Hamedan University of Technology
Iranian Association of Electrical and Electronics Engineers

Sustainable Energy and Artificial Intelligence

Online ISSN: 3060-8015

Journal Homepage: www.enai.hut.ac.ir

Identification of Human Emotions Using EEG Signals Based on an Intelligent Discriminative Sparse Model

Mohammad Bagher Khodabakhshi[1], Shahriar Jamasb[2*]

[1] *Department of Biomedical Engineering, Hamedan University of Technology, Hamedan, Iran*
[2] *Department of Biomedical Engineering, Hamedan University of Technology, Hamedan, Iran, 65169-13733*
** Corresponding Author: jamasb@hut.ac.ir*

Article Info

Article type:
Original Article

Article history:
Received 2024-10-15;
Revised 2025-02-26;
Accepted 2025-03-08.

How to cite this article:
khodabakhshi, M. B. and Jamasb, S. (2025). Identification of Human Emotions using EEG signals Based on an Intelligent Discriminative Sparse Model. *Sustainable Energy and Artificial Intelligence*, 1(4), 217-227.
DOI: 10.61186/seai.2410-1016

Abstract

Automatic categorization of human affective states is a significant challenge in the development of human-Machine interfaces (HMI), particularly due to the subject-dependent variability of electroencephalogram (EEG) signals. This study addresses this problem by proposing a novel sparse representation-based classification (en-SRC) framework that enhances the discriminative power of emotion recognition. The en-SRC model introduces class-specific dictionary atoms and integrates separate classifiers with a voting strategy to improve classification accuracy. Using EEG data from the DEAP dataset, our framework achieved classification accuracies of 90.47% for valence and 78.02% for arousal, surpassing several state-of-the-art methods. By employing dynamic windowing and leveraging both statistical and nonlinear features, the model effectively captures the temporal dynamics of emotional states. These findings highlight the en-SRC framework's potential for advancing emotion recognition in HMI systems, providing a robust solution for applications in affective computing and adaptive technologies.

Keywords: Emotions; Affective States; Sparse Representation Models; EEG Signals; Discriminative Information.

1. Introduction

Identification of emotional states is one of the main challenges encountered in the design of a human-computer interaction (HMI) system. In recent years, researchers have explored various methods for understanding human emotions by exposing subjects to different stimuli, such as images, sounds, and videos, and recording their responses using modalities such as facial expressions, speech patterns, and physiological signals [1-3]. To categorize the emotions, two general approaches have been adopted: discrete and dimensional models. The first scheme classifies emotions into predefined categories using such as sadness, happiness, anger, while the dimensional model represents emotions as points on a 2D plane, commonly referred to as Russell's 2D emotion model [4]. Notably, the latter approach locates human emotions on the Russel's 2D plane as specific points represented by the levels of valence and arousal corresponding to the *x* and *y* coordinates, respectively. The dimensional model is more popular than the discrete model, given the broader range of the basic emotions covered by the former.

Among physiological signals, electroencephalogram (EEG) has garnered significant attention for emotion recognition, as brainwave patterns directly reflect variations in emotional states and are not easily influenced by voluntary control [5, 6]. However, EEG signals present unique challenges: they vary significantly across individuals, exhibit dynamic behavior, and are often difficult to classify due to overlapping boundaries between emotional states. Additionally, the subjective nature of emotional labeling, often based on participants' self-reported ratings, further complicates the classification process. These challenges underscore the need for advanced machine-learning models that can effectively handle such variability and improve classification accuracy.

Recent studies have employed various pattern recognition and machine-learning techniques to tackle these challenges [7, 8]. To identify robust patterns associated with specific emotions, automatic classification schemes have utilized various features such as statistical measures, power spectral-based methods, and time-frequency approaches [9, 10]. The performance of various popular feature extraction and pattern classification methods such as power spectral density (PSD), differential entropy (DE), and differential caudality (DCAU) is systematically investigated in emotion analysis [11]. In addition, Bhattaacharyya, et. al, porpose a multivariate-multisclae approach for computing EECG Spectral and temporal complexity for emotion recognition [12]. Since EEG signals exhibit a nonlinear behavior, time and time-frequency analyses may not be applicable, given the dynamical regimes encountered. Accordingly, some studies have focused on the reconstruction of the signal's state space to permit application of chaotic analysis [13, 14].

Despite these efforts, current emotion recognition methods often struggle to produce robust features with sufficient discriminatory power, leading to unreliable classification results. Sparse representation models have shown promise in addressing such issues by treating input data as sparse linear combinations of dictionary atoms [1, 3]. These models have been successfully applied in various fields such as image restoration and pattern classification. In this approach, the atoms of the dictionary act as basis functions for representing input patterns. Unlike traditional orthogonal dictionaries, over-complete dictionaries allow greater flexibility and improved representation capability. The choice of dictionary significantly impacts the performance of sparse models. Conventional dictionaries based on analytic transforms, such as Fourier transforms or wavelets, often lack adaptability and are not optimized for specific classification tasks [15]. Therefore, although the reconstruction error is minimized, the dictionaries may not serve adequately for a specific classification task. In contrast, optimal dictionary learing methods such as (MOD) [16] and KSVD [17], have furnished high-performance dictionaries. However, the dictionaries developed by the KSVD method, which are only trained to represent the training samples, are not suitable for classification tasks. Therefore, exploiting the class discrimination information is pivotal for improving the classification performance. By employing the training samples as the atoms of the dictionary, Wright et al. [18] developed a novel method which classifies a query sample by evaluating the minimum reconstruction error. This approach, however, can lead to poor classification results due to the presences of noise and trivial information in the training patterns. The performance improvement achieved by a given model is strongly dependent on the incorporation of the class differences into the models. Accordingly, to improve the performance of the KVSD method, Zhang *et al.* introduced a more discriminative version of the method, incorporating the reconstruction and classification criteria in the formulation of the problem [19]. Also, Yang et al. have introduced a structured learning dictionary based on the Fisher discrimination criterion, in which the atoms have a one-to-one correspondence to the class labels [20].

