Neurocomputing 403 (2020) 452–461

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

A CNN-based comparing network for the detection of steady-state visual evoked potential responses

Jiezhen Xing[a,b,1], Shuang Qiu[a,1], Xuelin Ma[a,b], Chenyao Wu[a,b], Jinpeng Li[a,b], Shengpei Wang[a,b], Huiguang He[a,b,c,*]

[a] Research Center for Brain-inspired Intelligence and National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Science (CAS), Beijing 100190, China
[b] School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, China
[c] Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Beijing, China

ARTICLE INFO

Article history:
Received 1 July 2019
Revised 2 January 2020
Accepted 16 March 2020
Available online 5 May 2020

Communicated by Dr. Zhu Yu

Keywords:
Steady-State Visual Evoked Potentials
Brain-computer interface
Convolutional neural networks
Comparing network
Task-related component analysis

ABSTRACT

Brain-computer interfaces (BCIs) based on Steady-State Visual Evoked Potentials (SSVEPs) has been attracting much attention because of its high information transfer rate and little user training. However, most methods applied to decode SSVEPs are limited to CCA and some extended CCA-based methods. This study proposed a comparing network based on Convolutional Neural Network (CNN), which was used to learn the relationship between EEG signals and the templates corresponding to each stimulus frequency of SSVEPs. The effectiveness of the proposed method is validated by comparing it with the standard CCA and other state-of-the art methods for decoding SSVEPs (i.e., CNN and TRCA) on the actual SSVEP datasets collected from 23 subjects. The comparison results indicate that the CNN-based comparing network can significantly improve the classification accuracy. Furthermore, the comparing network with TRCA achieved the best performance among three methods based on comparing network with the averaged accuracy of 91.24% (data length: 2s) and 86.15% (data length: 1s). The study validated the efficiency of the proposed CNN-based comparing network in decoding SSVEPs. It suggests that the comparing network with TRCA is a promising methodology for target identification of SSVEPs and could further improve the performance of SSVEP-based BCI system.

1. Introduction

Brain-Computer Interfaces (BCIs) are communication systems that translate brain activities into computer commands without making use of peripheral nerve activity or muscles [1,2]. There are many applications such as spelling interfaces [3–5], playing computer games [6,7] and other assistive devices [8–10]. And it can provide environmental control capability for several disabled persons to improve life quality [11–14]. Recently, electroencephalogram (EEG)-based BCIs have achieved rapid progress in performance, functionality and practicality [15–19]. There are several typical EEG paradigms for BCIs, such as: sensorimotor mu/beta rhythms through motor imagine (MI) [16,17], P300 [18,41], and steady-state visual evoked potentials (SSVEPs) [19,26,40,43].

SSVEP-based BCI system has been widely used due to its high information transfer rate (ITR) compared with other EEG paradigms [20,21]. SSVEPs are brain responses elicited by stimulating the retina of the eyeball at a fixed frequency (1–100 Hz) [22]. The distribution of SSVEPs in the spectrum is multiplied by the frequency of visual stimulus. For example, when a 6 Hz fixed modulation frequency stimulates the eyeball, peaks at frequencies of 6 Hz, 12 Hz, and 18 Hz are observed in the spectrum [23].

So far, there are many methods for detecting SSVEPs. Based on the frequency feature of SSVEP signals, power spectrum density analysis (PSDA) through discrete Fourier transform (DFT) was widely used to detect target frequency of SSVEPs from each single channel [24,25]. Canonical correlation analysis (CCA) is the most popular method to detect SSVEPs tagged with frequency coding. It maximizes the correlation between EEG and sine-cosine reference signals in multi-channel [26]. Many modified CCA-based methods were proposed such as individual template-based CCA (IT-CCA) [27], multi-way CCA (MwayCCA) [28], L1-regularized multi-way CCA [29], and multi-set CCA (MsetCCA) [30], which showed better performance than standard CCA through optimizing the reference signals. In addition, Chen et al. proposed a filter bank

* Corresponding author at: Research Center for Brain-inspired Intelligence and National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, ZhongGuanCun East Rd. 95#, Beijing, 100190, China.
E-mail address: huiguang.he@ia.ac.cn (H. He).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.neucom.2020.03.048
0925-2312/

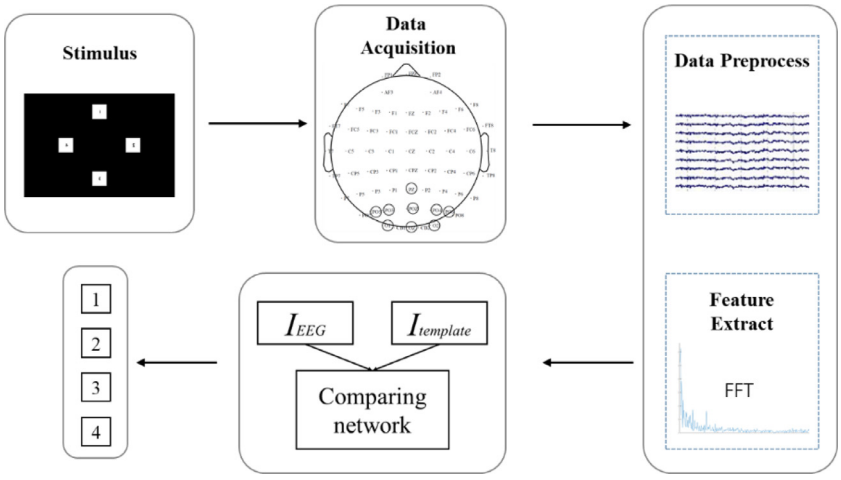

Fig. 1. The whole diagram of the experiment.

canonical correlation analysis (FBCCA) method to incorporate fundamental and harmonic frequency components to improve the detection of SSVEPs [31].

In 2013, Tanaka et al. proposed a new method which can extract task-related components from time-locked near-infrared spectroscopy (NIRS) data, named task-related component analysis (TRCA) [32]. TRCA can remove artifacts effectively to improve signal to noise ratio. By maximizing the reproducibility of time-locked activities across trials, the TRCA method showed good performance in extracting task-related components. SSVEPs are known as time-locked photic driving responses to repetitive visual stimulus [33,34]. Considering the stability of the latency in SSVEPs, the TRCA-based spatial filtering approach has great potential to improve the SNR of SSVEPs. In a recent study, Nakanishi et al. applied TRCA as a spatial filter to improve the SNR of SSVEPs, and the TRCA-based approach showed better performance than CCA [35].

Despite many examples of impressive progress to classify SSVEPs [36–38], there is still room for considerable improvement with respect to the accuracy. In recent years, convolutional neural network (CNN) is one of the largest advances in machine learning. CNNs allow automatic feature extraction within their layers, which means learning feature by itself instead of manually designing features. Moreover, SSVEPs has much abundant inner information. CNNs have advantages in dealing with this internal information compared with traditional methods. There are some advantages that traditional technology does not have: good fault tolerance, self-learning ability, good adaptive performance and high resolution. CNNs have better generalization than other methods. Thus, CNN has been used for decoding brain signals in some recent studies. In [37], Cecotti proposed a CNN architecture, which simulates FFT process, to classify time domain SSVEP signals. In 2017, Kwak et al. proposed a classifier based deep CNN that used frequency features as input for robust SSVEP detection and found highly encouraging SSVEP decoding results for the CNN architecture [39].

In this study, we aim to use deep learning methods to improve classification performance of SSVEPs. We combined advantages of deep learning and CCA and proposed a CNN-based comparing network to identify the target of SSVEPs. We designed templates to represent the prior information of the SSVEPs, such as SSVEP frequency, personal EEG pattern and so on. For our proposed network, these templates and SSVEP samples are the two kinds of inputs. Compared with [37], we use the frequency domain signal as the SSVEP input of the network. The feature extraction process begins with the frequency domain signal, so that our network can extract deeper feature of SSVEPs than the frequency domain features. Moreover, the network can learn the relationship between the SSVEP sample and template designed according to the prior information. Then, it solves the classification task of SSVEPs according to the relationship. Compared with other CNN methods [36–39], that the comparing network learns is the relationship between SSVEP sample and templates rather than the frequency feature of SSVEP.

