Class Discovery from Semi-Structured EEG Data for Affective Computing and Personalisation

Aladdin Ayesh
Faculty of Technology,
De Montfort University
Leicester LE1 9BH, UK
Email: aayesh@dmu.ac.uk

Miguel Arevalillo-Herráez
Departament dInformàtica.
Universitat de València
Email: miguel.arevalillo@uv.es

Pablo Arnau-González
University of the West of Scotland
Scotland, United Kingdom
Email: pablo@uws.ac.uk

***Abstract*—Many approaches to recognising emotions from metrical data such as EEG signals rely on identifying a very small number of classes and to train a classifier. The interpretation of these classes varies from a single emotion such as stress [24] to features of emotional model such as valence-arousal [4]. There are two major issues here. First classification approach limits the analysis of the data within the selected classes and is also highly dependent on training data/cycles, all of which limits generalisation. Second issue is that it does not explore the inter-relationships between the data collected missing out on any correlations that could tell us interesting facts beyond emotional recognition. This second issue would be of particular interest to psychologists and medical professions.**

In this paper, we investigate the use of Self-Organizing Maps (SOM) in identifying clusters from EEG signals that could then be translated into classes. We start by training varying sizes of SOM with the EEG data provided in a public dataset (DEAP). The produced graphs showing Neighbour Distance, Sample Hits, Weight Position are analysed holistically to identify patterns in the structure. Following that, we have considered the ground-truth label provided in DEAP, in order to identify correlations between the label and the clustering produced by the SOM. The results show the potential of SOM for class discovery in this particular context. We conclude with a discussion on the implications of this work and the difficulties in evaluating the outcome.

I. Introduction

There has been an increase in the use of Electroencephalography (EEG) like sensors provided by modern fitness devices, such as continuous heart rate monitoring. This naturally has lead to new possibilities of user-device interaction and equally to higher expectations by the users. Personalisation both by user preferences and by adaptation is now expected by many users as a standard function of many mobile and wearable devices. An emerging and ever expanding approach to such personalisation is emotion recognition in which the mood or affective state of the user is approximated and then used to modify or adapt the system functionality or appearance [1]. This is particularly ever more apparent in recommender systems, such as [28], [31], [33] few but to give examples.

Many approaches to emotion recognition from EEG signals rely on identifying a very small number of classes and to train a classifier. The interpretation of these classes varies from a single emotion such as stress to features of emotional model such as valence-arousal. There are two major issues here. First classification approach limits the analysis of the data within the selected classes and also highly dependent on training and limits generalisation. If we are to advance on personalised emotion models [6] we need more dynamic framework to model and identify emotions. This can then be naturally extended to include implicitly or explicitly other intertwining factors, such as personality, in representing and updating user affective states.

Second issue is that it does not explore the inter-relationships between the data collected missing out on any correlations that could tell us interesting facts beyond emotional recognition. This second issue would be of particular interest to psychologists and medical professions.

In this paper, we investigate the use of Self-Organizing Maps (SOM) in identifying clusters from EEG signals that could then be translated into classes. We start by training varying sizes of SOM with EEG data using a publicly available dataset DEAP [20]. The produced graphs showing Neighbor Distance, Sample Hits, Weight Position are analysed holistically to identify patterns in the structure. Following that, node density and sample clustering are compared to the sample classification that was provided with DEAP to identify correlation between the sample classification and the cluster. The results show the potential for class discovery. We conclude with a discussion on the implications of this work and the difficulties in evaluating the outcome.

The paper is organized as follows. First we start by giving background on the data used and how it was analysed and prepared. The experiments with SOM and the analysis of the results are presented and the main conceptual contribution of this paper discussed in detail. We then conclude the paper with a critical discussion covering outstanding research questions.

II. EEG Data

A. Motivation

In a number of neuropsychological studies, EEG data showed to exhibit correlates of emotion [19] e.g. event-related potentials (ERPs) [27] that can be analysed through the spectral power in several frequency bands [9], [10]. The results of these studies motivated the rapid development of emotion recognition techniques based on EEG data. The availability of public data sets has also played a role in advancing these techniques. One of these data sets is DEAP [20]. It stands out as one of the most frequently used to evaluate the performance of emotion recognition methods e.g. [5], [17], [25], [26], [37], [38]. Due to its wide use, it has become a common benchmark in this particular context. DEAP consists of EEG and peripheral physiological signals collected from 32 participants. These signals were recorded as each participant watched 40 one-minute long excerpts of music videos, which were selected in order to elicit emotions in each of the 4 quadrants of the Russells Circumplex Model [7], [13], [16], [30], [34].

