International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-4, November 2019

Emotional Interfaces for Effective E-Reading using Machine Learning Techniques

Indhumathi R, Geetha A

Abstract: ***Emotion-aware systems are very essential for effective e-reading. The aim of the proposed work is to detect and classify cognitive states from facial expressions of the students engaged in online learning which improves the e-reading process to a greater extent. In this proposed work the emotions such as happy, irritate, sleep and yawn that are mainly used for effective E-reading are taken into consideration. The Haar cascaded classifier is used to segment the facial regions from the input images. The Zernike moment features are extracted from the selected face regions. The extracted features are fit into Random Forest and Decision Tree machine learning models. The models classify the emotions. Finally the classified emotions are interfaced with e-reading. The proposed work is found to perform better than the existing methods.***

Keywords: Emotion classification, Haar cascaded classifier, Random Forest, Zernike moments, Gabor filter

I. INTRODUCTION

A. Emotion Classification

The term emotion is gotten from the Latin action word 'movere' which means stir up, agitate, disturb or move. In psychology, emotion is regularly characterized as a mind boggling condition of feeling that outcomes in physical and mental changes that impact thought and behavior. There are three components of emotions such as cognition, feeling and behavior. The cognition segment serves fundamentally to impact an assessment of given circumstance, inciting us to become emotional in one way or another, or not at all. In day by day life we consider emotions. The emotions are most promptly clear changes in an aroused individual. The behavioral component involves facial, postural, gestures and vocal responses [1].

During the 1970s, psychologist Paul Eckman distinguished six essential emotions that he proposed were all around experienced in every human culture. The emotions he recognized were happiness, sadness, disgust, fear, surprise, and anger. He later extended his rundown of fundamental emotions to incorporate such things as pride, disgrace, shame, and excitement [2]. Many researchers had taken these emotions for their research. Some mixed emotions are also used for research.

Manuscript published on November 30, 2019.
***** Correspondence Author

Indhumathi R[*], Research Scholar, Department of Computer and Information Science, Annamalai University, Annamalai Nagar, Tamilnadu, India.

Geetha A, Professor, Department of Computer Science and Engineering, Annamalai University,Annamalai Nagar, Tamilnadu, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license http://creativecommons.org/licenses/by-nc-nd/4.0/

The emotions are identified and applied to artificial intelligence. The emotions are applied to various applications in computer vision to create interactive applications.

The Fig1 shows the various emotions taken for this proposed work.

Fig.1a Happy

Fig.1b Sleep

Fig.1c Irritate

Fig.1d Yawn

Fig.1. Emotions for proposed work

B. Zernike Moments

Minute descriptors have been read for image recognition and computer vision since 1960s [3]. Teague originally presented the utilization of Zernike moments to conquer the weaknesses of data excess present in the well known geometric moments [4, 5]. Zernike moments are a class of orthogonal moments and have been demonstrated compelling as far as image representation. Zernike moments are rotation invariant and can be effectively developed to an arbitrary order. Although higher order moments convey all the more fine subtleties of an image, they are also more susceptible to noise. Consequently various requests of Zernike minutes are tested to decide the optimal order for our problem.

C. Decision Tree

Decision tree taxonomy is a flowchart-like tree structure where an internal node denotes feature (or attribute), the branch indicates a decision rule, and each leaf node represents the outcome. The topmost node in a decision tree is known as the root node or parent node. It learns to split on the basis of the attribute value. It partitions the tree in recursive manner called recursive splitting.

This flowchart-like structure helps in decision making which easily mimics the human level thinking. That is why decision trees are easy to understand and interpret [6].

Decision Tree is a white box kind of machine learning algorithm. It shares inner basic decision-making logic, which isn't accessible operating at a black box sort of algorithms, for example, Neural Network. Its training time is quicker contrasted with the neural network algorithm. The time multifaceted nature of decision trees is a component of the quantity of records and number of attributes in the given data. The decision is a conveyance free or non-parametric technique, which doesn't rely on probability distribution assumptions. Decision trees can deal with high dimensional information with great precision [6].

