Journal of Neural Engineering

ACCEPTED MANUSCRIPT

A novel hybrid deep learning scheme for four-class motor imagery classification

To cite this article before publication: Zhang Ruilong *et al* 2019 *J. Neural Eng.* in press https://doi.org/10.1088/1741-2552/ab3471

A novel hybrid Deep Learning scheme for four-class Motor Imagery classification

Ruilong Zhang, Qun Zong, Liqian Dou & Xinyi Zhao
School of Electrical and Information Engineering, Tianjin University, Tianjin, China
E-mail: douliqian@tju.edu.cn

March 2019

Abstract. Learning the structures and unknown correlations of motor imagery (MI) EEG signal is important for the classification. It is also a major challenge to obtain good classification accuracy from increased number of classes and increased variability from different people. In this study, four-class MI task is investigated. An end-to-end novel hybrid deep learning scheme is developed to decode MI task from EEG data. The proposed algorithm consists of two parts: *a.* One-versus-rest filter bank common spatial pattern (OVR-FBCSP) is adopted to preprocess and pre-extract the features of four-class MI signal. *b.* A hybrid deep network based on CNN and LSTM network is proposed to extract and learn the spatial and temporal features of MI signal simultaneously. The main contribution of this paper is to propose a hybrid deep network framework to improve classification accuracy of four-class MI-EEG signal. The hybrid deep network is a subject-independent shared neural network which means it can be trained by using the training data from all the subjects to form one model. The classification performance obtained by the proposed algorithm on BCI competition IV dataset 2a in terms of accuracy is 83% and Cohen's kappa value is 0.80. Finally, the shared hybrid deep network is evaluated by every subject respectively, and the experiment results illustrate that the shared neural network get a satisfactory accuracy. Thus, the proposed algorithm could be of great interest for real-life brain-computer interfaces (BCIs).

Keywords: brain-computer interface, four-class motor imagery, OVR-FBCSP, convolutional neural network, long short-term memory

Submitted to: *J. Neural Eng.*

1. Introduction

Brain-computer interface (BCI) system provides an alternative pathway of human-computer interface (HCI) [1]. The important applications of BCI are to analyze the EEG signal for people with neuromuscular disabilities and understand their intentions [2–4]. Among several BCI signal, P300 evoked potentials, steady-state visual evoked potentials (SSVEP) and motor imagery (MI) are the most popular research paradigms of EEG signal analyzing, where MI is the only one based on spontaneous potential without any external stimulus. MI is a mental process by which an individual rehearses or simulates a given action [5]. For different movement task, the changing of neurons firing pattern can be reflected by the changing of power spectrum of different frequency bands. This phenomenon is called event-related synchronization (ERS) and event-related desynchronization (ERD) [6]. In MI tasks, mu band (8-14Hz) and beta band (14-30Hz) are the main spectrum of ERS and ERD. The goal of MI tasks is to classify activities in order to recognize the imaged movement of body.

Pattern recognition plays one of the most important roles in EEG signal analyze. Common spatial patterns (CSP) [7,8] is a popular method in extracting different MI features. Dozens of extension methods of CSP have also been applied successfully in MI tasks [9,10]. In the classification part, support vector machine(SVM) is one of the most popular classifiers in traditional algorithms [8,10] and other traditional classifiers including linear discriminant analysis (LDA) [11] and Bayesian classifier [12] have also been employed in many studies.

A large amount of BCI data can be provided by BCI devices. As we know, deep learning method is a better classifier for dealing with the complex data, and increasing size of training data can make deep learning method more effective. Compared with the traditional machine learning method, the most powerful advantage of deep learning method is excellent fitting ability, hence it can approximate any complex function. So deep learning method can learn the features which human cannot describe into mathematics. So far, deep learning methods are successfully applied in computer vision and natural language processing and understanding [13–16]. This makes deep learning methods be a great choice to process the EEG signal of MI. In the study by Lu et al. [17], a deep belief network (DBN) based on restricted Boltzmann machine (RBM) combined with fast Fourier transformation (FFT) was applied for two-class MI classification and performance of result is more successful than traditional method. DBN is also used widely in MI tasks [18] and other BCI studies [19] and has achieved good results as well. Convolutional neural network (CNN) is also used in EEG researches. Tabar et al. [20] combine the 1D CNN and Stacked autoencoder (SAE) to classify two-class MI signal. Yang et al. [10] investigated classification of multi-class MI of EEG signal based on CNN and augmented CSP features. Sakhavi et al. [21] propose a new CNN architecture to introduce the temporal representation of the data for MI classification. Schirrmeister et al. [22] use two basic, shallow and deep ConvNet architectures recently shown to decode task-related information from EEG. Zhou et al. [23] propose a novel method based on wavelet envelope analysis and long-term short-term memory (LSTM) classifier which consider the amplitude modulation characteristics and time series information of MI-EEG to classify EEG signal into multiple classes. Many other deep learning framework methods [24–26] are also used in MI classification. However, studies based on deep learning methods only deal with the two-class MI classification, while the research on the four-class MI is still few. Moreover, EEG is a signal with temporal features and LSTM is an excellent network for processing time series signal but rarely use in four-class MI signal. Some recent deep learning methods for MI task are subject-dependent individual networks, such as [20, 22, 27, 28]. These networks need a lot of time and computer memory to train for each subject. So we want to design a subject-independent shared neural network which is trained by using the training data from all the subjects to form one common model.

Motivated by the observations, a deep learning scheme based on CNN-LSTM for four-class MI is developed. The novelty of the approach lies in three components. Firstly, we present a new hybrid deep learning framework by combining CNN and LSTM to learn the spatial features and temporal features of MI signal. Secondly, One-versus-rest filter bank common spatial pattern (OVR-FBCSP) is introduced before classification in order to extract the four-class MI features of EEG signal by one FBCSP filter. Thirdly, the proposed hybrid neural network is also a subject-independent shared neural network which is trained by merged data from all subjects and evaluated by every subject respectively. The evaluation result shows the shared network can get a good accuracy classification and even better than individual neural network for corresponding subject. The proposed approach is analyzed and evaluated by using BCI Competition IV dataset [29] and the performance results of proposed approach in this study are compared with the results of current state of art algorithms in this field. The experiments results are presented considering classification accuracy and kappa value [30].

In practice, MI classifier is used for identifying the subject EEG signal online in most cases, so it is important to design a classifier which can classify the MI signal quickly and accuracy. Hence, we provided a time courses experiment to validate our method also has a good classification effect with a short time EEG signal series. We extract the 0.5s to 4s (intervals of 0.5s) seconds date from the whole datasets respectively and classify the extracted data by proposed method. The result shows that our method can classify the short time EEG signal series with good accuracy.

