LIVE FACIAL EXPRESSION TO EMOJI CONVERSION

Anurag Kumar[1], Chandana S R[2], Rishabh Mishra[3], Sreshta Reddy N[4], Prof. Prathima M G[5]
Bangalore Institute of Technology,
anuragkumar219.ak@gmail.com[1], chandanastu729@gmail.com[2], rishabhmishra879@gmail.com[3],
sreshtareddy001@gmail.com[4], prathimamg@bit-bangalore.edu.in[5]

Abstract—The system proposes an application inconverting live facial expression to emoji using Deep Learning. It includes facial expressionand emotion detection. It provides a way to classify human facial expressions to filter and map them to corresponding emojis.In this paper, deep learning is used to recognise human emotions through facial expressions.Here, in order to experiment with and train a deep convolutional network, the Kaggle's FER2013 dataset has been used. This work has been successfully implemented in real time system.

Keywords—Desktop Application,Face Detection, Feature Extraction, Emotion Recognition, Haar Cascade, CNN.

I. INTRODUCTION

In today's digital communication, Emojis have become an essential part.They are used to express the emotions of a person through text in a way that is not possible with just words.Emotion is one of the fundamental expressions of human beings. Facial expression recognition is a hot topic in computer vision. As a direct expression of human emotions, facial expressions are a form of nonverbal communication.

Human facial expressions convey a lot of information visually rather than articulately and plays a crucial role in the area of human-machine interaction. Automatic facial expression recognition system has many applications including, but not limited to human behaviour understanding, detection of mental disorders, and synthetic human expressions. Recognition of facial expression by computer with high recognition rate is still a challenging task. Two popular methods utilized mostly in the literature for the automatic FER systems are based on geometry and appearance. Facial Expression Recognition usually performed in four-stages consisting of pre-processing, face detection, feature extraction, and expression classification. In this project we applied various deep learning methods (convolutional neural networks) to identify the key seven human emotions: anger, disgust, fear, happiness, sadness, surprise and neutrality. The best technique that helps the people in reducing the time, from manually checking/searching the appropriate emoji which expresses the non-verbal communication.

Keeping all these points in mind, we propose the use of Deep Learning techniques such as Face detection, emotion recognition, feature extraction which are used in the application to convert live facial expression to emoji where the user can get the mapped emoji based on the input image

The purpose of making this application is to provide an option, to send the emoji/avatar according to the live expression given by the user. Today many social media applications have options to send emoji, but they cannot recognize the real time face expressions of the users. Therefore, with this project we are helping the users to do the same.As a result, it provides a solution for people to reduce the time, from manually checking/searching the appropriate emoji.

II.MOTIVATION AND BACKGROUND STUDY

We were inspired to take on this project because of the following factors. The existing systems provide the options to send a variety of emoji's but they don't provide any option to send emoji's by capturing the real time emotions of the user. Moreover, the user sometimes struggles a bit to find an appropriate emoji from a huge list of emoji. To provide ease to the

users, we thought of building a desktop app that can overcome these difficulties of existing system.

Several works of facial expressions recognition were studied. As mentioned in the paper [1], Proposed system is a Convolutional Neural Network (CNN) based LeNet architecture for facial expression recognition. First of all, 3 datasets (JAFFE, KDEF and our custom dataset) were merged. Then LeNet architecture was trained for emotion states classification. In this study, the accuracy achieved is 96.43% and validation accuracy is 91.81% for classification of 7 different emotions through facial expressions.

