

Real-time Facial Recognition using Haar Cascade

Vibhor Chandra Gupta¹, Raj Bhagwani²

Bachelors of Technology Student, Department of Computer Engineering, Mukesh Patel School of Technology Management and Engineering, NMIMS University, Mumbai, Maharashtra, India¹

Bachelors of Technology Student, Department of Computer Engineering, Mukesh Patel School of Technology Management and Engineering, NMIMS University, Mumbai, Maharashtra, India²

ABSTRACT: Facial Recognition is a category of biometric software that maps an individual's facial features mathematically and stores the data as a faceprint. The software uses algorithms to compare a live capture or digital image to the stored faceprint in order to identify individuals present in the frame. The Human face recognition procedure is twofold, namely face detection followed by recognition. In the first stage, the objective is to locate the position of the face (region of interest), separating the ROI from the background. A training model is then Developed from the dataset which acts as a model for facial image recognition. There are three kinds of methods that are currently popular in developed face recognition pattern namely, Eigenface method, Fisherface method and LPBH(Local Binary Pattern Histogram) method. This paper proposes an automatic and robust method to detect human faces from the background in real time, and is capable of processing images rapidly while achieving high detection rates from video sequences. Highlight of the face detection system is to identify and locate all faces regardless of their position, scale, orientation, lighting conditions, expressions.

KEYWORDS: Eigenface, Face Detection, Face Recognition, Fisherface and Local Binary Pattern Histogram Method.

I. INTRODUCTION

Human face detection has been playing an important part in human-machine interaction and computer vision-based applications. A humans' face identity is unique and non-replicable. Face detection in image sequence has been an active research area in the computer vision field in the past and recent years due to its potential applications such as monitoring and surveillance, human computer interfaces, smart rooms, intelligent robots and biomedical image analysis.

Face detection algorithms begin by searching for human eyes. Eyes constitute what is known as a valley region and are one of the easiest features to detect. Once eyes are detected, the algorithm might then attempt to detect facial regions including eyebrows, the mouth, nose, nostrils, and the iris. Once the algorithm surmises that it has detected a facial region, it can then apply additional tests to validate whether it has, in fact, detected a face.

The detection methods learns statistical models of face and non-face images, and then applies the two-class classification rules to discriminate between face and non-face patterns. Occlusion and lighting distortions and illumination conditions can also change the overall appearance of face.

This paper, proposes an automatic and robust method to detect human faces from Video sequences. Each frame of the variables that were considered while developing the algorithm include variations in background illumination, expressions, and other uncertainties.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 11, November 2019

Local binary pattern (LBP) for classification of image textures, is a grayscale irrelevant texture operator. It has powerful discrimination and has also been used in human face detection and recognition. However, although LBP features have great discriminative power, they miss the local structure under some certain circumstances. A cascade classifier algorithm, such as Haar features for human face, in conjunction with using detection through haar-transforms is found to be a more efficient way.

II. RELATED WORK

Background

Much of the work in computer recognition of faces has focused on detecting individual features such as the eyes, nose, mouth, and head outline, and defining a face model by position, size, and relationships between these features. Such approaches have proven difficult to extend to multiple views, and have often been quite fragile, requiring a good initial guess to guide them. Research in human strategies of face recognition, moreover, has shown that individual features and their immediate relationship comprise an insufficient representation to account for the adult human face recognition. Nonetheless, this approach to human face recognition remain the most popular one in the computer vision literature.

About OpenCV

Open Source Computer Vision is a popular computer vision library started by Intel in 1999. It uses machine learning algorithm to search for faces within a picture based on matching multiple patterns and features.

A face typically has about 5000 classifiers, all of which must match for a face to be detected. Due to this, 5000 or more tests need to be conducted per block, which results in a high calculation complexity. OpenCV uses cascades, which breaks the problem of detecting faces into multiple stages and performs a detailed test for each block. The algorithm may have 30 to 50 of these stages or cascades, and it will only detect a face if all stages pass.

III. METHODOLOGY

Face detection uses classifiers, which are algorithms that detects what is either a face(1) or not a face(0) in an image. Classifiers have been trained to detect faces using thousands to millions of images in order to be accurate. There are 2 types of classifiers in OpenCV namely, LBP (Local Binary Pattern) and Haar Cascades. We shall be focussing on the latter in this paper.