The paper is organized as follows. Section 1 is the introduction. Section 2 describes proposed deep network for four-class MI classification. Experiments and results are given in Section 3. Discussion and conclusions can be found in Section 4.

2. Deep learning framework

The datasets we used in this work is BCI Competition IV dataset 2a provided by Graz University [29] including recordings from twenty two electrodes during left/right hand, foot, tongue MI task. The architecture of the proposed algorithm is illustrated in Fig.1. It contains three consecutive stages: OVR-FBCSP,CNNs and LSTM. Every trail of EEG signal is segmented into n time windows (for example, in the paper, we define $n = 5$, so that each time window is 0.8s). We respectively deal with these n time windows by OVR-FBCSP and CNNs to get spatial features of EEG. Then these spatial features in n time windows as a time series feed into LSTM to extract temporal feature of signal. Each stage of the algorithm is explained in more details in following sections.

2.1. Preprocess

A time window is employed to preprocess the EEG trial of each MI task. Each 4s EEG trail of BCI Competition IV dataset 2a is decomposed into 5 time windows. In case of 250Hz signal that is corresponding to 200 samples in each time window.

2.2. OVR-FBCSP

A filter bank is employed to decompose the EEG in every time window into 16 frequency pass bands by using causal Chebyshev Type II filter [8]. The filter bank,

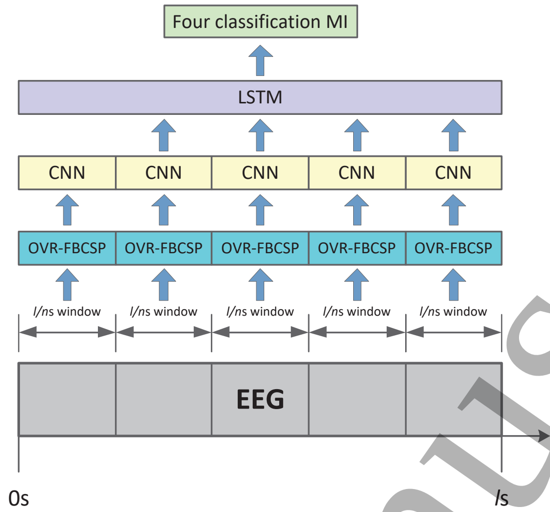

Figure 1. The proposed CNN-LSTM hybrid deep neural network.The parameters of OVR-FBCSP, CNN, LSTM are adjusted for each subject using training data labeled with the respective MI tasks.

which is including 16 filters, is from 4Hz to 38Hz and each filter has a bandwidth of 4Hz and overlap of 2Hz. A spatial filtering using the OVR-FBCSP algorithm is performed. CSP is a successful algorithm for motor imagery-based BCI in order to detect event-related desynchronization and synchronization (ERD/ERS) [7]. The OVR-FBCSP approach can extract the CSP features to discriminate one class from the rest of classes in multi-class MI task classification [31]. A four-class filter can be obtained by combining 4 two-classes filters [32]. EEG can be linearly transformed by OVR-CSP algorithm using

$$\boldsymbol{Z}_{b,t} = \boldsymbol{W}_b^{\mathrm{T}} \boldsymbol{X}_{b,t}, \tag{1}$$

where $\boldsymbol{X}_{b,t}$ denotes the EEG of bth band-pass filter of tth time window, $\boldsymbol{Z}_{b,t}$ is the OVR-FBCSP features, $\boldsymbol{W}_b^{\mathrm{T}} = \left[\boldsymbol{W}_{b,1}^{\mathrm{T}}, \boldsymbol{W}_{b,2}^{\mathrm{T}}, \boldsymbol{W}_{b,3}^{\mathrm{T}}, \boldsymbol{W}_{b,4}^{\mathrm{T}}\right]$ presents the weight of OVR-FBCSP filter and $\boldsymbol{W}_{b,j}^{\mathrm{T}}(j = 1, 2, 3, 4)$ is the CSP filter matrix of one class versus others, respectively. Each $\boldsymbol{W}_{b,j}^{\mathrm{T}}$ can be obtained by solving the eigenvalue decomposition problem [33, 34]

$$\boldsymbol{C}_{b,j}\boldsymbol{W}_{b,j} = (\sum_{j=1}^{4}\boldsymbol{C}_{b,j})\boldsymbol{W}_{b,j}\boldsymbol{E}_{b,j}, \tag{2}$$

where $\boldsymbol{C}_{b,j}$ is the covariance matric after bth band-pass filtered signal of jth class MI signal, $\boldsymbol{E}_{b,j}$ denotes the diagonal matrix which contains the $\boldsymbol{C}_{b,j}$ eigenvalues. For each class, we choose 2 pairs of CSP features of the tth time window for bth band-pass filtered EEG signal [8]. The CSP features are then given by

$$f_{b,t} = \log \frac{\operatorname{diag}(\hat{\boldsymbol{W}}_b^{\mathrm{T}}\boldsymbol{X}_{b,t}\boldsymbol{X}_{b,t}^{\mathrm{T}}\hat{\boldsymbol{W}}_b)}{\operatorname{tr}(\hat{\boldsymbol{W}}_b^{\mathrm{T}}\boldsymbol{X}_{b,t}\boldsymbol{X}_{b,t}^{\mathrm{T}}\hat{\boldsymbol{W}}_b)}, \tag{3}$$

where $\hat{\boldsymbol{W}}_b$ denotes combining matrix of $\boldsymbol{W}_{b,j}(j = 1, 2, 3, 4)$ which select first 2 and last 2 columns, $\operatorname{diag}(\cdot)$ is diagonal elements of the matrix and $\operatorname{tr}(\cdot)$ is the trace of the matrix, $f_{b,t}$ is the output of OVR-FBCSP.

2.3. Convolutional neural network (CNN)

After OVR-FBCSP processing, we can obtain 16×16 (bankfilters $\times$ features) format signal in one time window. The proposed convolutional neural network (CNN) model combined 3 hidden layers are presented in Fig.2. In each convolution layer, the input signal is convolved with the convolution kernel which is a 3×3 size filter and an activation function R is used to transform the linear operation into nonlinear [35]. The output of each convolution layer is given as

$$hc_l = R(\operatorname{conv}(W_l, x_l) + b_l), \tag{4}$$

where conv denotes convolutional operator, x_l is the input of l hidden layer, W_l is the weight matrix, b_l is the bias value and hc_l is the output of l hidden layer. The activation function R is selected as rectified linear unit (RELU) function [36]. RELU function is defined as

$$\mathrm{RELU}(a) = \max(0,\ a). \tag{5}$$

After each convolutional layer, the max-pooling layer with kernel size of 2×2 is applied to reduce the size of the feature matrix. Through the CNN, we can get $2 \times 2 \times 16$ size hc_3 and reshape the size into 1×64 size output o_t to fit input shape of LSTM, where t denotes tth time window.

