Providing a Four-layer Method Based on Deep Belief Network to Improve Emotion Recognition in Electroencephalography in Brain Signals

Abstract

Background: One of the fields of research in recent years that has been under focused is emotion recognition in electroencephalography (EEG) signals. This study provides a four-layer method to improve people's emotion recognition through these signals and deep belief neural networks. **Methods:** In this study, using DEAP dataset, a four-layer method is established, which includes (1) preprocessing, (2) extracting features, (3) dimension reduction, and (4) emotion identification and estimation. To find the optimal choice in some of the steps of these layers, three different tests have been conducted. The first is finding the perfect window in feature extraction section that resulted in superiority of Hamming window to the other windows. The second is choosing the most appropriate number of filter bank and the best result was 26. The third test was also emotion recognition that its accuracy was 92.93 for arousal dimension, 92.64 for valence dimension, 93.14 for dominance dimension in two-class experiment and 76.28 for the arousal, 74.83 for the valence, and 75.64 for dominance in three-class experiment. **Results:** The results of this method show an improvement of 12.34% and 7.74% in two- and three-class levels in the arousal dimension. This improvement in the valence is 12.77 and 8.52, respectively. **Conclusion:** The results show that the proposed method can be used to improve the accuracy of emotion recognition.

Keywords: *Deep belief neural network, deep neural network, electroencephalography, emotion recognition, independent component analysis*

Seyed Mohammad Reza Mousavinasr[1], Ali Pourmohammad[2], Mohammad Sadegh Moayed Saffari[1]

[1]*Department of Computer Engineering, Malek Ashtar University of Technology,* [2]*Department of Electrical Engineering, Amirkabir University of Technology, Tehran, Iran*

Introduction

In today's world that various waves and signals surround humans, the use of these signals and waves in different cases is inevitable. In many cases, they are the solution of problems in different areas. As a result of activities, thoughts, ideas, and feelings of human beings, signals are generated in the brain. These signals can be analyzed and used in various ways. There are various types of human brain signals, of which electroencephalography (EEG) signals are the most widely used. This signal is an electrical signal that is generated as a result of brain activity. The signal also contains various information that can be used in different contexts. Among these contexts is individuals' emotion recognition against various sensory stimuli that are usually induced through observation and sometimes by ear and through human hearing.

During these stimulations, different feelings of individuals emerge in EEG brain signals. These signals that contain useful information about people's feelings can be analyzed and emotions can be extracted. The importance of this task becomes clear when emotions, while only are imagined, appear in brain signals without moving a muscle. It can be easily inferred that this is so useful for people who lack mobility and suffer muscle diseases such as sclerosis, or hardening of the brain tissue, and muscle paralysis.

In this study, important papers presented between 2010 and 2015 are reviewed. These articles are about emotion recognition through brain signals using different methods. Majority of them have used two-dimensional (2D) arousal-valence model proposed by Russell[1] or 3D arousal-valence-dominance model which has dominance for controlling the emotions. Considering to previous works, deep learning is a way to improve the accuracy of emotion recognition in EEG brain signals.

Address for correspondence:
Dr. Ali Pourmohammad, Department of Electrical Engineering, Amirkabir University of Technology, Tehran, Iran. E-mail: pourmohammad@aut.ac.ir

Access this article online

Website: www.jmssjournal.net

DOI: 10.4103/jmss.JMSS_34_17

Quick Response Code:

In this method, according to previous research conducted at Malek Ashtar University, it is required to remove noise from signals to increase the accuracy of diagnosis in different applications. Imani *et al.*[2] conducted a study as "investigating independent-component analysis (ICA) algorithm to distinguish between two conceptual groups of risk and informing words using traffic signs" and proposed ICA algorithm as one of the best methods to eliminate noise from EEG brain signals. This method has also been used by Imani *et al.*[2] in their study entitled as "evaluation of the performance of the separator and the extracted features to differentiate patterns of brain related to mental activities associated with the four cardinal directions," who presented more appropriate and acceptable results compared to previous research; accordingly, it is decided to benefit from this method in the current study.

Research Significance and Objective

The main objective of this study is to improve emotion recognition through aroused brain signals using deep learning, which is one of the identification and classification methods.

Given the application of this research, the significance of this study can be studied from various aspects. One of these aspects is helping solve some problems for amyotrophic lateral sclerosis patients. These patients lack muscle mobility and this method can be used to help them express their feelings. Another aspect is this method's military use for the diagnosis of mental states such as distress and anxiety in the soldiers. On the other hand, it is possible to detect a liar, which is one of the most used one in the detection of emotions. Further, given that this research field is new, this research can be suitable for research and development in other fields related to this issue.

10–20 standard

In 1949, 10–20 electrode positioning method was known as an international standard that could cover almost all the areas of head by the electrodes. Electrodes are chosen based on particular parts of the skull. The electrodes are placed in the areas of intersection of the skull, and other central electrodes are arranged based on the 10% and 20% of the whole distance [Figure 1].[2]

Emotion recognition

Happiness, sorrow, fear, sadness, and more! All these words are derived from a concept called emotion. It can be said that this concept is in all creatures on earth. These feelings are stimulated and aroused by different situations. To stimulate and arouse emotions, it is needed to stimulate person by a stimulus so as to arouse feelings in the person. The stimulus can arouse emotions through all five senses of sight, taste, hearing, smell, and touch.

