



An innovative emotion assessment using physiological signals based on the combination mechanism

M. Amjadzadeh* and K. Ansari-Asl

Department of Electrical Engineering, Faculty of Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran.

Received 1 January 2015; received in revised form 7 May 2016; accepted 25 July 2016

KEYWORDS

Emotion assessment;
EEG;
Peripheral signal;
Feature extraction;
Classification.

Abstract. The main purpose of this paper is the assessment of emotions using Electroencephalogram (EEG) and peripheral physiological signals and improvement of recognition accuracy of emotional states using combination mechanism. In the first step, according to the type of signals, effective features were extracted in the time and frequency domains; then, by using the Fisher's Linear Discriminant (FLD) method, the most effective features were selected. Based on these features, six classifiers were used: Support Vector Machine (SVM), Nearest Mean (NM), K-Nearest Neighborhood (K-NN), 1-Nearest Neighborhood (1-NN), FLD, and Linear Discriminant Analysis (LDA). They classified emotions in two classes (low and high) through arousal, valence, and liking dimensions. The Leave-One-Out Cross-Validation (LOOCV) method has been implemented to evaluate the performance of classifiers. To enhance the accuracy of classification, combination at feature and classifier levels was performed. Via the concatenation method, combination at feature level was done. Then, by Majority voting, Fixed and Stacking algorithms, combination at classifier level was implemented. Results showed that these classifiers were selected properly and, thanks to them, good improvements were achieved compared with previous studies. Finally, by using combination methods, obtained recognition accuracy was much more reliable and combination at classifier level resulted in significant improvement.

© 2017 Sharif University of Technology. All rights reserved.

1. Introduction

Emotions are subjective activities associated with the human brain that influence important processes such as memory, concentration, thinking, and decision-making. Emotions play important roles in human communications and can be expressed either verbally through the emotional vocabulary or by expressing non-verbal cues such as pitch tone of voice, facial expressions, and body language. So, in this context, decoding of emotional cues is essential to understand a message [1]. But,

emotion is not only what arises; psychologists have founded the exact boundary between what arises and what is felt by a person [2]. Changing the face, voice, and body movements of a person shows outer aspect of emotions, but activities of the central nervous system (brain) and peripheral nervous system are indications of emotions' inner aspect [3,4].

Although emotions play important roles in people's life, but scientific knowledge about human emotion detection is still limited. For example, many Human-Machine interaction systems do not have the ability to interpret the human emotional information and use the emotional intelligence. In other words, they cannot recognize human emotion states and cannot use this information to decide on the appropriate action. The purpose of computation for recognizing human emotion is filling this lack with detecting emotional

*. Corresponding author. Fax: 009 86133336642
E-mail addresses: amjadzadeh@yahoo.com (M. Amjadzadeh); karim.ansari@scu.ac.ir (K. Ansari-Asl)

cues that occur during human-computer interaction [1]. In recent decades, most studies have focused on analysis of facial expression and voice to detect emotional states such as fear, anger, relaxation, excitement, and surprise. Many efforts have been done in this area, but these ways did not give enough information to the researchers about human emotions, and so they focused only on their outer aspects. Physiological activities also include emotional information that can be used to assess emotions, but less attention was done to them.

In 1998, researchers recorded peripheral signals including Heart Rate (HR), Galvanic Skin Resistivity (GSR), respiration rate (RSP), and electromyogram (EMG) of an actress and assessed eight different emotional states [5]; results showed that distinction between the eight emotional states was not done well. In 2004, Takahashi et al. assessed emotions by recording peripheral signals including electrocardiogram (ECG), GSR, Blood Volume Pressure (BVP), and EEG signals simultaneously [6]. In 2006, the simultaneous use of EEG and peripheral signals is performed by Chanel et al. at University of Geneva. This group used IAPS pictures as stimuli to assess arousal dimension in two levels (high and low) [7]. Another group of researchers used IAPS pictures as stimuli and assessed valence and arousal dimensions based on EEG signals [8]. In 2010, Koelstra et al. used emotional video films as stimuli and assessed emotions of five participants in two dimensions (arousal and valence) by using EEG and peripheral signals [9]. In 2011, Soleymani et al. prepared a public database, including EEG and peripheral signals of 32 participants, and used 40 video episodes as emotional stimuli. Then, they classified emotions in high and low levels of valence, arousal and liking dimensions using Bayes classifier [10]. In 2013, Mikhail and El-Ayat worked on brain signals produced from a new elicitation technique to detect emotion using minimal number of electrodes. They used SVM classifier to detect four emotional classes consisting of joy, anger, fear, and sadness [11]. Also, Liu et al. proposed a real-time Fractal Dimension (FD) based valence level recognition algorithm from EEG signals with 10 selected participants. They used DEAP dataset [12]. In 2014, Robert et al. worked on feature extraction methods for emotion recognition from EEG based on 33 subjects. Five different emotions (happy, curious, angry, sad, and quiet) were selected to cover a good part of the Valence-Arousal-Dominance (VAD) space. Emotions were induced using IAPS pictures. To evaluate and compare the proposed feature selection methods, classification was performed by means of Quadratic Discriminant Analysis (QDA) with diagonal covariance estimates (i.e., Naïve Bayes) [13]. Jirayucharoensak et al. worked on EEG signals to detect emotion in two dimensions: valence and arousal. They used Deep Learning Network (DLN) to discover correlations of

unknown features between input signals. DLN could classify three different levels of valence and arousal, and they enhanced classification accuracy by the principal component-based covariate shift adaptation [14].

The most important purpose of this paper is emotion recognition using peripheral and central nervous system activities and improvement of recognition accuracy of emotional states using combination mechanism. Emotions are assessed in high and low levels of valence, arousal, and liking dimensions. As mentioned earlier, most studies have focused on outer aspects of human emotion, such as facial expression and voice, to detect emotional states. On the other hand, physiological activities also include useful emotional information that can be used to assess emotions, but less attention was done to them. For this reason, this paper started its work with the analysis of physiological signals. EEG and peripheral signals are among the most important and the best means to recognize feelings. Most of the previous studies have used EEG signals for emotions assessment and few of them worked on peripheral signals. For this reason and considering the highly worthwhile emotional information on physiological signals, it was decided to use both EEG and peripheral signals and their combination. To find out the emotional information on physiological signals, it was attempted to extract the most efficient and suitable feature for physiological signals in time and frequency domains. Therefore, in addition to features mentioned in DEAP paper [10], some more efficient features, such as mobility, correlation coefficient, complexity, spectral entropy, spectral skewness, and spectral kurtosis, were extracted. Feature selection is an important issue in pattern recognition. Fisher linear discriminant is one of the best methods to select the most discriminant and appropriate features. In this work, many features were extracted from physiological signals, and the criterion of Fisher linear discriminant was used as an efficient method to reduce the computational load and processing time and also to increase accuracy of results. It should be noted that for testing and training classifiers, the Leave-One-Out Cross-Validation (LOOCV) method was used in order for results to be independent of individual. Although it was very difficult to implement this method due to very large dimension of processing data and too much analysis time on them, this method (LOOCV) was used for having better and independent results and also for a good comparison with papers of other researchers. Implementation of the method was successful. Some of the best classifiers (such as SVM, KNN, etc.) were selected; they are both suitable and efficient to classify emotional classes and are appropriate to be combined to form a new classifier. Combination was conducted at two levels: feature and classifier. Combination, in feature level, was done via concatenation mechanism

through algorithms of majority voting, stacking and fixed at classifier level. Fixed algorithms consist of minimum, maximum, mean, median, and product. Implementing combination at classifier level with the best methods is the most prominent work in this paper because it can improve results significantly.

