

Emotion Analysis Through EEG and Peripheral Physiological Signals Using KNN Classifier



Shourya Shukla and Rahul Kumar Chaurasiya

Abstract Emotions are the characteristics of human beings which are triggered by the mood, temperament or motivation of an individual. Emotions are nothing but the response to the stimuli that are experienced by the brain. Any changes in one's emotional state results in changes in electrical signals generated by the brain. The emotions can be explicit or implicit, i.e. either emotion may be expressed or remain unexpressed by the individual. As these emotions are experienced by the individual as the result of the brain stimulus, we can observe Electroencephalogram (EEG) signal to classify the emotions. Some of the physiological signals may also be taken into account as any change in emotional state result in some physiological changes. For the analysis, we have used the standard DEAP dataset for emotion analysis. In the dataset, the 32 test subjects are shown with 40 different 1-minute music videos and the EEG and other physiological signals are recorded. On the basis of the Self-Assessment Manikins (SAM), we classify the emotion state in the valence arousal plane. The K-Nearest Neighbour classifier is used to classify the multi-class emotions as higher/lower levels of the valence arousal plane. The comparison of KNN with other classifiers depicts that KNN has produced best average accuracy of 87.1%.

1 Introduction

Electroencephalogram is the electrical signals produced by the brain. EEG signal are used to analyse the emotional state of the subject under study. Emotions are triggered in the brain by stimuli such as video signals or audio signals etc., [1]. Depending upon the mood, temperament and motivation of the subject, different emotions are experienced. The brain works in different frequency bands, namely delta, theta, alpha, beta and gamma [2, 3]. Beta (12–30 Hz) band is emitted when someone is conscious

S. Shukla (✉) · R. K. Chaurasiya (✉)
National Institute of Technology, Raipur, India
e-mail: shourya.shukla2000@gmail.com

R. K. Chaurasiya
e-mail: rkchaurasiya@nitrr.ac.in

© Springer Nature Switzerland AG 2019
D. Pandian et al. (eds.), *Proceedings of the International Conference on ISMAC in Computational Vision and Bio-Engineering 2018 (ISMAC-CVB)*, Lecture Notes in Computational Vision and Biomechanics 30,
https://doi.org/10.1007/978-3-030-00665-5_11

and alert, and the subject is thinking or concentrating. Alpha band (8–12 Hz) is most active when the subject is in the state of physical and mental relaxation but in aware and conscious. Theta (4–7 Hz) band is associated with daydreaming or sleepy state. It is also called the creative state. The delta band (0.1–4 Hz) is the lowest frequency state related to deep sleep. At the upper highest frequencies, that is the gamma band (>30 Hz) the subject is in deep meditation, mainly found significant in Buddhist monks.

The human–human interaction is easier as humans are more aware of the sentiments of the human they are interacting, but in the case of the Brain–Computer interaction (BCI), it is more complicated as the computer is unaware of the emotions of the human subject. Here, we present a method which uses a signal-based approach by extracting information from the EEG and peripheral physiological signals. These peripheral physiological signals are useful in predicting the emotional state of the subject as they complement EEG signals in the emotional analysis. These peripheral signals are significant in different frequency bands and hence, the Discrete Wavelet Transform (DWT) of the signals is taken into consideration. The DWT of a signal splits the signal into the higher frequency detail (D) and the lower frequency approximation (A) coefficients [2].

The features are extracted from the wave decomposed signals, and the KNN classifier is used to classify the emotions as low valence low arousal, low valence high arousal, high valence low arousal and high valence high arousal. The valence indicates the pleasing level of the brain. A high valence level indicates happy or elated emotion, whereas a low valence indicates sad or stressed behaviour. The arousal indicates activeness of the brain. A higher arousal indicates alert or excited response, whereas the lower level of arousal indicates uninterested or bored response of an individual.

2 Related Works

The EEG signals were first recorded by the English scientist Richard Caton [4] in the year 1875. The study of the EEG signals was first explored by Hans Berger in 1920 [4]. The first study of emotion and physiological signals goes back to the year 1941 by Hadley, J. M. in which the author described the relationship between the EEG and peripheral physiological signals while performing multiplications of varying difficulties [5]. Plutchik, R first associated with high sound intensities on the performance feelings and physiology of the subject [6]. Moon et al. proposed the method for video preferences based on the extracted using EEG-based responses quadratic–discriminant–analysis-based model using BP features [7]. There are many proposed ways to classify the emotion of the subject using EEG. Ekman and Friesen, 1987 were the first to propose the six emotions existing that can be classified using facial signals [8]. The valence arousal plane is used to separate different emotional state as proposed by Sander Koelstra et al. (2011). A Database for Emotion Analysis using Physiological Signals (DEAP dataset) uses the EEG and other physiological

signals to classify the emotions in the valence arousal plane using F1 score and naïve Bayes classifier is performed [9].