Emotion recognition can be done in different ways on different grounds. Emotion recognition can also be used
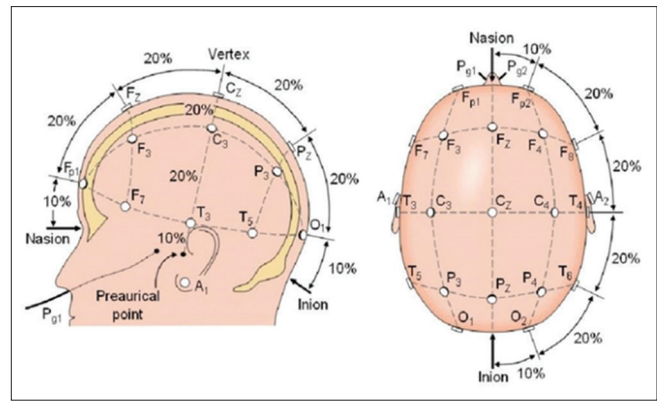

Figure 1: Electrodes positions in 10–20 international standards

in various areas and affect them. One of the applications of emotion recognition is helping those who are not able to mobilize and are paralysis, or it can be helpful in the treatment of depression. Emotion recognition, as noted, can be done in several ways. One approach is using brain signals for emotion recognition after recording brain signals. Given that all activities occurring in the body are processed in the brain, electrical brain signals that are created as a result of brain nerve cells' activities can be used for emotion recognition. The most important model for emotion recognition is arousal-valence model proposed by Russell.[1] This model is a 2D model that includes all the feelings. In addition, another model that was more comprehensive was proposed by those who proposed the 2D model, in which emotion control aspect was also added. This model is used as a reference model for emotion recognition either through brain signals or through other methods.

Arousal-valence model

The model was presented in an article published in 1980 by Russell.[1] In this model, all the feelings of people are placed on a page that has 2Ds of valence and arousal. This page is divided into four parts by these dimensions. People's different emotions are placed in these four areas and given the amount of valence and arousal of each emotion; the feeling is placed in a level of this model. Vertical dimension is arousal and horizontal dimension is valence [Figure 2].[1]

For example, since excitement is an inner sense with high valence and great arousal, it is located in the top right quadrant of the model. The opposite of excitement is depression which is located in the bottom left quadrant of the model. This model has been used in most of the researches carried out in the field of emotion recognition. It can be divided into various levels according to valence and arousal of different emotions. Given the complexities of emotion recognition, emotions are usually recognized in two or three levels in different studies. For instance, Hatamikia *et al.*[3] in 2014 recognized emotions in both two and three levels. There is another dimension which called
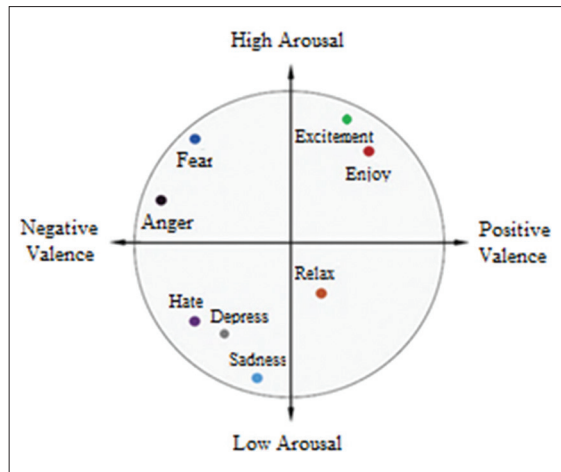
Figure 2: Two-dimensional arousal-valence model

dominance in this model too. This dimension indicates user's control over his/her emotions that affect performance and his/her manifestation in the brain signals. This model is used less than the 2D model.

Challenges

Despite the various studies conducted on emotion recognition, there are still major challenges. Among all challenges in this field, how to remove unwanted noise in the recorded signal signals is one of the important issues. To remove unwanted noises, some methods have been used in different studies, but it has been proven that using combined methods such as the combination of wavelet decomposition and ICA has superior performance compared to the native wavelet or ICA separately.[2] Combination of two methods such as empirical mode decomposition and canonical correlation analysis has a better performance in removing noises compared to using canonical correlation analysis solo. Moreover, combination of empirical mode decomposition and ICA has had better performance than ICA.[2]

The main challenge in these systems is improvement of performance, especially accuracy of emotion recognition. This is related to different categories and just one factor cannot be effective in improving the performance of emotion recognition in brain signals. Among the factors that contribute to this improvement, selection of the optimal feature as well as the selection of appropriate category to classify the feelings extracted from brain signals is more important to others. These two factors are somehow linked together and we can say that proper selection of features will improve the classification performance by sure.

Another factor that affects the accuracy of the results is the number of levels in the arousal-valence model of emotions. As was mentioned earlier, the reference model used for emotion recognition is 2D arousal-valence model proposed by Russell in 1980.[1] The number of levels studied in this model can affect the performance and accuracy. In

this regard, the study of Hatamikia *et al.*[3] in 2014 on DEAP datasets well represents this issue. The accuracy of recognition of two levels in the dimension of valence in their study was equal to 72.23 and was 74.20 in the dimension of arousal; however, when they were attempted to recognize at three levels, the accuracy values were reduced to 61.1 and 65.16, respectively.