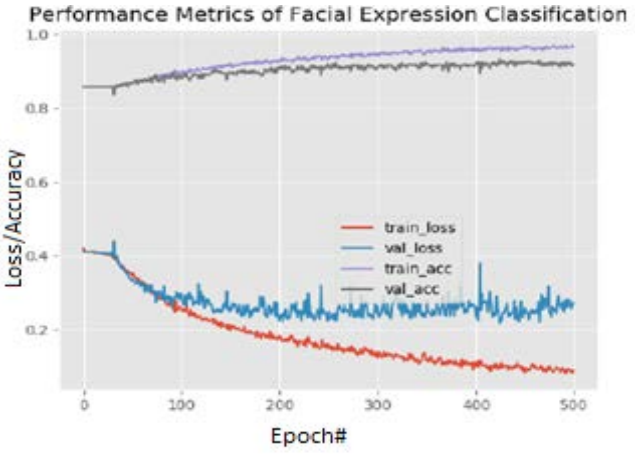


Fig1: Performance of the proposed system

As per the paper [2], To avoid the complex process of explicit feature extraction in traditional facial expression recognition, a face expression recognition method based on a convolutional neural network (CNN) and an image edge detection is proposed. Firstly, the facial expression image is normalized, and the edge of each layer of the image is extracted in the convolution process. The extracted edge information is superimposed on each feature image to preserve the edge structure information of the texture image. Then, the dimensionality reduction of the extracted implicit features is processed by the maximum pooling method. Finally, the expression of the test sample image is classified and recognized by using a SoftMax classier. To verify the robustness of this method for facial expression recognition under a complex background, a simulation experiment is designed by scientifically mixing the Fer-2013 facial expression database with the LFW data set. The experimental results show that the proposed algorithm can achieve an average recognition rate of 88.56% with fewer iterations, and the training speed on the training set is about 1.5 times faster than that on the contrast algorithm.

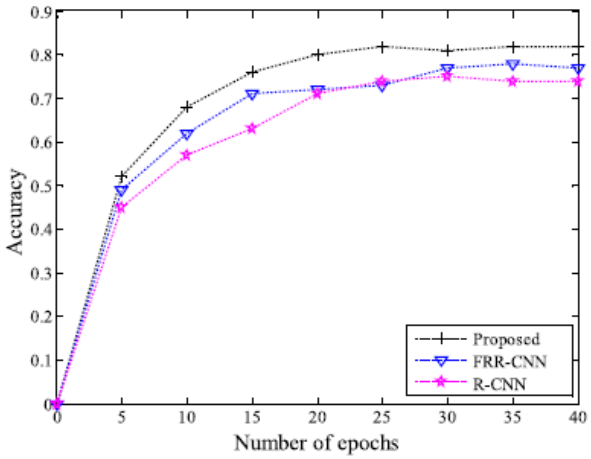


Fig 2: Expression recognition rate of different methods
in complex background

According to paper [3], The study is carried out for the facial expression recognition from facial expression dataset using Convolutional Neural Networks (CNN). In addition, intermediate outputs of CNN were also analysed. As a result, the obtained score for emotion recognition is a score of about 58%; two emotions (Happiness, Surprise) recognition score was about 70%. Study also confirmed that specific unit of intermediate layer have learned the feature about Happiness. This paper details these experiments and investigations regarding the influence of CNN learning from facial expression.

The article in [4] is an experiment on the FER2013 dataset, the purpose is to get the facial expression attributes from the facial image. Because the pictures in this dataset have low resolution, and some pictures have no faces at all. This reduces the accuracy of facial expression recognition. In this paper, proposed system is a robust improved model. In this model, they introduced attention mechanism and separable convolution to improve the extraction of image features, and use data argumentation techniques to enhance the generalization ability of the model. The model obtained 65.2% test set accuracy on the FER2013 dataset.

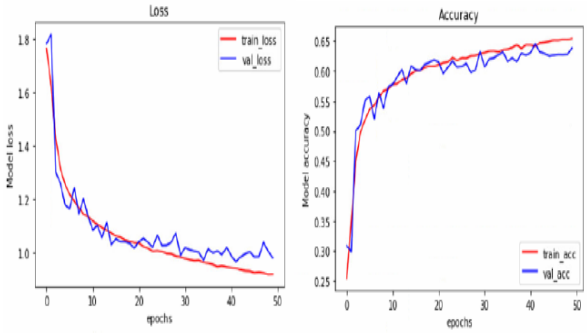


Fig 3: Model Loss and Accuracy curve

The paper [5] proposes a CNN for designing a real time classification system for face detection, gender classification and emotion recognition. Each of the three tasks will have a separate model trained for the specific problem of either face detection, gender classification or emotional classification. The three models will then be pipelined so get one common output for all the three tasks. The architecture of our model is designed so that it can give acceptable performance even on low end systems which lack powerful hardware.

