Scientia Iranica (2025) **32**(1): 6664

Sharif University of Technology

Scientia Iranica

Transactions on Computer Science & Engineering and Electrical Engineering

https://scientiairanica.sharif.edu

Temporal analysis of functional brain connectivity for EEG-based emotion recognition

Ensieh Khazaei, Hoda Mohammadzade*

Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran.

Received 2 April 2022; received in revised form 17 May 2023; accepted 21 February 2024

KEYWORDS

Emotion recognition;
Functional brain connectivity;
EEG signals;
Temporal analysis;
RNNs;
Transformers.

Abstract. Electroencephalogram (EEG) signals in emotion recognition absorb special attention owing to their high temporal resolution and their information about brain activity. Different brain areas work together and the activity of brain changes over time. In this study, we investigate the emotion classification performance using functional connectivity features in different frequency bands and compare them with the classification performance using Differential Entropy (DE) feature, which has been previously used for this task. Moreover, we investigate the effect of different time periods on classification performance. Our results on SJTU Emotion EEG Dataset (SEED) dataset show that as time goes on, emotions become more stable, and the classification accuracy increases. Among different time periods, we achieve the highest classification accuracy using the time period of 140 s-end. In this time period, the accuracy is improved by 4 to 6% compared to the entire signal. Pearson correlation coefficient, coherence, and phase locking value features and Support Vector Machine (SVM) obtain the mean accuracy of about 88%. Using the proposed framework, functional connectivity features lead to better classification accuracy than DE features (with the mean accuracy of 84.89%). Finally, using the best time interval and SVM, we achieve better accuracy than using Recurrent Neural Networks (RNNs) which need large amounts of data and have high computational costs.

1. Introduction

Emotions play an important role in human decisions. Emotion recognition has numerous applications in the enhancement of human life and human ability [1]. This interdisciplinary field plays a critical role in the development of psychology, neuroscience, cognitive science, and computer science [2,3]. In the field of computer science, automatic emotion recognition is used to improve the human-computer interface [4,5]. Other applications of emotion recognition include lie detection, behavior prediction, and health monitoring. Among these various applications, the human-computer interface is particularly important [3,6]. In this application, machines get the ability to understand the emotional states of people and so can interact better with them.

There are different approaches to emotion recognition which can be divided into two categories. The first category is based on non-physiological signals such as facial expressions, body movements, and intonation [3–7]. The second category is based on physiological signals such as electroencephalogram

*. *Corresponding author. Tel.: +98 21 66165927*
E-mail addresses: ensieh.khazaei@ee.sharif.edu (E. Khazaei); hoda@sharif.edu (H. Mohammadzade)

To cite this article:

E. Khazaei and H. Mohammadzade "Temporal analysis of functional brain connectivity for EEG-based emotion recognition", *Scientia Iranica* (2025), **32**(1): 6664 https://doi.org/10.24200/sci.2024.60201.6664

(EEG), electrocardiography, heart rate, and respiration signals [3,4]. Physiological signals provide more comprehensive and complex information and their results are more accurate [3,4]. There are a lot of physiological signals and brain imaging techniques such as functional Magnetic Resonance Imaging (fMRI) [8], Positron Emission Tomography (PET), Magnetoencephalography (MEG), and EEG [9] that express brain states, and they are used in emotion recognition [10–12]. Among these methods, although fMRI and PET have a high spatial resolution, their temporal resolution is not as well as EEG and MEG signals. On the other hand, MEG signal has a high temporal resolution as well as EEG signals, however, MEG instrumentation represents very high technology like expensive superconducting technology and heavy magnetic shielding. Therefore, MEG instrumentation is more expensive than EEG instrumentation with the same number of channels. In addition, MEG recording restricts the movements of the head and causes static tension in the muscles [13]. Therefore, EEG signals receive more attention because they have a high temporal resolution, which is important for our research with a focus on the temporal analysis of brain activities in different emotional states, besides they are cheap and easy to record [6].

There are different models for expressing human emotions which can be generally divided into two categories: discrete basic emotions and continuous emotions. In the discrete model, emotions are classified into a set of discrete labels including six basic emotions: happiness, sadness, surprise, fear, anger, and disgust [14]. In the continuous model, emotions are expressed using two dimensions: valence and arousal [15]. Valence indicates how much emotion is positive or negative, and arousal indicates how much a person is excited or indifferent [6,16].

Up to now, a wide range of research has been done in the field of EEG-based emotion recognition. However, most recent research in the field of emotion recognition has focused on finding the strong connections, the important frequency bands, and the important electrodes. Unfortunately, little research has been done to examine the significant time periods and the role of functional brain connectivity features in emotion classification which are the focus of this study. In this paper, we look for the optimal time period for emotion classification. Moreover, we examine the performance of Recurrent Neural Networks (RNNs) to find temporal patterns in emotion classification.

The structure of this article is as follows. Section 2 gives a brief overview of related research on emotion recognition using EEG signals. Section 3 describes the dataset, preprocessing methods, feature extraction, and classification procedure. The results of RNNs, Transformer, and the optimal time period are given in Section 4. In Section 5, we discuss our results and compare them against previous work. In Section 6, the conclusion is presented.

2. Related work

In recent years, due to the development of dry electrode techniques [17–19], EEG signals, possessing high temporal resolution, are frequently used for emotion recognition. In this section, we review some of the previous research in the field of emotion recognition using EEG signals. In [2], Differential Entropy (DE), Power Spectral Density (PSD), differential asymmetry (DASM), rational asymmetry (RASM), and differential causality (DCAU) in the five frequency bands, which are delta (1-3 Hz), theta (4-7 Hz), alpha (8-13 Hz), beta (14-30 Hz) and gamma (31-50 Hz), are extracted from the public database [20] and then emotion classification is performed using deep SJTU Emotion EEG Dataset (SEED) neural networks. The results of this paper show that beta and gamma perform better than other frequency bands or in other words, there exist specific neural patterns in high-frequency bands for positive, neutral, and negative emotions through time-frequency analysis. This paper also concludes that utilizing the electrodes that are located in the temporal area (FT7, FT8, T7, T8, C5, C6, TP7, TP8, CP5, CP6, P7, and P8) can increase classification accuracy by 2.66% compared to using all electrodes. In [3], by examining the DE and energy spectrum features extracted from their collected dataset, the authors infer that there is a close relationship between the emotional state of individuals and information of the gamma band. In [5], a number of functional brain connectivity features like Phase Locking Value (PLV), Pearson correlation coefficient, and Phase Lag Index (PLI) [21] are extracted and then the resulting connectivity matrices are fed to neural networks using different methods for the spatial arrangement of electrodes. The purpose of this paper is to investigate the effect of spatial information in emotion classification. In [5], the arrangement of electrodes in the connectivity matrix is investigated and it is discovered to be effective in classification accuracy. In [22], the graph of PLV is obtained and then four features are extracted from the graph. The authors then examine the simultaneous use of four graph features and local features such as DE and PSD.

They observed that adding graph features to local features leads to an increase in classification accuracy. Furthermore, this paper analyzes different frequency bands from which the PLV graph features are extracted, showing that utilizing the gamma and beta bands results in better performance compared to other frequency bands. In [23], the effect of familiar stimuli on emotion classification is investigated. The stimulus was a selected collection of music, and participants

had an education in music. Each participant listened to 8 familiar music and 8 unfamiliar music. Finally, results show that the classification accuracy is higher for unfamiliar music.

In [24], the stable patterns of EEG signals over time in emotion recognition are studied. The results are shown as follows:

(a) The lateral temporal areas are more active in positive emotions than negative ones in beta and gamma bands;

(b) Brain activity in a neutral state has a greater alpha response in the parietal and occipital regions;

(c) Negative emotions are more active in the delta band in the parietal and occipital regions;

(d) Activity of the gamma band in the prefrontal region is more in negative emotions.