2. Database

In DEAP database [10], 40 one-minute emotional videos were used for stimulating emotions. 32 participants, ranging in age from 19 to 37 years (16 people were women), participated in the experiment. At the end of each experiment, participants filled a questionnaire about their real feelings in each video. During watching each video, brain and peripheral signals of participants were recorded by 40 electrodes. In this database, there are a total of 1280 experiments with 8064 sample for each channel. From DEAP database, this three-dimensional dataset (i.e. $1280 \times 40 \times 8064$) was used in the paper for emotion recognition through arousal, valence, and liking dimension values. In each dimension, two classes consisting of high and low levels were considered.

3. Description of method

The most conventional scheme for detecting emotions is shown in Figure 1. The first step is the preprocessing of physiological signals that improves accuracy and processing speed in later steps. In the second step, extracting the features from these signals is done. The further the extracted features are, the longer the processing time will be. So, selecting the best features

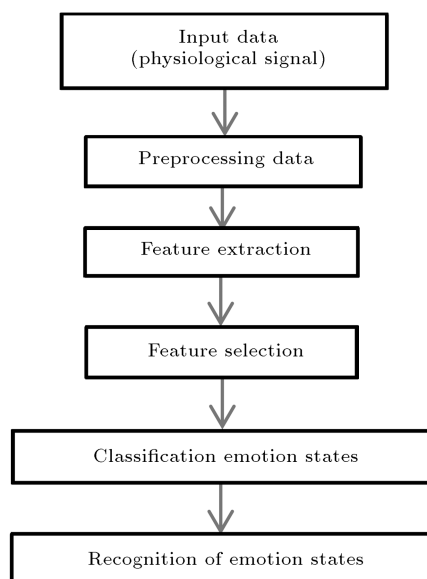


Figure 1. Framework for emotion detection by using physiological signals.

sounds crucial, considered in the third step. In the final step as the most important one, the classification is done using selected features and emotional classes are distinguished.

3.1. Feature extraction

Most current theories about emotions agree with this fact that physiological activities are important components of emotions. So, emotion recognition systems can recognize emotions by recording physiological signals, extracting appropriate features, and classifying using the best features. In addition to EEG, the following peripheral signals are considered (see Table 1):

- GSR is a physiological signal related to electrical resistance between two points on the skin. Secretion of sweat glands can decrease GSR average value. The mean value of the GSR signal is associated with emotion arousal [15]. Figure 2 shows an exemplary sample of the GSR signal;
- EMG signal is the generated electrical potential by muscle cells recorded by electrodes attached to the skin overlying the muscle. Emotional states appear on the face through its muscles activation; for example, laughing activates a Zygomaticus muscle, and Frontalis muscle activation is a sign of surprise or stress [16];
- BVP represents peripheral capillary blood volume in the body that is measured by plethysmograph or photoplethysmograph. Plethysmograph includes an infrared emitter and a detector. Amount of infrared light reflected from the surface of the skin is associated with peripheral capillary blood volume in the body [16]. HR signal can be obtained via peaks detection of BVP signal. By calculation of HR signal changes, the HRV signal can be obtained. HR and HRV signals change with feelings. For example, feelings of pleasure and satisfaction can increase the peak of HR signal [17], and HRV signal is reduced

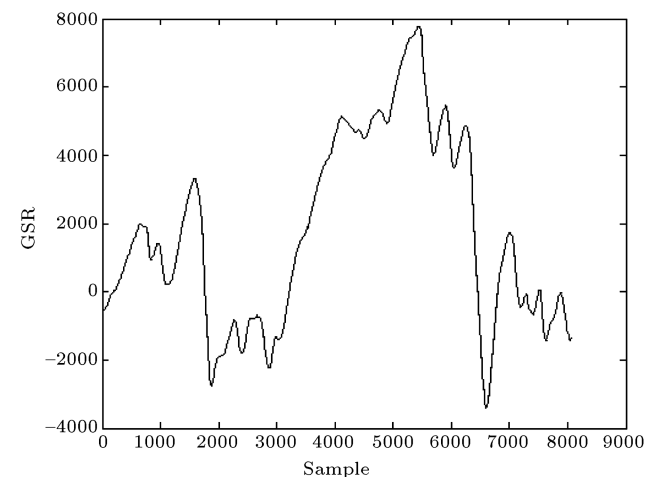


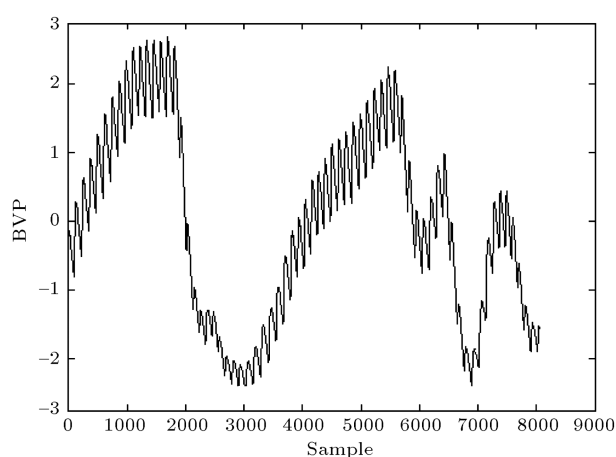
Figure 2. An exemplary sample of GSR signal.

Table 1. Extracted features from EEG and peripheral signals.

Signal	Extracted features
BVP	Average, standard deviation, mobility, complexity, correlation coefficient, frequency bands energy ratio (energy difference between [0.04-0.15] Hz and [0.1-0.2] Hz), frequency bands of spectral power ([0.1-0.2] Hz, [0.2-0.3] Hz, and then [0.3-0.4] Hz), spectral entropy (a mathematical measure of complexity, disorder or randomness degree of an event in a signal), and spectral skewness (asymmetry rate of the probability distribution function of the signal), spectral kurtosis (a measure of curve steepness at the maximum point).
RSP	Average, standard deviation, average of peak-to-peak time, mobility, median of peak-to-peak time, complexity, correlation coefficient, energy ratio of frequency bands ([0.04-0.15] Hz and [0.1-0.2] Hz), mean of derivative, dynamic range (difference between the slightest and greatest breaths), spectral centroid (mass center of signal in frequency domain), spectral entropy, 10 frequency bands of spectral power in [0-2.4] Hz, spectral kurtosis, and spectral skewness.
GSR	Average (average of derivative for values that are negative), mobility, proportion of samples that are negative in the derivative vs. all samples, complexity, number of local minima and average rising time in the GSR signal, correlation coefficient, spectral entropy, 10 frequency bands of spectral power in [0-2.4] Hz, spectral kurtosis, zero crossing rate of Galvanic Skin Slow Response (GSSR) [0-0.2] Hz, zero crossing rate of Galvanic Skin, Very Slow Response (GSVSR) [0-0.08] Hz, and spectral skewness.
STEMP	Average (average of STEMP signal derivative), mobility (standard deviation of signal derivative dividend into standard deviation of signal), complexity (mobility of signal derivative dividend into mobility signal), correlation coefficient, spectral entropy, frequency bands of spectral power ([0-0.1] Hz, [0.1-0.2] Hz), spectral skewness, and spectral kurtosis.
EMG and EOG	Energy of the signal, average and variance of the signal, mobility of signal, correlation coefficient, complexity of signal, spectral entropy, spectral kurtosis, and spectral skewness.
EEG	Spectral power in theta and slow alpha band, alpha band, beta and gamma band for each electrode, the spectral power asymmetry in between 14 pairs of electrodes in the theta, slow alpha, alpha, beta, and gamma bands.

with feelings of fear, sadness, and joy [18]. Figure 3 shows an exemplary sample of the BVP signal;

- Shallow breathing or breathing amplitude is obtained by measuring the expansion of chest or abdominal circumference. Using a flexible belt and

