Expert Systems With Applications 47 (2016) 35–41

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers☆

John Atkinson[a,*], Daniel Campos[b]

[a] Department of Computer Sciences, Faculty of Engineering, Universidad de Concepcion, Concepcion, Chile
[b] Artificial Intelligence Laboratory, Department of Computer Sciences, Universidad de Concepcion, Chile

ARTICLE INFO

Keywords:
Emotion recognition
Brain–Computer Interfaces
EEG
Feature selection
Emotion classification

ABSTRACT

Current emotion recognition computational techniques have been successful on associating the emotional changes with the EEG signals, and so they can be identified and classified from EEG signals if appropriate stimuli are applied. However, automatic recognition is usually restricted to a small number of emotions classes mainly due to signal's features and noise, EEG constraints and subject-dependent issues. In order to address these issues, in this paper a novel feature-based emotion recognition model is proposed for EEG-based Brain–Computer Interfaces. Unlike other approaches, our method explores a wider set of emotion types and incorporates additional features which are relevant for signal pre-processing and recognition classification tasks, based on a dimensional model of emotions: *Valence* and *Arousal*. It aims to improve the accuracy of the emotion classification task by combining mutual information based feature selection methods and kernel classifiers. Experiments using our approach for emotion classification which combines efficient feature selection methods and efficient kernel-based classifiers on standard EEG datasets show the promise of the approach when compared with state-of-the-art computational methods.

1. Introduction

Emotions play a critical role in rational decision-making, perception, human interaction, and human intelligence. Hence emotions are a fundamental component of being human as they motivate action and add meaning and richness to virtually all human experience. Traditionally, in *Human–Computer Interaction* (HCI), users must discard their emotional selves to work efficiently and rationality with computers (Sourina, Wang, Liu, & Nguyen, 2011; Wright, 2010).

Interfacing directly with the human brain is made possible through the use of sensors that can monitor some of the physical processes that occur within the brain that correspond with certain forms of thought. Researchers have used these technologies to build *Brain–Computer Interfaces* (BCIs), communication systems that do not depend on the brain's normal output pathways of peripheral nerves and muscles (Calvo & D'Mello, 2010). Instead, users explicitly manipulate their brain activity that can be used to control computers or communication devices.

State-of-the-art emotion recognition computational techniques have been successful on associating the emotional changes with the EEG signals, and so they can be identified and classified from EEG signals if appropriate stimuli are applied. However, automatic recognition is usually restricted to a small number of emotions classes mainly due to signal's features and noise, EEG constraints and subject-dependent issues.

Accordingly, in this research a novel feature-based emotion recognition model is proposed for EEG-based BCI interfaces. Unlike other approaches, our research explores a wider set of emotion types, claiming that combining a mutual information based feature selection method (i.e., *minimum-Redundancy-Maximum-Relevance*) and kernel classifiers may improve the accuracy of the emotion classification task.

This work is organized as follows: Section 2 describes the fundamentals and state-of-the-art emotion recognition techniques, Section 3 proposes a novel feature-based model for EEG emotion recognition, Section 4 discusses the main experiments conducted and the results for different model settings and finally, Section 5 highlights the main conclusions of the research and some further work.

2. Emotions recognition

Research of human emotional states via physiological signals involves recording and statistical analysis of signals from central and

☆ This research was partially supported by FONDECYT, Chile under Grant number 1130035: "*An Evolutionary Computation Approach to Natural-Language Chunking for Biological Text Mining Applications*."

* Corresponding author. Tel. :+56 412204305.
E-mail addresses: atkinson@inf.udec.cl (J. Atkinson), dancamposf@inf.udec.cl (D. Campos).

http://dx.doi.org/10.1016/j.eswa.2015.10.049
0957-4174/

parietal cortex. A popular physiological signal that is highly adopted for human emotion assessment is the EEG, etc. Unlike other physiological signals, EEG is a non-invasive technique with good temporal and acceptable spatial resolution. Thus, EEG might play a major role on detecting an emotion directly from the brain at higher spatial and temporal resolution (Yisi, Sourina, & Minh, 2010).

A major problem with recognizing emotions is that people have different subjective emotional experiences as responses to the same stimuli (Wright, 2010; Yisi et al., 2010). Accordingly, emotions can be classified into two taxonomy models:

(1) **Discrete model:** it is based on evolutionary features (Calvo & D'Mello, 2010) that include basic emotions (*happiness, sadness, fear, disgust, anger, surprise*), and mixed emotions such as Motivational (*thirst, hanger, pain, mood*), Self-awareness (*shame, disgrace, guilt*), etc.
(2) **Dimensional model:** it is expressed in terms of two emotions provoking people: *Valence* (*disgust, pleasure*) and *Arousal* (*calm, excitement*) Yisi et al. (2010).

