EEG-BASED EMOTION CLASSIFICATION USING DEEP BELIEF NETWORKS

Wei-Long Zheng, Jia-Yi Zhu, Yong Peng, and Bao-Liang Lu*

Department of Computer Science and Engineering
Key Lab. of Shanghai Education Commission for Intelligent Interaction and Cognitive Eng.
Shanghai Jiao Tong University, 800 Dong Chuan Rd., Shanghai, China

ABSTRACT

In recent years, there are many great successes in using deep architectures for unsupervised feature learning from data, especially for images and speech. In this paper, we introduce recent advanced deep learning models to classify two emotional categories (positive and negative) from EEG data. We train a deep belief network (DBN) with differential entropy features extracted from multichannel EEG as input. A hidden markov model (HMM) is integrated to accurately capture a more reliable emotional stage switching. We also compare the performance of the deep models to KNN, SVM and Graph regularized Extreme Learning Machine (GELM). The average accuracies of DBN-HMM, DBN, GELM, SVM, and KNN in our experiments are 87.62%, 86.91%, 85.67%, 84.08%, and 69.66%, respectively. Our experimental results show that the DBN and DBN-HMM models improve the accuracy of EEG-based emotion classification in comparison with the state-of-the-art methods.

Index Terms— EEG, Emotion Classification, Affective Computing, Deep Belief Network

1. INTRODUCTION

There are many early interdisciplinary researches about emotion in the past decades. And emotion related study is still very popular in many fields such as neuroscience, psychology and computer science. In computer science, the detection and modeling of emotion plays an important role in human machine interaction and tries to make machines more 'sympathetic'. In multimedia content analysis, affective characteristics of multimedia are important features for describing multimedia content [1]. However, due to the fuzzy boundaries of emotion, the detection and modeling of emotion using artificial intelligence and machine learning techniques still remains a big challenge and many open research questions are still remained [2].

Among the approaches to emotion recognition, methods based on electroencephalogram (EEG) signals are more reliable because of its high accuracy and objective evaluation compared to other external appearance such as facial expression and gesture. However due to the low signal to noise ratio (SNR), it is often very hard to analyze EEG signals 'by hand' even for neurophysiologists. Recent developing deep learning in machine learning community allows automated feature extraction and feature selection and eliminates the limitation of hand-crafted feature [3]. Besides the success in image and speech domain, deep learning methods have been introduced to process physiological signals recent years, such as electroencephalography (EEG), electromyogram (EMG), and electrocardiogram (ECG) [3].

Martinez *et al.* trained an efficient deep convolution neural network to classify four cognitive states (relaxation, anxiety, excitement and fun) using skin conductance and blood volume pulse signals [3]. Martin *et al.* applied deep belief nets (DBN) and hidden Markov model to detect sleep stage using multimodal clinical sleep datasets. Their results using raw data with a deep model were comparable to handmade features approach [4]. Soleymani *et al.* proposed a user-independent emotion recognition method with EEG and eye gaze data and they used logarithms of the power spectral density as EEG features [1]. Duan *et al.* firstly introduced differential entropy to emotion recognition and compared discriminative properties of different features. They used SVM as a classifier and achieved average accuracy of 81.17% [5]. Wang *et al.* systematically compared three kinds of EEG features (power spectrum feature, wavelet feature and nonlinear dynamical feature) for emotion classification. They proposed an approach to track the trajectory of emotion changes with manifold learning [6].

In this paper, we introduce recent advanced deep learning models to EEG-based emotion classification. The main contributions of this paper are as follows: First, we find that neural signatures associated with positive and negative emotions in beta and gamma frequency bands do exist. Second, we show that differential entropy (DE) features extracted from EEG data possess accurate and stable information for emotion classification. Finally, the paper compares the results between deep modals and shallow models like KNN, SVM and

GELM. Moreover, DBN-HMM performs well when compared with the state-of-the-art classification methods.

