# APPROACHES FOR HARVESTING ON EMOTION EXTRACTION FROM HUMAN FACIAL EXPRESSIONS

Dr. Nilesh Shelke
Assistant Professor, Department of Computer Science and Engineering,
Priyadarshini Indira Gandhi College of Engineering, Nagpur, Maharashtra, India
nileshrandd@gmail.com

Dr. Sachin Upadhye
Assistant Professor, Department of Computer Application,
Shri Ramdeobaba College of Engineering and Management, Nagpur, Maharashtra, India
upadhyesd@rknec.edu

Prof. Satyajit S. Uparkar
Assistant Professor, Department of Computer Application,
Shri Ramdeobaba College of Engineering and Management, Nagpur, Maharashtra, India
uparkarss@rknec.edu

Hemant Bawane
Research Scholar, Priyadarshini Indira Gandhi College of Engineering,
Nagpur, Maharashtra, India
bawanehemant5@outlook.com

Danish Shivvanshi
Research Scholar, Priyadarshini Indira Gandhi College of Engineering,
Nagpur, Maharashtra, India
danishshivvanshi7@gmail.com

Sujeet Kosekar
Research Scholar, Priyadarshini Indira Gandhi College of Engineering,
Nagpur, Maharashtra, India
Sujeetkosekar0@gmail.com

**Abstract**

**While artificial intelligence and machine learning is penetrating through its applications in all the domains, detection of human facial expression has significant role to deal with machines for the interaction. The study is for the development of detection of facial emotion recognition systems are heavily dependent on geometry and appearance. The emotion analysis from images overlooks the challenges of unbalanced dataset and training models to overfitting, it led to a misprediction of actual emotions to be classified. Proposed approach integrates the simplicity of execution, lower computational complexity, exceedingly competitive outcomes beyond numerous real-world emotion classification tasks. The proposed architecture uses ensemble learning techniques on the top of SVM, HMM, KNN, Haar Cascade, Random Forest for emotion classification. Evaluation is done on the popular datasets Ferg-dB, CK+, JAFFE and Fer2013. Study shall predict the universal human emotions more efficiently and precisely. CNN is used for emotion predictor as a model.**

*Keywords***: Haar cascade; Convolutional Neural Network; Facial emotion recognition; KNN, HMM; SVM; Ensemble learning.**

## 1. Introduction

Human emotions are revealed by the expressions present on the human face. From the facial expression we can get non-verbal scientific clues and it gets expressed on the human face as per the emotions. Automatic detection of emotions from facial expression can have several applications in artificial intelligence. This paper is intended to build a human Facial Expression Recognition [FER] system which will take input of human facial images containing several expressions and after being recognized it will classify it into seven different universal emotion.

In 2013, Du S et al. introduced twenty one compound expressions, in which six were initial expressions and fifteen were produced from the basic expressions. CNN was used in this research give a huge amount of positive output and provided some extra expression. These expressions included surprise, disgust, fear, happiness, sadness, anger, neutral, sad and disgust, sad and surprised, disgusted and surprised and hatred, angry and disgusted, angry and surprised, fearful and disgusted, fearful and angry, fearful and surprised, sad and angry, happy and surprised, sad and fearful, happy and disgusted. Typically, facial expression classification consists of three steps which include detecting faces, extracting features and recognizing facial expressions [1].

Another study revealed a method to classify images using different algorithms for better and precise performances. For better results, image has to be recognized in a precise way so that the emotion should be classified easily. To increase the recognition accuracy, the input test image is pre-processed. In the next step, the face are detected. Features of the face are extracted from a video or image. Output of a second step is the formation of a feature vector. In the last step, expressions are classified by the classifier. For extracting facial features different algorithms are employed. These algorithms go for eyes, eyebrows, mouth, and nose [2]. Sometimes with insufficient features, many classifiers cannot get precise recognition rates. Intensifying the preciseness of the feature extraction methods is a critical part when extracting the unique features representing the facial expression. Classifiers play a critical role in different recognition and classification systems. It classifies and also recognizes facial emotion.

Section 2 explores the literature survey whereas section 3 elaborates on proposed methods. Datasets used in the paper are discussed in section 4. Section 5 is for the experimental setup. The screenshots of the working model are shown in section 6. Section 7 is dedicated to the experimental results. Section 8, section 9 includes the conclusion and future scope.

## 2. Literature Survey

Currently, facial expression recognition is a rapidly growing research topic. It has many applications in the present era and it assists in the human computer interaction. Isha Talegaonkar et al. developed a CNN model based on haar features. The method uses the significant features and achieved a good test accuracy. The method achieved a validation accuracy of 89.78% and test accuracy of 60.12% on fer2013 dataset using CNN. But method had a smaller number of features extracted [3].
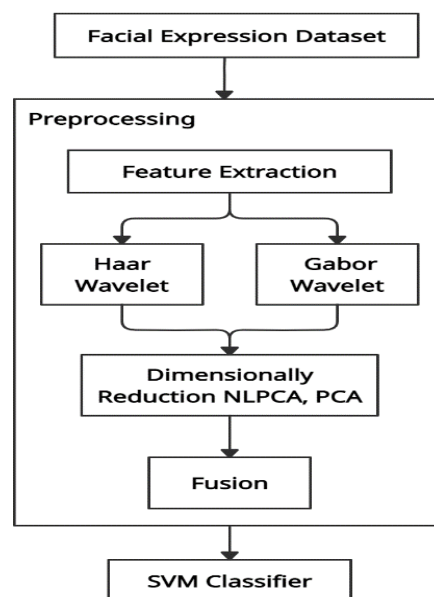


Fig. 1. Use of Haar and Gabor wavelet as feature extractor

Fig. 1 represents commonly used feature extraction technique which uses Haar wavelet, Gabor wavelet as a feature extraction. The feature vectors are high in dimensionality so transformation is required by PCA and NLPCA. These features extracted are concatenated with the help of fusion for better accuracy, SVM classifier is used for classification. The dataset used in the method is Ck+. But dataset contains small number of images and different model is trained by dividing data of the dataset which leads to learn the model itself [5].

Dan Duncan et al. developed a CNN model and uses two different datasets. The method uses trained VGG_S network with haar cascade filter. OpenCV provides this filter. It has prominently increased accuracy of both testing and training. The experimental results show that model had outperformed in terms of recognition rate. However, the model running at relatively slow frame rate which leads model to behave untenable [6].

Octavio Arriaga et al. has built a general CNN framework. These real time CNNs employed back propagation visualization technique. Guided back-propagation explores the dynamic of weights and evaluates the features. Two models had been used are mini-Xception and sequential fully-CNN. Their model scan be stacked for multi-class classification while maintaining real time interferences. It describes a complete pipeline mingled in a Care-O-bot 3 robot. However, increase in the number of parameters in a naive implementation caused to incompetent features and low performance of model [7].

Shervin Minaee et al. has employed deep learning strategy, which is able to focus on important parts of face. Their study concluded that different parts of the face are aligned with specific emotions. The method specifies important regions of face while classifying different facial expression. But method uses only single classifier over multiple datasets which leads to uneven predictions [8].

Minh-An Quinn et al. developed models capable of recognizing seven basic emotions from facial expressions. Their model achieved a test accuracy of 45.95% with SVM as a classifier and a test accuracy of 66.67% with the CNNs. Their model is very accurate on static images. But it doesn't reflect real-time conditions due to captured and non-posed images used in the dataset which leads to improper prediction [9].

J. Jayapradha et al. used the Viola Jones algorithm to implement recognition of individual identity of expression and to extract the features of the face. The experimental results show that an average accuracy of 70% on the happy, neutral, and sad emotion is achieved. Unfortunately, the method better predicts only three emotions and model is trained on six emotions which leads to inaccurate results [10].

D. Yang et al. developed a model to solve the problem of FER on facial recognition in virtual learning environments. It uses Haar cascade algorithm to detect eyes, mouth, and edge line, slant line. The method identifies all kinds of emotions through neural network method. However, the pre-processing of images doesn't removed illumination and pose of image which leads to improper features ranking during feature extraction [11].

Alex Martinet et al. has linearly combined a set of face spaces. This study defined few primary emotion categories. However, the study was not as good at anger and sadness and gives worst prediction on at fear and disgust emotions [12].

MalyalaDivya et al. presented live facial emotion detection through image processing. They have also applied the techniques of AI. The employed AdaBoost algorithm for this task. Best features out of 160000+ features. Method includes libraries like OpenCV, Tensor Flow, and Keras. The experimental results show that it achieved an emotion classification accuracy of 66% on each emotion. But the method had a prediction error due to lower accuracy on classification of emotions which leads analysis error [13].

Carlos Busso et al. analysed the strength and weakness of facial expression classifiers as well as acoustic emotion classifiers, also two fusion approaches were compared feature level and decision level, and observed that feature level is best unimodal system. The experimental system shows the performance of five facial blocks and the combined facial expression classifier. But the system confuses sadness with neutral state, which leads to poor performance during prediction analysis [14].

Nithya Roopa S. proposed a method to employ deep learning as well as image classification method to detect expression and classify the images. In the method Inception Net v3 model has been employed for expression recognition. Further analysed that using Inception Net for FER gives an accuracy of 35.6%, transfer learning has been used for the training of the model. It has produced high number of spectrograms for training. But a smaller number of images used in dataset for each class during training process which leads to less accuracy [15].