Emotion recognition enables systems to get non-verbal information from human subjects so as to put events in context based on underlying captured emotions. Humans are capable of recognizing emotions either from speech (voice tone and discourse) with an accuracy around 60% or from facial expressions and body movements with an accuracy of 78–90%. However, the recognition task is strongly dependent on the context and requires facial expressions to be deliberately performed or even in a very exaggerated manner, which is far away from the natural way a user interact with intelligent interfaces.

Other kinds of techniques use audio signals, obtaining classification accuracy close to 60–90% (Calvo & D'Mello, 2010), whereas some other methods use non-linguistic vocalizations (i.e., laughs, tears, screams, etc.) to recognize complex emotional states such as anxiety, sexual interest, boredom. Bi-modal methods also combine audio inputs and facial expressions based on the assumption that a human emotion can trigger multiple behavior and physiological responses whenever he/she experiences this emotion.

Nevertheless, most of these methods require humans to express their emotional (mind) states in a deliberated and exaggerated manner, so that emotions cannot spontaneously be expressed. On the other hand, extracting information from facial expressions requires monitoring a subject by using one of several cameras, whereas for audio-based approaches, emotions are very hard to recognize whenever a subject does not speak or produce any sounds (Giakoumis, Tzovaras, Moustakas, & Hassapis, 2011; Sourina et al., 2011).

A popular and effective non-invasive technique to measure changes on brain activity is called (EEG), which transforms brain activity into images of variations of electrical potential by using small low-cost devices (AlMejrad, 2010). There are several approaches for EEG-based emotion recognition which are usually based on four main tasks (Calvo & D'Mello, 2010):

(1) *Signal preprocessing:* an EEG device can directly get signals from the brain. However, there are some noise sources that are not neurologically produced known as artifacts (i.e., blinking, muscular effects, vascular effects, etc.), so digital signal processing techniques must be applied to represent signals using frequencies and harmonic functions (Petrantonakis & Hadjileontiadis, 2010; Yisi et al., 2010).
(2) *Feature extraction:* EEG signals are highly dimensional so computational processing becomes very complex. Hence different features must be extracted in order to simplify the further emotion classification task so to create input *Feature Vectors* (FV). Typical methods include *statistical* metrics of the signal's first difference (i.e., median, standard deviation, kurtosis symmetry, etc.), **spectral density** (i.e., EEG signals with specific frequency bands) Zhang, Yang, and Huang (2008), **Logarithmic Band Power (Log BP)** (i.e., power of a band within the signal based on its oscillatory processes) Brunner, an C. Vidaurre, and Neuper (2011), **Hjorth parameters** (i.e., EEG signals described by *activity, mobility* and *complexity*) Zhang et al. (2008), **wavelet transform** (i.e., decomposition of the EEG signal) Petrantonakis and Hadjileontiadis (2010), **fractal dimension** (i.e., complexity of the fundamental patterns hidden in a signal) Zhang et al. (2008).
(3) *Feature selection:* one little used technique of feature selection for emotions recognition combines a metaheuristic method known as **Genetic Algorithms** (GA) and a **Support Vector Machines** (SVM). This **GA-SVM** approach heuristically searches for the best sets of features initially represented as chromosomes of features which evolves as the GA goes on, so that these can then be provided as an input to an SVM classifier (Wang et al., 2011). A major drawback with this method is the time spent to converge toward good results and the redundancy of the selected features assessed in each iteration of the GA.
In order to deal with this issue, other EEG feature selection technique known as **minimum-Redundancy-Maximum-Relevance** (*mRMR*) selects the features that correlate the strongest with a classification variable, reducing information redundancy. This method selects features that are mutually different from each other while still having a high correlation make up the selection task of *mRMR* (Polat & Cataltepe, 2012), by reducing redundancy between bad and good features using *Mutual Information* (MI) methods, so that a subset of features that represents best the dataset can be obtained.
(4) *Emotions classification:* once the FVs are extracted from the previous task, emotions must be classified according to previously identified classes of emotions. Despite the large number of features used by these methods, no feature selection is usually carried out. There are plenty of state-of-the-art classifiers for automatic emotion identification. For example, Nearest Neighbor classifiers used features such as FFT and Wavelets to recognize 4 types of emotions (i.e., *joy, sad, angry, relaxed*) achieving accuracies ranging from 54% to 67%. On the other hand, statistical methods such as *Quadratic Discriminant Analysis* (QDA) used several statistical features for negative and positive arousal levels with an average accuracy of 63% (Koelstra et al., 2012; Petrantonakis & Hadjileontiadis, 2010; Wu et al., 2010; Yisi et al., 2010).

