International Conference on Machine Learning and Computational Intelligence-2017
International Journal of Scientific Research in Computer Science, Engineering and Information Technology

Enhanced EEG-Based Emotion Detection Technique using Deep Belief Network and Wavelet Transform

Sahar Jodat, Khosrow Amirizadeh

ABSTRACT

Today's, the role of emotion in communication , brain-computer interface, brain diseases and mental states, car driver monitoring and recommendation systems is proven. Therefore, automatic emotions detection has become one of the most challenging issue. Until now, numerous studies have been addressed different technique on improving automatic emotion detection.In this study, to achieve bether validation in classification of emotion by EEG signals, we combined wavelet transform with deep belief network. For, non-stationary and time-varying are the most important properties of EEG signals, we decided to use discrete wavelet transform (sym8) for extracting features such as power, then applied deep belief network as a classifier to classify emotions according to two-dimensional arousal-valence model. To examine the effectiveness of the method, we used DEAP database and mapped different emotions on two different classes of valence and arousal. Final results show an acceptable enhancement with the accuracy of 75.52% and 81.03% for valence and arousal, respectively.

Keywords: EEG signals, Discrete Wavelet Transform, Deep Belief Network, two-dimensional arousal-valence model, DEAP

I. INTRODUCTION

Emotions express mental state of the mind and thought process that can be perceived conscious or unconscious in different situations. Introducing different methods for processing signals, easy usage of ellectrodes in collecting data, various classification methods and real-world applications of computer and human interaction for normal people have provided the possibility of different emotion detection by intelligent devices. Generally, there is three different approaches in this case. In the first method, emotions are classified according to analysis of facial expressions or speech [2-4]. The second method take the periphery physiological signals, such as electrocardiogram (ECG), skin conductance (SC), respiration and pulse into account in classification[5-7]. In the last method, brain signals captured from central nervous system such as electroencephalograph (EEG), electrocorticography (ECOG) and functional magnetic resonance imaging (fMRI) become the seat of researchers attention [8-10]. As modern equipment such as electrodes provide collecting EEG signals easily, among all of mentioned methods recognition by EEG signals is a trustable method as these signals contain information of central neural system related to brain activities and have short time answering in detection. EEG signals reflect brain activities and can be acquired by electrodes according to 10-20 system.

In 1949, international standard of 10-20 [11] was introduced to determine the place of electrodes on the scalp. This method provides the possibility of comparing the results of the recording and processing of brain signals of diffierent people at any time, illustrated in Figure 2.

One of the problems with the classification of emotions and their naming is that the distinction between the boundaries of different emotions is not clear, since different individuals express their feelings differently, modeling emotions seems to be difficult. To tackle this problem, researchers have used two different methods for emotion modeling. The first method of emotion modeling is to consider them as separate and discrete senses.The second method is to consider feelings in a multidimensional space. The discrete model is considered as a complete set for describing emotions, such as happiness, sadness, surprise, anger, fear, disgust and the rest of the emotions are derived from the basic emotions. But the main problem with this type of model is that how many and which of the emotions are chosen as the main prototype. For example, Weiner considered only happiness and sadness as basic emotions [12], whereas Kemper suggested fear, anger, depression and satisfactions to be basic [13]. To overcome these problems, multiple dimensional or scale to categorize emotions become popular. For example, Russell describe two dimensional (2D) model for specified emotions by their position [1]. Two dimensional model described by two axes of valence and arousal. Valence represents Positive or negative emotional state of the individuals or, in other words, the rate of pleasure or unpleasantness (horizontal vector). Arousal refers the degree of excitement that a person feels generally, changes from calm to excitement (vertical vector), shown in Figure 3. Different emotions can be labeled in various positions in 2D models. In addition to the 2D model, in some models, the third dimension, is also considered, which represents the degree of dominance and its range from weak to strong.

As automatic emotion detection system can be used in real world and real-time applications, improving the accuracy is an important issue in this case. Until now, a large number of classifiers have been used in this field. For example, support vector machine (SVM), neural network(NN), k-nearest neighbor (KNN), deep belief network (DBN) and so on. Considering this point into account that, deep learning algorithm is capable to represent and classify a set of data in hence of their hierarchical structure and provides comprehensive presentation in comparison with shallow structures, we selected DBN among other classifiers. Then, chose discrete wavelet transform, for extracting statistical features like, power. In this study we used the data of DEAP1 dataset for testing the accuracy of proposed model [16]. In this dataset (is explained in the next section completely), two kinds of data exist, raw data and processed data. We applied processed data that down-sampled (to 128 Hz) and EOG artefacts are removed [17]. The data are labeled in 4 categories, valence, arousal, dominance and liking. We used just valence and arousal dimensions for categorize emotions. We achieved the accuracy of 75.52% for valence and 81.03% for arousal. The final results show acceptable progress compared with other experiments that are explained in section 4. Different steps in our experiment is illustrated in Figure 1.