Furthermore, we developed three different kinds of templates to build the comparing network. Each kind of template has four classes corresponding to each class of SSVEP. The effectiveness of the proposed comparing network was validated by comparing it with the traditional method CCA and a CNN-based method [39]. We proposed three kinds of templates, sinusoidal signal template, individual average template and TRCA (task-related component analysis) template, for the comparing networks. And comparison among three kinds of templates was taken to evaluate the performance of different kinds of templates. The performance was evaluated using a four-target SSVEP dataset recorded from twenty-three subjects.

2. Experiment

2.1. Subjects

Twenty-three healthy subjects (5 females, aged 20–28 years) with normal or corrected-to-normal vision participated in the experiment. Each subject was asked to read and signed an informed consent in advance. This study was approved by Research Ethics Committee of Institute of Automation, Chinese Academy of Science.

2.2. Stimulus design

This study used a sampled sinusoidal stimulation method [44,46] to present the visual flicks on a 24-inch LED monitor with resolution 1920 × 1080 pixels and a refresh rate of 60 Hz. The user interface is shown in the stimulus part of Fig. 1. There are 4 stimulus (4 cm × 4 cm), denoted by label 1, 2, 3 and 4. In order to produce the visual stimulus for the SSVEP, the squares are blinking at frequencies 7 Hz, 8 Hz, 9 Hz and 10 Hz. The stimulus sequence $s(f, i)$ can be generated by modulating the luminance of the screen, where 0 represents dark and 1 represents the highest luminance. The formula of $s(f,i)$ is shown as follow:

$$s(f, i) = \frac{1}{2}\{1 + \sin[2\pi f(i/\text{RefreshRate})]\} \quad (1)$$

where $sin()$ generates a sine wave, and i indicates the frame index in the sequence. The stimulus program was developed using Psy-

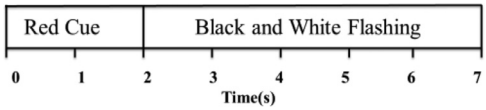

Fig. 2. Trial structure for collecting the SSVEPs data.

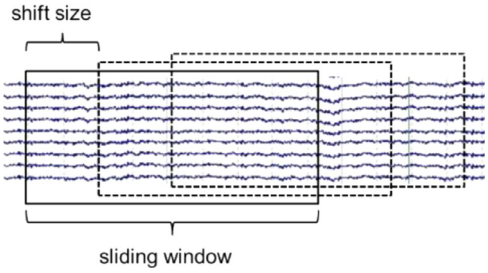

Fig. 3. The shift window and shift size for splitting SSVEP.

chophysics Toolbox Ver. 3 (PTB-3) [45] under Matlab (MathWorks, Inc.)

2.3. Data Acquisition

The EEG data were acquired by using a Synamp2 system (Neuroscan, Inc.) at a sampling rate of 1000Hz. The usable bandwidth of the system was 0.1–100 Hz. Nine electrodes (Pz, PO5, PO3, POz, PO4, PO6, O1, Oz, and O2) were placed over parietal and occipital areas according to the international 10–20 system (in the Data Acquisition part of Fig. 1). The reference electrode was M1 located in left ear mastoid. During the recording, impedances of all the electrodes were maintained below 10KΩ . A notch filter at 50 Hz was applied to remove power-line noise during recoding.

The experiment contains 20 blocks, and each block contains 20 trials (4 targets × 5) in a random order. As shown in Fig. 2, each trial lasted for 7 s. Each trial begins with a cue for 2 s, which is a red square at the target location with the same size as the stimulus. Subjects were asked to shift their gaze to the target as soon as possible. After the cue, visual stimulus starts to flicker for 5 s. There is no break between trials. After each block, there is a break for several minutes to avoid fatigue.

3. Method

3.1. Data preprocessing and input data organization

Data epochs comprising nine-channel SSVEPs were first extracted according to the event triggers. The frequency used is 7, 8, 9 and 10 Hz, so the data were filtered by a Butterworth bandpass filter from 5 to 70 Hz. This ensures that all harmonics are in the frequency band while removing noise. For easier analysis, all data were down-sampled to 250 Hz. And then, the data were segmented using a sliding window with a 40 ms shift size (Fig. 3).

The fast Fourier transform was adopted to transform the segmented data to frequency domain. In the frequency domain, 128 points (5–68 Hz frequency band) were selected and standardized by removing the mean and scaling to unit variance as the input data for CNN and compare methods.

$$x* = \frac{x-\mu}{\sigma} \tag{2}$$

3.2. Comparing network

In our study, we designed a comparing network, inspired by the CCA method. For the comparing network, the input layer has two different input data I_{EEG} and $I_{template}$. The network was used to learn the relationship between EEG signals and the template corresponding to one of four frequencies (7 Hz, 8 Hz, 9 Hz, 10 Hz).

3.2.1. Template design

As for $I_{template}$, we proposed three kinds of templates for the comparing network.

(a) Sinusoidal signal template. (Corresponding to COM-SIN)

For frequency f, we design the sinusoidal signal $s(f,t) \in \mathrm{R}^{N_s \times 1}$ as sinusoidal signal template by the following equation:

$$s(f,t) = \sum_{n=1}^{N_h} sin(2n\pi ft) + \cos(2n\pi ft) \tag{3}$$

where N_h is the number of harmonics ($N_h = 3$) Then, the sinusoidal signal above was processed using FFT and frequency band selection to form the sinusoidal signal template $\tilde{x}_f$ (size 128 × 1).

(b) Individual average template. (Corresponding to COM-IT)

This template was obtained by averaging multiple trials. Then, the averaged data was transformed to frequency domain.

$$\bar{x}_{IT,i} = FFT\left(\frac{1}{N_T}\sum_{n=1}^{N_T} x_{i,n}\right) \tag{4}$$

where N_T is the number of training trials; x_i is the EEG number of different trials. The size of $\bar{x}_{IT,i}$ is 128 × 9.

(c) TRCA (task-related component analysis) template. (Corresponding to COM-TRCA)

TRCA is the method that extracts task-related components efficiently by maximizing the reproducibility during the task period [32,42]. In previous study, Nakanishi et.al. used TRCA as a spatial filtering to improve the SNR of SSVEP[35]. We assume that EEG data consist of two parts: 1) task-related component $s(t)$; 2) task-unrelated component $n(t)$. The observed EEG signal $x(t)$ can be described as:

$$x_j(t) = a_{1,j}s(t) + a_{2,j}n(t), \quad j = 1,2,\ldots,N_c \tag{5}$$

where j is the index of channels, and $a_{1,j}$ and $a_{2,j}$ are mixing coefficients. The problem is to recover the task-related component $s(t)$ from a linear sum of observed signals $x(t)$ as:

$$y(t) = \sum_{j=1}^{N_c} w_j x_j(t) = \sum_{j=1}^{N_c} (w_j a_{1,j} s(t) + w_j a_{2,j} n(t)) \tag{6}$$

The task is to recover the task-related signal $s(t)$ from the observed signal $x(t)$. In reality, the correlation of the task-related signals between different periods is very large, while the task-independent signals are irrelevant. It is assumed that the correlation coefficient of the task-related data in different periods is a constant, while the correlation coefficient between the unrelated signals and between the task-independent signal and the task-related signal is 0. Then the solution to this problem is to maximize the covariance of the data between the trials.

