FEATURE EXTRACTION WITH DEEP BELIEF NETWORKS FOR DRIVER'S COGNITIVE STATES PREDICTION FROM EEG DATA

Mehdi Hajinoroozi[1]*, Tzyy-Ping Jung*[2]*, Chin-Teng Lin*[3]*, and Yufei Huang*[1]

1. Department of Electrical and Computer Engineering, University of Texas at San Antonio, TX, USA
2. Institute for Neural Computation, University of California, San Diego, CA, USA
3. Brain Research Center, National Chiao Tung University, Hsinchu, Taiwan

ABSTRACT

This study considers the prediction of driver's cognitive states from electroencephalographic (EEG) data. Extracting EEG features correlated with driver's cognitive states is key for achieving accurate prediction. However, high dimensionality and temporal-and-spatial correlations of EEG data make extraction of effective features difficult. This study explores the approaches based on deep belief networks (DBN) for feature extraction and dimension reduction. Experimental results of this study showed that DBN applied to channel epochs (DBN-C) produces the most discriminant features and the best classification performance is achieved when DBN-C is applied to the time-frequency and independent-component-analysis transformed EEG data. The results suggested that DBN-C is a promising new method for extracting complex, discriminant features for EEG-based brain computer interfaces.

Index Terms— Deep belief network, feature extraction, Classification

1. INTRODUCTION

This study considers a brain computer interface (BCI) system for driver's cognitive state prediction, where electroencephalography (EEG) measurement of driver's brain signals is collected continuously. The goal of this study is to construct a classifier that can detect driver alertness/drowsy from the EEG measurements [1], [2]. One of the key considerations is extraction of alertness/drowsy-related EEG features from the high-dimensional raw EEG data, which span in channels and over time. A large number of methods have been proposed in the literature for feature extraction for EEG-based cognitive event detection including driver's drowsiness (alertness) or attention detection. Among them, the most widely applied features include time-domain raw EEG, time-domain features of components obtained from independent component analysis (ICA), time-frequency power spectrum features by a wavelet transform, or a combination of the three [1-3]. To reduce the feature dimension, principle component analysis (PCA) or ICA is often performed but other methods such as moving average have also been employed [4]-[7]. Among all these methods, ICA and time-frequency power spectrum features have shown to provide better representation of underlying cognitive events. However, the selection of the discriminant independent components and frequency bands often needs prior knowledge about event-related brain activities, which are not always available.

Previously, we have shown that deep learning models including deep neural networks and deep stacking networks can better characterize important discriminant EEG features that correlate with underlying cognitive activities [8]. However, applying deep learning models on EEG features directly does not scale well for large amount of input EEG epochs and the training would also be likely to fail as training samples are often limited. This work explores the feature extraction and dimension reduction with deep belief networks (DBN). Deep belief network is an unsupervised deep learning model, which maps the input EEG data into multiple layers of hidden units. These hidden units represent the underlying EEG patterns with different degree of abstraction. To assess DBN's ability to represent temporal and spatial correlations in EEG data and examine appropriate dimension reduction practices with DBN, this study proposed and evaluated several feature extraction schemes. This study also tests to what extent the extracted features can improve the prediction performance.

2. DESCRIPTION OF EEG DATA

The EEG data were collected from 17 subjects each performed a lane-keeping driving task in a virtual reality interactive driving platform with a 3-D highway scene [1]. Perturbations to the car were introduced into driving path every 8 to 12 seconds and driver's reaction time and the amount of the lane deviation were measured to assess the degree of driver's drowsiness. Each experiment lasted one and half hours during which subjects' EEG were measured from 30 electrodes.

The reaction time (RT) is defined as the time between the onset of the lane perturbation and the moment when the subject starts steering the car. RT was used to define the drowsy or alert state of the driver. Particularly, when the reaction time is below or equal to 0.7s, the driver is considered as alert, whereas when the reaction time is higher or equal to 2.1s, the driver is considered as drowsy [2].

This work used EEG epochs 1s before the onset of the perturbation as data for predicting the "drowsy" or "alert" state of the drivers. There is a total of 2,796 (764 drowsy and 2,032 alert) epochs from the 17 subjects. Because the sampling rate is 250Hz, the dimension of 1s EEG epoch is $250 \times 30 = 7{,}500$.

3. PROPOSED FEATURE EXTRACTION METHODS WITH DBN

The workflow for driver's cognitive states classification includes feature extraction by DBN and training a classier using the DBN features. DBN are constructed by stacking layers of restricted Boltzmann machines and trained with a greedy layer wise procedure [9]. This work investigates how DBN as an unsupervised feature extractor can provide features which can be used for improving the performance of driver's cognitive state classification. Apparently, due to the high dimension of the 1s EEG epoch (7,500), training DBN directly on epoched data is not feasible. However, we do not wish to apply any dimension reduction algorithms such as PCA before DBN training because such an algorithm could distort the information in the data and thus restrict the ability of DBN to fully make use of the cognitive-state-related information in the data. To address this problem, we propose three different approaches, which will be discussed in detail next. We also discuss how the proposed the approaches can be applied to the ICA and time-frequency transformed data.