Hongli Zhang et al. uses a technique based on CNN and image edge recognition. This approach automatically learn pattern features and reduce the incompleteness caused by artificial design features. The convergence speed of model is faster in complex background environments. However, the overall recognition rate of fer2013 dataset is low which leads to analysis error [16].
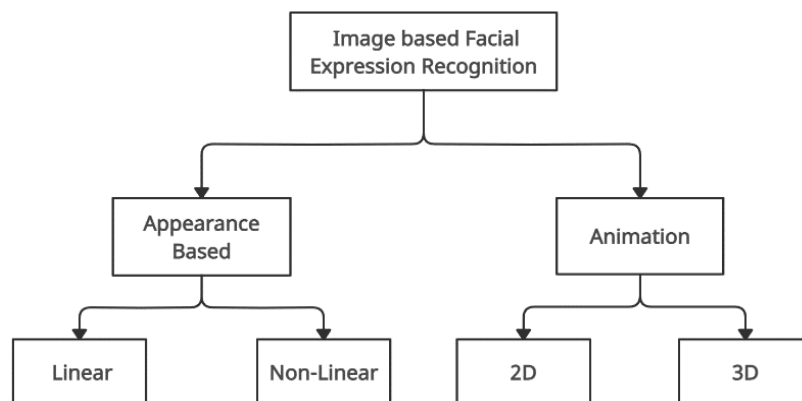


Fig. 2. Image based FER

Fig. 2 represents the image based facial expression recognition in which images are of two types: appearance based and model-based images. Appearance based images can be linear and non-linear whereas model-based images can be 2D and 3D. However, the study shows that model-based images have low learning rate [17]. Mohammad Rami Koujan et al. constructed a big quantity facial videos dataset. They train this dataset through a deep CNN. The method is executed at the rate of 50 frames per second & robustly estimating parameters of expression. The method had reconstructed 3D face from images and videos which consists of combined identity and expression 3D face modelling, initialization stage of estimating camera parameters. They have used ground truth creation, DeepExp3D network, back end emotion classifier and 5 different datasets. However, the method had a less recognition rate for some emotions, which leads to inaccurate prediction [4].

### 3. Dataset

To study the emotion recognition from the face four datasets are exploited. These four datasets consist of facial images. From datasets images are extracted in binary array format and binary value. The whole data is normalized. This data is taken as input and there are seven facial expressions as output.

### 3.1. *Fer2013*

It is an open-source dataset created by the P. L. carrier and A. Courvile for project workafter that it is publicly shared for Kaggle competition. This dataset consists 35887 number of labelled images.
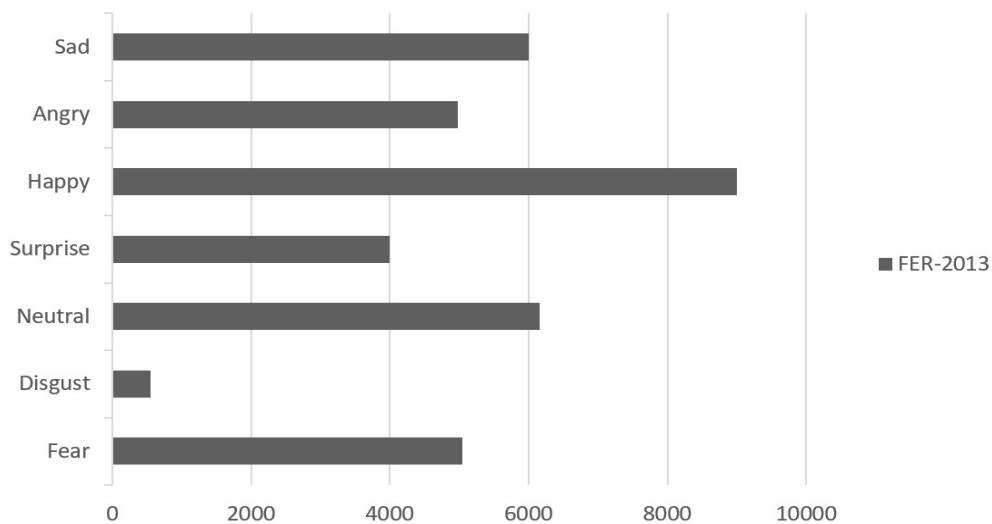


Fig. 3. Graph of distribution of images

This dataset has images of 48 x 48 pixels and all the images are black and white (Grayscale images). Samples from dataset are given below
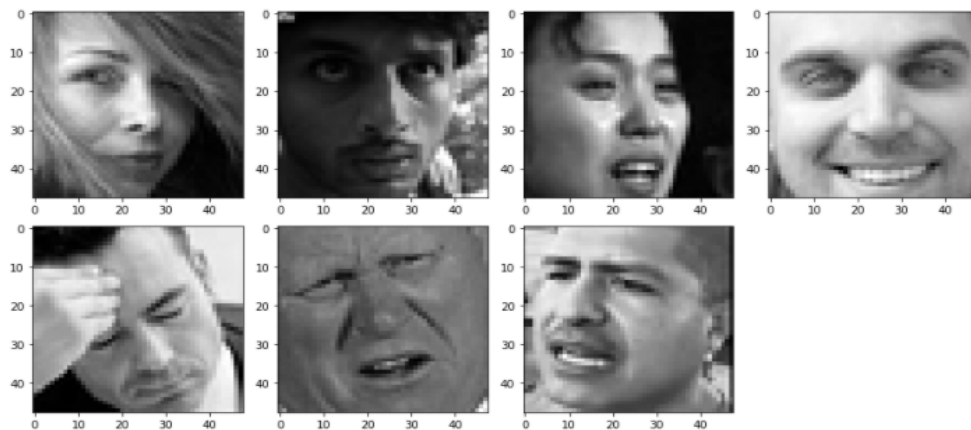


Fig. 4. Sample images from dataset

### 3.2. *CK+*

CK+ contains facial expression images. Dataset is developed by Lucey, P. Cohn, J. F. Kanade for their research [21]. This dataset is available on the Kaggle website publicly. This unbalanced image dataset consists of 981 facial images. Fig. 5 shows the number of facial images is classified as per the emotions.
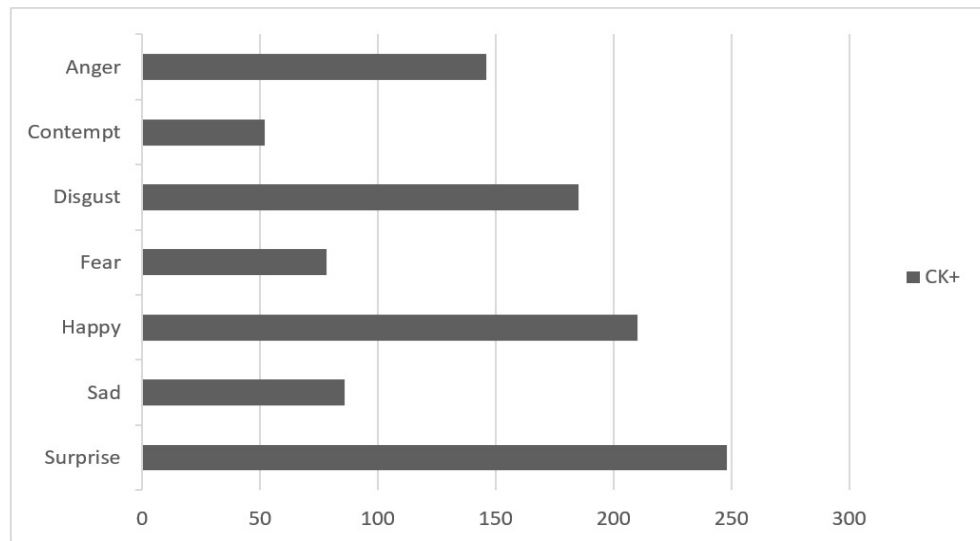


Fig. 5. Number of facial emotion Image in CK+

Facial image has resolution 640 x 490 pixel. Images are labelled with seven basic expressions i.e., sad, surprise, fear, anger, contempt, disgust, happy. The images in the dataset are black and white (Grayscale images). Some of the pictures from the dataset is shown below



Fig. 6. Sample images from CK+ dataset

### 3.3. *Ferg-db*

FERG-DB (Facial Expression Research Group 2D Database) is a dataset of 2D images of stylized characters. with an annotated facial expression. It has images of six stylized characters (3 males and 3 females) - Ray, Malcolm, Jules, Bonnie, Mery and Aia. The characters were modelled using the MAYA software and rendered out in 2D to create the images. Each character has seven types of expressions - anger, disgust, fear, happy, neutral, sadness and surprise. The database has 55767 annotated face images.

Fig. 7. Samples from Ferg-db dataset

### 3.4. *JAFFE*

JAFFE (Japanese Female Facial Expression) dataset contains 10 Japanese female expressers, 7 posed facial expressions (6 basic facial expressions + 1 neutral) and several images of each expression for each expresser, it contains 213 total images, each image has averaged semantic ratings on 6 facial expressions by 60 Japanese viewers, images having resolution of 256x256 pixels, and 8-bit grayscale.



