Signal Processing: Image Communication 47 (2016) 549–555

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

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

EEG-based prediction of driver's cognitive performance by deep convolutional neural network

Mehdi Hajinoroozi [a,*], Zijing Mao [a], Tzyy-Ping Jung [b], Chin-Teng Lin [c], Yufei Huang [a]

[a] Department of Electrical and Computer Engineering, University of Texas at San Antonio, USA
[b] Institute for Neural Computation, University of California, San Diego, CA, USA
[c] Brain Research Center, National Chiao Tung University, Hsinchu, Taiwan

ARTICLE INFO

Article history:
Received 30 October 2015
Received in revised form
27 May 2016
Accepted 28 May 2016
Available online 30 May 2016

Keywords:
Deep neural network
Convolutional neural network
Cognitive states

ABSTRACT

We considered the prediction of driver's cognitive states related to driving performance using EEG signals. We proposed a novel channel-wise convolutional neural network (CCNN) whose architecture considers the unique characteristics of EEG data. We also discussed CCNN-R, a CCNN variation that uses Restricted Boltzmann Machine to replace the convolutional filter, and derived the detailed algorithm. To test the performance of CCNN and CCNN-R, we assembled a large EEG dataset from 3 studies of driver fatigue that includes samples from 37 subjects. Using this dataset, we investigated the new CCNN and CCNN-R on raw EEG data and also Independent Component Analysis (ICA) decomposition. We tested both within-subject and cross-subject predictions and the results showed CCNN and CCNN-R achieved robust and improved performance over conventional DNN and CNN as well as other non-DL algorithms.

1. Introduction

Motor vehicle accidents are a serious economic and societal problem in US and around the world. The National Highway Traffic Safety Administration (NHTSA) estimated that motor vehicle crashes result in a staggering $871 billion per year in economic loss and societal harm. Among the leading causes of the crashes, driver's fatigue and drowsiness contributed to a conservative estimate of over 100,000 crashes annually, which result in an estimated 1500 deaths and $12.5 billion monetary losses. Therefore, developing technologies to monitor and predict driver's performance or the ability to safely drive the vehicle will have significant economic and societal payoff. For such prediction, different data sources including physiological measurements and video images of the driver have been considered in the literature, we focus in this paper on the prediction of driver's cognitive performance based on continuously measured electroencephalograph (EEG) brain signals.

The prediction of driver's cognitive performance based on EEG measurements can be treated as a classification problem, where the goal is to discriminate between the driver's states related to good or poor performance. For this classification, extracting discriminative features from the EEG data is an important part and developing approaches that consider high dimensionality as well as spatial and temporal correlations of EEG data is essential. Existing research has provided a number of solutions that utilise information from raw EEG signals and EEG signals in a transferred domain as features of classifiers. For instance, independent component analysis (ICA) was applied to the EEG data and classifiers that used independent components (ICs) selected based on the prior knowledge were demonstrated in a test involving 10 subjects to provide good performance in predicting driver's fatigue [1,2]. Time-frequency EEG power spectrum was also demonstrated as useful features in [3] on data from 16 subjects. A combination of raw EEG and time-frequency power spectrum features was also investigated for data from 4 subjects in [4].

Despite these advancements, robust and accurate prediction of driver's cognitive performance from EEG data still remains a challenge. First of all, while most of the existing work concerns prediction of driver's fatigue, prediction of driver's cognitive performance is much more difficult because driver's performance corresponds to short-duration, fast-varying brain states. However, such prediction is a more useful indication of driving safety than drowsy/alert prediction because being alert might not perform. Secondly, most of the existing results were obtained based on a very limited sessions from a small number of subjects, it is unclear how robust a proposed algorithm is especially for cross-session/cross-subject predictions. Thirdly, feature extraction methods and classifiers that can better capture the characteristics of EEG data are clearly needed to further improve the prediction performance. Particularly, a method that can reliably and automatically extract and select the best discriminative features from the EEG signals is a highly desirable. To address these issues, we examined 3 different

* Corresponding author.

http://dx.doi.org/10.1016/j.image.2016.05.018
0923-5965/

publications [1,2,5] that performed the same experiments on driver's fatigue and assembled EEG data from 37 subjects including 80 sessions. This represents one of the largest collections of EEG samples for driver's fatigue study. To take advantage of this large collection, we investigate solutions by deep learning. Deep learning is a collection of machine learning models that possess a unique layered architecture [6,7,18]. The most popular DL algorithms include deep neural networks (DNN) and convolutional neural networks (CNN) and they have been shown especially in computer vision and speech processing to have superior ability to learn sophisticate features from data [8,9]. Application of DL to EEG-based neuroscience and brain-computer interface is still limited and all the existing work focused on DL solutions for target event detection in a rapid serial visual presentation (RSVP). Particularly, in [10] DNN was shown to achieve superior performance in both cross-participant and cross-session target event detection and in [11], CNN coupled with well-known spatial filters for detection of visual evoked events including xDAWN and CSP was shown to provide robust performance than other existing non-DL based algorithms. However, DL for driver's performance detection has not been investigated in the literature and the filters for detecting visual evoked events might not be appropriate for a CNN solution in this case. In this paper, we propose a novel channel-wise convolutional deep neural network (CCNN) architecture and discuss two specific algorithms. We investigate the new CCNN algorithms on raw EEG and ICA-transformed data and test them for both within-subject and cross-subject predictions. The results show robust and improved performance of CCNN over conventional DNN and CNN as well as other non-DL algorithms.

