Deep Belief Networks for EEG-Based Concealed Information Test

Qi Liu[1(✉)], Xiao-guang Zhao[1], Zeng-guang Hou[1] and Hong-guang Liu[2]

[1] The State Key Laboratory of Management and Control for Complex Systems
Institute of Automation, University of Chinese Academy of Sciences
Beijing, 100190, China
popofay17@gmail.com
[2] Institute of crime, Chinese People's Public Security University
Beijing, 100038, China

Abstract. This paper introduces a deep learning approach to the feature extraction of P300 cognitive component existing in electroencephalogram signals collected in an autobiographical paradigm test. A deep belief network is used to extract deep features instead of raw feature vectors to train the classifier. It is shown that the classification accuracy is satisfactory by learning deep from the training data. The effectiveness of the algorithm has been validated through experiments. A high accuracy of recognizing concealed information with a single electroencephalogram channel is obtained. Moreover, performances of support vector machine with different feature extraction methods are compared.

Keywords: Electroencephalogram, Concealed Information Test, Deep Feature Extraction, Deep Belief Networks

1 Introduction

In recent years, EEG-based concealed information test has drawn considerable attention in the field of criminal investigation. Many effective methods have been used for EEG signal analysis in Concealed Information Test (CIT) [1]. Compared to traditional methods based on physiological responses which are easily affected by emotions and stress, cognitive behavior based polygraph is considered more reliable and scientific that can reduce the risk in false positive errors [2]. In addition, EEG is more convenient, more harmless and more economical than other brain activity monitoring methods such as PET, MEG and fMRI [3].

Due to the complexity and particularity of actual criminal investigation tasks and poor signal-to-noise ratio (SNR) of raw EEG signals, improving recognition performance remains a live problem. In which, methods based on machine learning algorithms have achieved the most effective results. Numerous feature extraction approaches have been adopted in machine learning algorithms such as time or frequency methods [4], model parameter methods [5], and wavelet decomposition -based methods [6], etc. [7]. However, the distinguishability of a certain feature is uncertain in different tasks, which may lead to a failure of recognition. Therefore, feature extrac-

adfa, p. 1, 2011.

tion methods with good ability of feature self-learning are necessary to be studied in this field. Recently, deep learning strategy has made great progress and the related algorithms have been applied to many fields including EEG signal processing [8]. It can be viewed as a computational intelligent method since its similar mechanism to human brain. To improve the generalization performance of EEG feature, deep belief networks (DBN) is adopted to learn features automatically.

In this paper, we use the CIT technique and focus primarily on the feature extraction process of different brain waves evoked by relevant stimulus and control stimulus. DBN was applied to self-learn features of EEG signals. Then support vector machine (SVM) was implemented as the classifier. The classification performance is satisfactory and the runtime is acceptable.

2 Data and Methods

2.1 Data Description

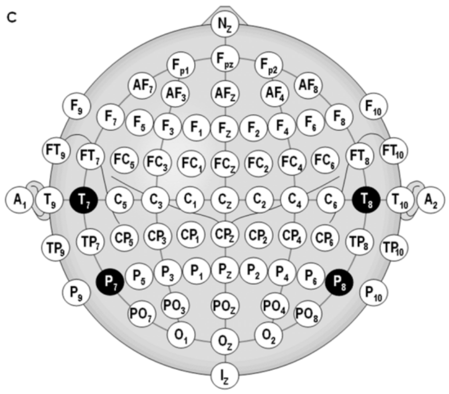

Fig. 1. The 10–20 system of electrode placement

The data used in this paper were recorded during an autobiographical paradigm test [9]. There were eleven subjects in total participated. They are all male volunteers between the ages of 22 and 35. They are all right handed and had normal or corrected-to-normal vision. They are not familiar with the scientific basis of the test and only have knowledge about how to perform the test. Each subject was asked to provide five numbers (all 4 digits long), one of them being their year of birth. The subjects did not reveal to the experimenter which one of the numbers is their birth date until the end of experiment. Each subject participated in 2 experimental runs, except for subject 11, who participated in 3 runs. For subject 1, 3, and 7, one run was discarded because of incorrect target stimulus counting (see below). Therefore, a total of 20 experimental runs were used in this study. In each run, each number was displayed to the subject randomly with thirty repetitions, resulting in a total of stimuli. Each number was displayed for one second and between the numbers, the screen was blank for two seconds. The subjects did not respond to the items, but were instructed to count the number of times the target stimulus (year of birth) was presented (they were unaware that all stimuli were repeated 30 times in each run). EEG signals were recorded at frontal (Fz), central (Cz), and parietal (Pz) electrode positions of the 10–20 international electrode placement system (Fig. 1). All electrodes were referenced to linked mastoids. Vertical EOG was also recorded for blink artifact detection. EEG signals were digitally sampled at 256 Hz [10].