3.1. Deep belief network on time samples (DBN-T)

The first approach is to train a DBN using individual time samples of EEG data. In this case, a time sample includes measurements from 30 channels and we pool together samples from all 2796 epochs, resulting in a total of $250 \times 2{,}796$ time samples to train a DBN. The advantage of this approach is obvious: it not only significantly reduces the dimension of the input data (from 7,500 to 30) to the DBN but also increases the number of training samples. However, the price to pay for this approach is its inability to model temporal correlations of EEG samples because this approach essentially assumes the time samples to be independently identically distributed (i.i.d.). Omitting temporal correlation might not be able to produce robust features as compared to other approaches listed below, especially in the presence of higher noise level and complex cognitive behaviors.

After a DBN is trained, we replace each time sample with the corresponding hidden units of the last DBN layer, which is normally smaller than 30. Then, we train a classifier using the 1s epochs of hidden units. Specially, if there are m hidden units in the last layer, the feature dimension for classification is $250m$, which achieved a 30/m times reduction of dimension over the original 1s epoch data.

3.2. Deep belief network on channel epochs (DBN-C)

Instead of using time samples, DBN-C treats individual channel epochs as i.i.d. samples and the DBN is trained with individual channel epochs from all the subjects. The dimension of an input channel epoch to DBN is 250 and there are $30 \times 2{,}796$ training samples. In contrast to DBN-T, DBN-C preserves the temporal dynamic patterns but ignores the channel correlations in the EEG.

After DBN training, each channel epoch is replaced by the corresponding DBN hidden units from the last layer. If the size of hidden units is m, then the dimension of a hidden-unit epoch is $30 \times m$ and these epochs are used to train a classifier. The achieved dimension reduction for classification is 250/m, which can be significant when m is small.

3.3. Deep belief network on windowed samples (DBN-W)

In the last approach, an epoch is divided into overlapping time windows and a DBN is trained with these time-windowed data. Suppose that there are p windows in an epoch, each with a length of w. Then, the dimension of a windowed sample is $30 \times m$ and there are totally $p \times 2{,}796$ DBN training samples. As in DBN-T and DBN-C, m DBN hidden units from the last layer replace the corresponding windowed data and the epochs of hidden units, each of dimension $m \times w$ data are used to train a classifier.

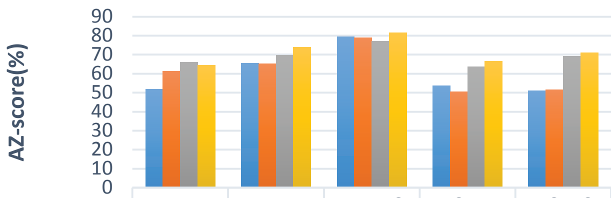
It is easy to see that DBN-T is a special case of DBN-W with w=1. Like DBN-T, DBN-W is intended to retain channel correlations in EEG data. However, for more general cases where w>1, DBN-W also retains a portion of temporal dynamics of an epoch. The larger the w, the more complete temporal dynamics it retains. The fold of dimension reduction for classification is 7,500/($m \times w$), which can be very high for small m and w.

4. EXPERIMENTAL RESULTS

This study conducted experiments to examine the performance of the classifiers trained using DBN-T, DBN-C, or DBN-W features. For each DBN training, a heuristic search of the number of layers and hidden units in each layer was conducted and the number of layers and hidden units that produced the best classification performance were reported in the following sections. This study compared four different classifiers including Linear Discriminant Analysis (LDA), Support Vector Machine, boosting and bagging classifiers [10], [11]. All classifiers are trained through a 10-fold cross validation. Area Under the receiver operating characteristic, or AZ score is used as a measure of classification performance.

4.1. Classification using time-domain DBN features

We first examined the proposed DBN features extracted directly from the raw EEG features. DBN-T, DBN-C, and DBN-W were trained first. For DBN-W, the window size (w) was 50 and the adjacent windows overlapped by 30 points. Therefore, each epoch has p =11 overlapping windows. The classification performances were examined for different number of layers and hidden units for the three approaches. The best performances were achieved at using 2 layers for all three approaches and the numbers of hidden units for the first and second layers are (10, 10) for DBN-T, (1100, 100) for DBN-W, and (230, 5) for DBN-C. As a result, the dimension of the feature vector for classification, or an epoch of the hidden units, is $250 \times 10=2500$, $5 \times 30 = 150$, and $100 \times 11=1100$ for DBN-T, DBN-C, and DBN-W, respectively. Compared with the original epoch dimension of 7500, DBN-T, DBN-C, and DBN-W achieved 3, 50, and ~7 times reduction, respectively. As a baseline performance, PCA was applied for dimension reduction and two types PCAs were obtained, i.e., PCA-T (PCA reduction of time samples) and PCA-C (PCA reduction of channels).