The remaining part of the paper is organised as follows. In Section 2, the experiments, the information of the EEG data, and the prediction goal are discussed in detail. In Section 3, the CCNN architecture is proposed and the derivation of the backpropagation formula for CCNN-R is detailed. In Section 4, the test results for within-subject and cross-subject prediction are presented and in Section 5, the conclusion is drawn.

2. Description of the data and formulation of the problem

EEG measurements were collected from three studies of driver's cognitive performance conducted in a virtual reality dynamic driving simulator [1–3], where the designs of all studies are the same. Thirty electrodes are used with a modified version of the international 10–20 system (standard electrode positioning nomenclature, American Encephalographic Association). The impedance of each electrode was adjusted below 5 k Ω. The EEG signals were down-sampled to 250 Hz and filtered using a band-pass FIR filter (1–50 Hz) in this analysis. Specially, the virtual reality environment includes a car fixed on a platform that can be adjusted for either the motion or motionless mode. In case of the motion mode, the platform's movement is synchronized with the driving scenes. The virtual reality driving scenes simulate a night time driving with 100 km/h on a highway with two straight lanes in each direction. A perturbation is introduced into driving path every 8–12 seconds to push the car off the lane to either left or right randomly. As soon as a perturbation is applied, the driver is required to steer the car back to the driving lane. Each session lasted about one and half hours, during which subjects' EEG were measured by NeuroScan EEG headset with 30 electrodes (channels) and 250 Hz sampling rate. The entire dataset includes 80 sessions from 37 subjects. Driver's reaction time (RT) and the amount of the lane deviation were also measured to assess the degree of driver's driving performance. RT is defined as the time between the onset of the lane perturbation and the moment when the subject starts steering the car. When $RT \leq 0.7$ s, the driver is considered with a good driving performance, whereas when $RT \geq 2.1$ s, the driver is considered with a poor driving performance [2].

The goal of this paper is to predict driver's poor vs. good driving performance based on EEG data. Because the data collected right before perturbation can best measure the concerned cognitive states, EEG data samples from 5 s before each perturbation that led to either a poor or a good driving performance state were retained. Furthermore, following the practice introduced in [2], we divided the retained samples into 1-second (1 s) epochs and labelled each one as either poor or good performance based on the driver's state after the corresponding perturbation. Because the sampling rate is 250 Hz and there are 30 EEG channels, the dimension of a 1 s EEG epoch is $250 \times 30 = 7,500$ and out of the 80 session, we extracted a total of 60,000 epochs (35,000 good driving performance vs 25,000 poor driving performance epochs). Our goal is to train a classifier by using these 1 s EEG epochs to best predict the cognitive performance. We discuss in the following a new deep learning model for this purpose.

3. Proposed channel-wise convolutional neural network for EEG-based driver's cognitive state prediction

A straightforward approach to apply DL to a 1 s EEG epoch is to stack the EEG channels into a $7,500 \times 1$ feature vector and treat this vector or certain transformation of the vector as the visible units for a DNN. However, an alternative DL architecture that considers the characteristics of EEG data inside its architectural design could potentially lead to better performance. For an EEG epoch, samples from each channel records 1 s time series of brain dynamics, whereas samples cross channels possess different degrees of correlations due to the spatial locations of the channels on the headset. To capture these unique characteristics, we resorted to CNN [20]. CNN was originally designed for image recognition and in CNN, local correlations of the image pixels are being taking into account by convolving small 2-D patches with filters of shared weights. Inspired by this feature of CNN, we sought to design a filter that better fits EEG within CNN. However, the conventional filter that applies to an image patch cannot be readily applied to an EEG epoch, even though an EEG epoch is naturally presented in a 2 dimensional (channel $\times$ time) format. This is because the arrangement of channels in an EEG epoch cannot reflect their spatial adjacency and therefore direct filtering across channels in an epoch cannot correctly capture its spatial correlation. We have already investigated channel-wise, time-wise, and windowed filters on EEG epochs for deep belief network and showed that the channel-wise filter provided the best performance [18]. Based on this preliminary result, we propose next a channel-wise convolutional neural network (CCNN).

3.1. Channel-wise convolutional neural network (CCNN)

Just like CNN, CCNN also includes multiple convolutional layers followed by fully connected layers. However, in each convolutional layer of CCNN, instead of a filter applied to a 2-D patches as in CNN, a 1-D filter convolved with each channel (256×1 in our case) is applied. Because, the filter is shared for all channels, CCNN has much less parameters than a DNN to estimate. Pooling layers and dropout layers can also be included. The output of the last convolutional layer will be concatenated and then fed into fully connected layers. Since this concatenated output would normally be of much smaller size than that of the input features, the fully connected layers would need much less connections than a DNN trained directly on the input features, thus making CCNN computational also more appealing than DNN. Training CCNN weights

also uses backpropagation. We refer the architecture as the basic CCNN when a convolutional filter such as the Gaussian or Xavier filter is employed. The basic CCNN filter can be readily implemented using existing computational frameworks such as Caffe. However, more sophisticate models can be used in place of a convolutional filter but detailed solutions for backpropagation need to be derived. We discuss next a CCNN that uses Restricted Boltzmann Machine (RBM) in its convolutional layer.

