CSE4076-Image and Video Analytics

**J Component – Project Report**

**FACIAL RECOGNITION & DETECTION IN REAL TIME SCENARIO**

***By***

***19MIA1006 Shivani Gokulram Naidu***

***19MIA1066 Madasu Deepika***

***19MIA1069 G.Harinisri***

***M.Tech CSE with Specialization in Business Analytics Submitted to***

***Dr.Maheswari S, SCOPE, VIT, Chennai***

***School of Computer Science and Engineering***



# ABSTRACT

## Face recognition has existed for a long time. A step further, human emotion can be roughly predicted from facial expressions and brain activity that have been recorded in video, electric signal (EEG), or image form.

## Human emotion recognition is urgently needed so that cutting-edge artificial intelligence systems can mimic and predict facial reactions. Making educated decisions about the determination of purpose, the advertising of offerings, or security-related threats might benefit from this. While recognising emotions from photos or video is a simple operation for the human eye, it is extremely difficult for machines to do and necessitates the use of numerous image processing techniques for feature extraction.

## For this task, a variety of machine learning techniques are appropriate. Machine learning algorithms must first be trained before being tested against an appropriate dataset for any detection or recognition.This research investigates feature extraction methods that would aid in the precise identification of human emotions.

**KEYWORDS:** DeepFace, Crowd Behaviour Analysis, Deep Learning, Keras, Computer Vision.

**INTRODUCTION**

Researchers have focused on this topic since being able to recognise one's expressionsaids in human-computer connection, corrects advertising campaigns, and culminates in augmented and enhanced human communication through modifying

emotional intelligence (EQ) in people.

The examination of human expression recognition can be done in a variety of

methods, including by looking at facial expressions, body posture, speech tones, etc.

We have emphasised the recognition of facial expressions in this research.

Numerous developments, including machine to human interaction and automatic

translation systems, are taking place in the burgeoning study field of facial

emotion recognition (FER).

Face acquisition, face detection, and face recognition are the three main steps of

face recognition, respectively. The process where the system needs to acquire a human

face image starts at the face acquisition step. Following the acquisition of the face is a

face detection procedure, which is crucial. It will now go through a few additional steps

like facial localization and feature extraction. There are various methods for locating

faces, and employing a segmentation method is one of them.

To interpret the input image into a reduced form for feature extraction, it commonly separates the frontal face and backdrop from the image. The process of feature extraction has grown to be essential to image processing, especially when conducting object detection. The system's final step is face recognition, after which it will classify the input and make a choice about how to interpret it.

Automatic video surveillance is getting more and more attention, including crowd behaviour analysis. This field seeks to understand how people behave in huge groups and to get useful information from recordings that contain large groups of people. For instance, monitoring the behaviour of huge numbers of spectators in a sports stadium or automatically analysing the motion flow of pedestrians at a congested pilgrimage site are essential for spotting potentially harmful circumstances before they turn into disasters.

**CHALLENGES**

Emotion recognition faces many of the same difficulties as object detection in video, including object identification, continuous detection, incomplete or unpredictable actions, etc.

**Data Augmentation:** ER solutions need a large amount of training data, just like any machine learning and deep learning algorithms. Videos with varied frame rates, viewpoints, backdrops, and people of diverse sexes, races, and nations must all be included in this data. In terms of racial and gender diversity, the majority of datasets fall short and only include a small number of emotional expressions.

**Face occlusion and lighting issues**: For motion recognition in video, occlusion caused by changes in stance is a frequent problem, especially when working with unprepared data. In an open space, varying lighting and contrast are also fairly common.

**Identifying features**: Faces are scanned by an emotion detection tool for features including the lips, chin, nose, eyes, and other facial structures. The distance between features, the size of the features, and the colour of the skin can sometimes make this detection difficult.