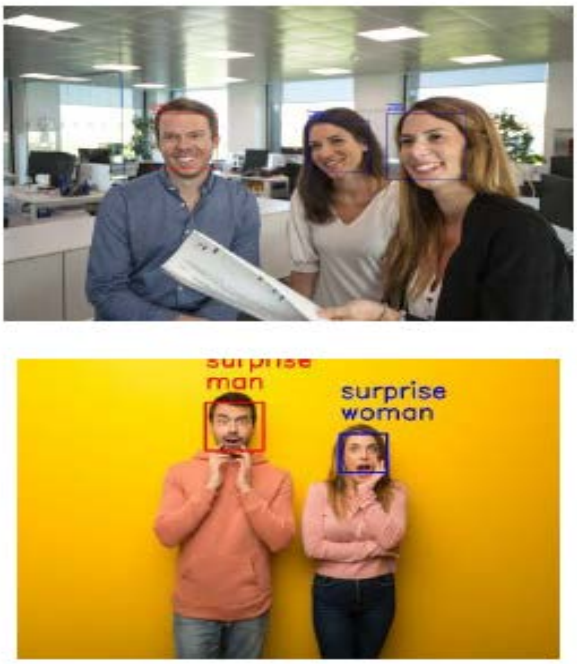


Fig 4: Results of the classification model.
Blue colour represents the assigned class woman and red the class man.

According to paper [6], the present FER (Facial Expression Recognition) system uses still images, this faces a complex problem in discriminating foreground from background cluster without motion information. The FER in video motion is implemented to overcomethis problem of existing systems.This paper brings an approach of Real-time Facial Expression Recognition using HAAR cascading classification for face detection followed by Convolutional Neural Networks (CNN) for classification of expressions.This model uses web-cam of the system and dynamically display the emotion in a text format.

With an accuracy of 58% on test data, we were successfully able to classify seven different human emotions: Angry, Disgust, Fear, Happy, Sad, Surprise, Neutral. Facial emotion recognition can be used in many real-time applications, like airport security, trading, patient monitoring, and others.

	Angry	Disgust	Fear	Happy	Sad	Surprise	Neutral
Angry	214	9	53	30	67	8	86
Disgust	10	24	9	2	6	0	5
Fear	45	2	208	29	89	45	78
Happy	24	0	40	696	37	18	80
Sad	65	3	83	56	285	10	151
Surprise	7	1	42	27	9	303	26
Neutral	45	2	68	65	88	8	331

Fig 6: The confusion matrix

III. SYSTEM DESCRIPTION

A. Architecture

The architecture is divided into capturing picture, image pre-processing, CNN training model and mapping emotion to emoji. It is designed in such a way that the live picture is captured through camera and then the image pre-processing takes place where the extracted live input image data is suppressed and features are extracted for further processing. The captured feature is trained, classified and the emotion is detected from the saved model. The detected emotion searches the corresponding emoji for the detected emotion from the stored emoji images. Finally, the output emoji is displayed and emoji is shared immediately when a person changes his/her emotion.

The module Face detection using Haar Cascade is defined as OpenCV Haar Cascade is used for face detection. It is done by capturing the face inside a rectangular box. Haar Cascade method is used to detect face in each frame of the webcam feed.

The trained neural network is the module which defines an emotion recognition system from facial expression is used for recognizing expressions from the facial images and

classifying them into one of the seven basic emotions. It is used to classify the input image into one of the expressions (angry, sad, happy, surprise, disgust, fear and neutral). The Keras library in Python makes it pretty simple to build a CNN. Computers see images using pixels. Pixels in images are usually related. It is used to extract features of the live input image provided by the user and map it to corresponding emoji.

The GUI module is the user interface module consisting of Tkinter is used for GUI. It provides various controls, such as buttons, slider, search field, labels and textboxes used ina GUIapplication. Tkinter is Python's de-facto standard GUI (Graphical User Interface) package. Streamlit – Open source python framework. Tkinter is a standard python interfaceto the Tk GUI toolkit shipped with python. Python with Tkinter is the fastest and easiest to create GUI applications.