Li et al.[11] worked on EEG-based emotion recognition based on multi-task learning with a capsule network (CapsNet). Employment of Multi-task learning obtains more data from different tasks of high/low arousal, high/low valence, and high/low dominance classifications, and results in improvement of generalization and robustness. Furthermore, in [11], an attention mechanism is used to capture the intrinsic relationship between different channels of raw EEG signals. Finally, the attention mechanism adapts the weight of different channels to extract important information from raw EEG signals. By simultaneous use of multi-task learning and attention mechanism, they achieve higher accuracy in arousal, valence, and dominance classifications for a Dataset for Emotion Analysis using Physiological Signal (DEAP) [25] and DREAMER [26] datasets. In [12], a Transformer neural architecture search based on a multi-objective evolution algorithm is proposed to find out the best MLP model for EEG-based emotion recognition. Their end-to-end deep learning method is directly applied to the raw EEG signals instead of their extracted features. Their work focuses on three binary classification tasks: high/low arousal, high/low valence, and high/low dominance.

Although a lot of works have been done on emotion recognition using EEG signals, a small number of them studied functional brain connectivity and the importance of different time periods during the stimuli which are the focus of this paper. The framework of our proposed approach is depicted in Figure 1. First of all, we processed raw EEG signals and extracted some features from the preprocessed signal. Then a feature selection method was employed to decrease the number of features and select the most informative features. Finally, we did a temporal analysis and recognized the emotion using the obtained information in the temporal analysis step. All of these steps will be explained in detail in Section 3.

3. Methods

3.1. Data acquisition

We use the public SEED database [20] in this study. This database was recorded from 15 participants (7 males and 8 females, age range: 19-28 years) with self-reported normal or corrected-to-normal vision and normal hearing. All participants were selected using the Eysenck Personality Questionnaire (EPQ). The EPQ is a questionnaire to assess the personality traits of a person devised by Eysenck et al. [27]. This questionnaire conceptualizes each personality based on three independent dimensions of temperament: Extraversion/Introversion, Neuroticism/Stability, and Psychoticism/Socialization. The subjects who are extraverted and have stable moods tend to elicit the right emotions throughout the emotion experiments. Therefore, from the feedback of the EPQ questionnaires, fifteen subjects are chosen to participate in the emotion experiments [2].

The stimuli in this database are 15 Chinese clips that create three emotions positive, negative, and neutral in participants; there are five clips for each emotion. The criteria for video selection were as follows:

1. The clips should not be too long to tire the participants;
2. The meaning of the videos can be understood without any explanation;
3. All the moments of the video convey only one emotion to the subject.

The videos are approximately 4 minutes, and they presented in such a way that two clips that target the same emotion are not shown consecutively. The EEG

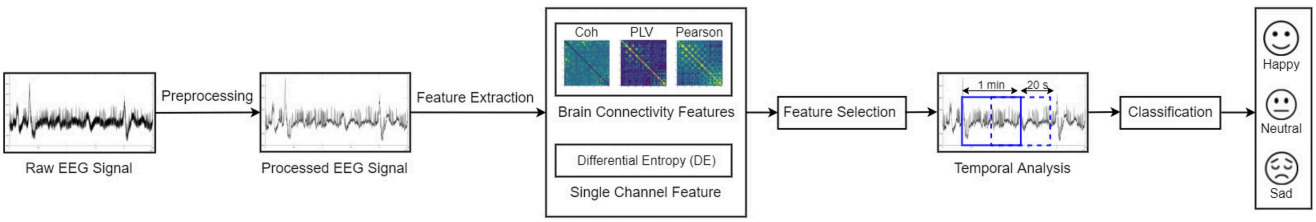

Figure 1. The framework of our proposed approach for emotion recognition using raw EEG signals.

signal of the participants was recorded for the whole duration that a clip is shown to the participant using 62 electrodes placed on the head according to the 10-20 system [24]. Electrooculogram (EOG) signal was recorded to remove the eye movement artifacts [2]. The location of the electrodes is shown in Figure 2.

The sampling rate of the EEG recording was 1000 Hz. The signals were recorded in three sessions, where the time interval between two consecutive sessions was at least 7 days. The stimuli were the same in all three sessions; three sessions were recorded from each subject, and in each session, the participant watched all 15 video clips. Figure 3 shows a detailed protocol for each session.

The participants were alerted 5 s before watching the video. After watching each video, an assessment form was given to the participant to express their emotions in 45 s. 15 s after the assessment was intended for rest [24]. Figure 3 shows the protocol of EEG recording in each session for SEED database.

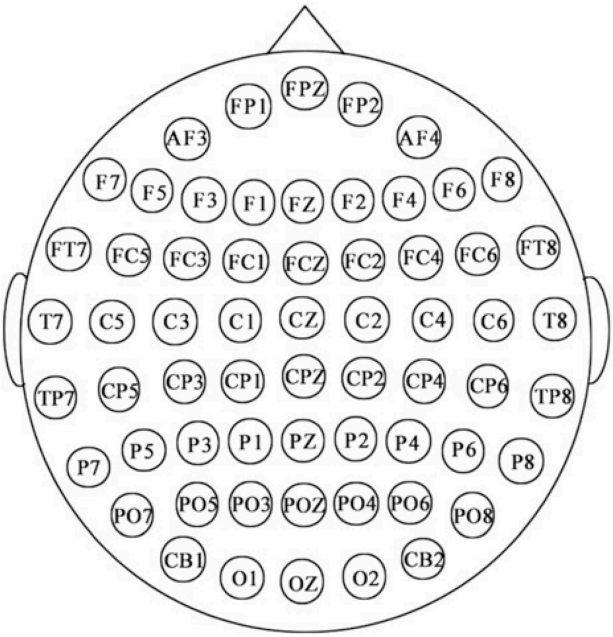

Figure 2. Location of 62 electrodes in SEED database [2].

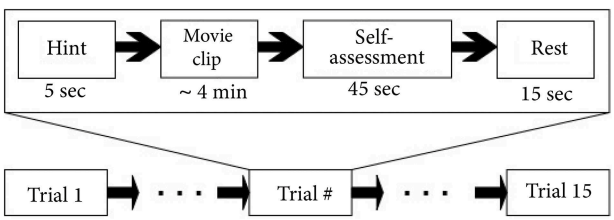

Figure 3. Protocol of EEG recording in each session for SEED database [2].

3.2. Preprocessing

The raw EEG signals are not appropriate for feature extraction and a series of preprocessing steps are needed beforehand. The raw EEG signals are downsampled to 200 Hz. The EEG signals are visually checked in order to remove the parts of the signal that are heavily contaminated with EMG and EOG [2]. The EOG signals recorded during the experiment are also used to detect blink artifacts [28]. EEG signals are processed with a bandpass filter with a frequency band between 0.3 Hz and 50 Hz [2]. Finally, a fifth-order Butterworth filter is used to divide the signals into delta (1-4 Hz), theta (4-8 Hz), alpha (8-14 Hz), beta (14-31 Hz), and gamma (31-50 Hz) frequency bands and then in the time domain, the signal of each band is divided into time windows of 2 s without overlapping. Time windowing is helpful to use the stationarity assumption in each time window because when the length of a time series is shorter, this assumption is more valid. Finally, features will be extracted from each of the 2 s windows.

3.3. Features extraction

In previous studies, various features such as DE, PSD, DASM, RASM, etc. have been used in emotion recognition. These features represent the local activity of electrodes and do not take into account the relationship between the electrodes. Among the local features, DE has the best performance in emotion classification [2,6,24]. Functional brain connectivity features show the relationship between different brain region's activities which are recorded by different electrodes. Pearson correlation coefficient, coherence, and PLV which are three important connectivity features are extracted in this study. These three features are non-directed and as a result, their connectivity matrix is symmetric [29].