Fig. 8. Samples from JAFFE dataset

### 4. Proposed System

In the proposed system, algorithms employed are SVM, Haar Cascade, HMM, KNN and Random Forest.



Fig. 9. Proposed Architecture

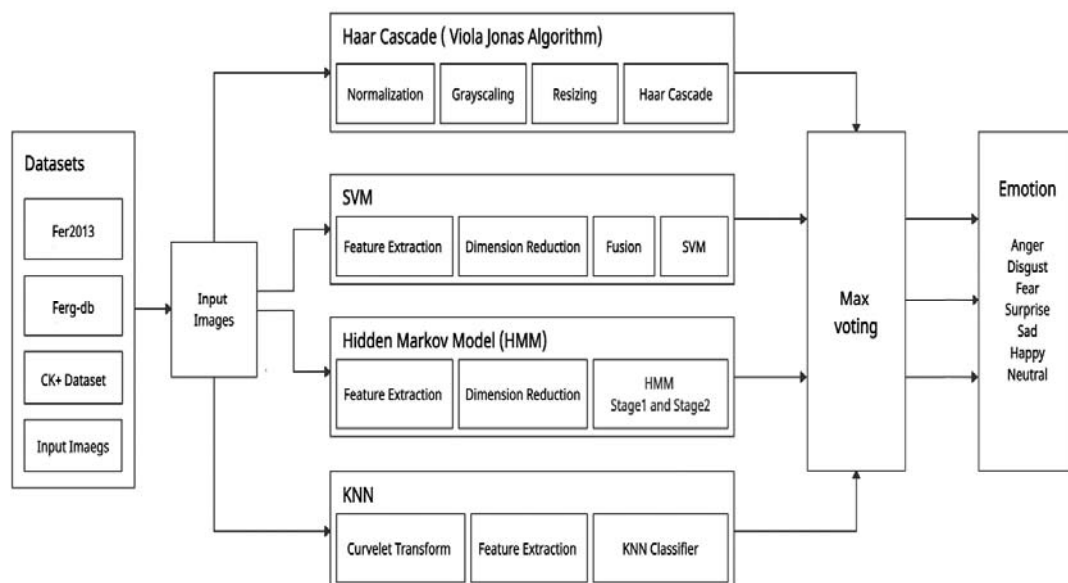In all these methods, four datasets are employed Fer2013, Ferg-Db, CK+ and JAFFE. This system will detect the facial expressions of human from the given seven expressions namely sad, angry, happy, surprise, neutral, disgust and fear. However, the emotion analysis from facial images faces the challenges of unbalanced dataset and training models to overfitting that often led to a misprediction of emotion. The developed model is capable of recognition of facial expression using real time webcam. Given below is the proposed architecture of system. Images from the datasets are fed to proposed four methods. In Haar cascade it will normalize the image converting to grayscale, resizing the image. In SVM dimension reduction and fusion will take place. In HMM dimension reduction of images takes place. In KNN curvelet transformation of images are done. In Random Forest decision class is used.

### 4.1. SVM

Support-vector machine (SVM) is a supervised machine learning model with combined learning algorithms that scrutinize information for classification. In SVM operations are done in input space by kernel function, but not in attribute space, hence represents greater dimensionality. Computational complexity decreases when attributes are mapped of the input space to the attribute space by the kernel function.

SVM with the polynomial kernel function has been exploited for image processing for getting a better accurate model. This kernel function can work more precisely on the data so that underfitting and overfitting can be avoided. The mathematical form of the kernel equation is listed. Polynomial equation is

$$y(x_i, x_j) = (\gamma x_i, x_j + 1)^d \tag{1}$$

Here $y(x_i, x_j)$ is the kernel function; $\gamma$ is the gamma term in the kernel function for all kernel types. $\gamma$, d are user-controlled parameters

### 4.2. HMM

The HMM is based on augmenting the Markov chain. A Markov chain is a model which informs us something about the probability, order of arbitrary variables, states, where each can take on values from some set. The stage before the latest stage has no effect in the future except via the latest stage. It's as if to predict tomorrow's weather you could examine today's weather but you weren't allowed to look at yesterday's weather. Formally, an HMM is mark out by

$$\lambda = [M, N, \Pi, A, B] \tag{2}$$

The classical constraint of HMM are:

$$\sum_{j=1}^{N} a_{ij} = 1, \forall i \in M \tag{3}$$

$$\sum_{j=1}^{N} b_j(k) = 1, \forall k \in N \tag{4}$$

$M$ $\quad$ (m$_1$…m$_s$) set of s>0 states of model, state at time t is referred by q$_t$ $\epsilon$ M.
It is the number of possible corresponding observation value of every state.

$N$ $\quad$ {n$_1$,….n$_v$}, discrete set of the V observable symbols, a symbol at state at t is referred by
$o_t \in N$ .
It is the number of state in Markov chain.

$\pi$ $\quad$ {n$_1$, …,n$_s$) original possibility distribution vector $i(1 \leq i \leq s)$ .

| A | $\{a_{ij}\}, 1 \leq i, j \leq s$ |
|---|---|

Where $a_{ij} = P(q_{t+1} = s_j \mid q_t = s_i)$ for models of order 1,

It is possibility matrix for sate transfer.

| B | $\{b_{j(k)}\}, 1 \leq j \leq s, 1 \leq k \leq v$ |
|---|---|

Where

$$b_j(k) = P(O_t = m_k \mid q_t = n_j)$$

It is a group of exist able function of observed value.

### 4.3. *KNN*

K-nearest neighbour algorithm is a method for classification depending upon the nearest training instances in the classification method.

In KNN data can be converted each image to a vector of fixed-length with real numbers, then the most common distance function for KNN is Euclidean distance. The mathematical expression for Euclidean distance is listed:

$$d(X, Y) = \sqrt{\sum_{i=1}^{k}(X_i - Y_i)^2} \tag{5}$$

Where d(X, Y) is Euclidean n-space with two points $X_i$, $Y_i$ are Initial points, Euclidean vectors k is a k-space.

### 4.4. *Haar cascade algorithm*

Face recognition is very important task in facial expression detection, for face detection OpenCV has a Haar cascade algorithm. Haar cascade algorithm is also known as Viola-Jones algorithm. This algorithm is trained to detect the objects in an image, these algorithms are trained in a set of positive and negative facial images. Haar cascade algorithm is proved very effective for the detection of a face from images. Haar cascade algorithms use face features like eyes, nose, eyebrows, and mouth. It takes an image as an input from the dataset and detects the edge features, line features, and four-rectangle features from the face. It also checks for whether an image contains a face or not.
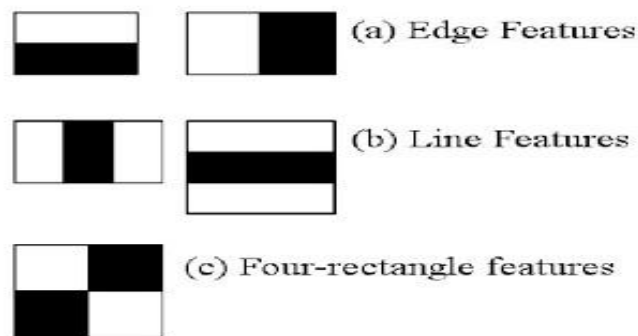


Fig. 10. Haar features

### 4.5. *Random Forest*

In this study to improve the method of classification, regression at training time the random forest is used. It operates by constructing a number decision tree model and outputting the decision class i.e., mode of classification of an individual tree. Random decision forests correct for decision trees habit of overfitting to their training set. Random forest is a collection of many decision trees trying to perform a successful implementation. In this study, random forest has been implemented on each random dataset. Random forest is used for a probability selection-based method to determine all the individual decision classes. The mathematical representation of averaging voting is shown in the following Fig. 11.
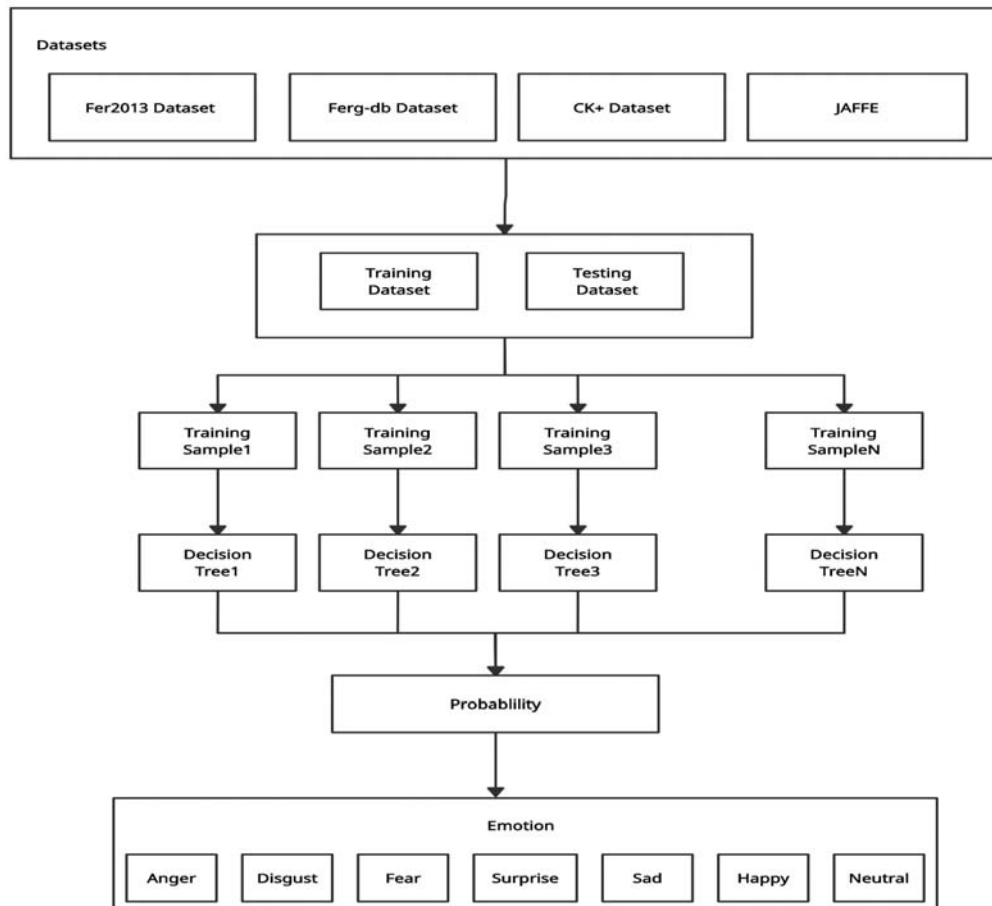
Fig. 11. Random forest