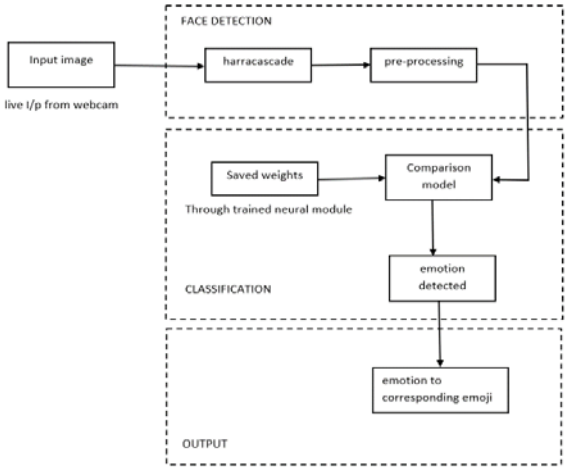


Fig 7: System Architecture

The system design is designed in such a way that the live picture is captured through camera andIt is done by capturing the face inside a rectangular box using Haar Cascade.

i. Haar Cascade method is used to detect face in each frame of the webcam feed.

ii. The region of the image containing the face is resized to 48x48 and is passed as input to the CNN.

iii. Then the image pre-processing takes place where the extracted live input image data is suppressed and features are extracted for further processing.

• The captured feature is trained, classified and sent to the comparison model where the emotion is detected from the saved model (i.e. already trained data is present) and compared with those of extracted features

• The detected emotion searches the corresponding emojifrom the stored emoji images.

• Finally, the output emoji is displayed and emoji is shared immediately when a personchanges his/her expression.

B. Activity Diagram

In the beginning, the users need to pass a simple authentication process. The authentication page displays a Drop-down menu that has three options. The first option helps the users to see brief information about the app. The second option provides an option to the users to register themselves on the app. The third option provides an option to the user to login into the app.

If the users have already registered then they can directly go for the login to start using the app. If the users haven't registered, they first need to go to the registration page and enter some basic details for their registration. If the users are able to pass this simple authentication process, they will get a button to start the application and by clicking that button the app will be started.

When the application starts, it will turn on the webcam of the device and user has to provide an emotion for the app to recognize it. On the captured image, face detection will happen by using the OpenCV Haar Cascade and some pre-processing will be applied. This pre-processed image will be sent to the trained CNN model. Model will classify the emotion into one of the seven categories which are happy, angry, surprised, neutral, disgust, sad, fear.

Based on the category of the emotion, app will display the corresponding emoji next to the video of the user. If the user changes his/her emotions then emoji will also change immediately. If the users want to quit the application, they can press quit button at the bottom of the UI.After pressing the quit button, the app will again ask the users if they really want to quit. If they say Yes, the app will be closed immediately and if they say No, the app will continue to recognize the expressions of the users.

