Intelligent Face Recognition System

Submitted By

Riddhi Sarsavadia

16MCEN17

DEPARTMENT OF INFORMATION TECHNOLOGY

INSTITUTE OF TECHNOLOGY

NIRMA UNIVERSITY

AHMEDABAD-382481

May 2018

Intelligent Face Recognition System

Major Project

Submitted in partial fulfillment of the requirements

for the degree of

Master of Technology in Computer Science and Engineering (Networking Technologies)

Submitted By

Riddhi Sarsavadia

(16MCEN17)

Guided By

Assitant Prof. Usha Patel

DEPARTMENT OF INFORMATION TECHNOLOGY

INSTITUTE OF TECHNOLOGY

NIRMA UNIVERSITY

AHMEDABAD-382481

May 2018

Certificate

This is to certify that the major project entitled **"Intelligent Face Recognition System"** submitted by **Riddhi Sarsavadia (16MCEN17)**, towards the partial fulfillment of the requirements for the award of degree of Master of Technology in Computer Science and Engineering (Netwoking Technologies) of Nirma University, Ahmedabad, is the record of work carried out by her under my supervision and guidance. In my opinion, the submitted work has reached a level required for being accepted for examination. The results embodied in this major project part-II, to the best of my knowledge, haven't been submitted to any other university or institution for award of any degree or diploma.

Prof Usha Patel
Guide & Assistant Professor,
CE / IT Department,
Institute of Technology,
Nirma University, Ahmedabad.

Dr. Gaurang Raval
Associate Professor,
Coordinator M.Tech - CSE (Networking Technologies)
Institute of Technology,
Nirma University, Ahmedabad

Dr. Madhuri Bhavsar
Professor and Head,
IT Department,
Institute of Technology,
Nirma University, Ahmedabad.

Dr Alka Mahajan
Director,
Institute of Technology,
Nirma University, Ahmedabad

Certificate

This is to certify that the major project entitled**Intelligent Face Recognition System** submitted by **Riddhi Sarsavadia (16MCEN17)**, towards the partial fulfillment of the requirements for the award of degree of Master of Technology in Computer Science and Engineering(Networking Technologies) of Nirma University, Ahmadabad, is the record of work carried out by her under my supervision and guidance. In my opinion, the submitted work has reached a level required for being accepted for examination.

Mr. Deepak Sharma
Dy. Manager,
IT Department.
J K Lakshmi Cement,Kalol.

Statement of Originality

I, **Riddhi Sarsavadia**, **16MCEN17**, give undertaking that the Major Project entitled "**Intelligent Face Recognition System**" submitted by me, towards the partial fulfillment of the requirements for the degree of Master of Technology in **Computer Science & Engineering (Networking Technologies)** of Institute of Technology, Nirma University, Ahmedabad, contains no material that has been awarded for any degree or diploma in any university or school in any territory to the best of my knowledge. It is the original work carried out by me and I give assurance that no attempt of plagiarism has been made.It contains no material that is previously published or written, except where reference has been made. I understand that in the event of any similarity found subsequently with any published work or any dissertation work elsewhere; it will result in severe disciplinary action.

_____

Signature of Student

Date:

Place:

Endorsed by

Asst.Prof.Usha Patel

(Signature of Guide)

Acknowledgements

It gives me immense pleasure in expressing thanks and profound gratitude to **Prof. Usha Patel**, Assistant Professor, Computer Engineering Department, Institute of Technology, Nirma University, Ahmedabad for her valuable guidance and continual encouragement throughout this work. The appreciation and continual support she has imparted has been a great motivation to me in reaching a higher goal. Her guidance has triggered and nourished my intellectual maturity that I will benefit from, for a long time to come.

It gives me an immense pleasure to thank **Dr. Madhuri Bhavsar**, Hon'ble Head of Information Technology Department, Institute of Technology, Nirma University, Ahmedabad for her kind support and providing basic infrastructure and healthy research environment.

