IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

# AN ANALYSIS OF THE INFLUENCE OF SUBJECTIVE EMOTION ON COGNITIVE STYLE USING DATA MINING TECHNIQUES

A. Oviya[1]

M. Phil (Computer Science)

Pioneer College of Arts and Science

Coimbatore

K.Manickaraj[2]

Assistant Professor in Computer Science,

Pioneer College of Arts and Science,

Coimbatore

**ABSTRACT**

Facial expressions communicate non-verbal clues, which are crucial for interpersonal communication. The method of determining a person's emotional state is done by the Cognitive Emotion AI system. The major goal of our research is to create a reliable system that can recognise and detect human emotion in live broadcast. A few feelings, such as anger, sadness, joy, surprise, fear, disgust, and neutrality, are shared by all people. The facial detection process is completed by extracting a face's Haar Cascade features using the Viola Jones algorithm, and the emotion is then confirmed and recognized using artificial intelligence techniques. The system will receive an image or frame as input, and upon receiving it, the model will preprocess it and choose its features before predicting its emotional state.

**Keywords :** Facial Expressions, Deep Learning, Image Pre-processing

## 1. INTRODUCTION

Humans frequently express their emotions through their facial expressions. One of the most effective, direct, and natural ways for people to express their feelings and intentions is through their faces. Because people often have to hide their feelings, as in hospitals or when they're sick, better understanding of other people's emotions would facilitate more effective conversation. The science of the twenty-first century is artificial intelligence. A machine having artificial intelligence (AI) is one that can "think or act like a human being or rationally." The enormous volume of data can now be processed by machines in real time, and they can react accordingly. But while having high IQs (Intelligence Quotients), these machines lacked emotional intelligence (EI) or emotional quotient (EQ) at all times (EQ). There is concern that we may lose personal connection and communication as technology develops and the world becomes more and more virtual; but, what if our devices could take the place of those interactions? Whether we can create machines that can recognise human emotions is the era's central conundrum.

Emotional research and analysis are not new fields of study. The Expression of the Emotions in Man and Animals, written by Charles Darwin in 1872, stated that "Facial displays of emotion are universal, not learnt differently in each culture." The topic's foundational research was conducted by psychologist Paul Ekman. He was the first to categorise feelings.

1. **Joy (happiness)** is represented by tightening the eyelids and lifting the corners of the mouth into an evident smile.

2. **Surprise** is represented by the jaw dropping slightly, the brows arching, and the eyes opening wide and showing more white.

3. **Sadness** is shown by lowering the corners of the mouth, lowering the brows to the inner corners, and drooping eyelids.

4. **Anger** is represented by lowering the brows, firmly pressing the lips together, and bulging the eyes.

5. Fear is shown by the upper eyelids lifting, the eyes opening, and the lips stretching horizontally.

6. Disgust is indicated by the upper lip lifting, the nose bridge wrinkling, and the cheeks lifting.

7. Contempt is represented by the upper lip's half tightening, and frequently the head is thrown back somewhat.


**Figure 1.2: Example of CK+ subjects belonging to the same class of disgust but with various head shapes and cultures.**

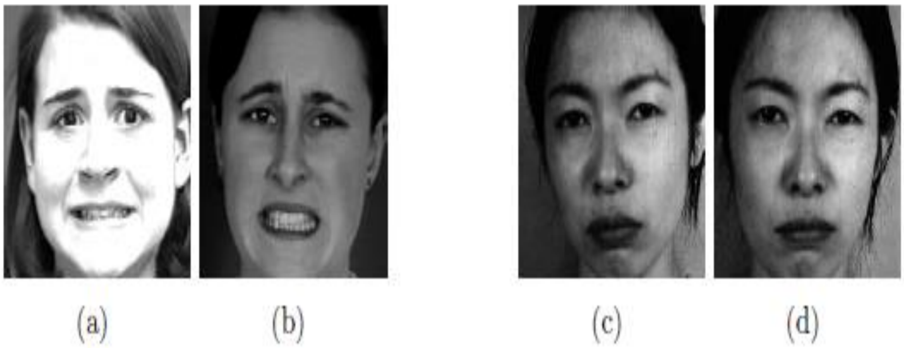


**Figure 1.3: Examples of different lighting conditions for subjects in the same dataset are shown in (a) and (b) in the CK+ and (c) and (d) in the JAFFE.**


**Figure 1.4: An illustration of how closely two kinds of face emotions resemble one another**

## 2 RELATED WORKS

The expression of emotions by a person through the movement of their facial muscles is referred to as facial expression. It reveals details regarding that person's mental health. A person experiences emotion as a mental condition. It is his internal reaction to what is happening to him outside of himself. In many circumstances, a person's facial appearance reveals information about their mental state. Numerous fields, including robotics, art, and lie detectors, use facial expression analysis [1].

