WILEY

Emotion detection from EEG recordings based on supervised and unsupervised dimension reduction

Jingxin Liu | Hongying Meng | Maozhen Li | Fan Zhang | Rui Qin | Asoke K. Nandi

Department of Electronic and Computer Engineering, Brunel University London, Uxbridge, UB8 3PH, UK

Correspondence
Hongying Meng, Department of Electronic and Computer Engineering, Brunel University London, Uxbridge, UB8 3PH, UK.
Email: hongying.meng@brunel.ac.uk

Summary

In recent years, researchers have been trying to detect human emotions from recorded brain signals such as electroencephalogram (EEG) signals. However, due to the high levels of noise from the EEG recordings, a single feature alone cannot achieve good performance. A combination of distinct features is the key for automatic emotion detection. In this paper, we present a hybrid dimension feature reduction scheme using a total of 14 different features extracted from EEG recordings. The scheme combines these distinct features in the feature space using both supervised and unsupervised feature selection processes. Maximum Relevance Minimum Redundancy (mRMR) is applied to re-order the combined features into max-relevance with the labels and min-redundancy of each feature. The generated features are further reduced with principal component analysis (PCA) for extracting the principal components. Experimental results show that the proposed work outperforms the state-of-art methods using the same settings in the publicly available DEAP data set.

KEYWORDS

affective computing, EEG, emotion detection, feature dimension reduction, feature selection

1 | INTRODUCTION

Human emotion is a complex phenomenon that comes from human brain, but there is no clear knowledge on its generation mechanism. Physiologists and computer scientists have been studying it for decades. For example, Ekman et al[1] proposed the notion of 6 basic emotions that were universal and found across cultures. Posner et al[2] proposed a two-dimensional model in which emotions were given coordinates denoting the degree of valence (the positive or negative quality of emotion) and arousal (how responsive or energetic the subject is). Other models include the Plutchik wheel of emotions,[3] a tree of emotions,[4] etc.

In relation to engineering applications, the definition and classification of emotions are of importance in deciding what variables should be considered and what measurements are required during the design of a system. A person's emotions can be gauged from external features such as facial expressions and the tone of voice. These can be captured via photographs, video, and audio recordings. However, emotion also results in other bodily changes such as variation in heart rate, muscle tension, respiration and skin conductance, and, of course, brain activity.

Emotion detection through electroencephalogram (EEG) signals has a wide variety of practical applications, evidenced by medicine and scientific research, and the field of affective computing.[5] The latter refers to the incorporation of emotions in human-computer interaction giving machines a degree of emotional intelligence. It has been proposed, for the use of these machine learning systems, to include multimedia environments that recognize the emotions of the users such as recommendation and tagging systems, games and films that respond to the user emotions, and biofeedback devices that can be worn in the manner of headsets and might help users gain control over their emotional states. This echoes objections made at a much earlier stage by Picard et al[5] to emotion recognition based on external displays. Among these are that emotions are frequently concealed or masked or even unknown to the subjects themselves, the disparity between posed and spontaneous emotion, the practical and ethical obstacles to recording spontaneous emotion, as well as matters that can limit applications such as privacy in relation to video-based detection. They recorded a

variety of peripheral physiological signals including heart rate, respiration, skin conductance, blood volume pressure, and facial muscle tension and achieved 81% accuracy, which was amongst the highest at that time.

In this paper, a new emotion detection system is proposed. Firstly, multiple feature extraction methods are used to produce different types of features from different domains. Secondly, we applied a new hybrid dimension feature reduction scheme, which used the combination of supervised and unsupervised reduction methods to fuse the different features in order to get the best feature. Advanced machine learning methods are used and evaluated on a public available dataset DEAP (Database for Emotional Analysis using Physiological Signals).[6] Experimental results are given on all different features and different feature selection methods for the emotion information extraction from EEG signals. It is compared with other state-of-the-art methods at the same setting up on the public DEAP dataset.

The rest of the paper is organized as follows. Section 2 gives a review on related work. The proposed method is introduced in details in Section 3. Section 4 presents the experimental results, and section 5 gives the conclusion.

2 | RELATED WORK

An emotion detection system from EEG signals can be treated as a classification problem since the goal of the system is to predict the correct label of emotion. It is thus often a supervised learning task since labels are already assigned to the data by humans, although clustering methods have also been employed.[7] Detailed information about the current research in this area can be found in the work of Alarcao and Fonseca.[8]

An important part of the study of emotion via machine learning involves the choice of features. Researchers have made use of a variety of features from EEG recordings. Jenke et al[9] made a survey on feature selection and extraction across a variety of studies and classified these as time domain, frequency domain, time-frequency domain, and multi-electrode features. Time domain features include event-related potentials, signal statistics, Hjorth features, non-stationary index, fractal dimension, and higher-order crossings. Frequency domain features include band power and higher-order spectra, time-frequency domain features include the Hilbert-Huang spectrum and discrete wavelet transforms, and multi-electrode features include magnitude-squared coherence estimate and differential and rational asymmetries. Frequency domain features are prevalent and appear in the majority of the studies surveyed in the paper, in particular spectral power, but it was also found that its performance scores is lower compared to other features.

A very high level of performance was achieved in the study by Valenzi et al[7] who analyzed EEG data from nine participants in response to video stimuli intended to induce the emotional states of amused, disgusted, sad, and neutral. A key difference in the use of video stimuli in this study was that between stimuli, a distraction task rather than a relaxation task was used to neutralize the emotional state of the participant and was considered to be more effective than a relaxation task. Data were recorded from 32 electrodes. The features extracted from the EEG were spectral power in delta (0.16-4 Hz), alpha (8-13 Hz), lower beta (14-21 Hz), upper beta (21- 30 Hz), and gamma (30-40 Hz) bands. A linear discriminant analysis was used to reduce the dimensionality of the feature space. Both supervised and unsupervised learning methods were used. Supervised learning methods were the error back propagation and support vector machine (SVM). Unsupervised learning algorithms used were vector quantization, fuzzy c-means clustering (FCM), k-means, and k-medians. A maximum average accuracy of 97.2% was achieved for supervised learning (for SVM) and a maximum average accuracy of 95.2% was achieved for unsupervised learning (for FCM). The average EEG power was computed across the stimuli for the different electrodes and showed larger frontal right symmetry for negative emotions. An electrode reduction was attempted, using only 8 electrodes (6 frontal and 2 temporal), yielding a best rate (using SVM) of 92.5% for individual classification and an average classification rate of 87.5%. It was noted, however, that the method in its current state was designed only to work offline.