In this work, we propose a novel sparse representation-based classification (en-SRC) framework inspired by structured dictionary learning approaches. The en-SRC model employs structured over-complete dictionaries with class-specific atoms to enhance between-class discrimination. Additionally, we introduce a voting strategy that combines decisions from multiple classifiers, further improving classification accuracy. To handle the temporal variability of EEG signals, the framework utilizes a dynamic windowing mechanism, segmenting the signals into overlapping windows to capture fine-grained temporal dynamics. Our proposed methodology is evaluated on the DEAP dataset, which includes EEG recordings of emotional responses to audiovisual stimuli [21]. We analyze statistical features extracted from the EEG signals and compare the performance of the en-SRC framework with state-of-the-art methods. The results demonstrate the effectiveness of the proposed model in improving classification accuracy for valence and arousal levels. This work contributes to the development of robust emotion recognition systems, with significant implications for real-world HMI applications in affective computing, adaptive systems, and healthcare technologies.

The rest of this paper is organized as follows: Section 2 describes the structured dictionary-based sparse model. Section 3 presents our methodology including the dataset description, the extracted features, and the experimental results, while Section 4 provides a discussion on the model's performance. Finally, relevant conclusions are drawn in Section 5.

2. Sparse Representation-Based Classifier

In this section, we review the preliminaries of the sparse representation model, and present our algorithm to improving the classification performance using sparse representation. First the given n_i training samples from i'th ($i = 1,2, \dots, K$) class are rearranged as columns of the matrix $A_i = [\mathbf{x}_{i,1}, \mathbf{x}_{i,2}, \dots, \mathbf{x}_{i,n_i}] \in \mathbb{R}^{N \times n_i}$. In the context of emotion recognition, each column of A_i is corresponds to the features extracted from the EEG signal belonging to class i. We employ the sparse representation-based classification (SRC) method introduced by Wright et al. [18] defining the structured dictionary based on the above definition as $\mathbf{D} = [A_1, A_2, \dots, A_K]$. Sparse represention of the input query on $\mathbf{D}$, yields the corresponding reconstruction coefficients. We can summarize the SRC procedure as follows:

a) Obtian sparse representation of $\mathbf{y}$ over dictionary $\mathbf{D}$, by solving the ℓ^1-minimization problem using:
$$\widehat{\boldsymbol{\alpha}} = \arg\min_{\alpha} \{ \|\mathbf{y} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 + \gamma \|\boldsymbol{\alpha}\|_1 \} \quad (1)$$
where $\widehat{\boldsymbol{\alpha}} = [\widehat{\boldsymbol{\alpha}}_1; \widehat{\boldsymbol{\alpha}}_2; \dots; \widehat{\boldsymbol{\alpha}}_K]$ denotes the reconstruction vector and γ is a constant scalar. To have the unit ℓ^2-norm, each column of $\mathbf{D}$ has to be normalized.

b) Compute the residuals $e_i = \|\mathbf{y} - \mathbf{D}_i \widehat{\boldsymbol{\alpha}}_i\|_2^2$ for the query sample $\mathbf{y}$. The query belongs to the class with the minimum residual:
$$\text{identity}(\mathbf{y}) = \arg\min_i \{e_i\} \quad (2)$$

In the sparse models, the number of atoms can be chosen to generate an over-complete dictionary. It has been proved that the ℓ^1-minimization problem in Eq.Error! Reference source not found.) is (Nondeterministic Polynomial-Time Hard) NP-Hard. To obtain the solution, The KSVD and the orthogonal matching pursuit (OMP) are reliable methods, which yield the solutions by approximating dictionary elements and the reconstruction vector, respectively. In othe words, KSVD and OMP are considered reliable methods offer approximate solutions efficiently, leveraging heuristic approaches that balance accuracy and computational cost. Accordingly, the process of training the model involves the iterative estimation of $\mathbf{D}$ and $\boldsymbol{\alpha}$ by KSVD and OMP until the convergence criterion is satisfied [17].

Although Wright *et al.* reported impressive results using this methodology [18], our findings indicate that for automatic recognition of emotions, relying on the discriminating information represented by the residual errors limits the classification performance. Therefore, we have introduced an alternative classification model in conjunction with the SRC termed the enhanced SRC (en-SRC) model given by:
$$\arg\min_{\alpha} \|\mathbf{y} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 + \gamma \|\boldsymbol{\alpha}\|_1 ; \quad st: \|f(\boldsymbol{\alpha}) - \mathbf{h}\|_2^2 \quad (3)$$
where f signifies the output label produced for the input vector $\boldsymbol{\alpha}$ by an external classifier, and $\mathbf{h}$ is a vector specifying the desired label. In Eq. Error! Reference source not found.) the reconstruction coefficients are obtained based on both the residual error minimization and classification accuracy. In other words, by adopting OMP, the coefficients $\boldsymbol{\alpha}$ are determined with the class-correspondence atoms of the dictionary remaining unchanged during the training procedure. The latter approach will preserve the discrimination property of the model, thus improving its classification performance. We examined the performance of different external classifiers, namely the support vector machine (SVM) and k-nearest neighbor (KNN), to categorize the sparse representation coefficients. Table 1 summarizes the training and testing procedures applied to the proposed model.

As indicated by Table 1 the dictionary atoms in the en-SRC model remain fixed during the training phase, undermining the ability of the model to minimize the reconstruction error.

3. Methodology

In this section, we evaluate the performance of the proposed model in distinguishing between different emotional states corresponding to EEG signals from the 'DEAP' dataset [21]. In this way, various characterizations of brain signals, including statistical and nonlinear features are employed as inputs of our models. Before describing how to extract the features, we briefly describe the contents of the 'DEAP' dataset. Then, the results obtained by our proposed models are compared with those presented in the literature.

