# Facial Expression Evaluation through Recurrent Neural Network

**Prof. Rupali Parte[1] Yogita Surve[2] Asiya Shaikh[3] Amruta Lohar[4] Pratiksha Lole[5]**
[1,2,3,4,5]Department of Computer Engineering
[1,2,3,4,5]JSCOE, Pune, India

*Abstract*— Facial expression is one of the most challenging and perplexing techniques ever attempted in the image processing paradigm. Humans express the majority of their emotions through their facial expressions, which can be used for other purposes like as identifying a person's mood. One of the most potent types of nonverbal communication is facial expression. The facial expression can be quite effective in precisely assessing an individual's mood. Humans can quickly identify facial expressions, while computers find it harder to do so. The recognition of mood via facial expression is a significant problem in computer vision. The facial expression is also extremely tough and subtle, which has shown to be the most difficult aspect of detection. Using machine learning approaches, this research study develops an excellent facial expression detection approach for mood diagnosis and Beverage Temperature detection. To achieve mood identification and Beverage Temperature detection, the described methodology employs Recurrent Neural Networks and Decision Tree. The accuracy of the methodology has been confirmed by the experimental results.

***Keywords:*** $YC_BC_R$ Model, Recurrent Neural Network, Decision Tree, Mood Identification

## I. INTRODUCTION

Facial expression is the result of a complicated interplay between muscles and muscle fibers in the human face. Various medical words and other associated techniques can be used to explain this physically. However, it is far more complicated and is largely used by people for communication. Language and facial expressions are used extensively in human communication and discussion. The majority of facial expressions and gestures offer a great deal of information about what the person is saying. Therefore, it is an essential component of communication and social dynamics, which have a significant effect on a person's overall well-being.

Humans are extremely sociable animals, which distinguishes us from the other species that inhabit this planet. Humans have complex relationships that are extremely beneficial to the mental and physical health of individuals as well as groups of persons. Humans are exceptionally skilled at detecting human faces thanks to their brains' built-in ability to detect faces quickly. Faces are made up of the bulk of human emotions that are employed for successful and insightful communication and engagement.

Facial expressions may be quite beneficial in identifying the conversation's intent and meaning on their own. Electronic systems are unable to distinguish a face or a facial image from any other image; hence this technique is not as developed. This is because the computer treats the image as a collection of pixel values rather than recognizing the image's objects. This is a very different mental process from how humans analyze and develop knowledge about the outside world and other people.

Therefore, there is a need for a system that can assess the expression as well as what it transmits based on facial traits. Many facial expressions in humans are linked to what the individual wishes to communicate or how they are feeling at the time. This serves as a marker for other people to recognize and interact with this person by utilizing their facial expressions as a cue for the emotions they are experiencing. Therefore, facial expressions must be discovered and employed efficiently for the aim of assessing an individual's mood with startling precision.

Multiple forms of studies have been conducted on this methodology for the objective of extracting an individual's mood through the usage of face photographs. For the sake of applying our methodology, these approaches were examined for their function in image processing and machine learning applications. This examination of various methodologies was crucial in the development of our methodology, which will be broadened in the future.

This research article dedicates section 2 for the evaluation of the past works under the name literature survey. And the implemented technique is broadly described under the section Implemented methodology which is numbered as 3. Section 3 discusses the obtained results from the experimental process. And finally, section 5 concludes this paper along with the scope for future enhancements.

## II. LITERATURE SURVEY

M.Azimi proposes a facial recognition system to analyze facial mood expressions and the emotional condition of the user. [1] Facial mood expressions can be classified into seven emotions angry, disgusted, fearful faces, happy, normal, sad, and surprised ones. The emotional communication of humans is directly attached to facial expressions. On face biometric recognition system's used three main reasons the samples are under the same illumination, everything remains constant unless facial expression and same ethnic and age group of users.

J.Inthiam expresses emotional assessment is a critical human social skill. Thus proposed paper implements Hidden Markov Model (HMM) there are two hidden states one is positive mood and negative mood and the second is recognized by certain bodily and facial expressions. In contemporary society robot technology is playing a major role. [2] Thus the proposed paper implements a convolution neural network (CNN) for recognizing emotion from the face. Thus proposed paper results be flexible in a more human-oriented way.