For an intelligent agent to interact with robots and humans in machine-human collaboration and robot-robot interaction, improvement in the capability of facial expression detection is necessary [2].

The progress made in research on this subject is noteworthy given that this field of study has been going on for years [3] [4]. Since they have a lower recognition rate for emotions like disgust and fear, fluctuation in recognition rates between classes is one of the problems for the majority of the research [5] [6].

This study aims to create a face expression detection system that can categorise an image into seven different emotional groups. Additionally, it will be addressed how to increase validation accuracy in comparison to other existing systems and keep each class's recognition rate commendable and equal or almost equal. Convolutional neural networks have been used in a few studies to recognise facial expressions, however inconsistent recognition rates between classes are one of the problems for the majority of these studies, since they have lower recognition rates for disgust and fear [5] [6] [7].

For many years, researchers have been studying face expressions. However, there was always room for advancement in every study project. Because of this, there are numerous opportunities related to this subject. The primary objective of their research in [5] is to increase the correctness of the FER2013 dataset. They used a convolutional neural network as the mechanism for categorising the seven fundamental emotions in their suggested model.

In order to increase the accuracy of biometric applications, convolutional neural networks have been successfully used in this study. However, because they were unable to maintain an equal or nearly equal recognition rate for each class, there is variability in the rate of each class's recognition. The recognition rate for identifying disgust and fear is just 45% and 41%, respectively, despite the fact that total accuracy has been attained at 91.12%.

Prior to this, in [6], researchers had created a face expression identification system using staged photos in a static setting. The RAF-DB dataset, however, which consists of photographs of people of various ages and stances in dynamic environments, was introduced by [6]. To categorise seven fundamental emotions, they used the deep locality-preserving CNN approach. The RAF-DB and CK+ datasets were used to train their suggested model. Despite having a 95.78% accuracy rate, the recognition rates for disgust and fear are just 62.16% and 51.25%, respectively.

They described this as an outgrowth of their earlier work in [7]. They used deep convolutional neural networks, a blend of convolutional neural networks and deep residual blocks, to classify the six primary emotions. They used the Extended Cohn Kanade (CK+) and Japanese Female Facial Expression datasets to train their model (JAFFE). Their research has been deemed successful when it performs better than state-of-the-art methodologies and has higher accuracy than competing models. 95.24% accuracy has been attained. Their system is biassed toward those datasets because it is based on just two of them. Additionally, they were unable to identify emotions from an image with faces that are geometrically distorted.

They compared two different approaches for extracting facial features in [8]. The Gabor-wavelet coefficient fetch method is one, while geometric fiducial point positions are the other. They have demonstrated through this comparison that the gaborwavelet coefficient fetch method outperforms the other one. They were able to determine that five to seven hidden layers are necessary to attain a greater recognition rate. 90.1% accuracy has been attained. They acknowledged having a lower recognition rate in the class on fear even though they haven't presented individual class recognition rates. Using the Spyder IDE, the application was created in the Python programming language.

For this experiment, you'll need the libraries Keras, Tensorflow [21], numpy, PIL, OpenCV, and matplotlib. While Keras assisted the system by offering built-in features like activation functions, optimizers, layers, etc., Tensorflow served as the system's backend. For image preprocessing tasks like face detection (using the Cascade Classifier), grayscale conversion, and picture normalisation, OpenCV was mostly employed. Keras API did data augmentation. Confusion matrix has been produced using Matplotlib. Using HTML, CSS, and JavaScript, a Graphical User Interface (GUI) to display outputs of the proposed model has been created.

The stored Neural Network model is accessible via a Python Flask server and accepts input in the form of a picture from the user's local device. The class of the image is displayed after execution when the user hits the "Predict" button. Fig. 5 includes a few examples of graphical user interface screenshots. To demonstrate that the model can reliably anticipate unseen images, real-time photos have been given as system input.