# LITERATURE SURVEY

# The papers that we have referred is given below:

# This format can help us to guide through all the papers we have studied and also their best findings from it in a brief way .

|  |  |  |  |
| --- | --- | --- | --- |
| **S.NO** | **TITLE** | **AUTHOR , YEAR OF PUBLICATION** | **FINDINGS** |
| 1 | **Modelling of Facial Images for Analysis of Recognition System[1]** | H, Assyakirin & Basah, Shafriza & Yazid, Haniza & S, Fathinul & S,Muhammad., **2021** | **They developed a synthetic model to represent facial images to be used as a platform for performance analysis of**  **facial recognition systems. The model includes 5 face types with the ability to vary all parameters that are affecting recognition performance –**  **measurement noise, face size and**  **face-background intensity differences. It results this model is important as it provide an avenue for performance analysis of facial recognition systems.** |
| 2 | **An Improved Artiﬁcial Neural**  **Network Design for Face Recognition utilizing Harmony Search Algorithm[2]** | **Maryam Mahmood Hussein, Ammar**  **Hussein Mutlag and Hussain Shareef**  **2019** | **Artiﬁcial neural network (ANN) has been suggested to achieve the face recognition process.This research**  **introduces an improved ANN design for face recognition using a meta-heuristic optimization algorithm.**  **The results revealed that the proposed system (HSA-ANN) achieved lower MSE**  **compare with the ANN. Furthermore, the HSA-ANN gives a better face recognition rate than traditional ANN.** |
| 3. | **Real-Time Video Emotion Recognition based on**  **Reinforcement Learning and**  **Domain Knowledge[3]** | **Ke Zhang, Yuanqing Li, Jingyu Wang, Member, IEEE, Erik Cambria, Fellow,**  **IEEE, Xuelong Li, Fellow, IEEE** | **Multimodal emotion recognition in**  **conversational videos (ERC) develops rapidly in recent years. Emotion recognition model for conversational videos based on reinforcement learning and domain knowledge (ERLDK) is**  **proposed in this paper. In ERLDK, the reinforcement learning algorithm is**  **introduced to conduct real-time ERC with the occurrence of conversations. The experimental results on datasets** |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  | **show that ERLDK achieves the**  **state-of-the-art results on weighted average and most of the speciﬁc emotion categories.** |
| 4. | **Facial Emotion Recognition in**  **Continuous Video[4]** | **Albert Cruz, Bir Bhanu and Ninad Thakoor Center for Research in**  **Intelligent Systems, University of**  **California, Riverside, CA, USA** | **This paper is about facial emotion recognition where a subject can freely express emotions, a derivative of features was more suitable than using**  **the features themselves. The derivative was estimated with histogram differencing and DoG features. A HMM fused the output of SVM matchers.**  **Cooccurrences were demonstrated to exist between feature derivatives and features, and their impact on**  **classiﬁcation was positively demonstrated. The proposed approach increased classiﬁcation results on the non-trivial AVEC2011 data set** |
| 5. | **Deep Learning of Human Emotion Recognition in Videos[5]** | **Yuqing Li Uppsala University** | **This research involves 3 models. The ﬁrst one is audio-SVM model. Audios will be used to extract audio features with OpenSMILE as feature extractor and classiﬁed with SVM. The second**  **one is CNN-LSTM model. This involves three steps: train the feature extractor, CNN model(Inception-ResNetv2 model); use CNN model to extract deep features from face images cropped from video**  **frames; and use LSTM as classiﬁer to**  **integrate the deep features and classify the emotion. The third one is video-C3D model. This model use face frames from videos as input and C3D model as both feature-extractor and classiﬁer.** |
| 6. | **Emotion AI,**  **Real-Time Emotion Detection using CNN[6]** | **Tanner Gilligan**  **M.S. Computer Science**  **Stanford University Baris Akis**  **B.S. Computer** | **CNN-based emotion detection model is proposed that utilizes**  **facial-detection software and cloud**  **computing to accomplish its task. The ﬁnal model resulted in accuracies**  **comparable to the state-of-the-art** |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **Science**  **Stanford University** | **papers in the ﬁeld. Exhibits more balanced accuracy results across the emotion spectrum. The proposed model still worked signiﬁcantly well with**  **non-actor subjects, especially for physically expressive emotions like sadness, happiness and surprise.** |
| 7. | **Emotion recognition using facial expressions[7]** | **Pawel Tarnowski, Marcin Kolodziej, Anderzej Majkowski, Remigiusz J Rak Warsaw University of Technology, Warsaw, Poland** | **Coeﬃcients describing elements of facial expressions, registered for six subjects, were used as features. The features have been calculated for**  **three-dimensional face model. The classiﬁcation of features were**  **performed using k-NN classiﬁer**  **and MLP neural network. The result was obtained for MLP classiﬁer and “natural” division of data for all users**  **(subject-independent). Experiments were carried out under the same**  **conditions and at a ﬁxed position of a user. The classiﬁcation accuracy was inﬂuenced by the way users play**  **speciﬁc facial expressions. In real**  **conditions the classiﬁcation accuracy can be affected by many additional**  **factors.** |