T.Yang presents a database of facial expressions that attached to the emotional stimuli with bipolar disorder (BD) patients and unipolar depression (UD). [3] For creating facial action unit (AU) profiles they have used a support vector machine to train the database where the feature extraction features are stored. Thus to predict the mood disorder where the facial expressions feature in the CHIMEI database generated by the AU profiles. Thus the proposed framework

All rights reserved by www.ijsrd.com 

implements a mood identification module for assisting doctors with diagnosis.

M.Han creates emotional expressions to combine robot emotion, mood, and personality. In the proposed model they have implemented by keeping a five-factor model created by the psychologists for the personality is programmed. [4] To extract the effectiveness of the proposed technique by using the artificial face simulator. The proposed mood state transition technique is used for the user's emotional changes frequently. They have also implemented a 2-D emotional framework to group the robot emotion and mood

M.Azimi states facial changes can affect the trend of the genuine score for both males and females because of the hypothesis tests. In the proposed paper the author implements numerical experiments by using PICS database two states such as Verilook and python lib to check the gender dependency for the expression intense. [5] The authors try to explain that facial expressions can affect the reliability of the system to show intense expressions as compared to men. Thus the proposed module shows an effective module.

K.Huang expresses facial expressions and speech responses based on unipolar disorder (UD) and bipolar disorder (BD). [6] By using the speech and facial expression depending recognition of mood disorder and emotion recognition. Thus in the proposed paper the developing the mood database is dependent on the eliciting emotional videos. This mood database can be used for accessing the relationship between the elicited facial expressions and the speech responses.

Q.Hong describes Mood disorder as a mental illness in which the individual can suffer from the loss the balance of the brain and it may lead to emotional dislocation. They express mood disorder has been a diagnosis in a portion of bipolar disorder patients. [7] Thus by observing macroscopic fluctuations mood disorders are investigated by facial expressions. To characterize mood disorder interval-based they have implemented using an artificial neural network. Based on the product rule and the geometric average is calculated among each emotion.

M.Su narrates based on facial expressions and temporal variations the mood of the individual is calculated. Thus in the proposed paper, they have implements cell-coupled LSTM. Thus in the proposed paper, they have used one-time diagnosis elicited from patients for six emotion-eliciting videos. Thus now global prevalence is given to mental illness such as Mood Disorder. Thus the proposed methodology is used to upgrade the database bias and it is an effective framework.

Z.Shah there is a lot of research made in the field of computer vision and there is a lot of advancement made in image processing such as pattern recognition and automatic detection for the classification of various facial expressions.[9] In one-to-one interaction, the most powerful communication is done by the human face. In communication messages, 55% of communication is made by facial expressions. They have implemented Region of Interest ROI for a better recognition rate.

D.Chiang develops an application for information and communication technology (ICT) which plays a music list based on emotion. To ease emotion music plays a major role to regenerate the brain cells. The application will automatically play the music based on the mood. [10] Thus the author of the proposed paper implemented the machine learning algorithm for emotion recognition and the emotion is predicted by the deep learning model. Thus the proposed module is known as effective Real-time Facial Expression Recognition.

Q.Hong describes if any individual suffering from a mental illness can affect the mood for a long period. [11] Based on facial expression fluctuations for mood disorder classification. Thus the proposed method is implemented based on AUs for observing the temporal transformation of muscles. Based on macroscopic and microscopic expressions for achieving classification accuracy of 63.9% and 66.7%.

A. Raheel narrates to facilitate humans by using electroencephalography (EEG) to obtain the brain signals. To recognize their facial expressions brain signals of individuals supervising different systems. [12] Thus the proposed system can be used for the paralyzed patients to know the mood and mental state of the individual. According to the author, there are thirteen different statistical features for expression as a smile, looking up, looking down, eye wink, and eye blink by using a K-nearest neighbor classifier with the combination of the Naïve Bayes, support vector machine.