It should be noted that the system pre-processes images in the same manner as when the model has been trained whenever a user inputs an image. This means that anytime a user inputs an image of any size, the system initially turns it to a 48*48 sized image. Afterward, using Cascade Classifier [17],

## 2.1 EXISTING SYSTEM

Through social media posts, depression was also early on foreseen. In addition to modeling depression, the audio-text technique can also be utilized to diagnose anxiety and mood disorders by scanning patient facial expressions and using cross validation for more accurate results issues brought on by a small amount of data points. Participants are unwilling to answer too many questions in research that rely on questionnaires or interviews, which is another issue. As was previously mentioned, persons who experience anxiety, sadness, or stress are frequently reluctant to talk to close friends, family members, or healthcare professionals and prefer to express their concerns in an anonymous manner. People who experience anxiety, despair, or stress are frequently reluctant to confide in close friends, family members, or medical professionals and prefer to express their concerns in an anonymous manner. Therefore, the easiest way to gather their comments is over the internet. Due to this, the data set used in this study was obtained from online surveys that were completed by various Participants, both male and female, aged 20 to 60, as well as employed and jobless individuals.

# 3. PROPOSED SYSTEM

## 3.1 OVERALL ARCHITECTURE

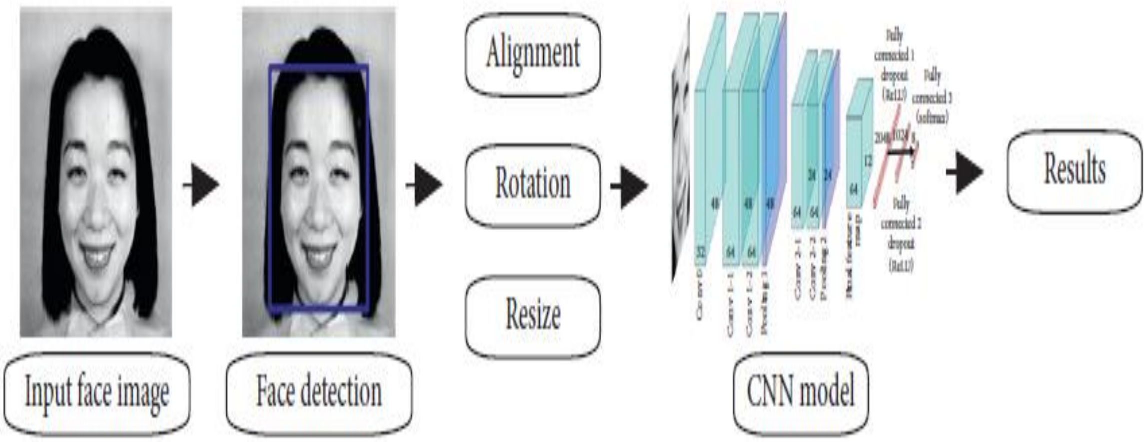


**Figure 3.1 Proposed Overall Architecture**

An effective CNN-based architecture is employed for the training and testing of face mask recognition after face detection. For training purposes, there are already a number of architectures that are available; these were already covered in the literature review portion. In this study, we proposed a unique architecture to determine whether or not someone is wearing a face mask. Face mask identification is carried out via facial feature analysis in the suggested work. The proposed lightweight CNN model consists of four convolutions, one fully connected layer, and the output layer. Thresholding operations are carried out after each convolution layer using a nonlinear ReLU activation function. Neurons with values less than zero are removed from the network by the ReLU activation function, whereas neurons with positive values remain. The proposed model uses two maximum pooling layers to reduce dimensionality, which has an impact on how long the network must train. Due to the model and data's simplicity, overfitting can become an issue during model training. We utilised the dropout and batch normalisation mechanisms to lessen the overfitting issue. The number of classes that the network can recognise is represented by the output neurons in the final fully connected layer. Finally, the softmax classifier is used to categorise the input into the appropriate class, such as face mask detected and no face mask detected.

The first convolution layer's input image in the suggested design is 128 by 128 and has RGB channels with 32 distinct filters (kernels). These filters have a 3-by-3-by-3 size with a 1-pixel stride. After the normalising and pooling layers, the output of the first convolution layer serves as the input for the second convolution layer. The second convolutional layer consists of 64 kernels with a size of 3 3. Without any batch normalisation or subsampling in between, the remaining convolutional layers are connected. There are 128 kernels in the third convolutional layer, with each kernel having a size of 3 3 3. The final convolutional layer includes 256 kernels that are each three by three by three in size. Totaling 128 neurons, the completely connected layers have four layers. A softmax layer receives the output of the fully connected layer and provides a distribution over the two classes of labels, mask detected and mask not detected. Figure 3 includes descriptions of the future layers as well as every aspect of the proposed system in its entirety.