A technique for identifying and verifying a person's identification using their face is facial recognition. It is beneficial to be able to recognize people in pictures, movies, and in person. Using biometric software, it maps a person's facial traits and records them as a faceprint.

Emotion detection, on the other hand, recognizes and examines human emotions. By obtaining details about someone's emotional state, it is possible to determine how they are feeling.

Identifying a person's mood depends heavily on their emotions. Typically, there are six basic emotions: joy, sorrow, rage, fear, surprise, disgust, and contempt. A system raising the alert in the case of a person who is feared might be helpful. Recognizing the emotions also plays a crucial function in camera surveillance to catch the culprits. Systems for recognising emotions can be utilized as a component of many different applications, such as different camera surveillance systems.

Due to the advent of new methodologies, crowd analysis has steadily improved throughout time. Due to their ability to discriminate between different inputs and their effective functional extraction, deep learning techniques are being used more and more in a variety of

applications. Due to these restrictions, many techniques utilised in conventional crowd analysis weren't appropriate for contemporary surveillance. The operational conditions of surveillance equipment and the extreme uncertainty and dynamicity in crowd mobility patterns are typical characteristics of contemporary surveillance systems.

The application of existing tools for the monitoring and analysis of the dense crowd can be complicated by this characteristic's diversity. To address the issue in the new environment, where computer vision is increasingly needed to monitor and analyse numerous people from video feeds of the security cameras in real time, researchers in crowd analysis need develop fresh methodologies.

Estimating both the crowd's diversity and the density distribution throughout the full collection region are included in this. Identifying places above protection can assist in sending out earlier alerts and help avoid crowd encirclement. Estimating the size of the crowds also aids in evaluating the event's significance as well as its infrastructure and logistics.

**PROPOSED STATEMENT**

In computer vision and artificial intelligence, identifying and detecting human emotion is a significant difficulty. Human communication is heavily influenced by emotions. Emotions are primarily used in communication. The major objective of our research is to create a reliable system that can recognize and detect human emotion from live broadcast. A few feelings, such as anger, sadness, joy, surprise, fear, disgust, and neutrality, are shared by all people. This has applications in fields of security, surveillance, robotics etc.

The implementation of the whole system in software will be completed around the network training and expression recognition using pandas, numpy, matplotlib, sklearn etc to complete the design of the application layer interface. The overall system has been developed using tools like python with support of Jupyter notebook and Visual Studio Code.

But here I want to highlight three important libraries that have been used

• OpenCV: This is one of the library which is widely used in processing images, particularly real time images.

• Keras: This is one of the library which is used to code deep learning models. In its back-end it uses Tensorflow.

• DeepFace: This is one of the library which is used to face recognition and facial attribute analysis library for Python.