Using all the features, an average classification accuracy of 87.5% was achieved when features from all the bands were used. For individual bands, higher frequency alpha, beta and gamma bands yielded better results compared with lower frequency bands. After the feature reduction, the classification accuracy, in fact, increased slightly when using the top one hundred subject-independent features and was at 89.2%. It was noted that the selected subject independent features were mainly in the higher frequency bands, and this was consistent with studies relating human emotional response to these bands.

Noticed by Othman et al,[10] one possible application for this area is intervention in cases of brain developmental disorders such as ADHD and autism. Their study involved 5 child participants, who were shown emotional faces whilst EEG recordings were made. Two different dimensional models were used for the classification of the emotions known as rSASM and 12-PAC. Recordings were taken only from the F3 and F4 frontal electrodes, and only the theta and alpha bands were considered. For the purposes of feature extraction, Mel-frequency cepstral coefficients and kernel density estimation were used, and a multi-layer perception was used for classification. The best performance achieved was for the 12-PAC with Kernel Density Estimation, at a mean squared error range (among the participants) of 0.07 to 0.09.

Another study used the connectivity between the EEG channels. Chen et al[11] considered the relationship between each channel, and calculated the mutual information, person correlation coefficient,[12] and phase coherence connectivity[13] between each channels and use these information as the extracted features and then used the Fisher linear discriminant method as the feature selection method, which is at the same setting with the one in the DEAP paper dataset.[6] They use an SVM as the classification method and obtained 76% on valence and 73% on arousal. Meanwhile, Gupta et al[14] also applied connectivity features for the DEAP dataset.[6] This method is inspired by the work of Pablo et al.[15] For their work, Welchs t-test and PCA used the two-step feature reduction methods, which applied for the fusion of spectral power and mutual information. At the DEAP

dataset, a total of 880 features were extracted from 32 channels. For Welch's t-test, they used 0.01 increase step as the threshold, and best features were selected. Then, they applied PCA for the second step reduction. In order to reduce the effect of different classifiers, they used SVM-RBF, SVM-Sigmoid, and Naive Bayes as the classification methods. The presented results show that arousal reaches 67.7 ± 11.3% and valence reaches 69.6 ± 9.3% for the best classification accuracy.

Inspired by the work of Arnau-González et al,[15] we make the further improvements: 1) we produce more different types of features, which potentially contain more information for different domain; 2) the mRMR method is used not only to consider the features significantly different from each other but also to keep the most useful information for the classification of the relationship between features and the labels.

3 | PROPOSED METHOD

3.1 | Overview

One of the most important factors that influence the performance of classification is the selection from different types of highly dimensional features of different classes. Certainly, the original features contain more information that can be fed into the classification. However, it will potentially decrease the positive impact when all types of features are used simultaneously. In this case, we applied the fusion work on the different types of features combined, which fused all types of features with the combination of supervised and unsupervised feature reduction method to achieve the best performance. As shown in Figure 1, we applied 14 different kinds of feature extraction methods from different domains in order to produce more information form the EEG data, and then, as described above, the feature reduction method is applied for these features combined in order to keep the most value information for further work. After the reduction work, we feed the valuable features into the different classifiers to do the binary classification work. The feature extraction methods and reduction methods are shown below.

3.2 | Feature extraction

A wide range of features were used for emotion recognition from EEG that have been proposed in the past. We generally distinguish the feature extraction methods into time domain, frequency domain, time-frequency domain, multi-electrode features, and connectivity features. Overall, 14 different features have been used.

1) *Time domain features:*
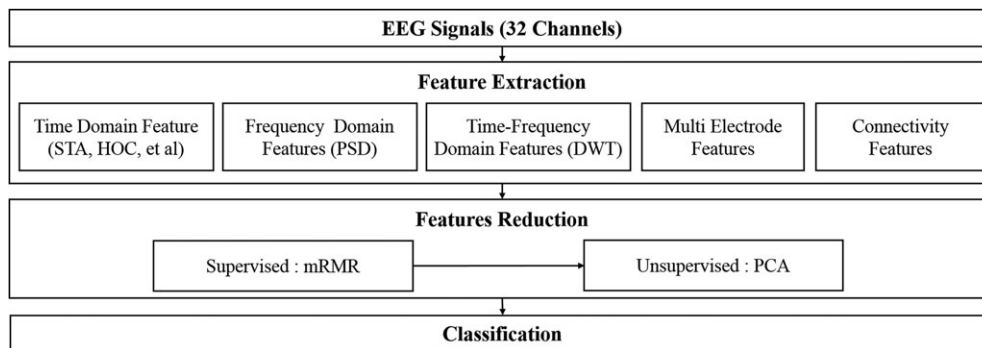
- Statistics features (STA)

There are 7 different features proposed in this method. They are straightforward to be calculated according to the formulas given in the work of Jenke et al,[9] for the signal $S(t), t = 1, 2, 3, \ldots, T$ as shown below:
Power:

$$P_S = \frac{1}{T} \sum_{-\infty}^{\infty} |S(t)|^2; \tag{1}$$

Mean:

$$\mu_S = \frac{1}{T} \sum_{t=1}^{T} S(t); \tag{2}$$

EEG Signals (32 Channels)				
Feature Extraction				
Time Domain Feature (STA, HOC, et al)	Frequency Domain Features (PSD)	Time-Frequency Domain Features (DWT)	Multi Electrode Features	Connectivity Features
Features Reduction				
Supervised : mRMR		Unsupervised : PCA		
Classification				

FIGURE 1 Overview of the proposed method. Multi-channel EEG signals are sent for distinct feature extraction, and the hybrid feature dimension reduction scheme is applied for emotion detection based on classification

Standard deviation:

$$\sigma_S = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(S(t) - \mu_S)^2}; \tag{3}$$