2.2 Methods

For the complexity and weak anti-interference capability of EEG, it is very difficult to recognize useful information from raw signals. Fig. 2 shows the raw waveforms. It is observed that the potential offset value of each sample belonging to the same category is quite different and there is no obvious distinction between samples belonging to separate categories.

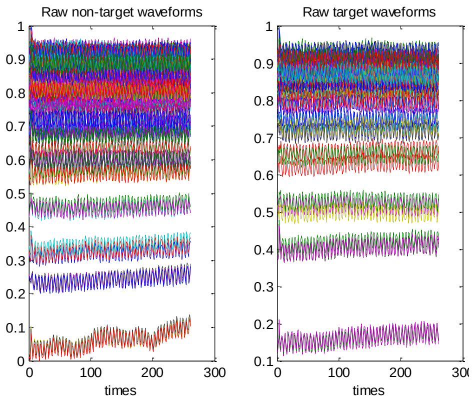

Fig. 2. Raw EEG waveforms

As shown in Fig. 3, EEG signal processing mainly includes data collection, preprocessing, feature exaction and pattern classification.

1) Pre-processing

This process comprises selection of electrodes, signal segmentation, superposition and filtering. For low SNR of EEG signals, the repetitive stimulations are superimposed to reduce the interference signals and enhance the desired information. Since the frequency of P300 is mainly distributed in low frequency area, a 6-order band pass Chebyshev Type I filter with cut-off frequencies 0.5 and 35 Hz is designed to filter each epoch. Moreover, the data matrix is mapped into a bound between 0 and 1 according to equation (1).

$$\mathbf{x}_{norm} = \mathbf{x} - \mathbf{x}_{min} / \mathbf{x}_{max} - \mathbf{x}_{min} \quad (1)$$

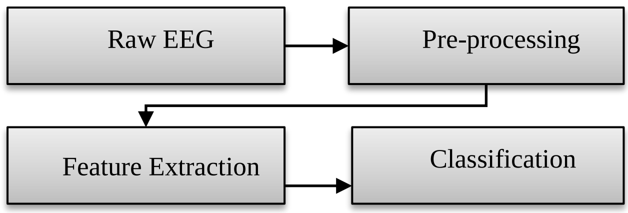

Fig. 3. The flowchart of signal processing

2) Deep Feature Extraction

To begin with, k-means method is adopted to represent features preliminary as described in [11]. Using subject 1 as an example, some differences between the two categories can be seen in Fig. 4 after the initial feature extraction. However, the difference is still too small to distinguish samples. Further feature extraction is implemented as following.

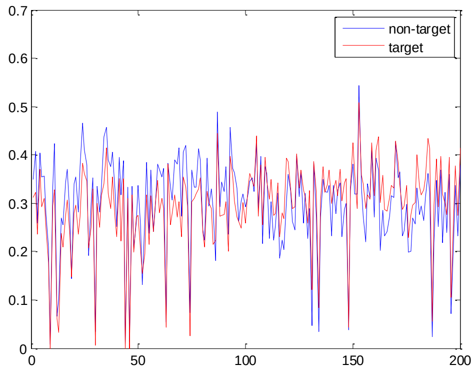

Fig. 4. Comparison of mean values of the two categories

DBN could be viewed as a stack of restricted Boltz-man machines (RBMs), which are motivated from the idea of equilibrium from the statistical physics literature [12]:

$$E(\mathbf{v},\mathbf{h};\boldsymbol{\theta}) = -\sum_j a_j v_j - \sum_i b_i h_i - \sum_{i,j} v_j h_i w_{ij} \tag{2}$$

where w_{ij} is the symmetric interaction term between visible unit v_j and hidden unit h_i, a_i and b_j are the bias term. $\boldsymbol{\theta} = \{\mathbf{w},\mathbf{a},\mathbf{b}\}$ is the model parameter need to be learned.