**ALGORITHM**

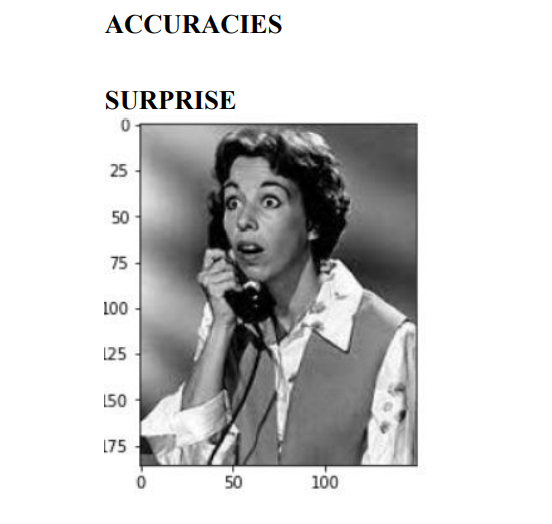
**IMPLEMENTATION**

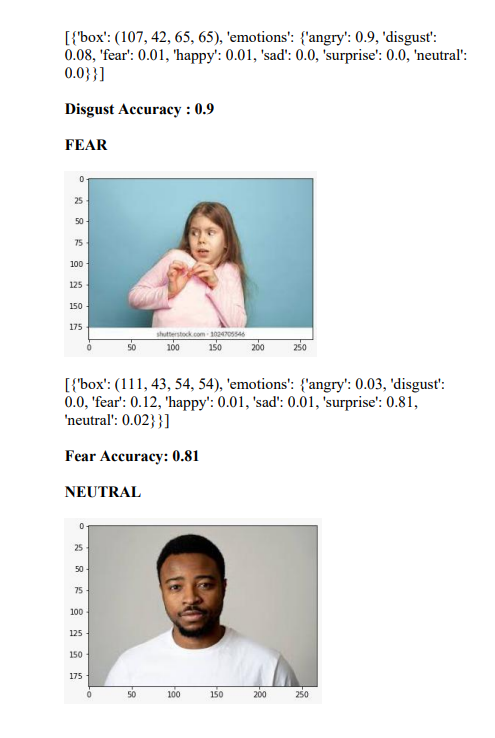
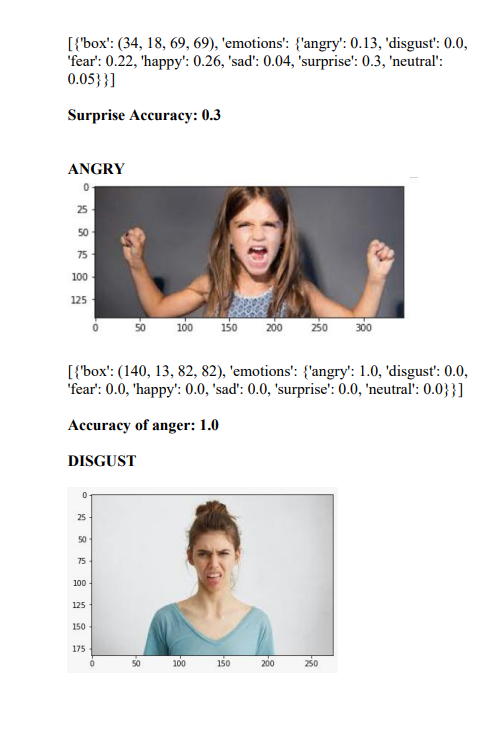
**IMAGE**

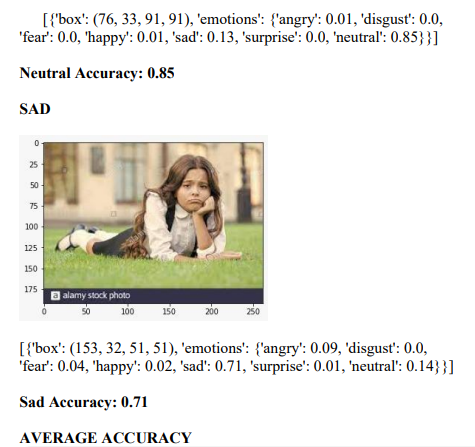
**Using Keras**

**This is the output that we have got for the images showing emotions**

* **Surprise**
* **Neutral**
* **Happy**
* **Sad**
* **Angry**
* **Disgust**
* **Fear:**