Our approach for face recognition can be represented as 3 well-defined steps –

- Training Data Gathering: Gather face data (face images in this case) of the persons you want to recognize
- Training of Recognizer: Feed that face data (and respective names of each face) to the face recognizer so that it can learn.
- Recognition: Feed new faces of the persons and see if the face recognizer you just trained recognizes them.

We will be using 3 different face recognizers, along with haarcascades –

- Eigenface recognizer
- Fisherface recogniser
- LBPH face recognizer

Haar Cascades

This approach is based on the Haar Wavelet technique to analyse pixels in the image into squares by function. It is a machine learning approach where a cascade function is trained from multiple positive and negative images from the training data. Haar feature extraction filters which are as shown below, are used to detect various informative components of the face which help define the features of the particular individual.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 11, November 2019

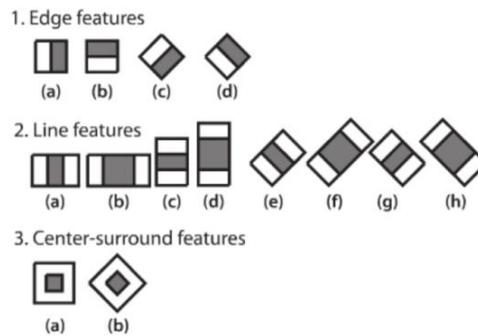


Fig. 1. Face Features (1)

Each feature is a single value obtained by subtracting sum of pixels under the white rectangle from sum of pixels under the black rectangle.

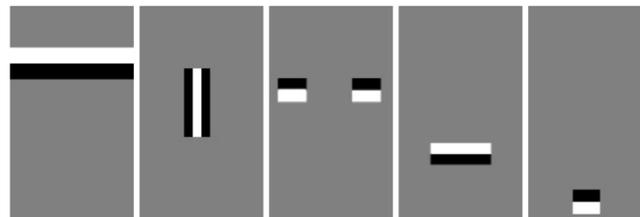


Fig. 2. Face Features (2)

We apply each feature on all the training images. For each feature, it finds the best threshold which will classify the faces to positive and negative. Features with minimum error rate are selected which accurately classify the face and non-face images.



Fig. 3. Haar Filter Overlay on Face

In an image, most of the image is non-face region. So, it is a better idea to have a simple method to check if a window is not a face region. If it is not, discard it in a single shot, and don't process it again. Instead, focus on regions where there can be a face. This way, we spend more time checking possible face regions.

Instead of applying all thousands of features on a window, Cascade of Classifiers are used, which groups features into different stages of classifiers and then these classifiers are applied one-by-one. If a window fails the first stage, it is

discarded along with its remaining features. If it passes, the second stage of features is applied. This process is continued. The window which passes all stages is finally classified as a face region.

Local Binary Pattern Histogram Face Recognizer(LBPH)

Local Binary Pattern (LBP) is a simple yet very efficient texture operator which labels the pixels of an image by thresholding the neighbourhood of each pixel and considers the result as a binary number.

The LBPH Face Recognizer Process –

Take a 3×3 window and move it across one image. At each move (each local part of the picture), compare the pixel at the center, with its surrounding pixels. Denote the neighbours with intensity value less than or equal to the center pixel by 1 and the rest by 0.

After you read these 0/1 values under the 3×3 window in a serial order, you will have a binary pattern like 10001101 that is local to a particular area of the picture. When you finish doing this on the whole image, you will have a list of local binary patterns.

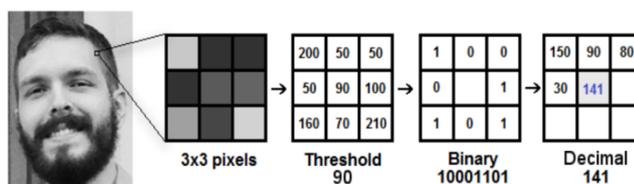


Fig. 4. Sample 3x3 Grid of a Recognizer

LBP conversion to binary. Source: López & Ruiz; Local Binary Patterns applied to Face Detection and Recognition.