Equation (2) could be optimized in a tricky way by contrastive divergence that is commonly used to approximate the expectation by a sample generated after a limited number of Gibbs sampling iterations [13].

The joint probability distribution over $\mathbf{v}$ and $\mathbf{h}$ is:

$$P(\mathbf{v},\mathbf{h}) = \frac{1}{z} e^{-E(\mathbf{v},\mathbf{h})} \tag{3}$$

where Z is a normalizing factor. Then

$$P(\mathbf{v}) = \sum_{\mathbf{h}} P(\mathbf{v},\mathbf{h}) = \frac{e^{-F(\mathbf{v})}}{z} \tag{4}$$

where

$$F(\mathbf{v}) = -log \sum_{\mathbf{h}} e^{-E(\mathbf{v},\mathbf{h})} \tag{5}$$

Model (2) can be simplified by using binary input variables. The conditional probabilities can be formulated as:

$$\begin{aligned} P(h_i = 1|\mathbf{v}) &= sigm(b_i + w_i\mathbf{v}) \\ P(v_j = 1|\mathbf{h}) &= sigm(a_j + w_j^{'}\mathbf{h}) \end{aligned} \tag{6}$$

Then

$$F(\mathbf{v}) = -\mathbf{a}'\mathbf{v} - \sum_i log\left(1 + e^{(c_i + w_i\mathbf{v})}\right) \tag{7}$$

$$-\frac{\partial\, log\, P(\mathbf{v})}{\partial\theta} = \frac{\partial F(\mathbf{v})}{\partial\theta} - \sum_{\tilde{\mathbf{v}}} P(\tilde{\mathbf{v}}) \frac{\partial F(\tilde{\mathbf{v}})}{\partial\theta} \tag{8}$$

To make RBM stability, the energy of system should be the minimum. By the above formulas, $P(\mathbf{v})$ should be maximized. The partial derivative of loss function $-P(\mathbf{v})$ is calculated as:

$$\begin{aligned} -\frac{\partial\, log\, P(\mathbf{v})}{\partial w_{ij}} &= E_{\mathbf{v}}\left[P(h_i|\mathbf{v}) \cdot v_j\right] \\ &\quad - v_j^{(i)} \cdot sigm\left(w_i \cdot v^{(i)} + c_i\right) \\ -\frac{\partial\, log\, P(\mathbf{v})}{\partial c_i} &= E_{\mathbf{v}}\left[P(h_i|\mathbf{v})\right] - sigm\left(w_i \cdot v^{(i)}\right) \\ -\frac{\partial\, log\, P(\mathbf{v})}{\partial b_j} &= E_{\mathbf{v}}\left[P(v_j|\mathbf{h})\right] - v_j^{(i)} \end{aligned} \tag{9}$$

Thus, the parameter $\boldsymbol{\theta}$ corresponding to maximum $P(\mathbf{v})$ is obtained. DBN could then be trained in a greedy layer-wise manner [12]. Each RBM is trained greedily and unsupervised [14]. The posterior distribution of the first RBM is used as the input distribution of the second RBM. Then the weights are fine-tuned by back propagation (BP) neural network. The architecture of DBN model is shown in Fig. 5. Fig. 6 shows the comparison of mean values of the two categories. The difference is significant after feature learning by DBN.

