IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received July 10, 2019, accepted July 19, 2019, date of publication July 25, 2019, date of current version August 27, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2930958

A Deep Transfer Convolutional Neural Network Framework for EEG Signal Classification

GAOWEI XU[1], XIAOANG SHEN[1], SIRUI CHEN[1], YONGSHUO ZONG[1], CANYANG ZHANG[1], HONGYANG YUE[1], MIN LIU[1], (Member, IEEE), FEI CHEN[2], AND WENLIANG CHE[3]

[1]School of Electrical Information and Engineering, Tongji University, Shanghai 201804, China
[2]Department of Cardiology, Shanghai Tongji Hospital, Tongji University, Shanghai 200065, China
[3]Department of Cardiology, Shanghai Tenth People's Hospital, Tongji University School of Medicine, Shanghai 200072, China

Corresponding authors: Min Liu (lmin@tongji.edu.cn), Fei Chen (riverapt@126.com), and Wenliang Che (chewenliang@tongji.edu.cn)

This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 61573257 and Grant 71690234, and in part by the China Scholarship Council.

ABSTRACT Nowadays, motor imagery (MI) electroencephalogram (EEG) signal classification has become a hotspot in the research field of brain computer interface (BCI). More recently, deep learning has emerged as a promising technique to automatically extract features of raw MI EEG signals and then classify them. However, deep learning-based methods still face two challenging problems in practical MI EEG signal classification applications: (1) Generally, training a deep learning model successfully needs a large amount of labeled data. However, most of the EEG signal data is unlabeled and it is quite difficult or even impossible for human experts to label all the signal samples manually. (2) It is extremely time-consuming and computationally expensive to train a deep learning model from scratch. To cope with these two challenges, a deep transfer convolutional neural network (CNN) framework based on VGG-16 is proposed for EEG signal classification. The proposed framework consists of a VGG-16 CNN model pre-trained on the ImageNet and a target CNN model which shares the same structure with VGG-16 except for the softmax output layer. The parameters of the pre-trained VGG-16 CNN model are directly transferred to the target CNN model used for MI EEG signal classification. Then, front-layers parameters in the target model are frozen, while later-layers parameters are fine-tuned by the target MI dataset. The target dataset is composed of time-frequency spectrum images of EEG signals. The performance of the proposed framework is verified on the public benchmark dataset 2b from the BCI competition IV. The experimental results show that the proposed framework improves the accuracy and efficiency performance of EEG signal classification compared with traditional methods, including support vector machine (SVM), artificial neural network (ANN), and standard CNN.

INDEX TERMS Motor imagery (MI), electroencephalogram (EEG), signal classification, short time Fourier transform (STFT), VGG-16, transfer learning.

I. INTRODUCTION

Brain computer interface (BCI), also known as brain-machine interface (BMI), enables human brains to directly communicate with external computers or machines. Generally, researches on human BCIs mainly include four aspects: invasive BCIs, partially invasive BCIs, non-invasive BCIs and synthetic telepathy [1]–[3]. Despite there being so many BCI systems and techniques, non-invasive BCI through electroencephalography (EEG) has been the most extensively studied, because it is relatively cheap and simple to carry and use, but provides fine temporal resolution. Motor imagery (MI), as one of the most popular EEG-based BCI applications, can help the disabled and elderly people to perform a specific task through the imagination without physically performing any limb movements [4]. More specifically, in a MI BCI system, MI EEG signals indicating the people's intention or imagination are collected by electrodes placed on the scalp, then the BCI system processes and translates the collected signals into the commands to control external devices [5], [6].

The design of a typical EEG-based MI BCI system is shown in Fig. 1 [1]. Traditionally, the procedure of the MI BCI system mainly consists of five phases, i.e., signal data acquisition, data preprocessing, feature extraction, feature

The associate editor coordinating the review of this article and approving it for publication was Yongqiang Cheng.

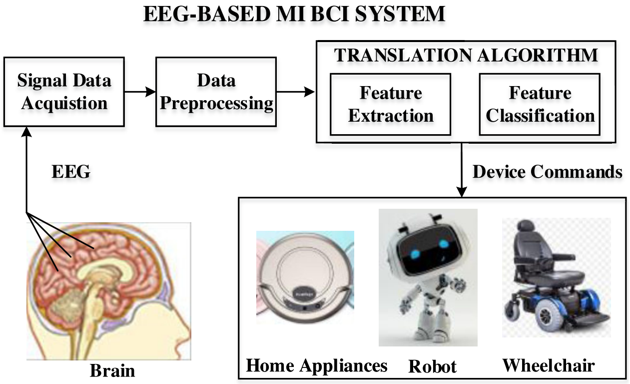


FIGURE 1. The design of a typical EEG-based MI BCI system.

classification and device control interface [7]. The data acquisition phase includes MI EEG signal collection, signal digitalization and data storage. The data preprocessing phase involves data filtering, data cleaning, data transformation and etc. The feature extraction phase extracts discriminative features containing useful information from EEG signal data. The feature classification phase utilizes the extracted features as input to train machine learning models. The trained models can classify different signals and MI tasks. Finally, in the device control interface phase, the categorized signals are translated into commands to control devices, such as robots, home appliances, wheelchairs [8]–[13].

During the past few years, many studies have focused on the feature extraction and classification phases since they can greatly impact the MI BCI system performance [14]–[16]. Common spatial patterns (CSP) algorithm is a classical and powerful method to extract discriminative features of raw EEG signals. A number of variants of the CSP algorithm have been developed and applied in many MI BCI problems [17], [18]. For instance, Aghaei *et al.* [18] designed a separable common spatio-spectral patterns algorithm for MI BCI and the designed algorithm required much lower computing resources. Ashok *et al.* [19] proposed two weighted CSP methods for MI task classification and achieved more accurate classification. Ang *et al.* [20] proposed a filter bank CSP algorithm and it obtained better MI signal classification results. In addition, many time-frequency signal processing methods, such as short time Fourier transform (STFT), empirical mode decomposition (EMD), continuous wavelet transform (CWT) [21]–[23], are also applied in the feature extraction phase. Tabar and Halici [21] used STFT to extract time, frequency and location information from raw EEG signals and convert signals to images at the same time. Lee and Choi [23] converted EEG signals to time-frequency spectrums with CWT. Sometimes, a feature selection process is added to get rid of redundant information in the extracted features in order to improve the computational efficiency. Linear discriminant analysis (LDA), artificial neural network (ANN), support vector machine (SVM) and other machine learning classifiers are frequently used in the feature classification phase [24]–[26]. Naseer and Hong [24] applied the LDA classifier to MI tasks based on two distinct features. Siuly and Li [25] designed a least square SVM method to classify two-class MI signals. However, these traditional methods rely heavily on the prior knowledge and expertise of EEG signal processing.