The analysis of the collected EEG signals produced a total of 216 features. These features were then used in a baseline experiment. In the experiment a binary classification setting using a Naive Bayes Classifier was applied using the ground truth labels extracted from the self-reported Self Assessment Manikin (SAM) ratings [14]. We expound the DEAP dataset and the analysis applied to it in section II-B.

Similar approaches have been reported in [3], [5], [17], [25], [26], [37]–[39], to give examples but a few. These approaches utilized typical classification frameworks in which the recorded EEG signals are pre-processed by using spatio-temporal filtering and noise reduction methods, to abate artefacts and enhance the Signal-to-Noise Power ratio (SNR). Relevant features were then extracted to provide training samples to a classifier. The samples were labelled according to a specific approach to describe emotions [35]. This may assume a set of distinct emotional categories, such is the case in Ekman's Basic Emotions Model [15]. Another approach is to describe the emotion as a point in a continuous multidimensional space where each dimension represents one aspect of the emotion such is the case in Millenson's Model [12]; other examples of this approach can be found in [7], [8], [16], [34]. The resulting model is used to predict the most likely emotional state by using the same description approach as for labelling the training samples.

B. Experimental Data Used

As aforementioned, the DEAP database [20] is a publicly available dataset that is utilized in several studies giving us a benchmark for comparisons, and thus we utilised it in testing our approach. DEAP dataset contains physiological recordings from 32 healthy participants divided equally 50% male and 50% female. The participants are aged between 19 and 37 with a mean of 26.9 years. The sensory data was collected whilst participants were watching 40 music videos. The video clips were 63 seconds long and carefully selected in order to elicit emotions in each of the quadrants presented in the Russells Circumplex Model [34]. The physiological recordings included Galvanic Skin Response (GSR), Respiration Amplitude, Skin Temperature, ElectroCardioGram (ECG), Blood Volume by Plethysmograph, ElectroMyoGrams (EMG) of Zygomaticus and Trapezius muscles, and ElectroOculoGram (EOG). EEG signals were recorded using 32 active AgCl electrodes at a sampling rate of 512 Hz.

Participants were asked to report, after watching each video, their emotion using Self Assessment Manikin (SAM) [14], in the range from 1 to 9. The self-reported ratings were stored along with the resulting signals. This data was then proposed as a ground truth after being converted into categorical variables (classes) with two possible values, namely low and high. On the nine points rating scales given in SAM, the threshold was placed in the middle. In addition, a series of power spectral features were extracted from the EEG signals. These were the logarithms of spectral power for each channel in each relevant frequency band, and the spectral power asymmetry between the 14 symmetrical pairs of electrodes in the same frequency bands. The five frequency bands were defined as θ (4-8 Hz), slow α (8-10Hz), α (8-12 Hz), β (12-30 Hz) and γ (30+ Hz). This data processing leads to extracting a total of 216 features ($32 \times 5 + 14 \times 4$). These features were then fed into a binary classification setting using a Naive Bayes Classifier to evaluate their performance in a baseline experiment. The ground truth labels extracted from the self-reported SAM ratings were used in the evaluation.

In contrast, our analysis has concentrated on studying implicit relations between valence and the EEG features extracted from the signals [2]. In similar fashion, we have extracted the spectral power for each channel in each relevant frequency band, from the defined four bands: α (8-13 Hz), β (14-30 Hz) γ (30-47 Hz) and θ (4-7 Hz); and the spectral power asymmetry between the 14 symmetrical pairs of electrodes in the same frequency bands. This leads to a total of 184 features ($32 \times 4 + 14 \times 4$). These are a set of commonly used features in the literature.

III. Class Discovery using SOM

A. Self-Organizing Maps

Regardless of the classification techniques that one may use, they are all reliant on labelling using feature-based representation to classify samples in a pre-defined set of classes. These pre-defined set of classes dependent on the emotional representation used on one hand whilst on the other hand reliant on measuring the emotional states including self-reporting, which is known to suffer from participant's interpretation at best and to be unreliable at worst. This makes any attempt to have a generalised model for emotion detection and analysis applicable to all users even more challenging if not impossible. Thus having a personalised emotional model that can adapt to each user physiological expression of emotions [6] being facial or EEG signals would be an ideal solution.