3. An adaptive BCI-based emotions recognition model

In this work, a novel approach that combines *minimum-Redundancy-Maximum-Relevance* (*mRMR*) based feature selection tasks and kernel classifiers for emotions recognition is proposed. The method takes EEG signals received from BCI devices and incorporates relevant features in order to detect several kinds of emotional states by using state-of-the-art classifiers. The main contribution of this research is that unlike other automatic emotion recognition methods our approach

(1) Incorporates a feature selection task into the classification task.
(2) Uses multi-label classifiers to simultaneously recognize a wider range of emotion types based on a dimensional model.

The overall model is composed of three tasks: signal preprocessing, feature extraction and selection, and emotions classification (see Fig. 1).

3.1. EEG signal preprocessing

In order to train the emotions classifier, a set of previously emotion-labeled EEG data extracted from subjects self-assessing

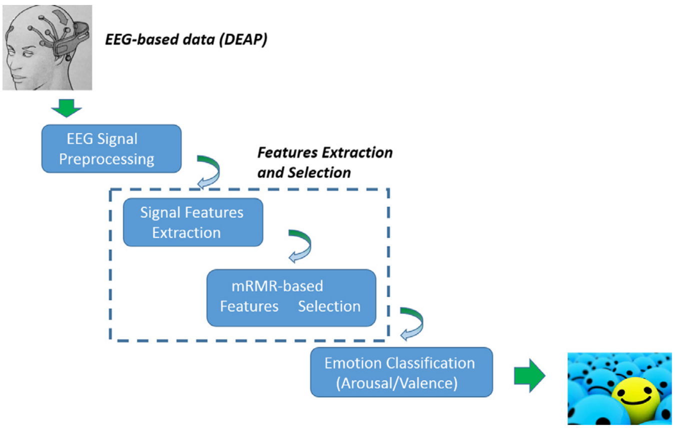

Fig. 1. Steps in our emotions recognition approach.

their emotional states was taken. It *Arousal* and *Valence* dimensions that were triggered from external stimuli. Since EEG brain signals contain much noise, the following basic preprocessing steps were performed:

- **Resolution reduction:** it optimizes the used memory by reducing a signal resolution. Since useful data for emotions recognition are found under 40 Hz (Yisi et al., 2010), resolution can be reduced to 128 Hz, preserving the original signal's information.
- **Electrooculography removal:** *electrooculography* (EOG) measures the corneo-retinal standing potential that exists between the front and the back of the human eye. In order to remove the noise produced from this kind of eyes movement, a method for removing EOG artifacts in the EEG called *Automatic Removal of Ocular Artifacts* is applied.
- **Band filter**: it filters EEG signals by generating bands that are useful for emotion recognition (e.g., 4 Hz–45 Hz).

3.2. Feature extraction and selection

EEG signals are highly dimensional data which may contain a lot of useless features. In order to reduce dimensionality, a large set of relevant features are extracted to create easy-to-process FVs for each stimuli. These included statistical features (S), band power (BP) for different frequencies, Hjorth parameters (HP) and fractal dimension (FD) for each channel. Statistical features included median, standard deviation, kurtosis coefficient, etc. Furthermore, bands of frequency for each EEG channel correspond to *theta* (4–8 Hz), *low alpha* (8–10 Hz), *alpha* (8–12 Hz), *beta* (12–30 Hz) and *gamma* (30–45 Hz).

In order to select a relevant set of features from the previously extracted candidate features so that further classification can be more accurate, the **minimum-Redundancy-Maximum-Relevance** (*mRMR*) method was used (Wu et al., 2010; Yisi et al., 2010). It selects the features that correlate the strongest with the classification variable, reducing information redundancy between bad and good features using *Mutual Information* (MI) methods, so that the best set of features can be selected. It is based on two underlying conditions: *minimum redundancy* and *maximum relevance*. Let S be a set of features, the *minimum redundancy* condition is defined as:

$$\min W_I, W_I = \frac{1}{|S|^2} \sum_{f_i, f_j \in S} I(f_i, f_j) \tag{1}$$

Where $I(f_i, f_j)$ is the MI between features f_i and f_j, and $|S| = n$ is the number of features from the set. The discriminant power of each feature regarding the emotion classes is then measured as the MI between features and classes. Since $I(C, f_i)$ expresses the relevance of feature f_i for a class C, the *maximum relevance* condition can be seen as:

$$\max V_I, V_I = \frac{1}{|S|} \sum_{f_i \in S} I(C, f_i) \tag{2}$$

Thus, finally obtained sets must accomplish the optimization conditions for Eqs. (1) and (2) simultaneously, into a single function, where the first and second condition are named `MID` and `MIQ`, respectively ($max(V_I - W_I)$ and $max(V_I/W_I)$). In addition, each feature was converted into a discrete value by using the transformation function of Eq. (3), where μ is the median of a subject's feature values, and σ is the standard deviation of values for the same feature (a common value of $\alpha = 0.5$ is used).