3.3.1. Pearson correlation coefficient

Pearson correlation coefficient is a simple criterion of linear correlation between two time series. This criterion calculates the linear relationship between the two signals x and y as follows:

$$\rho_{xy} = \frac{E[(x-\mu_x)(y-\mu_y)]}{\sigma_x \sigma_y}$$

$$= \frac{1}{N}\sum_{n=1}^{N}\frac{(x(n)-\mu_x)(y(n)-\mu_y)}{\sigma_x \sigma_y}, \quad (1)$$

where N is the number of time samples, μ_x, μ_y are mean of the signals x and y, respectively and σ_x, σ_y are standard deviation of the signals x and y, respectively. This criterion has a value between -1 and +1. The value of +1 in this criterion means a positive linear correlation between two signals and the value of -1 means a negative linear correlation. The value of zero

means that the two signals are not linearly related to each other.

3.3.2. Coherence

Coherence indicates a linear correlation between two signals in the frequency domain. First, the cross-spectral density function $S_{xy}(f)$, is calculated as:

$$S_{xy}(f) = \sum_{\tau=-\infty}^{+\infty} E\left[x_n y_{n+\tau}\right] e^{-j2\pi f\tau}, \tag{2}$$

where x_n and $y_{n+\tau}$ are the n-th sample of signal x and the $(n+\tau)$-th sample of signal y, respectively. Coherence is then calculated using the cross-spectral density function as follows:

$$C_{xy}(f) = \frac{S_{xy}(f)}{\sqrt{S_{xx}(f)S_{yy}(f)}}, \tag{3}$$

$$COH_{xy}(f) = |C_{xy}(f)|. \tag{4}$$

The cross-spectral density function of the signals x and y is normalized using the spectral density function of signal x and signal y. The value of coherence is between zero and +1. The value of +1 means the highest linear correlation at that frequency. Coherence with the value of zero means that two signals are linearly independent at that frequency.

3.3.3. Phase Locking Value

The most common feature for measuring phase synchronization between different areas of the brain is PLV. Suppose two signals x_1 and x_2 are filtered with a bandpass filter. An analytical signal z_i can be defined as:

$$z_i(t) = x_i(t) + jH(x_i(t)) = A_i(t)e^{j\phi(t)}, \tag{5}$$

where H is the Hilbert transform operator. The phase difference is calculated using:

$$\Delta\phi(t) = arg\left(\frac{z_1(t)z_2^*(t)}{|z_1(t)||z_2(t)|}\right). \tag{6}$$

PLV is then defined as [22]:

$$PLV = \left|\frac{1}{N}\sum_{t=0}^{N-1} e^{j\Delta\phi(t)}\right|, \tag{7}$$

where t is the time point and N is the number of time samples. PLV is between zero and one. The value of one indicates a phase lock, and the value of zero indicates a random phase distribution over time. The amount of PLV is independent of the signal amplitude and depends only on the phase difference between the two signals [30].

3.4. Classification

Most of the works that have used the SEED database have considered the first 9 trials of each session as training data and the final 6 trials of each session as test data. Due to the small number of trials in this database, it is better to use the Leave-One-Out cross-validation strategy. The Leave-One-Out cross-validation strategy decreases the probability of overfitting and has a higher degree of validity. In this strategy, we consider one trial as the test and the other 14 trials as training. We repeat this procedure until every trial is considered once as the test. Finally, we average the classification accuracies over all folds. We use Support Vector Machine (SVM) with linear kernel for classification and also the emotion classification is subject-dependent. The number of samples is equal to *number of time windows* * *number of trials* * *number of sessions*. The number of time windows depends on the length of time windows. In this study, the length of the time window is 2 s, and time windows are non-overlapping. Due to the fact that the length of the videos is not exactly the same, the total number of samples for happy, sad, and neutral emotions in this database is equal to 1767, 1686, and 1653, respectively. Therefore, the distribution of samples is balanced among different emotions. Given that our final purpose is to label the trials while samples are time windows, we use a simple voting scheme to determine the label of a trial using the label of its time windows; for instance, if there are n time windows in a trial, we will have a vector of length n of labels and then, we vote among all the labels in the vector to find the label of the trial.

3.5. Dimension reduction

It is known that an imbalance between the length of the feature vector and the number of samples can cause overfitting. The length of the feature vector for local features such as DE is equal to the number of channels which is equal to 62 in the SEED database. Coherence, PLV, and Pearson correlation coefficient are non-directed, and their connectivity matrix is symmetric. Therefore, the upper or lower half of the connectivity matrix is sufficient for classification. As a result, the length of the vector of connectivity features is equal to (62×61)/2=1891, which is large relative to the number of samples which is approximately 4700. Thus, we should use dimension reduction techniques to prevent overfitting. In this study, we use Fisher score to reduce the dimension. The Fisher score assigns a score to each feature where a higher score for a feature shows that the feature is more discriminative for classification. The score of a feature is obtained as follows [31]:

$$F(i) = \frac{\sum_{k=1}^{c}(\bar{x}_{k,i} - \bar{x}_i)^2}{\sum_{k=1}^{c}\frac{1}{n_k-1}\sum_{j=1}^{n_k}(x_{k,j,i} - \bar{x}_{i,k})^2}, \tag{8}$$

where $\bar{x}_i$ is the mean of the i-th feature in all samples, $x_{k,j,i}$ is the i-th feature of the j-th sample in the k-th class, $\bar{x}_{i,k}$ is the mean of the i-th feature in the k-th class and n_k represents the number of samples in the k-th class.

3.6. Recurrent Neural Networks (RNNs)

A RNNs is a type of neural network in which connections between nodes form a graph along a sequence [32]. RNNs [33] are widely used in text analysis [34,35], text generation [36,37], speech recognition [38], time series prediction [39], and processing of time signals such as EEG [40]. The main idea of RNNs is the use of hidden states which are used as the network memory and are responsible to store the information from past inputs. In RNNs, the output y_t at each time t is calculated by the hidden state h_t at time t. The hidden state h_t at time t is also updated with the input x_t at time t and the hidden state $h_{t\ 1}$ at time $t-1$. Mathematically, there are two important equations at each time step in RNNs as follows [32]:

$$h_t = \sigma(W_{hx}x_t + W_{hh}h_{t-1} + b_h), \tag{9}$$

$$y_t = \text{softmax}(W_{yh}h_t + b_y), \tag{10}$$

where W_{hx}, W_{hh}, and W_{yh} are the weight matrices. b_h and b_y are bias parameters which are utilized to learn the offset and σ is the sigmoid function.

These networks need a large number of samples for training. The number of our samples in the SEED database is small. Therefore, we select Gated Recurrent Units (GRUs) [41] which are a cell type of RNNs and their few parameters make them a suitable choice for small datasets [42]. In addition, we increase the number of samples by adding Gaussian noise with standard deviations of 0.001, 0.004, 0.008, and 0.012 to the samples according to [4]. The utilized network consists of a GRU layer that has 16 units and a fully connected layer for emotion classification. The filtered signal is then divided into time-sliding windows of 180 s with a step length of 2 s as the input of the network. We choose 180 s as the length of our inputs because the signals in the SEED database have different lengths but their minimum length of signals is 180 s. We also opt step length of 2 s to cover the end part of signals and also increase the number of inputs of the network. The Leave-One-Out cross-validation strategy is chosen for calculating accuracy in this section.

3.7. Transformer

Transformers have garnered significant attention in recent years due to their high speed and efficient performance, making them widely used in various domains such as natural language processing [43], image recognition [44], and time series analysis [45]. One of the key strengths of transformers lies in their ability to capture long-range dependencies in sequential data.