The most important classification methods used

To recognize emotions through brain signals, there is a need to methods that can be used to identify and classify signals related to emotions arouse by stimuli and displayed in brain signals. Machine learning methods can be used to detect emotions in the brain signals. One of the most important elements in this task is the classifier. The most widely used classifiers in different studies are described in [Table 1].

Materials and Methods

Based on previous works in this fields such as Imani *et al*[2], it decided to use ICA for removing the noises from the raw EEG signals in this study. According to other studies conducted on emotion recognition that mentioned before and also employment of deep learning in different studies for sound detection, a four-step process is proposed. The four steps include:
1. Preprocessing
2. Extracting features
3. Dimension reduction
4. Emotion identification and estimation.

Each step also has several substeps. To increase accuracy of emotion recognition, it is necessary to carry out each of these steps and their substeps in a way that the highest possible accuracy achieved. To find the optimal choice in some of the steps of these layers, different tests have been conducted.

The first step, preprocessing, includes the following four substeps:
1. Noise reduction
2. Selecting the appropriate signals
3. Filtering for extracting the five principal bands of EEG signal
4. Selecting the appropriate bands to recognize emotions.

The second step also includes the following steps for desired feature extraction, i.e., linear frequency cepstral coefficient (LFCC). The steps also include:
1. Framing
2. Windowing
3. Fast Fourier transform (FFT)
4. Apply filter bank
5. Logarithm
6. Discrete cosine transform (DCT).

After this, step 3 starts, i.e., dimension reduction using Kernel Fisher discriminant analysis (KFDA) algorithm.

Table 1: Major studies on emotion recognition in electroencephalography signals

	Features	Number of emotions	Participants	Accuracy (%)	Dataset	Authors
SVM	High-order crossings	6	16	83.3	Online sampling with stimulus images	[17]
	PSD30, DAMS12, RAMS12 and PSD24	4	Not mentioned	82.29	16 pieces of music tracks produced by Oscar film festival	[7]
	PCA and PSD	Arousal and valence	32	Arousal 67 Valence 76	Some emotional videos from YouTube	[8]
	Magnitude mutual information and squared coherence estimate	Arousal and valence	26	Arousal 75 Valence 81	IAPS	[8]
	Quality dimensions, approximate entropy, and wavelet entropy	Not mentioned	15	73.25	IAPS	[9]
	High-order crossing and cross correlation	Not mentioned	16	Independent participants 65.8 Dependent participants 94.40	Online sampling	[6]
	Fourier transform log band energy	Not mentioned	Not mentioned	87.53	4-min movie parts prepared from Oscar film festival	[10]
	PSD	Arousal and valence	30	Arousal 52.4 Valence 57.7	30 subjects prepared by film and photos	[4]
	Chebyshev II IIR	Arousal and valence	Not mentioned	Arousal 82.03 Valence 65.39	IIa and IIb data of BCI Competition data set	[11]
	Minimum, maximum, mean, standard deviation	Not mentioned	20	77.75	IAPS	[12]
	Fractal dimension, high-order crossing and 6 statistical features	8 4	32	4-electrod 53.7 4-electrod 87.02	IADS, IAPS, and DEAP	[13]
	Differential entropy of each 5 alpha, beta, gamma, delta and theta bands from 62 channels	Not mentioned	6	84.08	Displaying a series of pieces from famous films	[14]
	PSD and Fisher Kernel - PCA and LDA - SVM with imbalanced Quasiconformal kernel and imbalanced SVM	Arousal and valence	100 college students	Arousal 84.79 Valence 82.68	IAPS	[5]
	fractal dimension, high-order crossing, inter-class correlation coefficient, 6 statistical features	2 4	100 college students	71.75 49.63	IADS	[15]
	Hjorth parameters, i.e., activity, mobility and complexity for each 5 bands of five lobes of the brain	2 5	100 college students	67 34	IAPS	[16]
K-NN	High-order crossing	6	16	73.94	Online Sampling with stimulus images	[17]
	Magnitude mutual information, squared coherence estimate	4 emotions in arousal and valence	26	Arousal 79 Valence 82	IAPS database and unique pieces of music from bernard bouchard	[18]
	PSD, combined Gaussian model of EEG spectrums	Arousal and valence model	Not mentioned	Arousal Minimum=55.07, Maximum=67.0 Valence Minimum=58.8 Maximum=76	Some specific clips for emotion stimulation and signal recording	[19]
	Chebyshev II IIR	Arousal and valence	Not mentioned	Arousal 82.25 Valence 66.51	IIa and IIb data of BCI competition data set	[11]
	Minimum, maximum, mean, and standard deviation and ICA for noise removal	Not mentioned	20	70.37	IAPS	[12]

Contd...