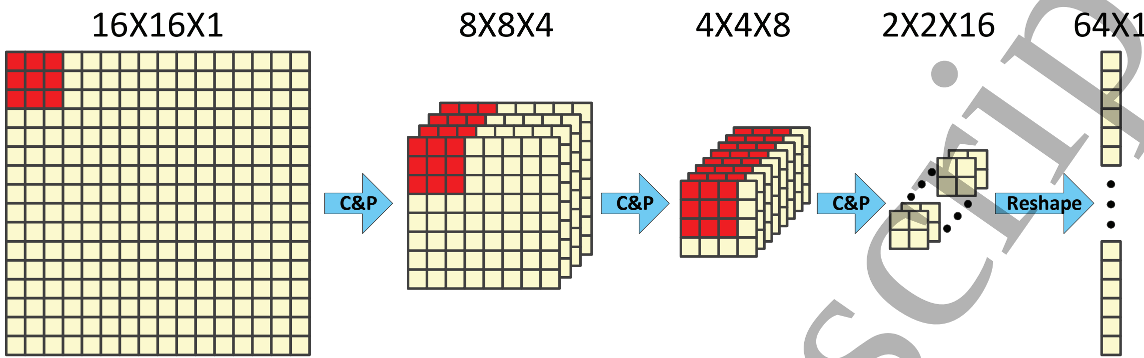

Figure 2. Proposed convolutional neural network (CNN) model. "C&P" defines convolutional and max-pooling operator for signal. The convolution kernel is 3×3 and the kernel of max-pooling layer is 2×2. Dropout strategy is adopted in the third layer while the neural network is trained.

In this paper, dropout strategy and zero-padding method is adopted on the proposed CNN model. The dropout strategy is used in the third layer of CNN while the neural network is trained. Dropout strategy is a regularization technique for reducing overfitting in CNN by preventing complex co-adaptations on training data [37]. As shown in Fig.3, when the neural network forward propagates, the activation value of some neuron stops working with a certain probability P. Dropout strategy can make the CNN model more generalized, because it can make the neural network ignore some local features. Here, dropout parameter P is selected as 0.5. Zero-padding method is adopted to prevent losing the feature dimension, which is shown in Fig.4. As we can see, the convolution output size can be same with input size after zero-padding.

2.4. Long Short-Term Memory (LSTM) network

EEG is a signal with temporal features, LSTM is an excellent way to process and predict time series signal [38]. Therefore, after extract spatial feature of EEG signal in 5 time windows by 5 CNNs, a deep LSTM network with one input layer, three hidden layers and one output layer is set behind CNNs. Deep LSTM network is consisted by many LSTM cells, the Deep LSTM network framework is shown in Fig.5 and the structure details of LSTM cell is presented in Fig.6.

Define the input signal $x_{1,t} = o_t, (t = 1, 2, \cdots, 5)$. Compared with normal

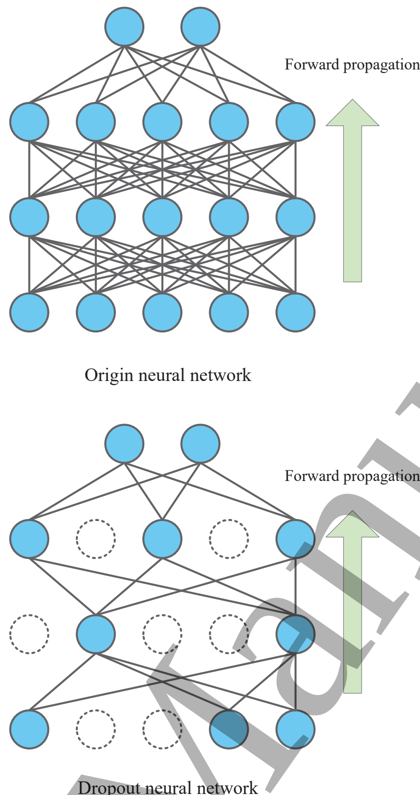

Figure 3. Dropout strategy

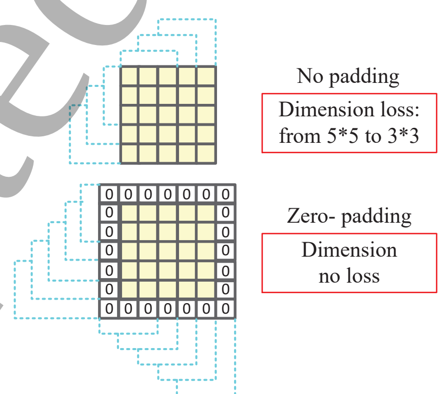

Figure 4. Zero-padding for CNN Model

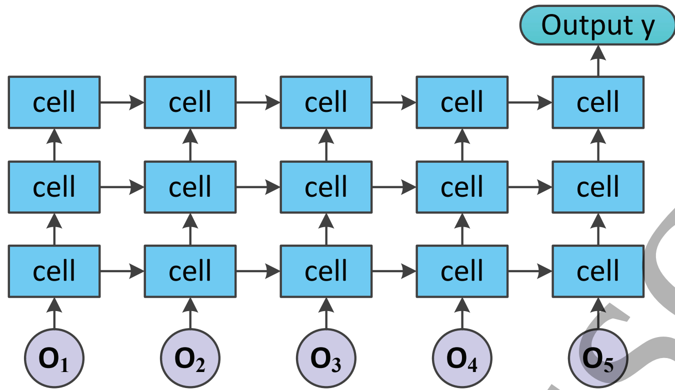

Figure 5. Proposed Long Short-Term Memory (LSTM) network. The architecture of LSTM cell is shown in Figure 4.