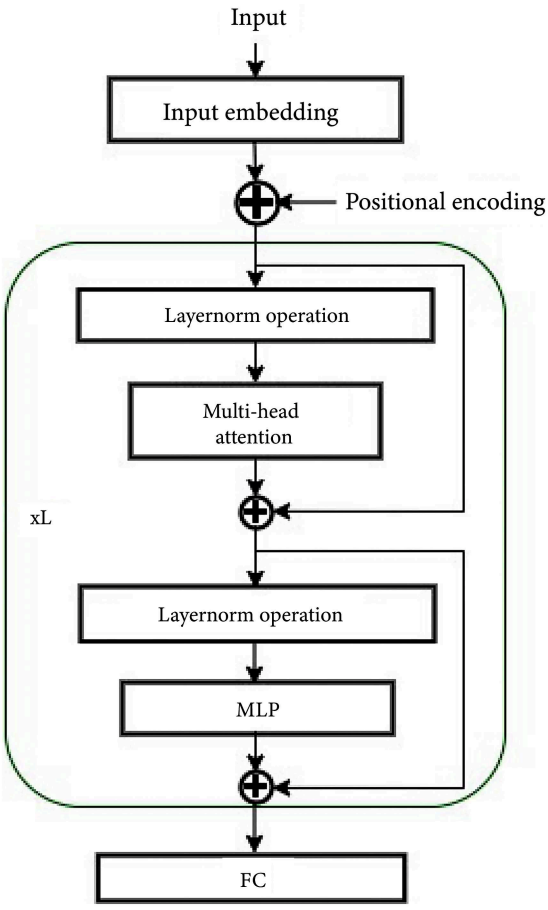

Figure 4. The overview of transformer. L is the number of encoders which is one in our case.

In a transformer model, the input sequence is divided into fixed-length segments or tokens, with each token associated with an embedding vector. The transformer architecture comprises multiple layers of self-attention and feed-forward neural networks, as depicted in Figure 4.

The self-attention mechanism in transformers operates by allowing each token to interact with all other tokens in the sequence. The importance or attention assigned to each token is determined based on its relevance to the others. This attention information is then used to compute a weighted sum of the embeddings of all tokens, resulting in a context-aware representation for each token. Mathematically, this process can be expressed as follows:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_h}}\right)V, \tag{11}$$

where Q, K, and V are query, key, and value matrices, respectively, with a size of $N \times d_h$, where d_h represents the dimensionality of Q, K, and V matrices. N represents the actual length of the input sequence. The attention mechanism enables the model to capture dependencies and relationships between different time intervals without relying on recurrent connections.

For this work, one encoder with two-head self-attention mechanism is employed to better capture the dependencies in the input sequence. Each attention head attends to a different subspace of the input representation and generates its own set of attention weights. Finally, the outputs from all attention heads are concatenated and linearly transformed to obtain the final representation for each token. It's important to note that the input format for the transformer model remains the same as that of the GRU, including the number of samples and their size.

4. Results

4.1. Investigation of frequency bands

Previous research in the field of emotion recognition has shown that higher frequency bands have more information about emotional states and the gamma band has the best performance in emotion classification [2,3,6,22,24,46]. However, most of them do not investigate the performance of different frequency bands for functional brain connectivity features. In this research, in addition to DE features, we investigate the performance of different frequency bands for coherence, PLV, and Pearson correlation coefficient. Table 1 shows the classification accuracy for our features in different frequency bands before dimension reduction.

As seen in Table 1, the gamma and beta bands have better classification accuracy rather than other bands. Specifically, DE achieves its maximum accuracy in the beta band while functional brain connectivity features perform best in the gamma band. The classification accuracy of DE feature in the gamma band is 0.88% less than the beta band while in functional connectivity features, the accuracy of the gamma band is at least 0.89% more than the beta band. Moreover, our focus in this research is on functional brain connectivity features, so we select the gamma band for further analysis. We also do dimension reduction for all features only on the gamma band.

4.2. Dimension reduction results

We apply the Fisher score for DE, Pearson correlation coefficient, coherence, and PLV features in the gamma band because, in Section 4, we investigate the performance of different frequency bands and we observe that the gamma band has better performance in emotion recognition. We calculate F-score for all features and then sort them in order to find the first 100, 200, 300, 400, 500, 600, 700, 800, and 1200 selected features for coherence, PLV, and Pearson correlation coefficient. We also calculate F-score for DE features and then sort them in order to find the first 10, 20, 30, 40, and 50 selected features for DE. The optimal number of selected features for DE, Pearson correlation coefficient, coherence, and PLV are 40, 700, 200 and 400, respectively. Classification accuracy versus the number of functional connectivity features and DE are shown in Figures 5 and 6, respectively. Figures 5 and 6 show the importance of feature selection; If the number of selected features is little, then the features will not be discriminative and the classification accuracy will be low. On the other hand, if the number of features is too large, the accuracy of classification is reduced because overfitting happens. Table 2 shows the classification

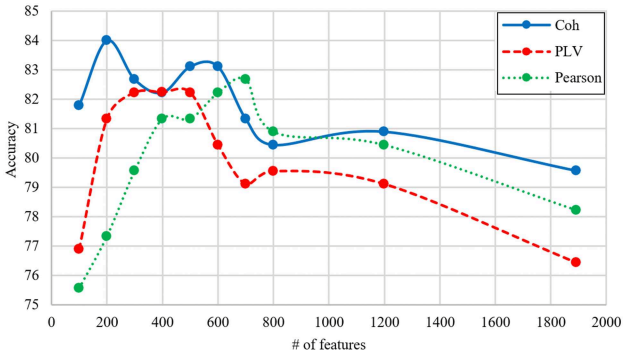

Figure 5. Average accuracy of classification over different subjects versus the number of selected functional connectivity features with Fisher score.

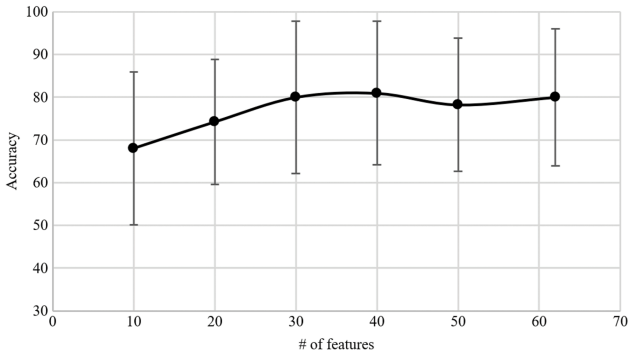

Figure 6. Average accuracy of classification over different subjects versus the number of selected DE features with Fisher.

Table 1. The mean accuracies and standard deviations (%) of different frequency bands.

	Delta	Theta	Alpha	Beta	Gamma
DE	64.44 ± 18.97	58.67 ± 15.97	65.33 ± 18.72	**81.32 ± 16.17**	80.44 ± 16.03
Coh	40.44 ± 17.36	51.11 ± 21.03	66.67 ± 19.52	77.78 ± 18.97	**79.56 ± 17.54**
PLV	26.67 ± 17.81	40.89 ± 21.21	60.44 ± 22.17	75.56 ± 18.11	**76.45 ± 17.97**
Pearson	35.11 ± 10.83	49.78 ± 19.00	65.78 ± 19.97	75.56 ± 17.02	**78.22 ± 18.25**

Table 2. The mean accuracies and standard deviations (%) before and after dimension reduction.

	DE	Coh	PLV	Pearson
Before Fisher	80.44 ± 16.03	79.56 ± 17.54	76.45 ± 17.97	78.22 ± 18.25
After Fisher	80.90 ± 16.87	84.00 ± 15.49	82.22 ± 16.65	82.67 ± 14.21

Table 3. The mean accuracies and standard deviations (%) of functional connectivity features and fusion of functional connectivity features with DE feature.

	Coh	PLV	Pearson
Brain connectivity	84.00 ± 15.49	82.22 ± 16.65	82.67 ± 14.21
Brain connectivity + DE	84.00 ± 15.28	83.11 ± 13.58	84.44 ± 13.72

Table 4. The mean accuracies and standard deviations (%) of different time intervals.