D. Random Forest

Random forest algorithm is a supervised machine learning classification algorithm. As the name recommends, this algorithm makes the forest with various trees. As a rule, the many trees in the forest the more powerful the forest resembles. Similarly in the Random forest classifier has higher the quantity of trees, higher the accuracy results. Following are the benefits of Random forest classification contrasted with other classification algorithms.

- The over fitting issue will never occur
- The same arbitrary woodland calculation can be utilized for classification and regression.
- The random forest algorithm can be accustomed to recognize the most vital features out of the accessible features from the training dataset.

II. RELATED WORK

Ai Sun et al. [7] proposed a deep learning method called convolutional neural network (CNN) for their research. Three databases (CK+, JAFFE and NVIE) are picked to train and test the model. 10-overlap cross validation technique is utilized to figure the accuracy. Aswin T S et a.l [8] proposed multiuser face detection based E-Learning system performed by support vector machine. Kiavash Bahreini et al. [9] present a framework for improving learning by using Webcams and Microphones. It offers timely and relevant feedback for learner's facial expressions and verbalizations. Krithika L.B et al. [10] proposed a framework that can distinguish and monitor emotions of the learners in an e-learning condition and give a real-time feedback mechanism to upgrade the e-learning helps for a superior substance conveyance. Recognition of eyes and head development can assist us with understanding student level of fixation. Salma Boumiza et al. [11] proposed a automatic tutor dependent on face recognition and emotion recognition framework for a distant learning stage. SL happy et al. [12] proposed a non-nosy, independent model for smart evaluation of emotion and alertness state just as age of proper criticism. Utilizing the non-meddling viewable signs, the framework characterizes emotion and alertness condition of the user, and gives suitable feedback as indicated by the recognized cognitive state utilizing outward appearances, visual parameters, stances, and signals.

III. SYSTEM DESIGN

A. Architecture for the proposed work

The overall architectural diagram of the proposed work is shown in Fig 2.

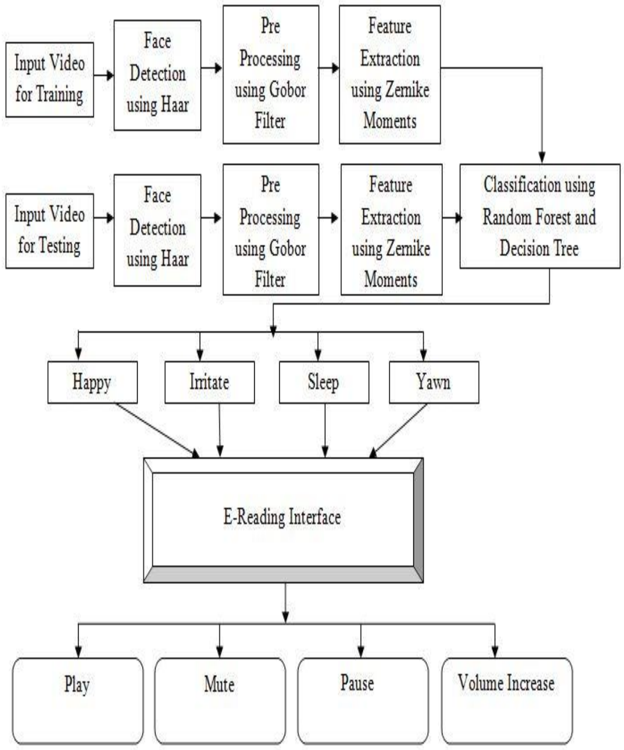

Fig.2 Architectural diagram of the proposed work

The phases of proposed work are as follows:

B. Face Detection

An input video is captured by web camera and then the videos are converted into frames. The faces are segmented by pre defined Haar cascaded classifier which is mostly used in python. In this work, Haar frontal face default classifier is used for face detection. The classifier was designed with a base size of 150 x 150 pixels, a scaling element of 1.3 and at any rate 5 neighborhoods. The greatest face in the image is picked as a region of interest.

C. Pre Processing

The RGB images are converted into grayscale format. Zernike moments are not invariant to scale and interpretation and in this way the images are first scaled and interpretation is standardized with the end goal that they are of a similar measurement and their centroids are situated at the origin point. Each face results in size of 200 X 100 pixels after scale normalization.