A special thank you is expressed wholeheartedly to **Dr. Alka Mahajan**, Hon'ble Director, Institute of Technology, Nirma University, Ahmedabad for the unmentionable motivation he has extended throughout course of this work.

I would also thank the Institution, all faculty members of Computer Engineering Department, Nirma University, Ahmedabad for their special attention and suggestions towards the project work.

- Riddhi Sarsavadia

16MCEN17

Abstract

Features of Human face can be used to identify human uniquely. This make Face recognition system(FRS) popular for many applications like security verification at gate in many organizations, for access control of confidential resources, identifying intruders by national defence, and many more. With time, researchers and practitioners are putting efforts to rectify and optimize Intelligent Face Recognition System for different perspective, e.g. optimizing for accuracy, time, and space complexity, accuracy for facial expression change with time, face captured at some degree of orientation, lightning condition,occlusions etc.. As a result, many algorithms are available for face recognition system. This dissertation work focus to optimize IFRS for four criteria: i) face captured at more than 45 orientation, ii) person wearing eyeglass of different shape and size. iii) Face recognition using IP camera. Hence, we name it as "Intelligent face recognition system".

Abbreviations

PCA	Principal Component Analysis.
LDA	Linear Discriminant Analysis.
ICA	Independent Component Analysis.
LBP	Linear Binary Patterns.
ASM	Active Shape Model.
EBGM	Elastic Bunch Graph Machine.
CNN	Convolution Neural Network.
RBF	Radial Basic Function.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Our faces are very complex items with facial components which can vary after some time. Humans have a very good capacity to recognize several faces learned throughout their life-time and can identify faces in milliseconds, even after years of separation, but the system does not have the same capacity. So, for this purpose, we need to create a face recognition system to make our system as intelligent as humans. Face recognition by systems can be signicant and has a huge amount of usage for security purpose for organizations, access control for higher level authorities, national defenses[1] etc. Face recognition incorporates fundamentally three-errand face detection, feature extraction and recognition.[1]

Figure 1.1: Face Recognition System

The first step of this process is face detection has its own challenges which are posture invariant,the existence or deciency of physical components, facial expression,occlusion,image orientation and imaging conditions [2].After detection of a face from an image,the next step would be feature extraction in which is used to extract only features that would be used for mapping with image database for face recognition. For face recognition step we have mainly three approaches which are the holistic approach, feature based and hybrid

approach.[1]

Using Face recognition approaches, we can get recognition rate of frontal face, oriented face(up to 30) and props(beard and eyeglasses with an existing database) up to 80 to 85%. To overcome these challenges we use hybrid approaches for face recognition. When we recognize face with eyeglasses using face recognition techniques, there will be some problems due to the thickness of eyeglasses, the size of a frame of eyeglasses,reective property of glasses etc. So, we use eyeglass removal[1] to remove eyeglasses after detecting a face. In this project work, the whole implementation would be done using OpenCV(Open Computer Vision) [3].

OpenCV(Open source Computer Vision) is a set of libraries with programming functions which has the main objective of real-time image processing [4].OpenCV was developed by Intel's research center by Willow Garage in 1999 at Nizhny Novgorod(Russia)[3].It is free for both commercial and non-commercial use. It is used for developing highly ecient computer vision applications.

1.1 Motivation

Due to advancement in technologies the method to sign in a register for official attendance in office, college etc was then substituted by the Password/PIN numbers. Through personal identification only a person could access something. Also, the token system supported by RFID technology is still also used, where you need a card to access something. Fraudulent, falsification, robbery, etc. are the potential misuses of such technology, biometric identification created abuzz due to its extremely secure interface. Fingerprints are an example of biometric but it requires clean hands and when ngers might be injured this will not work. So, face recognition system, iris scan, speech recognition etc are used for identication of a person. The main goal of face recognition system is to recognize authorized person within seconds.

1.2 Objective

The primary objective of this project is to develop highly efficient algorithm related to low computational techniques which would be able to identify a persons face immediately. IP cameras, web-cams and other flagship hardware will be used in order to cope with different posture, style and orientation of the face. Even though the person has beard or eyeglasses the algorithm would be able get the identity of the person.