The difficulties at labelling data in classifying EEG samples into emotional classes can be partially overcome if these emotional classes can, fully or partially, be discovered as part of the learning process. Unsupervised classification methods [11] attempt to infer the underlying structure of the data by analysing the existing relationships between the available data inputs, and do not require a previous labelling. One popular such technique is Self Organising Maps (SOM) [18], [22]. They were first proposed in the 1980s [21], [32]. Since then, they have been widely used for data analysis and visualization purposes on many and diverse application areas, ranging from engineering [23] to medicine [36]. SOM or SOFM, with *F* denotation features, used as a means to organize data according to their internal structure. In essence,they transform arbitrarily complex non-linear statistical relationships between high-dimensional data samples into simpler geometric relationships on a low-dimensional display.

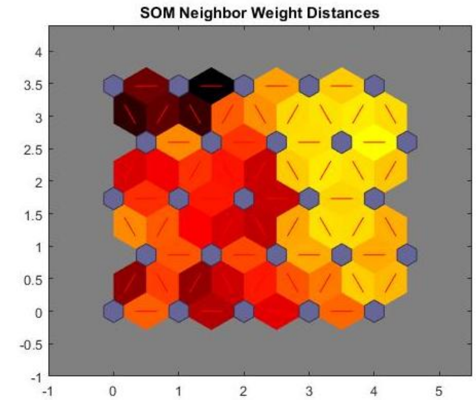
B. Experimental Setup

The EEG dataset of 1280 samples x 184 features was used to train Self-Organizing Maps with 5, 10 and 20 nodes. Once training is done the produced results such as neighbour distance, weight position, and sample hits for each node. Whilst each network produced different results and outputs were slightly different after training round, there was a clear visual pattern emerging in the output plots. If we look at the visualisation of the outputs in a more holistic qualitative approach than a precise quantitative or numerical views, we can see clear similarities between all the networks and across the different training rounds.

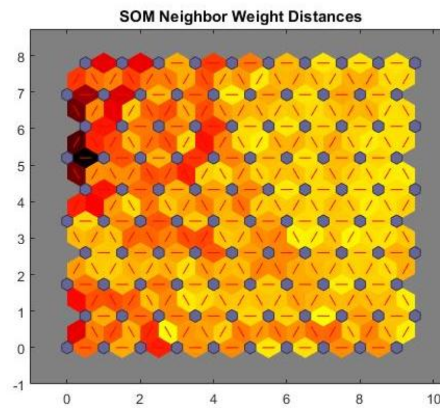
In the case of Neighbour Distance, 1, we can see concentration in certain regions. For example, we can observe such concentration in the upper left hand side of figure 1(a). Similar concentration can be observed in the same regions in the cases of 10, figure 1(b), and 20 nodes, figure 1(c), however, these concentrations are stretched in response to the larger structure. One can only draw a conclusion that in these clusters data is tightly connected with less use of further out clusters. This could imply the existence of interrelationships that enable identifying certain clusters as classes, which in turn supports the central hypothesis of this paper.

Similarly in the case of Weight Position, we can see similar concentration pattern looking at the sub-figures of figure 2. We can also observe the *exception* pattern especially in figures 2(b) and 2(c).

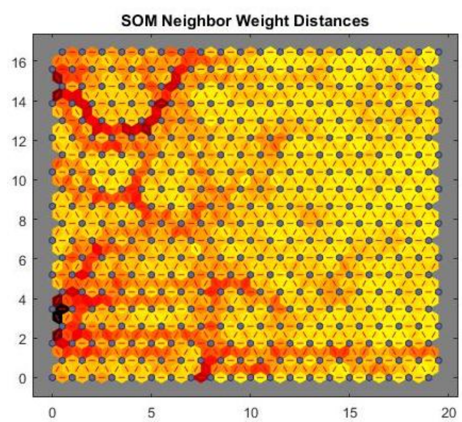
Exploring sample hits, figure 3, may give us the key to our class discovery approach. We can see concentrations of samples in certain nodes. The question would these nodes qualify to become classes? If we look at figure 3(a) for an example, we can see concentration of samples in number of classes to the right hand side, which may contradict our observations on figure 1(a). However, we can notice there are distinctively two clusters to the left hand side with 39 and 40 sample hits. These map perfectly to the concentration observed in 1(a).



