Chapter 2

Artificial Intelligence for Affective Computing: An emotion recognition case study

Pablo Arnau-González[1], Stamos Katsigiannis[1], Miguel Arevalillo-Herráez[2], and Naeem Ramzan[1]

This chapter provides an introduction on the benefits of artificial intelligence techniques for the field of affective computing, through a case study about emotion recognition via brain (electroencephalography - EEG) signals. Readers are first provided with a general description of the field, followed by the main models of human affect, with special emphasis to Russell's Circumplex model and the Pleasure-Arousal-Dominance (PAD) model. Finally, an AI-based method for the detection of affect elicited via multimedia stimuli is presented. The method combines both connectivity-based and channel-based EEG features with a selection method that considerably reduces the dimensionality of the data and allows for efficient classification. In particular, the Relative Energy (RE) and its logarithm in the spatial domain, as well as the spectral power (SP) in the frequency domain are computed for the four typically used EEG frequency bands (α, β, γ and θ), and complemented with the mutual information measured over all EEG channel pairs. The resulting features are then reduced by using a hybrid method that combines supervised and unsupervised feature selection. Detection results are compared to state-of-the-art methods on the DEAP benchmarking dataset for emotion analysis, which is composed of labelled EEG recordings from 32 individuals, acquired while watching 40 music videos. The acquired results demonstrate the potential of AI-based methods for emotion recognition, an application that can significantly benefit the fields of human-computer interaction (HCI) and of quality-of-experience (QoE).

2.1 Introduction

Human-Computer Interfaces have evolved enormously in recent years, with new modalities for human-computer interaction (HCI) becoming available at increasingly lower cost. The combination of these new HCI solutions with powerful Artificial Intelligence (AI) algorithms is providing the means to add the *smart* tag to many

[1]University of the West of Scotland, UK
[2]Universitat de Valéncia, Spain

solutions for everyday problems. Proof of this trend in the consumer market is the proliferation of devices that work as or embed smart-assistants, e.g. Google Home, Apple's Siri, Amazon's Alexa. Other examples of new HCI modalities are the recent developments on Brain-Computer-Interfaces (BCI). BCIs cover a wide spectrum of applications, spanning from medical purposes to educational uses, like for example detecting focus levels [1]. Furthermore, it has been suggested that the study of brain signals (electroencephalography - EEG) could lead to the detection of the emotional state of individuals at any given moment [2]. The understanding and modelling of the user affective state could lead to huge advances in the fields of HCI and quality of experience (QoE), as already pointed out by Dr Rosalind Piccard [3]. Having computers that are aware of the user's emotional state, would enable these computers to react to it, improving the user experience by providing more relevant content, in the case of a smart assistant, or help in one way or the other in the case of an Intelligent Tutoring System (ITS).

Regardless the application, it is clear that AI techniques, such as supervised classification, are essential for detecting different emotions from the acquired brain signals [2, 4, 5], while affect is key to improve the user experience in many different areas. Efficient affect detection from brain signals is currently an open problem with numerous research works being conducted every year. Apart from solutions based on brain signals, there are also video-based affect detection approaches [6, 7]. However, weak emotions, i.e. emotions that are shown with very little intensity, are difficult to capture from video sources [8]. Therefore, research is generally focused in recognising emotions from sources that are affected even when the emotion is not publicly shown, as is the case with brain signals. In this direction, AI techniques are vital to create user-specific models for recognising emotions.

2.2 Models of human affect

In order to measure or detect an emotion or the affective state of an individual, the literature proposes a number of human affect models. These models study emotion following two different approaches, either focusing on the the emotion itself or on characteristics of the emotion. These models can be categorised as either discrete or continuous.

2.2.1 Discrete models of affect

Theorists have long discussed a small set of categories for describing emotional states. In 1962, Tomkins suggested that there are eight basic emotions [9]. Plutchik later proposed a different set of eight basic emotions: fear, anger, sorrow, joy, disgust, surprise, acceptance, and anticipation [10]. More recently, Ortony, Clore, and Collins collected a summary of lists of basic emotions [11].