1.3 Outline of Thesis

Chapter 1 gives an introduction to face recognition system. Chapter 2 discusses the current system at J K Lakshmi Cement and the challenges of the current system. Chapter 3 discusses the literature that has been referred. The details of a case study on IFRS, which includes analysis and discussion of main three steps of IFRS.Comparison between face recognition algorithms using parameters which are used during recognition.Chapter 4 represents the proposed system which includes how face will be detected from an image and then from the detected face how eyeglass would be removed is shown and discussion of face recognition techniques,training the database and finally face recognition. Chapter 5 shows an experimental result of proposed system with the help of IP camera as well as Web-cam. The conclusion of this thesis is presented in chapter 6, which also provides the future scope of presenting work. Finally, the last section lists references used in this thesis.

Chapter 2

Project Details

2.1 Overview of Existing System

At present J K Lakshmi Cement owns face recognition system which is shown in fig 2.1 It is purchased from Shree Time in 2015. The system is built using ZK Face 5.0 algorithm on ZEM600 platform. It can store maximum of 700 faces and stores 3 images per person. This system is used for attendance of 50 employees.

Figure 2.1: Current Face Recognition System at J K Lakshmi Cement Pvt Ltd.

2.2 Challenges

- Only frontal face is recognized, if there is any orientation of face then face is not recognized.
- The distance should be maintained in the range of 1.5 meters between system and persons face.
- The face will be recognized if and only if it is in the controlled lightning conditions, the persons face cannot be recognized in dark lights.
- If all the images of the dataset are without spectacles and person comes with wearing spectacles, then persons face is not recognized.
- Accurate result should be obtained within seconds.

Chapter 3

Literature Survey

The survey is based on comparison of different algorithms, along with their pros and cons for different factors. A few important steps researched will be described below. Mainly, the face recognition systems architecture is divided in three major segments:

1. Face Detection
2. Feature Extraction
3. Face Recognition

3.1 Face Detection Approaches

Few of the important face detection methods are discussed below.

1. Knowledge-based: This technique is evolved on the regulations approved from the researchers' understanding of the structure of a human face. The hurdle with this method is the challenge in comprehending the human knowledge with respect to the predefined set of rules.
2. Feature-based: There are standard metrics to define a human face which would eventually help to identify the structure, skin tone, etc. However, these aspects can be easily manipulated with the noise, exposure, and illumination.
3. Template Matching technique: The input image of a human face is compared with a template. However, the major challenge here is that the template might not match with actual face structure, skin tone, scale, posture, etc.

4. Appearance-based technique: This interesting method includes the involvement of machine learning and artificial intelligence. The idea is simple. In the above method, there are predefined templates which are used as the benchmark to detect a human face. Here, the machine learns itself after scanning several images, applying statistical analysis and its intelligence in the method to enhance the overall process of face detection.

Important aspects of the primary face discovery techniques are discussed here.

1. Knowledge-based techniques: It construct strategies that are created with respect to the guidelines given by the specialists information of human countenances. Issue in this approach is difficulty in making an interpretation of human learning into very much defined rules.

2. Featured-based techniques: Invariant elements of appearances are utilized for identifying surface, skin shading. Be that as it may, highlights from such calculation can be extremely undermined because of enlightenment, commotion and impediment.

3. Template coordinating: Input image is contrasted and predefined confront layout. Be that as it may, the execution here suffers because of varieties in scale, posture and shape.

4. Appearance-based technique: In layout coordinating strategies, the formats are predefined by researchers. While, the formats in appearance based strategies are found out from cases in images. Statistical investigation and machine learning procedures can be utilized to find the significant qualities of face and non-confront images.