2. METHODS

Figure 1 shows the overview of the five setups for EEG-based emotion classification used in this work. After feature extraction from multichannel EEG data, we bulid the emotion models with different classifiers.
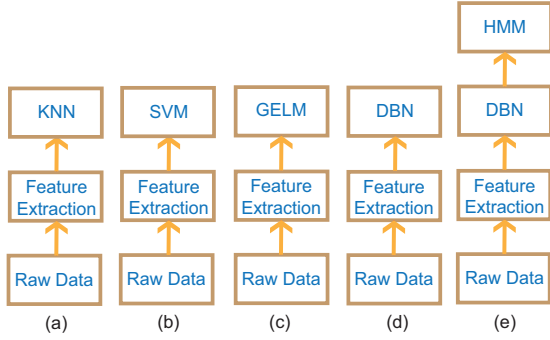
Fig. 1. Overview of the five setups for EEG-based emotion classification used in this work

2.1. Differential Entropy Feature Extraction

Differential entropy extends the idea of Shannon entropy and is used to measure the complexity of a continuous random variable. It has been proven that, for a fixed length EEG segment, differential entropy is equivalent to the logarithm energy spectrum in a certain frequency band [5]. So differential entropy can be calculated in five frequency bands (δ: 1-3Hz, θ: 4-7Hz, α: 8-13Hz, β: 14-30Hz, γ: 31-50Hz) with $O(KN\log N)$ time complexity, where K is the number of electrodes, N is the size of samples.

Since EEG data has the higher low frequency energy over high frequency energy, DE has the balance ability of discriminating EEG pattern between low and high frequency energy. For a specified EEG sequence, we use a 512-point Short-Time Fourier Transform with a non-overlapped Hanning window of 1s to extract five frequency bands signals of raw EEG signals and calculate differential entropy for each frequency band.

2.2. Classification with GELM

Extreme Learning Machine (ELM) is a single hidden layer feed forward neural networks (SLFNs). Peng *et al.* proposed a discriminative Graph regularized Extreme Learning Machine (GELM) based on the idea that similar samples should share similar properties and obtain much better performance in comparison with other state-of-the-art models for face recognition [7].

Given a training data set $L = \{(x_i, t_i)|x_i \in R^d, t_i \in R^m\}$, where $x_i = (x_{i1}, x_{i2}, \cdots, x_{id})^T$ and $t_i = (t_{i1}, t_{i2}, \cdots, x_{im})^T$. In GELM, the adjacent W is defined as follows:

$$x_i = \begin{cases} 1/N_t, & \text{if both } h_i \text{ and } h_j \text{ belong to the } t\text{th class} \\ 0, & \text{otherwise;} \end{cases} \quad (1)$$

where $h_i = (g_1(x_i), \cdots, g_K(x_i))^T$ and $h_j = (g_1(x_j), \cdots, g_K(x_j))^T$ are hidden layer output for two input samples x_i and x_j. Then we can compute the graph Laplacian $L = D - W$, where D is a diagonal matrix and each of the entries in D is the column sums of W. Therefore GELM incorporate two regularization terms into conventional ELM model. The objective function of GELM is as follows:

$$\min_{\beta} ||H\beta - T||_2^2 + \lambda_1 Tr(H\beta L\beta^T H^T) + \lambda_2 ||\beta||_2^2 \quad (2)$$

where $Tr(H\beta L\beta^T H^T)$ is the graph regularization term, $||\beta||^2$ is the l_2-norm regularization term, and λ_1 and λ_2 are regularization parameters to balance two terms.

By setting the differentiate of the objective function (2) with respect to β as zero, we have

$$\beta = (HH^T + \lambda_1 HLH^T + \lambda_2 I)^{-1}HT \quad (3)$$

In GELM, the constraint imposed on output weights enforces the output of samples from the same class to be similar. The constraint can be formulated as a regularization term to the objective function of basic ELM, which also makes the output weight matrix calculated directly.

2.3. Classification with DBN and HMM

Deep Belief Networks (DBN) is a probabilistic generative model with deep architecture, which characterizes the input data distribution using hidden variables. A DBN is constructed by stacking a predefined number of restricted Boltzmann machines (RBMs) on top of each other where the output from a lower-level RBM is the input to a higher-level RBM, see Figure 2. An efficient greedy layer-wise algorithm is used to pre-train each layer of networks.