$$f(x) = \begin{cases} 1 & \text{if } x \geq \mu + \dfrac{\sigma}{2} \\ 0 & \text{if } \mu - \dfrac{\sigma}{2} \leq x < \mu + \dfrac{\sigma}{2} \\ -1 & \text{if } x < \mu - \dfrac{\sigma}{2} \end{cases} \tag{3}$$

3.3. Emotions classification

In order to recognize different emotion classes, a multi-class *Support Vector Machine* (SVM) was trained for a set $E = \{\vec{x_i}, y_i\}_{i=1}^{N}$, where N is the number of samples built from previously selected features, $\vec{x_i}$ is composed of an FV and y_i the dimension class of $\vec{x_i}$ (i.e., *Arousal* and *Valence*). The classifier builds and trains $k(k-1)/2$ SVMs, where k is the number of classes for y_i. For our approach, three classes were considered for each dimension. Each of three SVM uses RBF kernels and the overall classification is then carried out by using a **One-versus-One** voting mechanism in which a finally assigned class label will become those having the higher accuracy among the voting SVMs. Classes produced for each dimension are divided according to a range of values for each dimension [1, 9], into three sets: [1, 3.66], [3.66, 6.33] and [6.33, 9] based on Eq. (4), where $r(i)$ indicates the point in which the i-th division of the range of values is created, $\max v - \min v$ is the difference between the maximum and minimum value for each dimension, and k is the number of sets to be created.

$$r(i) = \left(i * \frac{\max v - \min v}{k} \right) + \min v \tag{4}$$

Previously trained multi-class SVMs are then used to classify *Arousal* and *Valence* dimension classes for unseen FVs extracted from different EEG signals extracted from the same subject as our model is subject-dependent.

4. Experiments

In order to assess the accuracy of our emotion classification approach into different dimensions, a computational prototype was built and run on DEAP datasets. Different experiments were conducted in order to tune different parameters of the finally implemented model. In addition, comparisons with other state-of-art methods were also performed and discussed.

Accuracy for different classifiers configurations was measured as the proportion of correctly classified signals versus the total number of signals. In the case of the SVM classifier, the performance is based on different types of kernel functions (Sourina et al., 2011). In addition, the *mRMR* feature selection method was used for tunning and training purposes (Wang et al., 2011).

All the experiments used the standard DEAP (*Dataset for Emotion Analysis using EEG, Physiological and Video Signals*)[1] dataset (Koelstra et al., 2012) which contains a set of EEG physiological signals and video records of 40 stimuli tests for 32 human subjects, i.e., 1280

[1] http://www.eecs.qmul.ac.uk/mmv/datasets/deap.

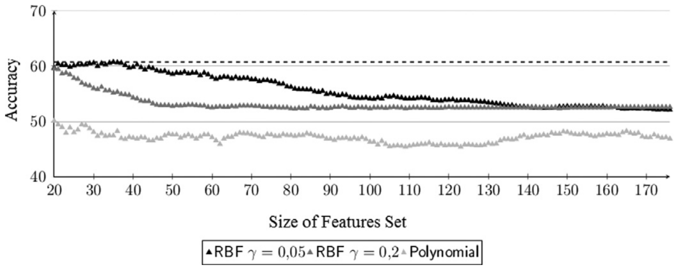

Fig. 2. Average classification accuracy for dimension *Arousal* and *mRMR*.

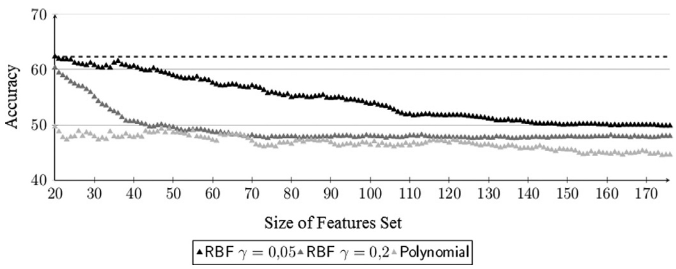

Fig. 3. Average classification accuracy for dimension *Valence* and *mRMR*.

stimuli tests each associated to their corresponding *Arousal* and *Valence* dimensions. Furthermore, training and testing the models were conducted by using *m-fold* cross-validation (with best results obtained for $m = 8$).