The main purpose of face discovery is to reveal that if there are any faces in the image or not and if face is present then return an extent of each face. For face detection step of our IFRS, we need to detect face using some facial components like eyes, nose, mouth, lips, eyebrows, face contour with some facial features like with beard or without beard, with eyeglasses or without eyeglasses and with an orientation of face. There are several techniques which includes these cases which are shown as below :

Table 3.1: Comparative analysis of face detection techniques

Papers	Technique	Components	Working Cases	Issues
"Human face detection in a complex background",G. Yang and T. S. Huang	Hierarchical Knowledge Based[2]	Eyes, Nose, Mouth	Frontal faces in uncluttered scenes.	Translating knowledge into rules, detect faces in different poses.
"Finding faces in cluttered scenes using random graph matching" T. Leung, M. Burl, and P. Perona	A computational approach to edge detection [5]	Edge of face,Nose, Eyes	Edges detected closed to true edges.	Real world conditions and scaling, thresholding of image
"An introduction to face recognition technology" S.-H. Lin	Integration of skincolor, size and shape[1]	Skin color, Edge of face,Size	Detect face at different orientations up to 45 degree andfor facial features (glasses and beards also).	Lightning conditions, Only single face.
"Automatic face identification system using flexible appearance models" A. Lanitis, C. J. Taylor, and T. F. Cootes	Deformable Templates: ASM [6]	Edge of face,Eyes, Eyebrows, Nose,Lips, Mouth	Single frontal face.	Multi-faces, Variation in pose, scale and shape.
"Eigenface based facial recognition" D. Pissarenko	Eigenfaces for Recognition [7]	Eyes, Nose, Mouth, Face Contour	Only frontal faces.	Lighting conditions, Facial expression, Orientation
"Human face detection in a complex background" G. Yang and T. S. Huang	Example based Learning for View-based Human face Detection[2]	Eyes, Nose, Mouth, Lips Eyebrows, Face contour	Faces are detected in frontal lightning only and less computational time than eigen face.	Lightning conditions
"Detecting faces in images: A survey" M.H. Yang, D. J. Kriegman, and N. Ahuja	Neural network based face detection[8]	Eyes, Nose, Mouth, Lips,Face contour	Complex class conditional density can be captured.	Requires external tuning
"Rapid object detection using a boosted cascade of simple features" P. Viola and M. Jones	Haar Classifier: Robust real-time face detection by Viola,Paul, and Michael J.Jones [9]	Eyes, Nose, Mouth, Lips, Face contour	Detect face faster than 15 times of other approaches with different facial conditions.	Requires more memory

As shown in the table 3.1, Haar classier can exist for multiple face detection with facial features like with eyeglasses and without eyeglasses, with beards and without beards can be detected as well as it is faster among previous approaches and we can detect faces with more than 45 orientation of faces and higher detection rate among all approaches discussed earlier.

3.2 Feature Extraction

The previous unit gives discussion about face detection, now, in this area, well discuss how to remove valuable and minimal highlights for further recognition procedure. Fig 3.1 shows basic flow of feature extraction after face detection.

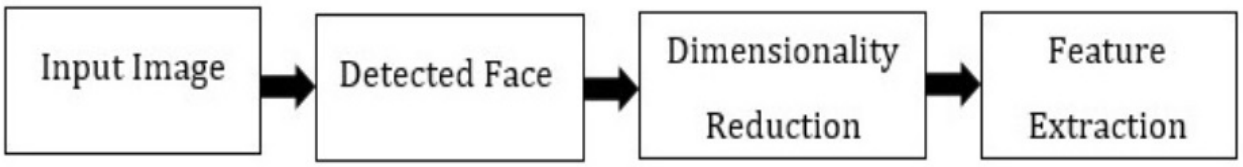

Figure 3.1: Flow of Feature Extraction

When input image to an algorithm is too large for the further processing then this step is necessary. Using feature extraction, we can get reduced size of an image for further procedure by which we can get lesser time for whole procedure as compared to without reducing an image. Most of the methods for feature selection and extraction are also used for later steps. Some of them are described in table 3.2.