With the advent of deep learning, it has replaced the traditional methods due to its automatic feature extraction ability [27]. Recently, many deep learning models-based EEG signal classification methods have demonstrated superior performance compared to traditional methods [21], [28]. CNN, as one of the most widely-used deep learning models, is always combined with the extracted features of EEG signal data to provide an improved classification result. Tabar and Halici [21] applied the CNN model to classify MI EEG signals and obtained 9% accuracy improvement compared to traditional methods. Yang *et al.* [28] proposed a multi-class MI EEG signal classification method based on augmented CSP features and CNN, the accuracy is high up to 69.27%. Although CNN has made considerable progress in the research field of EEG signal classification, there are still two limitations with these CNN-based methods. First, it has been proven that very deep CNN architectures with smaller convolution filters is beneficial for the classification accuracy in many research areas, including image recognition, natural language processing, biology, and etc., [29]. However, few studies have investigated the applications of very deep CNN architectures in EEG signal classification. This is because training a very deep CNN from scratch for EEG signal classification requires a large number of labeled EEG data. However, most of the collected signal data is unlabeled in practical applications. Labeling these data manually is very difficult or even impossible. Additionally, a small amount of data is very likely to cause over-fitting problem. Next, training a very deep CNN model from scratch successfully is very time-consuming and computationally expensive.

Transfer learning has emerged as an effective approach to overcome the above-mentioned limitations, and it is often expressed through the use of pre-trained models. A pre-trained CNN model can leverage the knowledge gained from a large dataset to solve a different but similar task with a small dataset more effectively and efficiently. Because the CNN model pre-trained on the benchmark dataset like ImageNet can extract universal low-level features, which are useful for most of image classification problems [30]. Instead of training the CNN model from scratch, fine-tuning the pre-trained CNN model on the target dataset can significantly reduce the training time and save the computing resources. A number of famous pre-trained CNN models including VGGNet, GoogleNet, ResNet and so on have been designed and applied to various domains. For example, Shao *et al.* [31] developed a deep transfer CNN for fault diagnosis, a pre-trained CNN model is used to accelerate the training process. Rahhal *et al.* [32] presented a pre-trained CNN-based transfer learning approach for electrocardiogram classification, the experiments conducted on three benchmark datasets proved the effectiveness of the presented method. Shi *et al.* [33] proposed a deep CNN-based transfer learning method for false positive reduction, all the pre-trained layers are transferred to target network and only the last fully connected layer is fine-tuned for the pulmonary nodule classification task.

In this paper, a transfer CNN framework based on VGG-16 is proposed for MI EEG signal classification. It consists of a pre-trained CNN model and a target CNN model used for MI classification. First, raw EEG signals from C3, C4 and Cz electrodes are converted to time-frequency spectrum images using STFT. Then, a VGG-16 CNN model is pre-trained on the ImageNet. Next, the structure and parameters of the pre-trained VGG-16 CNN model are transferred to the target CNN model. Finally, the target CNN model is fine-tuned on the target dataset for MI classification. In this paper, the target dataset is composed of the converted time-frequency spectrum images of EEG signals. Since the target dataset used for fine-tuning is not similar to the ImageNet dataset used for pre-training, only the front-layers are frozen and the later-layers of the target CNN are fine-tuned.

The rest of this paper is organized as follows. Section II briefly introduces the used dataset 2b from the BCI competition IV. Section III describes the conversion process of EEG signal data to time-frequency spectrums and presents the proposed transfer CNN framework for EEG signal classification. Section IV conducts experiments on the dataset 2b and analyzes results. Finally, Section V concludes the paper and discusses our future work.

II. DATASET DESCRIPTION

In this paper, the dataset 2b from the BCI competition IV is used to evaluate the superior performance of the proposed framework. The dataset 2b is composed of EEG signal data collected from nine subjects [34]. EEG signal data is collected under the sampling frequency of 250 Hz with three electrodes (C3, Cz and C4). A band-pass filter that only passes frequencies between 0.5 Hz and 100 Hz is employed to eliminate the signal noise. All the subjects are required to perform two different types of MI tasks, including the imaginations of left hand movement and right hand movement. There are five sessions for each subject, where the EEG signal data in the first two sessions are collected without feedback, while the EEG signal data in the remaining three sessions are collected with feedback. The timing scheme of each trail is shown in Fig. 2 [34]. Taking the first two sessions for example, the detailed timing scheme is illustrated as follows: a fixation cross appears on the screen at the start of each trail ($t = 0s$). Then, a cue in the form of arrow indicating different MI tasks appears on the screen from $t = 3s$ to $t = 4.25s$. After that, the subjects need to perform the corresponding MI tasks depending on the arrow direction from $t = 4s$ to $t = 7s$.

III. PROPOSED DEEP TRANSFER CNN FRAMEWORK FOR EEG SIGNAL CLASSIFICATION

In this section, firstly, a data preprocessing method based on STFT is proposed to convert EEG signals to a set of time-frequency spectrum images. Then, a deep transfer CNN framework is proposed for EEG signal classification, which takes the obtained images as the input. The proposed framework can achieve highly accurate and efficient EEG signal classification with the introduction of transfer learning.

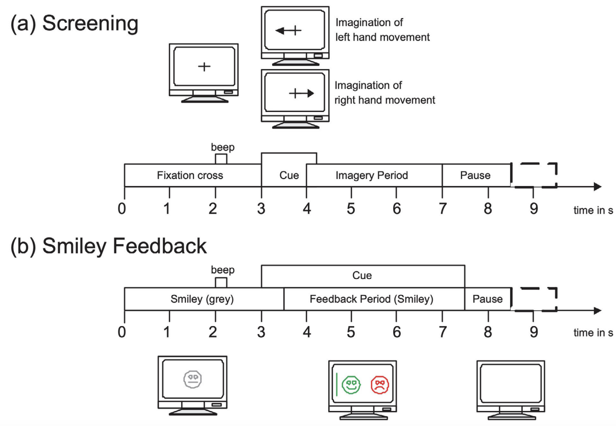


FIGURE 2. Timing scheme of each trail: (a) The first two sessions and (b) the remaining three sessions.