3.2. Channel-wise convolutional neural network with RBM (CCNN-R)

In CCNN-R, RBMs replace the convolutional filters and the convolutional layers become essential a deep belief network (DBN) [21,22]. However, different from a fully connected DBN that take an entire epoch as an input, RBMs in CCNN-R take only a channel as an input. This modification can be also viewed as first applying a DBN on channels as input feature and then using a fully connected DNN on the concatenated outputs of DBN. In CCNN-R we examined different activation functions like sigmoid function, hyperbolic tangent and rectified linear units (ReLUs) and the best classification performance was obtained from sigmoid function. Therefore, we used sigmoid functions in our CCNN-R model as activation function. In addition, one successful approach for generalisation in neural networks and avoiding overfitting is dropout technique [23]. In CCNN-R we are making use of dropout and our experimental results show that dropout works very well in CCNN-R and prevents overfitting in our model. In CCNN-R a 90% dropout achieved the best performance. We also applied max-pooling in CCNN-R but it did not improve the classification performance. Therefore we do not implement max-pooling in CCNN-R.

In order to train the CCNN-R, the detailed solutions for the backpropagation needs to be derived, due mainly to the unique structure of the convolutional part of CCNN-R. The backpropagation of CCNN-R includes backpropagation in the fully connected layer, backpropagation in the convolutional layer, and backpropagation from the convolutional layer to the input layer. Notice that the backpropagation in the fully connected layer is the same as that for DNN. Therefore, we focus our discussion next on the backpropagation of the convolutional layers. Fig. 1 shows the feed forward process of CCNN-R with dropout in all layers.

In this figure L is the layer size and X^l is the input of layer l, Y^l is the output of layer l, W^l is the weight between the neural units, where l is the current layer number. Following the feed forward and average error calculation with respect to the cost function $E=\Sigma_{p=1}^{k}\left(t_p^L-y_p^L\right)^2/2d$, where d is the number of training epochs, t_p^L is the true label, y_p^L is the estimated label and k is the label size, the error can be back propagated into the fully connected layers and convolutional layer of the CCNN-R. In order to update all the weights in the architecture, we follow the gradient descent method show in (1),

$$\omega^{new}=\omega^{old}-\gamma\left(\partial E/\partial\omega\right) \tag{1}$$

where γ is the learning rate. The derivatives of cost function w.r.t. ω is separated into three parts. The first part is weight ($w_{i,p}^{L-1}$) between two fully connected layers (layer $L-1$ and L). This weight updating is similar as that in DNN, which generates (2).

$$\delta_p^L=\frac{\partial E}{\partial y_p^L}=e_p\left(y_p^L\right)=\left(t_p^L-y_p^L\right)/d \tag{2}$$

The second part is connected weight ($w_{c,n,i}^{L-2}$) between fully connected layer and convolution layer (layer $L-1$ and $L-2$). Eq. (3) implements E w.r.t. this weight,

$$\frac{\partial E}{\partial w_{c,n,i}^{L-2}}=\frac{\partial E}{\partial x_i^{L-1}}\frac{\partial x_i^{L-1}}{\partial w_{c,n,i}^{L-2}}=\frac{\partial E}{\partial x_i^{L-1}}y_{c,n}^{L-2} \tag{3}$$

where i is the number of output units for $L-1$; x_i^{L-1} is the input element of $L-1$; and c is the channel number, n is the number of the output units of $L-2$. To calculate $\partial E/\partial x_i^{L-1}$ in (3), we used (4),

$$\frac{\partial E}{\partial x_i^{L-1}}=\frac{\partial E}{\partial y_i^{L-1}}\frac{\partial y_i^{L-1}}{\partial x_i^{L-1}}=\frac{\partial E}{\partial y_i^{L-1}}\sigma'\left(x_i^{L-1}\right)=\frac{\partial E}{\partial y_i^{L-1}}y_i^{L-1}(1-y_i^{L-1}) \tag{4}$$

where y_i^{L-1} is the output for $L-1$. To calculate $\partial E/\partial y_i^{L-1}$ in (4), we introduced (5), where δ_p^L is error from (2) and $w_{i,p}^{L-1}$ is weight connecting $L-1$ and L.

$$\frac{\partial E}{\partial y_i^{L-1}}=\Sigma_p\delta_p^L w_{i,p}^{L-1} \tag{5}$$

The third part is connected weight ($w_{n,m}^{L-3}$) between two convolution layers (layer $L-2$ and $L-3$). In convolutional layers, the errors are calculated for the shared weights w.r.t. each channel and at the end all the corresponding shared weights and errors are summed to result in the effective error of a particular weight. Eq. (6) implements E w.r.t. this weight,

$$\frac{\partial E}{\partial w_{n,m}^{L-3}}=\Sigma_c^C\frac{\partial E}{\partial x_{c,n}^{L-2}}\frac{\partial x_{c,n}^{L-2}}{\partial w_{n,m}^{L-3}}=\Sigma_c^C\frac{\partial E}{\partial x_{c,n}^{L-2}}y_{c,m}^{L-3} \tag{6}$$

where n is the number of the output units of $L-2$ and m is the number of the output units of $L-3$. These are originally RBM weights shared along channels. In order to calculate $\partial E/\partial x_{c,n}^{L-2}$ in (6), we used (7),

$$\frac{\partial E}{\partial x_{c,n}^{L-2}}=\frac{\partial E}{\partial y_{c,n}^{L-2}}\frac{\partial y_{c,n}^{L-2}}{\partial x_{c,n}^{L-2}}=\frac{\partial E}{\partial y_{c,n}^{L-2}}\sigma'\left(x_{c,n}^{L-2}\right)=\frac{\partial E}{\partial y_{c,n}^{L-2}}y_i^{L-2}(1-y_i^{L-2}) \tag{7}$$

where y_i^{L-2} is the output for $L-2$. To calculate $\partial E/\partial y_{c,n}^{L-2}$ in (6), we used (8),

$$\frac{\partial E}{\partial y_{c,n}^{L-2}}=\Sigma_{i=1}^{k}\frac{\partial E}{\partial y_i^{L-1}}w_{c,n,i}^{L-2} \tag{8}$$

where $\partial E/\partial y_i^{L-1}$ is from (5) and $w_{c,n,i}^{L-2}$ is weight connecting $L-2$ and $L-1$. It is noticeable that the weight connected two convolution layers is basically initialised by a restricted Boltzmann machine