In an RBM, the joint distribution $p(v, h; \theta)$ over the visible units $\boldsymbol{v}$ and hidden units $\boldsymbol{h}$, given the model parameters, is defined in terms of an energy function $E(v, h; \theta)$ of

$$p(v, h; \theta) = \frac{exp(-E(v, h; \theta))}{Z} \quad (4)$$

For a Bernoulli (visible) - Bernoulli (visible) RBM, the energy function is defined as

$$E(v, h; \theta) = -\sum_{i=1}^{I}\sum_{j=1}^{J} w_{ij}v_i h_j - \sum_{i=1}^{I} b_i v_i - \sum_{j=1}^{J} b_j h_j \quad (5)$$

where w_{ij} is the symmetric interaction term between visible unit v_i and hidden unit h_j, b_i and a_j are the bias term, I and J are the numbers of visible and hidden units, respectively. The conditional probabilities cans be efficiently calculated as

$$P(h_j = 1|v;\theta) = \sigma(\sum_{i=1}^{I} w_{ij}v_i + a_j) \quad (6)$$

$$P(v_j = 1|h;\theta) = \sigma(\sum_{j=1}^{J} w_{ij}h_j + b_i) \quad (7)$$

where $\sigma(x) = 1/(1 + exp(x))$.
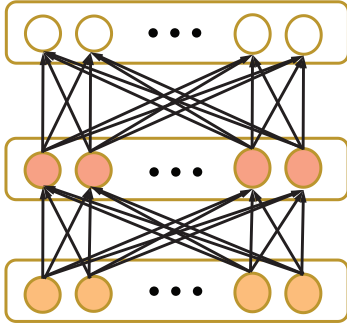
Fig. 2. The graphical depiction of DBN

Taking the gradient of the log likelihood $\log p(v;\theta)$, we can derive the update rule for the RBM weights as:

$$\triangle w_{ij} = E_{data}(v_ih_j) - E_{model}(v_ih_j) \quad (8)$$

where $E_{data}(v_ih_j)$ is the expectation observed in the training set and $E_{model}(v_ih_j)$ is that same expectation under the distribution defined by the model. But $E_{model}(v_ih_j)$ is intractable to compute so the contrastive divergence (CD) approximation to the gradient is used where $E_{model}(v_ih_j)$ is replaced by running the Gibbs sampler initialized at the data for one full step.

EEG-based emotion recognition is actually a sequential pattern recognition and emotional states change slowly. Therefore, we combine DBN and HMM to form a DBN-HMM model. The HMM, based on dynamic programing operations, can help port the strength of a static classifier to handle dynamic or sequential patterns. Combining DBN and HMM can help bridge the gap between static and sequence pattern recognition, which has been successfully used in sleep stage classification using EEG [4].
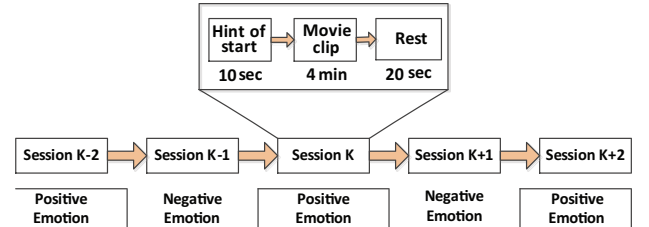
3. EXPERIMENTS

In the experiment, we choose some emotional movie clips to help subjects elicit their own emotion states. There are totally twelve clips (six for positive and six for negative) in one experiment and each of them last for about 4 minutes. All the movie clips are carefully chosen as stimuli to help elicit subjects' right emotion, which include Schindler's List, Titanic, The Sound of Music, Les Miserables, The Day After Tomorrow, Kung Fu Panda, and High School Musical. The selection criteria for movie clips are as follows: (a) the length of the whole experiment should not be too long in case it will make subjects fatigue; (b) the videos should be understood without explanation; and (c) the videos should elicit a single desired target emotion.

Three men and three women with self-reported normal or corrected-to-normal vision and normal hearing participated in the experiments. In advance, the participating test subject will be informed about the procedure. Subjects were instructed to sit comfortably, watch the forthcoming movie clips attentively, and refrain as much as possible from overt movements. Subjects got paid for their participation after the experiments. Each subject participated in the experiment twice at intervals of one week or longer.