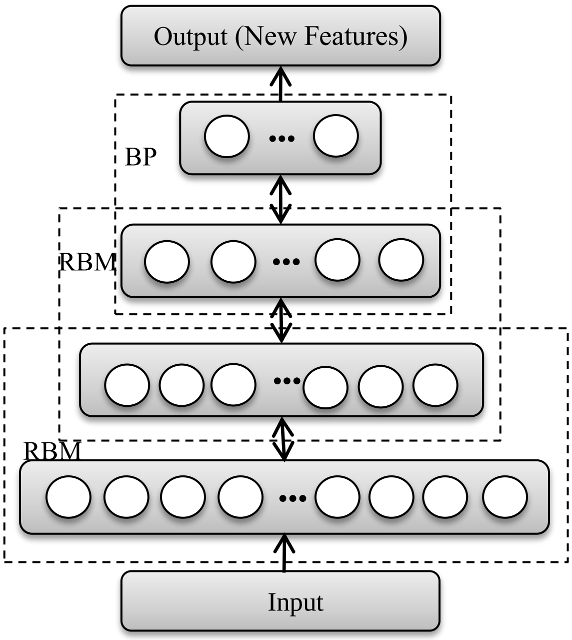

Fig. 5. The architecture of DBN model

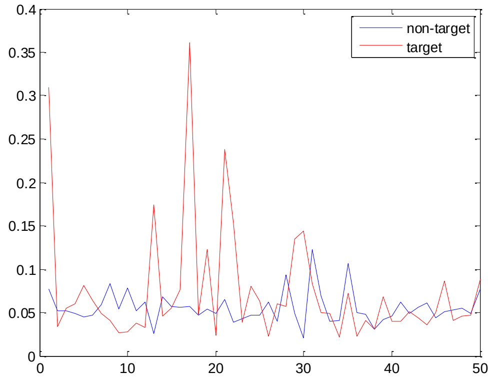

Fig. 6. Comparison of mean values of the two categories

3) Classification.

DBN model is viewed as a feature extraction system in this paper. Outputs of the last model were used as the new input feature vectors with labels of samples to train the SVM classifier.

3 Experiments

Responses to the birth year of the subject are expected to contain the P300 component which is considered as the most typical and common event-related potential (ERP) closely related to human cognitive process, P300 is a late positive component. For the time-locked assumption between the stimulus and the response [15], we took the values of the signal during 0-700ms after stimulus onset from selected electrode channels. The weights were randomly initialized and the turning parameters were set as: learning rate=0.07, momentum=0.95. For the first RBM, the number of the visible units is 200 and the number of the hidden units is 100. For the second RBM, the number of the visible units is 100 and the number of the hidden units is 50. The fifty-dimensional feature vector is input to libsvm.

To reduce the bias of training and testing data, a 10-fold cross-validation method was employed. According to this technique, the dataset was divided into ten subsets [16]. To improve the dependability, the 10-fold cross-validation procedure was performed 10 times. Each time, one of the ten subsets was utilized as the testing dataset and the other 9 subsets were put together to form the training dataset. In particular, the data from test fold is not be involved in the optimization procedure. All final results were averaged over the ten repetitions.

4 Results and Discussion

In this section, the classification performance of the DBN-SVM algorithm was tested on the dataset described in section 2.1. Table. 1 and Fig. 7 show the results. Specifically, Table. 1 shows the recognition accuracy and runtime over all eleven subjects. Fig. 7 compares performances of classifiers adopted different effective feature extraction methods for SVM classifier. All the experiments are repeated ten times, and the average results are reported.

From the effects of perspective, a high average accuracy is obtained. In addition, as shown in Fig. 7, compared with other features used methods, the performance of our approach is significantly better.

Moreover, it is worth noticing that no time-consuming pre-processing operations such as artifact removal or bootstrapping is required which makes the approach possible to be applied to actual tasks.

However, the complex application environment and unpredictable interference will definitely put forward higher requirements considering the practical applications in crime information identification tasks. As for future works, it would be interesting to investigate a way to overall fine-tune the weights of DBN model with regard to SVM learning rule [13].

Table 1. Performances of the algorithmover all subjects

Subject	Amount of samples	Accuracy (%)
S1	150	95.5
S2	300	98.9
S3	300	97.6
S4	150	96.7
S5	150	97.5
S6	300	98.0
S7	300	97.0
S8	150	96.3
S9	300	97.6
S10	300	96.2
S11	450	98.9
Average		97.3

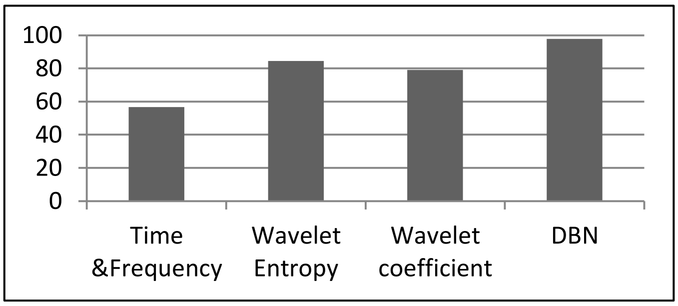

Fig. 7. Comparison of classification performances over different feature extraction methods