First difference:

$$\delta_S = \frac{1}{T-1}\sum_{t=1}^{T-1}|S(t+1) - S(t)|; \tag{4}$$

Normalized first difference:

$$\bar{\delta} = \frac{\delta_S}{\sigma_S}; \tag{5}$$

Second difference:

$$\gamma_S = \frac{1}{T-1}\sum_{t=1}^{T-2}|S(t+2) - S(t)|; \tag{6}$$

Normalized second difference:

$$\bar{\gamma} = \frac{\gamma_S}{\sigma_S}. \tag{7}$$

- Higher-order crossings (HOC)

The aim of the HOC feature is to try to capture the oscillatory pattern of the EEG waveform. The crossings are calculated by subtracting from the mean from the time series and then counting the number of sign changes. It is calculated only for the alpha and beta range, so the signal is first filtered through a tenth-order Butterworth band-pass filter. The highest order for which the number of crossings was calculated was 10. The first order is the original signal. For subsequent orders, the new signal is obtained by taking the difference between the consecutive points of the previous signal, and the number of crossings is then computed for this signal. When taking a difference, one point is lost, so, in order to retain the same number of points at each level, it is necessary to start (in this case) 10 points from the beginning of the signal.[16]

- Fractal dimension (FD)

This feature also seeks to capture information about the shape of the signal. The formal way of defining dimension is to consider the scaling relationship between units of measurement and the number of such units required to measure a shape, as shown in the following equation:

$$N \propto \epsilon^{-D}. \tag{8}$$

In the equation above, ϵ denotes the amount by which the unit of measurement is increased or reduced, N denotes the number of the newly scaled units of measurement required to measure the same shape, and D is the fractal dimension. For an ordinary line, reducing a unit of measurement by ϵ would mean that $N = 1/\epsilon$, so that $D = 1$. On the other hand, a fractal line reveals a higher degree of complexity at a higher resolution, so that more than $1/\epsilon$ units would be required every time the unit of measurement is reduced by ϵ. This complexity is quantified in the fractal dimension, which, for a line, is greater than 1.

The method used in the work of Jenke et al[9] was the Higuchi algorithm that is described in more detail in the work of Higuchi.[17] To use the method for the signal $S(t)$, a new time series is constructed as follows:

$$S_k^m = S(m), S(m + k), S(m + 2k), \ldots, S\left(m + floor\left[(N-m)/k\right] * k\right) \qquad m = 1, 2, \ldots, K, \tag{9}$$

where the function $floor[.]$ rounds down the value of the argument to the nearest integer. The length of the curve is then given by

$$L_m(k) = \left\{ \left(\sum_{i=1}^{floor\left[\frac{N-m}{k}\right]} |S(m + ik) - S(m + (i-1) * k)| \right) \frac{N-1}{floor\left[\frac{N-m}{k}\right] * k} \right\} / k \tag{10}$$

In the equation above, the term $\frac{N-1}{floor[\frac{N-m}{k}] * k}$ is a normalization factor for the curve length of subset time series. For each time interval k, the value of the length of the curve is averaged over all values of m to obtain the value $(L(k) \propto k^{-D})$, which is analogous to $N \propto \epsilon^{-D}$ above. We can estimate the

fractal dimension D by calculating $L(k)$ for different values of k, plotting it against k, and fitting a line to the points. The magnitude of the gradient of this line is the estimate for D.

- Hjorth feature (Hjorth)

These are simple statistical features computed using the following expressions, for the signal $S(t)$, where $\mu(S(t))$ is its gradient:
Mobility:

$$\sqrt{\frac{\sigma(\dot{S}(t))}{\sigma(S(t))}};\tag{11}$$

Complexity:

$$\sqrt{\frac{\mu(\dot{S}(t))}{\mu(S(t))}}.\tag{12}$$

As in the work of Jenke et al,[9] an additional feature, Activity' omitted because it is just square of the standard deviation (ie, the variance) and the standard deviation is already included among the statistical features above.

- Non-stationarity index (NSI)

NSI is a measure of fluctuation dynamics that is used to evaluate the change in time of the local average,[18] independent of the magnitude of the fluctuation.

2) *Frequency domain features:*

- Power spectral density (PSD)

As an initial choice of feature, the power spectral density was used. It is a commonly used frequency domain feature in studies on emotion recognition from EEG. It is usually computed for a number of frequency bands and used as an indicator of the extent of brain activity within each of these bands. The downloaded data have already been down-sampled to 128 Hz and low-pass filtered to remove frequencies above the desired range.[19]

3) *Time-frequency domain features:*

- Discrete wavelet transform (DWT)

The discrete wavelet transform (DWT) is an alternative method to power spectral density for measuring the prominence of different frequencies in the EEG. An important difference is that it preserves time domain information, which is lost in the power spectral density. The wavelet transform is an alternative to the Fourier transform, which decomposes a signal according to certain wavelet functions, rather than sine and cosine functions as in the case of Fourier transforms. In this case, the Daubechies 4 wavelet was used as in the works of Murugappan et al.[20,21] Using the detailed coefficients, three feature vectors were created, comprising the values related to the energy, root mean square (RMS), and entropy, for signal $S(t)$:
Band energy:

$$E_{\text{band}} = \sum_{-\infty}^{\infty} |S(t)|^2.\tag{13}$$

Total band energy across alpha, beta, and gamma bands:

$$E_{\text{total}} = E_{\text{alpha}} + E_{\text{beta}} + E_{\text{gamma}}.\tag{14}$$

From the energy values, the recursive energy efficiency (REE) was obtained for each of the alpha, beta, and gamma bands:

$$\text{REE} = \frac{E_{\text{band}}}{E_{\text{total}}}.\tag{15}$$

Two further values, log(REE) and the absolute value of log(REE) were also computed from the above equation. These three values were included together in a single feature vector.
RMS can be calculated from the decompositions $D_i(n)$ of the DWT in different layers as the following:

$$\text{RMS}(j) = \sqrt{\frac{\sum_{i=1}^{j} \sum_{n_i} D_i(n)^2}{\sum_{i=1}^{j} n_i}}.\tag{16}$$

The REE, RMS, and entropy features were obtained as three feature vectors to be used separately in the classifiers.