S. Bazrafkan[13] expresses to provide a suitable reaction in different conditions by determining the mood identification by using facial expressions and facing the challenges such as pattern recognition and machine learning technologies. To solve these problems and difficulties they have used Deep Neural Networks (DNN) in facial expression classification. To get the appearance-based methods by using extraction, feature selection, and classification methods thus to examine the facial expression

## III. IMPLEMENTED SYSTEM FOR FACIAL

### *A. Expression Evaluation*

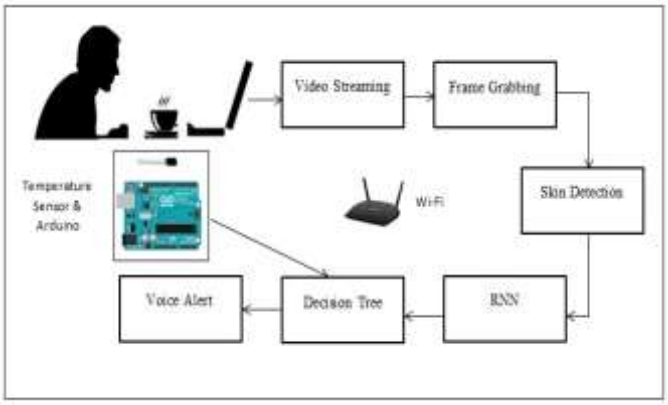


Fig. 1: Facial Expression evaluation System overview

The implemented methodology for facial expression detection for emotion identification has been illustrated in the figure 1 given above. The implemented methodology has been realized through the use of the steps given below.

1) Step 1: Video streaming and Frame Grabbing – The first step of the presented approach is to provide the system with the video containing the face of the person who is sitting to sip beverage. This is achieved through the use of an inbuilt web cam or an external camera.

All rights reserved by *www.ijsrd.com* 

This input video stream is provided to this step through the use of the OpenCV library which is an open source implementation that allows the java code to interface with the image capturing device for frame grabbing. The java code attains the facial frames of the video and effectively stores it in the jpg format in a particular location on the development machine.

The frames that are grabbed by the java code and stored in a particular location need to be provided to the python code for the purpose of further processing. This is done by the python code by scanning the particular location for the presence of the image. Once the image is encountered, it is read by the python code and utilized for the next step of the approach.

2) Step 2: Skin Detection –The java code provides the grabbed image through OpenCV library, to the python which performs the first logical step in the implemented methodology. This step is one of the most important parts of the implemented methodology that allows for an effective detection of the region of interest in the input image that is a person's face. The skin detection is the region of interest have been achieved by subjecting the image for haar cascade assessment which results in an XML file containing the cascade properties.

The position of the face in the image is extracted from the Haar features achieved previously in the input image. The image object is created by extraction of the facial coordinates to achieve the identification of the face in the image provided as an input. The YCbCr color model is being used to determine the color of the skin.

The component of the pixels in the image consists of values of Red, Green and Blue components. These values of Red, Green and Blue have been used to determine the color of the pixels and identify if the skin color is being detected. For this purpose, the values of the pixel are evaluated through the use of YCbCr color model. The red Chroma component and the blue Chroma component are extricated from the pixels through the use of the equations 1 and 2 given below.

$$Cb=-0.169 * R - 0.332 * G + 0.500 * B + 128 \quad (1)$$

$$Cr=0.500 * R - 0.419 * G - 0.081 * B + 128 \quad (2)$$

The blue Chroma component and the red Chroma component are being attained for the particular pixel is then put through the system for the purpose of classification of the values. For the skin detection purpose, the range of values between 137 and 177 are considered for the blue Chroma component, whereas the value of red Chroma component value between 127 and 77 are considered as the skin pixels. The skin pixels once detected, is colored white to identify the pixels.

The rest of the pixels which are non-skin are turned into black. This procedure is performed iteratively for all the pixels in the image. The resultant image is called a binary image, containing only two colors, white for the skin and black for the non-skin pixels. The process is elaborated in the algorithm 1 given below.