After we get a list of local binary patterns, you convert each one into a decimal number using binary to decimal conversion (as shown in above image) and then you make a histogram of all of those decimal values. A sample histogram looks like this:

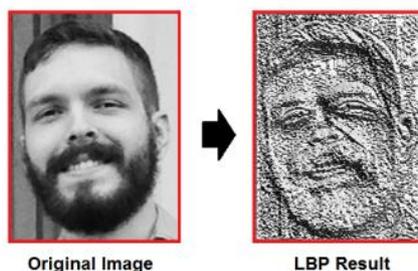


Fig. 5. Output Of LBP From Original Image

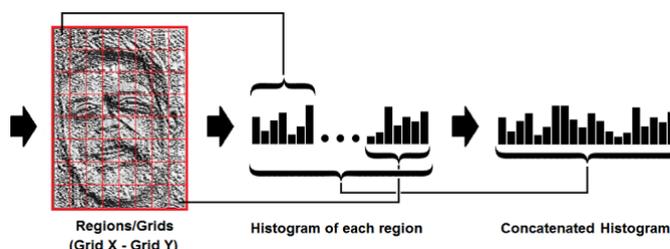


Fig. 6. Histogram Representation Of LBP Result

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 11, November 2019

In the end, you will have one histogram for each face in the training data set. That means that if there were 100 images in the training data set then LBPH will extract 100 histograms after training and store them for later recognition. The algorithm also keeps track of which histogram belongs to which person.

Both Eigenfaces and Fisherfaces are affected by light and in real life, perfect light conditions are not always available. LBPH face recognizer is an improvement to overcome this drawback.

Below is a group of faces and their respective local binary pattern images. You can see that the LBP faces are not affected by change in light condition.

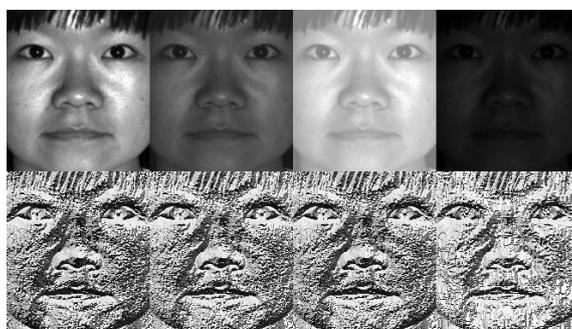


Fig. 7. Effect of Lighting on LBP faces

Other Methods

1. EigenFace

Eigenface Recognizer algorithm considers the fact that not all parts of a face are equally important and equally useful. An individual is recognized by their distinct features like eyes, nose, cheeks, forehead and how they vary with respect to each other. Areas of maximum change (i.e. change in variance) of the face, are of primary focus. There is a significant change from the eyes to nose and same is the case from nose to mouth. When multiple faces are compared, these are the parts of the faces are of utmost important. These components are useful since they catch the maximum change among faces, changes that help us differentiate one face from another.



Fig. 8. Amount of Variation in Faces

Eigen Faces Face Recognizer looks at all the training images of all the persons as a whole and tries to extract the components that have the maximum variance and discards the rest of the components. The extracted components are called principal components. EigenFaces Face Recognizer trains itself (by extracting principal components). Another point to be considered is that it also keeps a record of which principal component belongs to which person.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 11, November 2019

One thing to note is that EigenFaces algorithm considers illumination as an important feature. In consequence, lights and shadows picked up, are stored and classified as the feature of the face.

2. Fisherface

FisherFaces Face Recognizer algorithm is an improved version of EigenFaces. It looks at all the training faces of all the persons at once and finds principal components from all of them combined. By capturing principal components from all the of samples together, we focus on the features that represent all the persons in the training data.

The FisherFaces Face Recognizer approach has multiple drawbacks, for example, images with sharp changes (like light changes which are not a useful feature at all) may dominate the rest of the images and one may end up with features that are from an external source like light and are not useful for discrimination at all. Eventually, the principal components will represent light changes and not the actual facial features.



Fig. 9. Principal Components Of Samples

One thing to note here is that FisherFaces only prevents features of one person from becoming dominant, but it still considers illumination changes as a useful feature. We know that light variation is not a useful feature to extract as it is not part of the actual face.