2.2.1.1 Six Basic Emotions and FACS

Probably, the most prominent discrete model of emotions is the one proposed by Ekman. This model studies emotions via a discrete approach, suggesting that there is a

Table 2.1 Facial Action Coding System

Emotion	Activated AUs	Facial muscles contraction
Happiness	6+12	Cheek Raiser + Lip Corner Puller
Sadness	1+4+15	Inner Brow Raiser + Brow Lowerer + Lip Corner Depressor
Surpise	1+2+5+26	Inner Brow Raiser + Outer Brow Raiser + Upper Lid Raiser + Jaw Drop
Fear	1+2+4+5+7+20+26	Inner Brow Raiser + Outer Brow Raiser + Brow Lowerer + Upper Lid Raiser + Lid Tightener + Lip Stretcher + Jaw Drop
Anger	4+5+7+23	Brow Lowerer + Upper Lid Raiser + Lid Tightener + Lip Tightener
Disgust	9+15+16	Nose Wrinkler + Lip Corner Depressor + Lower Lip Corner Depressor
Contempt	12+14	Lip Corner Puller + Dimpler

reduced number of primary, or primitive, emotions, i.e. happiness, anger, fear, disgust, surprise and sadness [12, 13], and that all the other emotional states are nothing but combinations of these primary emotions. Ekman defines each of the basic emotions as not a single affective state but a family of states [14], where each member of the family shares certain characteristics. The justification of this affirmation is sustained in his previous work [15], where 60 different expressions of anger were specified. In that study, all the identified anger expressions shared a specific muscular pattern that was different from the patterns specified in other families, such as for disgust or happiness. Ekman also related the intensity of the emotion with the strength of the muscular contractions [16]. Under these assumptions, a model for mapping facial expression to emotions was proposed. The resulting Facial Action Coding System (FACS) [17] maps facial expressions, defined by fundamental contractions of facial muscles (Action Units, or AU). Table 2.1 provides a brief example of a FACS containing seven examples of emotions.

2.2.1.2 Plutchick's Wheel of emotion

Plutchick [10] proposed an alternative model of human affect. In his approach, emotions are categorised in three different categories (primary, secondary, and tertiary). The three categories are organised in a conical shape, turned upside down, where the emotions are located close, according to their relation (see Figure 2.1). Similarly to Ekman, Plutchik identified anger, disgust, sadness, surprise, fear, and joy, as primary emotions, but he also added trust and anticipation.

2.2.2 Continuous (dimensional) models of affect

Continuous models treat emotional states as states characterised by continuous variables in an N-dimensional space. Most notable models are Russell's Circumplex Model of affect, the Pleasure Arousal Dominance (PAD) model, and the Lövheim Cube of emotions.

2.2.2.1 Russell's Circumplex Model of Affect

Prior to Russell's proposal of a circumplex model of affect, other researchers had already proposed a similar system. Schlosberg [18] proposed that emotions are or-

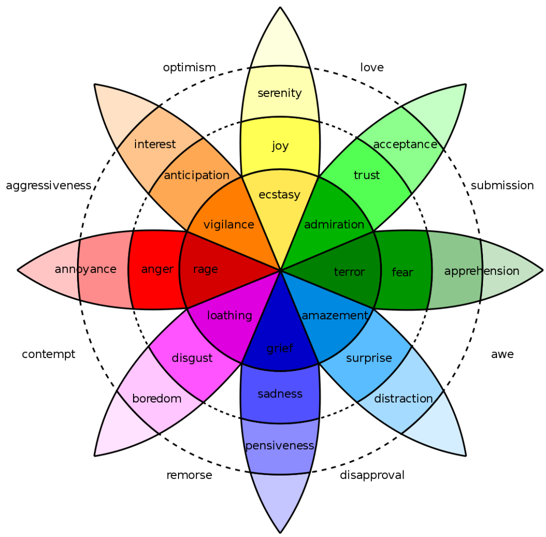


Figure 2.1 Plutchick's wheel of emotion, 2D projection. Source: Public domain image.

ganised in a circular manner, meaning that emotions were better represented in a bipolar space rather than in six mono-polar spaces, an assumption on which Russell based his model of affect, along with previous studies about the affective structure of the English language. Supporting the hypothesis of Schloberg that dimensions of evaluation, activity, and potency are major components of the meaning of the language [19], Russell's model of affect proposed that emotion can be located in a 2-Dimensional space defined by two traits: Arousal, ranging from inactivity to excitement, and Valence (positiveness), ranging from unpleasantness to pleasantness (see Figure 2.2).

2.2.2.2 The Pleasure-Arousal-Dominance model (PAD)

The Pleasure-Arousal-Dominance (PAD) model consists of a revision of 1979's Russell's circumplex model, representing a more modern version of it, with a few changes incorporated. Mehrabian and Russell [20] proposed that emotion can be located in a 3-Dimensional space. The new space is defined by the previously defined dimensions (Arousal, Valence) and includes a third dimension, Dominance, which defines the control the emotion has over the individual, i.e low dominance values would apply to emotions such as calm or joy, while strong emotions, such as love or fury, would have a high dominance value.