## 5. Experimental Setup
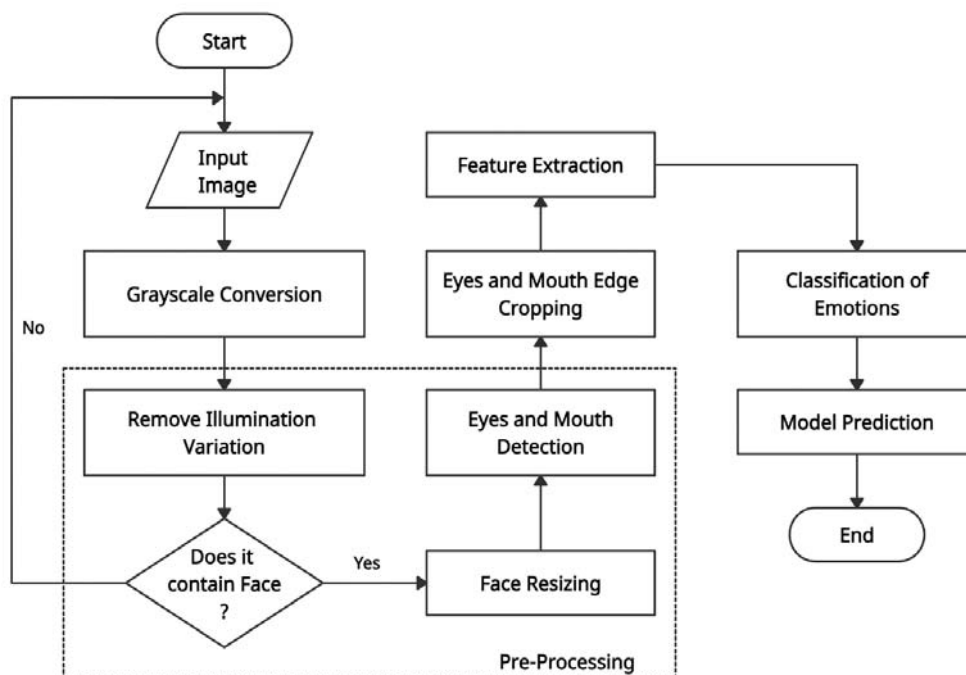


Fig. 12. Face expression detection method

**5.1 Pre-Processing**

The pre-processing step helps to bring the input data into a form that is predictable and analyzable by neglecting the unwanted parts of the data. The image from the datasets is raw image data which is easily corrupted by noise and hardware issues. In the proposed work, pre-processing undergoes for resizing the images, pixel transformation, removing illumination, dimensionality reduction, brightness transformation, image restoration. The mathematical representation of a pre-processing function is expressed as

$$\beta_{pr} = w_p[I_t] \tag{6}$$

Where, $\beta_{pr}$ is the output of pre-processing function, $I_t$ is the input image data and $w_p$ is the pre-processing function that is represented by

$$w_p = [w_{NZ}, w_{re}, w_{dr}, w_{da}] \tag{7}$$

Where $w_{NZ}$ is the normalization function, $w_{re}$ is the resize image function, $w_{dr}$ is the dimensionality reduction function and $w_{da}$ is the data augmentation function.

*5.1.1    Normalizing*

Data normalization is an important pre-processing task. It guarantees that each input parameter has a similar data distribution (pixel, in this case). It supports in rapid convergence. We subtract the mean from each pixel. Result is then divided by the standard deviation for the normalization task. Here, the image data can be represented in the form of image pixel in the array matrix.

$$P = \begin{matrix} x_{11} & .. & .. & x_{1j} \\ : & : & : & : \\ : & : & : & : \\ x_{i1} & : & : & x_{ij} \end{matrix} \tag{8}$$

Where P is represented as the image data store in form of an array matrix and $X_{ij}$ is an image pixel.

$$W_{nz} = \frac{P_{ij} - P_{min}}{P_{max} - P_{min}} \tag{9}$$

Where $P_{ij}$ is a represent the image data, $P_{min}$ is minimum array data of image, $P_{max}$ is a maximum array data of an image.

*5.1.2    Resize images*

It is important to do resizing the images in given data sets. Most of the images are different in pixel size that which causes the problem to extract features. In this study cropping of all the images up to that, it can show only essential features like face posture and many so it can help to predict more accurately.

*5.1.3    Dimensionality Reduction*

In the dataset, some of the images in RGB channels them is in the grayscale channel. In this study, dimension reduction can perform an important role to boost CNN model accuracy or make a model more traceable.

### 5.1.4    *Data Augmentation*

This step includes augmenting the existing dataset to the set of existing images. Like Scaling, rotations and other affine transformations are typical. It is done to expose the neural network to a wide variety of variations. Mathematical representation of the data augmentation.

$$W_{da} = [T_{rot}, T_{fip}, T_{b\&c}] \tag{10}$$

Where $W_{da}$ is represented as data augmentation function, $T_{rot}$ is image rotation techniques, $T_{fip}$ is image flipping technique, $T_{b\&c}$ is brightness & contrast adjusting technique.

## 5.2. Feature Extraction

This step involves the extraction of features from facial images using various feature extraction algorithms. Facial feature extraction mainly depends on the classification method and the different environments. In the feature extraction compression of information, reducing irrelevant information and also removing noise from the data. Eye and nose and edge and slant-line are extracted from the images. After feature extraction face is detected in an image.

$$\Phi_{ex}(w_{pr}) = [\Phi_i^E, \Phi_i^{Ed}] \tag{11}$$

$$\Phi_i^E = [\Phi_1^E, \Phi_2^E, \Phi_3^E ....\Phi_n^E] \tag{12}$$

$$\Phi_i^E = [\Phi_1^{Ed}, \Phi_2^{Ed}, \Phi_3^{Ed} ....\Phi_n^{Ed}] \tag{13}$$

Where, $\Phi_i^E$ and $\Phi_i^{Ed}$ is the eye and nose feature and edge and slant-line feature respectively and i=1,2,3….n.

## 5.3 Emotion Classification
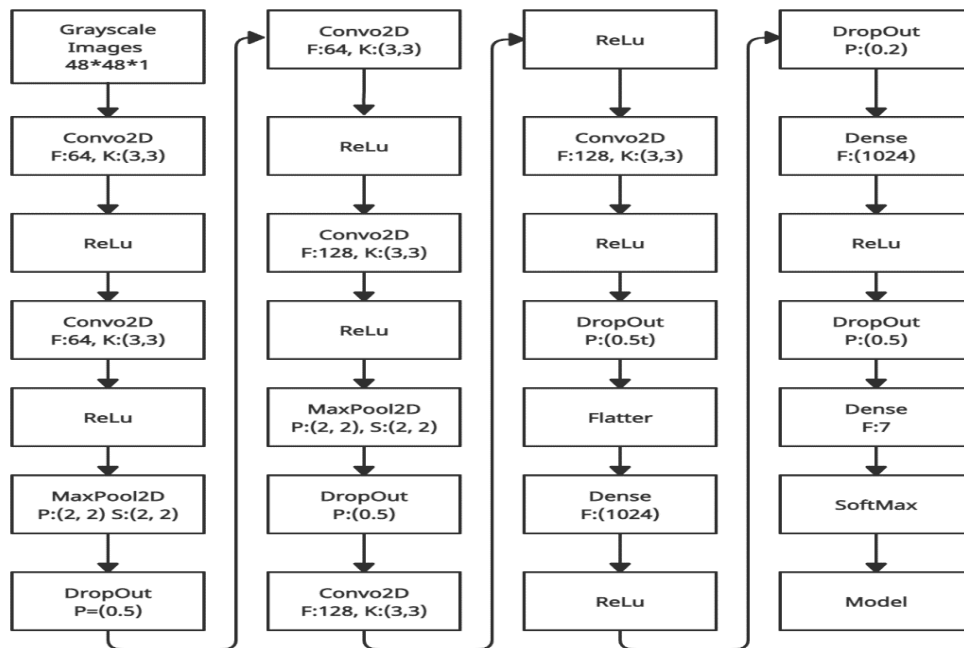


Fig. 13. CNN architecture

In emotion classification, the images from the dataset will be classified into seven universal emotions: happy, sad, disgust, angry, surprise, neutral, fear labelled as in the dataset. The training of the model is done on the CNN, Fig. 13 shows the CNN architecture used in the proposed system. The given CNN architecture takes a greyscale image of pixel 48x48x1 as an input, a convolutional 2D layer is used as 1st layer and ReLu activation function is used. ReLu function will output the input directly if it is positive, otherwise, it will output zero. Maxpool2D Layer reduces the dimensionality and allows assumptions to be made about features, dense layer performs a matrix-vector multiplication, dropout reduces unwanted feature of the images and it also helps to avoid overfitting in the CNN. Flatten layer is an output layer it converts data into a one-dimensional array. Grayscale images are provided to CNN architecture as input.