ALGORITHM 1: Skin Detection through $YC_BC_R$ model
//Input : Input image $I_{IMG}$
//Output: Skin Detected image $S_{IMG}$
// function: skindetection($I_{IMG}$)
1: Start
2: SIMG = ∅ , count=0
3: for i = 0 to size of breadth of $I_{IMG}$
4: for j=0 to size of Height of $I_{IMG}$
5: col = $I_{IMG}$ [i,j] ( PIX)
6: R= col[0]
7: G=col[1]
8: B=col[2]
9: Cb = -0.169 * R - 0.332 * G + 0.500 * B +128)
10: Cr = 0.500 * R - 0.419 * G - 0.081 * B + 128
11: if (Cr > 137 && Cr < 177), then
12: if (Cb < 127 && Cb > 77), then
13: t = Cb + 0.6 *Cr;
14: if (t > 190 && t < 215),then
16: SIMG [i,j] ( PIX )=[255,255,255]
17: end if,else
18: $S_{IMG}$ [i,j] ( PIX )=[0,0,0]
19: end if,else
20: $S_{IMG}$ [i,j] ( PIX )=[0,0,0]
21: end if,else
22: $S_{IMG}$ [i,j] ( PIX )=[0,0,0]
23: end if
24: endfor
25: endfor
26: return $S_{IMG}$
27: stop

3) Step 3: Recurrent Neural Network – This is one of the most essential steps of the presented approach for facial expression detection. This approach utilizes the Recurrent Neural Networks for the purpose of enabling an effective detection of the facial expression. The Recurrent Neural Network has been trained on a dataset containing various facial expressions. This dataset is downloaded from the URL - https://drive.google.com/file/d/1X60B-uR3NtqPd4oosdotpbDgy8KOfUdr/view. This leads to the Recurrent Neural Network providing an .h5 model file after training.

The dataset provided as an input consists of two different directories, training and testing. Both of these directories consists of a wide variety of directories, dedicated to particular expressions, namely, angry, disgusted, fearful, and happy, surprised, sad, and neutral. The RNN model is trained on these images effectively though the reduction in the image size to 48 x 48 and achieving an improvement in the approach through the use of 64 batch size and being run for 50 epochs. The Keras and TensorFlow libraries for python have been highly useful in design and deployment of the RNN model which utilizes the layers as given below in the figure 2.

All rights reserved by *www.ijsrd.com* 

| Layer | Activation |
|---|---|
| CONV 2D 32 X 3 X 3 | Relu |
| CONV 2D 64 X 3 X 3 | Relu |
| MaxPooling2D 2 X 2 | |
| Dropout 0.25 | |
| CONV 2D 128 X 3 X 3 | Relu |
| MaxPooling2D 2 X 2 | |
| CONV 2D 128 X 3 X 3 | Relu |
| MaxPooling2D 2 X 2 | |
| Dropout 0.25 | |
| Flatten | |
| Dense 1024 | Relu |
| Dropout 0.25 | |
| Dense 7 | Softmax |
| Adam Optimizer | |

Fig. 2: RNN network Architecture

The implemented RNN model stored trained data in .h5 file by executing for 50 epochs. The above architecture is being used to develop this model which is trained effectively. The resultant RNN model is then tested using the test images provided by the first step of the methodology.

4) Step 4: Decision Tree –The output obtained from the previous step through the execution of the .h5 RNN model file needs to be classified effectively through the use of Decision Tree approach. The captured image is subjected to this system for the emotion identification through the use of the if-then rules of the decision tree approach. The mood of the individual is effectively identified through the facial expressions; this has been then used to play the respective music that is stored according to the detected mood based directory. The decision tree also alerts the user on the beverage temperature if it varies drastically with the pre-determined threshold, that is, if it is too low or too high for consumption.

The alert message can be seen in the below mentioned figure.


Fig. 3: Alert Message for user

## IV. RESULTS AND DISCUSSIONS

The presented approach for mood detection through the use of facial expression identification through the use of RNN and Decision tree approaches has been attained through the use of a combination python and java programming language. As there are two different languages, two IDE's namely, NetBeans and Spyder have been utilized. For the development purposes, an Intel Core i5 processor along with 6GB of RAM and 1 TB of Hard disk is being used.

The facial images are extracted from the video stream captured using the inbuilt web camera or an external camera. The main model used for the purpose of facial expression analysis is used to determine the mood of the person. The evaluation of the performance is essential for the purpose of determining if the approach was implemented properly or not.

The real-time detected emotions can be seen in the below mentioned figure.


Fig. 4: Facial Expression Detection

– The evaluation process is detailed in the section below.