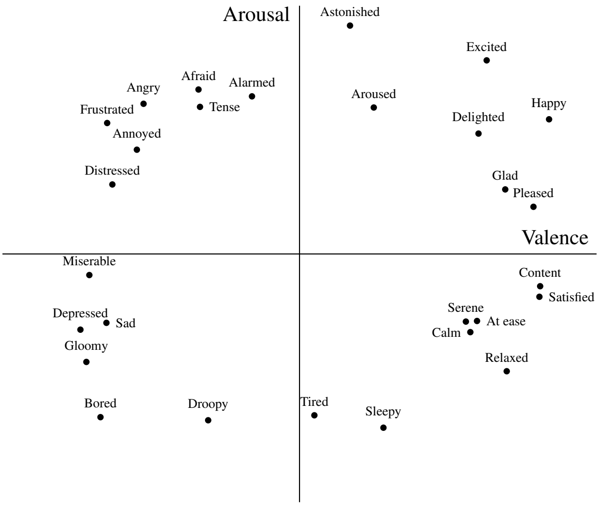


Figure 2.2 Russell's Circumplex model of Affect

2.2.2.3 Lövheim cube of emotion

A bio-chemistry-based model of affect was proposed by Lövheim [21] in 2011 that maps emotional states to a 3-Dimensional space defined by the combinations of the concentration levels of three mono-amines, i.e dopamine, noradrenaline, and serotonin. As discussed by the author in [21], the three dimensions do not match exactly with the ones described in Russell's model, but there are some common patterns observable (see Figure 2.3). This model, apart from being one of the newest in emotion theory, is very interesting since it explains emotion from a bio-chemical perspective.

2.3 Previous work on emotion recognition

One major difficulty in dealing with the evaluation of emotion recognition methods is related to the non-existence of a common ground truth that would allow a fair comparison between different proposals. This has lead many authors to test and report their results on proprietary datasets, limiting the impact of their proposals because of the intrinsic difficulty associated with assessing their performance in relation to other existing or newly developed methods. One major contribution in this direction was the DEAP dataset [22], a multimodal dataset specifically created for the analysis of human affective states. DEAP contains physiological recordings from 32 subjects while watching 40 music videos, which were selected in order to elicit emotions in each of the 4 quadrants of Russell's Circumplex Model [23]. Recordings were annotated with the associated emotional state via self-reporting, thus DEAP can be used as a baseline for benchmarking. The creators of DEAP performed an initial classification experiment for establishing a baseline performance [22]. They used the Spectral Power of single EEG channels and the Spectral Power Asymmetry from

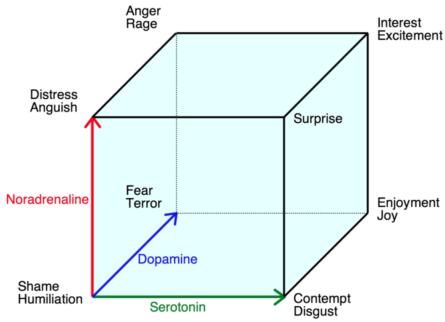


Figure 2.3 *Lövheim Cube of Emotion. Source: wikimedia.org. GNU Free Documentation License.*

14 pairs of electrodes, and selected the features according to a Fisher Discriminant Analysis with a threshold set at 0.3.

Previous works on emotion recognition from EEG signals mostly focused on channel-related features, e.g. [24]. As a first example, Liu et al. [25] proposed a single fractal model based on their observation that higher levels of arousal were usually related to higher values of the Fractal Dimension [26]; as much as valence levels relate to fractal dimension differences between concrete electrodes located in the right and left hemisphere of the scalp. This initial work was validated with their own data set, and later extended in [27] by using Higher Order Crossings [28] and features from the General Higuchi Fractal Dimension Spectra in order to understand EEG signals as multi-fractal signals. Other studies focused on a different type of features that considers the connectivity between the EEG electrodes. For example, Chen et al. [29] set a classification problem by extracting groups of such features and studied the performance of each set. In particular, they analysed Pearson's Correlation [30], Phase Coherence [31], and Mutual Information, which led to the best results. Other works, such as Gupta et al.'s [32], used graph-theoretic features to classify emotional states through support vector machines and relevance vector machine classifiers.