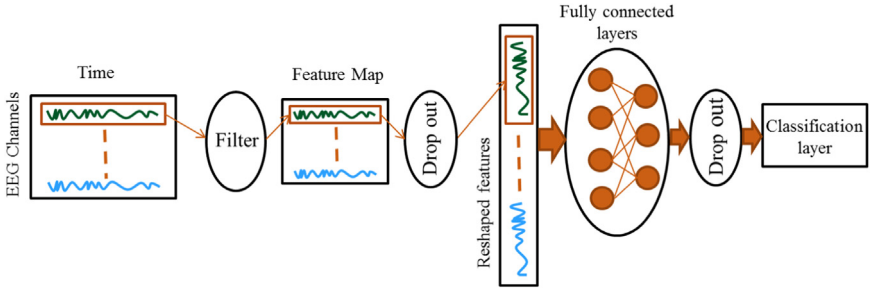

Fig. 1. Architecture of the proposed channel-wise convolutional neural network.

(RBM) [6] which operates on the channel of each epoch, in contrast to the conventional CNN which operates on n-by-n patches of input images.

4. Experimental results

4.1. Investigation of EEG data for poor/good driving performance features

We first performed global analysis of the 1 s epochs to examine if there exists potential discriminant features in both time and frequency domain. We first investigated event related potentials (ERPs) in the 1 s epoch raw EEG data for both poor/good driving performance states. They are obtained by averaging the multiple poor/good driving performance data epochs, respectively which can reduce the variability related with artefacts, noise or unrelated events while emphasising the prominent temporal patterns in each channel [12]. From ERP analysis we observe that except channel 6 and 17, which are ultrahigh frequency components on poor/good driving performance states, and channel 20, which shows no clear difference between the poor/good driving performance, ERPs in other channels have shown three obvious patterns. The first is a set of ERPs for poor driving performance (red curve) and good driving performance (blue curve) states that show increased amplitude with time for both poor/good driving performance states. The second includes a set of ERPs for poor/good driving performance states that show opposite trends. The third shows decreased amplitude with time for both poor/good driving performance state. Fig. 2 shows these patterns respectively.

The ability to capture all these channel-wise temporal ERP patterns is the key for an efficient classifier to discriminate poor driving performance and good driving performance states. Then, we analysed the average channel frequency spectrum of the 1 s epochs and associated topographical scalp maps (Fig. 3).

This analysis is aimed at interpreting data from a frequency aspect. As reported in [2], discriminate powers were observed in delta, theta, and alpha bands. Therefore, we particularly plotted the scalp maps around 5, 7 and 10.5 Hz, which roughly described most people's theta, delta and alpha band [13]. A palpable bump can be observed on both poor driving performance and good driving performance channel spectra around 10 Hz in many

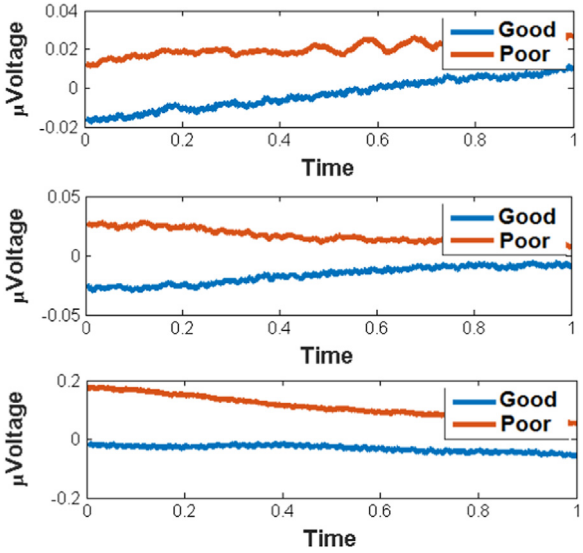

Fig. 2. ERP analysis. Each plot represents ERPs in a channel and red curve is for poor driving performance state, and blue curves for good driving performance state.

channels, and the scalp map revealed that the high power is notably concentrated over the occipital area, possibly indicating that subjects are fairly relaxed in the midst of seemingly intense experiments [14]. On the other hand, at 7 Hz theta band, the scalp map of the alert state clearly shows a higher power in the frontal region when compared with that in poor driving performance state. This potentially suggests that in the alert states, subjects are of higher cognitive levels (more focus) and therefore show higher power in the frontal region [14]. A similar pattern can be observed at both delta and alpha bands. Another important observation is that, except only a few channels (about 5 channels), most channels have been shown to be with highly congruent patterns across different frequency bands. This observation indicates that many channels are highly correlated, which also justifies the proposed channel-wise learning in CCNN.