	0-60 s	20-80 s	40-100 s	60-120 s	80-140 s	100-160 s	120-180 s	140 s-end
DE	65.33 ± 20.65	64.89 ± 18.07	69.78 ± 17.43	72.89 ± 16.61	77.33 ± 16.09	81.33 ± 14.29	81.78 ± 11.39	**84.89 ± 13.20**
Coh	61.33 ± 22.70	69.33 ± 18.82	74.22 ± 19.00	77.78 ± 15.46	82.67 ± 13.75	86.67 ± 15.11	86.22 ± 15.82	**88.00 ± 13.38**
PLV	59.56 ± 22.17	68.44 ± 19.75	72.00 ± 19.38	76.00 ± 16.67	81.33 ± 16.36	85.33 ± 16.56	85.33 ± 13.84	**88.44 ± 10.53**
Pearson	60.44 ± 20.84	69.78 ± 19.00	72.89 ± 18.59	77.33 ± 15.69	81.78 ± 16.22	86.22 ± 15.82	87.11 ± 13.20	**88.44 ± 12.20**

accuracy before and after performing Fisher on our features.

4.3. Emotion classification using fusion of DE and functional brain connectivity

According to our results in the previous part and a wide range of research in the field of emotion recognition, we carry out our experiments only in the gamma band. It is worth mentioning that from this part to the end of this study, we utilize all features in the gamma band after dimension reduction for our analysis. In this part, we investigate the effect of the fusion of functional connectivity and DE features on the classification accuracy in emotion recognition. This decision-level fusion uses the probability vectors of functional connectivity and DE features to make decisions. The probability vector of each feature is defined as $(P_{sad}, P_{neutral}, P_{happy})$, where P_{sad}, $P_{neutral}$, and P_{happy} express the probability that a trial belongs to a sad, neutral, or happy class, respectively. Then, the probability vector for the decision-level fusion is defined as follows:

$$P_{fusion} = 0.5 * (P_{sad}, P_{neutral}, P_{happy})_{connectivity\_features} + 0.5 * (P_{sad}, P_{neutral}, P_{happy})_{DE\_features}, \quad (12)$$

where *connectivity_features* can be Pearson correlation coefficient, coherence, and PLV features. Finally, the class which has the maximum probability in P_{fusion} vector is assigned to the trial label. Table 3 shows classification accuracy using a decision-level fusion of functional connectivity and DE features.

The classification accuracy using the DE feature in the gamma band is equal to 80.90% with 16.87% std. Table 3 shows that the fusion of functional connectivity and DE features improves classification accuracy. However, because of our focus on the temporal pattern of functional connectivity features, we do not fuse the functional connectivity and DE features in the rest of this article to purely observe the temporal pattern of functional brain connectivity features. It is also worth mentioning that we investigated the fusion of different connectivity features together but no improvement was observed.

4.4. Analysis of temporal pattern

Recent works mostly investigated the location and frequency of brain activity during emotional stimuli. To the best of the author's knowledge, brain activity during an emotional stimulation interval has not been yet studied in the literature. The length of stimulus in the SEED database is approximately 4 minutes. In this section, we want to find the temporal patterns over different time periods during each trial. To this end, we consider a one-minute sliding window with a step length of 20 seconds and calculate classification accuracy in each sliding window. Table 4 shows the classification accuracy in different time periods of signals.

As seen in Table 4, there is a temporal pattern in the classification accuracy in which, as time goes on, the mean accuracy increases. The best time period is 140 s-end in which the mean accuracy is maximum and the standard deviation is minimum. Comparing the results of Tables 3 and 4 shows that the mean accuracy in 140 s-end interval is increased compared to the entire signal by 3.99%, 4%, 6.22%, and 5.77% for DE, coherence, PLV, and Pearson correlation coefficient, respectively. The standard deviation in 140 s-end interval is also decreased compared to the entire signal.

It is noteworthy that the length of stimuli in many datasets for EEG-based emotion recognition such as DEAP [25] is less than two minutes while the

Table 5. The mean accuracies and standard deviations (%) of approach (1) averaging between subjects and then finding the best interval, and approach (2) obtain the optimal interval for each subject.

	DE	Coh	PLV	Pearson
Approach 1	84.89 ± 13.20	88.00 ± 13.38	88.44 ± 10.53	88.44 ± 12.20
Approach 2	86.22 ± 11.94	92.00 ± 11.32	91.56 ± 10.53	91.11 ± 10.59

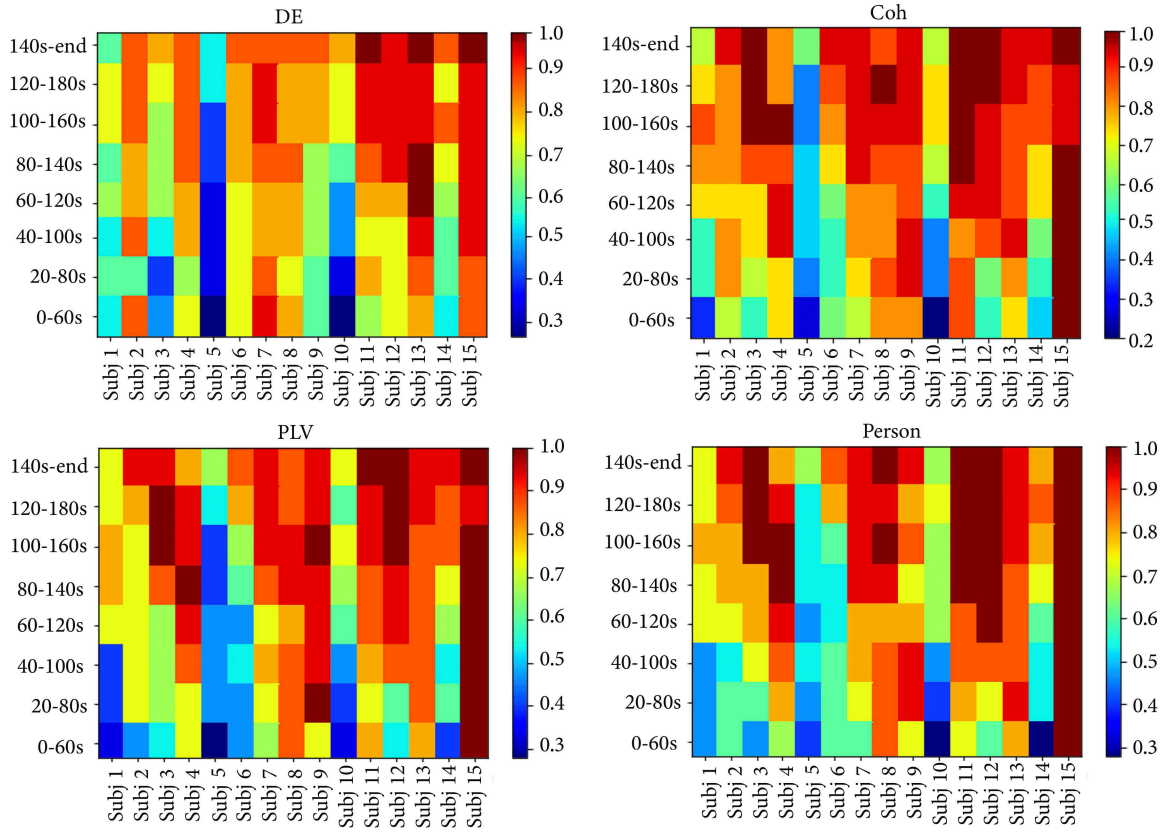

Figure 7. Visualization of the emotion classifications accuracy for each subject in different time intervals for DE, coherence, PLV, and Pearson correlation coefficient.

temporal analysis of SEED shows the importance of brain activities that occur after two minutes from the onset of stimuli. Therefore, we conclude that the appropriate length of stimuli in emotion recognition is about 3 to 4 minutes so that not only the subjects do not get tired but also there are the best periods for emotion recognition. In this study, we watched the videos and confirmed that all parts of the videos express the same emotion as the label of the video. Although we can understand the label of the video from the beginning part of it, over time the emotion becomes more stable in our mind, and consequently, the final intervals have better performance in emotion classification. We can also observe this phenomenon in our ordinary life, for instance, when a person tells a funny sentence, you feel happy and if that person tells a sequence of funny sentences, you feel happier and even a few minutes after telling those funny sentences, you still feel happy. Therefore, it seems that the length of stimuli in emotion recognition has an optimal interval at which the subject reaches the peak of his emotions. If the length of stimuli is less than the optimal value then, the subject's emotions are not well expressed, and if the length of stimuli is greater than the optimal length, the subjects get tired.