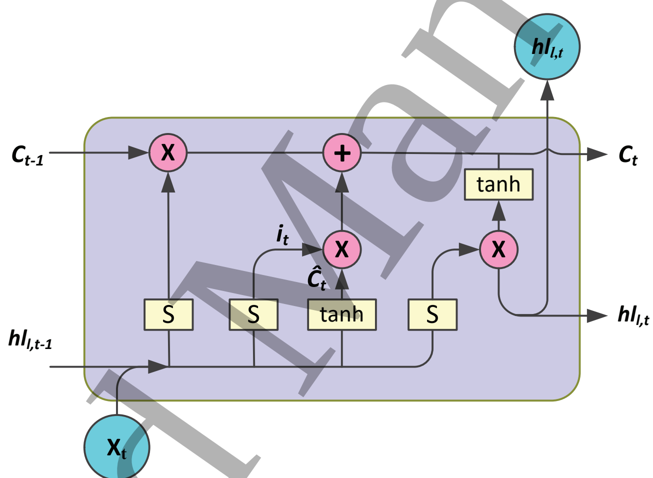

Figure 6. The architecture of LSTM cell, "S" denotes sigmoid operator, "tanh" denotes hyperbolic tangent operator, "+" is plus and "×" is multiplication. The "C_t" presents the state of LSTM cell at t moment

recurrent neural network (RNN), the main feature of LSTM is that there exist three gates: forget gate, external input gate and output gate. Forget gate decide the information which would be discard from prior cell and is given as

$$f_{l,t} = \sigma(\boldsymbol{W}_l^f \cdot [\boldsymbol{hl}_{l,t-1},\ \boldsymbol{x}_{l,t}] + \boldsymbol{b}_l^f), \tag{6}$$

where $\boldsymbol{h}_{l,t-1}$ denotes the output of prior cell, $\boldsymbol{x}_{l,t}$ is the input of the hidden layer, l denotes the lth hidden layer, $\boldsymbol{W}_l^f$ and $\boldsymbol{b}_l^f$ present the weight matrix and bias value respectively, σ is the sigmoid function to decide the degree of the information which would be forgotten.

External input gate operation has similar structure with forget gate and is set to learn the new knowledge to replace forgotten information

$$i_t = \sigma(\boldsymbol{W}_l^i \cdot [\boldsymbol{hl}_{l,t-1},\ \boldsymbol{x}_{l,t}] + \boldsymbol{b}_l^i), \tag{7}$$

$$\hat{\boldsymbol{C}}_t = \tanh(\boldsymbol{W}_l^c \cdot [\boldsymbol{hl}_{l,t-1},\ \boldsymbol{x}_{l,t}] + \boldsymbol{b}_l^c), \tag{8}$$

where $\boldsymbol{W}_l^i$, $\boldsymbol{W}_l^c$, $\boldsymbol{b}_l^i$ and $\boldsymbol{b}_l^c$ are weight matrix and bias value respectively. The state of LSTM cell $\mathbf{C}_{l,t}$ can be updated by

$$\boldsymbol{C}_{l,t} = f_{l,t} \times \boldsymbol{C}_{l,t-1} + i_{l,t} \times \hat{\boldsymbol{C}}_{l,t}. \tag{9}$$

Finally, the output $\boldsymbol{h}_{l,t}$ can be obtained through output gate

$$o_{l,t} = \sigma(\boldsymbol{W}_l^o \cdot [\boldsymbol{hl}_{l,t-1},\ \boldsymbol{x}_{l,t}] + \boldsymbol{b}_l^o), \tag{10}$$

$$\boldsymbol{hl}_{l,t} = o_{l,t} \times \tanh(\boldsymbol{C}_t). \tag{11}$$

The output vector of whole hybrid deep neural network y is equal with the output of LSTM which is $\boldsymbol{hl}_{3,5}$.

2.5. Training Process

The purpose of training process is to reduce the error between predicted value and actual label. Softmax function is selected to represent relative probabilities between different classes.

$$y_{p,m} = \frac{e^{y_m}}{\sum_m^T e^{y_m}}, \tag{12}$$

where m is the index of y, T is total number of classes. Moreover, the cross-entropy function is used as loss function by describing the distance between the probability distribution of the neural network prediction values $\boldsymbol{y}_p$ and the labels $\boldsymbol{y}_l$

$$L(\boldsymbol{y}_p, \boldsymbol{y}_l) = -\sum_m y_{p,m} \log y_{l,m}. \tag{13}$$

To minimize the loss function $L(\boldsymbol{y}_p, \boldsymbol{y}_l)$, adaptive moment estimation (ADAM) [39] is used to train the weights of the proposed CNN-LSTM algorithm. We define the learning rate $lr = 0.0001$ when we train the CNN-LSTM network.

3. Datasets and results

3.1. BCI Competition IV dataset 2a description

BCI Competition IV dataset 2a provided by Graz University [29] is used to train and test the proposed CNN-LSTM algorithm. The dataset contains EEG signal from 9 healthy subjects performing four different MI tasks: movement of the left hand, right hand, feet, and tongue. The train and test datasets of each subject includes 288 trials of data (72 for each of the four possible tasks) recorded with 22 EEG channels and 3 monopolar EOG channels. The signal were sampled at 250 Hz and band-pass filtered between 0.5 Hz and 100 Hz and notch filtered at 50Hz.

3.2. Result

In this study, the experiments were conducted in Matlab environment and Python environment on an Intel 4.00 GHz Core i7 and NVIDIA GTX 1070 PC with 16 GB RAM. Matlab was used to perform OVR-FBCSP for EEG signal and Tensorflow [40] was used for designing and testing the proposed CNN-LSTM network on Python platform. The weights of proposed CNN-LSTM network are initialed by truncated normal distribution function (mean = 0, std = 0.1).

In BCI Competition IV dataset 2a, the proposed algorithm was trained and tested for every subject. For each subject, datasets "T" was used to train the network and datasets "E" was used to evaluate. According to the position of the events which are stated in [41], we extract the valuable MI data in datasets "T" and datasets "E" of each subject (4s long, 22 EEG channels, 1,000 data points in each channel, total of 288 trials per session and 48 trials for each MI task in "T" and "E" respectively) and corresponding label for each sample. We train the hybrid network iteratively in 1000 times. The relationship between training error rate and training time is shown in Fig.7, and Fig.8 shows the signal sample visualization for the CNN-LSTM network perceived inputs for average class signal of subject 1. The performance of the proposed method was evaluated by accuracy achieved on the evaluation data. As we can see, the training error is less than 0.1 after 500 iterations about 14s, which can be considered acceptable, and total time of training process is 28s. After each training step, we also use evaluation data to test the neural network, and the evaluating accuracy curve is overlaid in Fig.7. After 500 iterations, the evaluating accuracy curve converges to around 0.9. The effect of different hidden layers of

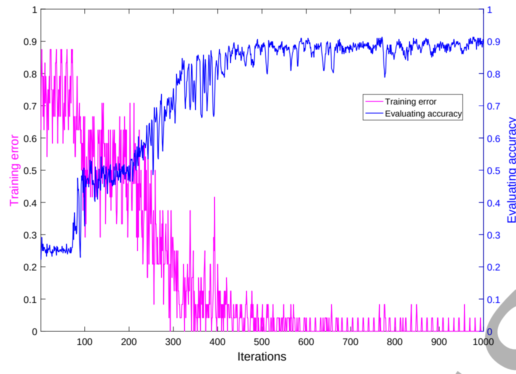

Figure 7. The relationship between training error, evaluating accuracy and iterations. We iterative training the hybrid networks in 1000 times. For each iteration, training batch is 24 sets of EEG training data and we also use evaluation dataset to evaluate the network after per step training. Total training time is about 28 seconds.