Finally, time-frequency analysis was performed to examine the temporal change of the power spectrum. Particularly, we investigated transient event-related spectral perturbation (ERSP) [15] and inter-trial coherence (ITC) (event-related phase-locking) events [15]. To obtain them, we applied time frequency analysis on multiple epochs and presented an averages time-frequency spectrum in Fig. 4. The ERSP image (Fig. 4A and B) shows different patterns for poor driving performance and poor driving performance states. There is a significant synchronization in power at about 400–550 ms across all the frequency bands and an increased in power from 30 to 40 Hz from 600 to 700 ms in the good driving performance spectrogram, which are depressed in poor driving performance state. Examining the ITC images (Fig. 4A.II and B.II) further reveal that there is no phase locking events because all the ITC values are below the significance threshold ($p = 0.01$). In conclusion, these analyses demonstrated that there are discriminant features related to performance in both time and frequency domain and these features are likely to be related to brain signals as opposed to artefacts or other non-EEG signals.

4.2. Prediction performance of CCNN and CCNN-R

Prediction performances of driver's cognitive states by the proposed CCNN and CCNN-R were investigated in this section. CCNN and CCNN-R were tested using raw EEG and IC as input features and compared with DNN, CNN, and other popular classifiers. In addition a feature extraction method proposed in [17] is considered and the feature vectors obtained from this method are classified with SVM. In order to obtain these frequency features and related feature vector, fast Fourier transform (FFT) with a Hamming window, considering 250 sample points of the EEG data on each channel of a 1-s EEG epoch is performed. The obtained power spectrum density function is separated into 4 parts based on the 4 standard EEG frequency bands of delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), and beta (13–20 Hz). Afterwards, for each frequency band four frequency features are obtained (Dominant frequency, Average power of dominant peak, Centre of gravity frequency, Frequency variability) and finally the resulting feature vector for each epoch would have a size of 30*16=480. For each case, performances of both within-subject and cross-subject prediction were evaluated. Within-subject prediction refers to the scenario where the epoch to be predicted can come from the same subject that the training epochs are obtained, whereas cross-subject prediction simulates the case where the predicting and training epochs come from different subjects and subjects unseen for training are used in testing. Cross-subject prediction is more difficult because of the cross-subject differences in cognitive behaviours. To obtain the within-subject test and cross-subject performance, the double cross validation is performed [19]. The double cross validation includes two nested 10-fold cross validation. Specifically, in the ten-fold cross validation, 10% of the data is

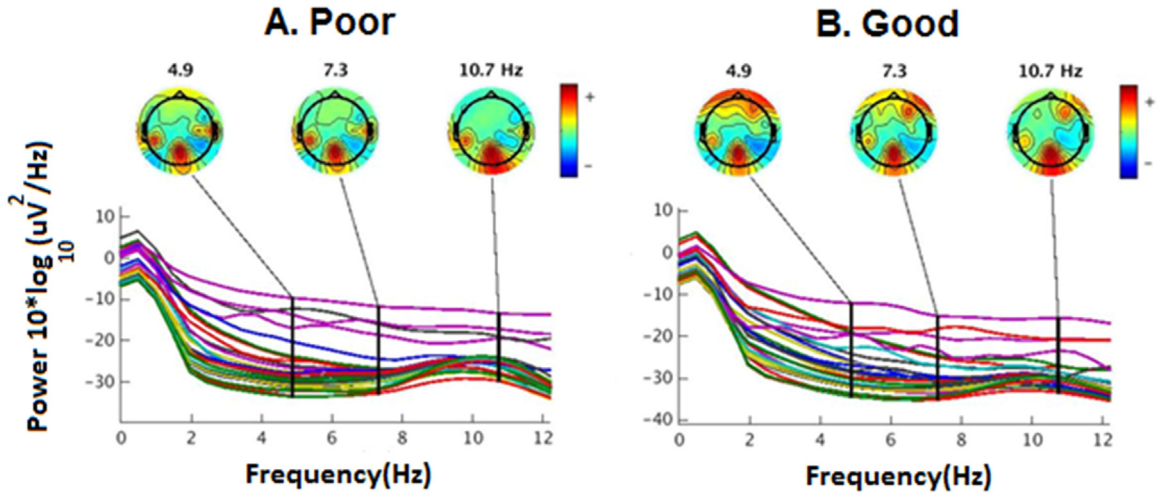

Fig. 3. Channel frequency spectrum analysis of channels for poor driving performance (A) and good driving performance (B) states. The horizontal axis of each figure describes frequency information, whereas the vertical axis denotes the power in decibel (dB). There are 30 curves in each figure, corresponding to the spectrum of each of the 30 channels.