The peripheral physiological signals play an important role in determining the emotion as they complement the EEGs and provide the information about the subject's reaction to a stimuli. Torres-Valencia et al. suggested a multimodal emotion recognition using the DEAP dataset by Hidden Markov Model (HMM) using the Galvanic Skin Resistance (GSR) and Heart Rate (HR) [10]. Ramasamy et al. defined the heart–brain interaction through the EEG and ECG involving the emotion through biofeedback system [11]. Li et al. related the EEG with peripheral physiological signals such as Electrooculogram (EOG), Electrocardiogram (ECG), Electromyogram (EMG), skin temperature variation and electrodermal activity in a brain–computer interface system to measure the attention level in ubiquitous environment [12].

3 Materials and Methods

3.1 Dataset for Experimental Analysis

In the last few years, the BCI is one of the most studied topics in the field of machine learning. The DEAP dataset is being used for research purpose by many scholars. We are also using the DEAP dataset which is available for the research work. An End-User License Agreement (EULA) is acquired to access the data. The dataset consists of the records of 32 patients (subjects). Out of 32 subjects, 22 subjects' facial recording is provided. Each subject's file consists of data and label file. The data file has the 32-channel EEG recordings along with 8 peripheral physiological signals. The physiological signals which are used to understand the emotional state of the subject are Galvanic Skin Resistance (GSR), Respiration Amplitude (RA), Skin Temperature (ST), Electrocardiogram (ECG), Blood Volume Pressure (BVP) and Electromyogram (EMG) of the zygomaticus and trapezoidal muscles. The eye-ball movement is captured by the Electrooculogram (EOG) signal. The signals are sampled at the rate of 512 Hz and the peripheral physiological are further down-sampled to 256 Hz. To record the EEG and other signals, visual stimuli are used. 40 videos were selected using the web-based survey and the stimulating 1 minute of each video is shown to the subject. The data is hence of the dimension $40 \times 40 \times 8064$, for 40 min, 40 signals for each minute and 8064 samples of each signal. 22 subjects' facial video was also recorded using SONY DCR-HC27E camcorder.

The EEG signals were obtained using a 10–20 electrode system. The EEG has 32-channel electrode having a 10–20 system with odd number of electrode placed on the left hemisphere and electrode with the number placed on the right hemisphere of the brain. The dataset used the Bio Semi Active Two System for recording the EEG signals.

The labelling of the data is done by the rating given by the individual. The process is named Self-Assessment Manikins (SAM). In the labelling process, after completion

of every 1-min video, the subjects are asked to rate the video between 1 and 9, where 1 being the lowest and 9 being the highest rating. The subject is asked to move the cursor horizontally and click on the rating bar to give the scores on the arousal and valence parameters. To define a binary class system, midpoint threshold were taken on the rating scale of 1–9 for arousal and valence. Other SAMs like dominance, liking and familiarities were also assessed. The dominance rating represents the feeling of being empowered, whereas for liking scale, thumbs up or thumb down symbols were used. Familiarity rating suggests that how well the subject knows or remembers the video stimulus. The labelling of data is done as high valence high arousal, high valence low arousal, low valence high arousal and low valence low arousal. Thus, multi-class classification is performed on basis of these four classes.

3.1.1 Preprocessing, Feature Extraction and Classifier Methodology

The DEAP dataset contains records of 32 test subjects, and the large dataset makes it difficult to work on the integrated data. We segmented the data for each minute recording, i.e. time signals were obtained for each minute video. Therefore, each subject's psychophysiological signals are divided into 40 segments. Then for each minute, we find the discrete wavelet transform coefficient. These coefficients are time signals with ascending order of frequencies. The length of each frequency band is given in the length coefficient matrix. The DWT gives the details and approximations as response to high-pass and low-pass filter, respectively. The five-level DWT is obtained for the EEG signals of the frequency range 0–30 Hz. The five-level DWT is also used for the peripheral physiological signals as the EMG signals works on the higher frequency range of 4–40 Hz.