A. DATA PREPROCESSING BASED ON STFT

Short time Fourier transform (STFT) is a Fourier-based time-frequency signal analysis and processing method and it is able to calculate the complex amplitude over time and frequency of a non-stationary signal. For STFT, the EEG signal $x(t)$ to be processed is multiplied by a short time window which slides along the time axis. Then, a set of windowed signal segments is generated. Finally, the Fourier transform is applied to each windowed signal segment respectively, generating two-dimensional time-frequency spectrums of the raw signal. Mathematically, STFT is defined as follows:

$$STFT(\tau, \omega) = \int_{-\infty}^{+\infty} x(t)w(t-\tau)e^{-j\omega t}dt \qquad (1)$$

where $w(t)$ is a window with a limited number of nonzero points, and τ is the window position on the time axis. The advantage of STFT is that it can extract time-domain features and frequency-domain features embedded in original signals simultaneously.

It was examined in [21] and [35] that the EEG signals from the C3, C4 and Cz electrodes can be obviously affected by performing MI tasks. More specifically, during MI tasks, there is an amplitude decrease in the mu band (8-13 Hz) of these EEG signals, which is named event related desynchronization (ERD). Oppositely, there is also an amplitude increase in the beta band (13-30 Hz), which is named event related synchronization (ERS). In view of this, STFT is performed on the signal recordings from C3, C4 and Cz electrodes and only the mu and beta frequency bands information of the obtained time-frequency spectrums is preserved. Besides, in this paper, 4-14 Hz frequency bands are used to represent the mu band and 16-32 Hz frequency bands are used

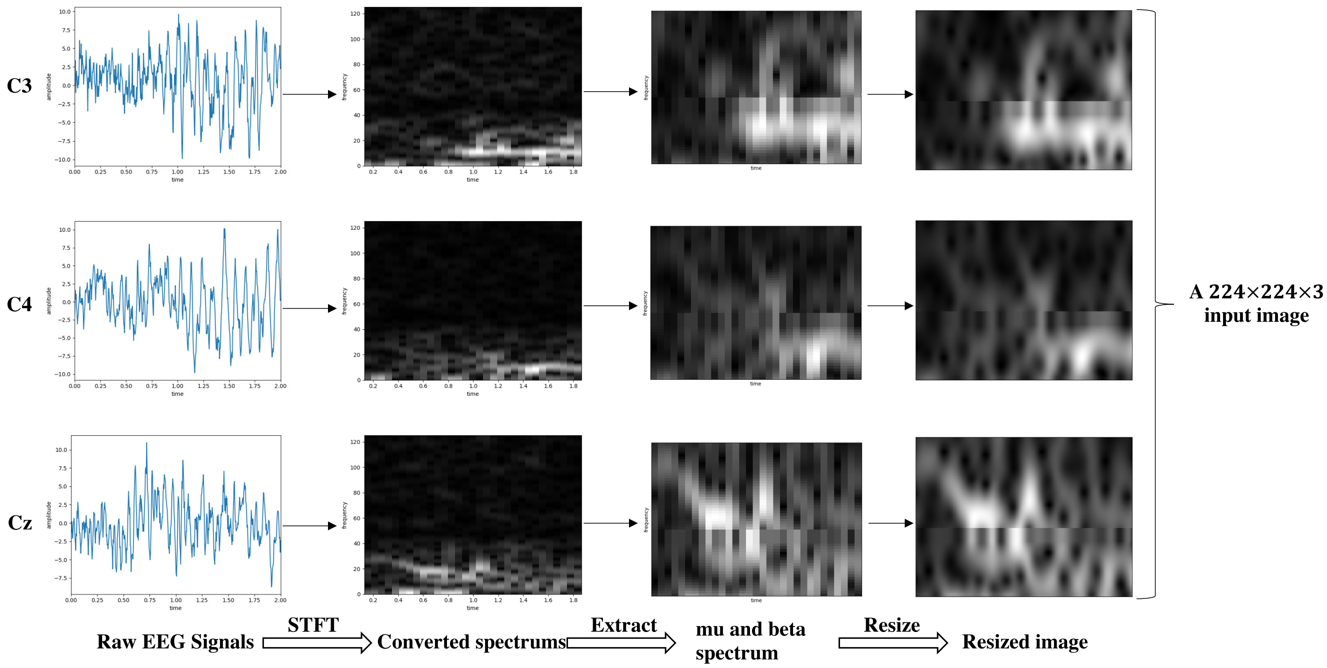


FIGURE 3. The detailed conversion process of EEG signals to time-frequency spectrum images.

to represent the beta band, because this results in a better performance in our verification experiments.

The detailed conversion process of EEG signals to time-frequency spectrum images is shown in Fig. 3. Firstly, for each trail, STFT is computed on the EEG signals from three electrodes, respectively. Each EEG signal is a 2s long time series with 500 signal points. Here, the window size of STFT is set to 64 and the number of points to overlap between segments in STFT is set to 50. This results in three spectrum images whose size are all 257×32, where 257 is the number of sample frequencies and 32 is the number of segment times. Then, mu and beta frequency bands of the spectrum images are extracted. The size of the extracted image for mu band is 20×32 and the size of the extracted image for beta band is 33×32. After that, in order to match the input of the proposed framework, the images for beta band and mu band are both resized to 112×224, thus generating a 224×224 spectrum image. Finally, three spectrum images converted from EEG signals at C3, C4 and Cz electrodes are combined together to form a $224 \times 224 \times 3$ input image.

B. PROPOSED DEEP TRANSFER CNN FRAMEWORK

After the EEG signal data are successfully preprocessed, a set of time-frequency spectrum images are generated. Then, the problem of EEG signal classification is solved by classifying these images. Due to the insufficient labeled data in practical applications and the long time used for training deep CNN, transfer learning technology is introduced in this paper to improve the training efficiency of CNN model with limited amount of labeled data. It has been proven in several previous studies [36]–[40] that 1) network-based deep transfer learning, which always transfer the partial structure and parameters of a pre-trained deep neural network in source domain to another network in target domain, is effective in reducing the training time and the need of labeled data in the target domain. 2) low-level features in the front-layers of CNN are universal for different but related tasks, while high-level features in the later-layers are specific for different tasks. Thus the front-layers of CNN are always regarded as a universal feature extractor. Based on those two remarks, a transfer CNN framework is proposed to classify these spectrum images, and the architecture of it is shown in Fig. 4. The proposed framework consists of a pre-trained VGG-16 CNN model and a target CNN model, where the pre-trained VGG-16 CNN model is used to extract universal features for common image classifications tasks, and the target CNN model aims to classify EEG signals efficiently and accurately with the aid of the pre-trained VGG-16 CNN. The detailed information about the pre-trained VGG-16 CNN, the target CNN and the training procedure of the proposed framework is as follows.