IV.IMPLEMENTATION AND OVERALL DESIGN

Approach

Face recognition system includes four main parts: information acquisition module, feature extraction module, classification module and training classifier database module. The image information collected by the information acquisition module will be used as a test sample for analysis. In the feature extraction module, a series of salient features which can represent human identity information are extracted and analysed. In the classification module, the classifier trained by database is used to classify the test samples to determine the identity information of the samples.

The system flow chart is shown in Fig. 1. A. Face Detection OpenCV provides a Haar cascade classifier, which can be used for face detection. The Haar cascade classifier detects multiple facial organs including the eye, nose, and mouth. First, it reads the image to be detected and converts it into gray image, then loads. Haar cascade classifier to judge whether it contains human face. If so, it continues to examine the eyes, nose, and mouth and draw a rectangular frame on the corresponding organ. If not, it continues to test the next picture. The LBP operator is used to describe the contrast information of a pixel to its neighbourhood pixels.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 11, November 2019

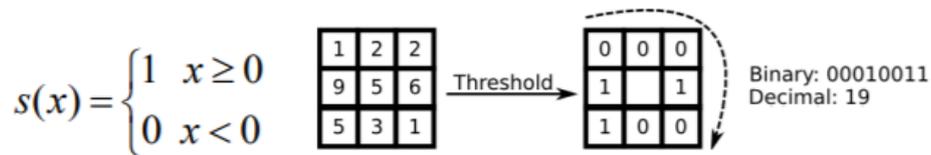


Fig. 10. Applying Threshold To Get A Binary Output

The LBPH algorithm uses the histogram of LBP characteristic spectrum as the feature vector for classification. It divides a picture into several sub regions, then extracts LBP feature from each pixel of sub region, establishing a statistical histogram of LBP characteristic spectrum in each sub region, so that each sub region can use a statistical histogram to describe the whole picture by a number of statistical histogram components. The advantage is to reduce the error that the image is not fully aligned in a certain range.