We performed the experiments in a quiet environment in the morning or early in the afternoon. EEG was recorded using an ESI NeuroScan System at a sampling rate of 1000 Hz from 62-channel electrode cap according to the international 10-20 system. To remove eye-movement artifacts, we also recorded the electrooculogram (EOG). A pressure sensor was employed to record the response from the subjects in the experiments. There is a 10s hint before each clips and 20s rest after each clip. Figure 3 shows the detailed protocol. For further analysis, the raw EEG data was first downsampled to 200Hz sampling rate. In order to filter the noise and remove the artifacts, The EEG data was then processed with a bandpass filter between 0.3Hz to 50Hz.

Fig. 3. Protocol of the EEG experiment

4. RESULTS AND DISCUSSION

4.1. Differential Entropy Feature

After transforming EEG sequences into frequency domain, we extract the differential entropy feature for each frequency band (Delta, Theta, Alpha, Beta and Gamma) and each channel (62 channels). We totally get 310 DE features for each sample. All the features used here are further smoothed by offline linear dynamic system (LDS) approach to filter out

noise that is unrelated to emotional states. The DE feature map of one experiment is shown in Figure 4.
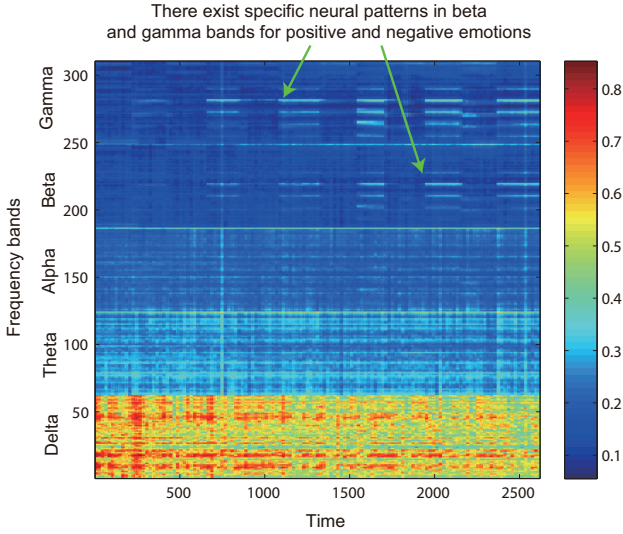
Fig. 4. The DE feature map in one experiment. Time frames are on the horizontal axis, and DE features are on the vertical axis.

Previous researches have shown that higher frequency brain activity reflects emotional and cognitive processes [8] [9]. The observed frequencies have been divided into specific groups, as specific frequency ranges are more prominent in certain states of mind. High frequency bands (Beta and Gamma) have more discriminative information for emotion recognition than lower ones. As shown in Figure 4, the high frequency oscillation has specified patterns for positive or negative tasks, which is very useful for emotion recognition.

4.2. Classifier Training

This study employed and evaluated five classifiers, KNN, SVM, GELM, DBN, DBN-HMM, for EEG-based emotion recognition. For training and testing, data from the first eight sessions of one experiment were used to train the model, and the data from the rest four sessions in the same experiment were used to test it. For KNN, we chosed k=1 for baseline. We also used SVM to classify the emotion state for each EEG segment. Here we used LIBSVM software to build the SVM classifier and employed linear kernel.

For GELM, there are three hyper-parameters: the number of hidden nodes, the parameter λ_1 for graph regularization and λ_2 for l_2-norm regularization. Previous GELM literature has shown that the performance of GELM is not sensitive to the number of hidden nodes. So we set this parameter as 10 times the dimension of input data in all our experiments. We just tuned the two regularization parameters in this study.

For Deep Belief Networks (DBN), we constructed a DBN with two hidden layers. The DBN structure for unsupervised

learning stage is 310-100-30 and for supervised learning stage is 310-100-30-2. The mini-batch size in unsupervised learning stage and supervised learning stage is 200. We set both the unsupervised learning rate and supervised learning rate as 0.05 in the experiment. Before putting DE features into DBN, Scaling to values between 0 and 1 is done by subtracting the mean, divided by the standard deviation and finally adding 0.5. We implemented DBN with the DBNToolbox Matlab code [4] in this study.