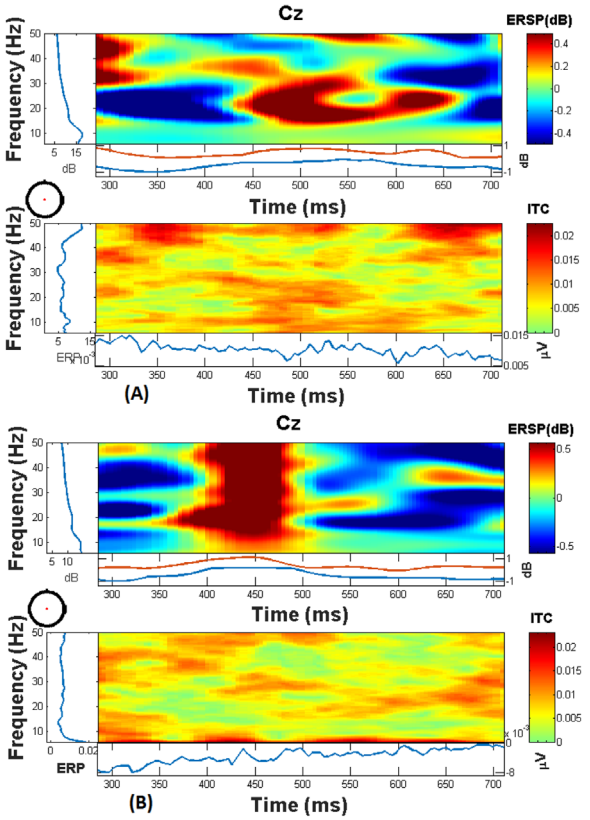

Fig. 4. A and B are time frequency analysis of a selected channel (channel 16). A and B show ERSP in power spectrum (from pre-stimulus baseline) at each time during the epoch and at each frequency <50 Hz. The upper left panel shows the baseline mean power spectrum, and the lower part of A and B are the ERSP envelope (low and high mean dB values, relative to baseline, at each time in the epoch).

used for validation and 90% is used for training and testing of the models. Training sets are used to train models and estimate the best model parameters and test sets are used to test the model performance and finally validation set assures the performance and validity of the selected model. Area under the receiver operating characteristic (ROC) curve (AUC), or Az score was evaluated for each test a measure of prediction performance. DNN, CNN and CCNN were implemented using Caffe, in which built-in Xavier filter has been applied for CNN and CCNN, but CCNN-R was programed based on Matlab® because of its unique architecture, which cannot be directly realized in Caffe. DL models were trained for each test, where the number of layers and hidden units were determined using the 1-fold holdout data. For easy composition, we use (x-y conv-z) to represent 2 fully connected and one convolutional layer each with x, y and, z hidden units, respectively.

4.2.1. Prediction based on raw EEG data

Prediction of driver's cognitive states from raw EEG data with minimum amount of the pre-processing is highly desirable for an online detection systems. In addition to DNN (500-100-2) and CNN ($5\times5\times10$ conv-500-100-2), the prediction performance of the popular classification methods including LDA and SVM was also performed. Because of the high dimensionality of the EEG epoch, dimension reduction using Principle Component Analysis (PCA) was also considered for LDA, SVM and DNN. Furthermore, the effect of xDAWN and CSP spatial filters on EEG data was also examined. The goal is to see if they preserved the ability to help extract discriminant features as for detecting visually evoked events. The spatial filters xDAWN and CSP are performed on the EEG epochs and then these filtered epochs are classified with deep neural network to measure the prediction performance. In addition, the extracted frequency features which already mentioned are also classified with SVM.

The test results for within-subject EEG epochs are shown in Fig. 5. Among 12 tested algorithms, the DL algorithms using minimally preprocessed data showed better performance and our proposed CCNN ($1\times250\times10$ conv-500-100-2) and CCNN-R ($1\times250\times80$ conv-100-100-2) achieved the best performance for within-subject testing. Specifically, CCNN reported the highest Az score at $86.08\pm0.46\%$. The test results for cross-subject EEG epochs are shown in Fig. 6. Our proposed CCNN-R (1×250-1×80 conv-100-100-2) achieved the best performance for cross-subject testing $76.72\pm3.7\%$. Tables 1 and 2 show the CCNN and CCNN-R parameters, respectively, where the values of the parameters for the lower (convolution) layer to the top (two fully connected and the classification) layer are shown from the left to the right of the tables.

In contrast, the performances of LDA and SVM with and without PCA are extremely low, which underscores the difficulty of the conventional classifiers in handling this prediction problem. The use of spatial filters xDAWN and CSP with DNN degraded as opposed to improving the performance, which suggests that they fail to extract the discriminative features of driver's cognitive states,

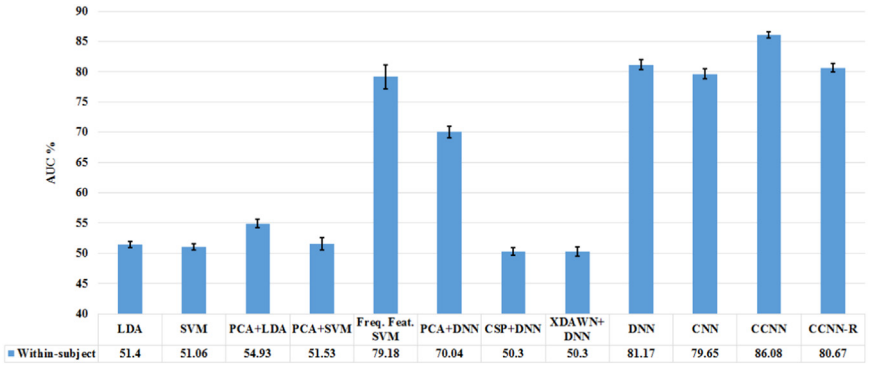

Fig. 5. Results using the raw EEG data for within- subject prediction.