1) PRE-TRAINED VGG-16 CNN

VGG-16 is a well-known CNN model with 16 layers proposed by Oxford Visual Geometry Group in 2014 and it has achieved remarkable performance in various image processing tasks [29]. VGG-16 replaces large-sized convolution filters with small-sized filters while increasing the depth of network. This is mainly because CNN with small filters will benefit the improvement of classification accuracy. The detailed configurations of all layers in VGG-16 can also be seen in Fig. 4. The VGG-16 CNN model used in this paper is pre-trained on the ImageNet dataset and the front-layers of the pre-trained CNN model can extract low-level universal

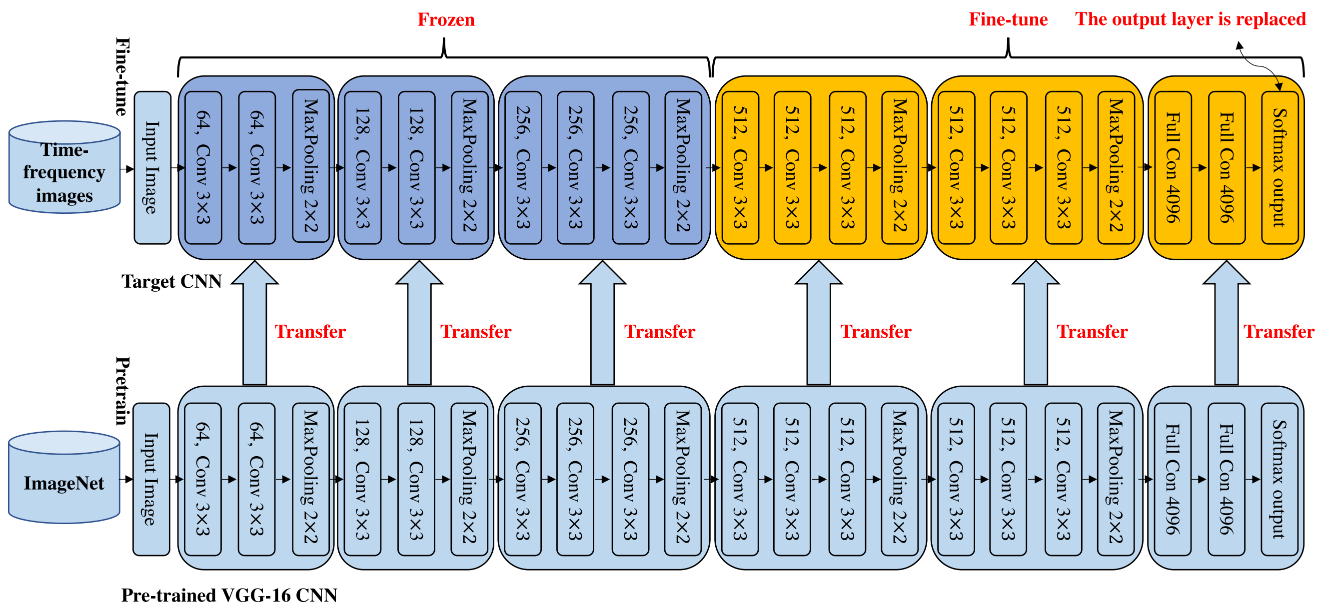


FIGURE 4. The proposed deep transfer CNN framework based on VGG-16.

features, which are appropriate for general image processing tasks.

2) TARGET CNN

The target CNN model has almost the same structure with the pre-trained VGG-16 CNN, the only difference between these two CNN models is the softmax output layer. The original output layer is removed and then a new softmax output layer is added, the number of neurons in the new output layer is exactly the same as the number of types of MI tasks. Here, the softmax classifier is used for EEG signal classification. The hyper-parameters, parameters and structures of the pre-trained CNN model are transferred to the target CNN model to improve its performance. Then, the target CNN model can be fine-tuned for the specific MI EEG signal classification task without the need to train the whole network from scratch. The loss function for fine-tuning the target CNN is softmax cross entropy, which is defined as

$$H(r,p) = -\sum_i r_i log(p_i) \tag{2}$$

where r is 1 when the predicted output is the same as the true label, otherwise, r is 0; p is the output probability.

3) TRAINING PROCEDURE

After the structures of the pre-trained VGG-16 CNN and the target CNN are successfully designed, the obtained time-frequency spectrum images are divided into training dataset and testing dataset. The training dataset is used to fine-tune the target CNN and the testing dataset is used to evaluate the performance of the target CNN. Since the target dataset used for fine-tuning is not similar to the ImageNet dataset used for pre-training, more layers of the target CNN need to be fine-tuned in this work. Consequently, as shown in Fig. 4, the front-layers from Layer 1 to Layer 11 of the target CNN are frozen; While the later-layers after Layer 11 are set to be trainable, these layers are fine-tuned on the target dataset. The specific training procedure of the proposed framework is given in Algorithm 1.

IV. EXPERIMENT

In this section, the accuracy and efficiency of the proposed framework are verified by the experiments conducted on the dataset 2b from the BCI competition IV. In addition, comparative experiments with traditional methods and standard deep learning-based methods are also conducted to verify the superiority of the proposed framework. All the experiments are carried out with the PyTorch deep learning framework on a high-performance computer, which is equipped with an Intel 12-core 3.5-GHz CPU, a GTX1080TI GPU, 256 GB SSD and 96 GB RAM.

A. EXPERIMENTAL DATA PREPARATION

For data preparation, only the EEG trails in the first three sessions of the dataset 2b are used in this paper. In order to verify the framework's performance, two subdatasets are generated as follows:

1) DATASET A

For each EEG trail, the duration between 0.5s and 3.5s after the arrow indicating MI tasks is displayed is extracted [21]. Each extracted 3s EEG signal is then converted to 11 2s signals with a step of 0.1s, and the specific conversion process is shown in Fig. 5. In this dataset, training data and testing data

Algorithm 1 Training Procedure

Input: Given the ImageNet dataset and the target training dataset $X_t\{x^t, y^t\}$ composed of time-frequency images

Output: The trained deep transfer CNN framework

Step 1: Establish the deep transfer CNN framework

The framework consisting of a pre-trained CNN and a target CNN is established, and the hyper-parameters and structures of these two CNNs are given.

Step 2: Initialize and pre-train the VGG-16 CNN

The VGG-16 CNN is randomly initialized and pre-trained on the ImageNet dataset, the objective is to extract common low-level features for image classification tasks.