3.1.2 Feature Extraction Using DWT

EEG being a non-stationary signal, Fourier transform is not a suitable transform for it. Hence, to find the feature matrix, five-level DWT is used. The DWT decomposes the wavelets in the time–frequency domain. The DWT decomposes the signal according to the increasing frequency bands. The frequency decomposition is obtained by selecting odd and even samples of the signal and then, these samples are passed through low-pass and high-pass filter, respectively, further, the filtered signals are downsampled by a factor of 2. The resulting high-pass downsampled signals is the detail signals and low-pass downsampled signals are the approximate signals. The multilevel DWT is performed by applying recursively DWT to the $(n - 1)$ th approximate samples [13]. For each minute video, the five-level DWT is applied on the psychophysiological signals, hence each signal is decomposed into five detail levels D1–D5 and an approximate A5, features are then extracted and classified by KNN classifier. As different signals works in different frequency ranges (EEG in the 0–30 Hz, whereas trapezoidal muscles in 4–40 Hz), all the DWT coefficients are taken into account.

Feature extraction is performed on the decomposed wavelets to study the EEG and other physiological signals in different frequency ranges. The EEG features extracted for classification of the emotional state of the subject are mentioned below:

- (1) First feature consists of the logarithmic values of power spectral densities of EEG samples of mean values of each band.
- (2) Second feature consists of the difference in power spectral densities of corresponding right and left hemispheric electrode of the EEG signal.

The GSR measures the resistance of the skin. It is related to the amount of perspiration the subject has while watch the video stimulus. The degree of perspiration provides information about the nervousness and anxiety of the subject. The resistance decreases with increase in perspiration. Lang et al. suggested that the arousal is correlated with the mean value of GSR [14]. The following features are extracted from the GSR signal:

- (1) The average values of each band D1–D5 and A5 are extracted.
- (2) The mean values of derivative of GSR is extracted for each band.
- (3) The mean values of derivative for negative values only. Hence, average decrease rate during perspiration is evaluated.
- (4) Ratio of number of negative sample to the total number of samples in each band.
- (5) Number of local minima of each band.
- (6) Zero-crossing rate of the GSR.

The BVP is measured using plythesmograph. It is the measure of the pressure by which heart exerts pressure into the arteries. The plythesmogarph is attached to thumb of the subject. The BVP can be used to find the heart rate and heart rate variability. The heart rate and its variability are correlated to the emotional state as the faster heart rate indicates subject under stress. The higher blood pressure indicates the sense of fear or surprise. The following are the feature extracted from the blood volume pressure of the subject undergoing visual stimuli:

- (1) Average value of BVP.
- (2) Standard deviation of BVP.
- (3) Average value and standard deviation of the heart rate (heart rate is identified with the help of local maxima, i.e. the heart beat and the heart beats per minute is the heart rate).
- (4) Heart rate variability (Iwona Cygankiewicz et al., suggested that HRV reflects beat to beat changes in R-R interval. Heart rate changes may occur as a response to the physical and mental stress [15]. And hence, related to the emotional state of the subject under study).
- (5) Inter-beat interval.
- (6) Energy of the blood volume level in each frequency band.
- (7) Energy ratio of blood volume level in consecutive frequency bands.
- (8) Power spectral densities of blood volume level.

The next physiological signal is the RA. The RA is recorded as the speed of respiration depends on the emotional state, a slow RA represents relaxation and fast

or irregular respiration implies a feeling of fear or anger. The RA has high correlation with the arousal. The feature extracted from RA is as follows:

- (1) Inter-band difference in energy of respiration signal.
- (2) Average respiration signal.
- (3) Mean of derivatives of each frequency band.
- (4) Standard deviation of RA. (This feature shows variation in respiration amplitude which in turn shows the change in mood of the patient).
- (5) Range or greatest breath time taken by the subject.
- (6) Spectral centroid of the respiration frequency bands.
- (7) Breathing rate.
- (8) Power spectral density of respiration amplitude.

ST is also used to extract the emotion and hence the following features are extracted:

- (1) Average value of each band of temperature.
- (2) Average of its derivatives are extracted to find the mean rate of change of ST
- (3) Power spectral densities

Muscle signals were obtained in the DEAP dataset to extract information about the facial expression and shoulder movement. The facial muscle plays an important role in expressing one's emotions. Hence, EMG of zygomaticus muscle was recorded. The shoulder muscle movement implies laughter (happy) emotion or anger as implicit tags. The features extracted are:

- (1) Energy of the muscle signals.
- (2) Mean of the EMG.
- (3) Variance of the zygomaticus and trapezoidal muscles.