4) *Multi-electrode features:* It is important that the existence of interconnections between different parts of the brain is also considered.

- Differential asymmetry and rational asymmetry (DA and RA)

    DA and RA are the difference and ratio of power bands of corresponding pairs of electrodes:

$$DA = X_R - X_L,\tag{17}$$

$$RA = \frac{X_R}{X_L}.\tag{18}$$

X_R and X_L represent the power spectral density feature for the right and left brain hemisphere symmetric pairs electrodes of the scalp.[9]

- Magnitude-squared coherence estimate (MSCE)

    This feature represents the two signals, $S1$ and $S2$, correspondence of each other.[22] It takes into account the cross-PSD between pairs of electrodes according to the following equation[23]:

$$C_{ij}(f) = \frac{|P_{ij}(f)|^2}{P_i(f)P_j(f)}.\tag{19}$$

Only the magnitude of the cross-power spectral density is required, and since $P_{ij} = Pij*$, $|Pji| = |Pij|$, so that $Cij = Cji$. In addition, the value of Cij for $i = j$ is simply Pi, the power spectral density, which is considered separately, so that is also neglected from this set of features.

5) *Connectivity between EEG Channels:* It is important that the existence of interconnections between different parts of the brain is also considered.

- Mutual information (MI)

    Mutual information[11] is a measurement of how informative a random variable is for another random variable. It is calculated on the basis of entropy, given the random X defined as:

$$I(X) = -\sum_{t=1}^{n} P_t \log(P_t).\tag{20}$$

Then, the mutual entropy between two random variables X and Y is defined as:

$$MI_{XY} = \sum_{ij=1}^{n} P_{XY}^{ij} \log\left(\frac{P_{XY}^{ij}}{P_X^i P_Y^j}\right).\tag{21}$$

where P_{XY}^{ij} is the joint probability. Two random variables X and Y are regarded as statistically independent if the mutual information MI_{XY} is zero.[24]

- Pearson correlation coefficient (PCC)

    Pearson correlation[11,12] measures the linear correlation between two variables, and the range is between -1 and 1, which represent negative or positive correlation. The Pearson correlation coefficient between two random variables X and Y is shown as:

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}.\tag{22}$$

where σ_{xy} stands for the covariance of two random variables X and Y. We consider the temporal sampling points which proposed by Chen et al.[11] The PCC is calculated as below:

$$\rho_{xy} = \frac{\sum_{i=1}^{t}(x_i - \bar{x})(y_j - \bar{y})}{\sqrt{\sum_{i=1}^{t}(x_i - \bar{x})^2}\sqrt{\sum_{i=j}^{t}(y_j - \bar{y})^2}}.\tag{23}$$

3.3 | Set of all features

EEG signals with whole duration except the 3 s prior to recording are used for extracting connectivity features. In order to settle the same set up with the dataset paper,[6] we use the last 30 s for extracting features and further steps.

We consider the time domain feature, frequency domain feature, time-frequency domain, and the connectivity feature, and we used the same setup when doing the feature extraction with the DEAP paper. Each feature is extracted from each participant.[6] Table 1 shows the dimension and type for each feature.

TABLE 1 Features and the number of components extracted from the EEG signals

Types	Features	
	Feature name	No. of components
Time domain	Statistical features	224
	Higher-order crossing	320
	Fractal dimension	32
	Hjorth features	64
	Non-stainarity index	32
Frequency domain	Power spectral density	128
Time-frequency domain	REE	288
	RMS	128
	entropy	128
Multi-electrode features	Differential Asymmetry	56
	Rational Asymmetry	56
	MSCE	496
Connectivity features	Mutual information	496
	Pearson correlation	496

3.4 | Hybrid dimension features reduction scheme

We proposed two steps of multiple highly dimensional features reduction in this scheme. We fused the highly dimensional features by the combination of Maximum Relevance Minimum Redundancy (mRMR) and principal components analysis (PCA) in a classification perspective.

- mRMR

We proposed to use mRMR as the first step for the feature reduction method for the combination feature of all these 14 kinds of features. This method uses mutual information to characterize the suitability of features proposed by Peng et al.[25] Mutual information between two variables is shown above. The mRMR method is used to optimize two criteria simultaneously: maximal-relevance criterion D, which aims to maximize average mutual information between each feature and the specific label. The minimum-redundancy criterion R means to minimize the average mutual information between two features.[26] The algorithm finds near-optimal features using forward selection. Given an already chosen set S_k of k features, the next feature is selected by maximizing the combined criterion $D - R$:

$$\max_{x_j \in (X - S_K)} \left[MI(x_j; y) - \frac{1}{k} \sum_{x_i \in (S_K)} MI(x_j; x_i) \right]. \tag{24}$$

At this stage, we applied mRMR to reorder the combined features based on the specific arousal and valence label for each subject. The new order is based on the mRMR theory that shows above the maximum the feature's relevance with the labels and the minimum the redundancy between each features.

- PCA

After the supervised step is the unsupervised stage in which PCA converts the new order features produced by mRMR into the next linearly uncorrelated set. From the first step, we keep the features that have maximum relevance with the labels but minimum redundancy with each other, so that, in this stage, all the features that applied for PCA are the max relevance label features. It potentially keeps the main information for all different types of features. The PCA step further reduces the new high-relevance dimensional features and retains most of its variance.

3.5 | Classification

The emotion information was represented by emotional dimensions such as arousal and valence. For the classification, we divided the values into two classes based whether the value was higher or lower than the midpoint value. For Arousal, we have two classes: high arousal (HA) and low arousal (LA). For valence, they are high valence (HV) and low valence (LV).