Step 3: Initialize the target CNN

The hyper-parameters, structures and parameters of the pre-trained CNN are transferred to initialize the target CNN.

Step 4: Fine-tune the target CNN

Froze the front-layers of the target CNN and fine-tune the later-layers using the training dataset $X_t\{x^t, y^t\}$, the objective of the fine-tuning process is to minimize the loss function (see equation (2)).

Step 5: Output the deep transfer CNN framework

After the target CNN is fine-tuned, the proposed framework is output and applied to classify time-frequency spectrum images of EEG signals.

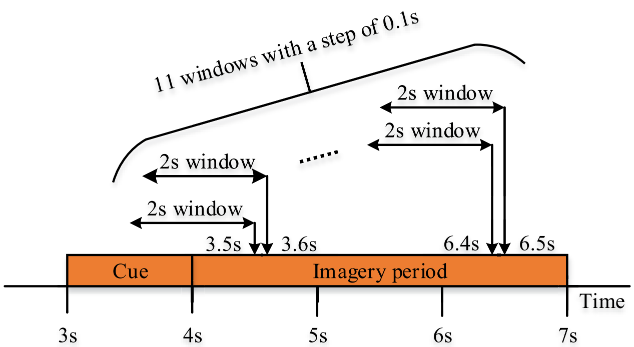


FIGURE 5. The specific process of data preparation.

are composed of time-frequency spectrum images of these 2s signals.

2) DATASET B

In this dataset, the duration between 1s and 4s after the arrow indicating MI tasks is displayed is extracted. Excepting this, the dataset B is generated in the same way as the Dataset A.

There are totally 4400 samples for each subject in both two subdatasets. Each sample is converted to a $224\times224\times3$ time-frequency image by the data preprocessing method. Then, 80% of all the converted images are randomly selected for training, and the remaining 20% are randomly selected for testing. Consequently, there are 3520 images in the training set and 880 images in the testing set of dataset A and dataset B for each subject. The converted time-frequency images of nine subjects in the dataset A are shown in Fig. 6.

TABLE 1. Average accuracy results of the proposed framework, CNN, SVM and ANN.

Subject	Dataset	Average accuracy(%)			
		Proposed framework	CNN	ANN	SVM
1	A	73.7	71.2	66.9	68.7
	B	72.6	71.2	65.2	68.4
2	A	59.2	58.7	45.6	52.2
	B	60.3	57.7	45.1	54.2
3	A	65.1	61.5	54.1	59.2
	B	66.9	64	54.2	58.7
4	A	92.8	90.8	86.1	90.3
	B	91.2	90.2	87.2	89.4
5	A	81.3	79.6	76.2	73.2
	B	80.6	78.3	72.2	77.5
6	A	70.8	69.2	62.8	68.8
	B	70.6	70.2	59.7	68.5
7	A	73.3	69.2	66	68.2
	B	73.2	70.7	66.1	67.8
8	A	78.1	71.9	62.1	70.2
	B	77.7	69.9	62.2	68.9
9	A	76.4	72.6	60.1	65.1
	B	71.2	68.5	59.2	62.6
Average	A&B	74.2	71.4	63.9	68.4

The MI tasks can be classified by classifying the converted images using the proposed transfer CNN framework. It can be seen from Fig. 6 that the converted images of two MI tasks for the same subject are different from each other, which offers an intuitive way of classifying them. This demonstrates the effectiveness of the data-processing method. Besides, the converted images of the same task from different subjects are also different from each other. Therefore, MI classification tasks across subjects are not taken into consideration in this paper, which is still a tough problem in the research area of MI BCI system.

B. EXPERIMENTAL RESULTS AND ANALYSIS

This subsection presents the accuracy and efficiency results of the proposed framework, together with the comparative results between the proposed framework and several existing methods.

1) ACCURACY

The target CNN model in the proposed framework is fine-tuned on the prepared training set, and then the fine-tuned target CNN model is used for EEG signals classifications. The accuracy performance of the proposed framework is evaluated on the prepared testing set. The performance of the proposed framework is also compared with that of SVM, ANN and standard CNN. Three points need to be pointed out: 1) SVM and ANN are two traditional machine learning-based methods, the performance of which is easily affected by the designed features and the configuration parameters. 2) The standard CNN is trained from scratch and the structure of it is shallower than the VGG-16 CNN model. 3) In this paper, dropout technique is introduced to prevent the target model from over-fitting during training and improve the generalization ability of the target model.

All the accuracy experiments are carried out ten times, and Table 1 presents the average accuracy results of the proposed