4.1. Data acquisition

Stimuli tests were conducted by using musical video records from DEAP as they are more suitable than other stimuli to evoke emotional reactions. The stimuli testing procedure was carried out as follows:

- Each subject watched a musical video as his/her EEG physiological signals and facial expressions are recorded.
- Each subject indicated his/her emotional state according to dimensions *Arousal* and *Valence*.

While DEAP data contain EEG signals extracted from a 32-channel BCI device, experiments only used information on 14 relevant channels.

4.2. Parameters tunning

In order to tune our model, *mRMR* was compared against other state-of-the-art features selection method such as GA-SVM. At the same time, three kernel configurations were tested for the SVM by using different kernel functions and degrees (RBF and $\gamma = 0.2$, RBF and $\gamma = 0.05$ and Polynomial and degree = 5).

Table 1
Parameters setting for *GA-SVM*-based feature selection.

Configuration	P_m	P_c	Selection	Crossover
C_1	1.0	0.80	*Roulette*	*One point*
C_2	0.8	0.66	*Roulette*	*One point*
C_3	1.0	0.80	*Tournament*	*One point*
C_4	0.8	0.66	*Tournament*	*One point*
C_5	1.0	0.80	*Roulette*	*Two points*
C_6	0.8	0.66	*Roulette*	*Two points*
C_7	1.0	0.80	*Tournament*	*Two points*
C_8	0.8	0.66	*Tournament*	*Two points*

Feature selection methods for different sets of candidate features were tested by its ability to find the best features for the same SVM classifiers. Overall results for *mRMR* can be seen in Figs. 2 and 3, for *Arousal* and *Valence* dimensions, respectively. Those results are obtained for SVMs using RBF kernels with $\gamma = 0.05$. Note the average accuracy slightly drops as a larger set of features is selected. The best classification accuracy was obtained for a set containing 35 and 29 features for *Arousal* and *Valence*, respectively (dotted lines), whereas worst accuracy for the configuration is obtained for set sized greater than 173 for both dimensions.

On the other hand, *GA-SVM* uses a GA to find an optimal subset of features, so the quality of evolved solutions is strongly dependent on how genetic operators modify initial hypotheses as the GA goes on: *mutation* (probability P_m), *crossover* (type and probability P_c), *parents selection* (i.e., *tournament, roulette*). Table 1 shows the different configurations for this task.

Preliminary runs showed the GA usually converges toward good solutions between 80 and 100 generations. Fitness evaluation of candidate solutions in the GA uses the classification accuracy (*acc*) of the SVM, so it is computed as seen in Eq. (5), where $fitness = accuracy$ where $w_a = 1$ (weight of *acc*), $w_f = 0$ (weight of the number and cost of the features), *n* is the number of features, c_i is the cost of extracting the *i-th* feature, and x_i represents the absence/presence of a selected featured (0 or 1).

$$\text{fitness} = w_a * acc + w_f * \left(\sum_{i=1}^{n} c_i * x_i \right)^{-1} \quad (5)$$

Classification results using *GA-SVM* feature selection for different population sizes, and RBF kernel ($\gamma = 0.05$) and dimensions, can be seen in Figs. 4 and 5, where C_i represents the $i-th$ configuration for the GA (Table 1). Unlike *mRMR* techniques, there is no a clear trend on the evolution of candidate solutions. However, as the size of the GA population increases, the average accuracy increases too. Hence best results were obtained for population sizes of 90 and 100 for *Arousal* and *Valence*, respectively. However, no significant differences are observed for both dimensions. Thus, best setting results are obtained using configurations C_2, C_4 and C_6.

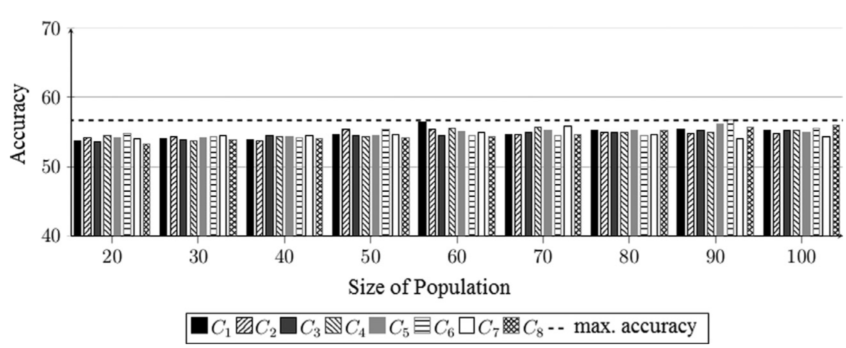

Fig. 4. Average accuracy for classifying *Arousal* dimension using *GA-SVM*.