	Features	Number of emotions	Participants	Accuracy (%)	Dataset	Authors
	Auto regression and KNN	2 classes arousal and valence	32	Arousal 74.2 Valence 72.33	DEAP	[3]
		3-classes arousal and valence		Arousal 65.156 Valence 61.1		
	Differential entropy of each 5 bands obtained 62 channels	Negative and positive poles	6	69.22	Some clips of famous films	[14]
Artificial neural networks	PSD30, RAMS12, DAMS12, PSD24	Not mentioned	26	51.52	16 pieces of music tracks produced by Oscar Film Festival	[7]
	Differential entropy of each 5 bands obtained 62 channels	Negative and positive poles	6	DBN 86.91 DBN-HMM 87.62	Some clips of famous films	[14]
	PSD of 32 EEG signal channels and PCA	Without compliance in arousal and valence	32	Arousal 74.2 Valence 72.33	IADS, IAPS	[20]
		Compliance based in arousal and valence		Arousal 65.156 Valence 61.1		
LDA	Frequency of EEG signal bands, brain asymmetry and dependence	Not mentioned	110	Frequency 38.8 Asymmetry 57.9 Dependency 55.3	Some Movies	[21]
	Auto regression, SFS, and Davies-Bouldin index	2 classes arousal and valence	32	Arousal 65.54 Valence 63.22	DEAP	[3]
		3-classes arousal and valence		Arousal 52.36 Valence 51.20		
QDA	High-order crossing	6	16	62.03	Online Sampling with stimulus images	[17]
	Auto regression, SFS, and Davies-Bouldin index	2 classes arousal and valence	32	Arousal 69.26 Valence 70.35	DEAP	[3]
		3-classes arousal and valence		Arousal 57.42 Valence 57.18		
Naïve Bays	Features extracted from IIR, i.e., CSP, ASP, and AF features and fisher linear discriminant for dimension reduction	Arousal and valence	Not mentioned	Arousal 66.24 Valence 83.10	IIa and IIb data of BCI Competition data	[11]
	Minimum, maximum, mean, and standard deviation and ICA for noise removal	Not mentioned	20	59.26	IAPS	[12]
	Fast Fourier transform analysis for feature extraction and features based on Pearson correlation coefficients	2 Classes arousal and valence	Not mentioned	Arousal 70.1 Valence 70.9	IAPS, IADS	[22]
		3 Classes arousal and valence		Arousal 55.2 Valence 55.4		
	Mean, standard deviation and linear fisher discriminant	Arousal and valence	32	Arousal 59 Valence 57	Some emotional videos from YouTube	[8]

SVM – Support vector machine; PSD – Power spectral density; PCA – Principal component analysis; ECG – Electroencephalography; IAPS – International affective picture system; IADS – International Affective Digitized Sounds; IIR – Infinite impulse response; BCI – Brain-computer interface; DEAP – Dataset for emotion analysis using EEG, physiological and video signals; K-NN – K-nearest neighbor; LDA – Linear discriminant analyzer; SFS – Sequential forward selection; ICA – Independent-component analysis; AF – Atrial fibrillation; QDA – Quadratic discriminant analysis; ASP – Asymmetric spatial pattern; CSP – Common spatial pattern; DBN-HMM – Deep belief network – Hidden markov model

Finally, in step 4, emotion detection and estimation, after testing both deep neural network and deep belief neural network (DBN), attempted to estimate and recognize emotions through EEG signals. Then, recognition accuracy of each deep learning methods is calculated using mean square error (MSE).

It should be noted that this study has been carried out at two two-class and three-class levels of Russell reference model.[1] In previous studies, which were described in detail, accuracy of three-class recognition was lower than two-class recognition, and it is one of the challenges of emotion recognition. In this study, both levels were tested by the proposed method [Figure 3].

DEAP dataset

This dataset prepared by Koelstra[8] in collaboration with four universities of EPFL, Genève, Twente, and Queens Marry for Emotion Analysis using EEG, physiological, and video signals. DEAP contains some video clips to stimulate emotions and recorded brain signals of the experimented individuals. It was the first standard dataset which attempted to incite people by video clips and induces emotions in the brain signals. The data has been prepared from 32 participants during the 40-min movie watching. Face of the 22 out of 32 participants was recorded as videos. This dataset is publicly available via a dedicated website at Queens Marry datasets.

Testing the best window for extracting linear frequency cepstral coefficient feature

In the step of feature extraction from DEAP dataset, i.e., LFCC feature extraction, there are some steps there, one of them called windowing.

In this part, four known and most widely used windows were tested to find the best window. The four windows are:
1. Hamming
2. Hanning
3. Black-Man
4. Rectangle.

The process of testing the best window is such that for each window, for all the individuals and tests conducted in DEAP dataset, LFCC feature extraction is done based on that window, and then, 3Ds of emotion, i.e., arousal, valence, and dominance, were estimated. The input data were first preprocessed to remove their noise. Then, 32 EEG channels were selected and applied filter to extract the main five EEG bands.