4.3. Classification Performance

The performance of different classifiers applied with DE features in twelve experiments of six subjects are shown in Table 1. We also compare the performance of DE feature on different frequency bands (Delta, Theta, Alpha, Beta, and Gamma). Figure 5 shows the plot of the average accuracies for different classifiers in five frequency bands. As we can see from Table 1 and Figure 5, Gamma and Beta frequency bands perform better than other frequency bands. The result shows that beta and gamma oscillation of brain activity are more related than other frequency oscillation.
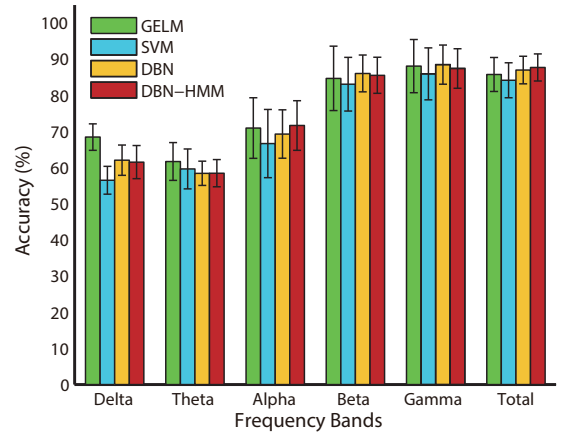
Fig. 5. The average accuracies for different classifiers in five frequency bands and total frequency bands

From the average accuracies and standard deviations in Table 1, we can see that GELM performs better than other classifiers in low individual frequency features, DBN performs better in high individual frequency features and all frequency features. This shows DBN has an ability to perform feature selection procedure to filter out the unrelated features and achieves a better result. Feature extraction and feature selection are crucial phases in the process of emotion modeling. The efficiency of DBN model can combine feature extraction and feature selection when doing unsupervised and supervised learning.

While the average accuracies for different classifiers in five frequency bands are demonstrated in Figure 5, it is also

Table 1. Classification accuracies using different classifiers and different frequency bands