$x^{(h)}(t)$ is the EEG data of h-th trial. And the estimated task-related signal is $y^{(h)}(t)$. T is the duration of trial. The covariance between h_1, h_2 trials is :

$$\begin{aligned} C_{h_1,h_2} &= Cov\left(y^{(h_1)}(t), y^{(h_2)}(t)\right) \\ &= \sum_{j_1,j_2}^{N_c} w_{j_1} w_{j_2} Cov\left(x^{(h_1)}(t), x^{(h_2)}(t)\right) \end{aligned} \tag{7}$$

All the possible combinations of trials are summed as:

$$\sum_{\substack{h_1,h_2=1\\h_1\neq h_2}}^{N_c} C_{h_1,h_2} = \sum_{\substack{h_1,h_2=1\\h_1\neq h_2}}^{N_t}\sum_{j_1,j_2=1}^{N_c} w_{j_1}w_{j_2} Cov(x_{j_1}^{h_1}(t), x_{j_2}^{h_2}(t)) = \mathbf{w^T S w} \tag{8}$$

To obtain a finite solution, the variance of y(t) is constrained as:

$$Var(y(t)) = \sum_{j_1,j_2=1}^{N_t} w_{j_1}w_{j_2}Cov(x_{j_1}(t), x_{j_2}(t)) = \mathbf{w^T Q w} = 1 \tag{9}$$

The problem can be solved by :

$$\hat{w} = \arg\max_{w} \frac{\mathbf{w^T S w}}{\mathbf{w^T Q w}} \tag{10}$$

The optimal coefficient vector is obtained as the eigenvector of the matrix $Q^{-1}S$.

The TRCA template has a spatial filter before FFT. For each template, N_T trials were used to train the template. Firstly, the preprocessed data was filtered by TRCA; secondly, we averaged the filtered data; at last, FFT and frequency band selection were used to form the TRCA template.

$$\bar{x}_{TRCA,i} = FFT\left\{\frac{1}{N_t}\sum_{n=1}^{N_T}[TRCA(x_{n,i})]\right\} \tag{11}$$

The size of $\bar{x}_{TRCA,i}$ is same as that of $\bar{x}_{IT,i}$ (128 × 9).

3.2.2. Architecture of the CNN

There are two CNN architectures used in the comparing network, named CNN-1 and CNN-2. The CNN-1 is shown in Fig. 4. Two convolutional layers are used to extract feature. The first convolutional kernel (size 1 × 9) extracts 9 feature maps (C1), and each map has 128 units. Then, the second convolutional kernel (size 11 × 1) is used and the feature map (C2) which has the same size as C1 is obtained.

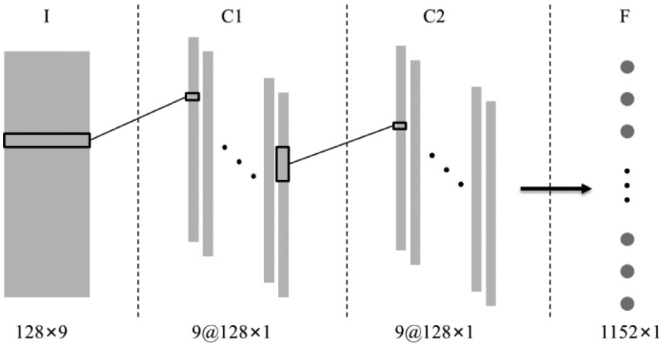

Fig. 4. The architecture of CNN in compare network.

In the first CNN layer, we convolute the data cross different channels to focus on the spatial feature. The second CNN layer is used to learn the frequency feature in the channel. After the extraction of two CNN layers, a full connected layer is used.

The structure of CNN-2 is simpler, which is adopted to sinusoidal signal template. The size of sinusoidal signal template is different from the others (size 128 × 1). So just a convolutional layer which convolutional kernel size is 11 × 1, and each map has 128 units.

3.2.3. The architecture of comparing networks

The architecture for different template is a little different from others. As for COM-SIN and COM-IT, the comparing networks are similar. As shown in Fig. 5(a), there are two kinds input data: one SSVEP data and four templates. After the feature extracting by CNN, the data is fully connected. The output layer has four nodes corresponding to four class. However, the CNN used for template is different. For COM-SIN, we adopt CNN-2, but for COM-IT, we adopt CNN-1.

The architecture for COM-TRCA is shown in Fig. 5(b). The branch $I_{template}$ has the same architecture. But for branch I_{EEG}, there are

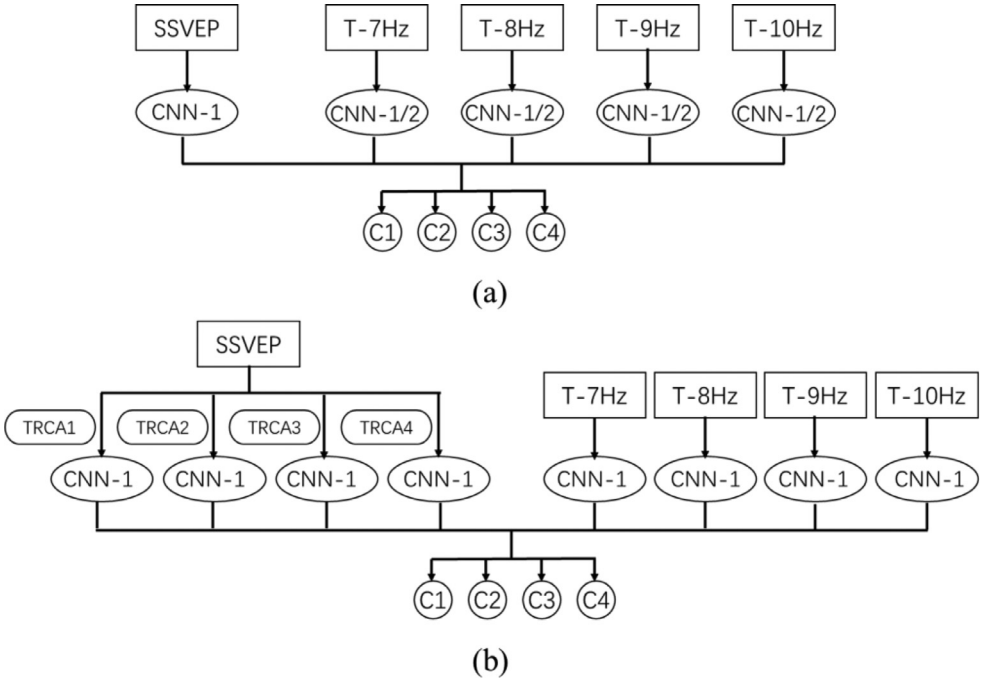

Fig. 5. The architecture of comparing networks.

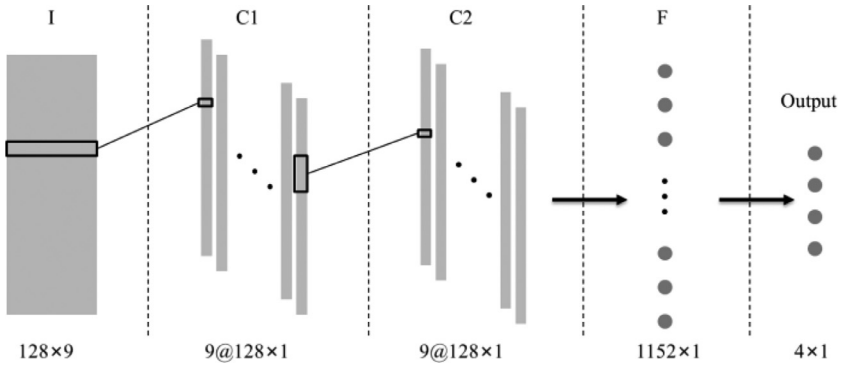

Fig. 6. The architecture of convolutional layer.

four inputs that are filtered by TRCA. After CNN, the data is fully connected and has four output nodes.

The gradient descent algorithm uses standard error backpropagation to update the network parameters. Except the output layer, the activation functions are ReLu. The output activation function is Softmax. The dropout rate is 0.5, and L2 regularization (coefficient is 0.5) is used.