### *A. Performance Evaluation through Root Mean Square Approach*

The error achieved by the implemented methodology through the use of Recurrent Neural Networks has been measured for the error achieved through effective experimentation. The error attained by the system is one of the most effective techniques for the purpose of achieving the accuracy of the implemented methodology for mood identification of the user.

The Root Mean Square Error or RMSE approach is highly useful paradigm that allows for the effective calculation of the error achieved by the presented methodology. The error of the technique for the determination of the facial expression through the use of Recurrent Neural Networks indicates the accuracy of the execution of the approach. The use of RMSE allows the extraction of the error achieved between variables that are continuous and correlated with one another. The variables in this methodology for this purpose are the inaccurate detection of the mood and accurate detection of the mood. These values are measured and the equation 3 is used to calculate the error.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(x_{1,i} - x_{2,i})^2}{n}} \quad \_ (3)$$

Where,

$\sum$ - Summation

$(x_1 - x_2)^2$ - Differences Squared for the summation in between the expected No. of mood identifications and the obtained No of mood identifications

All rights reserved by *www.ijsrd.com* 

n - Number of Trails

These two variables are calculated by execution of the system on 10 different users executing 10 trials each with the various facial expressions. The results of these trials are listed in the table 1 given below.

| User No. | Number of expected mood Identifications | Obtained no. Of mood Identifications | MSE |
|---|---|---|---|
| 1 | 10 | 10 | 0 |
| 2 | 10 | 8 | 4 |
| 3 | 10 | 7 | 9 |
| 4 | 10 | 8 | 4 |
| 5 | 10 | 6 | 16 |
| 6 | 10 | 5 | 25 |
| 7 | 10 | 7 | 9 |
| 8 | 10 | 6 | 16 |
| 9 | 10 | 8 | 4 |
| 10 | 10 | 9 | 1 |

Table 1: Mean Square Error measurement

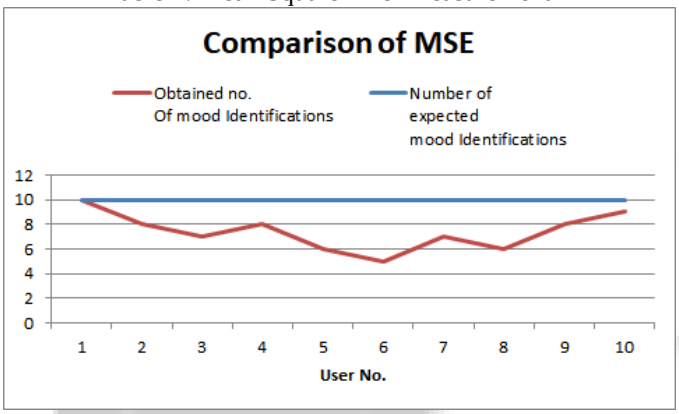


Fig. 5: Comparison of MSE in between Expected No of mood identifications V/s obtained No of mood identifications

The results of the evaluation are achieved through experimental evaluation. These tabulated values are used to construct the graph given in the figure 3 above. The graphical representation displays the miniscule amount of error achieved by the facial expression detection approach. This system has achieved exceptional results for the execution which can be attributed to the accurate implementation of the Recurrent Neural Networks and the Decision Tree implementations. This has been the reason for the very low error achieved by the implemented methodology. The values of MSE and RMSE achieved by the system are 8.8 and 2.96 respectively. This is a highly acceptable result that has been achieved for the mood identification approach through facial expression.

## V. Conclusion and Future Scope

The presented approach for the purpose of utilizing the facial expression for mood identification thorough the use of Recurrent Neural Networks and Decision Tree has been elaborated in this research article. The prescribed approach uses a live feedthrough a webcam to attain the frames consisting of the individual's face. These frames are effectively extracted and used for evaluation of the facial expression after effective preprocessing. The preprocessed frames are used to facial region identification through the Haar Features extracted into an XML file. These frames are then supplied to the next step for the purpose of skin detection. The YCbCr color model is used for the skin detection. This process generates a masked image that contains only two colors, black and white. The detected skin pixel is colored as white and the other pixels of the image are converted to black which leads to a binary image. This mask image is then used as an input to the RNN module to perform the facial expression detection. The results achieved through the model and the temperature sensor values are classified by the Decision Tree to achieve the accurate mood identification which is then used to play the respective song. The experimental evaluation has validated the accuracy of the mood detection and facial expression approach.