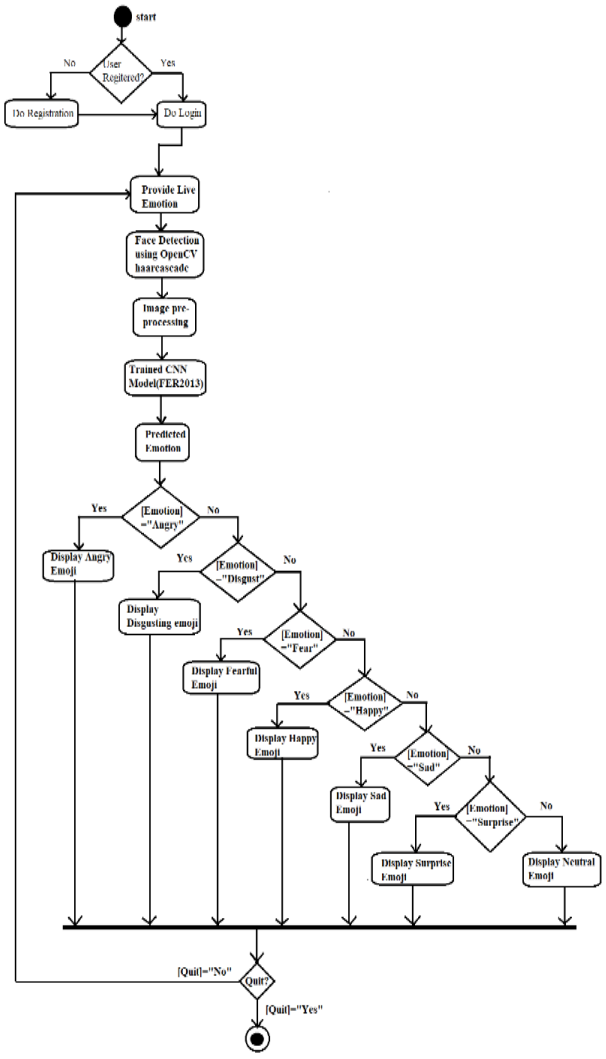


Fig. 8. Activity Diagram

IV. DATASET DESCRIPTION

FER2013 Dataset FER2013 is a well-studied dataset and has been used in ICML competitions and several research papers. It is one of the more challenging datasets with human-level accuracy only at 65±5% and the highest performing published works achieving 75.2% test accuracy. Easily downloadable on Kaggle, the dataset's 35,887 contained images are normalized to 48x48 pixels in grayscale. FER2013 is, however, not a balanced dataset, as it contains images of 7 facial expressions, with distributions of Angry (4,953), Disgust (547), Fear (5,121), Happy (8,989), Sad (6,077), Surprise (4,002), and Neutral (6,198).

IV. ALGORITHM

A. Algorithm for Pre-processing

Input: Full resolution of webcam
Output: a grayscale 48*48 picture
Step 1: Start.
Step 2: Receive the live expression.
Step 3: Received image is converted to grayscale.
Step 4: The received image which is converted to grayscale is rescaled using**ImageDataGenerator** class.
ImageDataGenerator(rescale=1./255) where Rescale 1./255is to transform every pixel value from range [0,255] -> [0,1].
Step 5: End.

B. Algorithm for Trained Neural Network

Input: Pre-processed image
Output: Weights
Step 1: Start
Step 2: Receive the pre-processed image.
Step 3: First layer will take input as image with relu activation function of 32 filters which enable model to detect boundary of the face.
Step 4: Second layer will take input as image with relu activation function of 64 filters which enable model to detect eyes, nose and mouth.
Step 5:MaxPooling2D() is added to reduce the spatial dimensionsof the output volume.
Step 6:Dropout(0.25) is added to maintain regularity in the modelin order to avoid overfitting.
Step 7:Add two more convolutional layers with relu activationFunction of 128 filters.
Step 8:Flatten() converts the pooled feature map to a single column that is passed to the fully connected layer.
Step 9:Two dense layers are added with relu and SoftMax activation functions

respectively, with resultant weights being seven.

Step 10:The SoftMax function presents itself as a probability for each emotion class. The model is able to show the detail probability composition of the emotions in the face.

Step 11: End.

C. Algorithm for Face Detection

Step 1: Start

Step 2: import cv2.

Step 3: if cv2 cannot pull the full resolution of webcam then print "cannot access camera" else capture the input image using read() in build function of cv2.videocapture and get flag value

Step 4: Cascade classifier is used to detect the face of the user by using parameter cv2.Cascade Classifier ('haarcascade_frontalface_default.xml')

Step 5: if flag is nonethen print “Face not detected " else

1) create a copy of the captured image

2) convert it into red, green and blue colorspace in such a way for enhancing thevisibility of the captured picture

3) map it to tkinter using PhotoImage function

4) display on the screen

Step 6: End

V. RESULTS

This section presents some results of the desktop application in which appropriate emoji is presented based on the live facial expression of the user. In our proposed model, the data is shuffled very time before testing and training dataset are split. This helps to make sure that same set of images are not used for training and testing images.

The image Fig. 9the main window of the application. The user can login and signup here.It is built using Streamlit library of python.After successful authentication, we get the option to launch the application.

