation, University of Jinan, Jinan, China, in 2019. She is currently pursuing the master's degree in control science and engineering with Fuzhou University. Her research interests include computational neuroscience and intelligent information processing.



QIAN ZHANG received the bachelor's degree in automation from the School of Electrical Engineering and Automation, Henan Polytechnic University, in 2017, and the master's degree in testing technology and automation equipment from Fuzhou University, in 2020. Her research interests include computational neuroscience, neuromusculoskeletal modeling, and intelligent information processing.



NAN ZHENG received the B.E. degree in automation from the College of Computer and Electronic Information, Guangdong University of Petroleum Technology, Maoming, China, in 2019. He is currently pursuing the master's degree in control science and engineering with Fuzhou University. His research interests include pattern recognition and feature extraction of EMG signals.