For this binary classification problem, many methods can be used for the classification such as k-nearest neighbour (kNN), support vector machine (SVM), Naive Bayes classifier, and random forest (RF).[27-29] In this study, SVM was chosen as it has achieved best performance in many binary classification tasks. We also add the RF classifier due to its super ability to automatically select best features in the classification process.
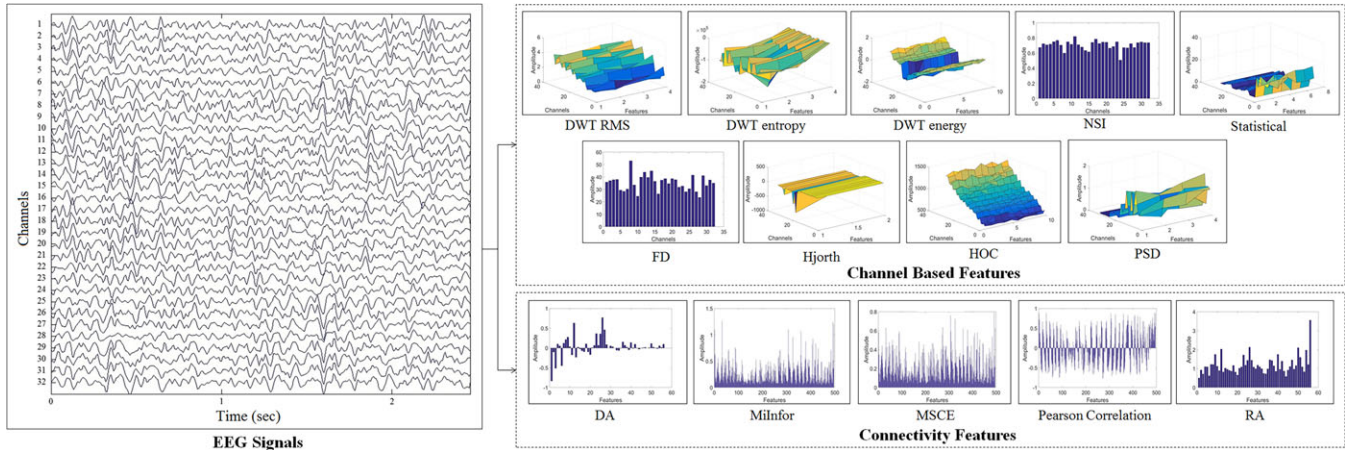
FIGURE 2 Distinct features extracted from one EEG recording with 32 channels in the DEAP dataset

4 | EVALUATION

The proposed method is evaluated on a public dataset and compared with the state-of-the-art performance achieved by other researchers in the same experimental setting condition.

4.1 | Performance measurement

For the classification performance measurement, accuracy is the most popular one that can identify how many samples are classified correctly. We use it here for all the binary classification tasks. In addition, another accuracy measure F1-score is also provided for summarizing the results of each method in considering the balance of a single number class. Furthermore, we carry out the paired t-test to evaluate the significance of the proposed method against other methods.

4.2 | Dataset

DEAP dataset[6] is a multi-modality dataset for the analysis of human affective states. EEG recordings from 32 channels, peripheral physiological signals, and frontal face videos were obtained from 32 participants whilst watching 40 music videos. The videos were selected to evoke one of the four of the following categories of emotion: 1) HV&HA; 2) HV&LA; 3) LA&HA; 4) LV&LA. The EEG data was processed by average referencing, down-sampling to 256 Hz and high-pass filtering to 2 Hz cut-off frequency. Changes in the power relative to the pre-stimulus period was computed and averaged over the Theta (3-7 Hz), Alpha (8-13 Hz), Beta (14-29 Hz), and Gamma (30-47 Hz) bands.

4.3 | Experiment settings

For the DEAP dataset, we address the binary classification problem after thresholding the self-assessments following the protocol in the work of Koelstra et al.[6] The affective label will be set to high if the rating is above 5. If the rating is equal or lower than 5, the corresponding affective label will be set to low. Thus, for each trial, binary labels were generated. High valence (HV) or low valence (LV) was used to describe the affective level in the valence space, and high arousal (HA) or low arousal (LA) was used to describe the affective level in the arousal space. The identification of valence and arousal levels are treated as two independent binary classification tasks in this paper.

4.4 | Feature extraction

For the DEAP dataset, we extract the time domain feature, frequency domain feature, time-frequency domain, and the connectivity feature from each EEG recording. Overall, 14 distinct features were extracted from each recording. Figure 2 shows an example of the features from one recording.

TABLE 2 Classification accuracy and its standard deviation on distinct features (%±Δ)

Features	RF (% ± Δ)		SVM (% ± Δ)	
	Arousal	Valence	Arousal	Valence
DA	62.4±7.2	62.1±7.1	61.4±7.5	59.5±6.8
DWTenergy	65.5±7.3	61.7±6.9	64.5±5.9	59.7±6.4
DWTentropy	62.0±7.4	63.6±7.1	61.0±5.2	60.1±7.6
DWTRMS	65.1±7.4	64.2±7.2	64.4±8.1	63.2±4.8
FD	63.9±6.8	63.3±7.5	62.2±6.7	62.3±7.5
HOC	68.1±5.9	65.7±7.6	66.1±6.3	64.7±6.0
MSCE	69.2±6.9	69.7±7.5	66.2±5.5	65.7±7.4
NSI	57.9±8.0	54.0±6.1	58.8±8.8	54.2±7.9
PSD	64.4±7.4	63.1±6.0	62.4±5.8	61.7±7.3
RA	61.4±7.4	62.2±7.0	62.1±6.8	61.1±5.7
Statistical	61.1±5.5	60.3±6.9	59.0±8.6	58.7±9.4
Hjorth	60.2±6.3	54.5±7.1	59.4±8.5	53.8±7.7
MI	**70.3±6.8**	**72.6±8.9**	**68.7±7.7**	**70.8±6.6**
PCC	69.6±5.9	70.1±8.7	68.2±7.5	68.8±8.4
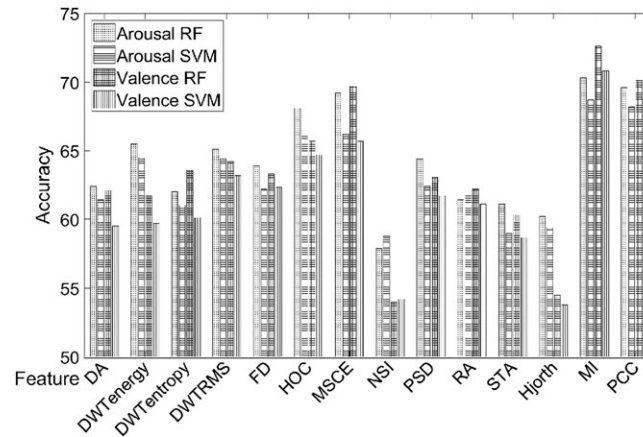
FIGURE 3 Classification accuracy on individual features using RF and SVM for arousal and valence emotional dimensions