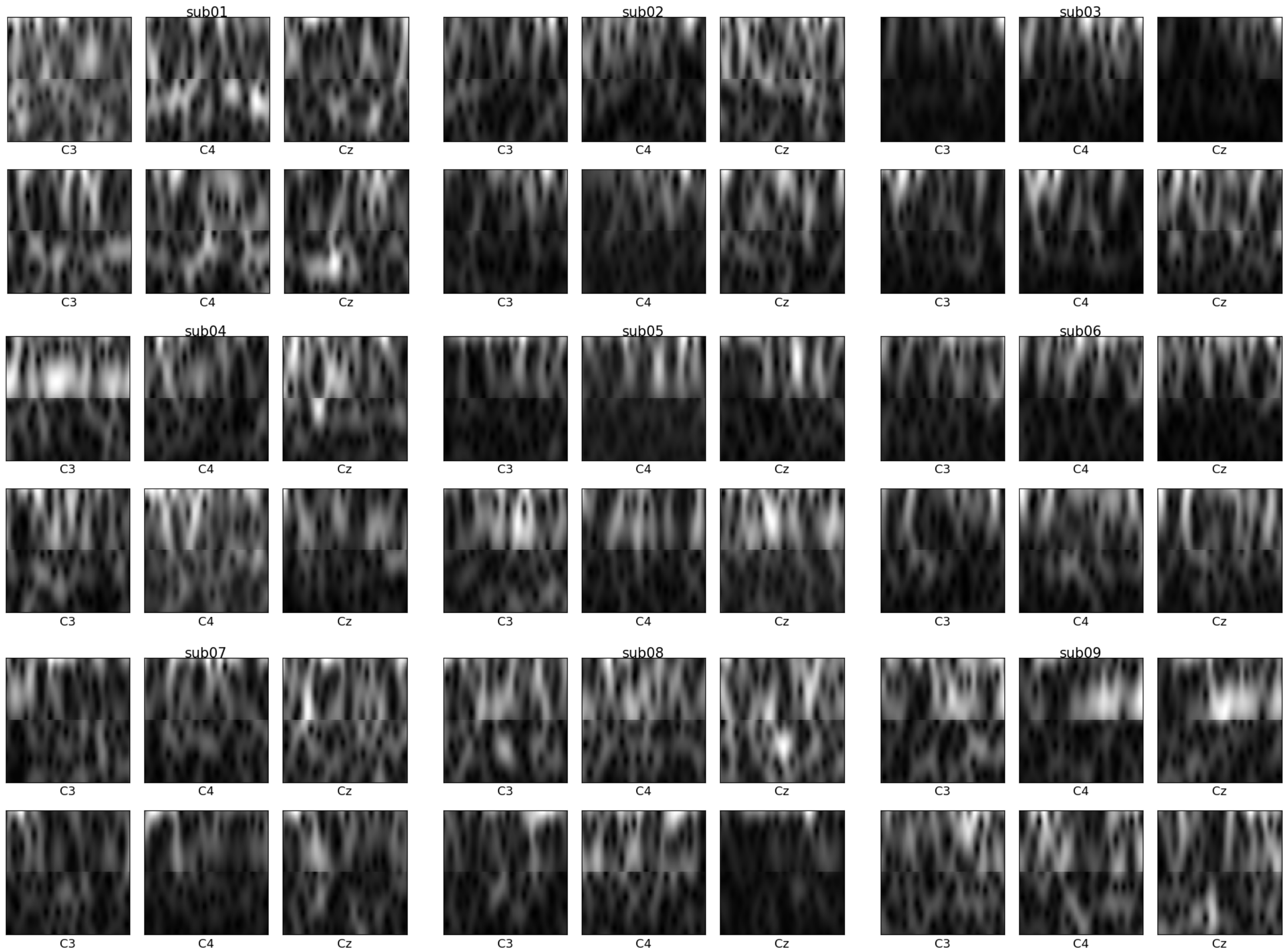


FIGURE 6. The converted time-frequency images of nine subjects in the dataset A.

framework, CNN, SVM and ANN. As shown in Table 1, the proposed framework achieves the best accuracy performance for all nine subjects compared with all the other methods. To be specific, the average accuracy of the proposed framework for all subjects is 74.2%. It achieves 2.8%, 10.3%, 5.8% average accuracy improvement compared with CNN, ANN, and SVM, respectively. It's remarkable that the standard CNN and the proposed framework both improve the classification accuracy a lot compared with traditional machine learning-based methods, proving the strong feature extraction and signal classification ability of deep learning-based methods. What's more, although the accuracy performance of the proposed framework is slightly superior to that of the standard CNN, it can eliminate the complexity of structure design and hyper-parameters tuning of the target CNN. This is because the hyper-parameters and structure of the pre-trained VGG-16 CNN model are directly transferred to the target CNN model.

2) EFFICIENCY

In order to compare the efficiency performance of the proposed framework with other methods, the time spent on training all the above-mentioned methods to achieve convergence are recorded. In essence, comparing with the traditional methods including ANN and SVM is unnecessary. Because selecting the optimal parameters for these two methods before training is required, which is extremely time-consuming. Therefore, only the efficiency comparison results between the proposed framework and the standard CNN trained from scratch are listed in Fig. 7.

From the results, it can be seen that training the proposed transfer CNN framework for all nine subjects is much faster than the standard CNN. More specifically, the time spent on training the standard CNN is two to three times longer than the time spent on the proposed framework. Although during the training process, the standard CNN has less parameters need to be updated than the proposed framework, training the CNN model with random initialization still costs more time. Besides, the CNN trained from scratch achieves the desired performance only when the hyper-parameters of it are optimized for the specific task, which is also very time-consuming. While the target CNN in the proposed framework already has the optimal hyper-parameters which are transferred from the pre-trained CNN. Therefore, it can be

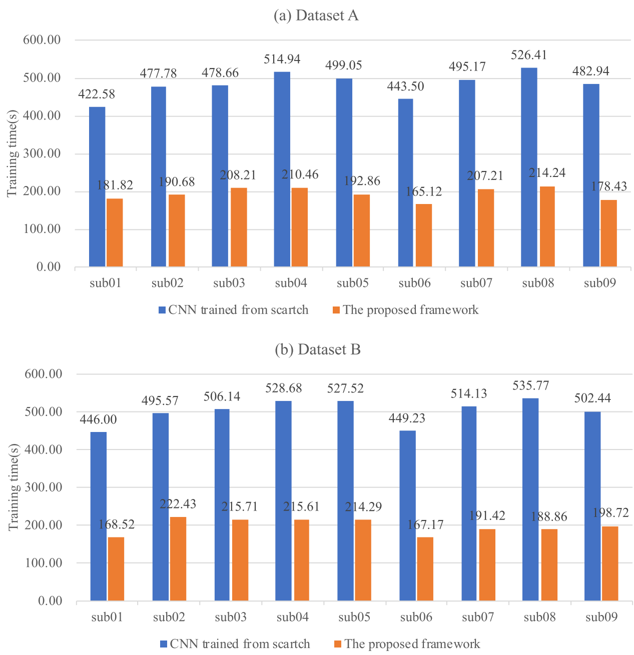


FIGURE 7. Training time of the proposed framework and the standard CNN trained from scratch: (a) Dataset A and (b) Dataset B.

concluded that the proposed framework can improve the accuracy performance and speed up the training process of the target CNN model in EEG signal classification applications.

V. CONCLUSION AND FUTURE WORK

In this paper, an EEG signal classification method based on a deep transfer CNN framework is proposed. The main contributions of this paper can be summarized as follows:

1) Designing a STFT-based data preprocessing method. The data preprocessing method is able to extract time-domain and frequency-domain features of raw signals simultaneously. At the same time, each EEG trail is converted to a $224\times224\times3$ time-frequency spectrum image after data preprocessing.

2) Proposing a deep transfer CNN framework. It consists of a pre-trained VGG-16 CNN and a target CNN, where the pre-trained CNN is trained using natural images and the target CNN is fine-tuned using the converted time-frequency spectrum images. The pre-trained parameters, structures and hyper-parameters of the pre-trained CNN are directly transferred to the target CNN in order to improve its classification accuracy and accelerate its training process.

3) Applying the proposed framework to MI EEG signal classification. The performance of the proposed framework is verified on a MI EEG signal dataset, and the efficiency and accuracy experiments are carried out on this dataset. From the experimental results, it can be concluded that the proposed framework possesses better classification accuracy and faster training speed compared with other methods including CNN, SVM and ANN.