Table 1. Classification based on the sparse representation with separate classifier (en-SRC)

Training procedure of the en-SRC
1. **Input:** input sample $\mathbf{x}_m \in \mathbb{R}^N, m = 1, 2, \ldots, M$ selecting dictionary atoms from each class $\mathbf{D} = [A_1, A_2, \ldots, A_K]$ for K classes.
2. Normalization of each column of $\mathbf{D}$.
3. Solving: $\widehat{\boldsymbol{\alpha}}_m = \arg\min_\alpha \{ \|\mathbf{x}_m - \mathbf{D}\boldsymbol{\alpha}_m\|_2^2 + \gamma\|\boldsymbol{\alpha}_m\|_1\}$ Estimation of the sparse representation coefficients $\widehat{\boldsymbol{\alpha}}_m$ over $\mathbf{D}$ through OMP for all training samples.
4. Training an external classifier using the solution obtained in step 3 for $\widehat{\boldsymbol{\alpha}}_m, m = 1, 2, \ldots, M$
Testing the performance of the model
1. **Input:** a query sample $\mathbf{y} \in \mathbb{R}^N$, and $\mathbf{D} = [A_1, A_2, \ldots, A_K]$ for K classes.
2. Estimating $\boldsymbol{\alpha}$ over $\mathbf{D}$ based on the OMP.
3. Applying $\boldsymbol{\alpha}$ to the trained classifier, leading to $identity\,(\mathbf{y}) = f(\boldsymbol{\alpha})$.

3-1. DEAP Dataset

One of the most famous datasets available for the analysis of emotions is ‘DEAP’ [21]. This dataset is captured from 32 participants while watching 40 0ne-minute long music videos. DEAP was created based on Russell’s emotion description model [8] in which the videos were scored based on arousal and valence. Also, all music-videos rated in terms of dominance, familiarity, and like/dislike. While watching videos, 40 physiological channels of data, including 32 channel EEG signals and other peripheral signals recorded. The trials take 63 seconds consisting of a 3-second pre-trial period and a 1-minute trial, each of which was terminated by filling a questionnaire based on the well-known self-assessment manikins (SAM) [22]. The pre-processed publicly available dataset contains 32-electrode EOG artifacts removal EEG signals which sampled at 128 Hz.

In the classification of human affective states, the feelings each volunteer experienced utilized as gold standards. SAM questionnaire receives the ratings as numeric values ranging from 1 to 9. So, in different studies, this range is discretized to implement the classification tasks. Here, we have considered binary partitioning by placing a threshold in the middle of the 9-point rating scales. In most emotion analysis systems, the recognition algorithms are often implemented to differentiate between the levels of arousal and valence. Herein, we have evaluated the performance of our introduced model with regard to its ability to distinguish between the levels of arousal and valence.

3-2. Feature Extraction

Two types of features were calculated to be used as inputs of the en-SRC and SRC models. The features are obtained from statistical and nonlinear analysis computed in all 32 EEG electrodes. First, we have introduced the features, and then its degree of distinction has been indicated by statistical measures. The statistical measures determine the complexity of the data distribution verifying the ability of the approaches in their classification tasks.

3-2-1. Statistical Features

According to [9], six well-known statistical features were extracted from EEG signals of all channels. A list of these six descriptive statistics is summarized as follows:

1) Mean: Mean of the time-series
2) STD: Standard deviation of EEG signals
3) AFD: Mean of the absolute values of the signal’s first differences
4) $\overline{\text{AFD}}$: Mean of the normalized absolute values of the signal’s first differences
5) ASD: similar to AFD for the signal’s second differences
6) $\overline{\text{ASD}}$: Same as $\overline{\text{AFD}}$ for the signal’s second differences

The listed features extracted from EEG time-series for all 32 channels creating the corresponding time-domain feature vectors.

3-2-2. Nonlinear Features

In recent years, several attempts have been made to extract nonlinear features from EEG signals to analyze emotions. As the EEG signals originated from the nonlinear nature of the human brain, these types of features are more likely to represent its different dynamical regimes. There are a couple of studies providing relevant insight into the nonlinear study of human emotions including the fractal dimension [23], correlation dimension [24, 25], approximate entropy [25], etc. Recurrence quantification analysis (RQA) [26] also employed as another chaotic approach, and it obtains from the signal's recurrence plot (RP). Motivated by the advantages of the RQA approaches, herein, we have examined the emotion recognition system by extraction of RQA features from EEG channels. Inspired by the results reported in [27,28], we have used the following five RQA features to produce a sequence of nonlinear features:

1) Length of longest horizontal (vertical) line: represents the maximum laminar phase in an RP
2) Laminarity: in an RP laminarity defined as the ratio of the number of vertical lines to the total number of recurrence points
3) Recurrence period density entropy: quantifies the normalized entropy of the recurrence time distribution in RP
4) Transitivity: defines as a probability that relates two nodes in an RP connected by an edge
5) Determinism: introduces a measure for determining the chaotic nature and unpredictability of the system.

3-2-3. Results

In this section, we will describe our methodology for the single-trial classification of the videos. The performance of the extracted features is compared with regard to their ability to distinguish between the levels of feelings using EEG data from the DEAP dataset. First, the features for all EEG channels were extracted for each trial. To provide an insight into how distinct the extracted features are, box plots associated with ASD feature are presented in Fig. 1. The plots in Fig. 3 verify the overlap between classes indicating the low degree of separation involved. Also, a significant difference between classes was identified based on a relatively small number of features.