The eyeball movement and tracking is involved in predicting one's emotions. The blinking rate is decreased to a large extent when the person experiences high arousal. The EOG signal can be related to anxiety of the person also. The features presented by the DEAP dataset are:

- (1) Energy of the EOG signals in each band.
- (2) Mean of the EOG signal in each frequency band.
- (3) Variance in eyeball movement.
- (4) Blinking rate of the eye. (The blinking rate is determined by the detectable peaks of the EOG).

3.2 Classification Using KNN Classifier

The KNN is a supervised learning algorithm, hence, it works on the given dataset directly. For any new instance, predictions are made on the search of the K most proximate instances. Here in the case of KNN, these instances are known as "Neighbours". To measure the parameter of the neighbour, we are using the Euclidean

distance. The K -nearest neighbour's Euclidean distance is calculated with respect to the instance data X . The majority voting is performed and the instance X is allotted the class of majority of K -nearest neighbour's class [16, 17].

4 Experiments and Results

In this paper, we have used a supervised machine learning approach to classify EEG and peripheral physiological signals to differentiate between different emotional states. The EEG and peripheral physiological signals were wavelet decomposed to five-level DWT using Haar wavelets. Thirty-four features are extracted using the detail and approximate subbands of the decomposed wavelets. Vectors are formed by concatenating these features. This feature matrix is classified using KNN classifier with fivefold cross-validation method. This method is proposed to find the accuracy of the emotional analysis using psychophysiological signals.

Different pilot experiments were done and $K = 10$ was chosen based on the performance in terms of accuracy.

The accuracy is computed using the confusion matrix. The multi-class confusion matrix is $4 * 4$ matrix with true positive are placed at $i = j$ cell of A_{ij} confusion matrix. The sensitivity and specificity is calculated using confusion matrix.

$$\text{SENSITIVITY} = \frac{\text{TRUE POSITIVE}}{\text{TRUE POSITIVE} + \sum_N \text{FALSE NEGATIVE}} \quad (1)$$

$$\text{SPECIFICITY} = \frac{\text{TRUE NEGATIVE}}{\text{TRUE NEGATIVE} + \sum_N \text{FALSE POSITIVE}} \quad (2)$$

The multi-class accuracy is obtained from the sensitivity and specificity of the classified results:

$$\text{ACCURACY} = \frac{\text{SENSITIVITY} + \text{SPECIFICITY}}{2} \quad (3)$$

The four classes are classified with the KNN classifier with 87.1% accuracy.

In Fig. 1, the confusion matrix of KNN classifier with data points is inserted as true class versus predicted class is shown. The diagonal elements of the $4 * 4$ matrix depicts that the data points are accurately predicted, whereas the other elements of the matrix depicts the missclassification of the data points.

In Fig. 2, the confusion matrix is shown in terms of percentage of data points accurately classified or misclassified. In Fig. 3, the region of convergence is shown for the KNN classifier.

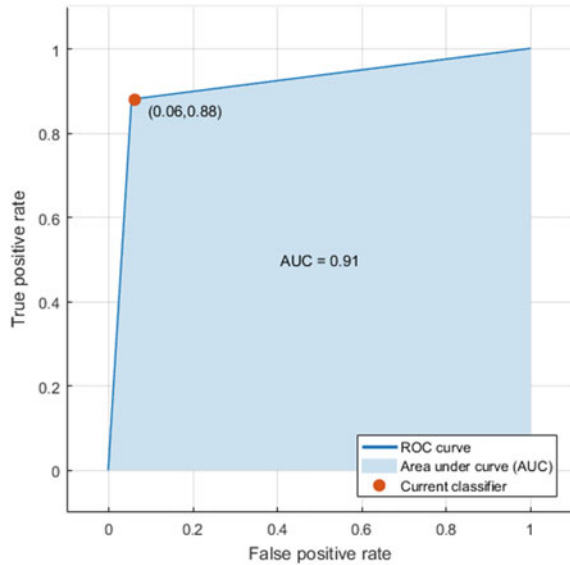
True class	1	3237	165	172	105
	2	167	2731	142	73
	3	180	133	2456	61
	4	121	77	62	1438
		1	2	3	4
		Predicted class			

Fig. 1 Confusion matrix in terms of data points of KNN classifier

True class	1	88%	4%	5%	3%	True Positive Rate	88%	False Negative Rate	12%		
	2	5%	88%	5%	2%					88%	12%
	3	6%	5%	87%	2%					87%	13%
	4	7%	5%	4%	85%					85%	15%
		1	2	3	4						
		Predicted class									