3.3. Related method

3.3.1. Standard CCA

In SSVEP-based BCIs, CCA has been widely used to detect the frequency of SSVEPs [26]. CCA is a statistical way used to measure the underlying correlation between two multi-dimensional variables. Considering two multi-dimensional variables *X, Y* and their linear combinations $x=X^TW_X$ and $y=Y^TW_Y$, CCA finds the weight vectors, W_X and W_Y , which maximize the correlation between x and y by solving the following problem:

$$\max_{W_X,W_Y} r(x,y) = \frac{E\left[W_X^T XY^T W_Y\right]}{E\left[W_X^T XX^T W_Y\right]E\left[W_X^T YY^T W_Y\right]} \tag{12}$$

The maximum of r with respect to W_X and W_Y is the maximum canonical correction vector. In SSVEP detection, X refers to multi-channel SSVEP signals and Y refers to reference signals that have the same length as X. To detect the frequency of SSVEPs in an unsupervised way, sinusoidal signals for frequency f_i can be used as the reference signals Y_{fi}:

$$Y_{f_i} = \begin{bmatrix} \sin(2\pi f_i n) \\ \cos(2\pi f_i n) \\ \vdots \\ \sin(2\pi N_h f_i n) \\ \cos(2\pi N_h f_i n) \end{bmatrix}, n = \frac{1}{f_s}, \frac{2}{f_s}, \cdots \frac{N}{f_s} \tag{13}$$

where N_h is the number of harmonics; N is the number of sampling points.

To recognize the frequency of SSVEPs, CCA calculates the canonical correlation vector r_{f_i} between multi-channel SSVEP signals and the reference signals Y_{f_i} at each stimulus frequency. And then, we select the maximal correlation r_{max,f_i} from the canonical correlation vector. The class of the SSVEP is the class of the maximal correlation value r_{max} which is selected from all r_{max,f_i}:

$$\tilde{f} = class\{max\, r_{f_i,max}\},\ i = 1, 2, \ldots, N_f \tag{14}$$

In a recent study, Nakanishi et al. applied task-related component analysis as a spatial filter to improve the signal-to-noise ratio of SSVEP [35]. After TRCA processing, CCA can be used for classification.

3.3.2. CNN

In 2017, Kwak et al proposed a CNN-based classifier to detect the dominant frequency of the EEG signal [39]. As shown in Fig. 6, the CNN network has three layers, and each channel has one or several maps containing frequency information for the different channels (similar to [41]). The input data is defined as $I_{p,j} \in R^{N_{fs}\times N_{ch}}$, where, N_{fs}=128 is the number of frequency samples and N_{ch}=9 is the number of channels. A one-dimensional (1D) kernel of $1 \times N_{ch}$ is applied along the row. The first and second hidden layers are composed of N_{ch} feature maps, and each map in the first layer has a size of N_{fs}. Each map in the second layer is composed of 128 units. The output of the second layer is connected to a dense layer with 4 units, which represent the four classes of the SSVEP signals.

3.4. Voting for result

In order to increase the training data, sliding window method was used. Thus, we would get one output every 40 ms. It means that we have a result every 40 ms. In order to apply in real-word SSVEP-BCI, we consider to get one output every 1 s using two method. One method is to get a final result every interval 1 s. The other is that we integrated 25 continuous results and voted for the final result. So that, the interval between two results was 1s.

3.5. Experimental evaluation

Based on the dataset, we evaluated the performance of our proposed methods under different data length. In this work, the data were segment using 7 lengths of sliding window separately (1 s, 1.5 s, 2 s, 2.5s , 3 s, 3.5 s, 4 s). The shift size was set to 40ms for both time windows. It means that the overlaps are 0.96s and 1.96s respectively. Our data consists of 100 trials for each target. Each accuracy is obtained by averaging the accuracies of three random samplings. For example, if training set has 80 trials, we select 80 trials three times randomly, and the remaining 20 trials made up testing set. We compared our result with CCA, TRCA and CNN method proposed in [39]. The difference between different methods were tested with paired t-tests. The statistical significance was set at $p<0.05$.

4. Result

The experiments were taken under the data length range from 1 s to 4 s with an interval 0.5 s. Table 1 shows the classification accuracy with various length of data. As for the result, we take a two-way repeated measures ANOVA to measure the effects of different factors. This result showed significant main effects of different data length ($F_{(6,\ 132)} = 61.11$, $p<0.0001$), different classification

Table 1
SSVEP classification accuracy by the data length.

Data length(s)	1	1.5	2	2.5	3	3.5	4
CCA	73.12±18.15***	80.88±16.78***	85.32±15.29**	88.06±13.92*	89.86±12.97*	91.15±12.02	92.13±11.22
TRCA	73.30±17.96***	81.07±16.74***	85.40±15.41**	88.29±13.58*	89.87±13.02*	91.39±11.73	91.98±11.33
CNN	74.06±19.62***	76.90±18.70***	81.57±17.01***	76.90±19.39***	78.64±18.01***	77.90±18.22***	81.54±17.29***
COM-SIN	81.63±15.79***	84.31±15.28***	88.18±13.40*	87.94±13.32*	89.59±11.97*	89.23±12.45*	89.53±12.96*
COM-IT	81.69±15.79**	84.33±15.29***	88.09±13.45*	87.87±12.96*	89.98±11.80*	89.14±12.73**	89.85±12.56*
COM-TRCA	86.15±14.81	89.39±12.66	91.24±11.79	91.87±11.01	93.36±9.67	93.09±10.68	94.31±8.27

The asterisks indicate significant difference between COM-TRCA and other methods by paired t-tests ($* p < 0.01$, $** p < 0.001$, $*** p < 0.001$).

methods ($F_{(5, 110)} = 27.7$, $p<0.0001$), and interaction effect between factors ($F_{(30, 660)} = 22.76$, $p<0.0001$).

For each kind of data length, a one-way ANOVA is taken to measure the effects of method. The results show that there is a significant effect of different method (1s: F=33.22 $p<0.0001$, 1.5s: F=27.05 $p<0.0001$, 2s: F=17.03 $p<0.0001$, 2.5s: F=26.91 $p<0.0001$, 3s: F=26.91 $p<0.0001$, 3.5s: F=30.96, 4s: F=21.04, $p<0.0001$) on classification accuracy in each kind of data length. For each kind of data length, we compared COM-TRCA and the other methods. Pos hos analysis shows COM-TRCA has better performance for each data length. Therefore, the COM-TRCA has the best performance no matter how long the data is. The performance of TRCA has improvement than that of CCA, while the increase of accuracy was less than 1%. In one previous study, Kwak et al reported that the accuracy of CNN was much higher than that of CCA by 16%. However, in our study, CNN did not perform better than CCA based on our recorded SSVEPs. It may not only due to the subject variance, but also due to a "trial-dependent" issue which was caused by the training and the testing data were from the same trial in Kwak et al.'s study [39].

For each kind of data length, every method based on comparing network (COM-SIN, COM-IT, COM-TRCA) significantly increased the classification accuracy of SSVEPs compared with CCA, TRCA and CNN ($p < 0.001$). Therefore, our comparing network showed better performance in classification of SSVEPs. Based on comparing network, three kinds of templates were applied to generalize the characteristics of SSVEPs, and then, to learn the deep relationship between templates and EEG input data. It indicates that combining deep learning and prior information derived from SSVEP paradigm would enhance target detections in SSVEP-based BCIs. In addition, the comparison of accuracy among three methods based on comparing network showed that the comparing network with TRCA template (COM-TRCA) performed best in each case. Since data processed by TRCA removed the task-unrelated EEG activities and increased the ratio of task-related EEG activities, it could increase the distance between target and not-target template feature values and decrease the distance between target and target temple feature values. Therefore, Compared with COM-SIN and COM-IT, COM-TRCA improved performance in target identification.