## 3.2 IMAGE ACQUISITION (DATA SET DESCRIPTION)

**Figure 3.2 Image Data Set Descriptions**

## 3.3.IMAGE PRE-PROCESSING

The pixel value for our input images spans from 0 to 255 because they are grayscale photos (256 Levels of grey). 0 stands for the colour white, and 255 for the colour black. The neural network computation of appropriate weights and bias will be too challenging with such large input values. Therefore, these image pixels with values between -1 and 1 need to be pre-processed. In order to transfer each of these pixel values to a value between -1 and 1, we first divide each by 255.

The CNN deep learning training model is trained using a set of labelled facial photos by the pre-train algorithm. The pre-train program's needed input and output file types are shown in Figure 3. Chapter 4 will address the option of using user-submitted images. An RGB camera is used by the recognizer to capture live video. The facial detector selects a facial image from the frame after identifying the largest face in the scene.

The facial emotion recognizer receives its input from the facial image. The command line output from the facial emotion recognizer reveals the outcome after processing the facial image. The RGB real-time video from a camera and the checkpoint file from the pre-train software serve as the recognizer's inputs. The facial expression recognizer uses the fixed-size image that the facial detector produces as an input to identify the facial expression. Using OpenCV's cascade classifier, the facial detector scans frames from the live stream and uses them to identify faces. A square facial picture of the largest face is taken from the frame while faces are being recognised. Figure 4 displays a sample facial recognition algorithm.

The facial image will be reduced in size to a fixed size and transmitted to the facial expression recognizer for analysis if one is available. The required input and the facial expression recognizer's output are shown in Figure 5. When the face expression recognizer has finished processing the current image, the facial detector will pass the image to it. The pre-train program's entire results are needed by the facial emotion recognizer to validate the facial image.