(a)

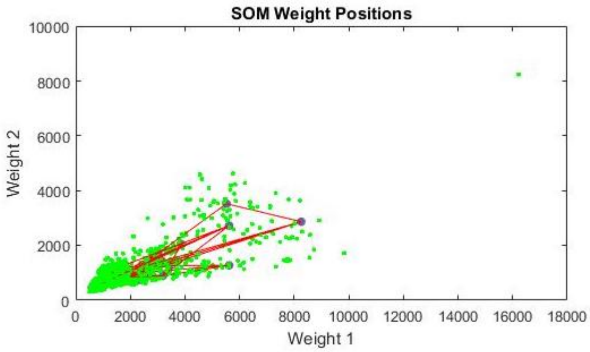


(b)

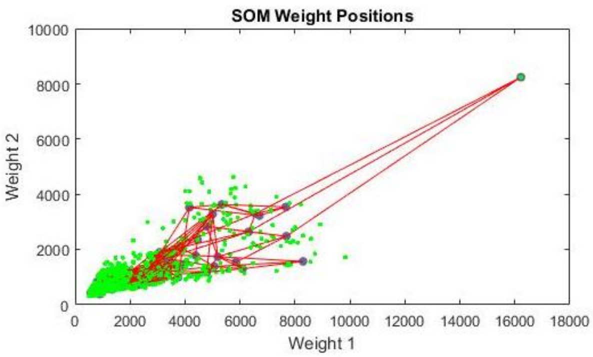


(c)

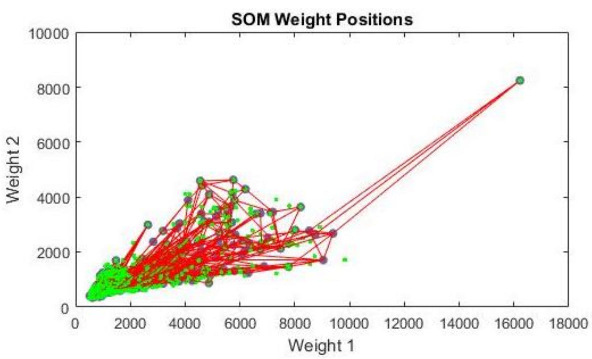
Fig. 1. Neighbour Distance with 5, 10, 20 Nodes SOM's



(a)

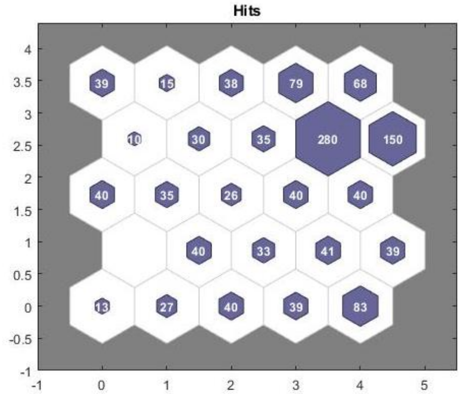


(b)

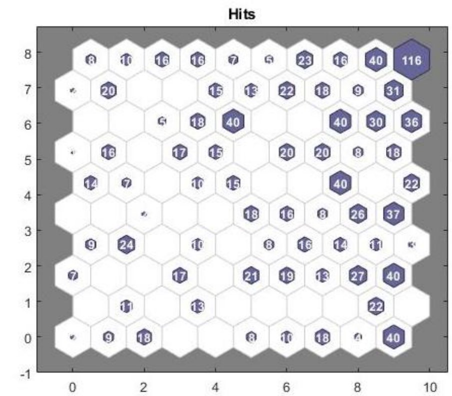


(c)

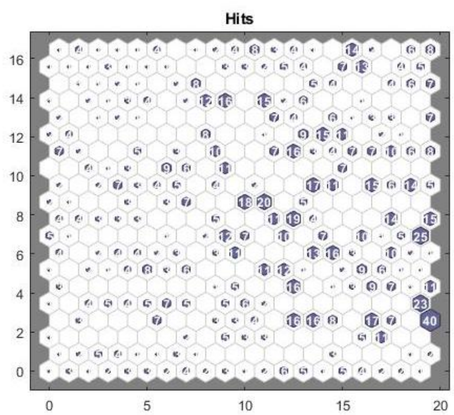
Fig. 2. Weight Position with 5, 10, 20 Nodes SOM's



(a)



(b)



(c)

Fig. 3. Sample Hits with 5, 10, 20 Nodes SOM's

C. SOM Analysis

To answer the questions imposed in the previous section, we used the first classification experiment performed by [20] as a baseline. In that work, an Arousal-Valence model of emotion was used. We focused on the classification relating to valence. The classification was made against two classes identified $\{0,1\}$ identifying low and high valence respectively. In that work, the samples were classified in one of the two classes. Taking each node of the network we examined the classification of each sample.