Flow Chart

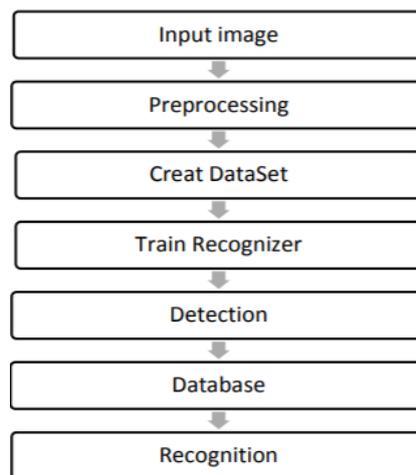


Fig. 11. System Block Diagram

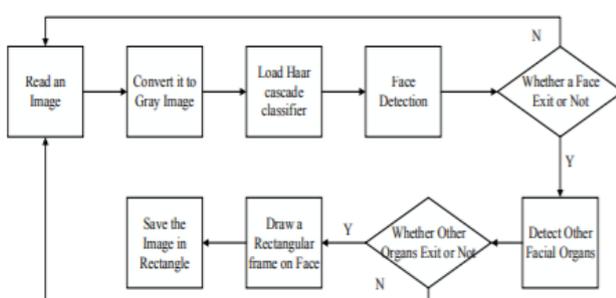


Fig. 12. Face Detection Flow Chart

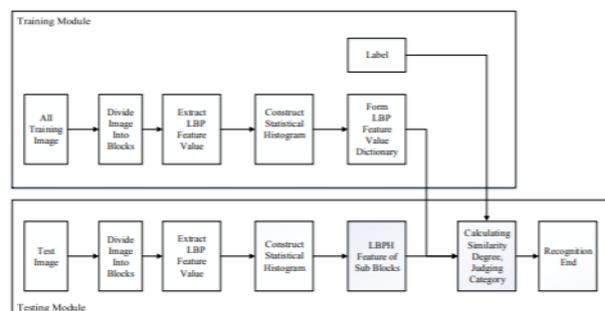


Fig. 13. LBPH Algorithm

V. EXPERIMENTAL OUTCOMES

Input and Output Images

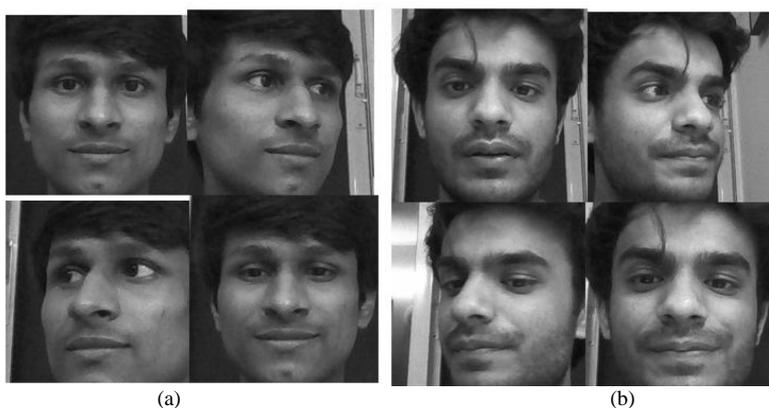


Fig. 14. (a) Input Subject-1, (b) Input Subject-2

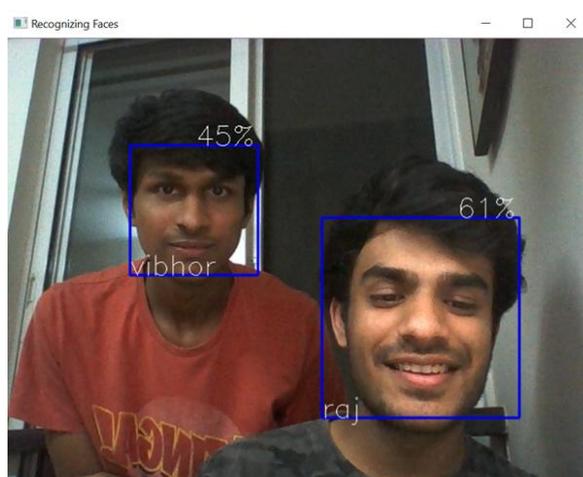


Fig. 15. Output

Figures and Tables

1. First observation table: Depict how variation in number of samples affect the accuracy and time factor of the algorithm.

Constants –

- Distance from the camera: 50cm
- Scale ratio: 1.2
- Minimum neighbours required: 5

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 11, November 2019

Table 1 – First Observation Table

No of samples	Match Percent	Training Time(s)	Capture Time(s)
10	44%	0.71	2.48
20	55%	1.19	4.48
50	65%	3.80	9.15
70	69%	4.43	13.73
100	69%	5.86	17.99

Conclusions from this table –

- As the number of samples decreases the amount of time to capture them and train them decreases too.
 - The capture and train time, both increases linearly along with the number of samples taken.
 - At lower number of samples we observe that the match percent ranges between 44% to 55%. As we approach 70 samples the match percentage goes up to 70%. There is no further increase in the match upon increasing the number of samples.
 - At 50 samples we get an optimal blend of match percent and low capture and train time.
2. Second observation table: Depicts how different values of scaling factor affects the recognition. It gives us the range of distance under which a face can be detected.

Constants –

- minimum numbers of neighbours required: 5
- number of samples taken: 50
- scaling factor > 1 [required]

Table 2 – Second Observation Table

Scaling factor	Max Distance	Match Percent	Tilt Range (degrees)
1.1	137.5	7%	NP
1.3	123	17%	10
1.5	115	21%	10
2.0	96	34%	30
2.5	75	41%	30

Conclusions from this table –

- A greater value of scaling factor can detect a face at a greater distance, however, the accuracy of the algorithm at this distance is very unreliable.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 11, November 2019

- If the value of scaling factor is above 2.0 faces won't be detected from a distance. Erratic behaviour is observed for close by faces. The tilt angle at the maximum distance is greater.
 - We can conclude that an optimum blend for our application can be formed at scale factor 1.3
3. Third observation table: Depicts how tilting face in various direction affect the outcome and match percentage during the recognition of the face.

Constants –

- Number of constants: 50
- Scaling factor: 1.2
- Minimum number of neighbours: 5

Table 3 – Third Observation Table

Tilt axis	Tilt angle (degree)	Match Percent	Max Distance
X	0	68%	118
	30	47%	54
	45	42%	48
Y	0	68%	120
	45	53%	78
	70	42%	67
Combination	NE	39%	42
	NW	42%	49
	SE	43%	46
	SW	42%	48

Conclusions from this table –

- The dataset contained samples having varying face profiles (angles) of the subject.
- This enabled us to get better and accurate results. When the experiment was performed with dataset consisting of only a single profile, the results were dis-satisfactory and inadequate.
- Faces are unable to be detected when the angle of tilt in x direction is more than 50 degrees and when the angle of tilt in y direction is more than 75 degrees.
- Hence, we can see that, when the dataset has a spread of different angles of the face, it performs better.

