ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 8, Issue 5, May 2021

Emotion Recognition Using Physiological Signals and Different Datasets

[1] Emily Joy, [2] Willson Joseph, [3] Dr. M. Rajeswari, [4] Dr. R. Sunder

[1][2][3][4] Department of Computer Science and Engineering, Sahrdaya College of Engineering and Technology, Kerala, India
Email: [1] emilyjoy21497@gmail.com, [2] willsonjoseph@sahrdaya.ac.in, [3] rajeswarim@sahrdaya.ac.in, [4] sunder@sahrdaya.ac.in4

***Abstract*----Now a day's emotion recognition has an important role in day-to-day life. Because emotions have greater value in human life. From the emotions, we can check whether the person is healthy or not and also the mental stability. For this mainly 8 basic emotions are fear, disgust, anger, happiness, sadness, acceptance, expectancy, and surprise. It also includes emotions like interest, guilt, shame, and neglect. To recognize these emotions mainly face recognition methods and speech recognition methods are used. But they are mainly in attention to facial expressions, speech, and gestures. By using visible signs of emotions cannot achieve the actual emotions of people. To gain true emotions, we are using Emotion Recognition using Physiological signals. In this paper, we use photoplethysmography (PPG), and galvanic skin response (GSR) signals as the data for recognizing the emotions and for classification. We use different classification methods to acquire the best accuracy. By this, we obtain emotions like sad, happy, and neutral. Using a modified random forest method, we achieve high accuracy of 97% for classification.**

***Keywords*--- Emotion recognition, Feature extraction, Machine learning, Stepwise regression, Modified random forest method**

I. INTRODUCTION

Now a day's emotion recognition has become an important one for human-computer interaction and as well as in medical fields. It is mainly used for health monitoring purposes. Because of these emotions, we can check a person's health condition. If a person has happy emotions then he is healthy. If a person has sad emotions then he is unhealthy because of their life situation. If he has neutral emotion then he is not mentally stable. So that emotions have a greater value in human-being. It is mainly helpful for paralyzed patients [14]. From those patients, it is very difficult to achieve emotions and also to check the health condition.

A human being contains 8 basic emotions like anger, disgust, sadness, happiness, acceptance, expectancy, fear, and surprise [13]. Aside from these emotions, there are emotions like guilt, shame, etc. These emotions are characterized into two attributes like two-dimensional and also three-dimensional emotions. In two-dimensional contains valence and arousal. In three-dimensional, additional to two-dimensional contain dominance. Valence is mainly pleasant feelings or positive feelings, whereas arousal is having some exciting moments.

These emotions are mainly dividing into two emotional models; the dimensional emotional model and the discrete emotional model. The dimensional emotional model contains different types of emotions but they are formed as dimensions like two-dimensional emotional space [15]. The discrete emotion model mainly contains basic emotions.

Recent studies of emotion recognition mainly by facial expression as well as speech recognition [7]. By these methods, we cannot achieve true emotions. So that emotion recognition using physiological signals has become important nowadays [10]. By using physiological signals, we can achieve true emotions because the signals are achieved directly from the central nervous system.

This paper, mainly explains emotion recognition using physiological signals by machine learning techniques. For this here we use PPG and GSR signals as our data for acquisition. The training and testing data of PPG and GSR signals from the IEEE data port dataset.

The feature extraction mainly performs in 35 subject's data. In training, the dataset contains 1374 signals and the testing dataset contains 700 signals. In both the train and test datasets, we extract the features from both the PPG signal and GSR signal. The feature extraction mainly performs according to frequency and time domain. We mainly extract 13 features from the PPG signal and 14 features from the GSR signal.

After feature extraction, we perform feature selection. From 27 features we select only some useful features. For this, we use stepwise regression. The selected features then proceed to different classification methods to obtain emotions.

To achieve the best accuracy for classification, as an addition we use a modified random forest method. we use

80% of training data and 20% of testing data. From this, we extract sad, happy, and neutral emotions.
By using the modified random forest method, we achieve high accuracy of 97% for classification when compared to other classification methods. We can recognize actual emotions from this method, so that they may help for medical purposes and human-computer intercommunication.