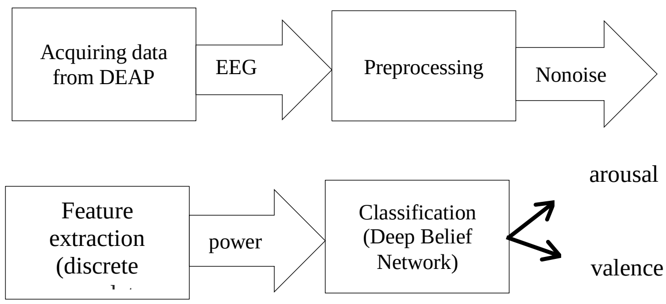

Figure 1. Different steps in emotion detection from EEG

The paper has divided into 5 sections. In section 2, we explained basic contents of applied classification method. Then we analyzed our different part of methodology in section 3. Section 4 is allocated to final results and comparison them with other methods. In last section summarized our work and represent our suggestion for future works.

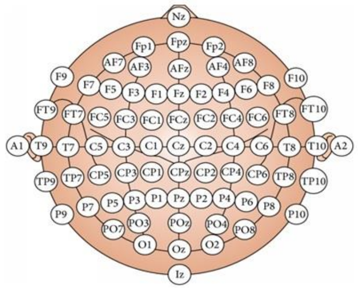

Figure 2. 10-20 international standard of electrodes placement [11]

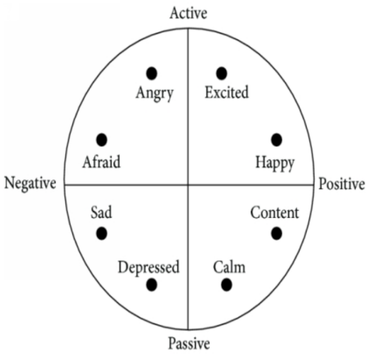

Figure 3. Two-dimensional emotion model [1]

II. DEEP BELIEF NETWORK (DBN)

Neural networks are the base of deep learning. These networks are formed of an input layer, one hidden layer and an output layer to model the human nerves system. Unfortunately, in implementing complicated model, with large number of nodes neural networks do not have goof performance. For solve this problem, deep neural networks are replaced. Generally, deep neural networks consist of a series of shallow networks (such as single-layer neural networks) that enable networks to learn and extract nonlinear hierarchical features. This feature has led to more automatic recognition methods towards deep learning methods. Although, using deep neural network is beneficial, but it has a main problem. Training all layers at once is difficult, with random initialization of weights, they do not converge to the correct answer and it increase the probability of get stuck in local minimum [19]. To overcome the difficulties, Hinton et al. proposed deep belief network. In fact, deep belief network is made up of few layers of the restricted Boltzmann machine (RBMs) [20]. RBMs stacked together with shared layers create a DBN and trained layer by layer in a greedy way [21], [22]. The process of training a deep belief network has two phases .The first step is the unsupervised pre-training, in which unlabelled data is used for training. The training starts from the lowest layer of the network (the first layer) and features are derived from raw input data. Then the training takes moves up to higher level (between the hidden nodes of the first layer and the second layer). The training of the hidden nodes of the first layer is a new input for obtaining the features in the second layer's hidden nodes. Greedy training continues to reach the topmost layers of hidden nodes. Finally, a productive model with weights between layers trained by using input data features. The greedy layer-to-layer training method will calculate the weights and biases of different layers. Fine-tuning weights and supervised learning are performed in the second phase of training at the upper layer. In this phase a new label layer will be added to the upper layer of the deep belief network and removes all links in the top-down direction. Now, the DBN becomes a feed-forward neural network, shown in Figure 4(b).Then the backward algorithm is used to learn the weights and biases that are trained based on labels. The goal of learning is to reduce the classification error from labeled examples. Weights and biases are initialized in the non-supervision training phase, except those that are randomly assigned in the upper layer. Figure 4 illustraites the structure of a DBN with three hidden layers.