Subject	Trial	Classifier	Delta(%)	Theta(%)	Alpha(%)	Beta(%)	Gamma(%)	Total¹(%)
#1	1	GELM	**77.33**	**66.28**	65.70	**89.97**	**85.32**	**77.33**
		SVM	53.83	65.96	60.21	67.87	47.02	65.53
		DBN	57.10	63.23	65.32	72.58	65.81	74.19
		DBN-HMM	57.90	62.26	**69.36**	71.45	65.32	74.35
	2	GELM	**64.83**	**70.49**	**72.97**	76.04	86.19	73.55
		SVM	47.87	66.11	54.89	65.11	75.53	78.30
		DBN	47.10	56.13	55.81	76.94	86.45	80.48
		DBN-HMM	47.10	59.52	57.26	**78.06**	**87.26**	**81.13**
#2	1	GELM	65.26	48.98	**97.24**	97.82	97.97	97.67
		SVM	62.55	68.72	93.62	97.45	97.45	97.45
		DBN	65.32	65.00	95.97	**98.06**	**98.06**	97.26
		DBN-HMM	**67.26**	**65.16**	96.29	**98.06**	**98.06**	**98.06**
	2	GELM	**77.03**	51.45	**93.75**	97.67	97.82	97.38
		SVM	62.55	**68.72**	93.62	97.45	97.45	97.45
		DBN	64.84	60.65	91.29	**98.06**	**98.06**	**98.06**
		DBN-HMM	66.13	58.87	92.58	**98.06**	**98.06**	**98.06**
#3	1	GELM	**65.26**	**64.97**	**71.51**	68.90	74.13	73.84
		SVM	45.32	44.26	45.74	60.85	75.32	77.66
		DBN	52.10	55.48	65.65	70.00	**83.06**	86.61
		DBN-HMM	52.26	54.68	70.81	**71.29**	82.58	**87.10**
	2	GELM	**81.70**	**61.00**	58.91	**78.81**	**87.32**	**79.78**
		SVM	47.92	53.39	55.47	63.02	84.38	77.08
		DBN	67.05	57.57	**61.40**	73.21	73.54	78.37
		DBN-HMM	67.89	54.74	55.24	73.54	73.71	79.70
#4	1	GELM	**64.85**	**65.17**	73.03	**94.06**	**94.54**	**90.37**
		SVM	41.93	48.44	48.44	89.84	95.05	77.60
		DBN	58.74	47.42	75.21	84.53	92.85	80.87
		DBN-HMM	58.90	46.26	**78.54**	86.69	93.18	82.70
	2	GELM	**71.11**	**77.21**	**76.73**	**94.22**	**97.59**	**92.13**
		SVM	63.54	66.93	60.68	90.88	91.41	94.27
		DBN	67.22	58.07	62.40	91.85	95.51	87.19
		DBN-HMM	70.22	59.57	64.56	91.68	96.01	89.35
#5	1	GELM	59.59	45.49	51.16	71.08	**87.94**	83.14
		SVM	64.89	**60.85**	**63.83**	79.15	79.57	84.26
		DBN	77.90	56.13	55.16	84.03	86.13	89.03
		DBN-HMM	**79.84**	56.13	57.58	**84.84**	86.29	**90.48**
	2	GELM	**59.30**	45.64	41.28	39.10	46.80	76.02
		SVM	56.38	54.68	51.06	72.98	**80.85**	82.13
		DBN	52.58	57.42	58.87	83.06	75.16	**84.19**
		DBN-HMM	52.10	**59.19**	**59.68**	**84.84**	75.97	83.71
#6	1	GELM	**75.92**	63.24	78.81	**97.43**	**95.67**	91.19
		SVM	54.69	42.45	**88.02**	96.61	95.05	86.38
		DBN	59.40	48.75	73.38	92.51	93.34	92.97
		DBN-HMM	59.40	48.92	76.21	91.18	93.68	**93.23**
	2	GELM	**67.90**	69.02	89.89	**97.59**	97.75	**95.67**
		SVM	54.17	**77.34**	**91.93**	96.61	91.93	90.89
		DBN	57.90	72.05	79.53	95.17	97.67	93.68
		DBN-HMM	58.40	75.37	80.70	95.67	**97.84**	93.51
Mean		GELM	**68.37**	**61.63**	70.88	84.60	87.99	85.67
		SVM	56.43	59.57	66.59	82.97	85.86	84.08
		DBN	61.98	58.36	69.22	**85.96**	**88.40**	86.91
		DBN-HMM	61.45	58.39	**71.57**	85.45	87.33	**87.62**
Std.		GELM	**7.35**	10.42	16.75	17.79	14.71	9.37
		SVM	7.70	11.02	18.84	14.84	14.36	9.66
		DBN	8.36	**6.69**	**13.40**	10.16	**10.84**	7.62
		DBN-HMM	9.16	7.52	13.72	**9.95**	10.93	**7.48**

¹ 'Total' represents the combination of five EEG frequency bands.

important to see classification accuracies for each subject. Table 1 shows classification accuracies for four classifiers on 6 subjects, each subject for 2 trials. From Table 1, we can see that while the accuracies vary between different subjects, DBN and DBN-HMM outperform over other existing methods for most subjects according to the results of total frequency bands. There are many factors that may affect the classification accuracies between the subjects, including subjects' edu-

cation background, sociability and their true evoked emotional state when participating in the experiments. The confusion matrix of different classifiers on one trial experiment for one subject is shown in Figure 6.

One of the good questions is whether there exists an efficient model that reliably and robustly identifies emotion in different time for each subject. From Table 1, we can see that our models can achieve similar prediction accuracies for each
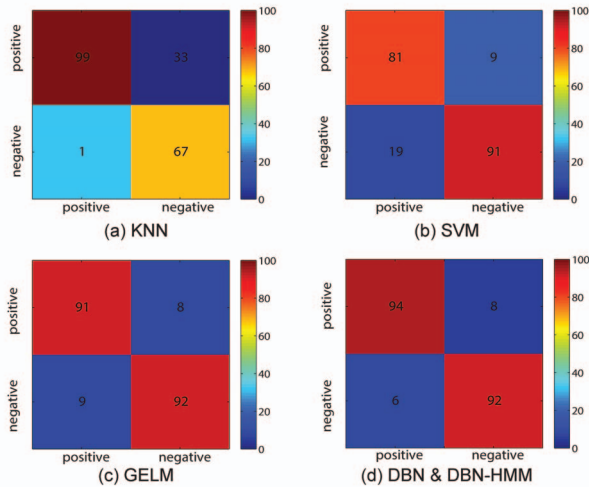
Fig. 6. The confusion matrix of different classifiers on one trial experiment for one subject (the numbers are shown in percents)

subject's two trials, despite manifest differences between people's psychology, which shows the potential strength of the proposed methods to identify emotion in different time.