Furthermore, a one-way ANOVA to measure the effects of data lengths for COM-TRCA. There is a significant effect of data length ($F = 21.48$, $p<0.0001$) on classification results. As the length of the data grows, the averaged classification accuracy increases. When data length shorter than 2 s, classification accuracy increases faster (about 5% improvement just with 1 s change). This is in consistence with the finding in previous studies that the accuracy of SSVEP detection increased as the data length increased [15,31,35]. More information is included in the samples with data length 2 s than those with data length 1 s. So, the accuracies of 2 s were higher than those of 1 s. But the accuracy rate does not increase much when the data length is more than 2 s (increase 3% from 2 s to 4 s).

For other methods, a one-way ANOVA is taken to measure the effects of data lengths. The results show that the data length has a significant influence on classification accuracy (CCA: F=66.99, $p<0.0001$, TRCA: F=59.65, $p<0.0001$, CNN: F=15.20, $p<0.0001$, COM-SIN: F=57.72, $p<0.0001$, COM-IT: F=39.46, $p<0.0001$).When the data length growing, except CNN, the classification effect of other methods is gradually improved. But CNN cannot maintain this trend when data length is longer than 2 s. CCA and TRCA have almost the same trend of change. When data length is 1 s, their accuracies is the lowest. But the accuracy will increase rapidly. As for COM-SIN and COM-IT, their changes are similar. When data length shorter than 2 s, their classification accuracy has increased rapidly, but when longer than 2 s, the accuracy increasing trend slows down.

As the data length becomes longer, a single sample contains more information. This is very beneficial to feature extraction and classification, resulting in the increasing accuracy for the 5 methods. However, as for deep learning method, CNN, COM-SIN, COM-IT and COM-TRCA, their performance can be affected by training set size. If fewer samples are trained for the deep learning method, they will not be trained appropriately. Under the influence of the above two factors, the four deep learning methods cannot get the same increase as CCA and TRCA, and the CNN even cannot keep increase.

While the longer data length, the higher classification accuracy, long data length is not suitable for BCI system to control other device. At present, many researches focus on the data length 2s and shorter [26–30].

The performance of different methods is shown in Fig. 7, where the classification confusion matrices of subject 3 is shown. The accuracy of CCA is based on all data, and the others are based on 80 trials training in 3 random sampling. The data length is 2 s. The values in diagonal of every matrix are the correct classification rate and the others are the misclassification rate. Fig. 7(a), (b) and (c) are the result from CCA, TRCA and CNN, respectively. Although the correct classification is majority, there are many misclassifications. Fig. 7(d), (e) and (f) are the result from COM-SIN, COM-IT and COM-TRCA, respectively. The misclassification is less than the other three methods. Evidently, the proposed methods significantly reduced the misclassifications, especially for COM-TRCA.

T-distributed Stochastic Neighbor Embedding (t-SNE) [47] was used to embed the data into two dimensions for drawing a scatter plot. Fig. 8 shows an example of visualization from one subject. As revealed in Fig. 8(a), we can see maximum overlap in t-SNE visualization for raw data. After extracting frequency feature, the samples of each class clustered into multiple block mass obviously but still cannot be linearly separated, which can be seen in Fig. 8(b). As the Frequency features are processed over the convolutional layers, the separation between four classes after the first convolutional layer is clearly visible in Fig. 8(c), and after the second convolutional layer, separation becomes bigger in Fig. 9 (d). The output of the third layer is linearly separable in Fig. 8(e). Thus, the visibility of separation increases from raw data to the output of the comparing network.

Fig. 9 shows the averaged SSVEP classification accuracy across all subjects obtained by CCA, TRCA, CNN, COM-SIN, COM-IT, COM-TRCA in three kinds of result output mode. A two-way repeated measures ANOVA as used to measure the effect of different factor.

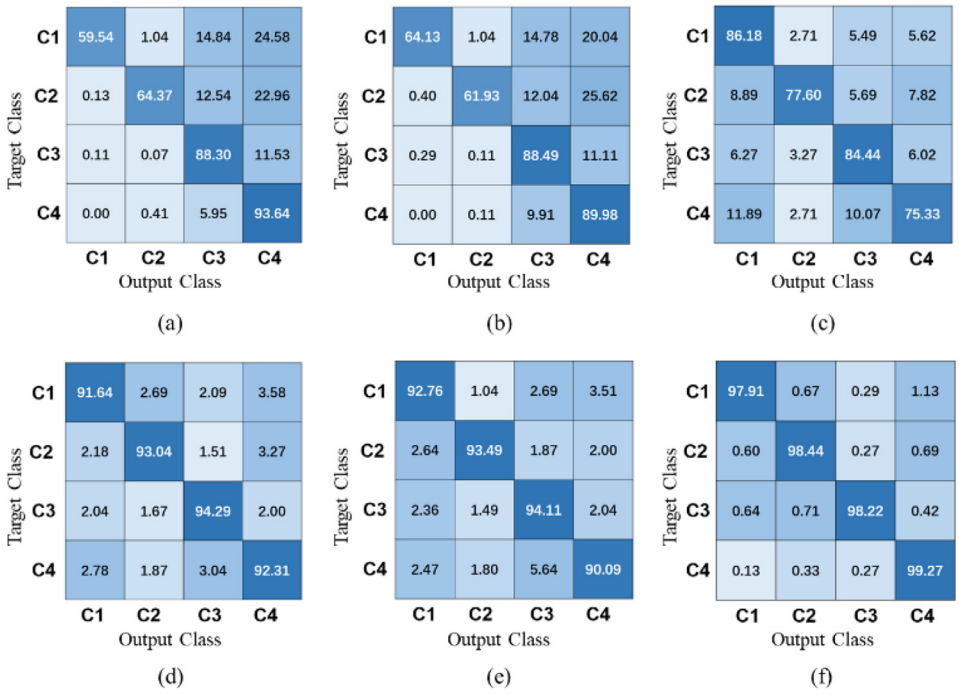

Fig. 7. The classification confusion matrices. (a) CCA method; (b) TRCA method; (c) CNN method; (d) COM-SIN method; (e) COM-IT method; (f) COM-TRCA method.