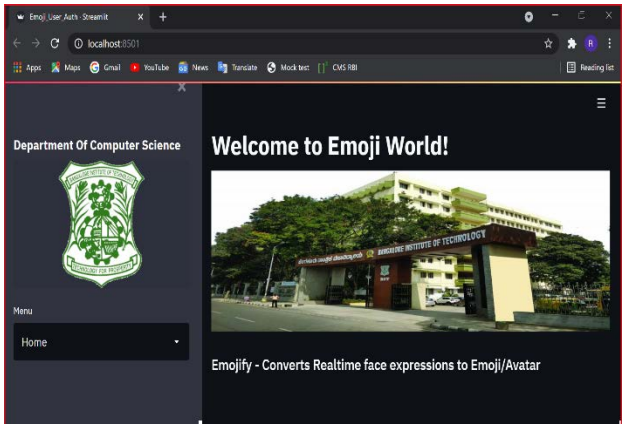
Authentication provides access control for systems by checking to see if a user's credentials match the credentials in a database of authorized users or in a data authentication server.

Fig. 10 shows the Signup page of the application where the user can register or signup in order to use the application.

Fig. 11shows the login page of the application where the user has to login after successful registration.

Fig. 12shows how CNN model is classifying the emotion correctlyand It is used to classify the input image into one of the expressions (angry, sad, happy, surprise, disgust, fear and neutral) and the figure shows for happy emotion.

In Fig.13shows the integrated face detection and CNN Model where the emotion is detected for the given input.



In Fig.14 GUI of the application is shown which is built using Tkinter library of python. Here, based on the live facial expression of user an appropriate emoji is generated. There is quit option, where the user can close the application by clicking on it.

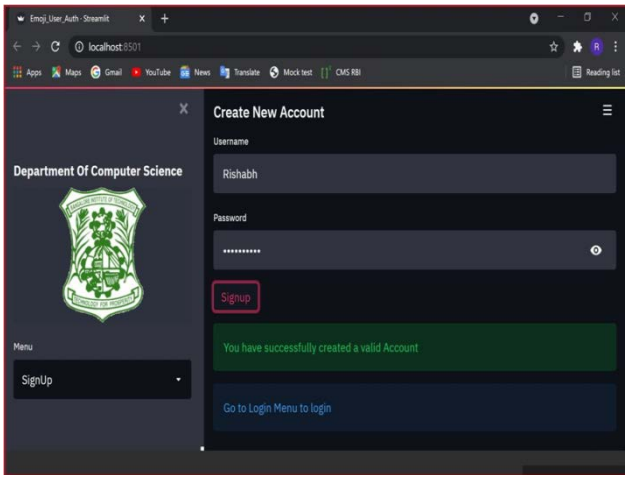


Fig. 10. Signup page

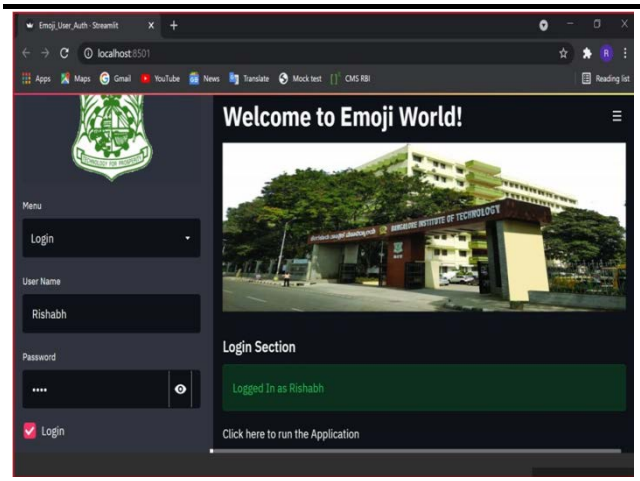


Fig. 11. Login page

Fig. 12. CNN Model classifying the happy emotion correctly

Fig. 13. Emotion of human face is detected using face detection and CNN Model

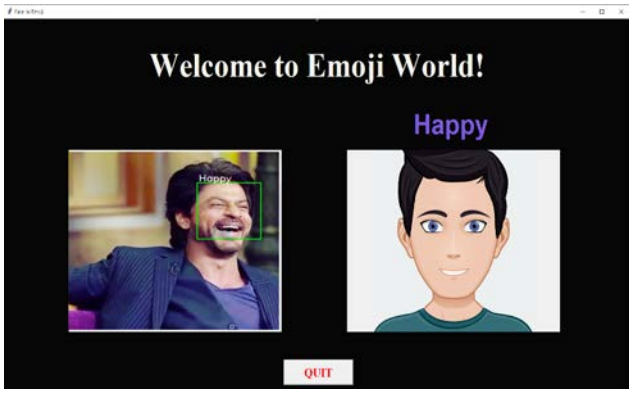


Fig. 14.GUI of the application

VI. APPLICATIONS

The proposed application can be used in social media applications. Users can choose the appropriate emoji based on his live expression. It can also be used in marketing activities. It may add extra level of relatability to the brand. This application can also be used for educational purposes. Students and teachers can be more expressive using the emoji.