4.5 | Results

The following results are obtained on the DEAP database under the same settings as the benchmark paper.[6] As in the work of Koelstra et al,[6] the leave-one-out validation method is applied for each subject to evaluate the performance for each method. It means that, at each step of the validation, one video is treated as the test and the rest, 39 videos, of the same subject are treated as training. For the random forest (RF) method, the parameter is set as 1000 trees. For the linear SVM classifier, a default parameter is used. Table 2 shows the average classification accuracy and its standard deviation over 32 subjects on 14 distinct under the same validation setting.

From Figure 3, it can be seen that the MI feature reaches the highest performance for both arousal and valence as 70.3% and 72.6%, respectively. All the connectivity features are significantly greater than other domain features. It indicates that the further performance improvement might be achieved by applying optimal work on connectivity features. Simultaneously, all types of features show to be potentially and relatively stable based on the classification methods. Additionally, it could be seen from Table 2 that the accuracy of valence levels identification is higher than that of the arousal. This is consistent with the results presented in the work of Lithari et al.[30] This Table presents the performance for different types of features and obviously illustrates the best performance for individual feature extraction method for further comparison.

The next step is the proposed method applied on the combination of all 14 types features. As the same setting with the work of Koelstra et al,[6] the reduction scheme will be applied to each validation step. After the mRMR method re-orders the features and gains the max-relevance min-redundancy attributes, the next is to reduce the features. However, we cannot determine the amount of features to choose. In this case, we proposed the full-range analysis, which means the mRMR will reduce from the last of the new order features and step by 0.01 until the number of dimensions that remained is strictly the best performance. Additionally, this last means at last of the new order features which means

TABLE 3 Average accuracy comparison on emotion dimensions between proposed fusion method with PCA and mRMR individual methods

	Features	RF (% ± Δ)			SVM (% ± Δ)		
		PCA	mRMR	Proposed	PCA	mRMR	Proposed
Arousal	PCC	68.4±7.1	69.8±8.3	71.4±6.2	65.8±8.9	69.6±9.4	70.5±8.4
	MI	68.4±8.7	71.6±7.1	73.6±5.8	66.2±8.8	71.9±7.9	72.1±6.7
	All combine	67.5±9.6	72.7±7.7	**74.3±8.4**	67.2±9.4	71.6±6.5	**72.4±8.8**
Valence	PCC	68.6±9.2	71.3±8.5	73.5±7.8	65.9±9.7	71.0±8.4	71.4±8.9
	MI	69.8±9.3	73.3±6.4	74.8±7.5	68.1±9.4	72.7±8.8	73.6±6.9
	All combine	67.9±9.9	74.2±8.1	**77.2±8.6**	67.2±9.6	73.8±8.3	**76.1±7.4**
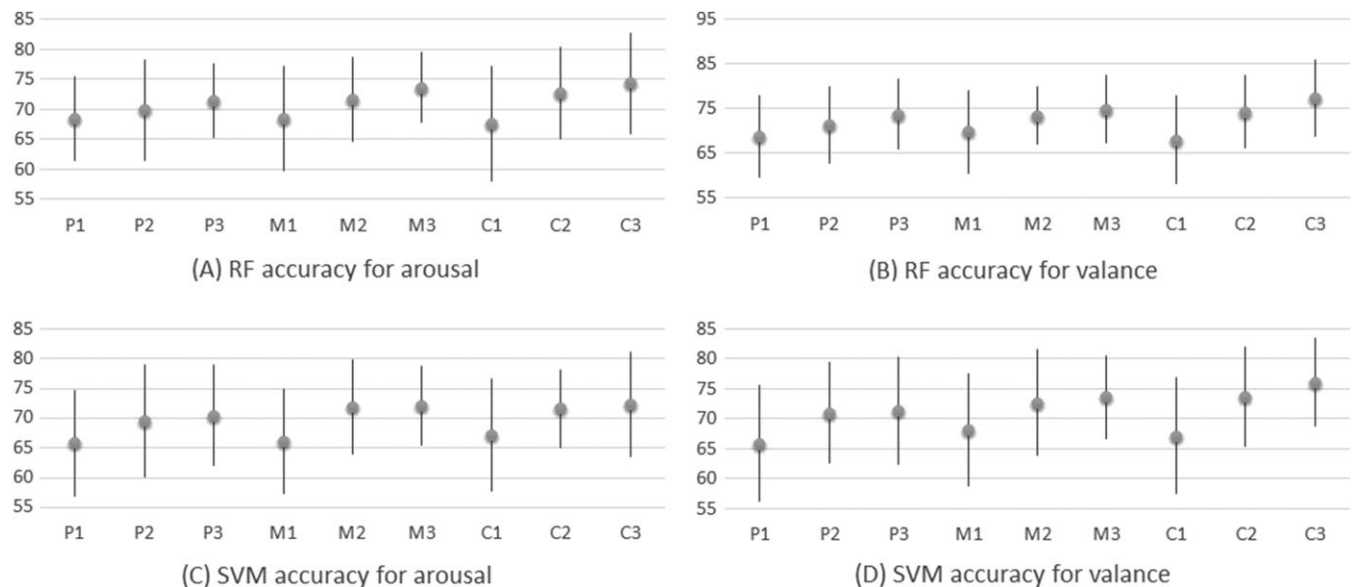
FIGURE 4 Classification accuracy (%) and standard deviation on individual and combined features using different feature reduction methods and classifiers on arousal and valence. (P1 stands for PCC PCA only, P2 stands for PCC mRMR only, P3 stands for Proposed; M1 stands for MI PCA only, M2 stands for MI mRMR only, M3 stands for Proposed; C1 stands for combine PCA only, C2 stands for combine mRMR only, C3 stands for combine Proposed only)

lowest relevance between features and the label and high redundancy of other features. The following presents the comparison between the proposed method and the best performance of other methods. At this stage, we applied two classification methods, SVM and RF, to do the comparison, which will possibly decrease the effect of different classifiers. Simultaneously, we also present the standard deviation, and F1-score represents the balance capability of each algorithm.