After that, two bands appropriate for emotion recognition were selected according to reference.[7] Then, feature extraction starts. At this point, the data are divided into 1-s frames, and on every frame, one window of above-mentioned four windows is applied every time. According to the research of Imani *et al*.,[2] window size was considered 512. After that, FFT is taken from results obtained and 21 filter banks with 50% overlapping are applied on it. In this experiment, the number of filter banks is fixed and considered to 21 to simplify calculations and
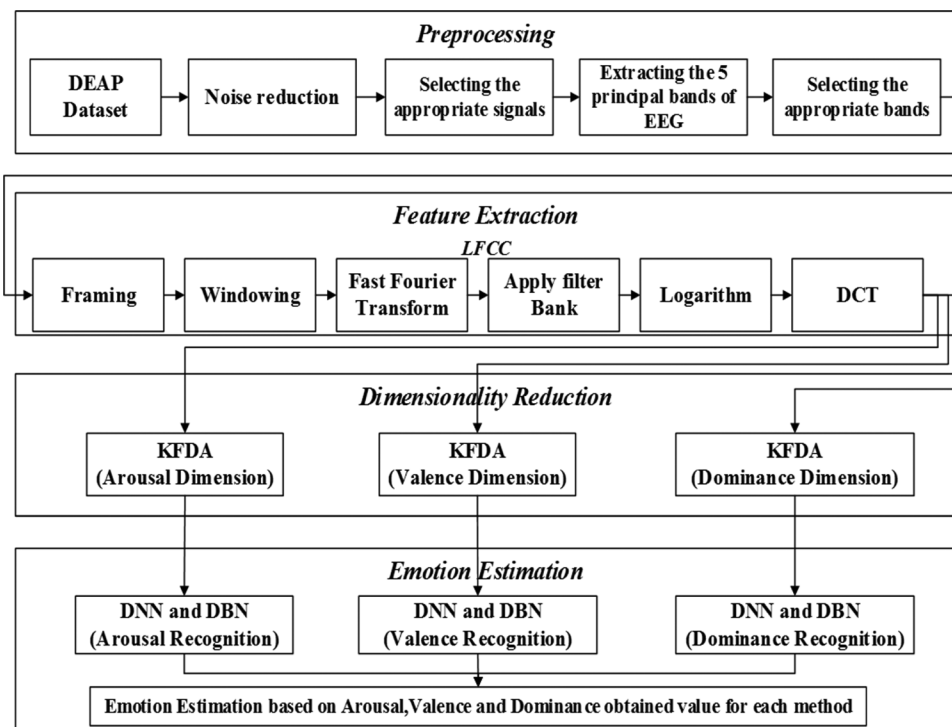

Figure 3: Proposed method for emotion recognition

carry out experiments in the same environment. Here, according to the range of 8–30 Hz, which is related to the alpha and beta bands, by applying a triangular filter with size 2 and an overlay of 50%, there are exactly 21 filters in this range.

After extracting the data of this step, logarithm and DCT is calculated for them. After that, according to paper,[5] KFDA with Gaussian Kernel is used for dimension reduction and emotion estimation for three dimensions of arousal, valence, and dominance is done. It should be noted that to obtain better features with fewer dependency to the other dimensions; dimension reduction is conducted for three times.

Three DBNs have been used for estimation so that each of them estimates one dimension and give better results as much as possible.

For selecting data for estimation, cross-validation method is used in this study. In this way, 25% of data are allocated for testing, 25% for validating, and the remaining 50% for training. Data selection process is done randomly.

After applying these settings, DBNs are trained and then tested. Finally, their results were compared based on different windows. This comparison was done as person-to-person and also as average of all patients. The accuracy of the results is calculated using the MSE. After that, considering the results of the comparison, the best window to estimate emotions is selected. It should be noted that this experiment was conducted in a two-class level.

The mean results of experiment on four tested windows are shown in Table 2.

Testing the number of filter banks

In this experiment, according to previous research on the use of filter banks whether with the use of Mel-frequency cepstral coefficients or other methods such as LFCC, a range of 10-number widely used filter banks are selected to choose the most appropriate number for emotion recognition and estimation. The range is from 20 to 29.

The process of testing the most appropriate number of filter banks for extracting features to recognize emotions in brain signals is the same as testing the most appropriate window with the exception that the window in this experiment is hamming and filter banks presented in Table 3 are tested for all the experiments of individuals in DEAP dataset. This means that after framing and applying Hamming on data and taking FFT, a different number of filter bank is applied based on Table 2 and then their logarithm and DCT are calculated. The experiment is done for all elements of the row of filter bank in Table 3.

Similar to window testing, KFDA with Gaussian Kernel is used for dimension reduction for all 3Ds, three times, and three DBNs are used for emotion recognition. The structure of neural networks used in this test is similar to the test

of the most appropriate window. Finally, the accuracy of the results obtained using MSE is calculated and results of each filter bank are compared to other results. It should be noted that experiment of this section is also two-class.

Results and figures related to mean recognition accuracy

According to tests performed by different studies, the mean recognition accuracy using different filter banks is required to make a decision to choose the most appropriate number of filter banks to increasing recognition rate. The mean results of recognition accuracy based on different number of filter banks are shown in Table 4.

Test of emotion recognition in brain signals

At this step, after setting the previous steps for emotion recognition, the test of the most appropriate window and the appropriate number of filter banks with the results of Kaiser with the parameters of Hamming window and 26 filter banks, deep neural networks and DBN were

Table 2: Average of accuracy for suitable window test

	Hamming	Hanning	Black-Man	Rectangle
Arousal	92.79	89.732	88.29	70.71
Valence	91.02	90.006	85.11	70.02
Dominance	92.66	88.45	85.25	71.45

Table 3: Tests numbers and correspond number of filter banks

Tests number	Number of filter banks
1	20
2	21
3	22
4	23
5	24
6	25
7	26
8	27
9	28
10	29

Table 4: Mean recognition accuracy based on the number of different filter banks

	Arousal	Valence	Dominance
The number of used filter banks			
20	91.975	90.347	91.51
21	92.212	91.767	92.596
22	91.83	90.884	92.334
23	91.514	91.695	91.147
24	91.522	91.967	91.469
25	92.974	91.399	92.369
26	93.925	92.644	93.142
27	92.089	91.926	92.284
28	92.546	90.802	91.02
29	92.125	91.641	90.753

tested to recognize emotions. The test is done in two two-class and three-class levels that cover two levels of the reference model of Russell and three levels of it. The process and results of these tests at two levels are as follows.