The best accuracy of all frequency-band features is achieved with the DBN-HMM, followed by DBN, GELM, SVM and lastly KNN. The means and standard deviations of accuracies in percentage (%) of KNN, SVM, GELM, DBN, and DBN-HMM are 69.66/19.80, 84.08/9.66, 85.67/9.37, 86.91/7.62, 87.62/7.48. The results show that DBN-HMM and DBN models outperform over other models with higher mean accuracy and lower standard deviations. The DBN model achieves 2.83% higher accuracy and 2.04% lower standard deviation than SVM. DBN-HMM model achieves 3.54% higher accuracy and 2.18% lower standard deviation than SVM. DBN-HMM perform slightly better than DBN. While GELM sometimes can achieve best performance for some trials, the results of GELM fluctuate more between subjects than DBN and DBN-HMM.

5. CONCLUSION

This study measures brain activity based on EEG and uses machine learning methods to accurately read emotions in individuals. This paper applied deep learning (DL) to the construction of reliable models of emotion recognition from EEG data. The algorithm was tested on 62 channels EEG signals for predicting the positive and negative emotional states while watching emotionally laden movie clips. The dataset was derived from 6 subjects, each for two trails at intervals of one week or longer. We also compared the performance with KNN, SVM and GELM in this study.

In this paper, the experiment results show that high frequency-band (beta and gamma) features are more related to

emotion recognition, which is consistent with previous work, proposing higher frequency brain activity reflecting emotional and cognitive processes [8] [9]. We also showed that the DBN and DBN-HMM methods obtained higher accuracy and lower standard deviation than GELM, SVM and KNN. The reliability of classifications achieved suggests that such neural signatures associated with positive or negative emotions do exist.

6. REFERENCES

[1] Mohammad Soleymani, Maja Pantic, and Thierry Pun, "Multimodal emotion recognition in response to videos," *IEEE Trans. on Affective Computing*, vol. 3, no. 2, pp. 211–223, 2012.

[2] Rafael A Calvo and Sidney D'Mello, "Affect detection: An interdisciplinary review of models, methods, and their applications," *IEEE Transactions on Affective Computing*, vol. 1, no. 1, pp. 18–37, 2010.

[3] Héctor Perez Martinez, Yoshua Bengio, and Georgios N Yannakakis, "Learning deep physiological models of affect," *IEEE Computational Intelligence Magazine*, vol. 8, no. 2, pp. 20–33, 2013.

[4] Martin Längkvist, Lars Karlsson, and Amy Loutfi, "Sleep stage classification using unsupervised feature learning," *Adv. in Artificial Neural Systems*, vol. 2012, pp. 5, 2012.

[5] Ruo-Nan Duan, Jia-Yi Zhu, and Bao-Liang Lu, "Differential entropy feature for eeg-based emotion classification," in *Proc. of 6th International IEEE/EMBS Conference on Neural Engineering,*. IEEE, 2013, pp. 81–84.

[6] Xiao-Wei Wang, Dan Nie, and Bao-Liang Lu, "Emotional state classification from EEG data using machine learning approach," *Neurocomputing*, vol. 129, pp. 94–106, 2014.

[7] Yong Peng, Suhang Wang, Xianzhong Long, and Bao-Liang Lu, "Discriminative graph regularized extreme learning machine for face recognition," *Neurocomputing* (to appear).

[8] William J Ray and Harry W Cole, "EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes," *Science*, vol. 228, no. 4700, pp. 750–752, 1985.

[9] Mu Li and Bao-Liang Lu, "Emotion classification based on gamma-band EEG," in *Proc. of 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2009, pp. 1223–1226.