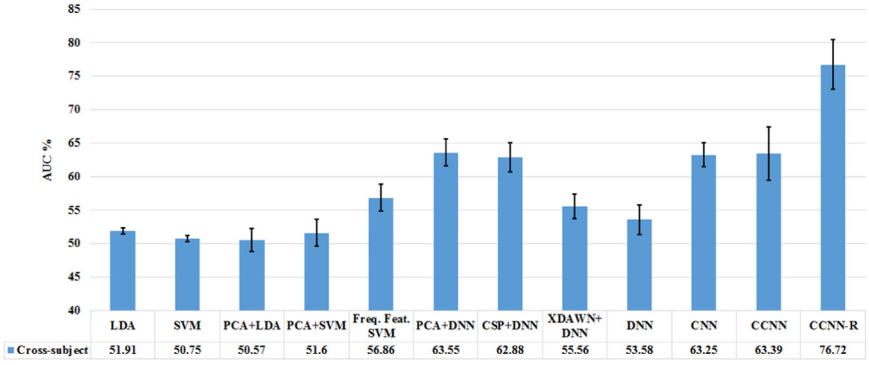

Fig. 6. Results using the raw EEG data for cross- subject prediction.

Table 1
CCNN parameters.

Kernel size in convolution layer	Feature-maps	Max pooling	Drop-Out	Fully-connected layer 1	Fully-connected layer 2	Drop-out	Classification layer
1×250	10	1×2	50%	500 nodes	100 nodes	50%	2 nodes

Table 2
CCNN-R parameters.

Kernel size in convolution layer	Feature-maps	Max pooling	Drop-out	Fully-connected layer 1	Fully-connected layer 2	Drop-out	Classification layer
1×250	1	1×2	90%	100 nodes	100 nodes	90%	2 nodes

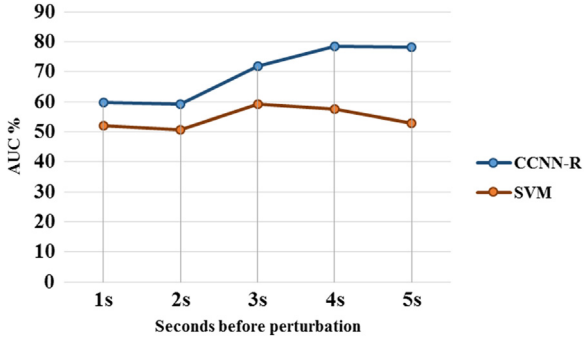

Fig. 7. Cross-subject prediction performance on 1 s epochs 1 s to 5 s before perturbation.

even though they have shown to be very effective in exacting EEG features for visually evoked events. Taken together, these results show that our proposed channel-wise DL algorithms are effective and robust in working with raw EEG data to predict driver's cognitive states. It is notable that CCNN-R has better performance than CCNN for cross-subject. This suggests that the use of RBM in

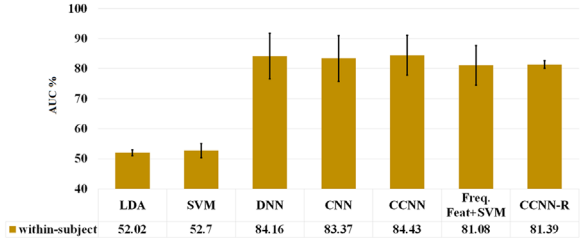

Fig. 8. Results for ICA features of within-subject.

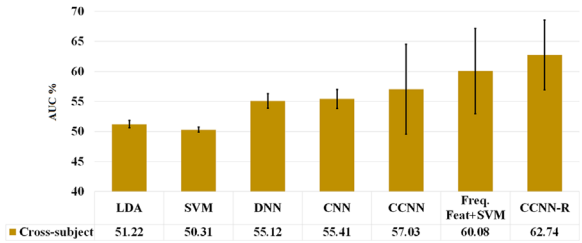

Fig. 9. Results for ICA features of cross-subject.

CCNN-R helps learning the better common features shared in cross subjects.

Next, we also investigated the change of prediction performance over time. Specifically, we carried out the cross-subject predictions using CCNN-R and SVM with frequency features on testing epochs that are 1 s, 2 s, 3 s, 4 s, and 5 s before perturbation separately. The results are shown in Fig. 7. First of all, as expected, CCNN-R has much better performance than SVM for all five cases. However, what is interesting is that for both classifiers, prediction performance of the epochs closer to injection of the perturbation is lower than the epochs further away. This seemingly counter-intuitive behaviour might be a result of the repeated injection of perturbation every 8–12 s for 1.5 h; the subjects could get well used to the cycle of perturbation and the brain signals corresponding to this cycle interferes more with the signals of performance as it gets closer to the perturbation, leading to decreased prediction of performance.

4.2.2. Prediction based on independent components

Independent component analysis (ICA) have widely been used in EEG analysis and shown in many studies including those for driver's fatigue to provide useful discriminate information. Therefore, we investigate the impact of ICs on the performance of CCNN. To this end, each EEG epoch was projected with ICA to a 30×250 epoch containing 30 independent components. Note that the projection matrix was obtained using only the training data and was applied to the validation data. ICA decomposition was first applied to the original epoched data, which resulted in 30 ICs spanning 250 time samples for each epoch. To obtain the frequency domain features, the Morlet wavelet transform was then applied to the 250 samples of each IC separately.

Seven methods including LDA, SVM, DNN (500-100-2), CNN ($5 \times 5 \times 50$-2×2-$3 \times 3 \times 100$-2×2-500-100-2), CCNN ($1 \times 250 \times 10$ conv-500-100-2) and CCNN-R ($1 \times 250 \times 80$ conv-100-100-2) in addition to extracted frequency features were trained and tested on ICA features obtained from within-subject and Cross-subject cases. Fig. 8 depicts the Az scores for these algorithms. CCNN reached highest Az score ($84.43 \pm 6.65\%$) for within-subject testing.