4. Fourth observation table: This observation table shows us how the output is affected if alter the value of minimum neighbours required parameter of this algorithm. Everything else except minimum neighbour parameter is kept constant while making this observation.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 11, November 2019

Table 4 – Fourth Observation Table

Minimum Neighbour	Observation
2	When the value in minimum neighbour field was taken as 2 the application detected anything which was remotely close to a face.
8	When the value of minimum neighbour was taken as 8 it was hard for the application to identify actual faces as faces too

Conclusions from this table –

- A value of 5 for our algorithm, will work best.
5. Fifth Observation Table: This table gives us information about the outputs when the images in the sample dataset were captured with different types

Table 5 – Fifth Observation Table

Dataset Type	Observation
Emoting	Better recognition while the subject was speaking in conversation, gives a more realistic model
Occluded Face	Blurry results since the face features could not be captured accurately to train the model
Different Angles	Helped in more successful recognition when the subject was not directly facing the camera

Conclusions from this table –

- The dataset should have a mix of the above types of samples, so as to give us overall improvement.
6. Sixth observation table: Depicts the output when we vary multiple parameters together.

Table 6 – Sixth Observation Table

Number of Samples	Scaling factor	Min Neighbours	Identified correctly
20	2	8	58%
40	1.8	7	72%
50	1.4	6	84%
70	1.2	5	92%

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 11, November 2019

We checked each test case with 50 times with our face being at slightly different positions and calculated the number of times it successfully recognised the person.

VI. LIMITATIONS

Face recognition is unpredictable since faces tend to vary in their orientation, expressions and lighting conditions.

Face recognition performance could be drastically be affected by lighting, orientation, scale. Features recognized as parts of the face may sometimes be of some background object. This results in erroneous outputs, devising a strategy to mitigate this still remains a primary concern.

In the practical application of face-recognition technology, the effect of face-recognition is decided not only by static face- recognition algorithm, but also by dynamic face-recognition algorithm. The hardware of the image acquisition equipment and face-recognition algorithm assembly will also affect the actual recognition rate and recognition effect.

VII. CONCLUSION

Our approach distinctly captures the most prominent features within the given facial images, so as to uniquely identify individuals amongst the given set. This paper proposes a 3-phase facial recognition method, using Haar cascade classifiers and Local Binary histogram pattern recognizer, to help detect a face and isolate it from the background. Observations were made about the recognition capacity of the algorithm at the output on changing various input parameters.

In the future, we need more study for efficient construction of training model. Along with the Haar cascade classifiers using we can use additional classifiers, to help give a more intricate and holistic approach to define the faces in an image. Additional weak classifiers are subsequently appended to the primitive Haar-like features based cascaded classifiers.

Possible classifiers can be based on –

- Human skin hue histogram matching.
- Feature detection relative to Eye position
- A more progressive mouth and nose detector

REFERENCES

- [1] Ali Sharifara, Mohd Shafry Mohd Rahim and Yasaman Anisi- A General Review of Human Face Detection Including a Study of Neural Networks and Haar Feature-based Cascade Classifier in Face Detection
- [2] Yea-Shuan Huang and Suen-Yu Chen - A geometrical-model-based face recognition
- [3] XueMei Zhao, ChengBing Wei -A Real-time Face Recognition System Based on the Improved LBPH Algorithm
- [4] Hyung-Il Kim, Seung Ho Lee, and Yong Man Ro- Investigating Cascaded Face Quality Assessment for Practical Face Recognition System
- [5] Dr. H. S. Fadewar-Face Recognition: A Holistic Approach Review
- [6] Teddy Mantoro, Media A. Ayu-Multi-Faces Recognition Process Using Haar Cascades and Eigenface Methods
- [7] Jalendu Dhamija, Tanupriya Choudhury, Praveen Kumar, Yogesh Singh Rathore - An advancement towards efficient Face Recognition using live video field