**Figure 3.** An exemplary sample of BVP signal.

a piezoelectric crystal sensor, the RSP signal is recorded. Slow breathing is related to relaxed state of feeling; irregular and rapid change and stopping of breathing is related to more aroused feelings such as anger or fear [17,18];

- Skin temperature (STEMP) also varies in accordance with emotional states [19]. Although temperature changes more slowly compared to other signals' changes, but it can also be an indicator of emotional responses [20]. Skin temperature is measured by placing a temperature sensor directly in contact with the skin of a person. Figure 4 shows an exemplary sample of the STEMP signal;
- Electrooculography (EOG) is a technique for measuring the electrical potential generated by the electrical activity of the eye muscles. Two pairs of electrodes are placed in both vertical and horizontal alignment (above, below, right, and left) that can measure the electrical potential difference between two points on the eye muscles [16]. Table 1 shows

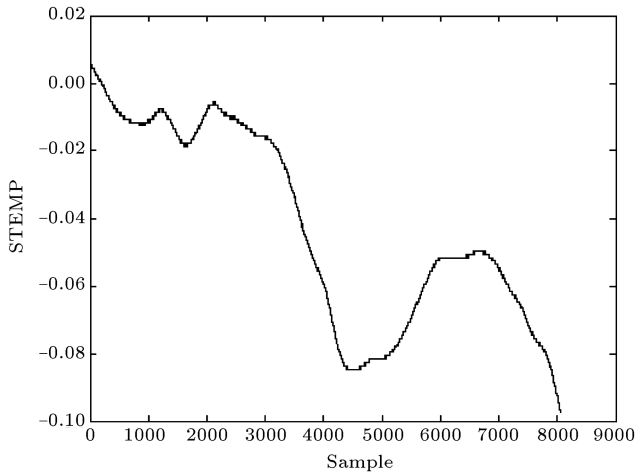


Figure 4. An exemplary sample of STEMP signal.

extracted features from these EEG and peripheral signals.

3.2. Feature selection

Many features are extracted from signals, but the computational load increases with feature number getting incremented. So, it is necessary to select the best discriminating features by an appropriate method. For this purpose, the Fisher linear discriminant criterion is used:

$$J(f) = \frac{|\mu_1 - \mu_2|}{\sigma_1^2 + \sigma_2^2}, \quad (1)$$

where for each feature f , μ_1 and μ_2 are the mean of class 1 and class 2; σ_1^2 and σ_2^2 stand for the variances of class 1 and class 2. Feature f is selected if $J(f)$ is more than 0.3 (this value was obtained empirically) [10]. Figure 5 shows steps of the feature selection method implemented in this paper.

3.3. Classification

After determining appropriate features and forming a feature vector, the classification of emotion states must be done. At first, classifiers are trained, and then they determine the class of an unknown feature vector. To evaluate the performance of classifiers, train and test data sets are determined by using Leave-One-Out (LOO) cross-validation scheme. Therefore, one from the data set is left as test data, and the rest of the data are used as training data. In the following, some of the best classifiers used in this work are described.

3.3.1. Support Vector Machine (SVM)

SVM works based on marginal hyper planes. This boundary, as shown in Figure 6, can be described using a linear equation. The method for finding such a marginal hyper plane (in the training step) is such that the plane should have the maximum distance or maximum boundary from the data of both classes. The hyper plane equation can be written in the following form:

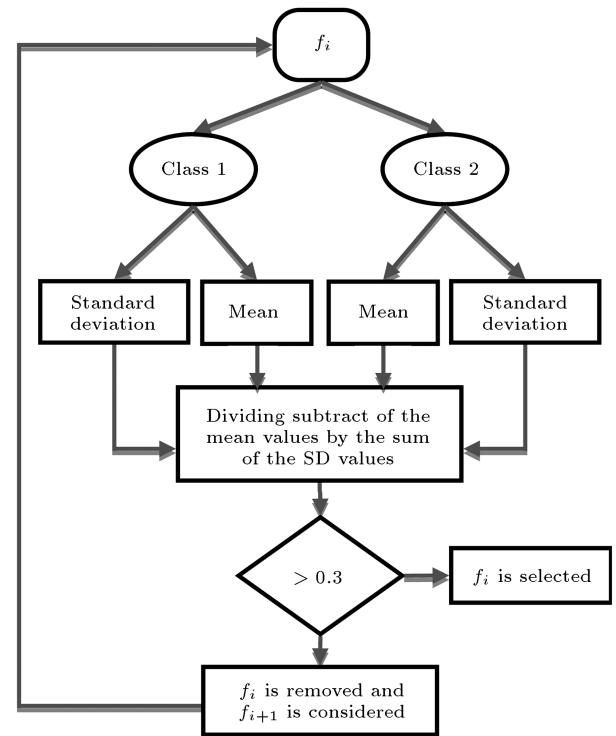


Figure 5. Flowchart of feature selection using the Fisher linear discriminant.

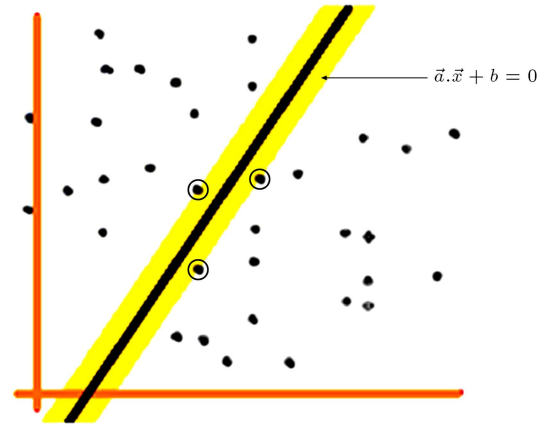


Figure 6. Hyperplane separator.

$$\vec{a} \cdot \vec{x} + b = 0, \quad (2)$$

where a is the perpendicular vector to the plane or weight vector, b is a threshold value, and x stands for feature vector. Discriminant function can be written for two classes as follows [21,23]:

$$H(\vec{x}) = \text{sgn}[\vec{a} \cdot \vec{x} + b] = \begin{cases} 1 & \text{class 1} \\ -1 & \text{class 2} \end{cases} \quad (3)$$

where $\text{sgn}(\cdot)$ denotes the mathematical sign function.

3.3.2. Nearest Mean (NM)

At first, NM classifier calculates the average of each class using training data; later, for determining the

class of unknown vector, f , its Euclidean distance to the average of each class is calculated. Then, the class of that average, which has less distance, is attributed to f . This distance is computed as follows:

$$d_i = f - \mu_i = \sum_{k=1}^d (f_k - \mu_{ik})^2, \quad (4)$$

where f_k is k th feature of feature vector f , μ_j is average of class i , μ_{jk} is k th component of mean vector, and d_i is Euclidean distance of feature vector f to i th class mean [22,24].

3.3.3. K-Nearest Neighborhood (K-NN)

For determining the class of feature vector f , at first, the classes of k nearest training data vectors (based on Euclidean distance criterion) are determined; then, the class, to which the majority of vectors belong, is defined as the class of vector f [24].

3.3.4. Linear Discriminant Analysis (LDA)

LDA is the simplest and most widely used statistical classification. LDA is a binary classifier in its standard form that is used in statistics, pattern recognition, and machine learning. The classifier, with assuming Gaussian model for conditional probability function and considering covariance matrix as equal, finds a linear combination of features that can specify or separate two or more classes [23].

3.3.5. Fisher's Linear Discriminant (FLD)

This classifier is based on portraying the classes; suppose that there are two classes and their data are two-dimensional. FLD finds a line in two-dimensional space by portraying data of two classes, thereby the two classes are separable from each other in an appropriate manner [25].