D. Gabor filter

"In this proposed work, Gabor filter is used with Zernike moments. This filter is used to pre process the face image. A Gabor filter responds to edges and texture changes. At the point when a Gabor channel is applied to an image, it gives the most elevated reaction at edges and at focuses where texture changes. The accompanying Fig 3 shows a test picture and its change after the filter is applied.

International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-4, November 2019

Fig.3 Image filtering using Gabor filter

There are certain parameters that affect the performance of a Gabor filter. In OpenCV Python, a Gabor kernel is created as "cv2.getGaborKernel (ksize, sigma, theta, lambda, gamma, psi, ktype)" with the following values.

- Ksize - size of Gabor filter (n, n) : 11×11
- Sigma - standard deviation of the Gaussian function : 5
- Theta - orientation of the normal to the parallel stripes : 45
- Lambda - wavelength of the sinusoidal factor : 8
- Gamma - spatial aspect ratio : 1
- Psi - phase offset value : 5
- Ktype - type and range of values that each pixel in: CV_32F

IV. FEATURE EXTRACTION

A. Zernike Moments

In this proposed work, Zernike moments are used to extract the features from face images. Different orders from order 1 to order 14 are evaluated with different image sizes to determine the optimal order for our problem. Zernike moments of order 8 with radius 21 which use only 25 features achieved better results than the other orders for image recognition. The recognition rate is found to be 89% for the image size 200 x 100. The image input is taken as gray image and the feature vector is displayed in the form of floating point values. A sample feature vector for single image with order 8, radius 21 with image size 200 x 100 is given below:

```
[ 0.31830989  0.3086861   0.56221425  0.11831021  0.28340694  0.19343704
  0.29430076  0.27417391  0.10862911  0.13383791  0.18800993  0.13895551
  0.0796991   0.18531897  0.24756046  0.22765364  0.05906204  0.09073461
  0.28134067  0.26975093  0.05881787  0.11253208  0.13179284  0.22308596
  0.13315095]
```

V. CLASSIFICATION

A. Decision Tree Classifier

In this proposed work, Decision Tree is used for classification. In python the sklearn package is used to import the Decision Tree classifier. There are three attributes such as information gain, gain ratio and gini index that are used for decision tree model based upon the requirement. In this proposed work the gini index is used for classification.

Steps for decision tree algorithm:

1. Select the Gini index attribute for Attribute Selection Measures (ASM). It is used to split the records.
2. Assure that attribute as decision attribute node and break the data into smaller subsets.
3. Start building the tree by recursively repeating this process for each child node until one of the following conditions matches:

- All the tuples belong to the similar attribute value.
- There are not any more remaining attributes.
- There are no more occasions.

The decision tree used three parameters for optimizing the performance i.e., criteria, splitter and max_depth. In this proposed work the gini index criteria is used for classification. It is shown below as:

$$\text{Gini }(D) = 1-\sum_{i=1}^{m} Pi^2 \qquad (1)$$

where pi is the probability that a tuple in D has a place with class Ci. The gini Index thinks about a twofold split for each attribute. A weighted aggregate of the impurity of each partition is processed. If a twofold split on attribute a partitions data D into D1 and D2, the Gini index of D becomes as below

$$\text{Gini}_{A}(D) = \frac{|D1|}{|D|}\text{Gini }(D_1) + \frac{|D2|}{|D|}\text{Gini }(D_2) \qquad (2)$$

If there should arise an occurrence of a discrete-esteemed attribute, the subset that gives the base gini index is chosen as a parting attribute. On account of continuous, the technique is to choose each pair of adjacent values as a conceivable split-point with littler gini index picked as below:

$$\Delta\text{Gini }(A) = \text{Gini }(D) - \text{Gini}_{A}(D) \qquad (3)$$

The attribute with minimum gini index is chosen as the splitting attribute for better classification.