****

****

****

**Average accuracy is found to be 0.7742**

**Using DeepFace**

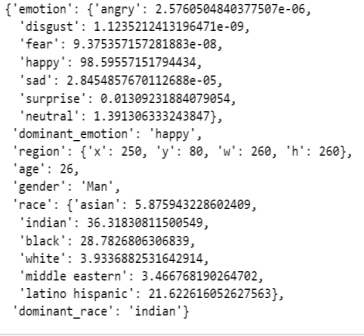
Here the predictions that are made for the picture are that the individual is 26 years of age, is of Indian origin and from his emotions we find that he is happy.

**algorithm:**

models = ["VGG-Face", "Facenet", "OpenFace", "DeepFace", "DeepID", "Dlib", "ArcFace"]

#face verification verification = DeepFace.verify("img1.jpg", "img2.jpg", model\_name = models[1])

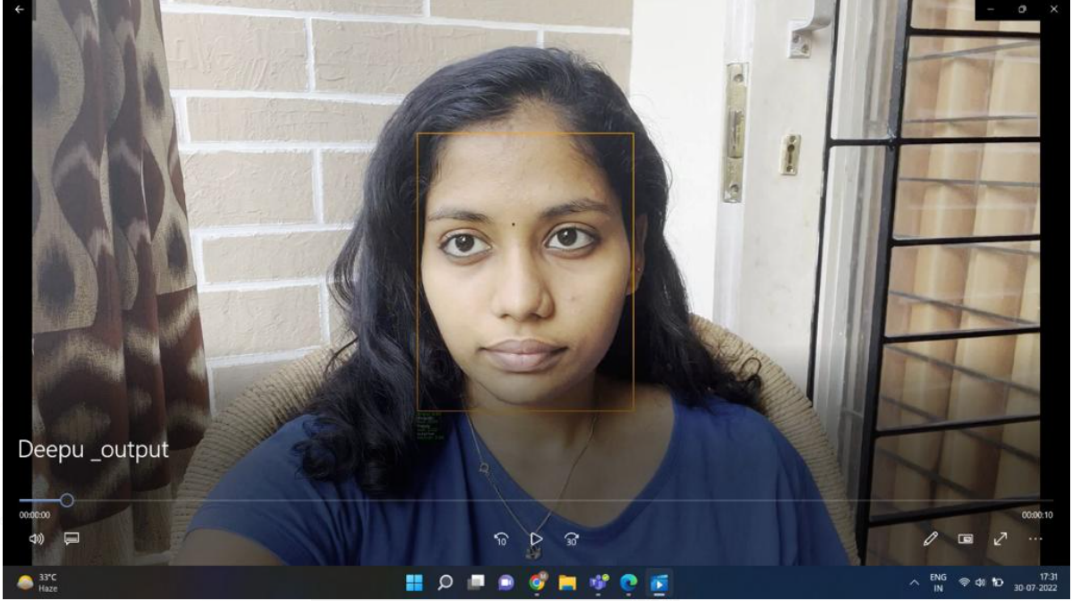
#face recognition recognition = DeepFace.find(img\_path = "img.jpg", db\_path = “C:/facial\_db", model\_name = models[1])