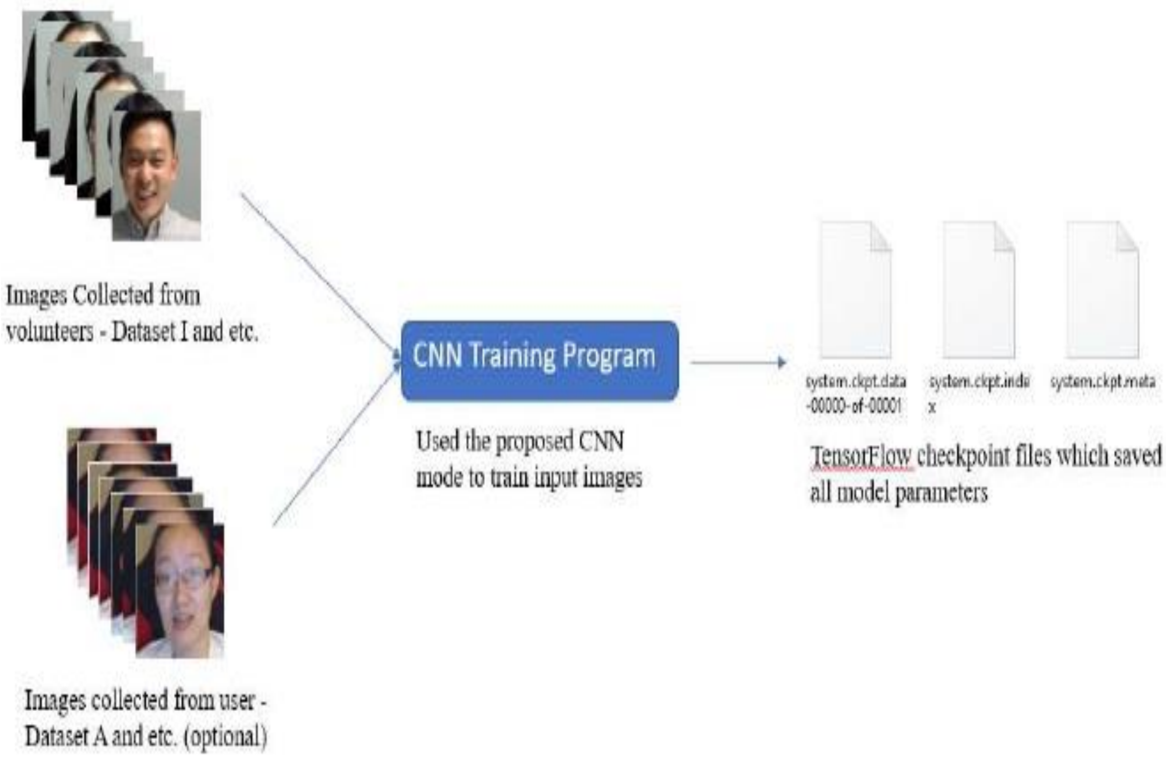


**Figure 3.3. The pre-train program's entire results are needed by the facial expression recognizer to validate the facial image.**

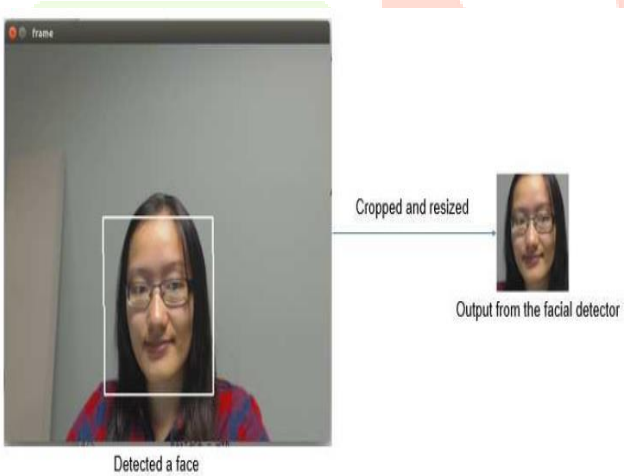


**Figure 3.4 Facial Detector - detect and extract image from video.**

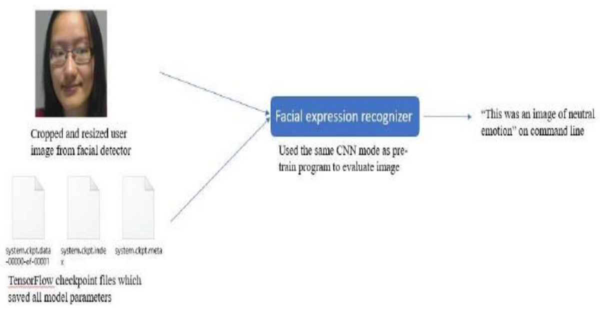


**Figure 3.5. Design of a facial expression recognizer. It assesses the facial detector image extraction results using checkpoint files from the pre-train programme.**

The pre-trained data is used by the facial expression recognizer to identify facial expressions. Deep learning technology is used to train the dataset in the pre-train system, and the results are saved in checkpoint files. When the facial expression recognizer receives an image input, it transforms the image into a tensor and guesses the image using information from the pre-train programme.

The ultimate result of the system is a single expression label, which is the output of the facial expression recognizer. The facial detector was created in Python using OpenCV, while the facial emotion recognizer was created in Python using Tensorflow and the deep learning method. The approach also detects a neutral face in addition to the six universal emotions of anger, contempt, fear, happiness, sadness, and surprise. The user of the suggested system has four possibilities. A train-only option with a command-python system is the initial choice. the input size for the py -t train file checkpoint.

The folder name used to store the tensorboad file and the name of the checkpoint file without an extension are both part of the checkpoint name. The second choice is a command-python system with a validation-only option. py -v file checkpoint name input validation For validation purposes, this option needs a checkpoint file containing finished training data. An image or a tfrecords file can be used as the input. With a command-python system, the third option is the train-validate option. py -b train file validation input size and file checkpoint name. This choice combines choices one and two. Live option using a command-line Python system is the final option. checkpoint name input size in Python. For live streaming, this option needs an RGB camera and a checkpoint file.

## 3.4 FACE DETECTION

- Face Registration is a computer technology that recognises human faces in digital photographs and is used in a range of applications.
- Faces are first located in the image during this face registration process utilising a set of landmark points known as face localization or face detection.
- Face registration is the process of geometrically normalising these recognised faces to match a template image.