4. Combination at feature level

By combining the extracted features, a complete and comprehensive feature vector of emotional information can be achieved to improve the accuracy of emotion recognition; however, the computational complexity is increased, too. Combination at feature level, via the concatenation method, joins together N feature vectors (f_i) and makes a new feature set called F :

$$F = [f_1, f_2, f_3 \cdots f_i, f_{i+1} \cdots f_N]. \quad (5)$$

In this paper, the concatenation method, through putting extracted features from EEG and peripheral signals together, is implemented.

5. Combination at classifier level

5.1. Majority voting algorithm

Majority voting method only requires output labels of classifiers and does not require their rates of certainties.

Suppose that i th classifier output is c -dimensional binary vector $[d_{i,1}, d_{i,2}, \dots, d_{i,c}]$; if $d_{i,j} = 1$, input vector x belongs to j th class on classifier D_i ; if $d_{i,j} = 0$, input vector x does not belong to j th class based on classifier D_i . So, it can be seen that for each classifier and for each input x , only one element is equal to one and the rest are zero.

In the majority voting method, the final output label consists in the class that has gained the most votes (Eq. (6)). Number of base classifiers for such a method should be odd:

$$\text{if } \max_j \left\{ \sum_{i=1}^L d_{i,j} \right\} = \sum_{i=1}^L d_{i,k} \Rightarrow \text{lable}(x) = k, \quad (6)$$

where L is the number of base classifiers, c is the number of classes, and k refers to the final label of unknown vector x [26].

5.2. Fixed algorithms

In the fixed algorithms (minimum, maximum, mean, median, and product), it is necessary to evaluate outputs of all base classifiers together. Thus, they are transferred to range of $[0,1]$ in a way that their sum becomes one. In other words, if the output of i th classifier is vector $[S_{i,1}, S_{i,2}, \dots, S_{i,C}]$ (assuming L base classifiers and C different classes), then its new output can be written $[S'_{i,1}, S'_{i,2}, \dots, S'_{i,C}]$ as follows:

$$S'_{i,j} = \frac{\exp[S_{i,j}]}{\sum_{k=1}^c \exp[S_{i,k}]}, \quad \begin{cases} j = 1, 2, \dots, C \\ i = 1, 2, \dots, L \end{cases} \quad (7)$$

So, the result will be as follows:

$$\sum_{j=1}^c S'_{i,j} = 1, \quad 0 \leq S'_{i,j} \leq 1. \quad (8)$$

In addition, mean, maximum minimum, product, and median of the fixed algorithms will be as follows:

$$\mu_{j1}(x) = \frac{1}{L} \sum_{i=1}^L S'_{i,j}, \quad (9)$$

$$\mu_{j2}(x) = \max_i [S'_{i,j}], \quad (10)$$

$$\mu_{j3}(x) = \min_i [S'_{i,j}], \quad (11)$$

$$\mu_{j4}(x) = \prod_{i=1}^L S'_{i,j}, \quad (12)$$

$$\mu_{j5}(x) = \text{median}_i [S'_{i,j}], \quad (13)$$

where μ_i is a new certainty rate of unknown feature vector x for joining j th class, and S'_{ij} is certainty rate of unknown feature vector x for joining j th class at the view of i th classifier. Finally, output label of the feature vector is determined by maximum certainty rate [26]:

$$\text{if } \max_j \{\mu_j(x)\} = \mu_d(x) \Rightarrow \text{labe}(x) = d. \quad (14)$$

5.3. Stacking algorithm

In this method, multiple classifiers classify input feature vector separately, which is known as 0-level classifier. Then, the output of each classifier is entered into another classifier that makes the final decision. The latter classifier is known as 1-level classifier [26]. The stacking algorithm is illustrated in Figure 7.

6. Results

At first, emotion recognition accuracy is obtained using base classifiers and features extracted from EEG signals, peripheral signals, and their combination. Then, recognition accuracy is achieved using the combination of classifiers and the aforementioned features. Fig-

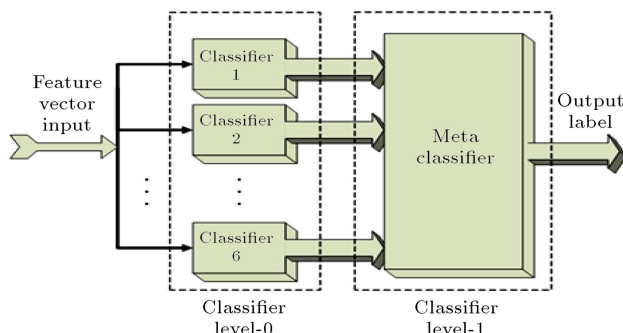


Figure 7. Combination of six classifiers using the stacking algorithm.

ures 8-13 show the results of classification for arousal, valence, and liking dimensions.

For the arousal dimension, by using EEG signals, the average of obtained accuracy from 6 base classifiers (accuracy of SVM, KNN, etc. is shown in Figure 8) is 57.20% and it is 63.93% for classifiers combination methods (accuracy of majority voting, stacking, etc. is shown in Figure 9). Subsequently, the average of obtained accuracy from base classifiers by using peripheral signals is 60.86%, and it is 64.16% for classifiers combination. Also, by using EEG and peripheral signals combination, the average of obtained accuracy from base classifiers is 57.30% and it is 63.30% for classifiers combination. Therefore, results show that, in the case of classifiers combination, emotion recognition accuracy has been a significant improvement by using EEG, peripheral signals, and combination of them. On the other hand, the above results show that, in the case of using base classifiers and classifiers combination, peripheral signals are more successful compared to EEG signals and signals combination (Tables A.1 and A.2 in the Appendix).

For valence dimension, by using EEG signals, the average of obtained accuracy from 6 base classifiers (whose accuracy is shown in Figure 10) is 55.48%, and it is 62.41% for classifiers combination methods (whose accuracy is shown in Figure 11). Subsequently, the average of obtained accuracy from base classifiers by using peripheral signals is 58.20%, and it is 60.05% for classifiers combination. Also, by using EEG and peripheral signals combination, the

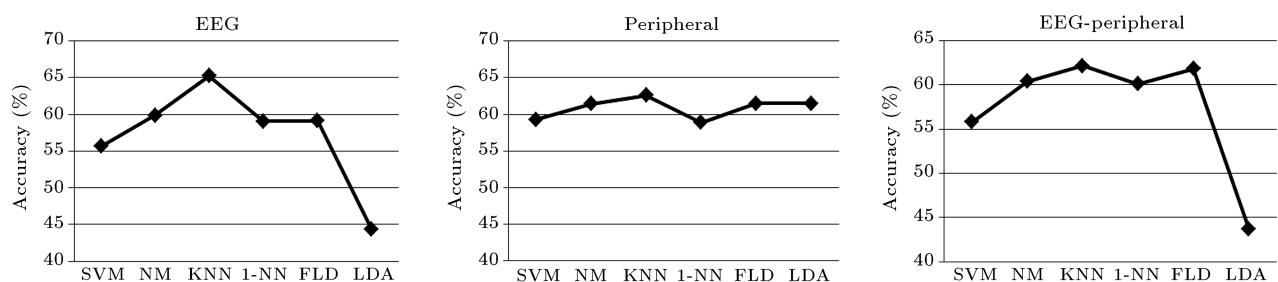


Figure 8. Arousal dimension results using base classifiers.