CNN and LSTM is shown in Fig.9 and Fig.10 by using kappa measures. Kappa is calculated as

$$\text{kappa} = \frac{p_o - p_e}{1 - p_e}. \tag{14}$$

In equation 14, p_o is the classification accuracy and p_e is the proportion of times the MI classes are expected to agree by chance alone. As shown in figure, the mean kappa of all subjects is higher than others for 3 CNN hidden layers and 3 LSTM hidden layers with a less computation time (about 28s). So in this paper, considering the computation time and the mean kappa of all subjects, we choose 3 CNN hidden layers and 3 LSTM hidden layers.

The classifiers based on only CNN, only LSTM and proposed hybrid CNN-LSTM networks were trained and tested separately for each subject in BCI Competition IV dataset 2a. It is obviously that each trained network is different because of the different initialization weights. So we ran 10 times the evaluation of each subject to obtain the statistics of the results by different initialization weights of CNN-LSTM neural networks. The accuracy results of performance are presented in Fig.11. We also applied SVM classifier for the EEG signal after OVR-FBCSP process in order to investigate the superiority of deep learning method. The average accuracy results of CNN, LSTM and CNN-LSTM methods are higher than SVM with the same input

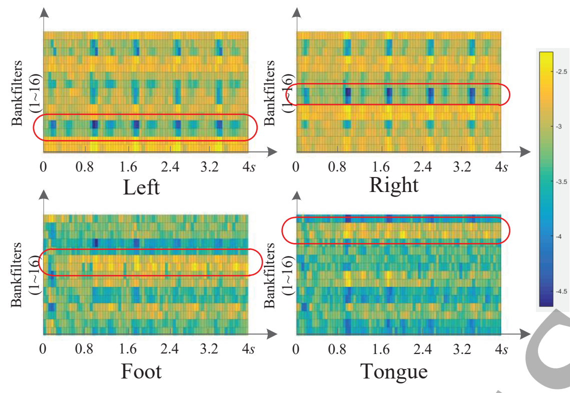

Figure 8. The visualization results for average class signal of 4-class MI task. The visualization results are the CNN-LSTM network perceived inputs. As can be seen, there are distinct signal characteristics which are marked in red circle on different MI task maps. These observations validate that the MI task feature information the hybrid neural network has learned is meaningful.

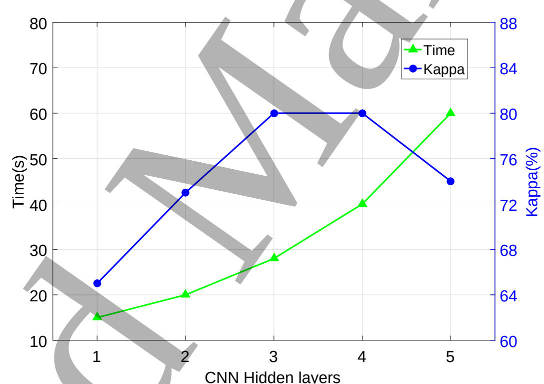

Figure 9. Effect of number of CNN hidden layers on training time and kappa value (3 hidden layers LSTM)

data. The average accuracy results of CNN-LSTM method are better than only CNN and only LSTM methods. For each subject, CNN-LSTM method is also the best way to classify MI signal.

Table 1 shows comparison of the proposed CNN-LSTM algorithm with the other methods [7, 21, 42–52] tested in BCI Competition IV Dataset 2a, including the winner of this Competition [8]. Moreover, LDA results were also presented as a baseline. The highest mean kappa value is highlighted in boldface. Although the

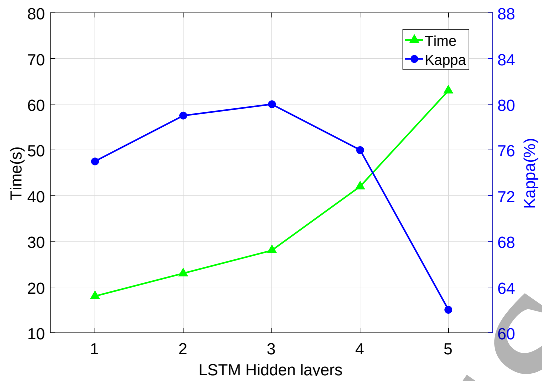

Figure 10. Effect of number of LSTM hidden layers on training time and kappa value (3 hidden layers CNN)

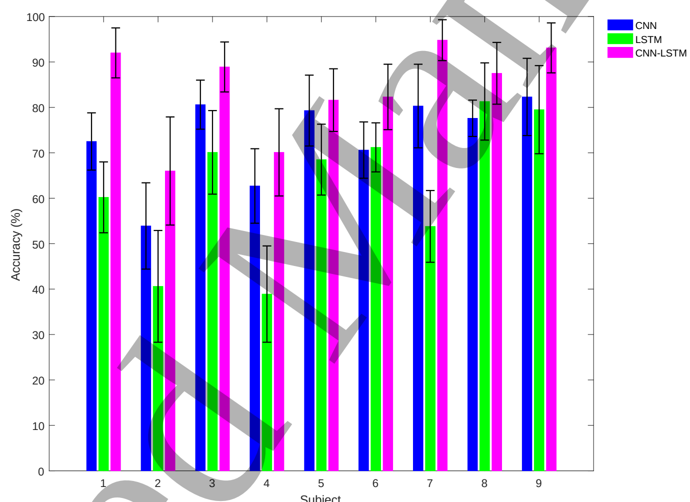

Figure 11. The accuracy results of classification by CNN, LSTM and CNN-LSTM

performance is difference between 9 subjects, our proposed method is superior to previously published methods in general. The corresponding mean Kappa values are 0.57 for FBCSP [8] and 0.60 for LDA, whereas it is 0.80 for CNN-LSTM, which demonstrates 40% improvement with respect to FBCSP and 33% improvement with respect to LDA on the accuracy of average mean Kappa values. This is showing that CNN-LSTM method is more effective and accurate to subjects than other methods.