According to Viola and Jones, the three essential steps of the original VJ detector method are the selection of Haar features using an IntegralImage representation, Adaboost learning, and the creation of cascade-classifiers. The two Haar features that Viola and Jones [43] found to be suitable describe the differences in brightness intensities between different parts of the human face, such as how the eyes are frequently darker than the top cheeks and the bridge of the nose. An integral picture format was suggested in order to make the computation of the sums of pixels under the black and white rectangles simpler. This allows the representation of two-rectangle and three-rectangle features to be computed in six and nine array references, respectively [43]. Each computation of the sum of pixels beneath a rectangle can then be reduced to four array references. After features are chosen, Adaboost learning is used to combine a number of weak classifiers into a final classifier that can discriminate between positive and negative images. The result is the fewest number of features that are optimally able to do this. Last but not least, classifiers are built in a "cascade" method to increase efficiency, with the latter classifier only being activated when the former classifier achieves a successful result. Due to this principle, the VJ face detector is known for its fast speed of feature computation and efficient feature selection, providing the capability for real-time analysis.

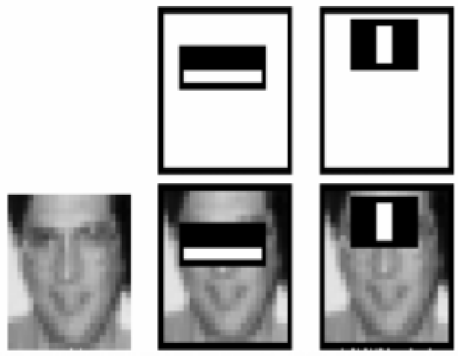

**Figure 4.6: Two Haar features selected for face detection**

In our study, the Haar cascade classifier is used for face detection in both offline and real-time systems, and since the FER component is the main focus, the non-frontal situation is ignored. The input photos are loaded, and the grayscale mode is applied. The four coordinates of the faces' rectangular region of interest (ROI) are returned by the classifier if it locates the faces. These four vertices are then used to trim the faces and remove the unimportant backgrounds after the ROI's placements have been determined. Image processing can then be dealt with. Figure 3 displays instances of faces from the two datasets (CK+ and JAFFE) with detected ROI. 3. Regarding the FER-2013 dataset, the disparate head angles and poor resolution of the photos (48 48 pixels) hinder accurate face detection.

## 4 RESULT AND DISCUSSION

The total number of accurate predictions divided by the total number of predictions made for a dataset is the accuracy. It can tell us right away if a model was properly trained and how it might perform in general. However, it doesn't provide specific details about how it applies to the problem. Precision, also known as PPV, is a good indicator, but false positives have a hefty cost. When there is a high cost associated with false negatives, recall is the model statistic used to determine the optimal model. Recall is beneficial while the cost of false negatives is costly. When you want to find symmetry between recall and precision, an F1-score is necessary. It serves as a general indicator of the model's precision. It combines recollection and precision. Low false positives and false negatives correlate with a good F1-score.

**True Positives (TP)**:- These are the accurately predicted positive values, indicating that both the actual class value and the projected class value are both true.

**True Negatives (TN)**:- These are the accurately predicted negative values, indicating that both the predicted class and the actual class have no value. False positives and false negatives, which occur when our actual class differs from the projected class, are these values.

**False Positives (FP):-** When actual class is no and predicted class is yes.

**False Negatives (FN)**:- When actual class is yes but predicted class in no.

$$\text{Accuracy} = \frac{(TP + TN)}{(TN + FN) + (FP + TP)}$$

$$\text{Recall} = \frac{TP}{(FN + TP)}$$

$$\text{Precision} = \frac{TP}{(FP + TP)}$$

$$F1 - \text{measure} = 2 \times \frac{(\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})}$$

Where,

- True positive (TP) = correctly identified
- False positive (FP) = incorrectly identified
- True negative (TN) = correctly rejected
- False negative (FN) = incorrectly rejected

**Precision**

Precision means to determine the number of positive class predictions that actually belong to the positive class.

**Precision = TP/TP+FP**

**Recall**

Recall means to determine the number of positive class predictions made out of all positive samples in the dataset.