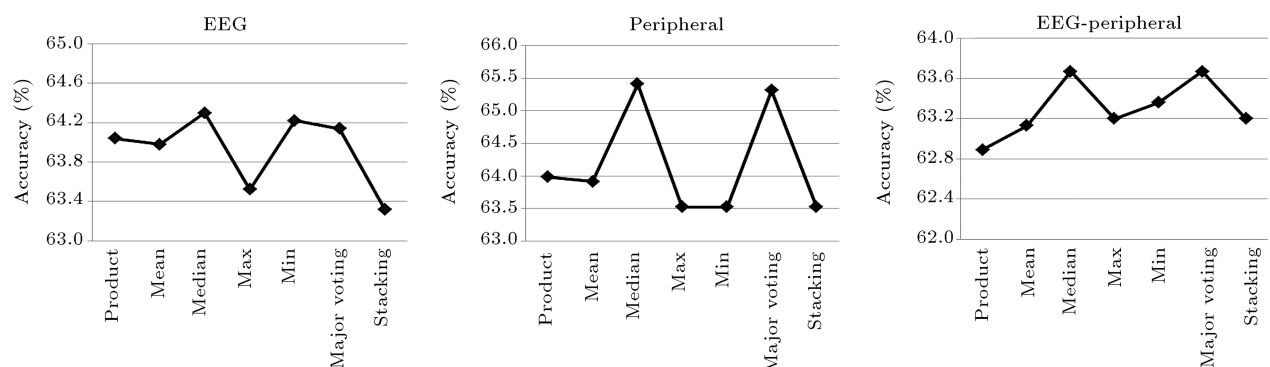


Figure 9. Arousal dimension results using classifiers combination.

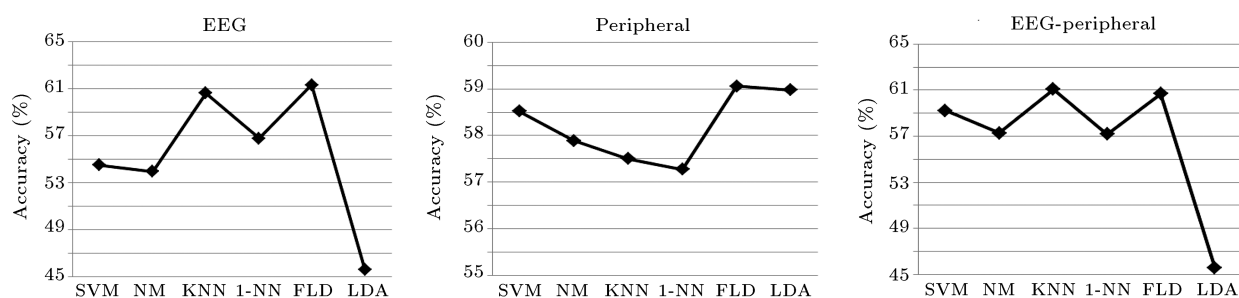


Figure 10. Valence dimension results using base classifiers.

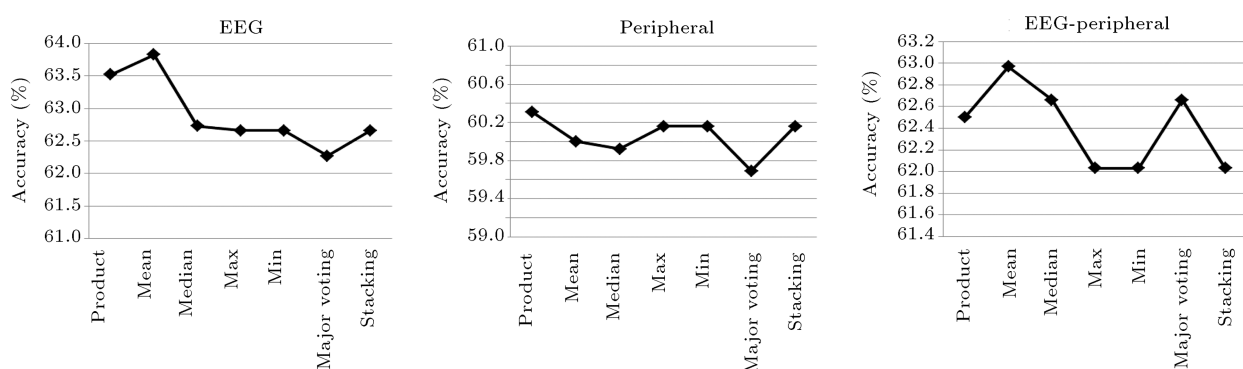


Figure 11. Valence dimension results using classifiers combination.

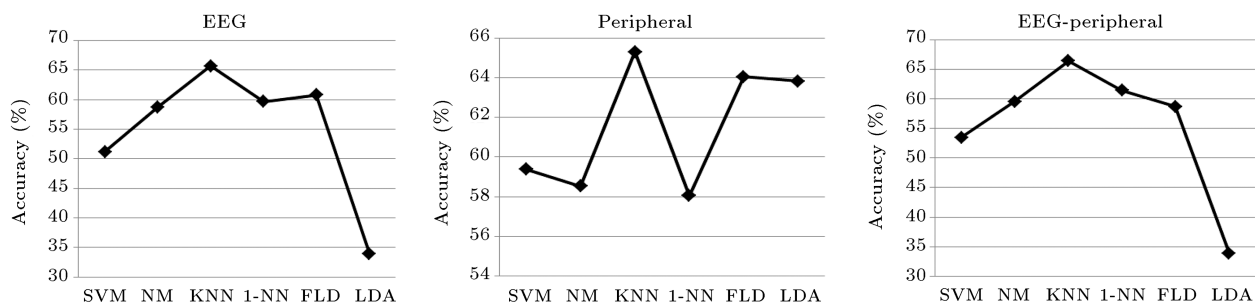


Figure 12. Liking dimension results using base classifiers.

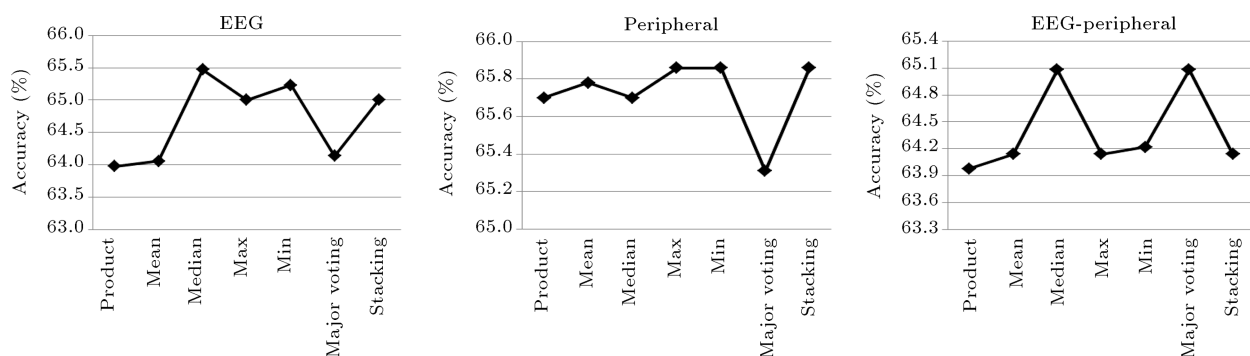


Figure 13. Liking dimension results using classifiers combination.

average of obtained accuracy from base classifiers is 56.83%, and it is 62.41% for classifiers combination. Therefore, results show that, in the case of classifiers combination, emotion recognition accuracy has been a significant improvement by using EEG, peripheral

signals, and their combination. On the other hand, the above results show that, in the case of using base classifiers, peripheral signals are more successful in emotion recognition; in the case of using classifiers combination method, EEG signals and signals combi-

nation present higher accuracy in emotion recognition compared to peripheral signals (Tables A.3 and A.4 in the Appendix).