All these results show that CNN-LSTM method provides more reliable classification with higher accuracy.

Table 1. Mean kappa values of the hybrid networks and competing methods on the BCI Competition IV Dataset 2a.

Methods	subjects									Average
	A01	A02	A03	A04	A05	A06	A07	A08	A09	
Ang et al. [8]	0.68	0.42	0.75	0.48	0.40	0.27	0.77	0.75	0.61	0.57
Gouy-Pailler et al. [42]	0.66	0.42	0.77	0.51	0.50	0.21	0.30	0.69	0.46	0.50
Wang [43]	0.67	0.49	0.77	0.59	0.52	0.31	0.48	0.75	0.65	0.58
Barachant et al. [44]	0.74	0.38	0.72	0.50	0.26	0.34	0.69	0.71	0.76	0.57
Wang et al. [45]	0.56	0.41	0.43	0.41	0.68	0.48	0.80	0.72	0.63	0.57
Kam et al. [46]	0.74	0.35	0.76	0.53	0.38	0.31	0.84	0.74	0.74	0.60
Asensio-Cubero et al. [47]	0.75	0.50	0.74	0.40	0.19	0.41	0.78	0.72	0.78	0.59
Asensio-Cubero et al. [48]	0.76	0.32	0.76	0.47	0.31	0.34	0.59	0.76	0.74	0.56
LDA	0.76	0.41	0.83	0.56	0.35	0.26	0.79	0.80	0.72	0.60
Blumberg et al. [7] (EM-LDA)	0.59	0.41	0.82	0.57	0.38	0.29	0.79	0.80	0.72	0.60
Vidaurre et al. [49] (PMean)	0.76	0.38	0.87	0.60	0.46	0.34	0.77	0.76	0.74	0.63
Luis et al [50].	0.83	0.51	0.88	0.68	0.56	0.35	0.90	0.84	0.75	0.70
Rebeca et al. [51]	0.84	0.55	**0.90**	0.71	0.66	0.44	0.94	**0.85**	0.76	0.74
Sakhavi et al. [21]	**0.88**	**0.65**	**0.90**	0.66	0.62	0.45	0.89	0.83	0.79	0.74
Ai et al. [52]	0.77	0.54	0.84	0.70	0.63	0.61	0.77	0.84	0.86	0.73
CNN-LSTM	0.85	0.54	0.87	**0.78**	**0.77**	**0.66**	**0.95**	0.83	**0.90**	**0.80**

Finally, we merged the data from all the 9 subjects in datasets "T" to train a subject-independent shared neural network to form one model and evaluate the shared neural network on datasets "E" for every subject respectively. The performance of mean accuracy and mean kappa value is shown in Table 2. Compared with the performance in Table 1, we can find that the shared network can get a good accuracy classification and even better than individual neural network. We think the reasons why the proposed shared network showed better performance are twofold: the training dataset is a bigger dataset and different subject EEG signal can reduce the overfitting of deep neural network. Moreover, the mean kappa values of time courses from 0.5s to 4s EEG data throughout evaluation session by CNN-LSTM for each subject are shown in Fig.12, respectively. The proposed CNN-LSTM algorithm also has a good classification effect on MI signal with a short time series.

4. Discussion and conclusions

In this paper, we propose a hybrid deep neural networks based on OVR-FBCSP, CNN and LSTM to discriminate four-class imagery motor tasks. The reasons of