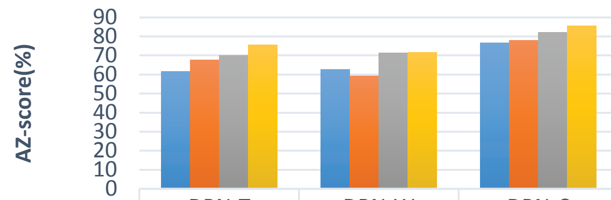
	DBN-T	DBN-W	DBN-C	PCA-T	PCA-C
LDA	51.62	65.48	79.44	53.49	50.94
SVM	61.28	65.22	78.91	50.36	51.36
Boosting	65.83	69.58	76.88	63.6	69.13
Bagging	64.48	73.83	81.36	66.58	70.98

Figure 1. AZ scores of the classifiers based on proposed DBN features and PCA features.

In each case, only the two components with the largest eigenvalues were retained as features. The four classifiers were also trained using PCA-T and PCA-C as features. The AZ scores of the four classifier using each of the three DBN features as well as PCA features are shown in Fig. 1. Among the four classifiers, it is not surprising that Bagging reported the best AZ scores for all cases. It is surly encouraging to see that, compared with the PCA features, three DBN features clearly improved classification performances. When comparing the performance among the three DBN approaches, we found that while DBN-T and DBN-W achieved comparable performances for all classifiers, DBN-C clearly outperformed both of them and the improvements were significant. In the case of Bagging, DBN-C improved about 17 and 8 percentage points in AZ scores over DBN-T and DBN-W, respectively. Recall that DBN-T completely ignores temporal dynamics, DBN-W models partial temporal dynamics, and DBN-C retains complete temporal dynamics, the better performance of DBN-C implies that temporal dynamics in EEG data are more important than channel correlations in differentiating the drowsy and alert states. Taken together, these results suggest that DBN features can improve the classification performance over the traditional PCA features and DBN-C resulted in the best performance among the three proposed DBN features.

4.2. Classification using IC-domain DBN features

We next examined if the classification performance could be further improved by using IC-domain DBN features. Briefly, ICA decomposition was first applied to the original epoched data and an epoch of ICA transformed data has 30 ICs spanning 250 time samples.



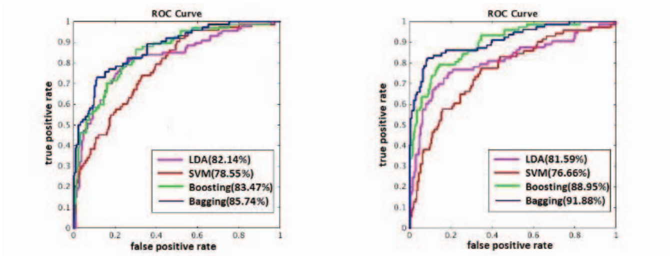
	DBN-T	DBN-W	DBN-C
LDA	61.53	62.54	76.61
SVM	67.61	59.32	77.86
Boosting	70.01	71.24	82.07
Bagging	75.65	71.55	85.57

Figure 2. AZ-scores of the three different features extraction methods, DBN-C, DBN-W and DBN-T after applying the ICA.

The three proposed DBN methods were subsequently applied to the ICA-transformed data, where DBN-C is essentially DBN-IC. For DBN-W, the same window setting of size 50 with 30 overlapping points was used. The trained DBNs all included two layers and the number of the hidden units for the first and second layers are (10, 5) for DBN-T, (60, 60) for DBN-W, and (10, 10) for DBN-C. The AZ scores of the four classifiers trained using each of the features are shown in Fig. 2. Once again, the best performance was achieved with DBN-C. Comparing with the performances in Fig. 1, we see that IC domain DBN features can lead to further improvement over time-domain features (~ 4% for Bagging).