## 5. Screenshots
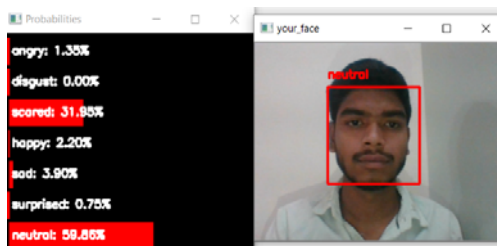


Fig. 14. Testing implemented model (Neutral)
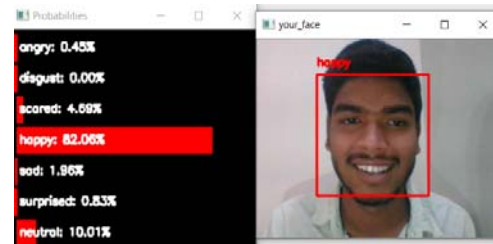
Fig. 15. Testing implemented model (Happy)

## 6. Result analysis and discussion

To access the performance of our models, a comparison of our models with the existing researches has been done, also combination of precision, recall and f1-score and accuracy to compare with existing researches. Accuracy measures the proportion of results amongst the evaluated set, precision shows the positive predictive value, and recall captures the sensitivity or true positive rate of models. Datasets has been tested with different methods. Graphs of training accuracy, validation accuracy, and confusion matrix of datasets are:

| EMOTION | MODEL | Fer2013 DATASET | | |
|---|---|---|---|---|
| | | P | R | F |
| ANGER | SVM | 0.79 | 0.85 | 0.82 |
| | HAAR | 0.53 | 0.77 | 0.63 |
| | HMM | 0.75 | 0.84 | 0.79 |
| | KNN | 0.50 | 0.62 | 0.55 |
| | ESM | 0.85 | 0.89 | 0.85 |
| | RF | 0.89 | **0.92** | 0.91 |
| | B0 | 0.71 | 0.69 | 0.70 |
| | B1 | 0.67 | 0.67 | 0.67 |
| | B2 | **1** | 0.85 | **0.92** |
| | B3 | 0.85 | 0.81 | 0.82 |
| DISGUST | SVM | 0.68 | 0.79 | 0.73 |
| | HAAR | 0.61 | 0.56 | 0.58 |
| | HMM | 0.73 | 0.61 | 0.66 |
| | KNN | 0.75 | 0.78 | 0.76 |
| | ESM | 0.82 | 0.89 | **0.86** |
| | RF | 0.85 | 0.89 | **0.86** |
| | B0 | **0.93** | 0.71 | 0.81 |
| | B1 | 0.71 | 0.63 | 0.67 |
| | B2 | 0.66 | **1** | 0.80 |
| | B3 | 0.73 | 0.79 | 0.75 |

| EMOTION | MODEL | Fer2013 DATASET | | |
|---|---|---|---|---|
| | | P | R | F |
| SURPRISE | SVM | 0.78 | 0.69 | 0.73 |
| | HAAR | 0.89 | 0.73 | 0.80 |
| | HMM | 0.74 | 0.79 | 0.76 |
| | KNN | 0.67 | 0.59 | 0.62 |
| | ESM | 0.89 | **0.92** | **0.90** |
| | RF | **0.90** | 0.85 | 0.87 |
| | B0 | 0.84 | 0.83 | 0.83 |
| | B1 | 0.81 | 0.89 | 0.85 |
| | B2 | 0.87 | 0.89 | 0.88 |
| | B3 | 0.72 | 0.79 | 0.75 |
| SAD | SVM | 0.85 | 0.87 | **0.86** |
| | HAAR | 0.74 | 0.88 | 0.80 |
| | HMM | 0.78 | 0.79 | 0.79 |
| | KNN | 0.86 | 0.89 | 0.87 |
| | ESM | **0.91** | 0.88 | 0.89 |
| | RF | 0.88 | **0.90** | 0.89 |
| | B0 | 0.64 | 0.59 | 0.61 |
| | B1 | 0.68 | 0.54 | 0.60 |
| | B2 | 0.50 | 0.84 | 0.62 |
| | B3 | 0.85 | 0.79 | 0.82 |
| HAPPY | SVM | **0.87** | 0.84 | 0.86 |
| | HAAR | 0.68 | 0.79 | 0.73 |
| | HMM | 0.77 | 0.82 | 0.79 |
| | KNN | 0.59 | 0.53 | 0.56 |
| | ESM | 0.79 | 0.83 | 0.81 |
| | RF | 0.90 | 0.89 | 0.89 |
| | B0 | 0.84 | 0.81 | 0.82 |
| | B1 | 0.82 | 0.84 | 0.83 |
| | B2 | 0.86 | **0.93** | **0.89** |
| | B3 | 0.8 | 0.81 | 0.8 |
| FEAR | SVM | 0.79 | 0.76 | 0.77 |
| | HAAR | 0.57 | 0.63 | 0.59 |
| | HMM | 0.72 | 0.63 | 0.67 |
| | KNN | 0.63 | 0.74 | 0.68 |
| | ESM | **0.93** | 0.79 | 0.85 |
| | RF | 0.92 | **0.96** | **0.94** |
| | B0 | 0.62 | 0.70 | 0.66 |
| | B1 | 0.64 | 0.69 | 0.67 |
| | B2 | 0.81 | 0.79 | 0.80 |
| | B3 | 0.8 | 0.81 | 0.8 |

| EMOTION | Fer2013 DATASET | | | |
|---------|------|------|------|------|
|         | MODEL | P | R | F |
| NEUTRAL | SVM | 0.75 | 0.69 | 0.71 |
|         | HAAR | 0.89 | 0.79 | 0.84 |
|         | HMM | 0.73 | 0.85 | 0.79 |
|         | KNN | 0.63 | 0.59 | 0.61 |
|         | ESM | 0.84 | **0.92** | 0.88 |
|         | RF | **0.94** | **0.92** | **0.93** |
|         | B0 | 0.65 | 0.68 | 0.67 |
|         | B1 | 0.73 | 0.71 | 0.72 |
|         | B2 | **0.94** | 0.67 | 0.79 |
|         | B3 | 0.73 | 0.79 | 0.78 |

Table 1. Comparison of precision, recall, f1-score with existing approaches for Fer2013 dataset

Given methods are compared with Ref [22], Ref [23] for the precision, recall, f1-score as shown in above Table 2.

Baseline B0, B1 are Ref [22] introduced a deep neural network architecture to detect human face expression using HOG with CNN, HOG is used for feature extraction of images, also trained two models with CNN and HOG-CNN. It found that when image size is small and picture quality is unclear then results are not effective that CNN model.