In order to achieve an insight into how the introduced models can be generalized for an independent dataset, the K-fold cross-validation approach with K=4 is applied. Finally, the performance of the models is evaluated based on the percentage of the correctly classified trials in each of the classes. In the en-SRC model, for, the representation vectors corresponding were estimated to train the separate classifier. Afterwards, an estimated representation vector associated with an unseen sample for a given class is assigned by the classifier. Setting N= 4 and p= 10, the dimensions of the dictionary become 4 × 10, resulting in an over-complete dictionary. These parameters, are model hyperparameters that were selected through a trial-and-error process to achieve optimal performance. Half of the atoms are selected from one class and the rest of the atoms are chosen from the second one.

Table 2 reports the classification accuracy based on the residual errors. In addition, Table 2 also illustrates the mean squared error (MSE) between the desired inputs and their reconstructed version. Inspection of the results given in Table 2 indicates that the en-SRC method can accurately represent data with a minimum reconstruction error. Due to the fixed structure of the dictionary, en-SRC cannot minimize the reconstruction error leading to the poor classification results obtained based on this strategy. To search for the cause of this shortcoming, reconstruction errors generated using en-SRC method associated with the samples over sub-dictionaries $\mathbf{D_1}$ and $\mathbf{D_2}$ are examined in Table 3. According to Table 3, the reconstruction error of the data of class '1' over $\mathbf{D_2}$ is fewer than that achieved over $\mathbf{D_1}$. Therefore, most of the instances in class '1' fall into the wrong category. Similarly, class '2' data is not properly categorized.

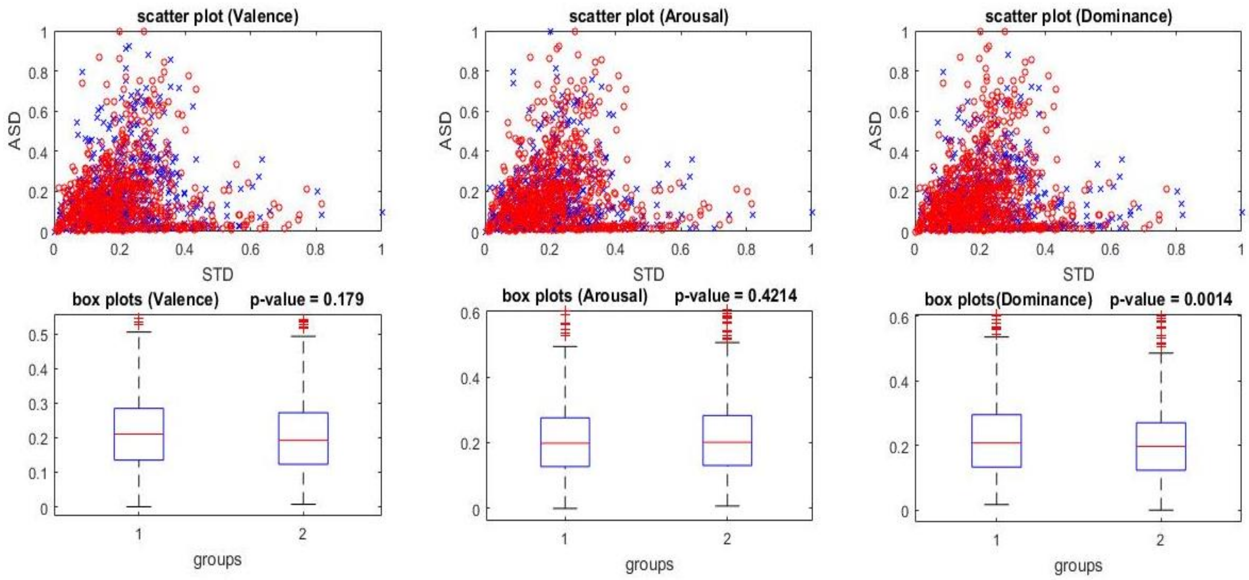

Fig. 1. Scatter and box plots of channel FP1 corresponding to the statistical features

Table 2. the classification rates in accordance with the residual errors.

model	Reconstruction error for class 1		Reconstruction error for class 2		Classification based on minimum reconstruction error	
	Train (MSE)	Test (MSE)	Train (MSE)	Test (MSE)	Train (%)	Test (%)
SRC	0.1468	0.1180	0.1630	0.1506	54.19	50.00
en-SRC	0.0022	0.0120	0.0047	0.0097	75.06	51.68

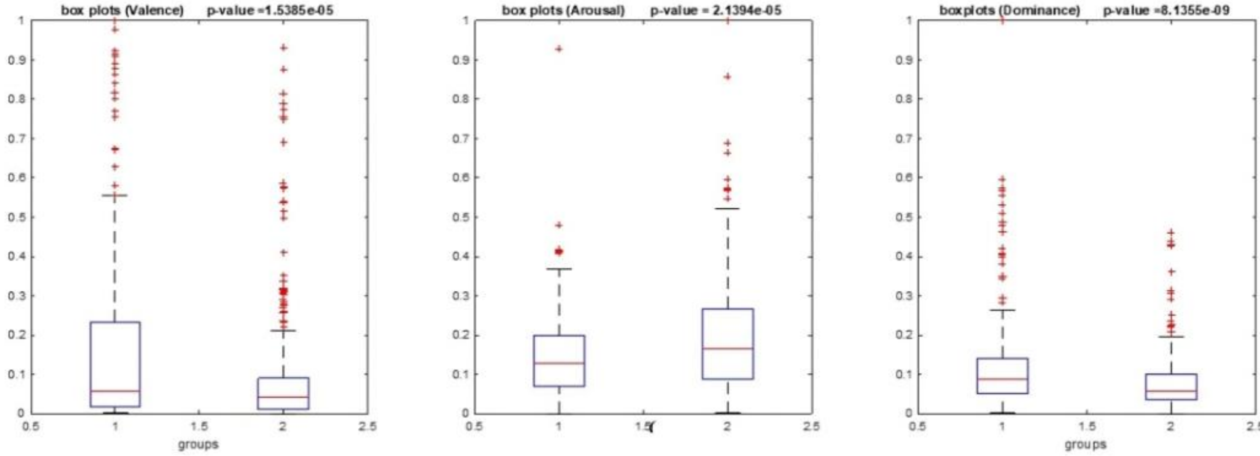

Fig. 2. Box plots of the reconstruction coefficients obtained by en-SRC

Table 3. The en-SRC model's output. the mean of reconstruction errors of each class data over the sub-dictionaries.