B. Random Forest Classifier

It is an ensemble strategy (in view of the divide-and-conquer approach) of decision trees created on a randomly split dataset. This gathering of decision tree classifiers is otherwise called the forest. The individual decision trees are produced utilizing an attribute selection indicator such as information gain, gain ratio, and gini index for each attribute. Random forest is considered as an exceptionally exact and hearty technique on account of the quantity of decision trees taking an interest all the while. It doesn't experience the ill effects of the over fitting issue. The main reason is that it takes the average of all the predictions, which cancels out the biases. The algorithm can be used in both classification and regression problems. Arbitrary backwoods can likewise deal with missing qualities.

Retrieval Number: D8391118419/2019©BEIESP
DOI:10.35940/ijrte.D8391.118419
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

There are two different ways to deal with these: utilizing middle qualities to supplant continuous variables, and processing the proximity weighted average of missing values.

Steps for Random Forest Algorithm:

1. Randomly select samples from a given dataset.
2. Construct a decision tree for each sample.
3. Get a predicted result from each decision tree.
4. Apply a vote for each predicted result.
5. Select the most voted prediction result as the final prediction.

Scikit-learn provide an additional variable with the model, which shows the relative significance or participation of each feature in the prediction.

C. Training

Face emotions are captured from web camera and classified as four types such as happy, irritate, yawn and sleep using Random forest and Decision Tree models. About 1200 training samples, 300 for each class are used for training. Every frame is resized as 200 x 100 and all images are converted as gray scale for better performance.

D. Testing

About 200 samples, 50 samples for each class are tested. The video frames are resized as 200 x 100 and converted into gray scale images. The models classify the frames as happy / irritate/ yawn/ sleep.

E. Environmental setup

The machine learning environment has been set by installing python 3.5 versions along with anaconda library. OpenCV (version 3.3.0) library is successfully linked with python 3.5 interpreter. The required libraries such as numpy, PIL, Scikitlearn, scipy, sklearn, pyautogui, glob and mahotas are imported using conda install command.

VI. PERFORMANCE ANALYSIS

The performance of the proposed system with Random Forest and Decision Tree models are measured using the following performance metrics.

A. Confusion Matrix

A confusion matrix is a table that is describes the performance of a classification model. The proposed system is tested with 50 samples for each gesture. The result of confusion matrix for Random Forest and Decision Tree are shown in table 1 and table 2 respectively.

Table -I: Confusion Matrix for Decision Tree

ACTUAL CLASS		PREDICTED CLASS			
		Happy	Irritate	Yawn	Sleep
	Happy	40	0	5	5
	Irritate	4	38	5	3
	Yawn	2	2	41	5
	Sleep	2	3	0	35

Table -II: Confusion Matrix for Random Forest

ACTUAL CLASS		PREDICTED CLASS			
		Happy	Irritate	Yawn	Sleep
	Happy	38	0	5	3
	Irritate	2	44	3	1
	Yawn	8	5	37	0
	Sleep	5	1	10	34

B. Classification Report

Various performance metrics are used to evaluate machine learning algorithms. Here the metrics used for evaluation are Precision, Recall, F-score and Accuracy.

- **Precision:**

Precision *P* is the number of correctly predicted positive examples divided by the sum of examples that are classified as positive as shown below:

$$P = \frac{TP}{TP+FP} \tag{4}$$

- **Recall:**

Recall *R* shown below is the number of correctly classified positive examples divided by the sum of actual positive examples in the test set.

$$R = \frac{TP}{TP+FN} \tag{5}$$

- **F-Score:**

It is hard to compare two classifiers using two measures. F-score shown below combines precision and recall into one measure.

$$F - score = \frac{2PR}{P+R} \tag{6}$$

- **Accuracy:**

Accuracy is calculated by the equation shown below. The accuracy shows how effectively the classifier predicts the testing data.

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \tag{7}$$

The classification reports for Random Forest and Decision Tree are given in table 3 and table 4 respectively.