Baseline B2, B4 are Ref [23] method used in the architecture is RNN and CNN, it uses two-class model architecture four and seven class model architecture, both the architecture utilizes CCE as loss function and Adam optimizers, it compares test cases of public and private test, the model works on a real time face emotion detection.

| EMOTION | Ferg-Db DATASET | | | |
|---------|------|------|------|------|
|         | MODEL | P | R | F |
| ANGER   | SVM | 0.85 | **0.89** | 0.86 |
|         | HAAR | 0.81 | 0.78 | 0.79 |
|         | HMM | 0.71 | 0.68 | 0.70 |
|         | KNN | 0.64 | 0.58 | 0.60 |
|         | ESM | 0.92 | 0.85 | 0.88 |
|         | RF | 0.91 | 0.87 | 0.89 |
|         | B0 | 0.71 | 0.69 | 0.70 |
|         | B1 | 0.67 | 0.67 | 0.67 |
|         | B2 | **1** | 0.85 | **0.92** |
|         | B3 | 0.80 | 0.81 | 0.80 |

| EMOTION | Ferg-Db DATASET | | | |
|---|---|---|---|---|
| | MODEL | P | R | F |
| DISGUST | SVM | 0.87 | 0.86 | 0.93 |
| | HAAR | 0.85 | 0.8 | 0.82 |
| | HMM | 0.77 | 0.67 | 0.78 |
| | KNN | 0.68 | 0.58 | 0.63 |
| | ESM | 0.89 | 0.89 | 0.89 |
| | RF | 0.90 | 0.94 | **0.92** |
| | B0 | **0.93** | 0.71 | 0.81 |
| | B1 | 0.71 | 0.63 | 0.67 |
| | B2 | 0.66 | **1** | 0.80 |
| | B3 | 0.85 | 0.69 | 0.76 |
| FEAR | SVM | 0.65 | 0.74 | 0.69 |
| | HAAR | 0.71 | 0.69 | 0.70 |
| | HMM | 0.71 | 0.68 | 0.69 |
| | KNN | 0.76 | 0.58 | 0.66 |
| | ESM | 0.79 | 0.86 | 0.82 |
| | RF | **0.91** | **0.89** | **0.90** |
| | B0 | 0.84 | 0.83 | 0.83 |
| | B1 | 0.81 | **0.89** | 0.85 |
| | B2 | 0.87 | **0.89** | 0.88 |
| | B3 | 0.64 | 0.89 | 0.74 |
| SURPRISE | SVM | 0.69 | 0.78 | 0.73 |
| | HAAR | 0.77 | 0.80 | 0.78 |
| | HMM | 0.71 | 0.57 | 0.63 |
| | KNN | 0.68 | 0.58 | 0.62 |
| | ESM | 0.86 | 0.78 | 0.81 |
| | RF | **0.94** | **0.88** | **0.91** |
| | B0 | 0.65 | 0.68 | 0.67 |
| | B1 | 0.73 | 0.71 | 0.72 |
| | B2 | **0.94** | 0.67 | 0.79 |
| | B3 | 0.73 | 0.79 | 0.76 |

| EMOTION | Ferg-Db DATASET | | | |
|---------|-----------------|-----|-----|-----|
| | MODEL | P | R | F |
| SAD | SVM | 0.77 | 0.68 | 0.72 |
| | HAAR | 0.65 | 0.77 | 0.70 |
| | HMM | 0.67 | 0.78 | 0.72 |
| | KNN | 0.78 | 0.58 | 0.66 |
| | ESM | 0.75 | **0.89** | 0.81 |
| | RF | **0.90** | 0.85 | **0.87** |
| | B0 | 0.64 | 0.59 | 0.61 |
| | B1 | 0.68 | 0.54 | 0.60 |
| | B2 | 0.50 | 0.84 | 0.62 |
| | B3 | 0.73 | 0.79 | 0.76 |
| HAPPY | SVM | 0.85 | **0.89** | 0.87 |
| | HAAR | 0.81 | 0.69 | 0.74 |
| | HMM | 0.71 | 0.79 | 0.75 |
| | KNN | 0.86 | 0.78 | **0.82** |
| | ESM | **0.90** | **0.89** | **0.90** |
| | RF | 0.88 | **0.89** | 0.88 |
| | B0 | 0.62 | 0.70 | 0.66 |
| | B1 | 0.64 | 0.69 | 0.67 |
| | B2 | 0.81 | 0.79 | 0.80 |
| | B3 | 0.80 | 0.81 | 0.80 |
| NEUTRAL | SVM | 0.76 | 0.84 | 0.80 |
| | HAAR | 0.65 | 0.78 | 0.71 |
| | HMM | 0.87 | 0.78 | 0.82 |
| | KNN | 0.68 | 0.58 | 0.69 |
| | ESM | 0.82 | 0.89 | 0.85 |
| | RF | **0.89** | **0.95** | **0.92** |
| | B0 | 0.84 | 0.81 | 0.82 |
| | B1 | 0.82 | 0.84 | 0.83 |
| | B2 | 0.86 | 0.93 | 0.89 |
| | B3 | 0.78 | 0.81 | 0.79 |

Table 2. Comparison of precision, recall, f1-score with existing approaches for Ferg-db dataset.

As discussed in Table 1.

e-ISSN : 0976-5166
p-ISSN : 2231-3850

| EMOTION | CK+ DATASET | | | |
| --- | --- | --- | --- | --- |
| | MODEL | P | R | F |
| ANGER | SVM | 0.89 | 0.90 | 0.89 |
| | HAAR | 0.92 | 0.88 | 0.90 |
| | HMM | 0.91 | 0.94 | 0.92 |
| | KNN | 0.96 | **0.98** | **0.97** |
| | ESM | 0.92 | **0.98** | 0.94 |
| | RF | 0.82 | 0.89 | 0.85 |
| | B0 | **1** | 0.50 | 0.67 |
| | B1 | 0.97 | 0.97 | **0.97** |
| | B2 | **1** | 0.85 | 0.92 |
| | B3 | 0.89 | 0.81 | 0.86 |
| DISGUST | SVM | 0.85 | 0.88 | 0.86 |
| | HAAR | 0.98 | 0.91 | 0.94 |
| | HMM | 0.88 | 0.98 | 0.93 |
| | KNN | 0.87 | 0.89 | 0.88 |
| | ESM | 0.92 | 0.89 | 0.90 |
| | RF | 0.89 | 0.94 | 0.91 |
| | B0 | **1** | 0.95 | **0.97** |
| | B1 | 0.81 | 0.83 | 0.87 |
| | B2 | 0.96 | **1** | 0.90 |
| | B3 | 0.93 | 0.89 | 0.87 |
| FEAR | SVM | 0.73 | 0.81 | 7.7 |
| | HAAR | 0.83 | **1** | 0.91 |
| | HMM | 0.81 | 0.98 | 0.89 |
| | KNN | 0.96 | 0.88 | 0.91 |
| | ESM | 0.93 | 0.98 | 0.95 |
| | RF | 0.98 | 0.95 | **0.97** |
| | B0 | **1** | 0.90 | 0.95 |
| | B1 | 0.91 | 0.89 | 0.85 |
| | B2 | 0.87 | 0.90 | 0.88 |
| | B3 | 0.89 | 0.91 | 0.93 |
| SURPRISE | SVM | 0.87 | 0.84 | 0.85 |
| | HAAR | 0.81 | **1** | 0.89 |
| | HMM | 0.81 | 0.83 | 0.82 |
| | KNN | 0.88 | 0.98 | 0.93 |
| | ESM | 0.91 | 0.95 | 0.93 |
| | RF | **0.95** | 0.92 | 0.94 |
| | B0 | **0.95** | 0.97 | **0.96** |
| | B1 | 0.83 | 0.81 | 0.82 |
| | B2 | 0.94 | 0.97 | 0.96 |
| | B3 | 0.82 | 0.87 | 0.89 |

| EMOTION | CK+ DATASET | | | |
| --- | --- | --- | --- | --- |
| | MODEL | P | R | F |
| SAD | SVM | 0.87 | 0.86 | 0.86 |
| | HAAR | 0.73 | **0.87** | 0.79 |
| | HMM | 0.88 | 0.78 | 0.86 |
| | KNN | **0.95** | 0.89 | **0.92** |
| | ESM | 0.85 | 0.92 | 0.88 |
| | RF | 0.89 | 0.91 | 0.90 |
| | B0 | 0.45 | 0.51 | 0.61 |
| | B1 | 0.88 | **0.94** | 0.90 |
| | B2 | 0.90 | 0.84 | **0.92** |
| | B3 | 0.93 | 0.89 | **0.92** |
| HAPPY | SVM | 0.89 | 0.79 | 0.83 |
| | HAAR | **1** | 0.95 | **0.97** |
| | HMM | 0.91 | 0.88 | 0.89 |
| | KNN | 0.76 | 0.84 | 0.80 |
| | ESM | 0.88 | 0.90 | 0.89 |
| | RF | 0.86 | 0.90 | 0.87 |
| | B0 | 0.95 | **0.97** | 0.96 |
| | B1 | 0.96 | 0.98 | **0.97** |
| | B2 | 0.81 | 0.89 | 0.84 |
| | B3 | 0.85 | 0.81 | 0.83 |
| NEUTRAL | SVM | 0.87 | 0.84 | 0.85 |
| | HAAR | 0.90 | 0.80 | 0.84 |
| | HMM | 0.95 | 0.87 | 0.90 |
| | KNN | 0.89 | 0.91 | 0.90 |
| | ESM | **0.89** | 0.92 | 0.90 |
| | RF | 0.96 | 0.98 | **0.97** |
| | B0 | 0.94 | **1** | 0.96 |
| | B1 | 0.82 | 0.84 | 0.83 |
| | B2 | 0.86 | 0.93 | 0.89 |
| | B3 | 0.88 | 0.85 | 0.87 |

Table 3. Comparison of precision, recall, f1-score with existing approaches for CK+ dataset.