For liking dimension, by using EEG signals, the average of obtained accuracy from 6 base classifiers (whose accuracy is shown in Figure 12) is 55.01%, and it is 64.69% for classifiers combination methods (whose accuracy is shown in Figure 13). Subsequently, the average of obtained accuracy from base classifiers by using peripheral signals is 61.52%, and it is 65.72% for classifiers combination. Also, by using EEG and peripheral signals combination, the average of obtained accuracy from base classifiers is 55.56%, and it is 64.39% for classifiers combination. Therefore, results show that, in the case of classifiers combination, emotion recognition accuracy has been a significant improvement by using EEG, peripheral signals, and their combination. On the other hand, the above results show that, in the case of using base classifiers and classifiers combination, peripheral signals are more successful compared to EEG signals and signals combination (Tables A.5 and A.6 in the Appendix).

7. Comparison of obtained results

In this section, a comparison of the reported results by the previous section curves will be performed for each dimension.

For arousal dimension, as shown in Figure 14, between six base classifiers, K-NN classifier is more successful in separating emotional classes with the average accuracy of 63.30%; between seven combination methods, the best recognition accuracy is also related to the median method with the average accuracy of 64.37% (Table A.7 in the Appendix).

For valence dimension, as shown in Figure 15,

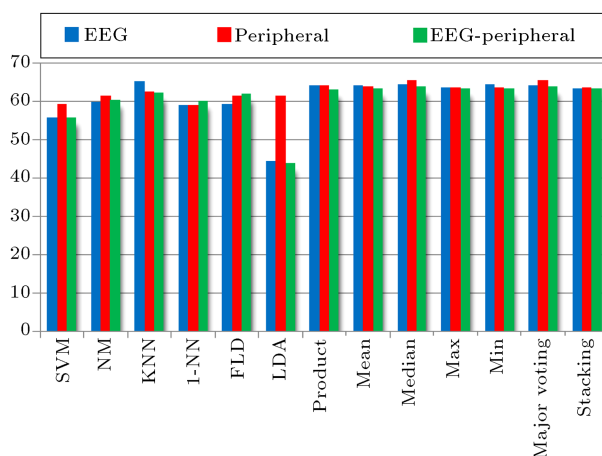


Figure 14. Arousal dimension-comparison of obtained results from extracted features using six classifiers and seven combination methods.

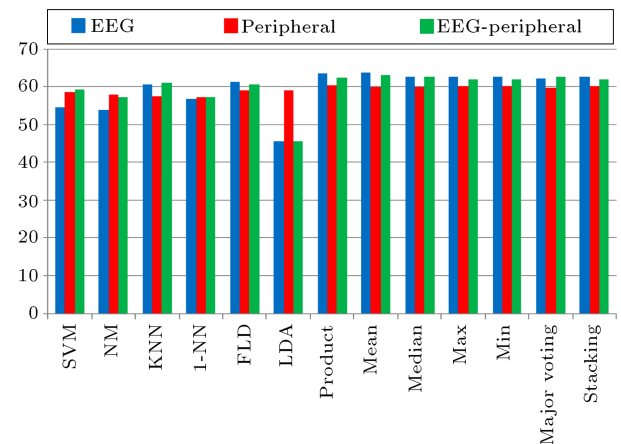


Figure 15. Valence dimension-comparison of obtained results from extracted features using six classifiers and seven combination methods.

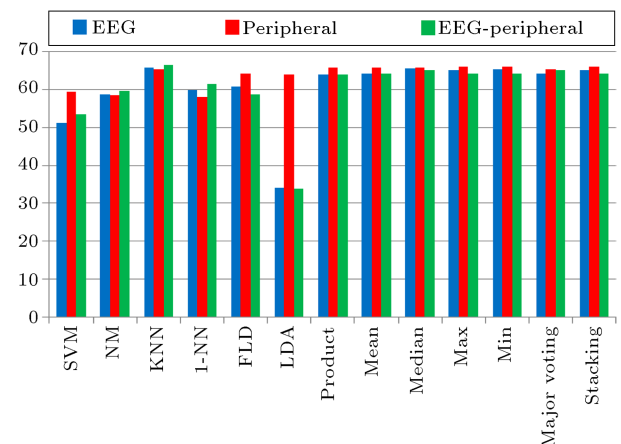


Figure 16. Liking dimension-comparison of obtained results from extracted features using six classifiers and seven combination methods.

between six base classifiers, FLD classifier is more successful in separating emotional classes with the average accuracy of 60.36%; between seven combination methods, the best recognition accuracy is also related to the mean method with average accuracy of 62.26% (Table A.8 in the Appendix).

For liking dimension, as shown in Figure 16, between six base classifiers, K-NN classifier is more successful in separating emotional classes with the average accuracy of 65.80%; between seven combination methods, the best recognition accuracy is also related to the median method with the average accuracy of 65.41% (Table A.9 in the Appendix).

Therefore, as Figures 14-16 show, with using base classifier method, the best recognition accuracy is related to KNN (for arousal and liking) and FLD (for valence) classifiers. On the other hand, the worst recognition accuracy is related to LDA classifier (for each 3 dimensions). But, with using classifiers

combination method, results significantly improve, and also they are more close to each other, and it can be said that they are independent of the type of combination method.

In Figures 17-19, the average results of six classifiers and seven combination methods for each of signals and their combination are, respectively, shown with

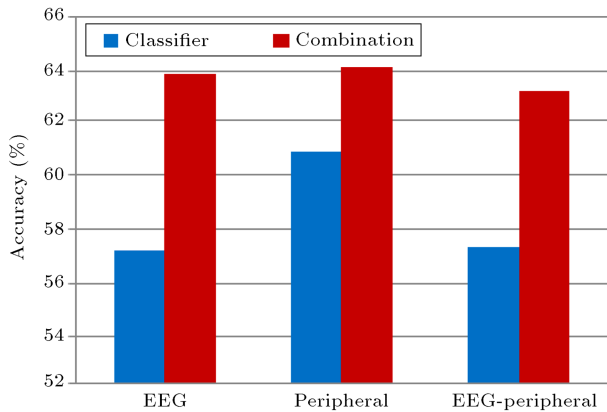


Figure 17. Arousal dimension-comparison of average results obtained from six classifiers and seven combination methods.

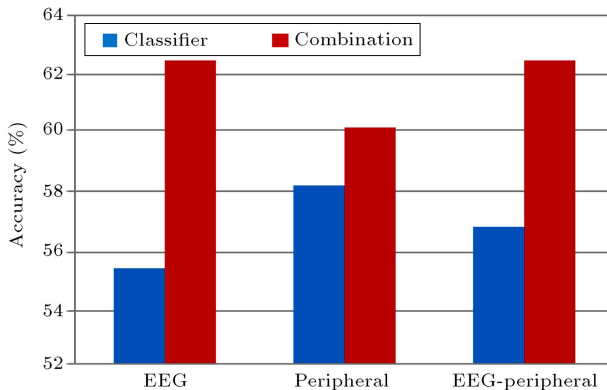


Figure 18. Valence dimension-comparison of average results obtained from six classifiers and seven combination methods.

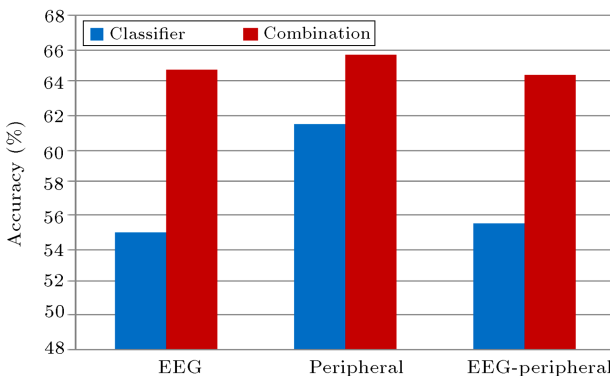


Figure 19. Liking dimension-comparison of average results obtained from six classifiers and seven combination methods.

blue and red rectangles. Comparison of the obtained results shows that:

- (i) Combination at classifier level has a tremendous impact on enhancing recognition accuracy;
- (ii) Peripheral signals are more successful in separating emotional classes;
- (iii) Combination at classifier level has the most effect on enhancing recognition accuracy by EEG signals;
- (iv) Emotion assessment in liking dimension is the easiest and the hardest in valence dimension (Tables A.10-A.12 in the Appendix).