Test conducted by deep neural network

In this test, LFCC features, extracted from the proposed method based on Kaiser Window with parameters of Hamming and 26 filter banks and after dimension reduction by KFDA algorithm with Gaussian Kernel, were entered to the deep neural network with the following structure to estimate and recognize emotions.

As seen in Figure 4, the network is composed of layers including an input layer with 32 neurons, three hidden layer with 70, 70, and 200 neurons, respectively, and an output layer with one neuron. This type of network is fitnet and its learning function is selected as trainscg.

To obtain appropriate results of each emotion dimensions, obtained LFCC features based on the dimension are trained, validated, and tested each time so as to obtain better accuracy for each dimension.

Dividing of input data or LFCC features extracted from data is done using cross-validation. By this method, 50% of extracted LFCC features are used for training, 25% for validation, and 25% for testing.

Finally, after receiving outputs of mentioned deep neural network, detection accuracy is measured using the MSE.

Test conducted by deep belief neural network

Like previous test, LFCC features, extracted from the proposed method based on Kaiser Window with parameters of Hamming and 26 filter banks and after dimension reduction by KFDA algorithm with Gaussian Kernel, were entered to the DBN with the following structure to estimate and recognize emotions.

As seen in Figure 5, the network is composed of three restricted boltzmann machine (RBMs) of size 50, 50, and 100.

To obtain appropriate results of each emotion dimension, obtained LFCC features based on the dimension are trained, validated, and tested each time so as to obtain better accuracy for each dimension.

Dividing of input data or LFCC features extracted from data is done using cross-validation. By this method, 50% of extracted LFCC features are used for training, 25% for validation, and 25% for testing.

For carrying out this test, three DBNs have been used and the structure of each of them is as follows:
1. All three DBNs used are function approximation
2. Sampling method for all of them is Free Energy in Persistent Contrastive Divergence
3. They are composed of three RBMs
4. The network units are probability type with actual quantities between [0, 1]
5. The first RBM is 50
6. The second RBM is 50
7. The third RBM is 100
8. All RBMs are of generative type
9. The maximum number of steps taken in the training process executed on the entire training data for each three RBMs is considered to be 50
10. DeebNet toolbox has been used for DBNs implementation.[23]

Finally, after receiving outputs of mentioned DBN, detection accuracy is measured using the mean square error.

Results

The results of this experiment are as follows:

Results of emotion recognition test using deep neural network

Based on the details given in previous sections, the results of conducting emotion recognition test at two-class level in Russell's model using the proposed method by deep neural network estimator is shown in Table 5.
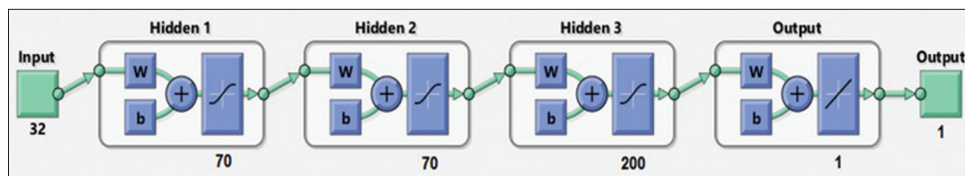

Figure 4: The structure of used deep neural network in emotion recognition tests
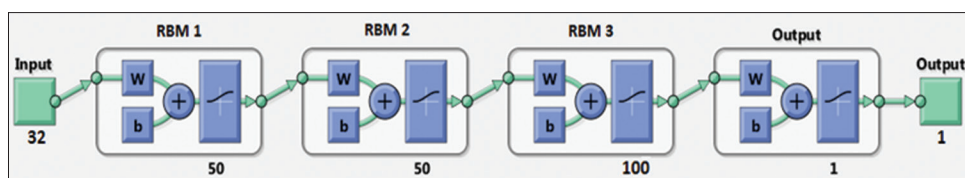

Figure 5: The structure of used deep belief neural network in emotion recognition tests

The results of conducting emotion recognition test at three-class level in Russel's model using the proposed method by deep neural network estimator is shown in Table 6.

Results of emotion recognition test using deep belief neural network

The results of conducting emotion recognition test at two-class level in Russell's model using the proposed method by DBN estimator is shown in Table 7.

The results of conducting emotion recognition test at three-class level in Russell's model using the proposed method by DBN estimator is shown in Table 8.

Finally, the recognition accuracy obtained from experiments conducted with the proposed method compared to most important studies using the same dataset is shown in Table 9.