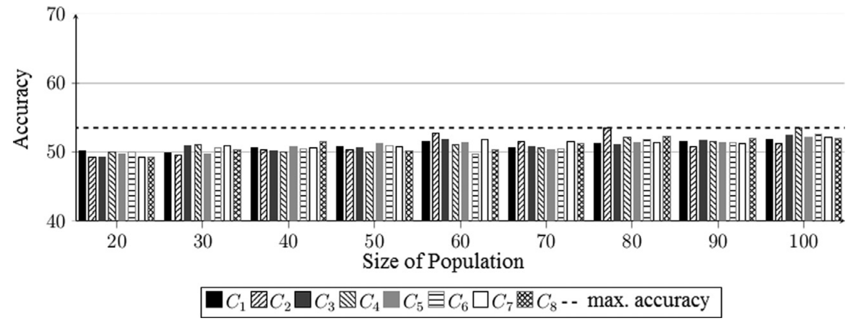

Fig. 5. Average accuracy for classifying *Valence* dimension using *GA-SVM*.

Table 2
Best results for setting parameters for different feature selection methods.

Method	Dimension	Accuracy (%)	Std. dev (%)	No. of features
mRMR	*Arousal*	60.72	9.08	35
mRMR	*Valence*	62.39	9.90	20
GA-SVM	*Arousal*	56.69	9.34	95
GA-SVM	*Valence*	53.46	9.05	94

Finally, Table 2 shows the best setting results for each dimension, feature selection method, and the number of selected features. The SVM classifier using RBF Kernel with $\gamma = 0.05$, produces the highest accuracy, and the performance of *mRMR* is better than *GA-SVM* for the selected features. Note also *mRMR* generates features sets that are smaller than for *GA-SVM*, which makes it more suitable for real-time applications as it requires less work to extract features and achieve good classification accuracy.

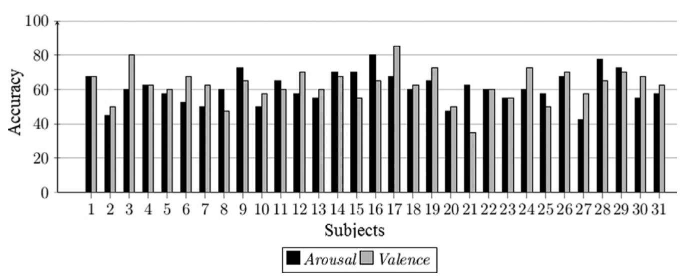

Fig. 6. Classification accuracy per dimension for each subject.

4.3. Overall evaluation

A final overall experiment compared our approach against some state-of-the-art methods. To this end, the best previously tunned configurations were used: *mRMR*-based feature selection, 35 features for dimension *Arousal* and 20 features for dimension *Valence*, and RBF kernel with $\gamma = 0.05$ for the SVM classifier. The model was then trained using 40 stimuli tests for each of the 31 subjects of the dataset.

Experimental results are shown in Fig. 6, indicating a median of 60.7% and 62.33%, for dimension *Arousal* and *Valence*, respectively (i.e., std. dev. of 9 is close to the median of both dimensions).

Graphic of Fig. 7 shows the classification accuracy for each dimension and subject. In addition, the lower row for each figure shows the number of subjects for whom certain features were selected, where darker points represent a larger number of subjects. This suggests there is no relationship between the selected features for one or other subject. Nevertheless, best selected features for both dimensions, correspond to the statistical measures extracted from each channel (left-hand side).

Classification accuracy of our model was also compared against other approaches, indicating very promising results when dealing with combination of methods and different classes of emotions as