II. COMPARISON OF DATASETS

For emotion recognition uses different types of datasets available on websites. These databases should vary with one another according to the number of participants, the number of channels of signals, and how each signal is recorded. These different types of datasets and their comparisons are discussed below:

A. DEAP

The database for emotion analysis using physiological signals (DEAP) is a dataset that accommodates 32 channels of EEG signals from 32 participants for 40 trials [11]. The signals are recorded by watching music videos for one minute. According to the three-dimensional dominance, the rating can be performed by using valence, arousal, and dominance. This emotional state is classified according to the valence and arousal into four classes containing HALV, HAHA, LALA, and LAHV. In this EEG signal is sampled into 128Hz. It is a three-dimensional emotional model.

B. TYUT 2.0

TYUT2.0 is a dataset that accommodates 64 channels of EEG signals from 16 participants for 250 trials. The signals are recorded by watching a speech clip for 0.5 to 3 sec. These emotions are classified into 5 states containing sadness, anger, happiness, surprise, and neutral. EEG signal is sampled at 1000Hz. It is a discrete emotional model.

C. MPED

Multi-modal physiological emotion database (MPED) [4] is a dataset that accommodates 62 channels of EEG signals from 23 participants. The signals are recorded by watching 1500 video clips. These emotions are classified into seven states containing joy, funny, disgust, anger, fear, sadness, and neutrality. It records different types of signals like EEG, ECG, RSP, and GSR. It samples the signals at a rate of 1000Hz. It is a discrete emotional model.

D. MAHNOB

MAHNOB is a dataset that accommodates videos, audio signals, eye gaze data, and physiological signals from 27 participants. The signals are recorded by watching 20 emotional videos. This emotion is classified according to the valence and arousal. It is a two-dimensional emotional model.

E. E4-dataset

E4-dataset [4] is a dataset that accommodates videos and physiological signals from 24 participants. The signals are recorded by watching 52 emotional film clips. This emotion is classified according to valence, arousal, and dominance. It is a three-dimensional emotional model.

F. SEED

The SJTU emotion EEG dataset (SEED) is a dataset that accommodates 62 channels of EEG signals from 15 participants for 15 trials [18]. The signals are recorded by watching 15 emotional film clips. These emotions are classified into three states containing positive, neutral, and negative. EEG signal is sampled at 1000Hz. It is a discrete emotional model.

III. RELATED WORKS

For recognizing emotion there are many methods like face recognition [8], speech recognition, etc. By using these methods, we cannot achieve actual emotions. Because these are visible signs of emotions so there is a chance to hide emotions [9]. In this, we mainly focus on emotion recognition using physiological signals to achieve actual emotion. We can increase the accuracy of emotion recognition in many ways like by using different datasets, feature extractions methods, etc.
In Emotion Recognition Based on DoubleTree Complex Wavelet Transform and Machine Learning in IoT [2], to increase the accuracy of classification they use Double Tree Complex Wavelet Transform and Machine Learning (DTCWT) for feature extraction method. It has properties like translation invariance and anti-aliasing effect. Here they use electroencephalogram (EEG) and electromyography (EMG) signals as training and test data. For classification, they use the support vector machine (SVM) method. They classify clam, happy and sad with an accuracy of 87 % for classification.
In Emotion feature analysis and recognition based on reconstructed EEG sources [3], we use a scalp EEG dataset because it contains assorted signals so we can extract from active source regions [16]. The database for emotion analysis using physiological signals (DEAP) and TYUT2.0 are the two datasets they used to achieve the data of scalp EEG signals. Here they use the sLORETA tool for extracting active sources of regions. Using this they extract 26 Brodmann areas as the region of interest. By using

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 8, Issue 5, May 2021

SVM as the classification they achieve high accuracy in both DEAP and TYUT.0 datasets.