**Recall = TP/TP+FN**

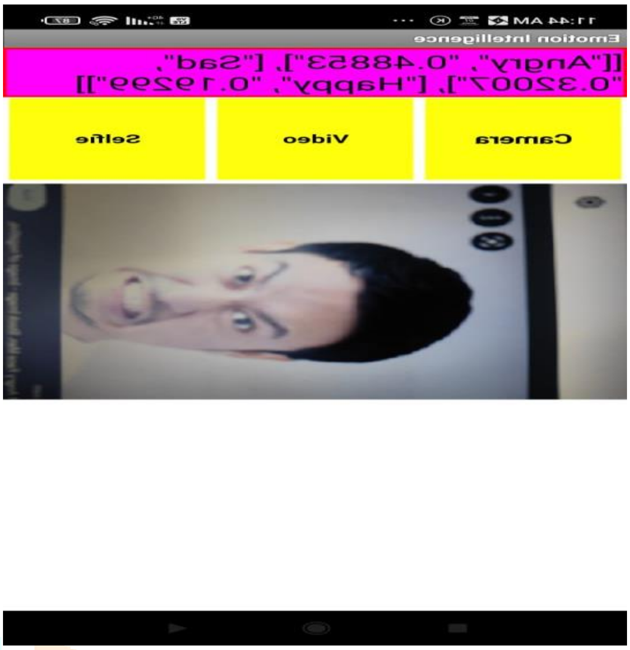


**Figure 4.1 Prediction Result Screen Shot for Video – Angry Face with Accuracy**

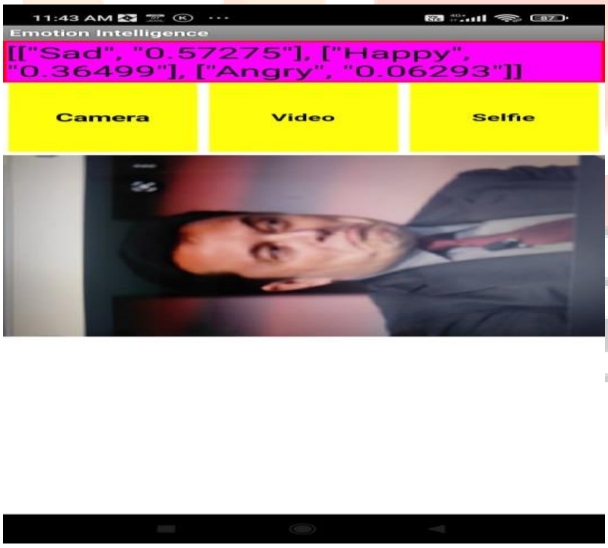


**Figure 4.2 Prediction Result Screen Shot for Video – Sad Face with Accuracy level**

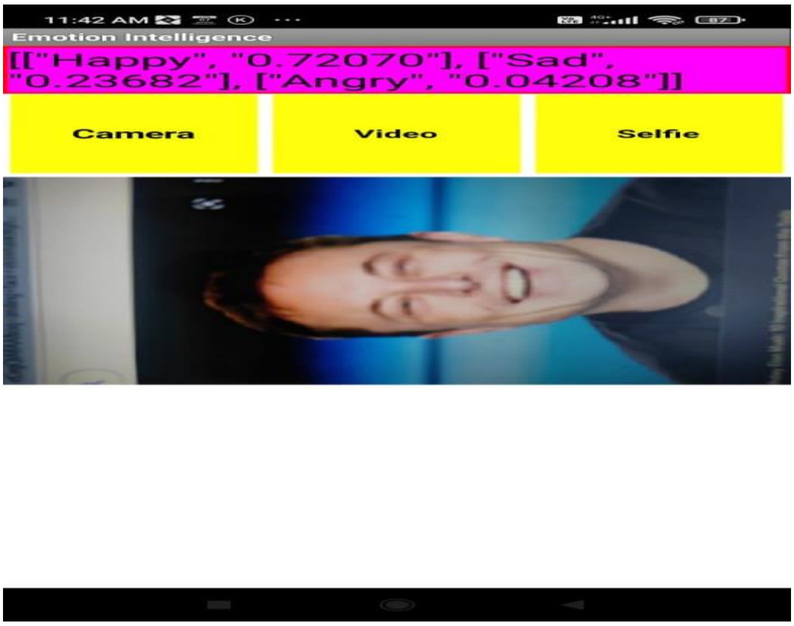


**Figure 4.3 Prediction Result Screen Shot for Video – Happy Face with Accuracy**

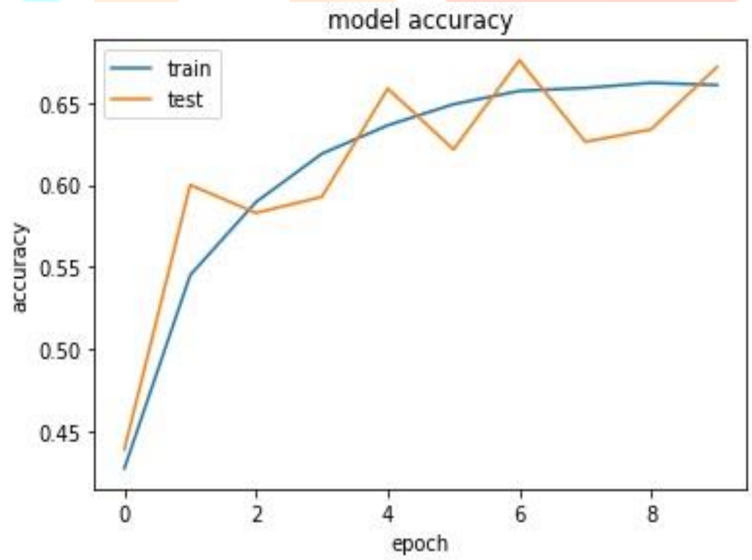


**FIGURE 4.4 PLOT FOR PREDICTIVE MODEL ACCURACY**

## CONCLUSION

Soft computing algorithms that predict facial expressions in people are crucial for human communication and engagement. They are also a crucial tool in behavioural research and medical rehabilitation. Techniques for non-invasive detection using facial image are quick and efficient. The idea behind this presentation was to use neural networks to build a smart system for classifying face image expressions.

**REFERENCES**

[1] A. Dhall, O. V. R Murthy, R. G. . J., and T.Gedeon. Video and image based emotion recognition challenges in the wild: Emotiw 2015. In the 2015 ACM on international Conference on Multimodal interaction (2015), ACM, pp. 423–426.

[2] Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., et al. Tensor- flow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467 (2016).

[3] Abaya, W. F., Basa, J., Sy, M., Abad, A. C., and Dadios, E. P. Low cost smart security camera with night vision capability using raspberry pi and opencv. In Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), 2014 International Conference on (2014), IEEE, pp. 1–6.

[4] Black, M. J., and Yacoob, Y. Tracking and recognizing rigid and non-rigid facial motions using local parametric models of image motion. In Computer Vision, 1995. Proceedings., Fifth International Conference on (1995), IEEE, pp. 374–381.

[5] Bradski, G., et al. The opencv library. Doctor Dobbs Journal 25, 11 (2000),120–126.