Table 5: The results of conducting emotion recognition test at two-class level using the method proposed by deep neural network

Arousal	Valence	Dominance
81.5839	79.8750	80.3595

Table 6: The results of conducting emotion recognition test at three-class level using the method proposed by deep neural network

Arousal	Valence	Dominance
68.54	66.31	66.92

Table 7: The results of conducting emotion recognition test at two-class level using the method proposed by deep belief neural network

Arousal	Valence	Dominance
93.9248	92.6444	93.1416

Table 8: The results of conducting emotion recognition test at three-class level using the method proposed by deep belief neural network

Arousal	Valence	Dominance
76.28	74.83	75.64

Conclusion

The aim of this study was to improve emotion recognition in EEG signals through deep learning. In this regard, as stated earlier, a four-step method was proposed. Among these four steps, there are some steps which needed further examination for more accurate selection to improve emotion recognition. One of these steps was feature extraction which LFCC features were extracted. Another one was windowing. According to tests conducted to determine the best window, among four windows, Hamming, Hanning, Black-Man, and Rectangle, Hamming was the best window for this section. According to the results of this experiment, Hamming window resulted in more accuracy in testing for the most appropriate two-class window. Mean accuracy achieved for this window in the dimension of arousal was 92.79, 91.02 in valence, and 92.66 in dominance. Likewise, for the Hanning window, these values were equal to 89.73, 90.006, and 88.45, respectively. Further, for Black-Man window, the stated values were 88.29 for arousal, 85.11 for valence, and 85.25 for dominance. Finally, for the fourth window, Rectangle, mean accuracy of the arousal was 70.71, 70.02 for valence, and 71.45 for dominance.

In another part of the feature extraction procedure, it was needed to determine how many filter banks are required for emotion recognition in EEG signals. Therefore, a test was designed to evaluate different values of the filter banks. Based on the results of the test described in detail in the previous sections, the most appropriate number of filter banks in LFCC procedure for feature extraction was 26. The mean results of the number of filter banks in two-class test were 93.92, 92.64 and 93.14 for arousal, valence, and dominance, respectively. After that, the second best choice for the number of filter banks was 25. Mean accuracy obtained for this number is 92.97, 91.39, and 92.36 for arousal, valence, and dominance, respectively. According to the result of the tests, the window which used for LFCC was Hamming.

Finally, to carry out the final test, i.e., emotion recognition in EEG brain signals using deep learning, two classification methods of deep neural network with five layers and DBN composed of three RBMs have been used.

Table 9: Compared table of obtained result in two-class and three-class with other outstanding studies

Level	Dimension	Hatamikia *et al.*[3]	Yoon and Chung[22]	Koelstra[8]	Jirayucharoensak *et al.*[20]	Proposed method by DNN (Me-DNN)	Proposed method by DBN (Me-DBN)
2-class	Arousal	74.2	70.1	-	-	81.58	93.9248
3-class		65.16	55.2	62	52.03	68.54	76.28
2-class	Valence	72.33	70.9	-	-	79.87	92.6444
3-class		61.1	55.4	57.6	53.42	66.31	74.83
2-class	Other dimensions	-	-	-	-	80.35	93.1416
3-class		-	-	55.4	-	66.92	75.64
	Details	-	-	Interest	±9.74	Dominance	Dominance

DBN – Deep belief neural network; DNN – Deep neural network

The mean results of emotion recognition test accuracy using deep neural network was 81.85, 79.87, and 80.35 for arousal, valence, and dominance in two-class experiment. The accuracy values in three-class experiment were 68.54, 66.31, and 66.92 for arousal, valence, and dominance. The results for DBN in two-class experiment were 93.92, 92.64, and 93.14 for arousal, valence, and dominance and in three-class experiment were 76.28, 74.83, and 75.64, respectively.

As can be seen, the results of this method compared with previous works, both at two-class and three-class levels, show better accuracy. These results show that the proposed method could solve this problem, with higher accuracy.

In the proposed method using deep neural network, the accuracy value obtained had an improvement of 7.38% in two-class level and 3.38% in three-class level for arousal compared to the best accuracy obtained so far, i.e., the study of Hatamikia *et al*.[3] This improvement is 7.54% and 5.21% for valence.

In the proposed method using DBN, the accuracy value obtained had an improvement of 12.34% in two-class level and 7.74% in three-class level for arousal. This improvement was 12.77 and 8.52% for valence.

Generally, the accuracy improvement can be the result of change in three different parts employed in this experiment compared to previous studies. These changes are:

1. Using all signals without downsampling and noise removal that although increased computation time, affected accuracy, and increased it to the same extent
2. Determining the most appropriate selections in feature extraction section that resulted in increased accuracy compared to other choices
3. Selecting classification that has had very good results in sound recognition section.

Financial support and sponsorship

None.

Conflicts of interest

There are no conflicts of interest.