2.4 Datasets for emotion recognition

Since EEG signals are very dependent on the individual, solutions proposed by different authors are potentially not generalisable for other individuals. For this reason, different benchmarking datasets have been made publicly available over time for providing a common ground for testing different proposals. Up to now, five major datasets have been publicly released: DEAP [22], MAHNOB [33], DREAMER [34], AMIGOS[35] and SEED [36].

These datasets have been created by recording EEG and other physiological signals of different individuals, while being exposed to different emotional stimuli in the form of video sequences. MAHNOB, DREAMER, SEED, and AMIGOS used film excerpts, while DEAP used music videos. There is also difference in the EEG

Table 2.2 Overview of the different available datasets

Dataset	**Subjects**	**Video No.**	**Video content**	**Video duration**	**Recording device**	**Channel No.**	**Sampling frequency**
DEAP	32	40	Music videos	60 s	Biosemi Active II	32	512 Hz (downsampled to 256 Hz)
MANHOB	27	20	Excerpts from movies	34.9-117 s ($\mu = 81$ s)	Biosemi Active II	32	512 Hz (downsampled to 256 Hz)
DREAMER	23	18	Excerpts from movies	60 s	Emotiv EPOC	14	128 Hz
AMIGOS	40	$16 + 4$ (Two protocols)	Excerpts from movies	45s Long videos	Emotiv EPOC	14	128 Hz
SEED	15	15	Excerpts from movies	240 s	ESI NeuroScan	63	1000 Hz (downsampled to 200 Hz)

signal acquisition devices. DEAP and MAHNOB used the Biosemi Active II system, a non-medical-grade high performance EEG recording and monitoring device, SEED used a similar EEG recording system, the ESI NeuroScan, while DREAMER and AMIGOS used the Emotiv EPOC wireless EEG headset. A brief overview of these datasets is provided in Table 2.2.

As can be observed from Table 2.2, DEAP, MAHNOB and SEED have been recorded using high quality EEG-recording devices, while DREAMER and AMIGOS datasets have been recorded using consumer-grade devices, namely the Emotiv EPOC. The experimental protocol applied for data acquisition was very similar for all datasets. The participant would watch the stimuli in random order and rate the felt emotion immediately after exposure to each stimulus. An exception to this protocol is found in the SEED dataset, where participants did not provide the subjective ratings, rather the dataset is annotated with the emotions each video is supposed to elicit and video sequences were presented in a specific sequence.

Since the main purpose of these datasets is to help researchers model human emotions from EEG sources, the samples are annotated according to emotional models. Excluding SEED, the examined datasets contain recordings annotated according to Russell's model (see section 2.2.2.1 for more details) or a revised version of the model, such as the PAD model (see section 2.2.2.2). Annotations were acquired through self-reporting using a standard scale designed for the self-report of emotions, i.e. the Self Assesment Manikin (SAM) scale [37]. SAM is composed by three different sets of 5 manikins each, with the central one corresponding to the neutral emotion and the extremes corresponding to the highest and lowest possible value for each of the scales.

2.5 Proposed methodology

Our proposal is based on using both *connectivity features* and *energy features* simultaneously. While *energy features* provide information about how the energy is distributed across the EEG signal bands, *connectivity features* study the interactions between different EEG channels and enrich the data provided by the former. The combination of both types of features, along with the use of a feature reduction scheme that allows the classification to be applied in a low-dimensional space, endows the proposal with the ability to distinguish low and high levels of arousal and valence more accurately than other state-of-the-art methods recently reported in the emotion recognition literature.

2.5.1 Connectivity Features

According to Chen et al. [29], *mutual information* is a good indicator of the connectivity between EEG channels. *Mutual information* measures how informative a random variable is to another. Its calculation is based in entropy, which is calculated as:

$$H(X) = -\sum p_i \cdot \log p_i \tag{2.1}$$

where p_i is the probability of the i-th element of time series X. This expression allows to compute the *mutual information* between two signals X and Y as:

$$MI(X;Y) = H(X) - H(X|Y) \tag{2.2}$$

or alternatively as:

$$MI(X,Y) = -\sum p_{ij}^{XY} \cdot \log\left(\frac{p_{ij}^{XY}}{p_i^X p_j^Y}\right) \tag{2.3}$$

where p_{ij}^{XY} is the joint probability of the i-th element of time series X and j-th element of time series Y.

In this work, we have used *mutual information* as implemented by Moddemeijer [38], and for the replication of Chen's experiment [29], we have used the same toolbox [39]. In particular, we have applied the floor function on the original signals, which rounds each value to the nearest integer that is lower or equal than it.