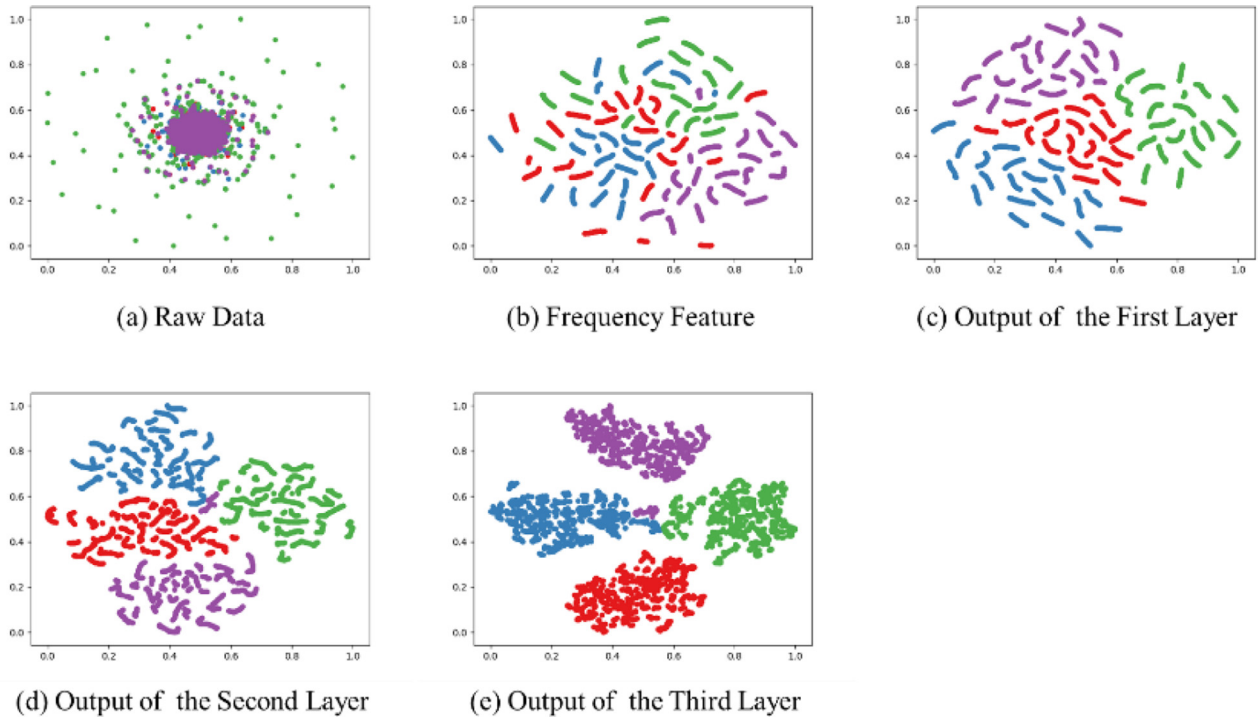

Fig. 8. An example of t-SNE Visualization for COM-TRCA: (a) raw data, (b) Frequency feature, (c) the output of first layer, (f) the output of second layer, (e) the output of third layer from one subject. Every point represents a sample and every color represents one class of SSVEP.

For data length 1 s, the result showed significant main effects of statistical ways ($F_{(2,\ 44)} = 46.01$, $p < 0.0001$), classification method ($F_{(5,\ 110)} = 15.78$, $p < 0.0001$) and interaction effect between factors ($F_{(10,\ 220)} = 9.305$, $p < 0.0001$). For data length 1 s, the result showed significant effects of statistical ways ($F_{(2,\ 44)} = 67.38$, $p < 0.0001$) and classification method ($F_{(5,\ 110)} = 22.5$, $p < 0.0001$). Post hoc pairwise comparisons reveal that vote method performed best among three result output ways. Therefore, voting is an effective method to increase the classification accuracy for SSVEPs.

Table 2

Different depth models	Accuracy
Model 1	91.27±12.06
Model 2	91.06±11.78
Model 3	90.60±12.02
Model 0	91.24±11.79

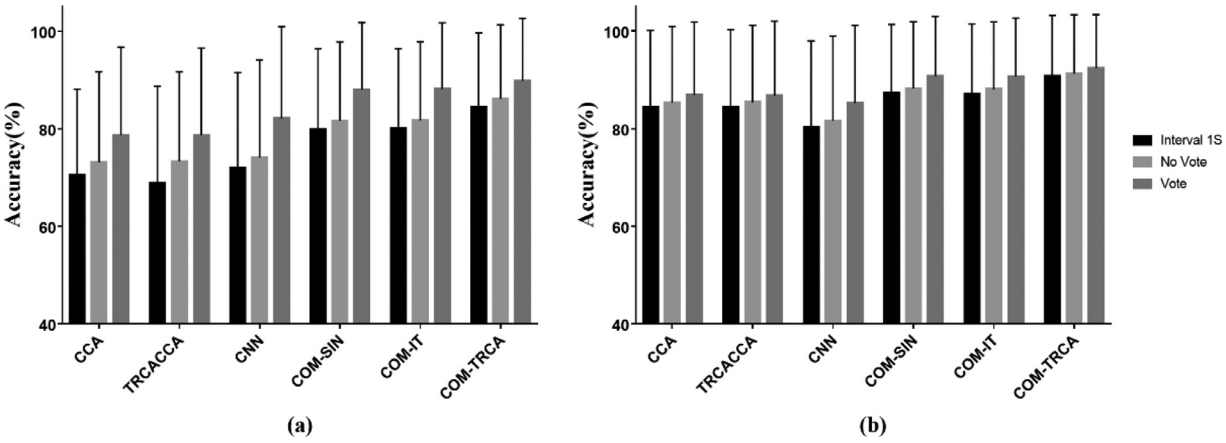

Fig. 9. Averaged accuracy comparison of six methods in three kinds of result output mode. (a) data length 1s; (b) data length 2s. Vote represents the final result voted among every 25 intermediate results; No Vote represents the intermediate results; Interval 1S represents the results got every 1S.

Table 3
SSVEP classification accuracy by the quantity of training data.

Trials quantity	30	40	50	60	70	80
TRCA	85.26±15.27	85.56±15.11	85.36±15.17	85.42±15.30	85.39±15.35	85.40±15.41
CNN	76.09±19.15	77.81±18.45	78.59±18.62	78.88±18.39	80.00±17.95	81.57±17.01
COM-SIN	84.51±15.78	85.65±15.04	86.14±14.52	86.72±14.19	86.86±14.46	88.18±13.40
COM-IT	84.65±15.58	85.63±14.98	86.41±14.29	86.87±14.18	86.92±14.43	88.09±13.45
COM-TRCA	89.93±12.91	90.51±12.33	91.01±11.87	90.69±12.41	91.58±11.29	81.24±11.78
CCA	85.32±15.29					

Based on the COM-TRCA, we have further studied the models with different depths under the condition of 80 training trails and 2-s data length. Three different changes are taken based on the COM-TRCA(marked as Model 0): 1) add a convolution layer for channels after the C1 layer (convolutional kernel is 1×9), marked as Model 1; 2) add a convolution layer for frequency after the C2 layer (convolutional kernel is 11×1), marked as Model 2; 3) Add a convolution layer after the C1 and C2 layers respectively, marked as Model 3.

Table 2 shows the classification accuracy of these four models. The differences between COM-TRCA (Model 0) and each other models are not significant. Although Model 1 has 0.03% higher mean accuracy than Model 0, the SD value of Model1 is higher than that of Model0. The accuracies of Model 2 and 3 are lower than that of Model 0. Since, Model 1, Model2, and Model 3 have more convolutional layers than model 0, which result in that more calculate costs. However, there is no significate improvement of classification accuracy. Therefore, Model 0 is optimal model to classify SSVEPs based on COM-TRCA.

Table 3 shows the classification accuracy of different quantity of training data with data length 2s. We conduct our experiment in the situation that training data ranged from 40 to 80 trials and the remaining data were used as testing data. Note that the CCA method stay constant, since no training phase is required. In CCA, the canonical correlations and synchronization index with reference signals are simply computed in order to find the maximum value.

A two-way ANOVA showed significant main effects of classification methods ($F_{(5,110)} = 20.81$, $p < 0.0001$), and quantity of training data ($F_{(4,88)} = 27.19$, $p < 0.0001$), and a significant interaction effect between factors (classification methods and quantity of training data), $F_{(20,\ 440)} = 6.338$, $p < 0.0001$. The data in Table 1 shows that as the amount of training data decreases, the accuracy of CNN, COM-SIN, COM-IT, and COM-TRCA results will decrease, which is expected. The smaller the amount of data, the less representative of the data, the lower the generalization ability of the trained model, and the lower the accuracy of the corresponding test. The results obtained by the TRCA method are almost unchanged and are very close to the CCA results. CNN shows the worst performance in either case. COM-SIN and COM-IT are similar in performance. When the training amount is greater than 50, the classification accuracies are signification higher than that of CCA. At 40 and 50, the result is higher than CCA, but the significance disappears. For any amount of training data, COM-TRCA maintains the best performance compared to other methods. When the amount of training data decreases, the accuracy of COM-TRCA decreases, but the change is small. Even at 30 training amount, COM-TRCA can significantly improve the classification accuracy compared to CCA. In summary, COM-TRCA has excellent performance in the case of less training data. Therefore, it is feasible in practical BCI system applications.

5. Conclusion and future work

In the present work, we proposed a novel architecture for SSVEP signal classification, where the prior knowledge of SSVEPs was taken into account. The results of this study indicate that comparing networks have better performance compared with standard CCA, TRCA and CNN. Furthermore, the comparing network with TRCA template had the largest improvement in SSVEPs classification. This research has thus made certain encouraging attempts for the application of deep learning in SSVEPs classification, and our results will be of significant interest to the BCI community.