In MPED: A multi-modal physiological emotion database for discrete emotion recognition [4], we increase the classification accuracy by developing an MPED dataset. This dataset contains EEG, GSR, respiration (RSP), and ECG signals [12], [17]. It can accommodate 62 channels of EEG signals from 23 participants [19]. From each signal, we extract different features. From this, they can classify different emotions; joy, funny, anger, fear, sadness, disgust, and neutrality. For classification, they use the SVM or KNN method. When compared SVM with K-nearest neighbor (K-NN), SVM has the highest accuracy of classification.

In towards user-independent emotion recognition using physiological signals [5], they use two datasets; MAHNOB and E4 dataset. E4 dataset is mainly developed to increase the accuracy. But MAHNOB has high accuracy. Here they use distributed approach in both data collection and preprocessing. A centralized approach is used in data fusion. So that in this they use a hybrid approach. For classification, they use weighted multi-dimensional dynamic time warping (WMD-DTW) and K-NN. They use different hyperparameter optimization methods; sequential model-based optimization (SMBO) and Gaussian process-based optimization (GP). During optimization, they perform a cross-validation method to check whether it performs well or not. They use a stacking classifier to merge the multiple models. so that we can increase the accuracy.

In Interpretable emotion recognition using EEG signals [6], they use DEAP and SEED dataset is used to achieve EEG signals. They use preprocessed data. So that there is no chance of noise in it. The feature extraction is performed in both datasets. For the DEAP dataset, they use both first and second derivatives. For the SEED dataset, they use differential entropy. They use an autoencoder to reduce dimensionality. For classification, they made the classification method through a decision tree, K-NN, and random forest method. At last for model assessment, they perform 10-fold cross-validation. In this, they mainly form an emotion activation curve from correlation, entropy, and weight coefficient. The weight coefficient form from both correlation and entropy coefficient. From the weight coefficient, they form high accuracy.

IV. METHODOLOGY

For an Emotion recognition using physiological signal mainly contain data collection, feature extraction, feature selection, classification, and at last model assessment. Fig.1 shows the process of emotion recognition using physiological signals.

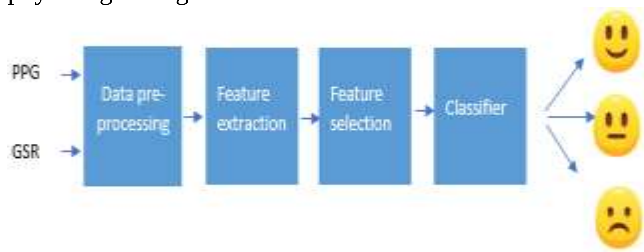

Fig. 1. Emotion Recognition using physiological signal Process [1].

In this, we use PPG and GSR signals for data acquisition. We achieve datasets from the IEEE data port. From this signal, we extract features according to frequency and time domain. Then these features are trained for classification. This signal processing was done in Matlab. From this system, we can identify three emotions; happy, sad, and neutral. Happy is a feeling many when watching a comedy scene. Sad is a feeling when watching an emotional scene. Finally, neutral is a feeling in between both happy and sad or no emotional feeling.

A. PHYSIOLOGICAL SIGNALS ACQUISITION

The PPG and GSR signals are used for data acquisition. For this, we have 80% of training data and 20% of testing data.

The galvanic skin response (GSR) is used to calculate the electrical conductivity of the skin. By using this we can measure the blood pressure. If skin conductivity increases, emotion will be positive. If skin conductivity decreases, emotion will be negative. In this, we use the data recorded from the middle and index finger.

Photoplethysmography (PPG) is used to detect any changes in blood volume. It can be used to calculate the heart rate and also oxygen saturation. Due to negative emotions, there is a chance of an increase in heart rate. If a person has a normal heart rate then he is healthy. In this, we use the data recorded from the ring finger.

For the train and test data, we use PPG and GSR data from the IEEE data port. The feature extraction mainly performs in 35 subject's data. In training, the dataset contains 1374 signals and the testing dataset contains 700 signals.