2.5.2 Energy Features

For each EEG signal X and frequency band $f = \{\ \alpha$ (8-13 Hz), β (14-30 Hz), γ (30-47 Hz), θ (4-7 Hz)$\}$, the energy was extracted as:

$$E_f(X) = \sum X_i^2 \tag{2.4}$$

where X_i is the i-th element of signal X filtered in the frequency band f.

The Relative Energy (RE) and the Logarithmic Relative Energy (LRE) for each combination of channel and frequency band were then extracted using the following formulas:

$$RE_f(X) = \frac{E_f(X)}{E_\alpha(X) + E_\beta(X) + E_\gamma(X) + E_\theta(X)} \tag{2.5}$$

$$LRE_f(X) = \log(RE_f(X)) \tag{2.6}$$

where f represents the frequency band $(\alpha, \beta, \gamma, \theta)$.

2.5.3 *Dimensionality reduction*

Once the Mutual Information and Energy Features were extracted from the EEG signals, an ad-hoc feature reduction scheme was applied prior to classification. To this end, a number of dimensions d is initially selected, and a one-way analysis of variance (ANOVA) is performed in order to detect which features are significantly different across the available classes and discard features with a p-value above a threshold. This threshold was set to 0.01 in this work. When less than d features are retained, this threshold is iteratively incremented by 0.01 until more than d features have been selected. A Principal Component Analysis (PCA) was then applied to further reduce the dimensionality of the remaining data to the established parameter d, which is set by integrating it within the grid search process required to tune the kernel-dependent SVM parameters, as explained in the following section.

2.6 Experimental results

Results obtained with the proposed method were compared to the ones obtained by using the Koelstra et al. [22] and the Chen et al. [29] methods, both implemented as indicated in their original publications. Koelstra et al. extracted the Spectral Power features for the different bands (alpha, slow alpha, beta, gamma and theta) and the asymmetry of 14 different pairs of those Spectral Power features, and then applied the Fisher Discriminant Analysis to reject features with $J_i \leq 0.3$, J_i being the Fisher linear discriminant for the i-th feature, before using a Naive Bayes classifier. Similarly, Chen et al. computed the mutual information between all pairs of EEG channels, without applying any filter in the frequency domain. Then, they applied the same feature selection method as in [22] to discard the less correlated features.

To provide a fair comparison, all competing methods were tested on the DEAP dataset and evaluated under the same experimental setting, namely a leave-one-out (LOO) cross validation scheme applied separately for each individual in the dataset. This process yielded a total of 40 different trained models per user (one per video). At each iteration of the cross validation procedure, one sample of an individual was used for testing the model and the rest of the individual's samples were used for training. Performance metrics were then averaged across all iterations in order to determine the classification performance for each individual. The class labels for each sample were computed as in the Koelstra et al. and Chen et al. works, by thresholding the original valence and arousal values to *Low* and *High* depending on

Table 2.3 F1-scores and standard deviation across individuals obtained with each method

	Koelstra et al. [22]	Chen et al. [29]	Proposed
AROUSAL	0.5333 (0.1009)	0.4167 (0.0860)	**0.5806** (0.1415)
VALENCE	0.6122 (0.1262)	0.6219 (0.1158)	**0.6715** (0.1077)

whether the original value was less than 5 or equal or higher than 5 respectively. Following this approach, the problem of predicting the valence and arousal values was converted to two binary classification problems. The Support Vector Machine (SVM) classifier with a Radial Basis Function kernel was selected in the proposed work and was applied using the libSVM interface for Matlab R2014a [40]. For estimating the SVM parameters, we used a grid-search for obtaining the best C and γ. Furthermore, we also added to the grid the optimal number of dimensions d for the dimensionality reduction as an additional parameter. Matlab R2014a was also used for applying the compared methods by employing the available implementation of Naive Bayes (fitNaiveBayes).

The conversion of the valence and arousal values to binary values resulted in the dataset being considerably unbalanced, for both valence and arousal. This makes the average classification accuracy a misleading measure of performance, given that a classifier that is biased towards the largest class may yield higher values. Hence, we have selected the average F1-score as the measure for evaluating the performance of the competing methods. The final F1-score was calculated by averaging the F1-scores obtained for each of the 32 individuals in the DEAP dataset. Table 2.3 presents the average F1-scores and the standard deviations achieved for the proposed and the examined methods.

As seen in Table 2.3, the proposed approach performed better than the Koelstra et al. and the Chen et al. methods for both valence and arousal. Furthermore, average F1-scores for valence are higher than for arousal for all methods, independently from the features and classification method used. This indicates that the original EEG recordings may be better suited to predict this variable.