Fig. 2 Confusion matrix in terms of percentage accuracy of KNN classifier

Fig. 3 Region of convergence curve for KNN classifier



5 Conclusion

In this paper, we have proposed a DWT-based approach to the emotion analysis of the humans using video stimuli. The proposed approach trains the psychophysiological signals using the five-level DWT and then, extracting features such as mean, standard deviation, power spectral density, average decrease rate during decay time, heart rate and heart rate variability and eye blinking rate to classify the emotion expressed explicitly or even implicit emotions are detected. The KNN classifies the emotions as into four classes that are high valence high arousal, high valence low arousal, low valence high arousal and low valence low arousal. A good classification accuracy of 87.1% is obtained using KNN classifier suggest that this algorithm is suitable for training data to classify the emotions using visual stimuli.

Use of weighted KNN with feature reduction using principal component analysis will also be tried to reduce the system complexity. In future, the focus will be to use deep learning, Random forest etc., to further improve the performance.

References

1. Dasdemir Y, Yildirim E, Yildirim S (2017) Analysis of functional brain connections for positive–negative emotions using phase locking value. *Cogn Neurodyn* 11(6):487–500
2. Jahankhani P, Kodogiannis V, Revett K (2006) EEG signal classification using wavelet feature extraction and neural networks. In: International symposium on modern computing, 2006. JVA'06. IEEE John Vincent Atanasoff 2006. IEEE, New York, pp 120–124

3. Başar E, Başar-Eroglu C, Karakaş S, Schürmann M (2001) Gamma, alpha, delta, and theta oscillations govern cognitive processes. *Int J Psychophysiol* 39(2):241–248
4. Haas LF (2003) Hans Berger (1873–1941), Richard Caton (1842–1926), and electroencephalography. *J Neurol Neurosurg Psychiatry* 74(1):9
5. Hadley JM (1941) Some relationships between electrical signs of central and peripheral activity: II. During 'mental work'. *J Exp Psychol* 28(1):53
6. Plutchik R (1959) The effects of high intensity intermittent sound on performance, feeling and physiology. *Psychol Bull* 56(2):133
7. Moon J, Kim Y, Lee H, Bae C, Yoon WC (2013) Extraction of user preference for video stimuli using EEG-based user responses. *ETRI J* 35(6):1105–1114
8. Ekman P, Friesen WV, O'sullivan M, Chan A, Diacoyanni-Tarlatzis I, Heider K, Scherer K (1987) Universals and cultural differences in the judgments of facial expressions of emotion. *J Pers Soc Psychol* 53(4):712
9. Koelstra S, Muhl C, Soleymani M, Lee JS, Yazdani A, Ebrahimi T, Pun T, Nijholt A, Patras I (2012) Deap: a database for emotion analysis; using physiological signals. *IEEE Trans Affect Comput* 3(1):18–31
10. Torres-Valencia CA, García HF, Holguín GA, Álvarez MA, Orozco Á (2015) Dynamic hand gesture recognition using generalized time warping and deep belief networks. In: *International symposium on visual computing*. Springer, Cham, pp 682–691
11. Ramasamy M, Varadan VK (2017) Study of heart-brain interactions through EEG, ECG, and emotions. In: *Nanosensors, biosensors, info-tech sensors and 3D systems 2017*, vol 10167. International Society for Optics and Photonics, p 101670I
12. Li Y, Li X, Ratcliffe M, Liu L, Qi Y, Liu Q (2011) A real-time EEG-based BCI system for attention recognition in ubiquitous environment. In: *Proceedings of 2011 international workshop on ubiquitous affective awareness and intelligent interaction*. ACM, New York, pp 33–40
13. Hoa LT, Anh ND (2007) Orthogonal-based wavelet analysis of wind turbulence and correlation between turbulence and forces. *Vietnam J Mech* 29(2):73–82
14. Lang PJ, Greenwald MK, Bradley MM, Hamm AO (1993) Looking at pictures: affective, facial, visceral, and behavioral reactions. *Psychophysiology* 30(3):261–273
15. Cygankiewicz I, Wranicz JK, Bolinska H, Zaslonka J, Zareba W (2004) Relationship between heart rate turbulence and heart rate, heart rate variability, and number of ventricular premature beats in coronary patients. *J Cardiovasc Electrophysiol* 15(7):731–737
16. Zhang ML, Zhou ZH (2007) ML-KNN: a lazy learning approach to multi-label learning. *Pattern Recogn* 40(7):2038–2048
17. Parthasarathy G, Chatterji BN (1990) A class of new KNN methods for low sample problems. *IEEE Trans Syst Man Cybern* 20(3):715–718