B. FEATURE EXTRACTION

In this, we have 35 subjects' signal data. From one subject we take 59 signals. For training data 1374 signals and testing data 700 signals in the dataset. Before feature extraction, we remove the noise from the data. For that, we use different filtration methods for different signals.

For the PPG signal, we processed with a 100th order bandpass finite impulse response filter with distinct frequencies equal to 0.1 Hz and 10 Hz and for the GSR signal, we processed with low pass filtered with a 1000th order finite impulse response filter with a distinct frequency of 1 Hz.

To obtain the heart rate, we use short-time fast Fourier transform (ST-FFT). For obtaining an effective frequency resolution of 2 beats per minute (BPM), we use 5-s length windows and zero-padding.

For PPG and GSR, we extract different features according to the time domain and frequency domain. In PPG signals we extract 13 features. From time-domain features, we calculate root mean square differences of successive R-R intervals (HRRMSSD) and standard deviation of normal to normal R-R intervals (HRSDNN). From the frequency domain feature, we calculate the mean of heart rate (hrmean), the standard deviation of heart rate (hrstd), heart rate dynamic range (hrdr), heart rate mode (hrmode), harmonic distortion of the second, third, fourth and fifth harmonics of the PPG signal (THD2, THD3, THD4, THD5).

In the GSR signal, we extract 14 features. From time-domain features, we calculate mean (scrmean), standard deviation (scrstd), dynamic range (scrdr), mean of the derivative (scravd), negative values of the derivative (scraonv), the ratio of negative values over the total number of samples (scrpnv), empirical mode decomposition (EMD), energy (emf1, emf2, emf3, emf4) and the zero-crossing rate (crm1, crm2, crm3, crm4) of each mode.

C. FEATURE SELECTION

In feature extraction, we extract a total of 27 features according to the time and frequency domain. But from these, we only need some of the features for classification. For these many methods are their; stepwise regression (SW), random forest recursive feature elimination (RF-RFE), and genetic algorithms (GAs).

Here we use stepwise regression for feature selection. It is mainly used for linear regression models. This method can consist of regressing multiple variables. From this, we form only independent variables with non-zero coefficients. In this, we use SW with the Akaike information criterion (AIC) as the stop criterion. It is an estimator used to find the errors in the prediction. Here according to the quality of the model, we select the features.

D. CLASSIFICATION

From the feature selection process, we select features for classification. From this classify three emotions; happy, sad, and neutral.

For classification, we get hold of 80% of training data and 20% of testing data. To classify the emotion, we use different classification methods like Support Vector Machine (SVM), Decision Tree, K-Nearest Neighbors(K-NN). At last, we calculate the accuracy of each classification method by true positive (TP) and true negative (TN).

$$\text{Accuracy} = \frac{TP + TN}{P + N} \qquad \ldots\ldots (1)$$

Where P is positive emotions and N is negative emotions.

In this, we use the classification learner application in Matlab to achieve different classification methods. By using this application, it can explore supervised machine learning using different classifiers. They automatically train the data and classify them according to the classifier.

E. MODIFIED RANDOM FOREST METHOD

It is a random forest method with the monte carlo sampling technique and removing uninformative variables. By performing monte carlo sampling, we select informative variables from the dataset by giving a particular threshold value. After that uninformative variables will be removed by the Uninformative Variable Elimination method (UVE) [20]. Then perform the random forest method into an informative variable.

V. EXPERIMENTAL RESULTS

In emotion recognition using physiological signals, the accuracy of the model is very important. The accuracy of classification depends on the dataset and also the methods of feature extraction and feature selection. We get an accuracy of 97% for classification by using testing data using the modified random forest method. When compared to other classification methods, the modified random forest method form high accuracy.