The results of this examination showed that each significant node had a dominant class within its samples with a reference to the two types of class labels, namely "0" class and "1" class that are provided in the DEAP dataset. This means that there is a structural similarity between groups of samples that belong to the same cluster. Samples classified under the same SOM node tend to belong to the same class or label as it was reported in the dataset. This makes it potentially possible to use the class information on training data to develop a customised classifier based on the most likely label in each SOM node. Since the knowledge is tacitly built within the network structure, there is less need to identify explicitly similarity features such is the case in Associative Classifiers for example. In fact, SOM becomes a knowledge representation tool to inform any given classifier we may wish to use.

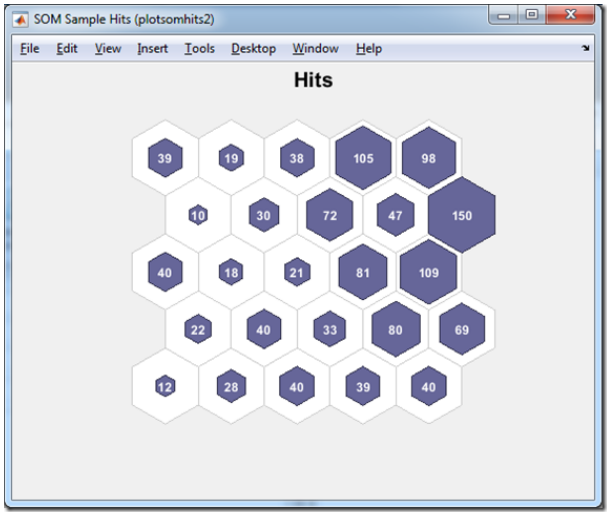


Fig. 4. Sample Hits in 5 nodes SOM from a different training round to fig. 3

IV. Critical Review

This work started with the aim of finding an effective data mining approach to explore EEG data for the purpose of emotion detection and sentiment analysis. The initial approach of using SVM classifier produced some good results [4], however, had the limitations of the pre-specified strict classes that does not allow the free exploration of the data. Thus it limited the emotion models that can be used and the interpretation of the data. In seeking an alternative, we identified the possibility of class discovery using Self-Organizing Maps. The main distinction here is that these discovered classes could be shown to map to clear distinctive classes with commonly shared features. This can then be used to map these classes to various theories of emotion for interpretation and subsequent use in affective computing with traditional classifiers for example.

There are still, however, several questions outstanding. For example, how do we identify true classes? What is the threshold of density required? Would this threshold changes greatly with different data sets? How can we evaluate the discovered classes and identify their features? Finally, can we merge classes to produce higher density or break them to have greater granularity? There may be a scope to inject the SOM with swarm intelligence algorithms enabling a more dynamic network structure in discovering classes.

Another question emerges between the manner by which class is treated here and rough sets [29]. An exploration of set theory alternatives may provide us with some answers to these questions.

V. Conclusion

In this paper, we investigated the use of Self-Organizing Maps (SOM) in identifying clusters from EEG signals that could then be translated into classes. We start by training varying sizes of SOM with EEG data that was provided in the publicly available dataset DEAP [20]. Then we analysed the visualised outputs holistically to identify patterns in the structure. Taking the first classification provided by [20] as ground truth, node density was produced for each node by identify correlation between the sample classification and the cluster. The results show potential for class discovery. We conclude with a discussion on the implications of this work and outstanding research questions.

Acknowledgment

The authors would like to thank De Montfort University for the support given to enable this research and acknowledge that this work has been partly supported by the Spanish Ministry of Economy and Competitiveness through project TIN2014-59641-C2-1-P

References

[1] Miguel Arevalillo-Herráez, David Arnau, Luis Marco-Giménez, José Antonio González-Calero, Salvador Moreno-Picot, Paloma Moreno-Clari, Aladdin Ayesh, Olga C. Santos, Jesus Boticario, Mar Saneiro, Sergio Salmeron-Majadas, Raúl Cabestrero, and Pilar Quirós. Providing personalized guidance in arithmetic problem solving. In *UMAP Workshops*, pages 42–48, 2014.