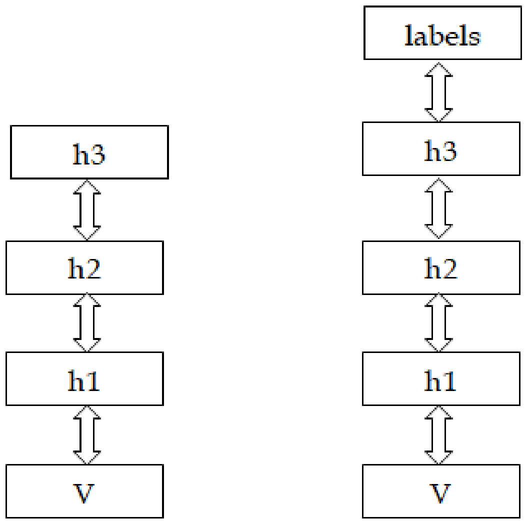

Figure 4. The structure of a DBN with three hidden layers: (a) The pre-training stage with un-labeled data and (b) The fine-tuning stage using the new layer added at the top of the network that is trained by labeled data[19].

2.1 Restricted Blotzmann Machine (RBM)

Restricted Boltzmann machine is an energy based generative model with two binary layers (visible and hidden). In other words, restricted Boltzmann machine is a two-part, weighted, non-directional, symmetrical graphical model that takes random decisions about the status of the nodes, whether on or off. A graphical model of an RBM is shown in Figure 5.

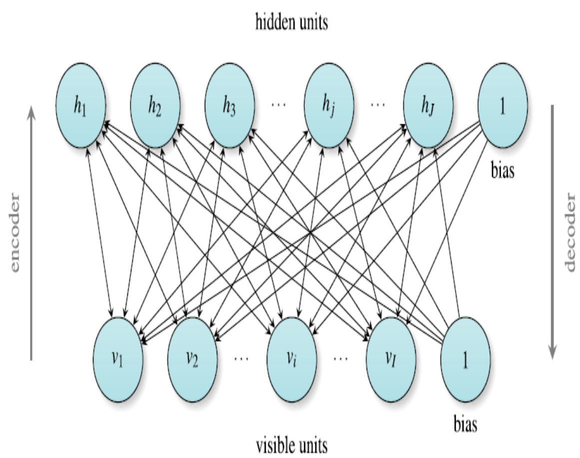

Figure 5. Restricted Boltzman machine [24]

The energy function for connecting visible layer to hidden layer obtain from the following equation [24] :

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

Where w_{ij} is the weight between visible unit i and hidden units j, a_i and b_j refer to biases in visible (v) and hidden (h) layer. The probability of given configuration is the normalized energy function [24] :

$$p(v,h) = \frac{e^{-E(v,h)}}{\sum_{u,g} e^{-E(u,g)}} \quad (2)$$

Since there is no direct connection between hidden units in an RBM, these units considered independent according to visible units. The binary state h_j of each hidden units j will be equal one (activated) with the below conditional probabilities :

$$P(h_j = 1|v) = g(b_j + \sum_i v_i w_{ij}) \quad (3)$$

Where g(x) is the logistic sigmoid function $g(x) = \frac{1}{1+e^{-x}}$.

Similarly, there is no direct connection between visible units in an RBM, with having a hidden vector, it will be easy to calculate unbiased the state of a visible unit (activated).

$$P(v_i = 1|h) = g(a_i + \sum_j h_j w_{ij}) \quad (4)$$

Restricted Boltzmann machines are trained to maximize the product of probabilities of a set of training examples X:

$$\mathrm{argmax}_w \prod_{x\in X} P(x) \quad (5)$$

or equivalently to maximize the log likelihood

$$\mathrm{argmax}_w \sum_{x\in X} \log P(x) \quad (6)$$

Unfortunately, calculating the gradient of the log likelihood is so difficult. Therefore, [22] proposed contrastive divergence (CD) by doing k iterations of Gibbs sampling to approximate it. Also, using CD method, enable an RBM to update weights according to Equation 7.

$$\Delta w_{ij} = \varepsilon(< v_i h_j >^0 - < v_i h_j >^k) \quad (7)$$

where $<\cdot>^m$ is the average in a contrastive divergence iteration m and ε is the learning rate.

2.2 Contrastive divergence

To solve the problem of calculating the log-likelihood gradient, Hinton proposed a contrastive divergence method in 2002. In this method, the state of visible units is initialized to training data. Then the binary state of the hidden units is calculated according to Equation 3. After the binary state of the hidden units is computed, the values of v_i will be update based on Equation 4. At the end, again, the probability of activation the hidden units is computed and the value of $<\cdot>^m$ will be calculated from the final values obtained from the hidden and visible units.