	Dictionary $\mathbf{D_1}$	Dictionary $\mathbf{D_2}$
Reconstruction error for class 1	0.0144	0.0065
Reconstruction error for class 2	0.0037	0.0020

Table 3 presents the average reconstruction errors for each data class across both sub-dictionaries. The results indicate that the queries associated with each class were not accurately reconstructed by their respective sub-dictionaries, leading to a high classification error.

To assess the capability of the en-SRC model in increasing the discriminative property of the input samples, boxplots of the reconstruction coefficients are illustrated in Fig. 2. The left, middle, and the right plot are the results corresponded to the valence, arousal, and dominance, respectively As indicated, the distribution of the data corresponding to the two classes in Fig. 2 is slightly more distinct compared with that of the samples in Fig. 1. This, in turn, influences the performance of the model by improving its ability to distinguish between the levels of the affective states.

Table 4 summarizes the classification results for the emotional states described in terms of accuracy for valence and arousal. The accuracy is calculated based on the correctly classified queries per the total number of samples. As can be seen there, the results for both sets of statistical and nonlinear features are listed. In Table 4, the performance of the studies reported in the literature is also listed. According to these findings, the highest classification accuracy levels achieved by the proposed en-SRC model in the testing phase were 86.65% and 77.80% for valence and arousal respectively. Also, Table 4 indicated that the en-SRC model did not comprehensively outperform the other approach. To increase the performance of the proposed model, inspired by the success of the methodologies presented by Zangeneh, et al. [27], we have applied ensemble strategy as external classifiers in our sparse model. To this end, the reconstruction coefficients obtained from SRC and en-SRC models were applied to different classifiers, and their decisions were combined through the voting algorithm. After that, the query sample belongs to the class with the maximum number of votes. Fig. 3 indicates the block diagram of our modeling approach in combination with the voting method. Since the participants' brain functions are changing while watching videos, we have considered the signals in a smaller period to handle their dynamical changes. For this purpose, EEG signals were split into a finite number of windows in which the statistical and nonlinear features were computed. The length of each window was selected to be 3 s with a 2 s shift between neighboring windows. Since the time duration of the pre-stimulus period of each trial is 3 s, we have selected 3 s window sizes. By combining the results obtained from different windows, the classification of each trial was made based on the maximum number of the assigned labels to all windows.

indicates that our model indicates promising outcomes compared to the other approaches in terms of classification accuracy in both arousal-valence states.

Table 4 indicated that the en-SRC model did not comprehensively outperform the other approach. To increase the performance of the proposed model, inspired by the success of the methodologies presented by Zangeneh, et al. [27], we have applied ensemble strategy as external classifiers in our sparse model. To this end, the reconstruction coefficients obtained from SRC and en-SRC models were applied to different classifiers, and their decisions were combined through the voting algorithm. After that, the query sample belongs to the class with the maximum number of votes. Fig. 3 indicates the block diagram of our modeling approach in combination with the voting method. Since the participants' brain functions are changing while watching videos, we have considered the signals in a smaller period to handle their dynamical changes. For this purpose, EEG signals were split into a finite number of windows in which the statistical and nonlinear features were computed. The length of each window was selected to be 3 s with a 2 s shift between neighboring windows. Since the time duration of the pre-stimulus period of each trial is 3 s, we have selected 3 s window sizes. By combining the results obtained from different windows, the classification of each trial was made based on the maximum number of the assigned labels to all windows.

We have divided the input samples into training and testing components, and applied the K-fold strategy with K=4. Table 5 lists the performance of the ensemble classifiers in combination with our sparse models. Following the approach presented in [27] and [9], we have employed several EEG channels for feature extraction. In particular, the signals of FP1, FP2, F3, and F4 have been chosen for statistical feature extraction [9], whereas the nonlinear features are computed using the signals of the following channels; FP1, T3, F4, O2, T6, and F7 [27]. Since each separate classifier provides different insights into the problem, combining their decisions improves the overall performance of the ensemvle en-SRC. The receiver operating characteristic (ROC) curve is plotted in Fig. 4. As is indicated, the combination of nonlinear and statistical features improves the performance of the model. In addition, the en-SRC model produces the best result when the voting strategy is employed. The performance of the en-SRC models is also improved when the ensemble classifier is adopted. In general, Fig. 4 demonstrates that the models are more successful in classification of arousal than valence.

Table 4. Results of the different methods for classification of dimensional basic emotions. The last row is devoted to the results of our proposed en-SRC

Method	Features	Classifier	# of Channels	Arousal (%)	Valence (%)
Nonlinear approach [25]	CD, FD, SaEn, ApEn	SOM	32	68.92	71.25
Recurrence plot [29]	RQA	KNN	32	64.53	56.80
DEAP [21]	Statistical and spectral	Naïve Bayes	32	65.10	62.70
Spectral power density (SPD) [9]	SPD	Random forest	4	67.90	86.60
Bispectrum analysis [8]	Bispectrum	SVM, ANN	32	64.84	61.17
Statistical methods [9]	Mean, ASD, AFD	SVM	4	65.00	61.20
SRC	Mean, ASD, AFD, RQA	SVM	6	79.39	70.01
en-SRC	Statistical, RQA	SVM	32	86.65	77.80