Given methods are compared with Ref [22], Ref [24] for the precision, recall, f1-score as shown in above Table 3 and as discussed in Table 2. Baseline B2, B4 are Ref [24] normalizes the data per image and per pixel, it also does data augmentation as there is small amount of training data and it uses combination of standard methods.

| EMOTION | JAFFE DATASET | | | |
|---------|---------------|---|---|---|
| | MODEL | P | R | F |
| ANGER | SVM | 0.86 | **0.95** | 0.90 |
| | HAAR | 0.89 | 0.89 | 0.89 |
| | HMM | 0.71 | 0.78 | 0.76 |
| | KNN | 0.87 | 0.88 | 0.87 |
| | ESM | 0.92 | 0.89 | 0.90 |
| | RF | 0.88 | **0.95** | 0.91 |
| | B0 | **1** | 0.71 | 0.83 |
| | B1 | 0.78 | 0.57 | 0.67 |
| | B2 | **1** | 0.85 | **0.92** |
| | B3 | 0.89 | 0.81 | 0.78 |
| DISGUST | SVM | 0.87 | 0.85 | 0.86 |
| | HAAR | 0.77 | 0.86 | 0.81 |
| | HMM | 0.76 | 0.78 | 0.77 |
| | KNN | 0.87 | 0.88 | 0.87 |
| | ESM | **1** | 0.86 | **0.92** |
| | RF | 0.95 | 0.90 | **0.92** |
| | B0 | **1** | 0.85 | 0.91 |
| | B1 | 0.71 | 0.73 | 0.72 |
| | B2 | 0.76 | **1** | 0.80 |
| | B3 | 0.78 | 0.79 | 0.78 |
| FEAR | SVM | 0.83 | 0.84 | 0.83 |
| | HAAR | 0.81 | 0.89 | 0.85 |
| | HMM | 0.71 | 0.88 | 0.78 |
| | KNN | 0.66 | 0.58 | 0.61 |
| | ESM | 0.79 | 0.89 | 0.84 |
| | RF | **0.90** | 0.89 | **0.89** |
| | B0 | 0.71 | **1** | 0.83 |
| | B1 | 0.81 | 0.89 | 0.85 |
| | B2 | 0.87 | 0.90 | 0.88 |
| | B3 | 0.77 | 0.79 | 0.78 |
| SURPRISE | SVM | 0.88 | 0.84 | 0.86 |
| | HAAR | 0.85 | 0.85 | 0.85 |
| | HMM | 0.76 | 0.79 | 0.78 |
| | KNN | 0.81 | 0.68 | 0.74 |
| | ESM | 0.88 | 0.90 | 0.89 |
| | RF | 0.90 | **0.92** | **0.91** |
| | B0 | 0.87 | 0.87 | 0.87 |
| | B1 | 0.73 | 0.78 | 0.78 |
| | B2 | **0.94** | 0.84 | **0.89** |
| | B3 | 0.73 | 0.79 | 0.76 |

| EMOTION | JAFFE DATASET | | | |
|---------|-----------|------|------|------|
| | MODEL | P | R | F |
| HAPPY | SVM | 0.85 | 0.83 | 0.84 |
| | HAAR | 0.81 | 0.90 | 0.85 |
| | HMM | 0.69 | 0.68 | 0.68 |
| | KNN | 0.86 | 0.88 | 0.87 |
| | ESM | 0.82 | 0.89 | 0.85 |
| | RF | **0.90** | 0.88 | 0.88 |
| | B0 | **0.90** | **1** | **0.95** |
| | B1 | 0.74 | 0.69 | 0.71 |
| | B2 | 0.81 | 0.79 | 0.80 |
| | B3 | 0.89 | 0.81 | 0.85 |
| NEUTRAL | SVM | **0.87** | 0.74 | 0.80 |
| | HAAR | **0.87** | 0.75 | 0.81 |
| | HMM | 0.75 | 0.78 | 0.77 |
| | KNN | 0.68 | 0.68 | 0.68 |
| | ESM | 0.86 | 0.90 | 0.88 |
| | RF | **0.87** | 0.94 | **0.90** |
| | B0 | 0.57 | **1** | 0.72 |
| | B1 | 0.85 | 0.84 | 0.84 |
| | B2 | 0.86 | 0.91 | 0.88 |
| | B3 | 0.78 | 0.81 | 0.79 |
| SAD | SVM | 0.87 | 0.65 | 0.74 |
| | HAAR | 0.75 | 0.86 | 0.80 |
| | HMM | 0.78 | 0.74 | 0.76 |
| | KNN | 0.68 | 0.88 | 0.77 |
| | ESM | 0.88 | 0.89 | 0.88 |
| | RF | 0.88 | **0.92** | **0.90** |
| | B0 | **1** | 0.66 | 0.80 |
| | B1 | 0.68 | 0.84 | 0.75 |
| | B2 | 0.70 | 0.64 | 0.67 |
| | B3 | 0.73 | 0.79 | 0.76 |

Table 4. Comparison of precision, recall, f1-score with existing approaches for JAFFE dataset.

As discussed in Table 3

| MODEL | ACCURACY | | | | |
|-------|----------|--------|------|------|------|
| | EMOTION | FER2013 | FERG-DB | CK+ | JAFFE |
| SVM | ANGER | 0.82 | 0.86 | 0.89 | **0.90** |
| | DISGUST | 0.73 | **0.93** | 0.86 | 0.86 |
| | FEAR | 0.77 | 0.69 | **0.93** | 0.83 |
| | SAD | **0.86** | 0.72 | **0.86** | 0.74 |
| | SURPRISE | 0.73 | 0.73 | 0.85 | **0.86** |
| | HAPPY | 0.86 | 0.87 | 0.83 | **0.93** |
| | NEUTRAL | 0.71 | **0.80** | 0.85 | 0.69 |

| MODEL | ACCURACY | | | | |
| | EMOTION | FER2013 | FERG-DB | CK+ | JAFFE |
|---|---|---|---|---|---|
| HAAR | ANGER | 0.63 | 0.79 | **0.86** | 0.72 |
| | DISGUST | 0.58 | **0.82** | 0.73 | 0.73 |
| | FEAR | 0.59 | 0.70 | 0.86 | **0.92** |
| | SAD | 0.8 | 0.70 | 0.79 | **0.80** |
| | SURPRISE | 0.80 | 0.78 | **0.89** | 0.85 |
| | HAPPY | 0.73 | 0.74 | **0.97** | 0.85 |
| | NEUTRAL | 0.81 | 0.71 | **0.84** | 0.81 |
| HMM | ANGER | 0.79 | 0.70 | **0.92** | 0.76 |
| | DISGUST | 0.66 | 0.78 | **0.93** | 0.77 |
| | FEAR | 0.67 | 0.69 | **0.89** | 0.78 |
| | SAD | 0.79 | 0.72 | **0.86** | 0.76 |
| | SURPRISE | 0.76 | 0.63 | **0.82** | 0.78 |
| | HAPPY | 0.75 | 0.69 | **0.89** | 0.68 |
| | NEUTRAL | 0.79 | 0.82 | **0.90** | 0.77 |
| KNN | ANGER | 0.55 | 0.60 | **0.97** | 0.87 |
| | DISGUST | 0.76 | 0.63 | **0.88** | 0.87 |
| | FEAR | 0.78 | 0.66 | **0.91** | 0.61 |
| | SAD | 0.87 | 0.68 | **0.92** | 0.77 |
| | SURPRISE | 0.62 | 0.87 | **0.93** | 0.74 |
| | HAPPY | 0.82 | 0.51 | 0.80 | **0.87** |
| | NEUTRAL | 0.61 | 0.69 | **0.90** | 0.68 |
| ESM | ANGER | 0.85 | 0.88 | **0.94** | 0.90 |
| | DISGUST | 0.86 | 0.89 | 0.90 | **0.92** |
| | FEAR | 0.85 | 0.82 | **0.95** | 0.89 |
| | SAD | **0.89** | 0.81 | 0.88 | 0.88 |
| | SURPRISE | 0.90 | 0.81 | **0.93** | 0.89 |
| | HAPPY | 0.82 | **0.89** | 0.89 | 0.85 |
| | NEUTRAL | 0.86 | 0.85 | **0.90** | 0.88 |

| MODEL | ACCURACY | | | | |
| | EMOTION | FER2013 | FERG-DB | CK+ | JAFFE |
|---|---|---|---|---|---|
| RF | ANGER | **0.91** | 0.89 | 0.89 | **0.91** |
| | DISGUST | 0.86 | 0.92 | **0.91** | 0.92 |
| | FEAR | 0.94 | 0.90 | **0.97** | 0.90 |
| | SAD | 0.89 | 0.87 | **0.9** | **0.90** |
| | SURPRISE | 0.92 | 0.91 | **0.94** | 0.91 |
| | HAPPY | **0.89** | 0.88 | 0.87 | **0.89** |
| | NEUTRAL | 0.93 | 0.92 | **0.97** | 0.90 |

Table 5. Comparison of accuracy for different datasets.

## 7. Conclusion

In the given paper facial expression recognition is done by ensemble learning. Five different methods are implemented in this study, which are SVM, Haar cascade, HMM, KNN and Random Forest for the Fer-2013, Ferg-db, CK+, JAFFE dataset and found an average accuracy of 79% all the methods on Fer-2013 dataset, average accuracy of 78.5% on Ferg-db dataset, average accuracy of 89.3% on CK+ dataset and average accuracy of 82.2% on JAFFE dataset. Comparisons of precision, recall, f1-score with existing research is done in the paper. Also, Comparisons of results with existing research is done in the paper. On evaluating models found that random forest performs well as compared to other models given. The given model is capable of detecting emotions of people from live webcam i.e., it can detect expression from live webcam. Comparisons lead to finding that CK+ dataset gives maximum accuracy due to it contains small number of images. While Fer2013 accuracy may vary due to imbalanced dataset i.e., it contains large number of images for one emotion class and a smaller number of images for different emotion class.