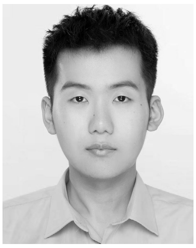
Although there have been some progress in this paper, two limitations still have to be addressed in the future. Firstly, the transfer-ability of different layers of the pre-trained CNN model hasn't been studied in this paper. Secondly, training the proposed framework is still a time-consuming process even with the GPU-accelerated computing. Therefore, the future works will focus on the following two aspects: firstly, an additional analysis on the effect of different transferred layers on the classification accuracy and efficiency should be studied. The aim of this analysis is to find the most appropriate layers of the pre-trained CNN to be transferred to the target CNN. Secondly, the architecture of the proposed framework should be future optimized to speed up the training process.

REFERENCES

[1] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "BCI2000: A general-purpose brain-computer interface (BCI) system," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1034–1043, Jun. 2004.
[2] C. Huang, X. Shan, Y. Lan, L. Liu, H. Cai, W. Che, Y. Hao, Y. Cheng, and Y. Peng, "A hybrid active contour segmentation method for myocardial D-SPECT images," *IEEE Access*, vol. 6, pp. 39334–39343, 2018.
[3] C. Huang, Y. Xie, Y. Lan, Y. Hao, F. Chen, Y. Cheng, and Y. Peng, "A new framework for the integrative analytics of intravascular ultrasound and optical coherence tomography images," *IEEE Access*, vol. 6, pp. 36408–36419, 2018.
[4] G. Pfurtscheller, C. Brunner, A. Schlögl, and F. H. L. da Silva, "Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks," *Neuroimage*, vol. 31, no. 1, pp. 153–159, May 2006.
[5] S.-M. Zhou, J. Q. Gan, and F. Sepulveda, "Classifying mental tasks based on features of higher-order statistics from EEG signals in brain–computer interface," *Inf. Sci.*, vol. 178, no. 6, pp. 1629–1640, 2008.
[6] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for EEG-based brain–computer interfaces," *J. Neural Eng.*, vol. 4, no. 2, pp. R1–R13, 2007.
[7] A. Subasi, "EEG signal classification using wavelet feature extraction and a mixture of expert model," *Expert Syst. Appl.*, vol. 32, no. 4, pp. 1084–1093, May 2007.
[8] X. Jiang, S. Wei, S. Zhang, L. Zhao, F. Liu, H. Tang, Z. He, and C. Liu, "Comparison of photoplethysmographic signal features between healthy and sleep apnea patients during five sleep stages," *J. Med. Imag. Health Informat.*, vol. 9, no. 1, pp. 63–69, 2019.
[9] J. Zhang, L. Yang, Z. Su, X. Mao, K. Luo, and C. Liu, "Photoplethysmogram signal quality assessment using support vector machine and multifeature fusion," *J. Med. Imag. Health Inform.*, vol. 8, no. 9, pp. 1757–1762, 2018.
[10] F. Liu, C. Liu, L. Zhao, X. Zhang, X. Wu, X. Xu, Y. Liu, C. Ma, S. Wei, Z. He, and J. Li, "An open access database for evaluating the algorithms of electrocardiogram rhythm and morphology abnormality detection," *J. Med. Imag. Health Informat.*, vol. 8, no. 7, pp. 1368–1373, Jul. 2018.
[11] Y. Liu, S. Wei, S. Zhang, F. Liu, J. Li, and C. Liu, "Signal quality index-based two-step method for heart rate estimation by combining electrocardiogram and arterial blood pressure signals," *J. Med. Imag. Health Inform.*, vol. 8, no. 7, pp. 1502–1507, 2018.
[12] S. Lu, Z. Lu, P. Phillips, S. Wang, and Y. Zhang, "Pathological brain detection in magnetic resonance imaging using combined features and improved extreme learning machines," *J. Med. Imag. Health Informat.*, vol. 8, no. 7, pp. 1486–1490, 2018.
[13] Z. Zhao, C. Liu, Y. Li, Y. Li, J. Wang, B.-S. Lin, and J. Li, "Noise rejection for wearable ECGs using modified frequency slice wavelet transform and convolutional neural networks," *IEEE Access*, vol. 7, pp. 34060–34067, 2019.
[14] W.-Y. Hsu, "Single-trial motor imagery classification using asymmetry ratio, phase relation, wavelet-based fractal, and their selected combination," *Int. J. Neural Syst.*, vol. 23, no. 2, 2013, Art. no. 1350007.
[15] K. K. Ang, C. Guan, K. S. G. Chua, B. T. Ang, C. W. K. Kuah, C. Wang, K. S. Phua, Z. Y. Chin, and H. Zhang, "A large clinical study on the ability of stroke patients to use an EEG-based motor imagery brain-computer interface," *Clin. EEG Neurosci.*, vol. 42, no. 4, pp. 253–258, 2011.