[6] Bradski, G., and Kaehler, A. Learning OpenCV: Computer vision with the OpenCV library. ” O’Reilly Media, Inc.”, 2008.

[7] Burkert, P., Trier, F., Afzal, M. Z., Dengel, A., and Liwicki, M. Dexpression: Deep convolutional neural network for expression recognition. arXiv preprint arXiv:1509.05371 (2015).

[8] Chakraborty, A., Konar, A., Chakraborty, U. K., and Chatterjee, A. Emotion recognition from facial expressions and its control using fuzzy logic. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans 39, 4 (2009), 726–743.

[9] Chan, K.-P., Zhan, X., and Wang, J. Facial expression recognition by correlated topic models and bayes modeling. In Image and Vision Computing New Zealand (IVCNZ), 2015 International Conference on (2016), IEEE, pp. 1–5.

[10] Cheng, F., Yu, J., and Xiong, H. Facial expression recognition in jaffe dataset based on gaussian process classification. IEEE Transactions on Neural Networks 21, 10 (2010), 1685–1690.

[11] Coelho, G., Kougianos, E., Mohanty, S. P., Sundaravadivel, P.,and Albalawi, U. An iot-enabled modular quadrotor architecture for real-time aerial object tracking. In Nanoelectronic and Information Systems (iNIS), 2015 IEEE International Symposium on (2015), IEEE, pp. 197–202.

[12] Ekman, P., and Friesen, W. V. Manual for the facial action coding system. Consulting Psychologists Press, 1978.

[13] Gavrilescu, M. Proposed architecture of a fully integrated modular neural network-based automatic facial emotion recognition system based on facial action coding system. In Communications (COMM), 2014 10th International Conference on (2014), IEEE.

[14] Guo, Y., Tao, D., Yu, J., Xiong, H., Li, Y., and Tao, D. Deep neural networks with relativity learning for facial expression recognition. In Multimedia & Expo Workshops (ICMEW), 2016 IEEE International Conference on (2016), IEEE, pp. 1–6.