The results shown in Table 4 are the averaged classification accuracies over different subjects. Another approach is to obtain the best interval for each subject separately. By using the second approach, we found that the best interval is not the same for all subjects, however, the best interval is 140 s-end for 13, 11, 10, and 11 subjects (out of 15 subjects) in DE, coherence, PLV, and Pearson correlation coefficient features, respectively. The best interval for other subjects occurs 120 s, 100 s, 80 s and 100 s after the onset of stimulation for DE, coherence, PLV, and Pearson correlation coefficient features, respectively. To visualize the results of the temporal analysis, the accuracy using each time interval is shown in Figure 7, separately for each feature. In this figure, squares with

Table 6. The mean accuracies and standard deviations (%) of SVM and GRU for functional connectivity features.

	Coh	PLV	Pearson
SVM	88.00 ± 13.38	88.44 ± 10.53	88.44 ± 12.20
GRU	84.00 ± 11.37	87.11 ± 12.46	81.78 ± 11.94
Transformer	**88.44 ± 12.46**	**88.89 ± 11.27**	**88.89 ± 13.27**

higher intensity are indicators of better classification accuracy.

Table 5 shows the comparison between the results of these two approaches: (1 averaging between subjects and then finding the best interval (first row of Table 5), 2) finding the best interval for each subject and then averaging between the maximum accuracy of each subject (second row of Table 5).

The results of Table 5 show that the average accuracy in the second approach compared to that in the first approach increases by 1.33% to 4%. Therefore, regarding the different characters and reactions of different subjects, it is a wise idea to find the best time interval for each subject independently and then predict the subjects emotions.

4.5. Classification using RNN and transformer

In the previous section, we found the best time interval using SVM. In this section, we investigate the performance of RNNs [32] and transformer in finding the temporal pattern of brain activity during emotional stimulation. In order to provide a fair comparison, we augment the training set by adding Gaussian noise to the training samples and then adding them to the training set. We also employ GRUs with few parameters for compatibility with the size of the dataset. Data augmentation and reduction of the number of parameters of the network prevent overfitting of the classifier to train samples. Table 6 compares the results of the RNN and transformer with the results of the previous section obtained using SVM.

According to the results presented in Table 6, the Transformer model demonstrates the highest performance among all brain connectivity features analyzed in this study. However, it is worth noting that its performance is closely comparable to that of the SVM model. This is significant because neural networks typically demand substantial amounts of data and come with a high computational cost. Therefore, the proposed time-interval selection method provides both advantages of relatively high accuracy and low computational cost for the task of emotion recognition.

5. Discussion

In this study, our aim was to examine the functional connectivity features in more detail as well as the temporal analysis of the brain's response to emotional stimuli, which have been rarely studied in the field of emotion recognition. In this section, we compare our results with previous works. However, it is difficult to compare the classification accuracy of our method with previous works because they have used different datasets with different numbers of subjects and various methods of representing emotions. Moreover, even in the works that use the SEED dataset for their experiments, the first 9 trials are considered as training and the final 6 trials as test data. As a result, due to the small size of this dataset, the accuracy reported by those methods may not be much reliable. In order to obtain more reliable results, we used the Leave-One-Out strategy on the trials to calculate the classification accuracy for each subject. Consequently, our classification accuracy might be lower than some of those reported in the literature due to our different strategies for constructing train and test sets but it is more valid. Therefore, in this comparison, it is important to note that firstly, we used a more reliable method to measure the classification accuracy, and secondly, our goal was to find optimal time intervals and analyze temporal patterns, but not to propose a method to increase the classification accuracy in emotion recognition.

We summarize the accuracy rates of emotion classification in previous works in Table 7. In [2,3,6,24], frequency analysis for some features like DE and PSD have been investigated, but the functional connectivity features have not been investigated except in [22]. In [22], it has been shown that using the gamma and beta bands results in better performance compared to using the other frequency bands for emotion classification. As discussed in Subsection 4.1., the results of our frequency analysis on the functional connectivity features using the SEED database confirm the results of previous articles. It can be concluded that these two frequency bands contain more discriminative information related to the emotional states of individual's brains.

Regarding temporal analysis, unfortunately, none of the previous works has studied the temporal patterns during stimuli, but we examined this for the first time. We observed that each subject has an optimal time interval in which the classification performance is higher than the accuracy of using the whole signal. It is noteworthy that, as discussed in Subsection 4.4. the optimal time interval for all subjects in this database occurs 80 s after the onset of stimuli. Therefore, the

Table 7. The accuracy of emotion classification in previous works.

Study	Year	Data	Features	Frequency band	Classifier	Number of classes	Description	Accuracy
Zheng et al. [2]	2015	SEED	DE	Gamma band	DBN	3	-	79.19%
Duan et al. [3]	2013	Collected	DE	Gamma band	SVM	2: positive and negative emotions	-	84.25%
Wang et al. [4]	2018	SEED	DE	Concatenation of all frequency bands	ResNet	3	–	75%
Zheng et al. [18]	2019	SEED	DE	Concatenation of all frequency bands	GELM	3	–	91.07%
Song et al. [6]	2020	SEED	DE	Gamma band	GCNN	3	–	83.36%
Our method	–	SEED	DE	Gamma band	SVM	3	–	84.89%
Moon et al. [5]	2018	DEAP	PLV, PCC	Concatenation of all frequency bands	CNN (two layers)	2: low and high valence	–	96.62% 93.80%
Li et al. [22]	2019	SEED	ENP (extracted from PLV graph)	Gamma band	SVM	3	-	33%*
Wu et al. [48]	2019	SEED	Three topological features (for PCC and Coh indices)	Concatenation of all frequency bands	SVM	3	Feature-level fusion method was used to fuse topological features	79.16% 78.15%
Our method	–	SEED	Coh, PLV, Pearson	Gamma band	SVM	3	–	88.00% 88.44% 88.44%

PCC stands for Pearson correlation coefficient

* They achieved 79% classification accuracy using GELM.

different subjects behave relatively similarly to each other.

Using DE features extracted from the gamma band we obtained a mean accuracy of 84.89%. More details regarding the comparison between our result and the previous works are as follow. In [2], the accuracy of 79.19% using DE features extracted from the gamma band and Deep Belief Networks (DBN) as the classifier has been achieved. In [24], the mean accuracy of 91.07% using the concatenation of DE from all frequency bands and Graph regularized Extreme Learning Machine (GELM) has been obtained. Despite the good accuracy obtained in [24], the computational complexity of GELM is higher than SVM which is utilized in our study. In [3], the accuracy of 84.25% using DE features from the gamma band has been achieved. In [3], the authors collected their own dataset which contains only two positive and negative emotional states, while in our research, the number of classes is three. In [6], the average accuracy of 83.36% using DE features from the gamma band and Graph Convolutional Neural Networks (GCNNs) has been obtained on the SEED dataset. Despite the good performance of Graph Neural Networks, their computational cost is higher than SVM which is used in our work. In [4], the maximum accuracy of 75% has been reached using DE features and ResNet [47] on the SEED dataset. Due to the small size of the SEED dataset, in [4] first data augmentation methods are used to provide sufficient data for training of the network.