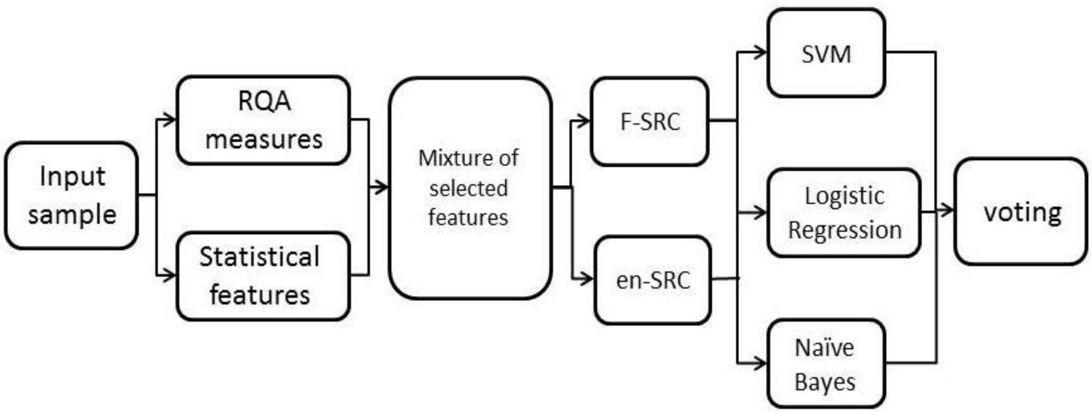

Fig. 3. Block diagram of our approach for combining features and classification approaches

Table 5. Classification performance of the proposed methods based on the dynamical variations of the signals and ensemble classifiers. Compared to Table 4, we have added ensemble strategy to our proposed en-SRC.

Method	Features	Classifier	# chann.	Arousal (%)	Valence (%)
Statictical analysis	Mean, ASD, AFD	SVM	4	71.90	60.63
SRC	Mean, ASD, AFD, RQA	SVM	6	79.39	70.01
SRC	Mean, ASD, AFD, RQA	ensemble	6	80.16	70.24
en-SRC	Mean, ASD, AFD, RQA	SVM	6	86.65	77.80
en-SRC	Mean, ASD, AFD, RQA	ensemble	6	90.47	78.02

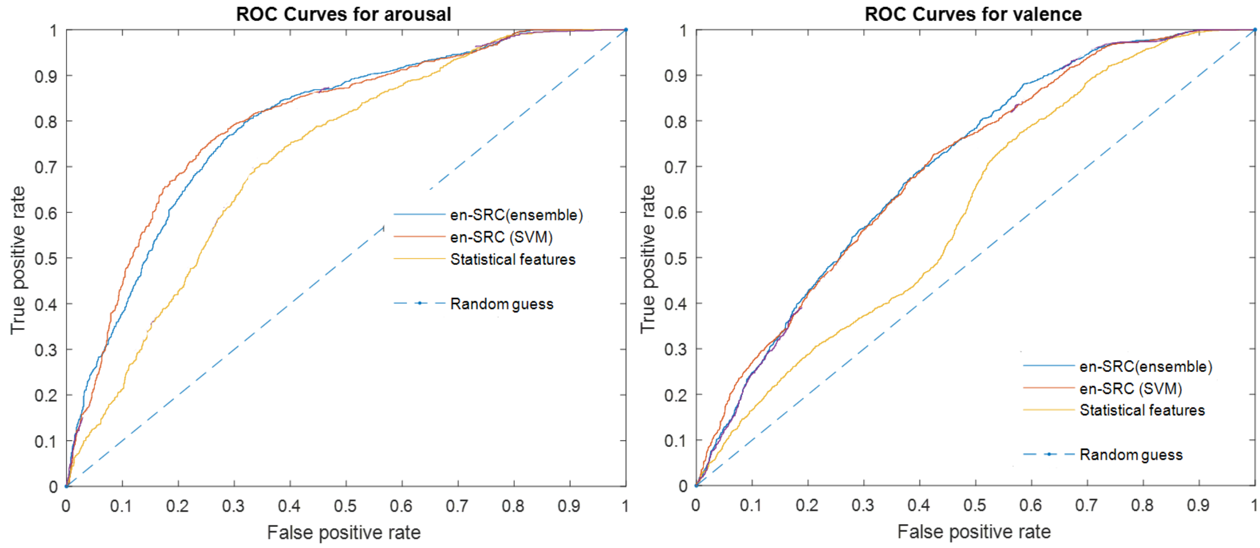

Fig. 1. ROC curves for different methods

Fig. 4 shows two Receiver Operating Characteristic (ROC) curves, one for arousal (left) and one for valence (right), comparing the performance of different classifiers in distinguishing between classes. The x-axis represents the false positive rate (FPR), and the y-axis represents the true positive rate (TPR). The curves include results for multiple classifiers: en-SRC (ensemble), en-SRC (SVM), and Statistical features againt the Random guess baseline represented by the diagonal line. The ROC curve for arousal demonstrates that en-SRC, outperform other methods, showing a higher true positive rate across various false positive rates. Similarly, for valence, the ensemble-based classifiers exhibit superior performance compared to SVM and statistical features, but with closer competition between methods. The curves indicate that en-SRC is generally more effective, achieving higher discrimination performance, as evidenced by their upward trend toward the top-left corner of the ROC space.

4. Discussion

The results of our investigation generally indicate that the best performance can be achieved by the sparse models when a couple of classifiers were jointly employed. Previous studies have primarily focused on the extraction of statistical features, power-band waves, as well as non-linear features. Most of these approaches lead to the representation of non-distinctive features, which are not useful for classification of the affective states. However, estimation of reconstruction coefficients based on the proposed models allowed extraction of more distinctive features compared with the previous approaches as confirmed by Figs. 1 and 4 characterizing the emotional dimensions.