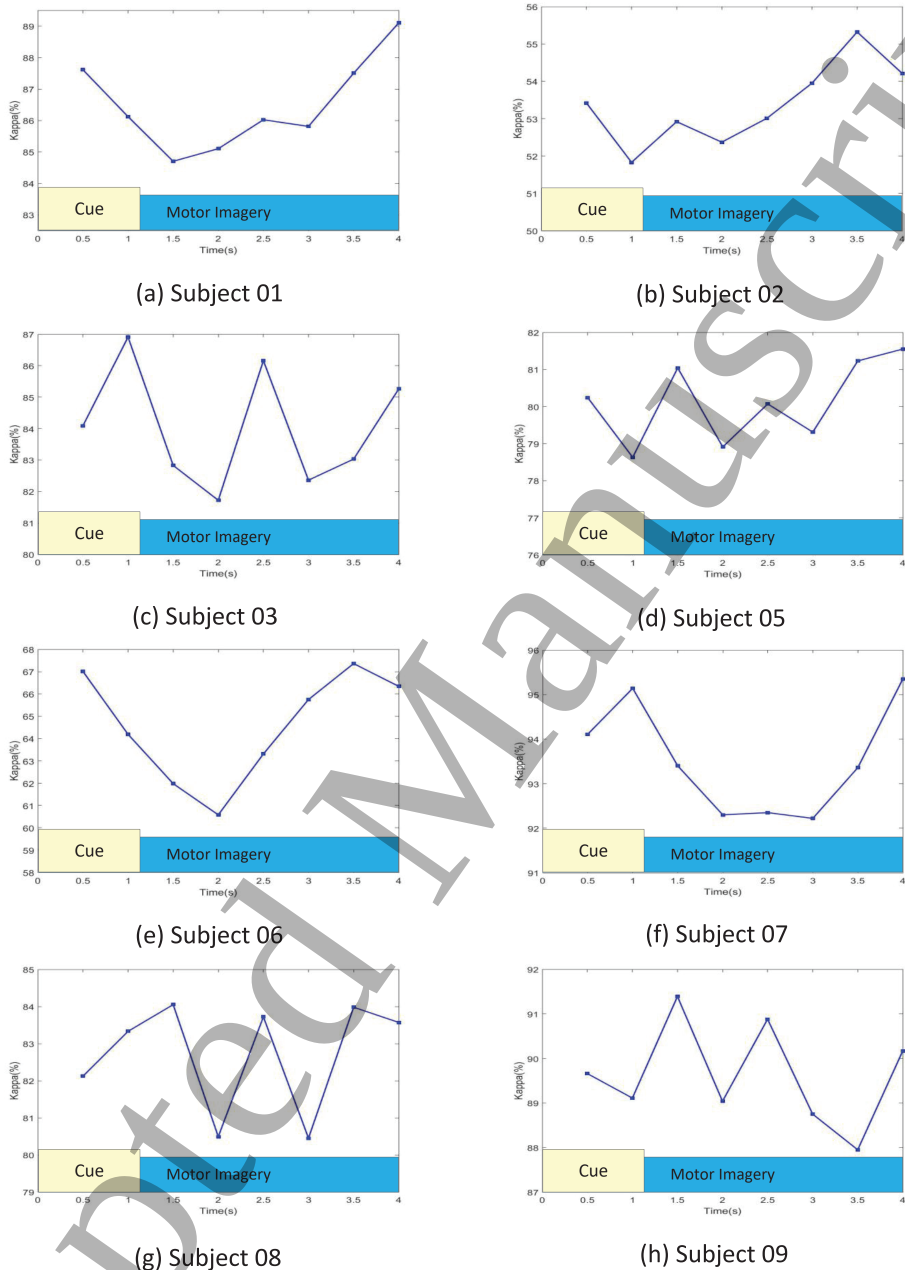

Figure 12. Time course of the mean kappa value for CNN-LSTM. Mean kappa value is computed obtaining the kappa time course for each subject from 01 to 09 (Because some data of 04 is missed in BCI Competition IV dataset 2a, we had no statistics on him).

Table 2. Mean accuracy and kappa values of merged training data

statistics	subjects									Average
	A01	A02	A03	A04	A05	A06	A07	A08	A09	
Accuracy	0.89	0.69	0.92	0.82	0.84	0.68	0.95	0.90	0.92	0.84
Kappa	0.87	0.59	0.90	0.76	0.82	0.66	0.95	0.89	0.89	0.81

good performance of proposed method maybe lie as following: Firstly, OVR-FBCSP make the features of each class more prominent. Then, the spatial and temporal features of MI signal can be learned effectively by the proposed hybrid deep learning scheme. Compared with other art algorithms, the experimental results indicate that the proposed scheme is very effective to improve the discriminant analysis results and also has a good classification effect on MI signal with a short time series.

We design a unified, end-to-end classification framework which combines two stages which are OVR-FBCSP and CNN-LSTM network. Exactly as the conclusion in [21], the OVR-FBCSP method is not affected by the network optimization and in turn, the network is forced to work with an input it has no control over. The OVR-FBCSP method is an excellent way to make the features more prominent and assist our proposed neural network to get a better accuracy of MI task classification.

Furthermore, many authors regard the EEG signal as an image and they dealt with EEG signal as dealing with images, like [17, 20, 26], so that they may lose temporal information of EEG signal. We not only extract the spatial features, but also mine the temporal features. We use a CNN network to extract the spatial features in detail and utilize LSTM to explore temporal correlation and sequence features of MI signal.

In practice, training a specific network for individual subject can be cumbersome, because it requires not only a lot of time but also computer memory. So, we merge the training data of all subjects to train a shared neural network which can learn the common features of all subjects. The evaluation result which is shown in Tab.2 shows the shared neural network is successful. The reasons of shared neural network better than individual neural network maybe lie as following: Firstly, merged dataset is a bigger dataset [53]. Bigger training dataset which contains not only the own subjects MI EEG features but also other subjects MI features can improve the classification accuracy of deep neural networks. Secondly, reducing overfitting of neural networks [54]. The merged data contains not only common features of MI task but also individual differences. These individual differences can make the neural networks reduce the overfitting effect thus to learn more common features.

In the future, we will be dedicated to the study of online MI classification algorithms and the application on controlling unmanned aerial vehicles by combining hybrid brain-computer-interface with virtual reality (VR) system.

5. Acknowledgments

This research was supported in the National Natural Science Foundation of China (No. 61673294, No. 61773278, No.61573060); the Ministry of Education Equipment Development Fund (6141A02022328, 6141A02033311).

References

[1] Graimann B, Pfurtscheller G and Allison B 2010 *Brain-Computer Interfaces* (Springer Berlin Heidelberg)
[2] Yuan H and He B 2014 *IEEE Transactions on Biomedical Engineering* **61** 1425–1435 ISSN 0018-9294
[3] Long J, Li Y, Wang H, Yu T, Pan J and Li F 2012 *IEEE Transactions on Neural Systems & Rehabilitation Engineering* **20** 720–729 ISSN 1534-4320
[4] Ma T, Li H, Deng L, Yang H, Lv X, Li P, Li F, Zhang R, Liu T and Yao D 2017 *Journal of Neural Engineering* **14** 026015
[5] Decety J and Ingvar D H 1990 *Acta Psychologica* **73** 13 – 34 ISSN 0001-6918
[6] Pfurtscheller G and Fh L D S 1999 *Clinical Neurophysiology* **110** 1842–1857 ISSN 1388-2457
[7] Blankertz B, Tomioka R, Lemm S and Kawanabe M 2008 *IEEE Signal Processing Magazine* **25** 41–56 ISSN 1053-5888
[8] Keng A K, Yang C Z, Wang C, Guan C and Zhang H 2012 *Frontiers in Neuroscience* **6** 39
[9] Kumar S, Sharma A and Tsunoda T 2017 *Bmc Bioinformatics* **18** 545
[10] Yang H, Sakhavi S, Kai K A and Guan C 2015 On the use of convolutional neural networks and augmented CSP features for multi-class motor imagery of EEG signals classification *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* pp 2620–2623 ISSN 1094-687X
[11] Novi Q, Guan C, Dat T H and Xue P 2007 Sub-band Common Spatial Pattern (SBCSP) for Brain-Computer Interface *2007 3rd International IEEE/EMBS Conference on Neural Engineering* pp 204–207 ISSN 1948-3546
[12] He L H and Liu B 2014 *Motor Imagery EEG Signals Analysis Based on Bayesian Network with Gaussian Distribution* (Springer International Publishing)
[13] Guo Y, Liu Y, Oerlemans A, Lao S, Wu S and Lew M S 2016 *Neurocomputing* **187** 27 – 48 ISSN 0925-2312 recent Developments on Deep Big Vision

[14] Sun X, Wu P and Hoi S C 2018 *Neurocomputing* **299** 42 – 50 ISSN 0925-2312

[15] Dong C, Chen C L, He K and Tang X 2016 *IEEE Transactions on Pattern Analysis & Machine Intelligence* **38** 295–307 ISSN 0162-8828

[16] Zeiler M D and Fergus R 2014 Visualizing and understanding convolutional networks *European conference on computer vision* (Springer) pp 818–833

[17] Lu N, Li T, Ren X and Miao H 2017 *IEEE Transactions on Neural Systems & Rehabilitation Engineering A Publication of the IEEE Engineering in Medicine & Biology Society* **25** 566–576 ISSN 1534-4320