4.3. Classification using frequency-domain DBN features

We finally investigated the classification performance of using frequency-domain DBN features. Because DBN-C has consistently achieved the best performance for time- and IC-domain feature, we choose to focus this investigation on the DBN-C approach. To obtain the frequency domain features, the Morlet wavelet transform was first applied to the time domain epochs and 11 frequency bands were evenly sampled in a logarithm-scale from 1-12Hz for each channel. This frequency range was chosen following the results in [2]. As a result, the dimension of the transformed epoch was $11 \times 30 \times 250$. For DBN-C, the input visible units were time-frequency powers of the all channels in an epoch the size of each is $11 \times 250 = 2750$ units. The training resulted in a DBN with two layers including 50 and 5 hidden units in the first and second layer respectively. The ROC curves and their AZ scores of the four classification algorithms are shown in Fig 3-A. The best AZ score (85.74%) is again obtained by the Bagging algorithm but it is comparable to that using IC-domain features (85.57%; Fig. 2). To see if additional improvement could be achieved, we applied the Morlet wavelet transform to the ICA transformed epochs and extracted the DBN-C features. The classification performances are shown in Fig 3-B. This time, the Bagging algorithm achieve an AZ score of 91.88%, which is about ~6% point improvement over that using either IC-domain or frequency-domain features alone. Taken together, these results suggest that DBN-C features extracted from the spectra of ICs provided the best representation of driver's cognitive states.



A B

Figure 3. ROC curves and AZ-scores of classifiers using frequency-domain DBN-C features. A. DBN-C features extracted from wavelet-transformed EEG data. B. DBN-C features extracted from wavelet-transformed ICs of EEG data.

5. CONCLUSION

This study investigated for the first time how DBN can be used to extract effective features from high-dimensional EEG signals that can lead to high-performance classification of driver's cognitive states. Three methods including DBN-T, DBN-W, and DBN-C were proposed and their extensions to IC and frequency domains were discussed. The experimental results showed that DBN features can improve the performance of traditional PCA features and DBN-C consistently achieved the best performance among the proposed methods. Furthermore, DBN-C features extracted from the frequency domain information of ICs produced the best performance over those from time-, frequency-, or IC-domain alone. In conclusion, DBN-C is a promising new approach for extracting complex and discriminant features for EEG-based brain computer interfaces.

6. REFERENCES

[1] Lin, Chin-Teng, Ruei-Cheng Wu, Sheng-Fu Liang, Wen-Hung Chao, Yu-Jie Chen, and Tzyy-Ping Jung. "EEG-based drowsiness estimation for safety driving using independent component analysis." Circuits and Systems I: Regular Papers, IEEE Transactions on 52, no. 12 (2005): 2726-2738.

[2] Chuang, Chun-Hsiang, Li-Wei Ko, Yuan-Pin Lin, Tzyy-Ping Jung, and Chin-Teng Lin. "Independent Component Ensemble of EEG for Brain–Computer Interface." Neural Systems and Rehabilitation Engineering, IEEE Transactions on 22, no. 2 (2014): 230-238.

[3] Correa, Agustina Garcés, Lorena Orosco, and Eric Laciar. "Automatic detection of drowsiness in EEG records based on multimodal analysis." Medical engineering & physics 36, no. 2 (2014): 244-249.

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

[5] Lin, Chin-Teng, Yu-Chieh Chen, Teng-Yi Huang, Tien-Ting Chiu, Li-Wei Ko, Sheng-Fu Liang, Hung-Yi Hsieh, Shang-Hwa Hsu, and Jeng-Ren Duann. "Development of wireless brain computer interface with embedded multitask scheduling and its application on real-time driver's drowsiness detection and warning." Biomedical Engineering, IEEE Transactions on 55, no. 5 (2008): 1582-1591.

[6] Heger, Dominic, Felix Putze, and Tanja Schultz. "Online Workload Recognition from EEG data during Cognitive Tests and Human-Computer Interaction." Inproceedings of 33rd Annual German Conference on Artificial Intelligence 2010. 2010.

[7] Lin, Chin-Teng, Ruei-Cheng Wu, Tzyy-Ping Jung, Sheng-Fu Liang, and Teng-Yi Huang. "Estimating driving performance based on EEG spectrum analysis."EURASIP Journal on Applied Signal Processing 2005 (2005): 3165-3174.

[8] Mao, Zijing, Vernon Lawhern, Lenis Mauricio Merino, Kenneth Ball, Li Deng, Brent J. Lance, Kay Robbins, and Yufei Huang. "Classification of non-time-locked rapid serial visual presentation events for brain-computer interaction using deep learning." In Signal and Information Processing (ChinaSIP), 2014 IEEE China Summit & International Conference on, pp. 520-524. IEEE, 2014.

[9] Hinton, Geoffrey, Simon Osindero, and Yee-Whye Teh. "A fast learning algorithm for deep belief nets." Neural computation 18.7 (2006): 1527-1554.

[10] Dietterich, Thomas G. "Ensemble methods in machine learning." In Multiple classifier systems, pp. 1-15. Springer Berlin Heidelberg, 2000.

[11] Bari, Mehrab Ghanat, Xuepo Ma, and Jianqiu Zhang. "PeakLink: a new peptide peak linking method in LC-MS/MS using wavelet and SVM." Bioinformatics (2014): btu299.