VII. CONCLUSION

In the proposed system we found that the physiologicalcharacteristics of the human face with relevance to various expressions such as happiness, sadness, fear, anger, surprise and disgust are associated with geometrical structures which restored as base matching template for the recognition system.The behavioural aspect of this system relates the attitude behind different expressions as property base. The training set evaluates the expressional uniqueness of individual faces. In the future, AR emoji feature can be added to the existing application where AR(Augmented Reality) emoji is a feature which creates an animated version of yourself which can then be used to record your movements and facial expressions.

REFERENCES

[1] Facial expression recognition using Gabor filter and multi-layer artificial neural network, Authors –Kunika Verma, AjayaKhunteta, Year Published – 2017,IEEE

[2]Feature Acquisition from Facial Expression Image Using Convolutional Neural Networks, Authors- TaikiNishime, Satoshi Endo, Koji Yamada, Naruaki Toma, Yuhei Akamine, Year Published – 2018,IEEE

[3] A Face Emotion Recognition Method Using Convolutional Neural Network and Image Edge, Authors - Hongli Zhang, Alireza Jolfaei and MamounAlazab, Year Published – 2019,IEEE

[4] Facial Expression Recognition with Deep Learning, Authors - Amil Khanzada, Charles Bai, Ferhat TurkerCelepcikay, Year Published – 2019,IEEE

[5]Authors Real Time Facial Emotion Recognition using Deep Learning, Authors - Avigyan Sinha, RP Aneesh, Year Published – 2019,IEEE

[6] Facial Expression Recognition Based on a Hybrid Model Combining Deep and Shallow Features, Authors - Xiao Sun, Man Lv, Year Published - 2019

[7] Convolutional Neural Networks Architectures for Facial Expression Recognition Authors - George Porusniuc, Florin Leon, Radu Timofte, CasianMiron, Year Published - 2019

[8]Real Time Emotion Recognition from Facial Expressions Using CNN Architecture Authors - Mehmet Akif Ozdemir, BerkayElagoz, Reza Sadighzadeh, Aydin Akan, Year Published - 2019,IEEE

[9]Face Detection using Haar Cascades to Filter Selfie Face Image on Instagram, Authors - Adri Priadana, MuhammadHabibi, Year Published – 2019, IEEE

[10]Facial Expression Recognition Using Neural Networks, Authors - Yoshitha Gali, Kevin Sanghoi and Mrinalini Chava, Year Published – 2019,IEEE

[11]Emotion and gender classification in real-time, Authors - Utkarsh Agrawal, Anjnesh Sharma, K. Senthil Kumar, Year Published – 2019

[12] Face Recognition with Age, Gender and Emotion Estimations, Authors - Dongqin Gu, Miaomiao Zhu, Lifeng Zhang, Seiichi Serikawa, Year Published – 2020,IEEE

[13]Facial Expression Recognition Using Convolutional Neural Network, Authors - Weier Wan, Chenjie Yang, Yang Li, Year Published – 2020,IEEE

[14] A Robust Improved Network for Facial Expression Recognition, Authors-Hao Gao, Bo Ma, Year Published- 2020

[15]Real-Time Facial Emotion Recognition with Deep Convolutional Neural Network

Authors – Lakshmi Bhadana, P.V. Lakshmi, D. RamaKrishna,G. SuryaBharti, Y. VaibhavYear Published- 2020,IEEE

[16] Human Emotions Detection Using Hybrid CNN Approach, Authors - Nehmat Sandhu; Aksh Malhotra; Mandeep Kaur Bedi, Year Published – 2020,IEEE