In this work, the CNN used to build comparing network is a two-layer CNN structure. In the future, we will develop different networks from the perspective of the architecture.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Jiezhen Xing: Conceptualization, Methodology, Data curation, Writing - original draft. **Shuang Qiu:** Formal analysis, Investigation, Writing - review & editing. **Xuelin Ma:** Data curation. **Chenyao Wu:** Data curation. **Jinpeng Li:** Methodology. **Shengpei Wang:** Writing - review & editing. **Huiguang He:** Funding acquisition, Writing - review & editing.

Acknowledgment

This work was supported by National Natural Science Foundation of China (Nos. 81701785, 61976209), CAS International Collaboration Key Project (No. 173211KYSB20190024), and Strategic Priority Research Program of CAS (No. XDB32040000).

Reference

[1] M.A. Lebedev, M.A.L. Nicolelis, Brain–machine interfaces: past, present and future, Trends Neurosci. 29 (9) (2006) 536–546.
[2] K.R. Müller, A. Kübler, Toward brain computer interfacing, Eur. J. Neurol. 16 (3) (2010) e52.
[3] K.R. Müller, B. Blankertz, Toward noninvasive brain-computer interfaces, IEEE Signal Process. Mag. 23 (5) (2006) 125–128.
[4] E. Donchin, K.M. Spencer, R. Wijesinghe, The mental prosthesis: assessing the speed of a p300-based brain-computer interface, IEEE Trans. Rehabil. Eng. 8 (2) (2000) 174–179.
[5] H.J. Hwang, J.H. Lim, Y.J. Jung, H. Choi, S.W. Lee, C.H. Im, Development of an ssvep-based bci spelling system adopting a qwerty-style led keyboard, J. Neurosci. Methods 208 (1) (2012) 0–0.
[6] M. Krauledat, K. Grzeska, M. Sagebaum, B. Blankertz, C. Vidaurre, K.-R. Müller, M. Schröder, Playing Pinball with Non-invasive BCI, (2009) 1641–1648.
[7] R. Krepki, B. Blankertz, G. Curio, Klaus-Robert Müller, The berlin brain-computer interface (bbci) – towards a new communication channel for online control in gaming applications, Multimed. Tools Appl. 33 (1) (2007) 73–90.
[8] L.R. Hochberg, M.D. Serruya, G.M. Friehs, J.A. Mukand, M. Saleh, A.H. Caplan, et al., Neuronal ensemble control of prosthetic devices by a human with tetraplegia, Nature 442 (7099) (2006) 164–171.
[9] J.L. Collinger, B. Wodlinger, J.E. Downey, W. Wang, E.C. Tylerkabara, D.J. Weber, et al., High-performance neuroprosthetic control by an individual with tetraplegia, Lancet 381 (9866) (2013) 557–564.
[10] J.H. Kim, F. Biessmann, S.W. Lee, Decoding three-dimensional trajectory of executed and imagined arm movements from electroencephalogram signals, IEEE Trans. Neural Syst. Rehabil. Eng. 23 (5) (2015) 1.
[11] A. Ramos-Murguialday, D. Broetz, M. Rea, L. Läer, Yilmaz Ö., F.L. Brasil, G. Liberati, M.R. Curado, E. Garcia-Cossio, A. Vyziotis, W. Cho, M. Agostini, E. Soares, S. Soekadar, A. Caria, L.G. Cohen, N. Birbaumer, Brain-machine interface in chronic stroke rehabilitation: A controlled study, Ann. Neurol. 74 (2013) 100–108.
[12] G. Morone, I. Pisotta, F. Pichiorri, S. Kleih, S. Paolucci, M. Molinari, et al., Proof of principle of a brain-computer interface approach to support post-stroke arm rehabilitation in hospitalized patients: design, acceptability, and usability, Arch. Phys. Med. Rehabil. 96 (3) (2015) S71–S78.
[13] A. Nilsson, K.S. Vreede, V.H. glund, H. Kawamoto, Y. Sankai, J. Borg, Gait training early after stroke with a new exoskeleton – the hybrid assistive limb: a study of safety and feasibility, J. Neuroeng. Rehabil. 11 (2014) 92.
[14] A. Tsukahara, Y. Hasegawa, K. Eguchi, Y. Sankai, Restoration of gait for spinal cord injury patients using hal with intention estimator for preferable swing speed, IEEE Trans. Neural Syst. Rehabil. Eng. 23 (2) (2015) 308–318.
[15] Y. Wang, X. Chen, X. Gao, S. Gao, A benchmark dataset for ssvep-based brain–computer interfaces, IEEE Trans. Neural Syst. Rehabil. Eng. 25 (10) (2016) 1746–1752.
[16] A. Venkatakrishnan, G.E. Francisco, J.L. Contreras-Vidal, Applications of brain—machine interface systems in stroke recovery and rehabilitation, Curr. Phys. Med. Rehabil. Rep. 2 (2014) 93–105.
[17] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, K.R. Muller, Optimizing spatial filters for robust EEG single-trial analysis, IEEE Signal Process. Mag. 25 (1) (2007) 41–56.
[18] B. Blankertz, S. Lemm, M. Treder, S. Haufe, K.R. Müller, Single-trial analysis and classification of ERP components—a tutorial, NeuroImage 56 (2) (2011) 814–825.
[19] G.R. Müller-Putz, R. Scherer, C. Brauneis, G. Pfurtscheller, Steady-state visual evoked potential (SSVEP)-based communication: impact of harmonic frequency components, J. Neural Eng. 2 (4) (2005) 123.
[20] X. Chen, Y. Wang, M. Nakanishi, X. Gao, T.P. Jung, S. Gao, High-speed spelling with a noninvasive brain–computer interface, Proce. Natl. Acad. Sci. 112 (44) (2015) E6058–E6067.
[21] S. Gao, Y. Wang, X. Gao, B. Hong, Visual and auditory brain–computer interfaces, IEEE Trans. Biomed. Eng. 61 (5) (2014) 1436–1447.
[22] D. Zhu, J. Bieger, G.G. Molina, R.M. Aarts, A survey of stimulation methods used in SSVEP-based BCIs, Comput. Intell. Neurosci. 2010 (2010) 1.
[23] M. Middendorf, G. McMillan, G. Calhoun, K.S. Jones, Brain-computer interfaces based on the steady-state visual-evoked response, IEEE Trans. Rehabil. Eng. 8 (2) (2000) 211–214.
[24] G. Yang, L. Xiao, H. Shen, Y. Bian, L. Zhao, S. Cui, Design and implementation of a brain-computer interface based on virtual instrumentation, Int. J. Adv. Mechatron. Syst. 2 (1-2) (2010) 36–45.
[25] Y. Wang, R. Wang, X. Gao, B. Hong, S. Gao, A practical VEP-based brain-computer interface, IEEE Trans. Neural Syst. Rehabil. Eng. 14 (2) (2006) 234–240.
[26] Z. Lin, C. Zhang, W. Wu, X. Gao, Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs, IEEE Trans. Biomed. Eng. 53 (12) (2006) 2610–2614.
[27] G. Bin, X. Gao, Y. Wang, Y. Li, B. Hong, S. Gao, A high-speed BCI based on code modulation VEP, J. Neural Eng. 8 (2) (2011) 025015.
[28] Y. Zhang, G. Zhou, Q. Zhao, A. Onishi, J. Jin, X. Wang, A. Cichocki, Multiway canonical correlation analysis for frequency components recognition in SSVEP-based BCIs, in: Proceedings of International Conference on Neural Information Processing, Springer, Berlin, Heidelberg, 2011, pp. 287–295.
[29] Y. Zhang, G. Zhou, J. Jin, M. Wang, X. Wang, A. Cichocki, L1-regularized multiway canonical correlation analysis for SSVEP-based BCI, IEEE Trans. Neural Syst. Rehabil. Eng. 21 (6) (2013) 887–896.
[30] Y.U. Zhang, G. Zhou, J. Jin, X. Wang, A. Cichocki, Frequency recognition in SSVEP-based BCI using multiset canonical correlation analysis, Int. J. Neural Syst. 24 (04) (2014) 1450013.
[31] X. Chen, Y. Wang, S. Gao, T.P. Jung, X. Gao, Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain–computer interface, J. Neural Eng. 12 (4) (2015) 046008.
[32] H. Tanaka, T. Katura, H. Sato, Task-related component analysis for functional neuroimaging and application to near-infrared spectroscopy data, NeuroImage 64 (2013) 308–327.
[33] Regan, D. (1989). Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine.
[34] F.B. Vialatte, M. Maurice, J. Dauwels, A. Cichocki, Steady-state visually evoked potentials: focus on essential paradigms and future perspectives, Prog. Neurobiol. 90 (4) (2010) 418–438.
[35] M. Nakanishi, Y. Wang, X. Chen, Y.T. Wang, X. Gao, T.P. Jung, Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis, IEEE Trans. Biomed. Eng. 65 (1) (2017) 104–112.
[36] V. Bevilacqua, G. Tattoli, D. Buongiorno, C. Loconsole, D. Leonardis, M. Barsotti, M. Bergamasco, A novel BCI-SSVEP based approach for control of walking in virtual environment using a convolutional neural network, in: Proceedings of the 2014 International Joint Conference on Neural Networks (IJCNN), IEEE, 2014, pp. 4121–4128.
[37] H. Cecotti, A time–frequency convolutional neural network for the offline classification of steady-state visual evoked potential responses, Pattern Recognit. Lett. 32 (8) (2011) 1145–1153.
[38] H. Cecotti, A. Graeser, Convolutional neural network with embedded Fourier transform for EEG classification, in: Proceedings of the 2008 19th International Conference on Pattern Recognition, IEEE, 2008, pp. 1–4.
[39] N.S. Kwak, K.R. Müller, S.W. Lee, A convolutional neural network for steady state visual evoked potential classification under ambulatory environment, PloS One 12 (2) (2017) e0172578.
[40] M. Attia, I. Hettiarachchi, M. Hossny, S. Nahavandi, A time domain classification of steady-state visual evoked potentials using deep recurrent-convolutional neural networks, in: Proceedings of International Symposium on Biomedical Imaging, 2018.
[41] H. Cecotti, A. Graser, Convolutional neural networks for P300 detection with application to brain-computer interfaces, IEEE Trans. Pattern Anal. Mach. Intell. 33 (3) (2010) 433–445.
[42] H. Tanaka, T. Katura, H. Sato, Task-related oxygenation and cerebral blood volume changes estimated from NIRS signals in motor and cognitive tasks, NeuroImage 94 (2014) 107–119.
[43] X. Chen, Z. Chen, S. Gao, X. Gao, A high-itr ssvep-based bci speller, Brain-Comput. Interfaces 1 (3-4) (2014) 181–191.
[44] D.H. Brainard, S. Vision, The psychophysics toolbox, Spatial Vis. 10 (1997) 433–436.
[45] F. Sung, Y. Yang, L. Zhang, T. Xiang, P.H. Torr, T.M. Hospedales, Learning to compare: relation network for few-shot learning, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 1199–1208.
[46] N.V. Manyakov, N. Chumerin, A. Robben, A. Combaz, M. van Vliet, M.M. Van Hulle, Sampled sinusoidal stimulation profile and multichannel fuzzy logic classification for monitor-based phase-coded SSVEP brain–computer interfacing, J. Neural Eng. 10 (3) (2013) 036011.
[47] L.V.D. Maaten, G. Hinton, Visualizing data using t-SNE, J. Mach. Learn. Res. 9 (Nov) (2008) 2579–2605.