[18] An X, Kuang D, Guo X, Zhao Y and He L 2014 A Deep Learning Method for Classification of EEG Data Based on Motor Imagery *International Conference on Intelligent Computing* pp 203–210

[19] Ahmed S, Merino L M, Mao Z, Meng J, Robbins K and Huang Y 2014 A Deep Learning method for classification of images RSVP events with EEG data *Global Conference on Signal and Information Processing* pp 33–36

[20] Tabar Y R and Halici U 2017 *Journal of Neural Engineering* **14** 016003

[21] Sakhavi S, Guan C and Yan S 2018 *IEEE Transactions on Neural Networks and Learning Systems* **29** 5619–5629 ISSN 2162-237X

[22] Schirrmeister R T, Springenberg J T, Ldj F, Glasstetter M, Eggensperger K, Tangermann M, Hutter F, Burgard W and Ball T 2017 *Human Brain Mapping* **38** 5391–5420

[23] Zhou J, Meng M, Gao Y, Ma Y and Zhang Q 2018 Classification of motor imagery eeg using wavelet envelope analysis and LSTM networks *2018 Chinese Control And Decision Conference (CCDC)* pp 5600–5605 ISSN 1948-9447

[24] Dose H, Muller J S, Puthusserypady S and Iversen H K 2018 A Deep Learning MI - EEG Classification Model for BCIs *2018 26th European Signal Processing Conference (EUSIPCO)* pp 1676–1679 ISSN 2076-1465

[25] Amin S U, Alsulaiman M, Muhammad G, Bencherif M A and Hossain M S 2019 *IEEE Access* **7** 18940–18950 ISSN 2169-3536

[26] Wu Y, Huang T H, Yi Lin C, Tsai S J and Wang P 2018 Classification of EEG Motor Imagery Using Support Vector Machine and Convolutional Neural Network *2018 International Automatic Control Conference (CACS)* pp 1–4

[27] Xu B, Zhang L, Song A, Wu C, Li W, Zhang D, Xu G, Li H and Zeng H 2018 *IEEE Access* **7** 6084–6093

[28] Ma X, Qiu S, Du C, Xing J and He H 2018 Improving EEG-Based Motor Imagery Classification via Spatial and Temporal Recurrent Neural Networks *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (IEEE) pp 1903–1906

[29] Tangermann M, Mller K R, Aertsen A, Birbaumer N, Braun C, Brunner C, Leeb R, Mehring C, Kai J M and Mllerputz G R 2012 **6** 55

[30] Cohen J 1960 *Educational and Psychological Measurement* **20** 37–46

[31] Dornhege G, Blankertz B, Curio G and Müller K R 2004 Increase information transfer rates in BCI by CSP extension to multi-class *Advances in neural information processing systems* pp 733–740

[32] Wu W, Gao X and Gao S 2006 One-versus-the-rest (OVR) algorithm: An extension of common spatial patterns (CSP) algorithm to multi-class case *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the* (IEEE) pp 2387–2390 ISSN 1094-687X

[33] Ramoser H, Mllergerking J and Pfurtscheller G 2000 *IEEE Transactions on Rehabilitation Engineering A Publication of the IEEE Engineering in Medicine & Biology Society* **8** 441

[34] Blankertz B, Losch F, Krauledat M, Dornhege G, Curio G and Mller K R 2008 *IEEE Transactions on Biomedical Engineering* **55** 2452–2462 ISSN 0018-9294

[35] Cun Y L, Boser B, Denker J S, Howard R E, Habbard W, Jackel L D and Henderson D 1990 *Advances in Neural Information Processing Systems* **2** 396–404

[36] Hara K, Saito D and Shouno H 2015 Analysis of function of rectified linear unit used in deep learning *International Joint Conference on Neural Networks* pp 1–8

[37] Hinton G E, Srivastava N, Krizhevsky A, Sutskever I and Salakhutdinov R R 2012 *Computer Science* **3** 212–223

[38] Graves A 2012 *Long Short-Term Memory* (Springer Berlin Heidelberg)

[39] Kingma D P and Ba J 2014 *arXiv preprint arXiv:1412.6980*

[40] Tensorflow homepage `https://tensorflow.google.cn/`

[41] Tangermann M, Muller K R, Aertsen A, Birbaumer N, Braun C, Brunner C, Leeb R, Mehring C, Miller K, Mueller-Putz G, Nolte G, Pfurtscheller G, Preissl H, Schalk G, Schlogl A, Vidaurre C, Waldert S and Blankertz B 2012 *Frontiers in Neuroscience* **6** 55

[42] Gouy-Pailler C, Congedo M, Brunner C, Jutten C and Pfurtscheller G 2010 *IEEE Transactions on Biomedical Engineering* **57** 469–478 ISSN 0018-9294

[43] Wang H 2011 *IEEE Transactions on Biomedical Engineering* **58** 1412–1420 ISSN 0018-9294

[44] Barachant A, Bonnet S, Congedo M and Jutten C 2012 *IEEE Transactions on Biomedical Engineering* **59** 920–928 ISSN 0018-9294

[45] Wang D, Miao D and Blohm G 2012 *Frontiers in Neuroscience* **6** 151

[46] Kam T E, Suk H I and Lee S W 2013 *Neurocomputing* **108** 58–68 ISSN 0925-2312

[47] Asensio-Cubero J, Gan J Q and Palaniappan R 2013 *Journal of Neural Engineering* **10** 046014

[48] Asensio-Cubero J, Gan J Q and Palaniappan R 2013 *Biomedical Signal Processing and Control* **8** 772 – 778 ISSN 1746-8094

[49] Vidaurre C, Kawanabe M, Von B P, Blankertz B and Mller K R 2011 *IEEE Transactions on Biomedical Engineering* **58** 587–597 ISSN 0018-9294

[50] Nicolas-Alonso L F, Corralejo R, Gomez-Pilar J and Hornero R 2015 *Neurocomputing* **159** 186–196 ISSN 0925-2312

[51] Nicolas-Alonso L F, Corralejo R, Gomez-Pilar J, lvarez D and Hornero R 2015 *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **23** 702–712 ISSN 1534-4320

[52] Ai Q, Chen A, Chen K, Liu Q, Zhou T, Xin S and Ji Z 2019 *Journal of neural engineering* **16** 026032

[53] Lecun Y, Bengio Y and Hinton G 2015 *Nature* **521** 436

[54] Schmidhuber J 2015 *Neural Netw* **61** 85–117