Table -III: Classification Report for Decision Tree

	Precision (%)	Recall (%)	F-Score (%)	Accuracy (%)
Happy	80	83	81	91
Irritate	76	88	82	91
Yawn	82	67	74	85
Sleep	70	72	71	86
Avg/ Total	77	78	77	88

Table -IV: Classification Report for Random Forest

	Precision (%)	Recall (%)	F-Score (%)	Accuracy (%)
Happy	76	71	73	87
Irritate	88	81	84	92
Yawn	74	67	70	85
Sleep	68	89	77	91
Avg/ Total	77	77	76	89

The Performance of Decision Tree is compared with Random Forest as shown below. The proposed produces 89% accuracy with Random Forests whereas 88% for the Decision Trees. The Random Forest produces better accuracy than decision trees.

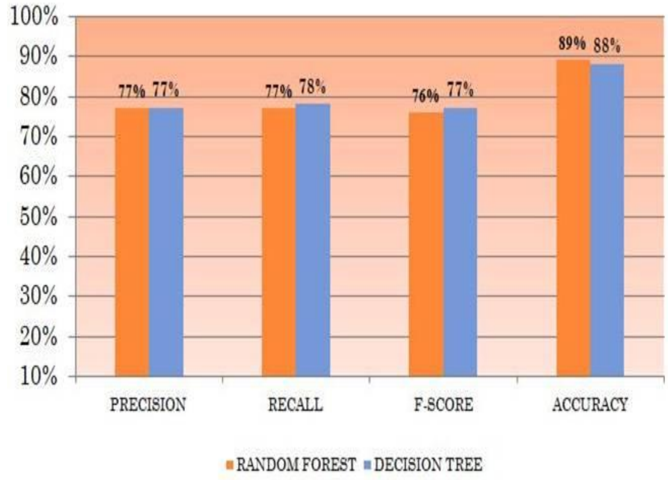

Fig.4. Performance Comparison

VII. OUTPUT

The screenshots below show the classification results of the classifiers used in the proposed work.

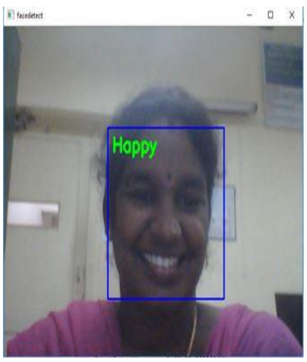

Fig.5a Output Happy

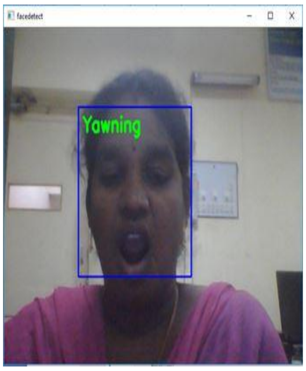

Fig 5b Output Yawning

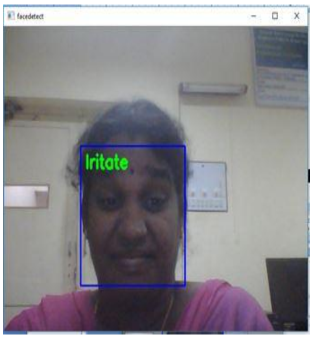

Fig.5c Output Irritate

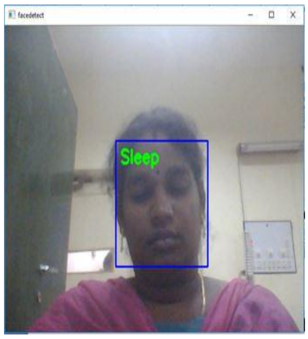

Fig. 5d Output Sleep

Fig.5 Output Screenshots for Classification

VIII. INTERFACING WITH E-READING

The e-reading can be more effective if the cognitive states of the learners are considered. E-reading system should automatically adapt to the cognitive states of the learners. In this proposed work, four emotions such as happy, irritate, sleep and yawn are considered. The computer providing online lesson videos acts like a mentor which automatically redirects and reinforces the learner based on their mood. The emotions classified from the models are interfaced with the e-reading system.

If the learners happy the machine assumes that the learner is interested to learn and proceeds with the consecutive pages of the lesson. If the learner is irritated then the machine understands the learner is not interested and reduces the volume of the video till he retains to the normal mood. If the learner is sleepy the video pauses for a while. If the learner yawns the volume of the video will be increased to get back the leaner in learning mood. The output screenshots of the interfacing with online video are shown below.