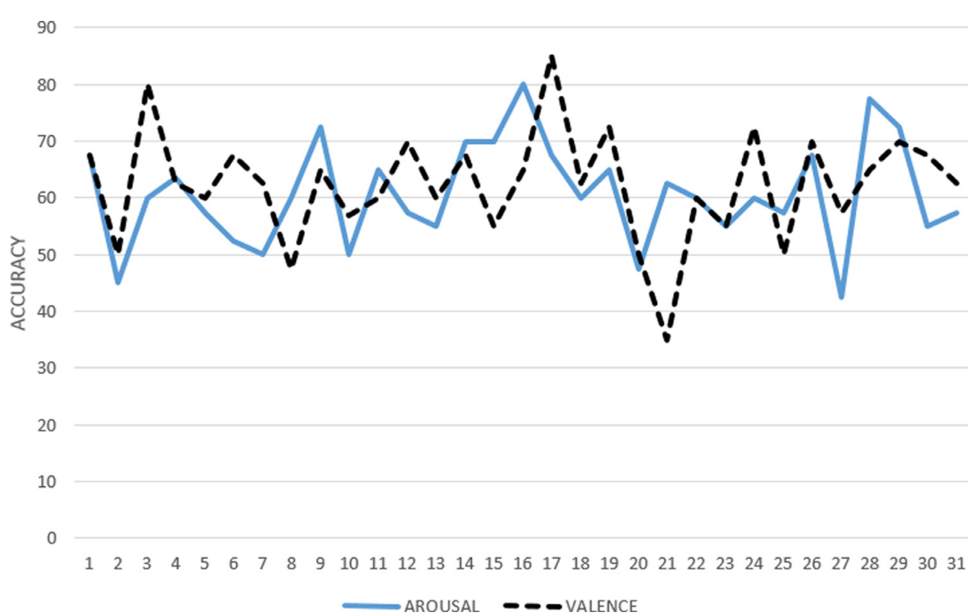

Fig. 7. Classification accuracy of the proposed model for each subject.

Table 3
Comparing our recognition approach and some state-of-the-art methods.

Method	No. of classes per dimension	Accuracy (Arousal) (%)	Accuracy (Valence)
Our model	2	73.06	73.14%
Our model	**3**	**60,7**	**62,33%**
Our model	5	46.69	45.32%
Spectral density and SVM	3	96.5	–
Band power and Naive Bayes	2	62	67.6%

seen in Table 3, even for recognition using two classes per dimension (high/low). Experimental results for different approaches for EEG emotion recognition show our model's performance is similar to others but for a higher number of classes.

The table includes results for classifying 2 and 5 classes per dimension, indicating promising results for fair parameters settings. It suggests results for two classes per dimension are better than other approaches using the same DEAP dataset. Note also that the table also indicates that our model is capable of recognizing a wider range of emotion classes based on the dimensions (*Arousal* and *Valence*), with no need to use additional emotions classifiers. Hence our technique can recognize several emotion types simultaneously by using a single multi-class kernel classifier. Note that spectral density and SVM does well for recognizing emotions within the *Arousal* dimension, but it has not been assessed for more than 5 classes/dimensions as for our case.

5. Conclusions

In this paper, an EEG feature-based emotion recognition method was proposed. Unlike other approaches, the approach uses the *mRMR* feature selection method as a signal preprocessing step so as to improve the predictive accuracy of an SVM emotion classifier based on two-dimension emotions model (i.e., *Valence* and *Arousal*). In addition, compared with state-of-the-art emotion recognition methods, our approach deals with a higher number of emotion classes (i.e., 8) on a standard DEAP dataset, which makes the problem more realistic but at the same time, the training task becomes more demanding.

Accordingly, one of the contributions of this research is that it incorporates a statistical-based feature selection task into the classification task. Furthermore, our approach which combines feature selection and kernel classifiers uses multi-label classifiers to simultaneously recognize a wider set of emotion classes based on a dimensional emotion model.

In order to assess the effectiveness of our kernel-based classifier, several preliminary experiments were conducted so as to produce the best parameters settings. It included tunning feature selection methods, emotion classifier, signal preprocessing tasks, etc. Preliminary experiments showed that our *mRMR*-based feature selection method outperformed the most popular feature selection strategy (GA-SVM) for both dimensions (*Arousal* and *Valence*) when classifying emotions (Accuracy of 60.72% and 62.4% versus 57% and 53.4%). In addition, for both dimensions, our method reduced the number of relevant features of almost to capable of reducing in 63% with a higher accuracy. Classification accuracy of our model was then compared against other competitive current approaches to emotion recognition: SVM-based spectral density and Bayes-based Band Power (BP). An important issue with these two techniques is that either the EEG signal they analyse must be within very specific frequence bands (i.e., spectral density) or the power of the frequence band within a signal is strongly dependent on its oscillatory processes. Hence they might not very effective when attempting to classify a wider set of emotion classes (i.e., a higher number of classes per dimension).

Overall results showed that our methods outperformed those of the state-of-the-art for the same number of classes per dimension (i.e., 73% versus 62%). In addition, our approach was capable of classifying a higher number of classes per dimension whenever no other state-of-the-art method did it for em Valence (i.e., 62.33% versus no accuracy known in spectral density for 3 classes per dimension). Note that spectral density does it well for recognizing emotions within the *Arousal* dimension, but it has not been assessed for more than 5 classes/dimensions as for our research. Thus, our *mRMR*-based emotion classification approach outperformed other state-of-the-art methods. Furthermore, the method is promising when considering a higher number of classes per dimension (i.e. 3 and 5), that had not been proven in the literature. This also showed our method recognizes a higher number of emotion classes without using additional emotions classifiers.