[2] Miguel Arevalillo-Herráez, Aladdin Ayesh, O. C. Santos, and P. Arnau-González. Combining supervised and unsupervised learning to discover emotional classes. In *Proceedings of the 25th ACM Conference on User Modeling, Adaptation and Personalisation (UMAP 2017)*, 9-12th July 2017.

[3] Pablo Arnau-González, Miguel Arevalillo-Herráez, and Naeem Ramzan. Fusing highly dimensional energy and connectivity features to identify affective states from EEG signals. *Neurocomputing*, 244:81–89, 2017.

[4] Pablo Arnau-Gonzalez, Naeem Ramzan, and Miguel Arevalillo-Herraez. A method to identify affect levels from eeg signals using two dimensional emotional models. In *The 30th European Simulation and Modelling Conference - ESM'2016,*, Spain, October 2016. Univ. of Las Palmas, Eurosis.

[5] John Atkinson and Daniel Campos. Improving bci-based emotion recognition by combining eeg feature selection and kernel classifiers. *Expert Systems with Applications*, 47:35–41, 2016.

[6] Aladdin Ayesh, Miguel Arevalillo-Herráez, and Francesc J. Ferri. Towards psychologically based personalised modelling of emotions using associative classifiers. *International Journal of Cognitive Informatics and Natural Intelligence*, 10(2):52–64, April 2016.

[7] Aladdin Ayesh and William Blewitt. Models for computational emotions from psychological theories using type i fuzzy logic. *Cognitive Computation*, 7(3):285–308, 2015.

[8] Aladdin Ayesh and William Blewitt. Models for computational emotions from psychological theories using type-ii fuzzy logic. *Cognitive Computation*, 7(3):309–332, 2015.

[9] Michela Balconi and Claudio Lucchiari. Eeg correlates (event-related desynchronization) of emotional face elaboration: a temporal analysis. *Neuroscience letters*, 392(1):118–123, 2006.

[10] Michela Balconi and Guido Mazza. Brain oscillations and bis/bas (behavioral inhibition/activation system) effects on processing masked emotional cues.: Ers/erd and coherence measures of alpha band. *International Journal of Psychophysiology*, 74(2):158–165, 2009.

[11] Sanghamitra Bandyopadhyay and Sriparna Saha. *Unsupervised classification: similarity measures, classical and metaheuristic approaches, and applications*. Springer Science & Business Media, 2012.

[12] William Blewitt, Aladdin Ayesh, Robert I. John, and Simon Coupland. A millenson-based approach to emotion modelling. In *Human System Interactions, 2008 Conference on*, pages 491–496, 2008.

[13] William F. Blewitt and Aladdin Ayesh. Modeling the emotional state of an agent through fuzzy logic with reference to the geneva emotion wheel. In Cyrille Bertelle and Aladdin Ayesh, editors, *European Simulation and Modelling (ESM'2008) Conference*, pages 279–283, Le Havre, France, 2008. EUROSIS.

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

[15] Paul Ekman. An argument for basic emotions. *Cognition & emotion*, 6(3-4):169–200, 1992.

[16] Andrew J Gerber, Jonathan Posner, Daniel Gorman, Tiziano Colibazzi, Shan Yu, Zhishun Wang, Alayar Kangarlu, Hongtu Zhu, James Russell, and Bradley S Peterson. An affective circumplex model of neural systems subserving valence, arousal, and cognitive overlay during the appraisal of emotional faces. *Neuropsychologia*, 46(8):2129–2139, 2008.

[17] Narendra Jadhav, Ramchandra Manthalkar, and Yashwant Joshi. Electroencephalography-based emotion recognition using gray-level co-occurrence matrix features. In *Proceedings of International Conference on Computer Vision and Image Processing*, pages 335–343. Springer, 2017.

[18] Jari Kangas and Teuvo Kohonen. Developments and applications of the self-organizing map and related algorithms. *Mathematics and Computers in Simulation*, 41(1-2):3–12, 1996.