The future research directions for this facial expression detection approach can be developed in the form of a mobile application for better reach.

## References

[1] M. Azimi, "Effects of Facial Mood Expressions on Face Biometric Recognition System's Reliability," 2018 1st International Conference on Advanced Research in Engineering Sciences (ARES), 2018, pp. 1-5, doi: 10.1109/ARESX.2018.8723292.

[2] J. Inthiam, E. Hayashi, W. Jitviriya and A. Mowshowitz, "Mood estimation for human-robot interaction based on facial and bodily expression using a Hidden Markov Model," 2019 IEEE/SICE International Symposium on System Integration (SII), 2019, pp. 352-356, doi: 10.1109/SII.2019.8700422.

[3] T. Yang, C. Wu, M. Su and C. Chang, "Detection of mood disorder using modulation spectrum of facial action unit profiles," 2016 International Conference on Orange Technologies (ICOT), 2016, pp. 5-8, doi: 10.1109/ICOT.2016.8278966.

[4] M. Han, C. Lin and K. Song, "Robotic Emotional Expression Generation Based on Mood Transition and Personality Model," in IEEE Transactions on Cybernetics, vol. 43, no. 4, pp. 1290-1303, Aug. 2013, doi: 10.1109/TSMCB.2012.2228851.

[5] M. Azimi and A. Pacut, "The effect of gender-specific facial expressions on face recognition system's reliability," 2018 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR), 2018, pp. 1-4, doi: 10.1109/AQTR.2018.8402705.

[6] K. Huang, C. Wu, Y. Kuo, H. Yen, F. Jang and Y. Chiu, "Data collection of elicited facial expressions and speech responses for mood disorder detection," 2015 International Conference on Orange Technologies (ICOT), 2015, pp. 42-45, doi: 10.1109/ICOT.2015.7498502.

[7] Q. Hong, C. Wu, M. Su and K. Huang, "Exploring Macroscopic Fluctuation of Facial Expression for Mood Disorder Classification," 2018 First Asian Conference on Affective Computing and Intelligent Interaction (ACII

All rights reserved by *www.ijsrd.com* 

Asia), 2018, pp. 1-6, doi: 10.1109/ACIIAsia.2018.8470337.

[8] M. -H. Su, C. -H. Wu, K. -Y. Huang and T. -H. Yang, "Cell-Coupled Long Short-Term Memory With L -Skip Fusion Mechanism for Mood Disorder Detection Through Elicited Audiovisual Features," in IEEE Transactions on Neural Networks and Learning Systems, vol. 31, no. 1, pp. 124-135, Jan. 2020, doi: 10.1109/TNNLS.2019.2899884.

[9] Z. H. Shah and V. Kaushik, "Performance analysis of canny edge detection for illumination invariant Facial Expression recognition," 2015 International Conference on Industrial Instrumentation and Control (ICIC), 2015, pp. 584-589, doi: 10.1109/IIC.2015.7150809.

[10] D. Chiang, J. Yang, Z. Huang and F. Lai, "Music Response Based on Real-time Facial Expression Recognition," 2019 29th International Telecommunication Networks and Applications Conference (ITNAC), 2019, pp. 1-3, doi: 10.1109/ITNAC46935.2019.9078004.

[11] Q. Hong, C. Wu, M. Su and C. Chang, "Exploring Macroscopic and Microscopic Fluctuations of Elicited Facial Expressions for Mood Disorder Classification," in IEEE Transactions on Affective Computing, doi: 10.1109/TAFFC.2019.2909873.

[12] A. Raheel, M. Majid and S. M. Anwar, "Facial Expression Recognition based on Electroencephalography," 2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), 2019, pp. 1-5, doi: 10.1109/ICOMET.2019.8673408.

[13] S. Bazrafkan, T. Nedelcu, P. Filipczuk and P. Corcoran, "Deep learning for facial expression recognition: A step closer to a smartphone that knows your moods," 2017 IEEE International Conference on Consumer Electronics (ICCE), 2017, pp. 217-220, doi: 10.1109/ICCE.2017.7889290.

All rights reserved by *www.ijsrd.com* 