8. Comparison with previous studies

In 2014, Suwicha [14] worked on EEG signals to detect emotion in valence and arousal dimensions. Their best accuracy rates were 55.07% and 52.56% for three different levels of valence and arousal dimensions, respectively. In comparison with the results obtained in [14], the best results of the present study are higher in accuracy rate; here, the results have been enhanced significantly by combination methods. Also, Robert et al. [13] studied feature extraction methods for emotion recognition from EEG. Five different emotions (happy, curious, angry, sad, and quiet) were selected to cover an important part of the Valence-Arousal-Dominance (VAD) space. Classification accuracy ranges from 25.0% to 47.5% dependent on subject and maximum average obtained was 36.8%. Therefore, accuracy rates in our work are higher than these rates.

In 2013, Mina Mikhail and Khaled EI-Ayat [11] worked on EEG signals to detect four emotional classes consisting of joy, anger, fear, and sadness; they reached average accuracy rates of 51%, 53%, 58%, and 61%, respectively. Therefore, the results of the present study, using combination of classifiers, are much better than their reported results. Also, Liu et al. [12] proposed a real-time Fractal Dimension (FD) based valence level recognition algorithm from EEG signals and obtained a mean accuracy of 63.04% for arousal-dominance recognition (four classes). They used the DEAP dataset, and it is noteworthy that the results of the present study, in classifier combination method, are higher than their work.

Tables 2 and 3 are also presented to compare the obtained results from EEG and peripheral signals with the proposed methods by Chanel et al. in 2009 [1], Koelstra et al. in 2011 [10], and Christy et al. in 2012 [27]. In [27], the authors used the DEAP database. The comparison of results shows that the proposed method in this study outperforms these previous studies.

Table 2. Comparison of obtained results with those of previous studies.

		Soleymani [10]	Chanel [1]	This work
Valence	EEG	57.6%	58%	63.83%
	Peripheral	61.7%	50%	60.31%
Arousal	EEG	62%	60%	65.23%
	Peripheral	57%	55%	65.31%
Liking	EEG	55.4%	—	65.7%
	Peripheral	59.1%	—	66.41%

Table 3. Comparison of obtained results with those of previous works studies.

EEG-Peripheral	Christy [27]	This work
Valence	62.19%	62.97%
Arousal	60.70%	63.30%

9. Conclusion

Emotion elicitation is a multi-modal process that involves several components of the organism. As a consequence, emotions are expressed through several channels, giving rise to many emotional cues that can be recorded by different sensors. Since the information recorded by those sensors can represent the activities of the different components involved in emotional processes; combining the information obtained from those sensors can improve the reliability of emotion assessment. Therefore, we decided to combine EEG and peripheral signals based on the concatenation method.

Six base classifiers (SVM, KNN, etc.) were used in order to evaluate their performance and accuracy and compared them with each other in classification of emotional classes for a common database. Then, preparing them to implement combination mechanism of classifiers was the next target. The best recognition accuracy among the base classifiers was related to KNN (for arousal and liking) and FLD classifiers (for valence). On the other hand, the worst recognition accuracy was related to LDA classifier (for each of three dimensions).

In addition to enhancing results and improving the quality of diagnosis, seven suitable combination methods were used until their obtained results were compared with each other and every one of six base classifiers. By using classifiers combination method, the results significantly improved, and also the obtained results from seven combination methods were similar to each other and their standard deviations were very low. On the contrary, the obtained results from six

base classifiers differed from each other. In addition, it is concluded that the obtained results from classifiers combination method are higher and independent of the type of combination method (product, major voting, etc.), but with using base classifier, results depend extremely on the type of classifier.

For valence, arousal and liking dimensions, when base classifiers (SVM, KNN, etc.) were used, peripheral signals were more successful compared to EEG signals; the combination of these signals led to better results. Classifiers combination significantly improved the recognition rates of emotions for EEG signals approaching recognition accuracy of peripheral signals or, in many cases, became more successful than peripheral signals.

Acknowledgement

We are grateful to Sajjad Hedayatipour, the expert of Khuzestan Regional Electric Co., for his sincere assistance with preparing the article.

References

1. Chanel, G. "Emotion assessment for affective computing based on brain and peripheral signals", Ph.D., Geneva, Switzerland (2009).
2. Savran, A., Ciftci, K., Chanel, G., Mota, C., et al. "Emotion detection in the loop from brain signals and facial images", *Proceedings of the eINTERFACE*, Dubrovnik, Croatia (2006).
3. Rolls, E.T., *The Brain and Emotion*, Oxford University Press (1998).
4. Jenkins, J.M., Oatley, K. and Stein, N.L., *Human Emotion: A Reader*, Blackwell Publisher, Malden, MA, Ed., 1 (1998).
5. Healey, J. and Picard, R. "Digital processing of affective signals", *Proc. Int. Conf. on Acoustics*, pp. 3749-3752 (1988).
6. Takahashi, K. "Remarks on emotion recognition from bio-potential signals", *Int. J. IEEE Trans, on Autonomous Robots and Agents*, pp. 186-191 (2004).

7. Chanel, G., Kronegg, J. and Pun, T. "Emotion assessment: Arousal evaluation using EEG's and peripheral physiological signals", *Proc. Int. Workshop MRCIS*, **4105** (2006).
8. Horlings, R., Datcu, D. and Rothkrantz, L.J.M. "Emotion recognition using brain activity", *Proc. Int. Conf. on Computer Systems and Technologies* (2008).
9. Koelstra, S., Yazdani, A. and Soleymani, M. "Single trial classification of EEG and peripheral physiological signals for recognition of emotions induced by music videos", *Proc. Int. Conf. on Brain Informatics, Computer Science*, Springer, **6**, pp. 89-100 (2010).
10. Koelstra, S., Muehl, C. and Soleymani, M. "DEAP: A database for emotion analysis using physiological signals", *Int. J. IEEE Trans. Affective Computing*, **3**, pp. 18-31 (2011).
11. Mikhail, M., El-Ayat, Coan, J.A. and Allen, J.J.B. "Using minimal number of electrodes for emotion detection using brain signals produced from a new elicitation technique", *Int. J. Autonomous and Adaptive Communications Systems*, **6**(1), pp. 80-97 (2013).
12. Liu, Y., Sourina, O. and Nguyen, M.K. "Real-time EEG-based emotion recognition and its applications", *Transactions on Computational Science XII*, **6670**, pp. 256-277 (2013).
13. Robert, J., Angelika, P. and Martin, B. "Feature extraction and selection for emotion recognition from EEG", *IEEE Transactions on Affective Computing*, **5**, pp. 327-339 (2014).
14. Jirayucharoensak, S., Pan-Ngum, S., and Israsena, P. "EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation", *Hindawi Publishing Corporation, The Scientific World Journal*, Article ID 627892, 10 (2014).
15. Dawson, M.E., Schell, A.M. and Filion, D.L., *The Electrodermal System*, (Ed.) New York, NY, US: Cambridge University Press, pp. 200-223 (2000).
16. Soleymani, M. "Implicit and automated emotional tagging of videos", PhD, Geneve (2011).
17. Rainville, P., Bechara, A., Naqvi, N. and Damasio, A.R. "Basic emotions are associated with distinct patterns of cardiorespiratory activity", *Int. J. of Psychophysiology*, **61**, pp. 5-18 (2006).
18. Kim, J. and André, E. "Emotion recognition based on physiological changes in music listening", *Pattern Analysis and Machine Intelligence, IEEE Trans.*, **30**, pp. 2067-2083 (2008).
19. McFarland, R.A. "Relationship of skin temperature changes to the emotions accompanying music", *Applied Psychophysiology and Biofeedback*, **10**, pp. 255-267 (1985).
20. Kreibig, S.D. "Autonomic nervous system activity in emotion: A review", *Biological Psychology*, **84**, pp. 394-421 (2010).
21. Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F. and Arnaldi, B. "A review of classification algorithms for EEG-based brain-computer interfaces", *Int. J. of Neural Engineering*, **4**, pp. R1-R13 (2007).
22. Sapankevych, N. and Sankar, R. "Time series prediction using support vector machines: A survey", *Computational Intelligence Magazine, IEEE*, **4**, pp. 24-38 (May, 2009).
23. Theodoridis, S. and Koutroumbas, K., *Pattern Recognition*, Second Edition, Academic Press (April 9, 2003).
24. Theodoridis, S. and Koutroumbas, K. "Pattern recognition and neural networks", *Machine Learning and Its Applications*, pp. 169-195 (2001).
25. Shashua, A. "On the relationship between the support vector machine for classification and sparsified Fisher's linear discriminant", *Neural Processing Letters*, **9**, pp. 129-139 (2000).
26. Kuncheva, L.I., *Combining Pattern Classifiers: Methods and Algorithms*, John Wiley & Sons, Inc. Hoboken, New Jersey (2004).
27. Christy, T., Kuncheva, L.I. and Williams, K.W. "Selection of physiological input modalities for emotion recognition", Technical Report, Bangor University, UK (2012).