For arousal, the proposed method reached an average F1-score of 0.5806, which is considerably higher than the one achieved by the second best method, which is Koelstra et al. It is noticeable that Chen et al. yielded an F1-score below 0.5, which indicates that the method performs worse than by systematically choosing the dominant class for each individual. Koelstra et al. performs only slightly higher than 0.5.

In terms of valence, all methods yield average F1-scores that are significantly higher than 0.5. The proposed method yielded an average F1-score of 0.6715, which is again significantly better than the next highest value, which is 0.6219 and corresponds to Chen et al. In this case, Koelstra et al. scores only slightly below (0.6122). With regard to the variance of the results across different subjects, there is not much difference between the three approaches, and only a slight advantage is observed in favour of the proposed method.

It is also worth mentioning that, apart from the lower F1-score compared to valence, the prediction of arousal using the proposed approach exhibited the highest

Table 2.4 p-values acquired using the one-tailed Wilcoxon's signed rank test

	Chen et al. [29]	Koelstra et al. [22]
AROUSAL	$< 10^{-4}$	0.0280
VALENCE	0.0009	0.0073

variance across different subjects, out of all entries in Table 2.3. This indicates that the performance for predicting the level of arousal differed considerably between individuals, supporting the previous argument that the method performs worse at predicting arousal compared to valence.

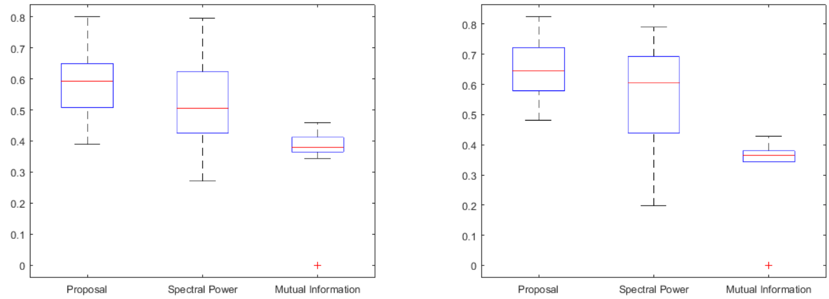
In order to test the statistical significance of the results acquired for the proposed method, we computed the significance statistic p of two-sampled tests, comparing our proposal to each of the competing methods. Due to the fact that none of the pair of distributions followed were normal or homoscedastic, as they did not fit a normal distribution ($p > 0.05$ for Lilliefors test) or they did not have similar standard deviations ($p > 0.05$ for Bartlett's test), we used a non-parametric alternative to Student's t-Test, the one-tailed Wilcoxon's signed-rank test to compute the significance statistic p of the F1-scores. The p-values obtained using the Wilcoxon's test are displayed in Table 2.4, showing that the improvements in performance in relation to the compared methods were statistically significant and outperformed Chen et al.'s and Koelstra et al.'s approach, with a significance level $p < 0.05$ in both cases.

To further demonstrate the difference in performance between the three examined methods, Figure 2.4 shows the boxplots for the distributions of F1-scores obtained across the different individuals, for both valence and arousal. The horizontal line within each box represents the median F1-score m, across all individuals. The bottom and top edges of the box refer to the 25^{th} and 75^{th} percentiles, $q1$ and $q3$, respectively. The whiskers represent the interval $[m - 1.57\frac{(q_3-q_1)}{\sqrt{(n)}}, m + 1.57\frac{(q_3-q_1)}{\sqrt{(n)}}]$. Any measurement outside this interval is considered as an outlier and is represented by using a cross.

Results presented on Figure 2.4 are consistent with the data shown in Table 2.3. Regarding arousal, the median F1-score for the proposed method appears significantly higher than for the other methods, but the bigger box indicates a higher variance. For valence, the three methods present different variances, and the median F1-score is again higher for the proposed approach, indicating a better overall performance.

2.7 Conclusions and discussion

The performance of classification-based approaches for emotion recognition depends on many different factors, such as feature extraction, feature selection, or the classifier used. In emotion detection from EEG signals, it is common to have a small number of training samples with a relatively high dimensionality, making the problem



a) Arousal b) Valence

Figure 2.4 F1-scores comparison

difficult to address. In this chapter, we have proposed a feature combination method for emotion recognition from EEG signals that uses both connectivity and channel-based features. In particular, the distribution of the energy between EEG frequency bands and the connectivity between EEG electrodes have been combined under a unified classification approach. To compensate the increase in the dimensionality of the data, the proposed method employs a feature reduction method that selects features according to their significance level as computed by a one-way ANOVA, and them applies PCA on the remaining set. Finally, classification takes place in a low dimensional space. Overall, the proposed method resulted in a considerable increase in performance at predicting both the valence and arousal dimensions of an emotional state, in terms of the average classification F1-score.