Jiezhen Xing received the B.E. degree in biomedical engineering from Beijing Institute of Technology, Beijing, China in 2016 and the M.S. degree in computer technology from Research Center for Brain-inspired Intelligence and National Laboratory of Pattern Recognition,Institute of Automation, Chinese Academy of Science (CAS), Beijing, China in 2019. His current research interests include pattern recognition, machine learning and their application in brain-computing interfaces.

Shuang Qiu received the B.S. and Ph.D. degrees in biomedical engineering from Tianjin University, Tianjin, China, in 2010 and 2017, respectively. From 2014 to 2016, she was a visiting student with the Department of Physical Medicine and Rehabilitation, Harvard Medical School, Boston, MA, USA. She is currently an Associate Professor with the Research Center for Brain-Inspired Intelligence, Institute of Automation, Chinese Academy of Sciences, Beijing, China. Her current research interests include biosignal processing, machine learning, and their applications in rehabilitation technology, with emphasis on humanmachine interface and neurofeedback treatment.

Xuelin Ma received the B.S. degree in School of Automation Science and Engineering, South China University of Technology, Guangzhou, in 2015. He is currently pursuing the Ph.D. degree with the Research Center for Brain-Inspired Intelligence, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences (CASIA), Beijing. He is also with the School of Artificial Intelligence, University of Chinese Academy of Sciences (UCAS), Beijing. His current research interests include pattern recognition, machine learning, deep learning and their applications in Brain-Computer Interface.

Chenyao Wu received the B.E. degree in Intelligence Science and Technology from the University of Science and Technology Beijing, Beijing, China, in 2019. She is currently pursuing the M.E. degree with the Research Center for Brain-Inspired Intelligence, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing. She is also with the School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing. Her current research interests include pattern recognition, machine learning, deep learning, and their applications in brain-computer interfaces.

Jinpeng Li received the B.E. degree (2012) and M.E. degree (2015) in Automatic Control from University of Science and Technology, Beijing (USTB), China. He received Ph.D. degree (2019) in Pattern Recognition and Intelligent Systems from Institute of Automation, Chinese Academy of Sciences (CASIA). He is now a researcher at HwaMei Hospital, University of Chinese Academy of Sciences, and Institute of Life and Health Science, University of Chinese Academy of Sciences. His research interests include pattern recognition, machine learning, deep learning, transfer learning algorithms and their applications in brain-computer interface and medical image analysis.

Shengpei Wang received the B.E. degree in measurement and control technology and instrument from the East China University of Science and Technology. He is currently pursuing the Ph.D. degree with Research Center for Brain-Inspired Intelligence, National Laboratory of Pattern Recognition, Institute of Automation Chinese Academy of Sciences, Beijing. His current research interests including pattern recognition, computational neurosciences, brain network, neuroimaging analysis calculation, and their applications in neurogenic diseases and brain-computer interfaces.

Huiguang He (M'04–SM'10) received the B.S. and M.S. degrees from Dalian Maritime University (DMU), Dalian, China, in 1994 and 1997, respectively, and the Ph.D. degree (Hons.) in pattern recognition and intelligent systems from the Institute of Automation, Chinese Academy of Sciences (CASIA), Beijing, China. From 1997 to 1999, he was an Associate Lecturer with DMU. From 2003 to 2004, he was a Post-Doctoral Researcher with the University of Rochester, Rochester, NY, USA. From 2014 to 2015, he was a Visiting Professor with the University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. He is currently a Full Professor with the Research Center for Brain-Inspired Intelligence, National Laboratory of Pattern Recognition, CASIA. He is also with the School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, and the Center for Excellence in Brain Science and Intelligence Technology, CAS. His research has been supported by several research grants from the National Science Foundation of China. He has authored or co-authored over 150 peer-reviewed papers. His current research interests include pattern recognition, medical image processing, and brain–computer interfaces. Dr. He was a recipient of the Excellent Ph.D. dissertation of CAS in 2004, the National Science and Technology Award in 2003 and 2004, the Beijing Science and Technology Award in 2002 and 2003, the K.C. Wong Education Prizes in 2007 and 2009, and the Jia-Xi Lu Young Talent Prize in 2009. He is an Excellent Member of Youth Innovation Promotion Association, CAS in 2016.