The results also indicate that ensemble-based sparse representation classifiers (en-SRC) demonstrate superior performance, particularly in distinguishing arousal levels, compared to other classification methods, including statistical features and traditional SVM-based approaches. Notably, 78% classification accuracy was achieved for arousal using the en-SRC model with statistical features, outperforming several alternative methods. However, the accuracy for valence was lower, which aligns with findings from previous studies, such as Menezes et al. [9], who achieved higher accuracy using spectral power density features and SVM classifiers. These findings underscore the importance of ensemble learning and sparse models in analyzing EEG-based emotional states, particularly in scenarios involving class overlap or imbalanced datasets. While the proposed method improves arousal classification, further refinements—such as integrating spectral or nonlinear features—are necessary for valence.

It is important to point out that the box-plots illustrated in Fig. 1 show considerable overlap between classes. Therefore, while the nonlinear, SPD, and the ensemble deep learning outperformed the proposed approach in classifying the valence dimension, we believe the categorization of the samples may not be regarded as a routine task. Notably, based on our assessment, the results of the latter three methods regarding the valence dimension was not robustly reproducible. Notably, given the imbalance associated with the classes depicted by the box plots of Fig. 2, the application of conventional classification schemes to the DEAP dataset is challenging. This assertion has also been confirmed in the original paper introducing the 'DEAP' dataset [21]. Moreover, we have applied the output of the sparse models to a couple of distinctive classifiers. As each classifier processes the inputs differently, combining their outcomes amounts to fusing of the outputs to obtain the final decision, which tends to improve the final classification accuracy. The proposed approach allowed identification of additional discriminative features to improve the classification accuracy. The additional discriminative features are incorporated by selecting an equal number of samples from each class as the atoms of the dictionaries, so as to increase the between-class variance of the samples. At the same time, the reconstruction coefficients are chosen such that they exhibit a structural relationship to the class of the data. Accordingly, in the en-SRC model, the lack of flexibility in the dictionary will prevent the reduction of the reconstruction error. As a result, by the proper selection of the atoms and fixing them during the training phase, we forced the model to maintain the discriminative property underlying the initial collection of atoms corresponding to the class labels. This approach, which is inspired by the work presented by Wright *et al.* [18] essentially combines separate classifiers. Table 2 indicates that the reconstruction error for both emotional levels is minimized removing the distinction between the two dimensions. Our findings, therefore, verify that the representation residual is not a reliable metric for classification of the DEAP dataset. To overcome this limitation, the proposd model forces the dictionary learning procedure to select the dictionary atoms from the most discriminant training samples. Finally, this work provides a foundation for developing more accurate and generalized emotion recognition systems, which have significant applications in areas like affective computing, brain-computer interfaces (BCIs), and mental health assessment.

5. Conclusion

A sparse model for automatic identification of human affective states has been proposed. The proposed sparse model is equipped with external classifiers such as SVM, naïve bayes, and KNN to increase the classification capability of the model. Sparse representation of input samples is shown to properly increase the discrimination property of the inputs. The structured dictionaries have been chosen based on the training samples leading to improved classification accuracy. The proposed model has enabled selection of features reflecting the dynamical variation of EEG signals arising from basic emotional dimensions, leading to promising classification results based on the DEAP dataset. To increase the between-class variance further our future investigations will be devoted to improving the classification accuracy by refining the approach to structured dictionary learning.

References

[1] Luo, X., Xie, T., Qin, W., Duan, Z., Tan, J., & Zhang, T. (2025). Combining hierarchical sparse representation with adaptive prompt for few-shot segmentation. *Expert Systems with Applications*, *260*, 125377.

[2] Faraji, P., & Khodabakhshi, M. B. (2023). CollectiveNet-AltSpec: A collective concurrent CNN architecture of alternate specifications for EEG media perception and emotion tracing aided by multi-domain feature-augmentation. *Neural Networks*, *167*, 502-516.

[3] Meng, M., Yin, X., She, Q., Gao, Y., Kong, W., & Luo, Z. (2021). Sparse representation-based classification with two-dimensional dictionary optimization for motor imagery EEG pattern recognition. *Journal of Neuroscience Methods*, *361*, 109274.

[4] Khodabakhshi, M. B., & Saba, V. (2020). A nonlinear dynamical approach to analysis of emotions using EEG signals based on the Poincaré map function and recurrence plots. *Biomedical Engineering/Biomedizinische Technik*, *65*(5), 507-520.

[5] Huang, W., Yan, G., Chang, W., Zhang, Y., & Yuan, Y. (2023). EEG-based classification combining Bayesian convolutional neural networks

with recurrence plot for motor movement/imagery. *Pattern Recognition*, *144*, 109838.

[6] Prabhakar, S. K., & Lee, S. W. (2022). Improved sparse representation based robust hybrid feature extraction models with transfer and deep learning for EEG classification. *Expert Systems with Applications*, *198*, 116783.

[7] André-Obadia, N., Zyss, J., Gavaret, M., Lefaucheur, J. P., Azabou, E., Boulogne, S., ... & Delval, A. (2018). Recommendations for the use of electroencephalography and evoked potentials in comatose patients. *Neurophysiologie clinique*, *48*(3), 143-169.

[8] Kumar, N., Khaund, K., & Hazarika, S. M. (2016). Bispectral analysis of EEG for emotion recognition. *Procedia Computer Science*, *84*, 31-35.

[9] Menezes, M. L. R., Samara, A., Galway, L., Sant'Anna, A., Verikas, A., Alonso-Fernandez, F., ... & Bond, R. (2017). Towards emotion recognition for virtual environments: an evaluation of eeg features on benchmark dataset. *Personal and Ubiquitous Computing*, *21*(6), 1003-1013.

[10] Yin, Z., Zhao, M., Wang, Y., Yang, J., & Zhang, J. (2017). Recognition of emotions using multimodal physiological signals and an ensemble deep learning model. *Computer methods and programs in biomedicine*, *140*, 93-110.