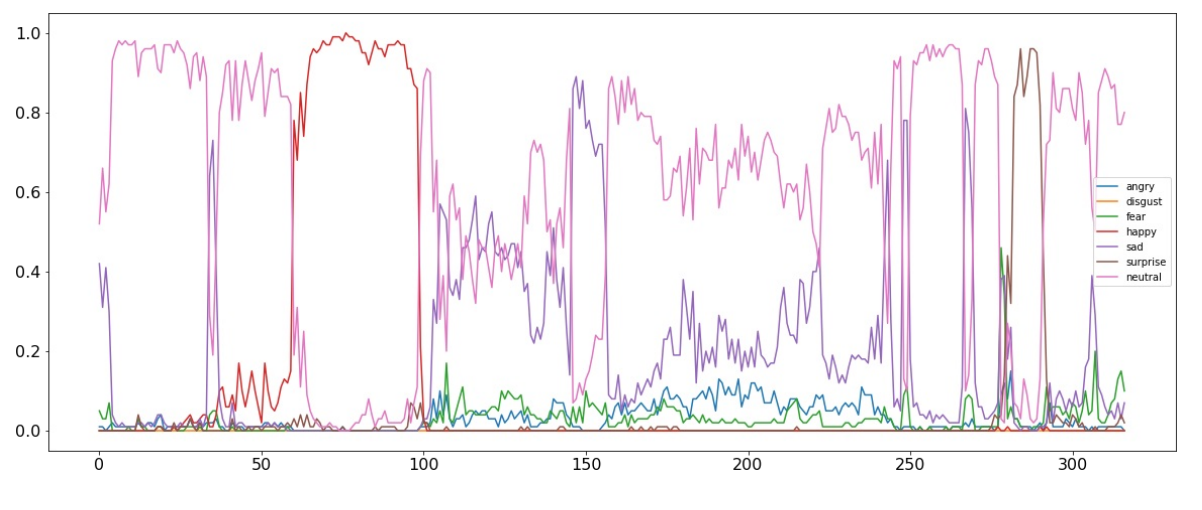
****

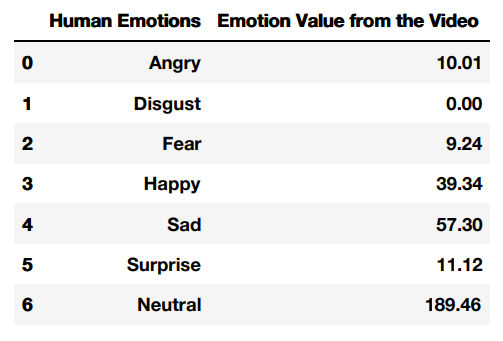
**VIDEO**

In this a video is recorded and then the emotion is recognized for it. the 7 basic emotions are presented in a graph for the video with a different color for each emotion and then the emotion value is provided in a table for the given video.

Here the emotion with highest frequency of 189.46 is neutral. Which was correctly predicted.

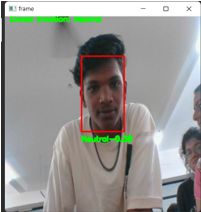
****

****

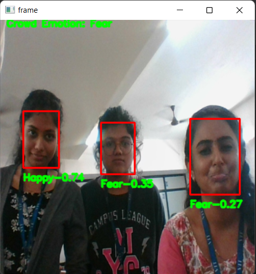
****

From the above graph and table , we are able to find neutral to have the highest emotional value and disgust to be the lowest.