As a shared limitation with most other existing methods, the classification instances are the entire EEG recordings. This limits the practical application of the approaches within a real setting where actions need to be triggered in response to an emotion. In these cases, the detection will suffer a delay that depends on the length of the recording, which may or may not be acceptable in each particular case. Windowed approaches have recently been attempted to solve this problem, e.g. [41]. This would allow a seamless integration with existing applications, including recommendation systems and intelligent tutors, e.g. [42].

Another limitation of the presented work is related to the reconstruction of the emotion by using two-dimensional mappings based on valence and arousal levels. This is also a challenging problem. Firstly, because it requires a simultaneous correct prediction of the two variables in order to locate the emotion in the right quadrant. Secondly, because different emotions in the same quadrant need to be distinguished by using quantifiable valence and arousal levels.

Finally, other works have already demonstrated the benefits of data fusion approaches for emotion recognition and used multimodal models to combine features and/or scores from different sources of information, e.g. [22]. We believe that the proposed feature reduction mechanism also has potential in such context.

Overall, the acquired results demonstrate the potential of AI-based methods for emotion recognition, an application that can significantly benefit the fields of human-computer interaction and of quality-of-experience.

Acknowledgements

We would like to thank Philippe Geril for allowing us to repurpose part of the materials previously presented in the 30th European Simulation and Modelling (ESM) conference [43]. This chapter was partially supported by the Spanish Ministry of Economy and Competitiveness through projects TIN2014-59641-C2-1-P and PGC2018-096463-B-I00.

References

[1] Huang J, Yu C, Wang Y, et al. FOCUS: enhancing children's engagement in reading by using contextual BCI training sessions. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM; 2014. p. 1905–1908.

[2] 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.

[3] Picard RW. Affective Computing. MIT Press; 2000.

[4] Arnau-González P, Arevalillo-Herráez M, Katsigiannis S, et al. On the influence of affect in EEG-based subject identification. IEEE Trans Affect Comput. 2018;In press. DOI: 10.1109/TAFFC.2018.2877986.

[5] Arnau-Gonzalez P, Katsigiannis S, Ramzan N, et al. ES1D: A deep network for EEG-based subject identification. In: 2017 IEEE 17th International Conference on Bioinformatics and Bioengineering (BIBE). IEEE; 2017. p. 81–85.

[6] Dhall A, Goecke R, Joshi J, et al. Emotiw 2016: Video and group-level emotion recognition challenges. In: Proceedings of the 18th ACM International Conference on Multimodal Interaction. ACM; 2016. p. 427–432.

[7] Noroozi F, Marjanovic M, Njegus A, et al. Audio-visual emotion recognition in video clips. IEEE Transactions on Affective Computing. 2017;10(1):60–75.

[8] Le Ngo AC, Oh YH, Phan RCW, et al. Eulerian emotion magnification for subtle expression recognition. In: 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE; 2016. p. 1243–1247.

[9] Tomkins S. Affect imagery consciousness: Volume I: The positive affects. Springer publishing company; 1962.

[10] Plutchick R. The nature of emotions. American Scientist. 2001;89:344–350.

[11] Ortony A, Clore GL, Collins A. The cognitive structure of emotions. Cambridge university press; 1990.