[19] Min-Ki Kim, Miyoung Kim, Eunmi Oh, and Sung-Phil Kim. A review on the computational methods for emotional state estimation from the human eeg. *Computational and mathematical methods in medicine*, 2013, 2013.

[20] S. Koelstra, C. Muhl, M. Soleymani, J. S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras. Deap: A database for emotion analysis using physiological signals. *IEEE Transactions on Affective Computing*, 3(1):18–31, January 2012.

[21] Teuvo Kohonen. Self-organized formation of topologically correct feature maps. *Biological cybernetics*, 43(1):59–69, 1982.

[22] Teuvo Kohonen. The self-organizing map. *Neurocomputing*, 21(1):1–6, 1998.

[23] Teuvo Kohonen, Erkki Oja, Olli Simula, Ari Visa, and Jari Kangas. Engineering applications of the self-organizing map. *Proceedings of the IEEE*, 84(10):1358–1384, 1996.

[24] Yee Mei Lim, Aladdin Ayesh, and Martin Stacey. Using mouse and keyboard dynamics to detect cognitive stress during mental arithmetic. In Kohei Arai, Supriya Kapoor, and Rahul Bhatia, editors, *Intelligent Systems in Science and Information 2014*, volume 591 of *Studies in Computational Intelligence*, pages 335–350. Springer International Publishing, 2015.

[25] Y. Liu and O. Sourina. EEG-based subject-dependent emotion recognition algorithm using fractal dimension. In *Proc. and Cybernetics (SMC) 2014 IEEE Int. Conf. Systems, Man*, pages 3166–3171, October 2014.

[26] Ahmet Mert and Aydin Akan. Emotion recognition from eeg signals by using multivariate empirical mode decomposition. *Pattern Analysis and Applications*, pages 1–9, 2016.

[27] Jonas K Olofsson, Steven Nordin, Henrique Sequeira, and John Polich. Affective picture processing: an integrative review of erp findings. *Biological psychology*, 77(3):247–265, 2008.

[28] A. H. Parizi and M. Kazemifard. Emotional news recommender system. In *Proc. Sixth Int. Conf. of Cognitive Science (ICCS)*, pages 37–41, April 2015.

[29] Zdzisław Pawlak and Andrzej Skowron. Rudiments of rough sets. *Information Sciences*, 177(1):3–27, jan 2007.

[30] Jonathan Posner, James A Russell, and Bradley S Peterson. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology*, 17(03):715–734, 2005.

[31] J. Qin and W. Zhang. Connotative space for music affective recommendation. In *Proc. SAI Computing Conf. (SAI)*, pages 502–506, July 2016.

[32] Helge Ritter and Teuvo Kohonen. Self-organizing semantic maps. *Biological cybernetics*, 61(4):241–254, 1989.

[33] R. L. Rosa, D. Z. Rodriguez, and G. Bressan. Music recommendation system based on user's sentiments extracted from social networks. *IEEE Transactions on Consumer Electronics*, 61(3):359–367, August 2015.

[34] James A. Russell. Affective space is bipolar. *Journal of Personality and Social Psychology*, 1979.

[35] Marko Tkalcic, Andrej Kosir, and Jurij Tasic. Affective recommender systems: the role of emotions in recommender systems. In *Proc. The RecSys 2011 Workshop on Human Decision Making in Recommender Systems*, pages 9–13. Citeseer, 2011.

[36] William W van Osdol, Timothy G Myers, Kenneth D Paull, Kurt W Kohn, and John N Weinstein. Use of the kohonen self-organizing map to study the mechanisms of action of chemotherapeutic agents. *Journal of the National Cancer Institute*, 86(24):1853–1859, 1994.

[37] Yuliyan Velchev, Svetla Radeva, Strahil Sokolov, and Dimitar Radev. Automated estimation of human emotion from eeg using statistical features and svm. In *Digital Media Industry & Academic Forum (DMIAF)*, pages 40–42. IEEE, 2016.

[38] H. Xu and K. N. Plataniotis. EEG-based affect states classification using deep belief networks. In *Proc. Digital Media Industry Academic Forum (DMIAF)*, pages 148–153, July 2016.

[39] Haiyan Xu and Konstantinos N Plataniotis. Affective states classification using eeg and semi-supervised deep learning approaches. In *Multimedia Signal Processing (MMSP), 2016 IEEE 18th International Workshop on*, pages 1–6. IEEE, 2016.