Using functional connectivity features, we have achieved mean accuracy of 88%, 88.44%, and 88.44% for coherence, PLV, and Pearson correlation coefficients, respectively. More details regarding the comparison between our results and the previous works are as follow. In [5], the accuracy rates of 93.80% and 96.62% using a two-layer Convolutional Neural Network (CNN) as the classifier and Pearson correlation coefficients and PLV matrices as features have been achieved, respectively. In [5], only the results of high and low valence classification are reported, which are on the DEAP dataset. They have not reported their accuracy No results on the arousal classification have been reported. In [22], 33% accuracy on the SEED dataset using the SVM classifier and ENP features extracted from the gamma band using the PLV graph has been achieved. In [48], first of all, the Pearson correlation coefficients and coherence were extracted then three topological features of strength, clustering coefficient, and eigenvector centrality were calculated for each connectivity indices. By using the feature-level fusion method, these three topological features were concatenated. Finally, they obtained an accuracy of 79.16% and 78.15% on Pearson correlation coefficients and coherence indices.

It can be seen that although there are a few works with better accuracy rates on the SEED dataset, the accuracy that we have obtained using the optimal time intervals and a simple SVM classifier is competitive with the results of previous works while it has less computational costs.

6. Conclusion

In this study, we examined the applicability of functional brain connectivity for the emotion recognition task. Our results indicate that by using functional connectivity features better emotion classification accuracy can be obtained compared to using Differential Entropy (DE) features. In order to study the temporal variation of different features in terms of classification accuracy, we used a one-minute sliding window on the signals. Classification accuracy increased with sliding window progress leading us to the conclusion that using the 140 s-end interval results in the best performance compared to using other intervals as well as compared to using the entire signal in the SJTU Emotion EEG Dataset (SEED) dataset. The mean accuracy and standard deviation using this interval (140 s-end) for DE, Phase Locking Value (PLV), coherence, and Pearson correlation coefficient are 84.89%/13.20%, 88.44%/10.53%, 88%/13.38%, and 88.44%/12.20%, respectively, showing 4-6% improvement compared to using the entire signal. Although the temporal behavior of each subject is different, the selected interval (140 s-end) is the best period for at least two-thirds of the subjects. The results show that by using the best time interval, we can achieve high accuracy with relatively low computational cost and a limited number of training samples.

7. Future works

In future works, a dataset with more participants can be collected because at the moment, there is not any large database for EEG-based emotion recognition. A larger dataset smooths the path to present results with a higher level of validity and generalization. In addition, a large dataset provides an opportunity for us to apply complex deep neural networks without concern about overfitting. Furthermore, we can investigate multimodality learning by using brain imaging techniques besides brain signals simultaneously. For instance, a combination of brain imaging techniques that have good spatial resolution such as fMRI images and EEG signals is utilized to have a better understanding of spatial and temporal patterns simultaneously, being very valuable in neuroscience.

Declaration

The authors declare no conflict of interest.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests.

Authors contribution statement

First name and last name of frist author Ensieh Khazaei: Conceptualization, methodology, formal analysis, validation, writing original draft.

First name and last name of second author Hoda Mohammadzade: Conceptualization, supervision, review and editing.

References

1. Chao, H., Zhi, H., Dong, L., et al. "Recognition of emotions using multichannel EEG data and DBN-GC-based ensemble deep learning framework", *Comput. Intell. Neurosci.*, **2018**, pp. 1–11 (2018). DOI: 0.1155/2018/9750904
2. Zheng, W.-L. and Lu, B.-L. "Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks", *IEEE Trans. Auton. Ment. Dev.*, **7**(3), pp. 162–175 (2015). DOI: 10.1109/TAMD.2015.2431497
3. Duan, R.-N., Zhu, J.-Y., and Lu, B.-L. "Differential entropy feature for EEG-based emotion classification", *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*, San Diego, CA, USA, pp. 81-84 (2013). DOI: 10.1109/NER.2013.6695876
4. Wang, F., Zhong, S., Peng, J., et al. "Data augmentation for EEG-based emotion recognition with deep convolutional neural networks", *MultiMedia Modeling*, **17**, pp. 82–93 (2018). DOI: 10.1007/978-3-319-73600-6_8
5. Moon, S.-E., Jang, S., and Lee, J.-S. "Convolutional neural network approach for EEG-based emotion recognition using brain connectivity and its spatial information", *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Calgary, AB, Canada, pp. 2556–2560 (2018). DOI: 10.1109/ICASSP.2018.8461315
6. Song, T., Zheng, W., Song, P., et al. "EEG emotion recognition using dynamical graph convolutional neural networks", *IEEE Trans. Affect. Comput.*, **11**(3), pp. 532–541 (2020). DOI: 10.1109/TAFFC.2018.2817622
7. Lin, Y.-P., Wang, C.-H., Jung, T.-P., et al. "EEG-based emotion recognition in music listening", *IEEE Trans. Biomed. Eng.*, **57**(7), pp. 1798–1806 (2010). DOI: 10.1109/TBME.2010.2048568
8. Menting-Henry, S., Hidalgo-Lopez, E., Aichhorn, M., et al. "Oral contraceptives modulate the relationship between resting brain activity, amygdala connectivity and emotion recognition, A Resting State fMRI

Study", *Frontiers in Behavioral Neuroscience*, **16**, 775796 (2022). DOI: 10.3389/fnbeh.2022.775796

9. Gannouni, S., Al-Edaily, A., Belwafi, K., et al. "Emotion detection using electroencephalography signals and a zero-time windowing-based epoch estimation and relevant electrode identification", *Scientific Reports*, **11**(1), 7071 (2021). DOI: 10.1038/s41598-021-86345-5

10. Hattingh, C.J., Ipser, J., Tromp, S.A., et al. "Functional magnetic resonance imaging during emotion recognition in social anxiety disorder: an activation likelihood meta-analysis", *Frontiers in Human Neuroscience*, **6**, 347 (2013). DOI: 10.3389/fnhum.2012.00347

11. Li, C.M., Wang, B., Zhang, S., et al. "Emotion recognition from EEG based on multi-task learning with capsule network and attention mechanism", *Computers in Biology and Medicine*, **143**, 105303 (2022). DOI: 10.1016/j.compbiomed.2022.105303

12. Li, C.M., Zhang, Z., Zhang, X., et al. "EEG-based emotion recognition via transformer neural architecture search", *IEEE Transactions on Industrial Informatics*, **19**(4), pp. 6016–6025 (2022). DOI: 10.1109/TII.2022.3170422

13. Malmivuo, J. "Comparison of the properties of EEG and MEG in detecting the electric activity of the brain", *Brain Topography*, **25**(1), pp. 1–19 (2012). DOI: 10.1007/s10548-011-0202-1

14. Van den Broek, E.L. "Ubiquitous emotion-aware computing", *Pers. Ubiquitous Comput.*, **17**(1), pp. 53–67 (2013). DOI: 10.1007/s00779-011-0479-9

15. Mehrabian, A. "Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in Temperament", *Curr. Psychol.*, **14**(4), pp. 261–292 (1996). DOI: 10.1007/BF02686918

16. Khosrowabadi, R., Heijnen, M., Wahab, A., et al. "The dynamic emotion recognition system based on functional connectivity of brain regions", *2010 IEEE Intelligent Vehicles Symposium*, La Jolla, CA, USA, pp. 377–381 (2010). DOI: 10.1109/IVS.2010.5548102

17. Chi, Y.M., Wang, Y.-T., Wang, Y., et al. "Dry and noncontact EEG sensors for mobile brain-computer interfaces", *IEEE Trans. Neural Syst. Rehabil. Eng.*, **20**(2), pp. 228–235 (2012). DOI: 10.1109/TNSRE.2011.2174652

18. Wang, L.-F., Liu, J.-Q., Yang, B., et al. "PDMS-based low cost flexible dry electrode for long-term EEG measurement", *IEEE Sens. J.*, **12**(9), pp. 2898–2904 (2012). DOI: 10.1109/JSEN.2012.2204339

19. Huang, Y.-J., Wu, C.-Y., Wong, A.M.-K., et al. "Novel active comb-shaped dry electrode for EEG measurement in hairy site", *IEEE Trans. Biomed. Eng.*, **62**(1), pp. 256–263 (2015). DOI: 10.1109/TBME.2014.2347318