In this, for testing we use few id's; 1,2,5,34,100, 250, 450, and 600. From this form happy and sad emotions at an accuracy of 97% for classification.

```
reading test dataset
Extracting ppg features of testing set..
Extracting gsr features of testing set..
Testing data..
Accuracy : 0.84143
Testing id : 1  Emotion: SAD
Testing id : 2  Emotion: SAD
Testing id : 5  Emotion: SAD
Testing id : 34  Emotion: SAD
Testing id : 100  Emotion: SAD
Testing id : 250  Emotion: HAPPY
Testing id : 450  Emotion: HAPPY
Testing id : 600  Emotion: HAPPY
```

Fig. 2. Accuracy of classification by using

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 8, Issue 5, May 2021

```
reading test dataset
Extracting ppg features of testing set..
Extracting gsr features of testing set..
Testing data..
Accuracy : 0.83
Testing id : 1  Emotion: SAD
Testing id : 2  Emotion: SAD
Testing id : 5  Emotion: SAD
Testing id : 34  Emotion: SAD
Testing id : 100  Emotion: SAD
Testing id : 250  Emotion: HAPPY
Testing id : 450  Emotion: HAPPY
Testing id : 600  Emotion: HAPPY
```

Fig. 3. Accuracy of classification by using K-NN.

```
reading test dataset
Extracting ppg features of testing set..
Extracting gsr features of testing set..
Testing data..
Accuracy : 0.95
Testing id : 1  Emotion: SAD
Testing id : 2  Emotion: SAD
Testing id : 5  Emotion: SAD
Testing id : 34  Emotion: SAD
Testing id : 100  Emotion: SAD
Testing id : 250  Emotion: HAPPY
Testing id : 450  Emotion: HAPPY
Testing id : 600  Emotion: HAPPY
```

Fig. 4. Accuracy of classification by using Decision Tree.

Classification methods	Accuracy	Training time
Complex tree	61.4	6.47 sec
Simple tree	65.6	4.79 sec
Medium tree	65.3	5.01 sec
Logistic regression	66.2	9.723 sec
Quadratic SVM	66.3	304.54 sec
Cubic SVM	61.7	363.02 sec
Fine gaussian SVM	64.1	9.2452 sec
Medium gaussian SVM	66.1	12.657 sec
Coarse gaussian SVM	66.2	11.006 sec
Fine KNN	55	11.983 sec
Medium KNN	64	12.44 sec
Coarse KNN	66.2	13.863 sec
Cosine KNN	63.7	13.335 sec
Cubic KNN	63.3	22.303 sec
Weighted KNN	61.6	14.272sec
Ensemble boosted trees	65.8	29.429 sec
Ensemble bagged trees	64.3	32.737 sec
Ensemble subspace discriminant	66.2	33.309 sec
Ensemble subspace KNN	58.4	37.595 sec
Ensemble RUS Boosted trees	51.8	43.552 sec

Fig. 5. Accuracy of classification by using classification learner application.

```
reading test dataset
Extracting ppg features of testing set..
Extracting gsr features of testing set..
Testing data..
Accuracy : 0.96857
Testing id : 1  Emotion: SAD
Testing id : 2  Emotion: SAD
Testing id : 5  Emotion: SAD
Testing id : 34  Emotion: SAD
Testing id : 100  Emotion: SAD
Testing id : 250  Emotion: HAPPY
Testing id : 450  Emotion: HAPPY
Testing id : 600  Emotion: HAPPY
```

Fig. 6. Accuracy of classification by using modified random forest method.

VI. CONCLUSION AND FUTURE WORK

In this paper, we discuss emotion recognition using physiological signals by machine learning techniques. This includes processes like data collection, data processing, feature extraction, feature selection, classification, and model assessment.

The PPG and GSR signals are taken as the data for acquisition. The training and testing data of PPG And GSR signals from the IEEE data port dataset. From this data, we extract 27 features using the feature extraction method. By performing stepwise regression, we achieve useful features. For the classification of these features, we use different classification methods. By this, we can extract happy, sad, and neutral emotions. In this, we achieve high accuracy of 94% for classification using the modified random forest method.

Future works include the development of more public datasets with several subjects and also an emotion recognition system that can predict emotions in real-time.

VII. ACKNOWLEDGEMENTS

We thank all the anonymous referees for their useful suggestions and supports.