2.3 Softmax classifier

The softmax classifier is used to estimate the probability of output values in a deep belief network. The method of this type of classification is to learn all the parameters of weights and biases using the featured learned from the last hidden layer. In the case of binary classification (k = 2), the softmax regression hypothesis output $h_\theta(x)$ is obtained from the following Equation :

$$h_\theta(x) = \frac{1}{e^{\theta_1^T x} + e^{\theta_2^T x^{(i)}}} \begin{bmatrix} e^{\theta_1^T x} \\ e^{\theta_2^T x} \end{bmatrix} \quad (8)$$

Softmax classifier can be generalized to be multiclass classification .The hypothesis will output a vector of k estimated probabilities, shown as follows:

$$h_\theta(x) = \frac{1}{\sum_{j=1}^{k} e^{\theta_j^T x^{(i)}}} \begin{bmatrix} e^{\theta_j^T x^{(i)}} \\ e^{\theta_j^T x^{(i)}} \\ \cdot \\ \cdot \\ \cdot \\ e^{\theta_j^T x^{(i)}} \end{bmatrix} \quad (9)$$

The softmax layer needs to learn the weight and bias parameters with supervised learning approach by minimizing its cost function, shown as follows:

$$\text{cost} = -\frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{k} 1\{y_i = j\}\log\frac{e^{\theta_j^T x^{(i)}}}{\sum_{l=1}^{k} e^{\theta_j^T x^{(i)}}} + \frac{\lambda}{2}\sum_{i=1}^{k}\sum_{j=1}^{n}\theta_{ij}^2 \quad (10)$$

where m is number of hidden units, n is number of inputs, k is number of classes, y is ground truth, and θ is weight of hidden nodes.

III. METHODOLOGY

3.1 DEAP Dataset

DEAP [16] is a a multi-model database designed to provide signals for emotion detection. In this database, video clips are used as visual stimulus for stimulation different emotions. This database contains a set of brain signals that are used to analyze emotions. The way of collecting information in this database is that in this way, 40 pre-selected video clips, each one for one minute, are displayed as emotional stimulus for 32 participants between 19 and 37 years old, of which 50% is female, and EEG signals and other fuzzy signal signals, such as ECG, EMG GSR and BVP, are collected from 40 channels during video viewing.

The ordering of videos is based on the code number of the test, not on the order in which it is displayed, which means that the first video clip is the same for each participant. It should be noted that the electrodes are arranged according to the standard 10-20. After the end of each clip, participants rate it according to arousal, valence, liking or not, level of dominance and familiarity.

The format of the data is 40*40*8064 that represent the concept of video/trial*channel*data, similarly the format of the labels are 40*4 (valence, arousal, dominance, liking).

Self-assessment manikins (SAM) [26], as shown in Figure 6, were used to visualize the scales. The scales between 1 and 9 for 2 different levels of valence and arousal are mapped in the order below. For valence dimension the numbers between 1 and 3 represents negative emotions, numbers between 4 to 6 represent neutral feelings and numbers between 7 to 9 represent positive emotions, while in the dimension of the arousal the numbers between 1 and 3, represent the inactive emotions, the numbers Between 4 to 6 neutral feelings and numbers 7 to 9 represent active emotions

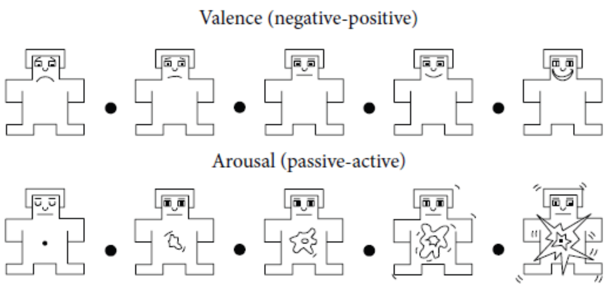

Figure 6. An example of a self-assessment. The first line above indicates the feelings of valence and the second line is allocated to arousal [26]

3.2 Channel selection

In order to reduce the number of the EEG channels as much as possible and implement, an emotion recognition method that would result in a more user-friendly environment in the future, the signals were acquired from fifteen positions only, according to the 10–20 system. According to essay [27] we chose the EEG signals recorded at positions AF3, F7, C3, T7, CP5, Pz, AF4, F8 FC6, FC2, CZ, C4, T8, CP2, O2. IN addition, the left frontal area is involved in the experience of positive emotions (high values of valence), such as joy or happiness (the experience of positive affect facilitates and maintains approach behaviors), whereas the right frontal region is involved in the experience of negative emotions (lower valence values), such as fear or disgust (the experience of negative affect facilitates and maintains withdrawal behaviors) [28].