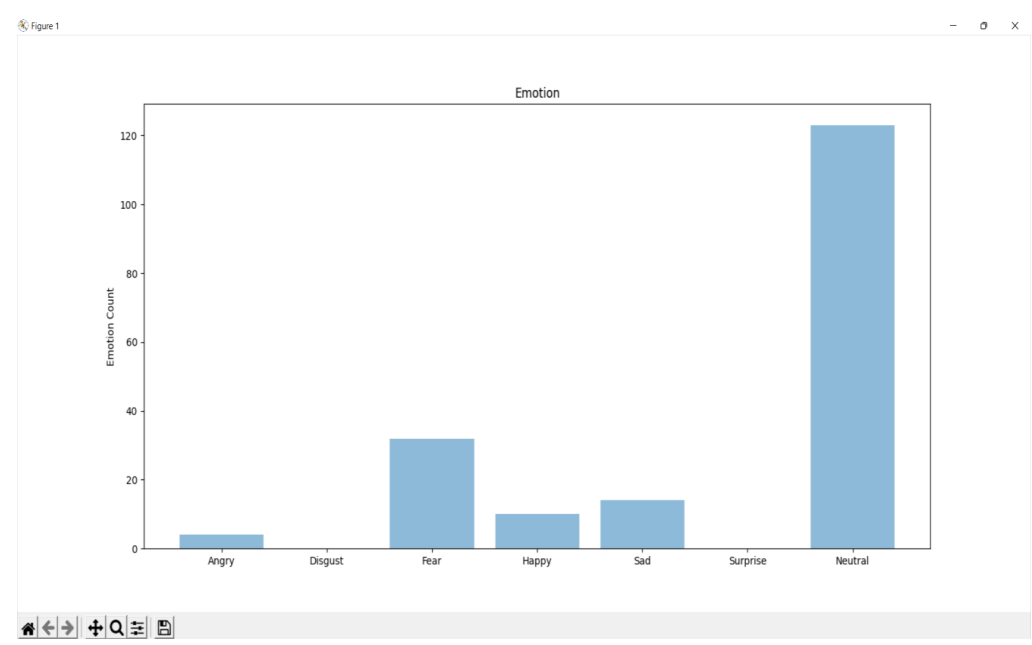
**CROWD BEHAVIOUR ANALYSIS**

****

Here a single individual is present and the emotion that is recognized here is neutral with a value of 0.38 which happens to be the correct prediction.



Here three individuals are present the 1st showing the emotion happy at 0.74. Individual 2 shows the emotion fear at 0.35 and last individual also shows fear at 0.27. Hence from this we gather that the crowd emotion is fear since 2 individuals show fear.

****

**CONCLUSION**

Emotions play an important role in human sphere of life. This project was comprehended to assess and gather the significant and efficient emotion recognition techniques developed. Today, we have a wide range of methods and considerable amount of work has been done in the fields of facial emotion recognition. The 7 elementary emotions which humans display are sadness, surprise, disgust, happiness, fear ,neutral and anger. Facial Emotion Recognition (FER) is largely done through two categories of methods, namely, feature and model based techniques.Here we are doing the feature technique.

In the carried out experiments, for 7 emotional states, we achieved a very good classification accuracy of emotions - 96% for the images. Experiments were carried out using the FER2013 dataset uploading it in google colaboratory and also using visual code for real time recognition . Certainly, the classification accuracy was influenced by the way users play specific facial expressions. In real conditions the classification accuracy can be affected by many additional factors. When you feel real emotions, facial expressions can vary greatly - may be exposed to a greater or lesser extent.

**REFERENCES**

1. H, Assyakirin & Basah, Shafriza & Yazid, Haniza & S, Fathinul & S, Muhammad. (2021). Modelling of Facial Images for Analysis of Recognition System. Journal of Physics: Conference Series. 2107. 012041.

10.1088/1742-6596/2107/1/012041.

1. Hussein, Maryam & Mutlag, Ammar & Shareef, Hussain. (2020). An Improved Artificial Neural Network Design for Face Recognition utilizing Harmony Search

Algorithm. IOP Conference Series: Materials Science and Engineering.

1. Lin, Shih-Lin & Wu, J. (2021). Face recognition unlocking uses principal component analysis to control the vehicle door system. Journal of Physics: Conference Series. 2020. 012028. 10.1088/1742-6596/2020/1/012028.
2. Deep Learning of Human Emotion Recognition in Videos Yuqing Li Uppsala University
3. Du Tran, Lubomir Bourdev, Rob Fergus, Lorenzo Torresani, and Manohar Paluri. Learning spatiotemporal features with 3d convolutional networks. In Proceedings of the IEEE international conference on computer vision, pages 4489–4497, 2015.
4. Yaser Yacoob and Larry S Davis. Recognizing human facial expressions from

long image sequences using optical flow. IEEE Transactions on pattern analysis and machine intelligence, 18(6):636–642, 1996.

1. Tanner Gilligan and Baris Akis (2017).Emotion AI, Real-Time Emotion Detection using CNN.
2. Tarnowski, Paweł & Kołodziej, Marcin & Majkowski, Andrzej & Rak, Remigiusz. (2017). Emotion recognition using facial expressions. Procedia Computer

Science. 108. 1175-1184. 10.1016/j.procs.2017.05.025.