[16] Siuly, Y. Li, and P. Wen, "Modified CC-LR algorithm with three diverse feature sets for motor imagery tasks classification in EEG based brain–computer interface," *Comput. Methods Programs Biomed.*, vol. 113, no. 3, pp. 767–780, 2014.
[17] Y. Wang, S. Gao, and X. Gao, "Common spatial pattern method for channel selelction in motor imagery based brain-computer interface," in *Proc. 27th Annu. Conf. IEEE Eng. Med. Biol.*, Jan. 2006, pp. 5392–5395.
[18] A. S. Aghaei, M. S. Mahanta, and K. N. Plataniotis, "Separable common spatio-spectral pattern algorithm for classification of EEG signals," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2013, pp. 988–992.
[19] A. Ashok, A. K. Bharathan, V. R. Soujya, and P. Nandakumar, "Tikhonov regularized spectrally weighted common spatial patterns," in *Proc. Int. Conf. Control Commun. Comput. (ICCC)*, Dec. 2013, pp. 315–318.
[20] K. K. Ang, Z. Y. Chin, C. Wang, C. Guan, and H. Zhang, "Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b," *Frontiers Neurosci.*, vol. 6, no. 1, p. 39, 2012.
[21] Y. R. Tabar and U. Halici, "A novel deep learning approach for classification of EEG motor imagery signals," *J. Neural Eng.*, vol. 14, no. 1, 2017, Art. no. 016003.
[22] C. Park, D. Looney, N. ur Rehman, A. Ahrabian, and D. P. Mandic, "Classification of motor imagery BCI using multivariate empirical mode decomposition," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 21, no. 1, pp. 10–22, Jan. 2013.
[23] H. K. Lee and Y.-S. Choi, "A convolution neural networks scheme for classification of motor imagery EEG based on wavelet time-frequecy image," in *Proc. Int. Conf. Inf. Netw. (ICOIN)*, Jan. 2018, pp. 906–909.
[24] N. Naseer and K.-S. Hong, "Classification of functional near-infrared spectroscopy signals corresponding to the right- and left-wrist motor imagery for development of a brain–computer interface," *Neurosci. Lett.*, vol. 553, pp. 84–89, Oct. 2013.
[25] S. Siuly and Y. Li, "Improving the separability of motor imagery EEG signals using a cross correlation-based least square support vector machine for brain–computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 20, no. 4, pp. 526–538, Jul. 2012.
[26] A. Subasi, "Automatic recognition of alertness level from EEG by using neural network and wavelet coefficients," *Expert Syst. Appl.*, vol. 28, no. 4, pp. 701–711, 2005.
[27] C. Huang, G. Tian, Y. Lan, Y. Peng, E. Y. K. Ng, Y. Hao, Y. Cheng, and W. Che, "A new pulse coupled neural network (PCNN) for brain medical image fusion empowered by shuffled frog leaping," *Frontiers Neurosci.*, vol. 13, p. 210, Mar. 2019.
[28] H. Yang, S. Sakhavi, K. K. Ang, and C. Guan, "On the use of convolutional neural networks and augmented CSP features for multi-class motor imagery of EEG signals classification," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 2620–2623.
[29] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: https://arxiv.org/abs/1409.1556
[30] V. Jayaram, M. Alamgir, Y. Altun, B. Scholkopf, and M. Grosse-Wentrup, "Transfer learning in brain-computer interfaces," *IEEE Comput. Intell. Mag.*, vol. 11, no. 1, pp. 20–31, Feb. 2016.
[31] S. Shao, S. McAleer, R. Yan, and P. Baldi, "Highly accurate machine fault diagnosis using deep transfer learning," *IEEE Trans. Ind. Inform.*, vol. 15, no. 4, pp. 2446–2455, Apr. 2019.
[32] M. M. Al Rahhal, Y. Bazi, H. AlHichri, N. Alajlan, F. Melgani, and R. R. Yager, "Deep learning approach for active classification of electrocardiogram signals," *Inf. Sci.*, vol. 345, pp. 340–354, Jun. 2016.
[33] Z. Shi, H. Hao, M. Zhao, Y. Feng, L. He, Y. Wang, and K. Suzuki, "A deep CNN based transfer learning method for false positive reduction," *Multimedia Tools Appl.*, vol. 78, no. 1, pp. 1017–1033, 2019.
[34] R. Leeb, C. Brunner, and G. Müller-Putz. *BCI Competition 2008—Graz Data Set B*. Accessed: Jun. 25, 2019. [Online]. Available: http://www.bbci.de/competition/iv/
[35] G. Pfurtscheller and F. L. Da Silva, "Event-related EEG/MEG synchronization and desynchronization: Basic principles," *Clin. Neurophysiol.*, vol. 110, no. 11, pp. 1842–1857, 1999.
[36] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, "A survey on deep transfer learning," 2018, *arXiv:1808.01974*. [Online]. Available: https://arxiv.org/abs/1808.01974
[37] H.-C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. M. Summers, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1285–1298, May 2016.
[38] S.-H. Wang, S. Xie, X. Chen, D. S. Guttery, C. Tang, J. Sun, and Y.-D. Zhang, "Alcoholism identification based on an AlexNet transfer learning model," *Frontiers Psychiatry*, vol. 10, p. 205, Apr. 2019.
[39] Z. Dong, X. Chen, W. Jia, S. Du, K. Muhammad, and S.-H. Wang, "Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation," *Multimedia Tools Appl.*, vol. 78, no. 3, pp. 3613–3632, 2019.
[40] S.-H. Wang, K. Muhammad, J. Hong, A. K. Sangaiah, and Y.-D. Zhang, "Alcoholism identification via convolutional neural network based on parametric ReLU, dropout, and batch normalization," *Neural Comput. Appl.*, pp. 1–16, Dec. 2018.

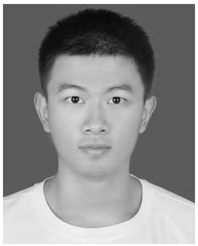
GAOWEI XU received the bachelor's degree from the Nanjing University of Information Science and Technology, Nanjing, Jiangsu, China, in 2015. He is currently pursuing the Ph.D. degree with the College of Electronics and Information Engineering, Tongji University, Shanghai, China. His research interests include machine learning, deep learning, transfer learning, and their applications in intelligent predictive maintenance and biomedical signal processing.



XIAOANG SHEN is currently pursuing the B.Sc. degree with Tongji University, Shanghai, China. His research interests include image processing, reconstruction, and 3-D visualization.

SIRUI CHEN is currently pursuing the B.Sc. degree with Tongji University, Shanghai, China. Her research interests include image processing and reconstruction.

YONGSHUO ZONG is currently pursuing the B.Sc. degree with Tongji University, Shanghai, China. His research interest includes image processing.



CANYANG ZHANG is currently pursuing the B.Sc. degree with Tongji University, Shanghai, China. His research interest includes image processing.

HONGYANG YUE is currently pursuing the B.Sc. degree with Tongji University, Shanghai, China. His research interest includes image processing.

MIN LIU received the bachelor's degree from the China University of Geosciences, Wuhan, in 1993, and the master's and Ph.D. degrees from Zhejiang University, China, in 1996 and 1999, respectively. He has been working on computer science and system engineering, collaborative MRO and intelligent manufacturing for about 14 years. He is currently a Full Professor with the School of Electronic and Information Engineering, Tongji University, Shanghai China. He has published over 60 papers in scientific journals and international conferences in the related areas.

FEI CHEN received the M.D. and Ph.D. degrees from the Medical College, Zhejiang University, Zhejiang, China. He is currently a Fellow Doctor with Tongji Hospital, Tongji University, Shanghai. His research interests include the pathophysiological proceeding of atherosclerosis, interventional treatment of CHD, image reconstruction of coronary stent, and 3-D visualization of coronary stent.

WENLIANG CHE received the Ph.D. and M.D. degrees from Tongji University, Shanghai, China. He is currently an Interventional Cardiologist with the Department of Cardiology, Shanghai Tenth People's Hospital, Tongji University School of Medicine, Shanghai. His research interests include atherosclerosis imaging, intra-coronary imaging in managing CHD, and quantitative assessment of coronary microvascular function.