## 8. Future Scope

In this study, as facial expression recognition is an elementary process, many tremendous potentials lie untapped in this domain, like to explore more facial and image features. Also, there would be an exploration on recognizing emotions in color images and performing these predictions across in a video. Also, one can explore the more detailed scope of deep learning in FER. Facial expression recognition would form a vital part for a multimodal system for emotion recognition as well.

## References

[1] Du S, Tao Y, Martinez AM. (2014): Compound facial expressions of emotion. Proceedings of the National Academy of Sciences of the United States of America, Vol. 111, No. 15, E1454–E1462.

[2] Pantic M, Rothkrantz (2000). LJM: Automatic analysis of facial expressions: the state of the art. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 12, 1424-1445.

[3] Anagha Kulkarni, Isha Talegaonkar, Kalyani Joshi, Shreya Valunj, Rucha Kohok. (2019).: Real Time Facial Expression Recognition using Deep Learning. International Conference on Communication and Information Processing

[4] Anastasius Roussos, Giorgos Giannakakis, Luma Alharbawee, Mohammad Rami Koujan, Nicolas Pugeault. (2020).: Real Time Facial Expression Recognition "In The Wild" by Disentangling 3D expression from Identity. Institute of Computer Science (ICS). [Online]. Available: https://arxiv.org/pdf/2005.05509 (current February 2021)

[5] Chitra Venkata Rami Reddy, Kolli Venkata Krishna Kishore, Uyyala Srinivasulu Reddy. (2019).: Facial Emotion Recognition Using NLPCA and SVM. International Information and Engineering Technology Association, Vol. 36, 13-22.

[6] Chris English, Dan Duncan, Gautam Shine: Facial Emotion Recognition in Real Time. Stanford University (2016). [Online]. Available: http://cs231n.stanford.edu/reports/2016/pdfs/022_Report.pdf (current February 2021)

[7] Matias Valdenegro, Octavio Arriaga, Paul G. Plunger. (2017).: Real Time Convolutional Neural Networks for Emotion and Gender Classification. Sankt Augustin Germany. [Online]. Available: https://arxiv.org/abs/1710.07557 (current March 2021)

[8] Amirali Abdulrashid, Mehdi Minaei, Shervin Minaee . (2021): Deep-Emotion: Facial Expression Recognition Using Attentional Convolutional Network. University of California, Riverside. [Online]. Available: https://www.mdpi.com/1424-8220/21/9/3046/pdf (current February 2021)

[9] Grant Sivesind, Guilherme Reis, Minh-An Quinn. (2017): Real time Facial Recognition From Facial Expressions. CS 229- Stanford University. [Online]. Available: http://cs229.stanford.edu/proj2017/final-posters/5147982.pdf (current March 2021)

[10] J. Jayapradha Soumya Sharma and Yash Dugar.(2018): Detection and Recognition of Human Emotion using Neural Network. International Journal of Applied Engineering Research, Vol. 18, 6472-6477.

[11] A.K. Singh, A. Elchouemi, Abeer Alsadoon, D. Yang, P.W.C. Prasad. (2017): An Emotion Recognition Model Based on Facial Recognition in Virtual Learning Environment. 6th International Conference on Smart Computing and Communications

[12] Alex Martinet, Shinchuan Du, (2013): A Model of the Perception of Facial Expressions of Emotion by Humans: Research Overview and Perspectives. Journal of Machine learning Research, 1589-1608.

[13] C Raghavendra, MalyalaDivya, R Obula Konda Reddy. (2019): Effective Facial Emotion Recognition using Convolutional Neural Network Algorithm. International Journal of Recent Technology and Engineering, Vol.8,

[14] Abe Kazemzadeh, Carlos Busso, Chul Min Lee, Murtaza Bulut Sungbok Lee, Serdar Yildirim, Shrikanth Narayanan, Ulrich Neumann, Zhigang Deng. (2004): Analysis of Emotion Recognition using Facial Expressions, Speech and Multimodal Information. Emotion Research Group. [Online]. Available: https://ecs.utdallas.edu/research/researchlabs/msplab/publications/Busso_2004_poster.pdf (current March 2021)

[15] Nithya Roopa S. (2019).: Emotion Recognition from Facial Expression using Deep Learning. International Journal of Engineering and Advanced Technology, Vol.8

[16] Alireza Jolfaei, Hongli Zhang and Mamoun Alazab. (2019): A Face Emotion Recognition Method Using Convolutional Neural Network and Image Edge Computing. IEEE Access. [Online]. Available: https://researchers.mq.edu.au/files/112802715/112715760.pdf (current March 2021)

[17] Mahesh M. Goyani, V Mistry, (2013): A literature survey in Facial Expression Recognition using Global features. Semantic Scholar. [Online]. Available: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.645.5162&rep=rep1&type=pdf (current March 2021)

[18] Jeffrey F. Cohn, Jason Saragih1, Patrick Lucey, Takeo Kanade, Zara Ambadar. (2010): The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression. [Online]. Available: https://www.computer.org/csdl/pds/api/csdl/proceedings/download-article/12OmNzZ5olI/pdf (current February 2021)

[19] Asif Khan and Adnan Zaidi, Fayyaz Ali, Irfan Ali Kandhro, Sahar Zafar Jumani, Subhash Guriro, (2019): Facial Expression Recognition with Histogram of Oriented Gradients using CNN. Indian Journal of Science and Technology, Vol.12, No. 24.

[20] Anders Skibeli Rokkones(2018): Real-Time Face Detection Identification and Emotion Analysis for Affective Human Robot Interaction. University of Oslo. [Online]. Available: https://www.duo.uio.no/bitstream/handle/10852/66497/master_thesis_final.pdf?sequence=1&isAllowed=y (current March 2021)

[21] Tao Zhao, Wen-Di Cai, Xin-Mei Wang, Yong-Bo Li, Zhong-Zhao Xie(2018): Convolutional Neural Networks for Facial Expression Recognition with Few Training Samples. Conference Paper. [Online]. Available: https://arxiv.org/pdf/1804.08348 (current March 2021).

## Authors Profile

**Dr. Nilesh Shelke** is a senior faculty at Computer Science & Engineering Department and Dean Academics at Priyadarshini Indira Gandhi College of Engineering, Nagpur. He holds MCA and M. Tech., M. Phil., PhD degrees in Computer Science & Engineering. He is Microsoft Certified Solution Developer and has more than 22 years of experience of imparting IT Training which includes learners from different streams, faculties and Microsoft Certifications to corporate employees. He has strong contribution in syllabus framing of various subjects at the university level. He has published patents, copyrights and sellable technical articles in the renowned journals. He has successfully completed many certifications from IIT through NPTEL including Machine Learning, Deep Learning, Data Science, Databases etc. He is also the resource person at the various events and authored a book *Introduction to Machine Learning*.

**Dr. Sachin Upadhye** is an Assistant Professor at Computer Application Department of Shri Ramdeobaba College of Engineering and Management, Nagpur. He holds MCA, and M. Tech. & PhD degrees in Computer Science & Engineering. He is an Oracle Certified Associate, IBM Rational Certified developer and has more than 20 research paper published in National & International Journals and conferences. He has 12 Years teaching experiences. He has successfully completed many certifications from NPTEL, IIT and also completed the research proposal and project.

**Satyajit Sidheshwar Uparkar** is a certified data scientist from International School of Engineering, Hyderabad. He is working as Assistant Professor in Department of Computer Application at Shri Ramdeobaba College of Engineering and Management, Nagpur, since last 11 years. He is a triple post graduate and expert in the domain of data analytics. His research area includes Data Mining, Scalable Data Science and Operation Research Modelling. He has bagged two best paper awards in International conferences. His research contribution includes 30 research papers in various National and International Journals. He has bagged two best paper awards in International conferences. He is a data science consultant and has provided consultancy to the local companies.

**Mr. Hemant G. Bawane** is a research scholar at Priyadarshini Indira Gandhi College of Engineering, Nagpur. His research lies in the domain of Image Processing. He is certified in Microsoft Azure Data Fundamentals. He is working at Celebal Technologies Private Ltd. as an Associate Data Engineer Intern. He has completed IIT certified courses Machine Learning, Deep Learning through NPTEL.

**Mr. Danish Shivvanshi** is a research scholar at Computer Science & Engineering department of Priyadarshini Indira Gandhi College of Engineering, Nagpur. He is working in the field of Image Processing. He is also an intern at Persistent Systems Pvt. Ltd. He has successfully completed certificate course in Python with Data Science, Machine Learning with TensorFlow from IIT Madras through NPTEL.

**Sujeet Kosekar** is research scholar at Computer Science & Engineering Department of Priyadarshini Indira Gandhi College of Engineering, Nagpur. He has completed different internships and real-time projects in python language. He has successfully completed many certifications from IIT through NPTEL including Machine Learning, Python etc.