20. https://bcmi.sjtu.edu.cn/home/seed/

21. Stam, C.J., Nolte, G., and Daffertshofer, A. "Phase lag index: Assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources", *Hum. Brain Mapp.*, **28**(11), pp. 1178–1193 (2007). DOI: 10.1002/hbm.20346

22. Li, P., Liu, H., Si, Y., et al. "EEG based emotion recognition by combining functional connectivity network and local activations", *IEEE Trans. Biomed. Eng.*, **66**(10), pp. 2869–2881 (2019). DOI: 10.1109/TBME.2019.2897651

23. Thammasan, N., Moriyama, K., Fukui, K.-I., et al. "Familiarity effects in EEG-based emotion recognition", *Brain Inform.*, **4**(1), pp. 39–50 (2017). DOI: 10.1007/s40708-016-0051-5

24. Zheng, W.-L., Zhu, J.-Y., and Lu, B.-L. "Identifying stable patterns over time for emotion recognition from EEG", *IEEE Trans. Affect. Comput.*, **10**(3), pp. 417–429 (2019). DOI: 10.1109/TAFFC.2017.2712143

25. Koelstra, S., Muehl, C., Soleymani, M., et al. "DEAP: A dataset for emotion analysis using physiological and audiovisual signals", *IEEE Transactions on Affective Computing*, **3**(1), pp. 18-31 (2012).

26. Katsigiannis, S., and Ramzan, N. "DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices", *IEEE Journal of Biomedical and Health Informatics*, **22**(1), pp. 98–107 (2018). DOI: 10.1109/JBHI.2017.2688239

27. Eysenck, S.B.G., Eysenck, H.J., and Barrett, P. "A revised version of the psychoticism scale", *Personality and Individual Differences*, **6**(1), pp. 21–29 (1985). DOI: 10.1016/0191-8869(85)90026-1

28. Singh, B., and Wagatsuma, H. "A removal of eye movement and blink artifacts from EEG data using Morphological Component Analysis", *Comput. Math. Methods Med.*, **2017**, 1861645 (2017). DOI: 10.1155/2017/1861645

29. Bastos, A.M. and Schoffelen, J.-M. "A tutorial review of functional connectivity analysis methods and their interpretational pitfalls", *Front. Syst. Neurosci.*, **9**, p. 175 (2015). DOI: 10.3389/fnsys.2015.00175

30. Lachaux, J.P., Rodriguez, E., Martinerie, J., et al. "Measuring phase synchrony in brain signals", *Hum. Brain Mapp.*, **8**(4), pp. 194–208 (1999). DOI: 10.1002/(SICI)1097-0193(1999)8:4/194::AID-HBM4ب3.0.CO;2-C

31. Chen, Y.-W. and Lin, C.-J. "Combining SVMs with various feature selection strategies", *Feature Extraction*, Berlin, Heidelberg: Springer Berlin Heidelberg, **207**, pp. 315–324 (2008). DOI: 10.1007/978-3-540-35488-8_13

32. Lipton, Z.C., Berkowitz, J., and Elkan, C. "A critical review of recurrent neural networks for sequence learning", *arXiv [cs.LG]* (2015).

33. Jain, L.C. and Medsker, L.R., *Recurrent Neural Networks: Design and Applications*, 1st Ed., Boca Raton, FL: CRC Press (1999).

34. Liu, Z., Yang, M., Wang, X., et al. "Entity recognition from clinical texts via recurrent neural network", *BMC Med. Inform. Decis. Mak.*, **17**(S2), pp. 53-61 (2017). DOI: 10.1186/s12911-017-0468-7

35. Wu, Y.-C., Yin, F., Chen, Z., et al. "Handwritten Chinese text recognition using separable multi-dimensional recurrent neural network", *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, Kyoto, Japan, pp. 79-84 (2017). DOI: 10.1109/ICDAR.2017.22

36. Sutskever, I., Martens, J., and Hinton, G.E. "Generating text with recurrent neural network", In *Proceedings of the 28th International Conference on International Conference on Machine Learning (ICML'11)*, Madison, WI, USA, pp. 1017–1024 (2011). DOI: 10.5555/3104482.3104610

37. Yang, Z.-L., Guo, X.-Q., Chen, Z.-M., et al. "RNN-stega: Linguistic steganography based on recurrent neural networks", *IEEE Trans. Inf. Forensics Secur.*, **14**(5), pp. 1280–1295 (2019). DOI: 10.1109/TIFS.2018.2871746

38. Lu, L., Zhang, X., and Renais, S. "On training the recurrent neural network encoder-decoder for large vocabulary end-to-end speech recognition", In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Shanghai, China, pp. 5060–5064 (2016). DOI: 10.1109/ICASSP.2016.7472641

39. Guo, T., Xu, Z., Yao, X., et al. "Robust online time series prediction with recurrent neural networks", In *2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, Montreal, QC, Canada, pp. 816-825 (2016). DOI: 10.1109/DSAA.2016.92

40. Vidyaratne, L., Glandon, A., Alam, M., et al. "Deep recurrent neural network for seizure detection", In *2016 International Joint Conference on Neural Networks (IJCNN)*, Vancouver, BC, Canada, pp. 1202–1207 (2016). DOI: 10.1109/IJCNN.2016.7727334

41. Dey, R. and Salem, F.M. "Gate-variants of Gated Recurrent Unit (GRU) neural networks", *2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)*, Boston, MA, USA, pp. 1597-1600 (2017). DOI: 10.1109/MWSCAS.2017.8053243

42. Gruber, N. and Jockisch, A. "Are GRU cells more specific and LSTM cells more sensitive in motive classification of text", *Front Artif Intell*, **3**, p. 40 (2020). DOI: 10.3389/frai.2020.00040

43. Nguyen, M.V., Lai, V.D., Veyseh, A.P.B., et al. "Trankit: A light-weight transformer-based toolkit for multilingual natural language processing", arXiv Preprint arXiv:2101.03289 (2021).

44. Dosovitskiy, A., Beyer, L., Kolesnikov, A., et al. "An image is worth 16×16 words: Transformers for image recognition at scale", *arXiv preprint arXiv:2010.11929* (2020).

45. Kostas, D., Aroca-Ouellette, S., and Rudzicz, F. "BENDR: Using transformers and a contrastive self-supervised learning task to learn from massive amounts of EEG data", *Frontiers in Human Neuroscience*, **15**, 653659 (2021). DOI: 10.3389/fnhum.2021.653659

46. Li, X., Song, D., Zhang, P., et al. "Exploring EEG features in cross-subject emotion recognition", *Front. Neurosci.*, **12**, p. 162 (2018). DOI: 10.3389/fnins.2018.00162

47. He, K., Zhang, X., Ren, S., et al. "Deep residual learning for image recognition", *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, pp. 770-778 (2016).

48. Wu, X., Zheng, W.-L., and Lu, B.-L. "Identifying functional brain connectivity patterns for EEG-based emotion recognition", *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*, San Francisco, CA, USA, pp. 235-238 (2019). DOI: 10.1109/NER.2019.8717035

Biographies

Ensieh Khazaei received her BSc degree of Science in Electrical Engineering and her MSc degree in Biomedical Engineering both from Sharif University of Technology, Iran, in 2018 and 2020, respectively. Her current research interests include computer vision, machine learning, neuroscience, and signal processing.

Hoda Mohammadzade is an Associate Professor at Sharif University of Technology, Department of Electrical Engineering. She received her BSc degree from Amirkabir University of Technology (Tehran Polytechnic), Iran, in 2004, the MSc degree from the University of Calgary, Canada, in 2007, and the PhD degree from the University of Toronto, Canada, in 2012, all in electrical engineering. Her research interests are machine learning, computer vision, biomedical signal processing, and biometrics. She has published several papers in international journals and conferences.