Appendix

Tables A.1 and A.2 are numeric values of Figures 8, 9, and 14; Tables A.3 and A.4 are numeric values of Figures 10, 11, and 15; Tables A.5 and A.6 are numeric values of Figures 12, 13, and 16.

Tables A.7-A.9 are average of obtained accuracy of EEG, peripheral, and EEG-peripheral signals for six base classifiers and seven combination methods.

Tables A.10-A.12 are numeric values of Figures 17, 18, and 19. They are comparison of average results obtained from six classifiers and seven combination methods for EEG, peripheral, and EEG-peripheral signals.

Table A.1. Arousal dimension results using base classifiers.

Classifier	EEG	Peripheral	EEG-peripheral
SVM	55.63%	59.30%	55.78%
NM	59.84%	61.41%	60.39%
K-NN	65.23%	62.58%	62.11%
1-NN	59.06%	58.91%	60.08%
FLD	59.14%	61.48%	61.80%
LDA	44.30%	61.48%	43.67%
Average	57.20%	60.86%	57.30%

Table A.2. Arousal dimension results using classifiers combination.

Combination method	EEG	Peripheral	EEG-peripheral
Product	64.04%	63.98%	62.89%
Mean	63.98%	63.91%	63.13%
Median	64.30%	65.41%	63.67%
Max	63.52%	63.52%	63.20%
Min	64.22%	63.52%	63.36%
Major voting	64.14%	65.31%	63.67%
Stacking	63.32%	63.52%	63.20%
Average	63.93%	64.16%	63.30%

Table A.3. Valence dimension results using base classifiers.

Classifier	EEG	Peripheral	EEG-peripheral
SVM	54.53%	58.52%	59.22%
NM	53.98%	57.89%	57.27%
K-NN	60.63%	57.50%	61.09%
1-NN	56.80%	57.27%	57.19%
FLD	61.33%	59.06%	60.70%
LDA	45.63%	58.98%	45.55%
Average	55.48%	58.20%	56.83%

Table A.4. Valence dimension results using classifiers combination.

Combination method	EEG	Peripheral	EEG-peripheral
Product	63.52%	60.31%	62.5%
Mean	63.83%	60.00%	62.97%
Medial	62.73%	59.92%	62.66%
Max	62.66%	60.16%	62.03%
Min	62.66%	60.16%	62.03%
Major voting	62.27%	59.69%	66.62%
Stacking	62.66%	60.16%	62.03%
Average	62.41%	60.05%	62.41%

Table A.5. Liking dimension results using base classifiers.

Classifier	EEG	Peripheral	EEG-peripheral
SVM	51.17%	59.38%	53.44%
NM	58.75%	58.52%	59.53%
K-NN	65.70%	65.31%	66.41%
1-NN	59.69%	58.05%	61.41%
FLD	60.78%	64.06%	58.67%
LDA	33.98%	63.83%	33.91%
Average	55.01%	61.52%	55.56%

Table A.6. Liking dimension results using classifier combination.

Combination method	EEG	Peripheral	EEG-peripheral
Product	63.98%	65.70%	63.98%
Mean	64.06%	65.78%	64.14%
Medial	65.47%	65.70%	65.08%
Max	65.00%	65.86%	64.14%
Min	65.23%	65.86%	64.22%
Major voting	64.14%	65.31%	65.08%
Stacking	65.00%	65.86%	64.14%
Average	64.69%	65.72%	64.39%

Table A.7. Arousal dimension.

Classifier & combination method	Average recognition accuracy
SVM	56.90%
NM	60.54%
KNN	63.30%
1NN	59.35%
FLD	60.80%
LDA	49.81%
Product	63.63%
Mean	63.67%
Median	64.46%
Max	63.41%
Min	63.70%
Major voting	64.37%
Stacking	63.34%

Table A.8. Valence dimension.

Classifier & combination method	Average recognition accuracy
SVM	57.42%
NM	56.38%
KNN	59.74%
1NN	57.08%
FLD	60.36%
LDA	50.05%
Product	62.11%
Mean	62.26%
Median	61.77%
Max	61.61%
Min	61.61%
Major voting	61.86%
Stacking	61.61%

Table A.9. Liking dimension.

Classifier & combination method	Average recognition accuracy
SVM	54.66%
NM	58.93%
KNN	65.80%
1NN	59.71%
FLD	61.17%
LDA	43.90%
Product	64.55%
Mean	64.66%
Median	65.41%
Max	65%
Min	65.10%
Major voting	64.84%
Stacking	65%

Table A.10. Arousal dimension-comparison of average results obtained from six classifiers and seven combination methods.

	EEG	Peripheral	EEG- peripheral
Base classifier	57.2%	60.86%	57.3%
Classifier combination	63.93%	64.16%	63.3%

Table A.11. Valence dimension-comparison of average results obtained from six classifiers and seven combination methods.

	EEG	Peripheral	EEG- peripheral
Base classifier	55.48%	58.2%	56.83%
Classifier combination	62.41%	60.05%	62.41%

Table A.12. Liking dimension-comparison of average results obtained from six classifiers and seven combination methods.

	EEG	Peripheral	EEG- peripheral
Base classifier	55.01%	61.52%	55.56%
Classifier combination	64.69%	65.72%	64.39%

Biographies

Marzieh Amjadzadeh was born in Ahvaz, Iran, in 1988. She received BSc and MSc degrees in Electronic Engineering from University of Shahid Chamran, Ahvaz, Iran, in 2010 and 2014, respectively. Her current research interests include signal processing, image processing, intelligent transportation system, and machine learning.

Karim Ansari-Asl received his BSc degree in Electronic Engineering from Semnan University, Semnan,

Iran, in 1995, MSc degree in Biomedical Engineering from Iran University of Science and Technology (IUST), Tehran, Iran, in 1999, and PhD degree in the Biomedical Signal Processing from University of Rennes 1, Rennes, France, in 2005.

From 2005 to 2007, he was a post-doctoral fellow at University of Rennes 1 and University of Geneva. Since 2008, he has been an Assistant Professor with the Electrical Engineering Department, Shahid Chamran University, Ahvaz, Iran. His research interests include biomedical engineering, digital signal, and image processing.