Accordingly, combining features-selection methods (*mRMR*) and SVM classifiers using RBF kernels yield significant improvements in accuracy. In words, our method requires less work to classify based on a smaller set of selected so as to achieve higher accuracy than other techniques.

5.1. Further research

There are some open issues which may be addressed so as to produce more accurate results and robust emotion recognition methods, including:

- *Parameters design and analysis:* running time was a constraint when assessed different configurations for our model as analyzing EEG signals became a very demanding task. As a consequence, experiments were designed only for a small set of settings. Hence more exhaustive setup evaluations may be required on the model's parameters to evaluate the extent to which different frequence bands, number of subjects, EEG oscillations, etc. affect the effectiveness of the approach.
- *Specific-purpose training dataset creation:* while there are some training corpus for emotion recognition purposes such as DEAP (Koelstra et al., 2012), large-scale and well-balanced dataset are required so as to avoid bias and overfiting of the classification task.
- *Higher number of classes per dimension:* while recognizing a higher number of emotion classes (i.e., 3 or 5) was a main purpose of the proposed method, it might be not enough to deploy real-world EEG-based emotion recognition applications (i.e., videogames, brain-controlled wheelchairs, etc.) as these must adjust to a significantly big set of emotional states. Hence further experiments may be needed to modify the model so that it can effectively recognize more than 3 levels per dimension.

References

AlMejrad, A. (2010). Human emotions detection using brain wave signals: a challenging. *European Journal of Scientific Research, 44*, 640–659.

Brunner, C., Vidaurre, C., Billinger, M., & Neuper, C. (2011). A comparison of univariate, vector, bilinear autoregressive, and band power features for brain–computer interfaces. *Medical & Biological Engineering & Computing, 49*(11), 1337–1346.

Calvo, R., & D'Mello, S. (2010). Affect detection: an interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing, 1*(1), 18–37.

Giakoumis, D., Tzovaras, D., Moustakas, K., & Hassapis, G. (2011). Automatic recognition of boredom in video games using novel biosignal moment-based features. *IEEE Transactions on Affective Computing, 2*, 119–133. http://doi.ieeecomputersociety.org/10.1109/T-AFFC.2011.4.

Koelstra, S., Muhl, C., Soleymani, M., Lee, J., Yazdani, A., & Ebrahimi, T. (2012). Deap: a database for emotion analysis; using physiological signals. *IEEE Transactions on Affective Computing, 3*(1), 18–31.

Petrantonakis, P., & Hadjileontiadis, L. (2010). Emotion recognition from brain signals using hybrid adaptive filtering and higher order crossings analysis. *IEEE Transactions on Affective Computing, 1*(2), 81–97.

Polat, D., & Cataltepe, Z. (2012). Feature selection and classification on brain computer interface (BCI) data. In *Proceedings of signal processing and communications applications conference (SIU)* (pp. 1–4). IEEE.

Sourina, O., Wang, Q., Liu, Y., & Nguyen, M. (2011). A real-time fractal-based brain state recognition from eeg and its applications. In F. Babiloni, A. L. N. Fred, J. Filipe, & H. Gamboa (Eds.), *Biosignals* (pp. 82–90). SciTePress.

Wang, L., Xu, G., Wang, J., Yang, S., Guo, L., & Yan, W. (2011). GA-SVM based feature selection and parameters optimization for BCI research. In *Proceedings of seventh international conference on natural computation (ICNC): vol. 1* (pp. 580–583).

Wright, F. (2010). *Emotional instant messaging with the epoc headset,* (Master's thesis). Baltimore County: University of Maryland.

Wu, D., Courtney, C., Lance, B., Narayanan, S., Dawson, M., Kelvin, S., & Parsons, T. (2010). Optimal arousal identification and classification for affective computing using physiological signals: virtual reality stroop task. *IEEE Transactions on Affective Computing, 1*(2), 109–118.

Yisi, L., Sourina, O., & Minh, N. (2010). Real-time EEG-based human emotion recognition and visualization. In *Proceedings of international conference on cyberworlds (CW)* (pp. 262–269).

Zhang, A., Yang, B., & Huang, L. (2008). Feature extraction of eeg signals using power spectral entropy. In *Proceedings of international conference on biomedical engineering and informatics, 2008. BMEI: vol. 2* (pp. 435–439). IEEE.