Again, seven methods including LDA, SVM, DNN (500-100-2), CNN ($5 \times 5 \times 50$-2×2-$3 \times 3 \times 100$-2×2-500-100-2), CCNN ($1 \times 250 \times 10$ conv-500-100-2) and CCNN-R ($1 \times 250 \times 80$ conv-100-100-2) in addition to extracted frequency features were trained and tested on ICA features obtained from cross-subject case. Fig. 9 depicts the Az scores for these algorithms. CCNN-R reached highest Az score ($62.74 \pm 5.8\%$) for cross-subject testing.

5. Conclusion

In this work, we proposed a novel channel-wise convolutional neural network (CCNN) for prediction of driver's cognitive states from EEG signals. CCNN's architecture considers the unique characteristics of EEG data. We also discussed CCNN-R, and derived the detailed algorithm. We investigated the new CCNN and CCNN-R on raw EEG, time-frequency power spectrum, and ICA-transformed data and tested them for both within-session and cross-session predictions. The results showed CCNN and CCNN-R achieved robust and improved performance over conventional DNN and CNN as well as other non-DL algorithms. Because it is designed for EEG data, CCNN has great potential to be a useful tool for other EEG analysis.

References

[1] Chin-Teng Lin, Ruei-Cheng Wu, Sheng-Fu Liang, Wen-Hung Chao, Yu-Jie Chen, Tzyy-Ping Jung, EEG-based drowsiness estimation for safety driving using independent component analysis, IEEE Trans. Circuits Syst. I: Regular Pap. 52 (12) (2005) 2726–2738.

[2] Chun-Hsiang Chuang, Li-Wei Ko, Yuan-Pin Lin, Tzyy-Ping Jung, Chin-Teng Lin, Independent component ensemble of EEG for brain–computer interface, IEEE Trans. Neural Syst. Rehabil. Eng. 22 (2) (2014) 230–238.

[3] S.F. Liang, C.T. Lin, R.C. Wu, Y.C. Chen, T.Y. Huang, T-P. Jung. Monitoring driver's alertness based on the driving performance estimation and the EEG power spectrum analysis, in: Proceedings of the 27th International Conference of the IEEE-EMBS 2005 Annual Engineering in Medicine and Biology Society, 2006, pp. 5738–5741.

[4] Agustina Garcés Correa, Lorena Orosco, Eric Laciar, Automatic detection of drowsiness in EEG records based on multimodal analysis, Med. Eng. Phys. 36 (2) (2014) 244–249.

[5] S.W. Chuang, L.W. Ko, Y.P. Lin, R.S. Huang, T.P. Jung, C.T. Lin, Co-modulatory spectral changes in independent brain processes are correlated with task performance, Neuroimage 62 (3) (2012) 1469–1477.

[6] Y. Bengio, Learning deep architectures for AI, Found. Trends® Mach. Learn. 2 (2009) 1–127.

[7] l Deng, D. Yu, Deep Learning for Signal and Information Processing, Foundations and Trends in Signal Processing, NOW Publishers, 2013.

[8] G. Hinton, L. Deng, D. Yu, G.E. Dahl, A.-r Mohamed, N. Jaitly, et al., Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups, Signal Process. Mag. IEEE 29 (2012) 82–97.

[9] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in: Proceedings of the Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.

[10] Z. Mao, V. Lawhern, L. M. Merino, K. Ball, L. Deng, B. J. Lance, et al., Classification of non-time-locked rapid serial visual presentation events for brain-computer interaction using deep learning, in: Proceedings of the IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP), Xi'an, China, 2014, pp. 520–524.

[11] H. Cecotti, M.P. Eckstein, B. Giesbrecht. Single-trial classification of event-related potentials in rapid serial visual presentation tasks using supervised spatial filtering, 2014.

[12] S. Makeig J. Onton, ERP features and EEG dynamics: an ICA perspective, Oxford Handbook of Event-Related Potential Components, New York, NY: Oxford, 2009.

[13] W. Klimesch, EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis, Brain Res. Rev. 29 (2) (1999) 169–195.

[14] O. Jensen, A. Bahramisharif, R. Oostenveld, S. Klanke, A. Hadjipapas, Y. O. Okazaki, M.A. van Gerven, Using brain–computer interfaces and brain-state dependent stimulation as tools in cognitive neuroscience, Front. Psychol. 2 (2011) 100.

[15] A. Delorme, S. Makeig, EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis, J. Neurosci. Methods 134 (1) (2004) 9–21.

[17] Mervyn V.M. Yeo, et al., Can SVM be used for automatic EEG detection of drowsiness during car driving? Saf. Sci. 47.1 (2009) 115–124.

[18] Hajinoroozi, Mehdi, et al., Feature extraction with deep belief networks for driver's cognitive states prediction from EEG data, in: Proceedings of the IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP), 2015.

[19] François Romain, Florent Langrognet. Double Cross Validation for Model Based Classification, Book of Abstracts, 2006.

[20] Yann LeCun, et al., Gradient-based learning applied to document recognition, Proc. IEEE 86.11 (1998) 2278–2324.

[21] Ruslan Salakhutdinov, Geoffrey E. Hinton. Deep boltzmann machines, in: Proceedings of the International Conference on Artificial Intelligence and Statistics, 2009.

[22] Yoshua Bengio, et al., Greedy layer-wise training of deep networks, Adv. Neural Inf. Process. Syst. 19 (2007) 153.

[23] Nitish Srivastava, et al., Dropout: a simple way to prevent neural networks from overfitting, J. Mach. Learn. Res. 15.1 (2014) 1929–1958.