Table 3 helps observe the relatively higher performance achieved by the proposed method in comparison with the PCA only and mRMR only on best individual feature extraction algorithm and all features combined. In arousal, the proposal presents the highest average accuracy which reach of 74.3% by RF. Figure 4 also presents the results in the more specific form. Therefore, the proposal performance is significantly higher than other individual features selection method on both single feature or all features combined. In valence, similarly to the previous result, RF achieved the best performance for 77.2%. Simultaneously, if only PCA is applied, for both individual or combined feature, the performance decreases. Moreover, if only mRMR is applied, the performance shows to be greater than original but not significantly. The proposed method reached both highest for arousal and valence.

Table 4 provides the F1-score for PCA-only, mRMR-only and proposed method on combined feature. The results support that a consequence of variety features combination and the proposed hybrid feature reduction method achieved the improvement. In addition, we conduct the paired t-test to check whether the proposed highlighted results are significantly better than the rest of methods in Table 3 for the combined feature. Both in arousal and valence terms, the p-value are presented in Table 5 which indicated that the results of proposed method are statistically significant with PCA only and mRMR only methods.

TABLE 4 Average F1-score accuracy comparison on emotion dimensions between proposed fusion method with PCA and mRMR individual methods

Emotion	Method	RF (% ± Δ)	SVM(% ± Δ)
Arousal	PCA	58.1±11.3	57.1±10.9
	mRMR	62.5±12.1	58.9±9.6
	Proposed Method	71.0±10.2	66.7±8.9
Valence	PCA	59.1±16.2	68.2±11.9
	mRMR	62.9±9.9	71.1±10.1
	Proposed Method	78.9±10.4	77.3±9.8

TABLE 5 The p-value of t-test between the proposed method with PCA and mRMR on all features combined

Emotion	Method	RF	SVM
Arousal	PCA	0.011	0.013
	mRMR	0.032	0.043
Valence	PCA	0.001	0.019
	mRMR	0.024	0.038

4.6 | Comparison with existing works

Finally, we compare our results with other existing work on DEAP dataset for a variation purpose which includes emotional state recognition,[31,32] EEG-based feature extraction,[9,33] and emotion detection system modeling.[34-36] We only present the comparison for the same binary affective levels identification in valence and arousal spaces using EEG signals as shown below in Table 6. Our proposed method reaches highest performance for both arousal and valence.

TABLE 6 Comparison on classification accuracy with the state of the arts

Methods	Arousal (% ± Δ)	Valence (% ± Δ)
Koelstra et al[6]	62.0	57.6
Chung and Yoono[37]	66.4	66.6
Naser and Saha[38]	66.2	64.3
Zhuang et al[39]	67.1±14.1	70.9±11.4
Bahari and Janghorbani[40]	64.6±10.7	58.1±9.3
Arnau-González et al[15]	67.7±11.3	69.6±9.3
Zhuang et al[36]	72.0±7.8	69.1±7.0
Chen et al[11]	73.6±7.9	76.2±6.8
Liu et al[41]	71.2	69.9
Proposed Method	**74.3±8.4**	**77.2±8.6**

5 | CONCLUSION

Emotion detection based on EEG signals is a comparatively new and developing research area. The main aim of this paper is to extract more useful information for the emotion detection from a variety of features and fusion them in an efficient way. From this work, a total of 14 features have been extracted from different domains of EEG recordings. Then, the two-step feature reduction method was used for the dimension reduction of the combine feature vectors. From the experimental results, it can be found that the proposed method was useful to detect the emotion information from EEG recordings with good accuracy. From the experimental results, it can be seen that combination of supervised and unsupervised feature dimension reduction methods can improve the performance by removing some irrelevant feature vectors and is better than using them individually. The final feature produces the best results within all the existing results in the same dataset.

For future improvements, more features can be added into the system, and the performance might be improved further. Especially, various deep learning models can be used for feature extraction and dynamic modeling, which will improve the performance. In addition, based on different kinds

of features, more advanced fusion methods might be used in the future. Furthermore, the contribution of each channel can be further analyzed, and channel selection might make further improvements.

ORCID

Hongying Meng http://orcid.org/0000-0002-8836-1382

REFERENCES

1. Ekman P, Friesen W V, O'Sullivan M, et al. Universals and cultural differences in the judgments of facial expressions of emotion. *J Pers Soc Psychol*. 1987;53(4):712-717.

2. Posner J, Russell JA, Peterson BS. The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. *Dev Psychopathol*. 2005;17(3):715-734.

3. Plutchik R. The nature of emotions: human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *Am Sci*. 2001;89(4):344-350.

4. Shaver P, Schwartz J, Kirson D, O'connor C. Emotion knowledge: further exploration of a prototype approach. *J Pers Soc Psychol*. 1987;52(6):1061-1086.

5. Picard RW, Vyzas E, Healey J. Toward machine emotional intelligence: analysis of affective physiological state. *IEEE Trans Pattern Anal Mach Intell*. 2001;23(10):1175-1191.

6. Koelstra S, Muhl C, Soleymani M, et al. DEAP: a database for emotion analysis; using physiological signals. *IEEE Trans Affect Comput*. 2012;3(1):18-31.

7. Valenzi S, Islam T, Jurica P, Cichocki A. Individual classification of emotions using EEG. *J Biomed Sci Eng*. 2014;7(8):604-620.

8. Alarcao SM, Fonseca MJ. Emotions recognition using EEG signals: a survey. *IEEE Trans Affect Comput*. 2017.

9. Jenke R, Peer A, Buss M. Feature extraction and selection for emotion recognition from EEG. *IEEE Trans Affect Comput*. 2014;5(3):327-339.

10. Othman M, Wahab A, Karim I, Dzulkifli MA, Alshaikli IFT. EEG emotion recognition based on the dimensional models of emotions. *Procedia Soc Behav Sci*. 2013;97:30-37.