3.3 Feature extraxtion

A wavelet transform is a variable-length window technique that uses a time-scale domain. A wavelet is a definite function with a mean of zero and with a limited period, and expansion is carried out based on transformation and scale. A wavelet transform with a multi-resolution analysis feature is suitable for analyzing the signal at different time and frequency bands. In fact, a wavelet is a mathematical transformation function that divides the signal into different frequency bands. Wavelet transform is the representation of a function by mother wavelets. ($\psi_{a,b}$, the mother wavelet).

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \tag{11}$$

Where a,b ∈ R (a>0), R is the wavelet space. Parameter 'a' is the scaling factor and 'b' is the shifting factor. The only limitation for choosing a prototype function as mother wavelet is to satisfy the admissibility condition.

The time-frequency representation is performed by repeatedly filtering the signal with a pair of filters namely high pass filter (H(n)) and low pass filter (L(n)), that cut the frequency domain in the middle. Specifically, the discrete wavelet transform decomposes the signal into an approximation coefficients (CA) and detailed coefficients(CD). The approximation coefficient is subsequently divided into new approximation and detailed coefficients. This process is carried out iteratively producing a set of approximation coefficients and detail coefficients at different levels or scales [29].

In this work, the multiresolution analysis of wavelet functions, namely sym8 was used to decompose the EEG signals into five different frequency bands delta(0-4)Hz, theta(4-8)Hz, alpha(8-12)Hz, beta(12-30)Hz and gamma(>30)Hz that the characteristics of each band can be utilized to estimate subject's cognition and emotion states. This wavelet functions was chosen due to it is near optimal time-frequency localization properties. Moreover, the waveforms of this wavelet was similar to the waveforms to be detected in the EEG signal. In order to analyze the characteristic natures of different EEG patterns, we derived linear feature (power). This feature was derived from the five frequency bands of EEG and was concatenated to form a feature vector. Table 1.

Table 1. Statistical feature used for emotion recognition and it's description

Feature	Formula	Description
Power	$P_j = \frac{1}{N}\sum_{k=1}^{N}(d_j(k)^2)$ $d_j(k)$ is the detail wavelet coefficient	Measure the squares of amplitude of EEG signals

J = decomposition level; k = No. of wavelet coefficient, varies from 1 to N

3.4 Classification

In this study, proposed DBN contains a layer as an input, three hidden layers with two softmax classifiers in output layers, one for valence and another for arousal. Here the data is divided in two parts 90 % of data is used for train the system and remaining 10% data is used for testing the data. The DBN uses unsupervised pre-training technique with greedy layer-wise training, starting from the input layer to the softmax layer. The first hidden layer is trained on the inputs' features extracted from data to learn the primary features of first hidden layer on these input features. Subsequently, the algorithm performs forward propagation by using the input features into this trained hidden layers to obtain the primary feature activations. The features, deriving from feedforward propagation of the 1st hidden layer, must be used to perform unsupervised pretraining in the second hidden layer. The algorithm computes its features in the same procedure from the learned features from the previous hidden layers. The weight and bias parameters of the softmax layer are trained by using a supervised learning approach. The output features of the last hidden layer are used as the input features of both softmax layers. We used a set of self-assessment emotion states (valence and arousal) of subjects as a ground truth. These softmax layers can be trained as the parameters concurrently. After the network finishes learning weight and bias parameters in both softmax classifiers, the algorithm has to perform fine-tuning of all weight and bias parameters in the whole network simultaneously. However, we are not able to use the same network parameters for two classifiers. We need to save the learned parameter outcomes of unsupervised pretraining and load the parameters for fine-tuning process of another softmax classifier. The fine-tuning process treats all layers of a stacked hidden layers and softmax layer as a single model and improves all the weights of all layers in the network by using backpropagation technique with supervised approach. The backpropagation process is used to learn the network weights and biases based on labeled training examples to minimize the classification errors. For evaluating the accuracy of the proposed system cross validation with six repetition has been done. We implemented the proposed model with DeeBNet toolbox [23], and for setting primary parameters performed like Hinton essay [30].

IV. RESULTS AND DISCUSSION

The proposed model is tested to classify EEG signals from DEAP dataset. The signals are collected from 15 channels out of 40. Whole data are divided into test and training set and the cross validation is used to validate the performance of classification results. Finally, the average of all (32) participants' classification accuracies was investigated. We categorized each emotional valence and arousal into two-level class according to the SAM ratings values on a scale of 1–9 [33].Table 2. The results for valence and arousal is shown in Table 3.