Fig.6a Happy – Play

Retrieval Number: D8391118419/2019©BEIESP
DOI:10.35940/ijrte.D8391.118419
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

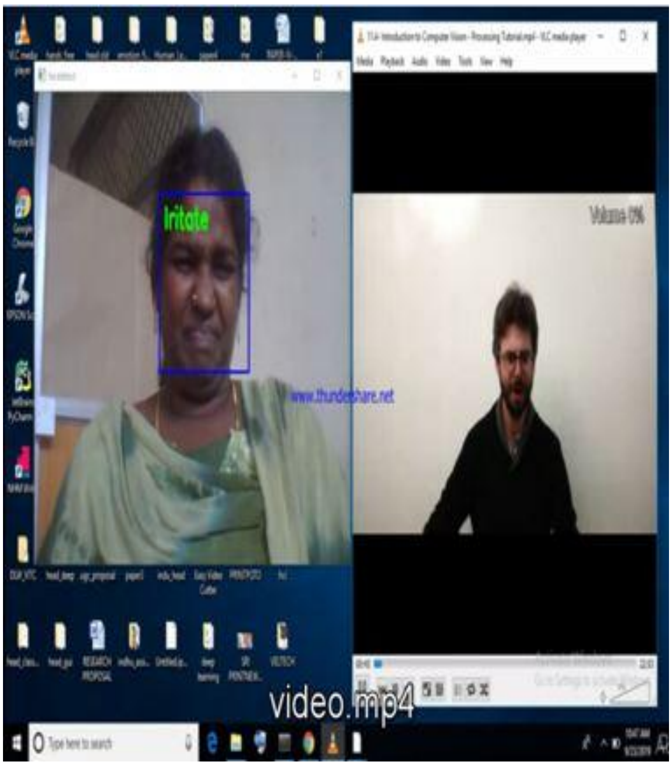

Fig.6b Irritate – volume reduce

Fig.6c Sleep – Pause

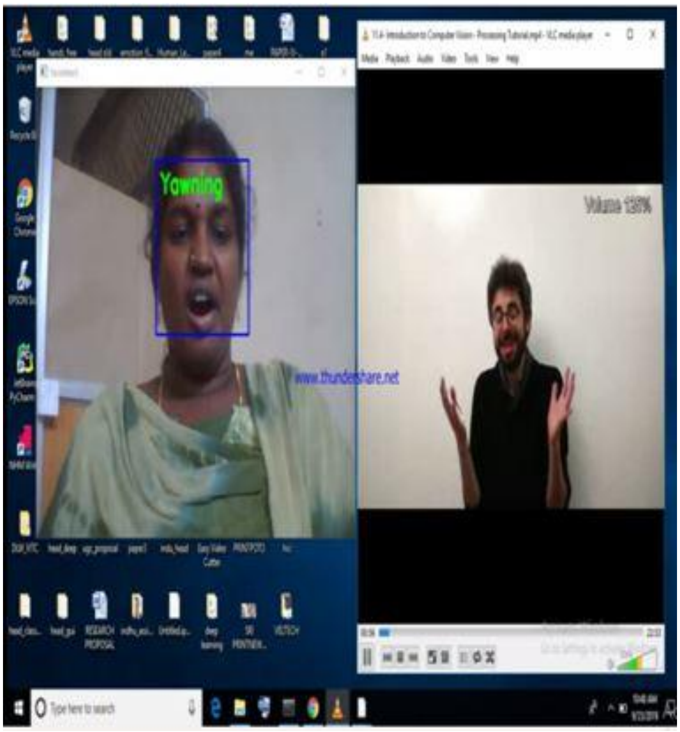

Fig.6d Yawning –Volume Increase

Fig.6. Interfacing with the E-Reading System

IX. CONCLUSION

In this proposed work, emotional interfacing for effective e-reading has been implemented. E-reading will be effective if the cognitive states of the learners are taken into account. The proposed system detects the faces of the learners using Haar cascaded classifier. The detected faces are pre-processed using Gabor filters. The Zernike moment features are extracted and modeled using decision trees and random forests. The models classify the emotions as happy, sleepy, yawn and irritate. These emotions are interfaced with the e-learning system and tested. The proposed system produces better performance with random forests than with decision trees. As a future work, more number of emotions can be taken into account for e-reading system and other e-applications.