[11] Zheng, W. L., Zhu, J. Y., & Lu, B. L. (2017). Identifying stable patterns over time for emotion recognition from EEG. *IEEE transactions on affective computing*, *10*(3), 417-429.

[12] Padhmashree, V., & Bhattacharyya, A. (2022). Human emotion recognition based on time–frequency analysis of multivariate EEG signal. *Knowledge-Based Systems*, *238*, 107867.

[13] Gruszczyńska, I., Mosdorf, R., Sobaniec, P., Żochowska-Sobaniec, M., & Borowska, M. (2019). Epilepsy identification based on EEG signal using RQA method. *Advances in medical sciences*, *64*(1), 58-64.

[14] Sohrabi, F., & Khodabakhshi, M. B. (2019). The trajectory intersection: An approach for nonlinear down-sampling. *Chaos, Solitons & Fractals*, *124*, 10-17.

[15] Marsousi, M., Abhari, K., Babyn, P., & Alirezaie, J. (2014). An adaptive approach to learn overcomplete dictionaries with efficient numbers of elements. *IEEE Transactions on Signal Processing*, *62*(12), 3272-3283.

[16] Engan, K., Aase, S. O., & Husoy, J. H. (1999, March). Method of optimal directions for frame design. In *1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No. 99CH36258)* (Vol. 5, pp. 2443-2446). IEEE.

[17] Aharon, M., Elad, M., & Bruckstein, A. (2006). K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on signal processing*, *54*(11), 4311-4322.

[18] Wright, J., Yang, A. Y., Ganesh, A., Sastry, S. S., & Ma, Y. (2008). Robust face recognition via sparse representation. *IEEE transactions on pattern analysis and machine intelligence*, *31*(2), 210-227.

[19] Jiang, Z., Lin, Z., & Davis, L. S. (2013). Label consistent K-SVD: Learning a discriminative dictionary for recognition. *IEEE transactions on pattern analysis and machine intelligence*, *35*(11), 2651-2664.

[20] Yang, M., Zhang, L., Feng, X., & Zhang, D. (2014). Sparse representation based fisher discrimination dictionary learning for image classification. *International Journal of Computer Vision*, *109*(3), 209-232.

[21] Koelstra, S., Muhl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T., ... & Patras, I. (2011). Deap: A database for emotion analysis; using physiological signals. *IEEE transactions on affective computing*, *3*(1), 18-31.

[22] Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, *25*(1), 49-59.

[23] Liu, Y., & Sourina, O. (2014, October). EEG-based subject-dependent emotion recognition algorithm using fractal dimension. In *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 3166-3171). IEEE.

[24] Soroush, M. Z., Maghooli, K., Setarehdan, S. K., & Nasrabadi, A. M. (2018). A novel method of eeg-based emotion recognition using nonlinear features variability and Dempster–Shafer theory. *Biomedical Engineering: Applications, Basis and Communications*, *30*(04), 1850026.

[25] Hatamikia, S., & Nasrabadi, A. M. (2014, November). Recognition of emotional states induced by music videos based on nonlinear feature extraction and som classification. In *2014 21th Iranian conference on biomedical engineering (ICBME)* (pp. 333-337). IEEE.

[26] Marwan, N., Romano, M. C., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics reports*, *438*(5-6), 237-329.

[27] Zangeneh Soroush, M., Maghooli, K., Setarehdan, S. K., & Motie Nasrabadi, A. (2018). Emotion classification through nonlinear EEG analysis using machine learning methods. *Int. Clin. Neurosci. J*, *5*(4), 135-149.

[28] Khodabakhshi, M. B., & Moradi, M. H. (2017). The attractor recurrent neural network based on fuzzy functions: An effective model for the classification of lung abnormalities. *Computers in biology and medicine*, *84*, 124-136.

[29] Bahari, F., & Janghorbani, A. (2013, December). Eeg-based emotion recognition using recurrence plot analysis and k nearest neighbor classifier. In *2013 20th Iranian Conference on Biomedical Engineering (ICBME)* (pp. 228-233). IEEE.

Biography

Mohammad Bagher Khodabakhshi recieved his Bachelor's degree in Biomedical Engineering from the Faculty of Engineering at Shahed University in 2008. He also received his Master's and Ph.D. degrees in 2011 and 2017 from Iran University of Science and Technology and Amirkabir University of Technology, respectively. During his Ph.D. degree Dr. Khodabakhshi also served as a researcher at Technology Services Complex at Sharif University. Currently, he is a faculty member of the Biomedical Engineering group and the head of the Medical Equipment and Bioinstrumentation lab at Hamedan University of Technology, Hamedan, Iran. In 2023, he joined the faculty of Shahed University, Tehran, Iran, where he is currently an Assistant Professor in biomedical engineering. His research interests include medical instrumentation, biological signal and image analysis, deep neural networks, and their applications in the medical signal and image processing.

Shahriar Jamasb received the B.S. and M.S. degrees in electrical engineering, and the Ph.D. degree in biomedical engineering from the University of California, Davis. His graduate study in electrical engineering concerned analog and digital integrated circuit design for implementation of a decision feedback equalizer, and his Ph.D. dissertation research focused on modeling and correction of drift in pH-sensitive ISFETs. He has over thirteen years of experience as a Device engineer, Manager of the device modeling group, Technical Manager, and Electronic design principal engineer at Samsung Semiconductor Inc., San Jose, CA, Cirrus Logic Inc., Fremont, CA, Commquest, an IBM company, Encinitas, CA, and Conexant Systems Inc., Newport Beach, CA respectively. After receiving his Ph.D. degree Dr. Jamasb also served as a lecturer at the University of California, Irvine, and the University of California, San Diego. In 2010, he joined the faculty of Hamedan University of Technology, Hamedan, Iran, where he is currently an Associate Professor in biomedical engineering. His principal research interests include integrated sensors for biomedical applications, solid-state device physics, instability phenomena, and applications of artificial intelligence in medicine.