Table 2. SAM rating for each emotion class- the conditions for categorizing the emotional class levels

	Two-level class	
	High	Low
SAM rating (S_R)	$S_R \geq 5$	$S_R < 5$

Table 3. Arousal and Valence classification accuracy (%)

Subjects	Valence	Arousal
S01	70.83	70.83
S04	91.67	83.33
S14	66.67	87.5
S15	79.17	75
S16	83.33	79.17
S24	83.33	95.83
S25	79.17	87.5
S26	62.5	66.67
S27	66.67	79.17
S28	87.5	79.17
S31	79.17	87.8
S32	62.5	83.33

According to the Table 3 the best result is for subject 04 with 91.67% in valence and 95.83% for subject 24 in arousal. However, the average of final results for 32 participants were 75.52% and 81.03%, in a row. In continues, to show that the suggested model has better accuracy, at first we compared the final results with the results in essay [18]. In this paper, the validity of valence and arousal were analyzed for 10 subjects separately. In the experiment, six statistical features such as mean, standard deviations, means of the absolute values of the first differences of the raw signals, means of the absolute values of the first differences of the normalized signals, means of the absolute values of the second differences of the raw signals, means of the absolute values of the second differences of the normalized signals are calculated. In addition, fractal dimension (FD) values are calculated and support vector machine classifier with polynomial kernel is used for classification. Researchers tested the performance of their proposed method by using the data of 10 subjects from DEAP database. In classification according to arousal dimensional they used the combination of arousal-dominance or high and low dominance states.

They propose a novel subject-dependent valence level recognition algorithm and apply it to recognize up to 16 emotions where 4 levels of valence are identified with each of the four arousal-dominance combinations, and to recognize up to 9 levels of valence states with controlled dominance level (high or low). In the proposed emotions recognition algorithm, first, four classes of combinations of high/low dominance and high/low arousal levels or two classes of high/low dominance are recognized.

The resulting accuracy using SVM for four arousal-dominance combinations is shown in Table.4. As we can see from the table, the best and worst accuracy obtained in recognition of four arousal-dominance combinations are 80.50% for subject 13 and 46.67% for subject 14, while in our study, the highest and lowest accuracy was 95.83% for subject 24 and 66.67% for subject 25. In addition, the average accuracy across all subjects was 63.04% which contrast highly with 81.03% for our proposed model.

Table 4. Arousal-Dominance recognition accuracy (%)

Subjects	Arousal
S01	63.25
S05	53.08
S07	74.17
S10	65.21
S13	80.50
S14	46.67
S16	67.12
S19	67.29
S20	58.41
S22	49.72
Avg	63.04

However, in the same essay researchers evaluated the performance of valence by six statistical features (mentioned above) and fractal dimension, received the accuracy approximately 50%. Table5 According to the information of the Table 5, fractal dimension features give better accuracy compared with other features in classification with SVM. The average value of accuracy for this feature are around 50%. According to Table 3, in our model the lowest value of accuracy is 62.5% that even is higher than the average in Table5. In our method, the rest of value of accuracy are more than 70%.

Table 5. Valence recognition accuracy (%), $\overline{\Delta\mu_x}$ (mean), $\overline{\Delta\sigma_x}$ (standard deviations), $\overline{\Delta\delta_x}$ (means of the absolute values of the first differences of the raw signals), $\overline{\overline{\Delta\delta_x}}$ (means of the absolute values of the first differences of the normalized), $\overline{\Delta\gamma_x}$(means of the absolute values of the second differences of the raw signals), $\overline{\overline{\Delta\gamma_x}}$(means of the absolute values of the second differences of the normalized Signals)

subject	$\overline{\Delta\mu_x}$	$\overline{\Delta\sigma_x}$	$\overline{\Delta\delta_x}$	$\overline{\overline{\Delta\delta_x}}$	$\overline{\Delta\gamma_x}$	$\overline{\overline{\Delta\gamma_x}}$	Fractal dimension
S01	50.00	45.16	63.71	53.23	55.65	50.8	45.16
S05	38.31	46.77	57.66	64.92	50.40	56.8	48.79
S07	49.19	58.06	50.00	47.58	45.97	50.8	56.45
S10	50.00	51.61	55.65	54.84	54.03	45.9	65.32
S13	42.74	35.48	50.00	42.74	53.23	47.5	34.68
S14	42.74	47.58	40.32	54.03	44.35	57.2	54.03
S16	48.39	47.58	37.10	54.42	42.74	49.1	54.84
S19	48.39	54.84	55.65	51.61	50.40	54.4	50.00
S20	41.94	43.55	46.77	41.13	45.97	44.3	52.42
S22	49.19	49.19	27.42	47.58	28.23	48.3	53.23
Avg	46.09	47.98	48.43	51.01	47.10	50.5	51.49

Finally, for more comparison, we explained three different other methods then illustrated the final outcome in Table 6.