11. Chen M, Han J, Guo L, Wang J, Patras I. Identifying valence and arousal levels via connectivity between EEG channels. Paper presented at: 2015 International Conference on Affective Computing and Intelligent Interaction (ACII); 2015; Xi'an, China.

12. Pearson K. Note on regression and inheritance in the case of two parents. *Proc R Soc Lond*. 1895;58:240-242.

13. Hassan M, Dufor O, Merlet I, Berrou C, Wendling F. EEG source connectivity analysis: from dense array recordings to brain networks. *PloS One*. 2014;9(8):e105041.

14. Gupta R, Laghari KR, Falk TH. Relevance vector classifier decision fusion and EEG graph-theoretic features for automatic affective state characterization. *Neurocomputing*. 2016;174:875-884.

15. Arnau-González P, Arevalillo-Herráez M, Ramzan N. Fusing highly dimensional energy and connectivity features to identify affective states from EEG signals. *Neurocomputing*. 2017;244:81-89.

16. Petrantonakis PC, Hadjileontiadis LJ. Emotion recognition from EEG using higher order crossings. *IEEE Trans Inf Technol Biomed*. 2010;14(2):186-197.

17. Higuchi T. Approach to an irregular time series on the basis of the fractal theory. *Physica D: Nonlin Phenom*. 1988;31(2):277-283.

18. Franses PH. Seasonality, non-stationarity and the forecasting of monthly time series. *Int J Forecast*. 1991;7(2):199-208.

19. Cajochen C, Brunner DP, Krauchi K, Graw P, Wirz-Justice A. Power density in theta/alpha frequencies of the waking EEG progressively increases during sustained wakefulness. *Sleep*. 1995;18(10):890-894.

20. Murugappan M, Rizon M, Nagarajan R, Yaacob S, Zunaidi I, Hazry D. EEG feature extraction for classifying emotions using FCM and FKM. *Int J Comp Commun*. 2007;1(2):21-25.

21. Murugappan M, Ramachandran N, Sazali Y. Classification of human emotion from EEG using discrete wavelet transform. *J Biomed Sci Eng*. 2010;3(4):390-396.

22. Khosrowabadi R, Quek HC, Wahab A, Ang KK. EEG-based emotion recognition using self-organizing map for boundary detection. Paper presented at: 2010 20th International Conference on Pattern Recognition (ICPR); 2010; Istanbul, Turkey.

23. Carter G, Knapp C, Nuttall A. Estimation of the magnitude-squared coherence function via overlapped fast Fourier transform processing. *IEEE Trans Audio Electroacoustics*. 1973;21(4):337-344.

24. Brown G, Pocock A, Zhao M-J, Luján M. Conditional likelihood maximisation: a unifying framework for information theoretic feature selection. *J Mach Learn Res*. 2012;13:27-66.

25. Peng H, Long F, Ding C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans Pattern Anal Mach Intell*. 2005;27(8):1226-1238.

26. Ding C, Peng H. Minimum redundancy feature selection from microarray gene expression data. *J Bioinform Comput Biol*. 2005;3(02):185-205.

27. Altman NS. An introduction to kernel and nearest-neighbor nonparametric regression. *Am Stat*. 1992;46(3):175-185.

28. Chang C-C, Lin C-J. LIBSVM: a library for support vector machines. *ACM Trans Intell Syst Technol*. 2011;2(3):27.

29. Ho TK. Random decision forests. Paper presented at: Proceedings of the Third International Conference on Document Analysis and Recognition; 1995; Montreal, Canada.

30. Lithari C, Klados MA, Bamidis PD. Graph analysis on functional connectivity networks during an emotional paradigm. Paper presented at: XII Mediterranean Conference on Medical and Biological Engineering and Computing; 2010; Chalkidiki, Greece.

31. Liu Y, Sourina O. EEG-based valence level recognition for real-time applications. Paper presented at: 2012 International Conference on Cyberworlds (CW); 2012; Darmstadt, Germany.

32. Liu Y, Sourina O. EEG databases for emotion recognition. Paper presented at: 2013 International Conference on Cyberworlds (CW); 2013; Yokohama, Japan.

33. Bastos-Filho TF, Ferreira A, Atencio AC, Arjunan S, Kumar D. Evaluation of feature extraction techniques in emotional state recognition. Paper presented at: 2012 4th International Conference on Intelligent Human Computer Interaction (IHCI); 2012; Kharagpur, India.

34. Akinci HM, Yesil E. Emotion modeling using fuzzy cognitive maps. Paper presented at: 2013 IEEE 14th International Symposium on Computational Intelligence and Informatics (CINTI); 2013; Budapest, Hungary.

35. Vijayan AE, Sen D, Sudheer AP. EEG-based emotion recognition using statistical measures and auto-regressive modeling. Paper presented at: 2015 IEEE International Conference on Computational Intelligence & Communication Technology (CICT); 2015; Ghaziabad, India.

36. Zhuang N, Zeng Y, Tong L, Zhang C, Zhang H, Yan B. Emotion recognition from EEG signals using multidimensional information in EMD domain. *BioMed Res Int*. 2017;2017:1-1.

37. Chung SY, Yoon HJ. Affective classification using Bayesian classifier and supervised learning. Paper presented at: 2012 12th International Conference on Control, Automation and Systems (ICCAS); 2012; JeJu Island, South Korea.

38. Naser DS, Saha G. Recognition of emotions induced by music videos using DT-CWPT. Paper presented at: 2013 Indian Conference on Medical Informatics and Telemedicine (ICMIT); 2013; Kharagpur, India.

39. Zhuang X, Rozgic V, Crystal M. Compact unsupervised EEG response representation for emotion recognition. Paper presented at: 2014 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI); 2014; Valencia, Spain.

40. Bahari F, Janghorbani A. EEG-based emotion recognition using recurrence plot analysis and K nearest neighbor classifier. Paper presented at: 2013 20th Iranian Conference on Biomedical Engineering (ICBME); 2013; Tehran, Iran.

41. Liu J, Meng H, Nandi A, Li M. Emotion detection from EEG recordings. Paper presented at: 2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD); 2016; Changsha, China.