References

1. Russell JA. A circumplex model of affect. J Pers Soc Psychol 1980;39:1161-78.
2. Imani E, Pourmohammad A, Bagheri M, Mobasheri V. ICA-based imagined conceptual words classification on EEG signals. J Med Signals Sens 2017;7:130-44.
3. Hatamikia S, Maghooli K, Nasrabadi AM. The emotion recognition system based on autoregressive model and sequential forward feature selection of electroencephalogram signals. J Med Signals Sens 2014;4:194-201.
4. Soleymani M, Lichtenauer J, Pun T, Pantic M. A multimodal database for affect recognition and implicit tagging. IEEE Trans Affect Comput 2012;3:42-55.
5. Liu YH, Wu CT, Cheng WT, Hsiao YT, Chen PM, Teng JT, *et al*. Emotion recognition from single-trial EEG based on kernel Fisher's emotion pattern and imbalanced quasiconformal kernel support vector machine. Sensors (Basel) 2014;14:13361-88.
6. Petrantonakis PC, Hadjileontiadis LJ. A novel emotion elicitation index using frontal brain asymmetry for enhanced EEG-based emotion recognition. IEEE Trans Inf Technol Biomed 2011;15:737-46.
7. Lin YP, Wang CH, Wu TL, Jeng SK, Chen JH. "EEG-based emotion recognition in music listening: A comparison of schemes for multiclass support vector machine," 2009 IEEE International Conference on Acoustics, Speech and Signal Processing 2009.
8. Koelstra S. Deap: A database for emotion analysis; using physiological signals. Affect Comput 2012;3:18-31.
9. Hosseini SA, Naghibi-Sistani MB. Emotion recognition method using entropy analysis of EEG signals. Int J Image Graph Signal Process 2011;5:30-6.
10. Nie D, Wang XW, Shi LC, Lu BL. EEG-Based Emotion Recognition During Watching Movies. 2011 5th International IEEE/EMBS Conference on Neural Engineering (NER); 2011. p. 667-70.
11. Huang D, Guan C, Ang KK, Zhang H, Pan Y. Asymmetric Spatial Pattern for EEG-based emotion detection. Proceeding International Joint Conference on Neural Networks; 2012. p. 10-5.
12. Sohaib AT, Qureshi S, Hagelbäck J, Hilborn O, Jerčić P. Evaluating classifiers for emotion recognition using EEG. Lect Notes Comput Sci 2013;8027:492-501.
13. Liu Y, Sourina O. EEG Databases for Emotion Recognition. Proceeding – 2013 International Conference on Cyberworlds (CW 2013); 2013.p. 302-9.
14. Zheng WL, Zhu JY, Peng Y, Lu BL. EEG-based emotion classification using deep belief networks. In 2014 IEEE International Conference on Multimedia and Expo (ICME); 2014. p. 1-6.
15. Lan Z, Sourina O, Wang L, Liu Y. Real-time eeg-based emotion monitoring using stable features, The Visual Computer, 2016;32:347-58.
16. Mehmood RM, Lee HJ. EEG based emotion recognition from human brain using Hjorth parameters and SVM. Int J Bio Sci Bio Technol 2015;7:23-32.
17. Petrantonakis PC, Hadjileontiadis LJ. Emotion recognition from EEG using higher order crossings. IEEE Trans Inf Technol Biomed 2010;14:186-97.
18. Khosrowabadi R, Heijnen M, Wahab A, Quek HC. The dynamic emotion recognition system based on functional connectivity of brain regions. Emotion 2010;639798:377-81.
19. Khosrowabadi R, Rahman AW. Classification of EEG Correlates on Emotion using Features from Gaussian Mixtures of EEG Spectrogram. Proceeding 3rd International Conference on Information and Communication Technologies. Moslem World; 2010. p. E102-7.
20. Jirayucharoensak S, Pan-Ngum S, Israsena P. EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation. ScientificWorldJournal 2014;2014:627892.
21. Park MS, Oh HS, Jeong H, Sohn JH. Eeg-Based Emotion Recogntion during Emotionally Evocative Films. In: 2013 International Winter Workshop on Brain-Computer Interface (BCI); 2013. p. 56-7.
22. Yoon HJ, Chung SY. EEG-based emotion estimation using Bayesian weighted-log-posterior function and perceptron convergence algorithm. Comput Biol Med 2013;43:2230-7.
23. Keyvanrad MA, Homayounpour MM. A Brief Survey on Deep Belief Networks and Introducing a New Object Oriented MATLAB Toolbox (DeeBNet), Preprint, arXiv: 1408.3264; 2014.

BIOGRAPHIES

Seyed Mohammad Reza Mousavinasr has received his BSc in Software Engineering from Mazandaran University of Science and Technology, Mazandaran, Iran in 2011. He also received his MSc from Malek-Ashtar University of Technology in 2015 in artificial intelligence. Research interest of him are EEG Signal processing, Machine Learning Information Security and Deep Leaning.

Email: Reza.Mousavinasr@gmail.com

Ali Pourmohammad was born in Azerbaijan (Tabriz City), IRAN. He has a Ph.D. in Electrical & Electronics Engineering (Applied Signal Processing) from the Electrical Engineering Department, Amirkabir University of Technology, Tehran, IRAN, Jan. 2012. He is an invited lecturer at this department now and teaches several courses as C++ Programming, Multimedia Systems, Microprocessor Systems, and Digital Signal Processing. His research interests include digital signal processing and applications (audio & speech, image & video, and biomedical & brain signals processing and applications).

Email: pourmohammad@aut.ac.ir

Mohammad Sadegh Moayed Saffari has received his BSc in Software Engineering from University of Science and Culture, Tehran, Iran in 2011. He also received his MSc from Malek-Ashtar University of Technology in 2015 in artificial intelligence. His research interests are Signal processing, Machine Learning Information Security and Deep Leaning.

Email: ms.moayedsaffari@gmail.com