In paper [31], two and three categories of the DEAP database are used to classify valence and arousal. AR regression coefficients have been calculated as features. Feature selection is done using sequential forward feature selection (SFS) to decrease the complexity of computing and redundancy of features. Then, the three KNN, LDA and QDA clasifiers are used for categorization and the final results are compared. The best results are between %72.33 and %74.2 for the classification of the two valence and arousal categories and 61.1 and 65.16 for the classification of the three classes.

In essay [32] researchers integrated singular value decomposition (SVD) and Deep Belief Network (DBN) to gain better results. For achieving their goal, they used signals of DEAP database, then extracted information of channels F3 and F4. They applied empirical mode decomposition (EMD) method for decomposing EEG signals to into a set of intrinsic mode functions (IMFs). Then, effective components of IMFs were selected by using SVD. Extracted features that reduced by standard deviation (SD) method were considered as an input for DBN. Finally, classifier mapped emotion to three classes of valence and arousal. The results show the accuracy of 57.25% and 56.70% in a row.

In paper [33], emotions are classified according to the model of arousal-valence by using Fast Fourier transform analysis to extract features and Pearson correlation coefficient method to feature selection. Then researchers used a probabilistic classifier based on Bayes theorem with supervised learning using a perceptron convergence algorithm. To verify the proposed methodology, they used an open database, DEAP. They achieved the average accuracy of the valence and arousal, 70.9% and 70.1%, respectively.

Table 6. The accuracy of valence and arousal with different method of feature extraction and classifier (%)

Method	Valence	Arousal	Reference
SFS & KNN	72.33%	74.20%	[31]
SVD & DBN	57.25%	56.70%	[32]
FFT & Bayes	70.9%	70.1%	[33]
Proposed model	75.52%	81.03%	

According to the Table 6, in the case of valence, our system has the highest value 75.52%, which was closely followed by [31], [32]. In fact there is a few difference with other two methods. But, in the case of arousal this difference become more, that show accuracy has improved noticeably.

V. CONCLUSION

Deep belief network as a classifier is capable of discovering unknown features coherences of input signals that is crucial for the learning task to represent such a complicated model. The DBN provides hierarchical feature learning approach. When learning algorithms process more data, they provide better performance. The key advantage of self-taught learning and unsupervised feature learning is that the algorithm can learn from unlabeled data, and then it can learn from massive amount of information. Consequently, DBN algorithm is suitable for problems where there are a plenty of sets of unlabeled data and a handful amount of sets of labeled data. According to this, We developed a DBN based on restricted Boltzmann machine for classifying emotions. Discrete wavelet transform was used to extract linear features such as, power from EEG signals of 15 channels. Extracted features were considered as an input vector of DBN. An open access DEAP database employed for evaluating the efficiency of model. Finally, we obtained the accuracy rate of 75.52%, and 81.03%, for valence and arousal, respectively. It is shown that the proposed method has better performance in comparison with mentioned methods in section 4.

In future work, The performance of DBNs on the raw data from more than 15 channels in the dataset, up to all the 40 channels, should be investigated. Also, development methods for selecting channels is necessary to improve the performance of the algorithm.

Secondly, we will develop the model from two dimensions to four dimensions . We will investigate dominance and liking, too. Thirdly, For real application, the accuracy should be further improved. The DBN has some parameters that could be effectively improved to get better result. In the future, we will work on improving DBN structure and use other features and feature extraction methods.

VI. REFERENCES

[1]. Russell, James (1980). "A circumplex model of affect". Journal of Personality and Social Psychology. 39: 1161–1178.

[2]. V. Petrushin, Emotion in speech: recognition and application to call centers, in: Proceedings of the Artificial Networks in Engineering Conference, 1999, pp. 7–10.

[3]. M. Black, Y. Yacoob, Recognizing facial expressions in image sequences using local parameterized models of image motion, Int. J. Comput. Vis. 25 (1997) 23–48.

[4]. K. Anderson, P. McOwan, A real-time automated system for the recognition of human facial expressions, IEEE Trans. Syst. Man Cybern. B Cybern. 36 (2006) 96–105.

[5]. J. Wagner, J. Kim, From physiological signals to emotions: Implementing and comparing selected methods for feature extraction and classification, in: Proceedings of IEEE International Conference on Multimedia and Expo, 2005, pp. 940–943.

[6]. K. Kim, S. Bang, S. Kim, Emotion recognition system using short-term monitoring of physiological signals, Med. Biol. Eng. Comput. 42 (2004) 419–427.

[7]. J. Brosschot, J. Thayer, Heart rate response is longer after negative emotions than after positive emotions, Int. J. Psychophysiol. 50 (2003) 181–187.