REFERENCES

[1] J. A. Dominguez- Jimenez, K. C. Campo- Landines, J. C. Martinez-Santos, E. J. Delahoz, S. H. Contreras-Ortiz, “A machine learning model for emotion recognition from physiological signals”, Elsevier, 2020.

[2] Xin Xu, Yiwei Zhang, Minghong Tangn, Hong Gu, Shancheng Yan, and JieYang, “Emotion recognition based on double tree complex wavelet transform and machine learning in the internet of things”, IEEE Access, 2019.

[3] Guijun Chen, Xueying Zhang, Ying Sun, and Jing Zhang, "Emotion feature analysis and recognition based on reconstructed EEG sources", IEEE Access, 2019.
[4] Tengfei Song, Wenming Zheng, Cheng Lu, Yuan Zong, Xilei Zhang, and Zhen Cui, "MPED: A multi-modal physiological emotion database for discrete emotion recognition", IEEE Access, 2019.
[5] Amani Albraikan, Diana P. Tobon and AbdulmotalebEl Saddik, "Toward user-independent emotion recognition using physiological signals", IEEE Sensors Journal, 2019.
[6] Chunmei Qing, Rui Qiao, Xiangmin Xu, and Yongqiang cheng, "Interpretable emotion recognition using EEG signals", IEEE Transactions on multimedia,2019.
[7] R. Cowie, R.R. Cornelius, Describing the emotional states that are expressed in speech, Speech Commun. 40 (1–2) (2003) 5–32.
[8] M.A. Turk, A.P. Pentland, Face recognition using eigenfaces, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1991. Proceedings CVPR'91, IEEE, 1991, pp. 586–591.
[9] A.P. Atkinson, M.L. Tunstall, W.H. Dittrich, Evidence for distinct contributions of form and motion information to the recognition of emotions from body gestures, Cognition 104 (1) (2007) 59–72.
[10] A. Heraz, M. Clynes, Recognition of emotions conveyed by touch through force-sensitive screens: observational study of humans and machine learning techniques, JMIR Mental Health 5 (3) (2018), e10104, http://dx.doi.org/10. 2196/10104.
[11] N. Sebe, I. Cohen, T.S. Huang, Multimodal emotion recognition, in: Handbook of Pattern Recognition and Computer Vision, World Scientific, 2005, pp. 387–409.
[12] M. Soleymani, M. Pantic, T. Pun, Multimodal emotion recognition in response to videos, IEEE Trans. Affect. Comput. 3 (2) (2012) 211–223.
[13] O. Alaoui-Ismaïli, O. Robin, H. Rada, A. Dittmar, E. Vernet-Maury, Basic emotions evoked by odorants: comparison between autonomic responses and self-evaluation, Physiol. Behav. 62 (4) (1997) 713–720.
[14] R.W. Levenson, L.L. Carstensen, W.V. Friesen, P. Ekman, Emotion, physiology, and expression in old age, Psychol. Aging 6 (1) (1991) 28.
[15] I.C. Christie, B.H. Friedman, Autonomic specificity of discrete emotion and dimensions of affective space: a multivariate approach, Int. J. Psychophysiol. 51 (2) (2004) 143–153.
[16] R.W. Picard, J. Healey, Affective wearables, Pers. Technol. 1 (4) (1997) 231–240.
[17] J. Scheirer, R. Fernandez, R.W. Picard, Expression glasses: a wearable device for facial expression recognition, in: CHI'99 Extended Abstracts on Human Factors in Computing Systems, ACM, 1999, pp. 262–263.
[18] A. Haag, S. Goronzy, P. Schaich, J. Williams, Emotion recognition using bio-sensors: first steps towards an automatic system, in: Tutorial and Research Workshop on Affective Dialogue Systems, Springer, 2004, pp. 36–48.
[19] T. Hui, R. Sherratt, Coverage of emotion recognition for common wearable biosensors, Biosensors 8 (2) (2018) 30, http://dx.doi.org/10.3390/ bios8020030.
[20] W. Cai, Y. Li and X. Shao, Chemom. Intell. Lab. Syst., 2008, 90, 188–194