REFERENCES

1. https://socialworkeducationbd.blogspot.com/2017/08/definition-of-emotion-and-its.html.
2. https://www.verywellmind.com/an-overview-of-the-types-of-emotions-4163976
3. Teh, C. and Chin, R.T., "On Image Analysis by the Methods of Moments", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 10, no. 4, pp. 496-513, 1988.
4. Lipscomb, J.S., "A Trainable Gesture Recognizer", Pattern Recognition, Elsevier, vol. 24, no.9, pp. 895-907, 1991.
5. Teague, M.R., "Image Analysis via the General Theory of Moments", Journal of the Optical Society of America, vol.70, no.8, pp. 920-930, 1980.
6. https://www.datacamp.com/community/tutorials/decision-tree-classification-python.
7. Ai Sun, Ying-Jian Li," Using facial expression to Detect Emotion in E-learning System: A Deep Learning Method", Springer International Publishing AG, 2017.

Retrieval Number: D8391118419/2019©BEIESP
DOI:10.35940/ijrte.D8391.118419
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

8. Aswin T S, Jijo Jose, Raghu G and Ram Mohana Reddy, “An E-learning System with Multiracial Emotion Recognition using Supervised Machine Learning”, IEEE Seventh International Conference on Technology for Education, 2015.
9. Kiavash Bahreini, Rob Nadolski and Wim Westera, “Towards Multimodal Emotion Recognition in E-learning Environment”, Taylor & Francis Interactive Learning Environments, 2014.
10. KrithikaL.B and Lakshmi Priya GG, “Student Emotion Recognition System (SERS) for E- learning Improvement on Learners Concentration Metric”, Published by Elsevier B.V, pp.1877-0509, 2016.
11. Salma Boumiza and Alexander Bekiarski,” Development of Model for Automatic Tutor in E-learning Environment based on Student Reactions Extraction using Facial Recognition”, IEEE 15-th International Conference on Electrical Machines, Drives and Power Systems (ELMA), 2017
12. .SL Happy, Anirban Dasgupt, Priyadarshi patnaik, “Automated Alertness and Eemotion Detection for Empathic Feedback during E-learning”, IEEE, 2013.

AUTHORS PROFILE

R. Indhumathi M.Sc., B.Ed., M.Phil, is a Research Scholar in Computer Science, Annamalai University, Tamil Nadu, India. She has been completed her Master degree in Computer Science in the year of 2012 at SASTRA University, Tanjore, Tamilnadu, India and completed her Master of Philosophy in Computer Science in the year of 2015 at Bharathidasan University, Trichy, Tamil Nadu, India. She also published four International journals. Her area of research is image processing and computer vision.

Dr. A. Geetha obtained her B.E degree in Computer Science and Engineering from Periyar Maniammai University, Thanjavur in the year 1994. She received her M.E degree in Computer Science from Regional Engineering College, Trichy in the year 1997. She has been working in Annamalai University, since 1999. She was awarded Ph. D in Computer Science and Engineering at Annamalai University in the year 2014. Presently she works as a Professor in the same department. She has published around 40 papers in national & international journals and conferences. She has coordinated a workshop in the field of speech and image processing. An international conference was also organized by her to her credit. Also she was one of the editors for International Journal of Applied Engineering Research which was published in the international conference. Two book chapters were published by her. She has won the best paper award in National Conference on Multimedia Signal Processing. She has delivered special lectures in several topics of her area, in workshops and staff development programme. Presently she is the coordinator of the Value Added Course, FEAT, and Annamalai University. She has visited and presented papers in international conferences conducted in Malaysia and Japan. Currently she is the member of Doctoral Committee in Annamalai University and Research Advisory Committee in Anna University. Her areas of research interest include pattern classification, image & video processing and human-computer interaction.