[8]. J.Coan,J.Allen,Frontal EEG asymmetry as a moderatoran dmediator of emotion, Biol.Psychol.67(2004)7–50.

[9]. P.Petrantonakis, L.Hadjileontiadis, A novel emotion elicitation indexusing frontal brain asymmetry for enhanced EEG-based emotion recognition, IEEE Trans. Inf.Technol. Biomed.15(2011)737–746.

[10]. X.Li,B.Hu,T.Zhu,J.Yan,F.Zheng,Towards affective learning with an EEG feedback approach, in: Proceedings of the 1st ACM International Workshopon Multimedia Technologies for Distance Learning, 2009,pp.33–38.

[11]. Klem, G. H., Lüders, H. O., Jasper, H. H., & Elger, C. (1999). The ten-twenty electrode system of the International Federation. Electroencephalogr Clin Neurophysiol, 52(3), 3-6.

[12]. B. Weiner, Attribution, emotion, and action, Handbook of Motivation and Cognition: Foundations of Social Behavior 1 (1986) 281–312.

[13]. T. Kemper, A Social Interactional Theory of Emotions, Wiley, New York, 1978z

[14]. R.Davidson,G.Schwartz,C.Saron,J.Bennett,D.Goleman,Frontalversus parietalEEGasymmetryduringpositiveandnegativeaffect,Psychophysiology 16(1979)202–203

[15]. P.Ekman,R.Davidson,TheNatureofEmotion:FundamentalQuestions,Oxford UniversityPress,1994

[16]. http://www.eecs.qmul.ac.uk/mmv/datasets/deap/

[17]. Koelstra, S., Muhl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T., ... & Patras, I. (2012). Deap: A database for emotion analysis; using physiological signals. IEEE Transactions on Affective Computing, 3(1), 18-31.

[18]. Liu, Y., & Sourina, O. (2013). Real-time fractal-based valence level recognition from EEG. In Transactions on Computational Science XVIII (pp. 101-120). Springer, Berlin, Heidelberg.

[19]. 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. 153, 2007.

[20]. P. Smolensky, "Information processing in dynamical systems: Foundations of harmony theory," Parallel distributed processing:

explorations in the microstructure of cognition, Cambridge, MA, USA: MIT Press, 1986, pp. 194-281.
[21]. G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," Science, vol. 313, pp. 504-507, 2006.
[22]. G. E. Hinton, "Training products of experts by minimizing contrastive divergence," Neural computation, vol. 14, pp. 1771-1800, 2002.
[23]. divergence," N Keyvanrad, M. A., & Homayounpour, M. M. (2014). A brief survey on deep belief networks and introducing a new object oriented toolbox (DeeBNet). arXiv preprint arXiv:1408.3264.
[24]. Lopes, N., & Ribeiro, B. (2015). Deep Belief Networks (DBNs). In Machine Learning for Adaptive Many-Core Machines-A Practical eural computation, vol. 14, pp. 1771-1800, 2002.
[25]. Koelstra, S., Muhl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T., ... & Patras, I. (2012). Deap: A database for emotion analysis; using physiological signals. IEEE Transactions on Affective Computing, 3(1), 18-31.
[26]. Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. Journal of behavior therapy and experimental psychiatry, 25(1), 49-59.
[27]. Kshirsagar, P., & Akojwar, S. (2016). Classification of Human Emotions using EEG Signals. International Journal of Computer Science, Volume146, (7).
[28]. Petrantonakis, P. C., & Hadjileontiadis, L. J. (2011). A novel emotion elicitation index using frontal brain asymmetry for enhanced EEG-based emotion recognition. IEEE Transactions on information technology in biomedicine, 15(5), 737-746.
[29]. Parameswariah, C., & Cox, M. (2002). Frequency characteristics of wavelets. IEEE Transactions on Power Delivery, 17(3), 800-804.
[30]. Hinton, G. (2010). A practical guide to training restricted Boltzmann machines. Momentum, 9(1), 926.
[31]. Hatamikia, S., Maghooli, K., & Nasrabadi, A. M. (2014). The emotion recognition system based on autoregressive model and sequential forward feature selection of electroencephalogram signals. Journal of medical signals and sensors, 4(3), 194.
[32]. Hosseini M, Pouyan A, Ferdosi S,Mashayekhi H,(2016),"Emotion Recognition of EEG data using Deep Belief Network and Empirical Mode Decomposition",submitted on journal of computational of Neuroscience
[33]. Yoon, H. J., & Chung, S. Y. (2013). EEG-based emotion estimation using Bayesian weighted-log-posterior function and perceptron convergence algorithm. Computers in biology and medicine, 43(12), 2230-2237.