5 Conclusion

In this paper, deep learning strategy is applied for signal processing in EEG-based concealed information test. The purpose of introducing DBN is to better express characteristics of different signals. We choose SVM as the classifier which can avoid overfitting effectively. The results show that our method achieves high recognition accuracy. The study in this paper suggests that it is valuable to do further development on deep learning or other computational intelligence strategies used in the field of EEG-based CIT and provide reliable supports to actual investigations in the future.

Acknowledgement

The authors would like to thank Dr.V.Abootalebi and the Research Center of Intelligent Signal Processing (RCISP), Iran, for providing the data. They also would like to thank Dr. Deng Wang with department of Computer Science and Technology, Tongji University, for his data support.

References

1. V. Abootalebi, M. H. Moradi, and M. A. Khalilzadeh. “A comparison of methods for ERP assessment in a P300-based GKT,” International Journal of Psychophysiology, vol. 62, no. 2, pp. 309-320, 2006.
2. M. Zhao, C. Zhang, and C. Zhao, “New approach for concealed information identification based on ERP assessment,” Journal of Medical Systems, vol. 36, no. 4, pp. 2401-2409, 2011.
3. J. R.Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfortscheller, and T. M. Vaughan, “Brain-computer interfaces for communication and control,” Clinical Neurophysiology, vol. 113, no. 6, pp. 767–791, 2002.
4. X. J. Guo, X. P. Wu, and D. J. Zhang, “Motor imagery EEG detection by empirical mode decomposition,” International Joint Conference on Neural Networks, pp. 2619–2622, 2008.
5. M.Y. Zhao, M.T. Zhou, Q.X. Zhu, “Feature extraction and parameters selection of classification model on Brain-Computer Interface,” IEEE 7th International Symposium on Bioinformatics and Bioengineering, pp. 1249–1253, 2007.
6. J. Sherwood, R. Derakhshani, “On classifiability of wavelet features for EEG-based Brain-computer Interfaces,” International Joint Conference on Neural Networks, pp. 2895-2902, 2009.
7. D. P. Subha, P. K. Joseph, U. R. Acharya, and C.M. Lim, “EEG signal analysis: A survey,” Journal of Medical Systems, vol. 34, no. 2, pp. 195–212, 2010.
8. S. Jirayucharoensak, S. Pan-Ngum, and P. Israsena, “EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation,” The Scientific World Journal, vol. 2014, pp. 1-10, 2014.
9. V. Abootalebi,M. H.Moradi, and M. A. Khalilzadeh, “Detection of the cognitive components of brain potentials using wavelet coefficients,” Iranian Journal of Biomedical Engineering, vol. 1, no. 1, pp. 25–46, 2004.
10. D. Wang, D. Miao, and G. Blohm, “A new method for EEG-based concealed information test,” IEEE Transactions on Information Forensics and Security, vol. 8, no. 3, pp. 520-527, 2013.
11. Coates. A, and Y. Ng. Andrew, "Learning feature representations with k-means." Neural Networks: Tricks of the Trade. Springer Berlin Heidelberg, pp. 561-580, 2012.
12. G. E. Hinton and R. R. Salakhutdinov, “Reducing the dimensionality of data with neural network,” Science, vol. 313, no. 5786, pp. 504-507, 2006.
13. G. E. Hinton, “Training products of experts by minimizing contrastive divergence,” Neural Computation, vol. 14, no. 8, pp. 1771–1800, 2002.
14. Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, “Greedy layer-wise training of deep networks,” Advances in neural information processing systems, vol. 19, pp. 146-153, 2007.
15. R. Q. Quiroga, “Quantitative analysis of EEG signals: time-frequency methods and chaos theory,” Institute of Physiology-Medical University Lubeck and Institute of Signal Processing-Medical University Lubeck, 1998.
16. B. D. Ripley, “Pattern Recognition and Neural Networks,” Cambridge University Press, 1996.