[12] Ekman P. An argument for basic emotions. Cognition & emotion. 1992;6(3-4):169–200.
[13] Ekman P. Emotion in the Human Face. Malor books reprint edition. Malor Books; 2013. Available from: https://books.google.co.uk/books?id=F8r0mwEACAAJ.
[14] Ekman P, Friesen WV. Measuring facial movement. Environmental psychology and nonverbal behavior. 1976;1(1):56–75.
[15] P Ekman W Friesen JH. Facial action coding system: A technique for the measurement of facial movement. Consulting Psychologists Press. 1978;.
[16] Ekman P, Freisen WV, Ancoli S. Facial signs of emotional experience. Journal of personality and social psychology. 1980;39(6):1125.
[17] Friesen WV, Ekman P. EMFACS-7: Emotional Facial Action Coding System; 1983.
[18] Schlosberg H. The description of facial expressions in terms of two dimensions. Journal of Experimental Psychology. 1952;44(4):229237. Available from: http://dx.doi.org/10.1037/h0055778.
[19] Russell J. A Circumplex Model of Affect. Journal of Personality and Social Psychology. 1980 12;39:1161–1178.
[20] Mehrabian A. Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in Temperament. Current Psychology. 1996;14(4):261–292.
[21] Lövheim H. A new three-dimensional model for emotions and monoamine neurotransmitters. Medical Hypotheses. 2012;78(2):341 – 348.
[22] Koelstra S, Muhl C, Soleymani M, et al. DEAP: A Database for Emotion Analysis Using Physiological Signals. IEEE Transactions on Affective Computing. 2012 Jan;3(1):18–31.
[23] Russell JA. Affective Space is Bipolar. Journal of Personality and Social Psychology. 1979;37(3):345–356.
[24] Valenza G, Greco A, Bianchi M, et al. EEG oscillations during caress-like affective haptic elicitation. Psychophysiology. 2018;55(10):e13199.
[25] Liu Y, Sourina O, Nguyen MK. Real-Time EEG-Based Human Emotion Recognition and Visualization. In: International Conference on Cyberworlds (CW); 2010. p. 262–269.
[26] Higuchi T. Approach to an irregular time series on the basis of the fractal theory. Physica D: Nonlinear Phenomena. 1988;31(2):277–283.
[27] Liu Y, Sourina O. EEG-based subject-dependent emotion recognition algorithm using fractal dimension. In: IEEE International Conference on Systems, Man and Cybernetics (SMC); 2014. p. 3166–3171.
[28] Kedem B, Yakowitz S. Time series analysis by higher order crossings. IEEE press New York; 1994.
[29] Chen M, Han J, Guo L, et al. Identifying valence and arousal levels via connectivity between EEG channels. In: IEEE International Conference on Affective Computing and Intelligent Interaction (ACII), 2015; 2015. p. 63–69.

[30] Pearson K. Note on Regression and Inheritance in the Case of Two Parents. Proceedings of the Royal Society of London. 1895;58(347-352):240–242.

[31] Pikovsky A, Rosenblum M, Kurths J. Synchronization: a universal concept in nonlinear sciences. vol. 12. Cambridge university press; 2003.

[32] Gupta R, Falk TH, et al. Relevance vector classifier decision fusion and EEG graph-theoretic features for automatic affective state characterization. Neurocomputing. 2016;174:875–884.

[33] Soleymani M, Pantic M, Pun T. Multimodal emotion recognition in response to videos. IEEE transactions on affective computing. 2012;3(2):211–223.

[34] Katsigiannis S, Ramzan N. DREAMER: A Database for Emotion Recognition Through EEG and ECG Signals From Wireless Low-cost Off-the-Shelf Devices. IEEE Journal of Biomedical and Health Informatics. 2018 Jan;22(1):98–107.

[35] Correa JAM, Abadi MK, Sebe N, et al. Amigos: A dataset for affect, personality and mood research on individuals and groups. IEEE Transactions on Affective Computing. 2018;.

[36] Zheng WL, Lu BL. Investigating Critical Frequency Bands and Channels for EEG-Based Emotion Recognition with Deep Neural Networks. IEEE Transactions on Autonomous Mental Development. 2015 Sept;7(3):162–175.

[37] Morris JD. Observations: SAM: the Self-Assessment Manikin; an efficient cross-cultural measurement of emotional response. Journal of advertising research. 1995;35(6):63–68.

[38] Moddemeijer R. On Estimation of Entropy and Mutual Information of Continuous Distributions. Signal Processing. 1989;16(3):233–246.

[39] Brown G, Pocock A, Zhao MJ, et al. Conditional likelihood maximisation: a unifying framework for information theoretic feature selection. The Journal of Machine Learning Research. 2012;13(1):27–66.

[40] Chang CC, Lin CJ. LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology. 2011;2:27:1–27:27.

[41] Salmeron-Majadas S, Arevalillo-Herráez M, Santos OC, et al. Filtering of Spontaneous and Low Intensity Emotions in Educational Contexts. In: Artificial Intelligence in Education - 17th International Conference, AIED, Madrid, Spain, June 22-26, Proceedings; 2015. p. 429–438.

[42] Arnau D, Arevalillo-Herráez M, González-Calero JA. Emulating Human Supervision in an Intelligent Tutoring System for Arithmetical Problem Solving. IEEE Transactions in Learning Technologies. 2014;7(2):155–164.

[43] Arnáu-Gonzalez P, Ramzan N, Arevalillo-Herráez M. A method to identify affect levels from EEG signals using two dimensional emotional models. In: 30th European Simulation and Modelling Conference-ESM'2016; 2016. .