Table 3.2: Analysis of techniques of feature extraction

Technique	Working Cases	Issues
PCA,Kernel PCA, weighted PCA [7]	Eigenfaces and dimensionality reduction	The transformed dimensions are uncorrelated with each other, Second order statistics are only used.
LDA,Kernel LDA [10]	Eigenvector based with supervised map,LDA with kernel method	Small sample size problem are solved.
ICA [11]	Transformed dimensions are independent and higher order statistics are used for dimensionality reduction.	Cannot linearly separate independent sources in smaller subspace
LBP [12]	Uniform Binary patterns will contain at most two-bitwise transactions from 0 to1 or vice versa	Blurred images

3.3 Face Recognition Approaches

As discussed in the previous section, the images which are reduced using feature extraction technique would be used for recognition process to decrease the comparison time of an image with the database. There are various techniques for FRS which can recognize human face within lesser time, using facial features like with eyeglasses or without eyeglasses, with bread or without bread, multiple face recognition at a time, orientation of face etc [13]. IFRS techniques which incorporates these characteristics are discussed in table 3.3:

Table 3.3: Analysis of Face Recognition Approaches

Approach	Technique	Working Cases	Issues
Holistic	PCA[14]	Requires full facial information	Facial features like with eyeglasses, with beard etc; and multiple faces at a time
Holistic	2-D PCA: a new approach to appearance based face representation and recognition[15]	Recognize single face with up to 45 degree orientation	Multiple faces
Holistic	Face recognition using LDA- based algorithms [10]	Recognize more than 45 degree oriented face	Recognize maximum 5 faces at a time
CNN	CNN: Training Convolution neural network: an application to face recognition[16]	CNN classifies facial features and train database	Lightning conditions and only frontal face recognition
Feature Based	Face Recognition by EBGM [17]	Faces as a graph	Recognize up to 22 degree oriented faces
Feature	Face Recognition by Active Appearance Models [18]	Facial shapes, facial expressions etc.	Lightning conditions

Comparison between face recognition techniques

There are various approaches for face recognition which are discussed in the previous section. Here, I have shown comparative analysis of face recognition techniques with some facial components like facial features, props, Orientation, lightning conditions etc. Here, I have found the presence of the facial components and props for the success of face recognition.

Table 3.4: Comparative Analysis of face Recognition Techniques

—Components						**Hybrid Techniques**	
	PCA	**ICA**	**LDA**	**LBP**	**EBGM**	**Fisherfaces**	**Proposed Sys**
Features							
Eyes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nose	-	-	-	-	Yes	-	-
Ears	-	-	-	-	Yes	-	-
Face Contour	-	-	-	-	-	-	-
Mouth	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Props							
Eyeglasses							
With Existing Dataset	Yes	Yes	Yes	-	-	Yes	Yes
Without Existing Dataset	-	-	-	-	-	-	Yes
Beards	-	-	-	-	-	-	-
Orientation							
Up to 45 degree	-	-		-	-	-	Up to 45 degree

Chapter 4

Proposed Methodology

The proposed system is as shown below:

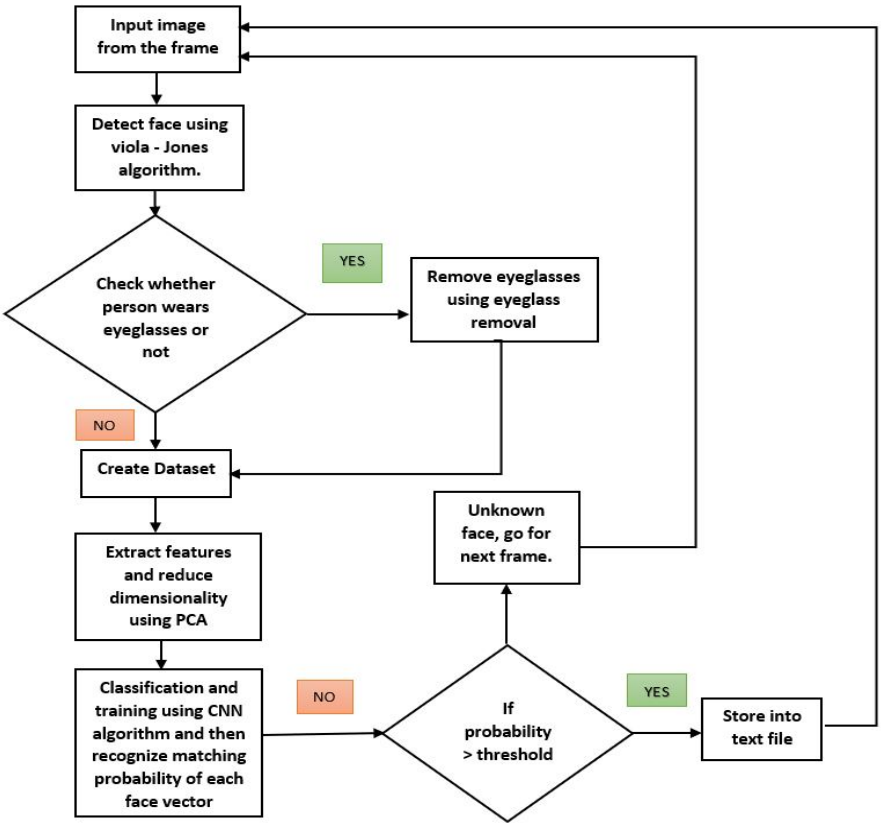

Figure 4.1: Proposed System

As shown in the figure 4.1, we will detect face from an image using Viola-Jones algorithm. Before we create a dataset, we will check whether a person wears eyeglasses or not if person wears eyeglasses then rest of all we will remove eyeglasses and then create dataset. At the time of dataset creation, if the person wears eyeglasses, then we will store 50 percent images with eyeglasses and another 50 percent images with eyeglass removal so that if that person will come with eyeglass or without eyeglass can be recognized. Then we will extract features using PCA algorithm and then using CNN we will train face vectors and recognize face which can recognize multiple face at a time.

4.1 Face Detection using Viola Jones Algorithm

As discussed earlier, Viola Jones [9] have proposed face detection method which is faster than other approaches and has 94% detection rate. Viola and Jones have 4 essential concepts in the design of their face detection system.[9]

- Haar Features
- Integral Image
- AdaBoost Algorithm
- Cascaded Classifier

4.1.1 Haar Features

Haar features are a basic process to distinguish a specific object within an image. Each Haar component interprets an object by depicting the difference in its intensities at a certain region. Haar features are combinations of binary rectangles that can coincide a particular feature in an object and then that features will be used to detect the existence of the feature in an image. We will obtain each feature as a single value by subtracting the sum of an intensity value of white region pixels and black region.

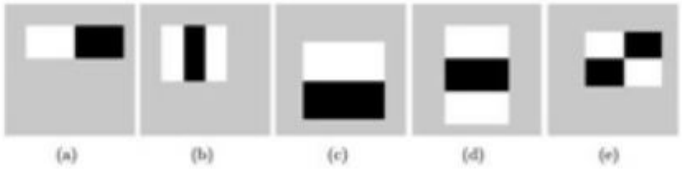

Figure 4.2: Haar Features
Source: https://www.youtube.com/watch?v=$_{Q}ZLbR67fUU/Viola - Jones_{e}xplained$

For an example, eye region contains darker intensity values than the nose bridge, so we will give black-white-black (g-d) pattern for this region. In Viola-Jones algorithm,[9] uses a 24x24 window to classify these features in an image. If we consider all possible parameters like size, position and type we end up with more than 160,000 features in a window. Because of this, we make the use of integral images for low computation cost.

4.1.2 Integral Image

The integral image is a matrix which consists of the sum of all pixels in the upper left region for several pixel[?]. We can find the mean value of pixels at particular region by following formula:

Where i(x,y) is the intensity of the pixels x, y in the input image. This is illustrated below in figure:

$$IImage(x, y) = \sum_{i=1}^{x} \sum_{j=1}^{y} i(x, y) \tag{4.1}$$

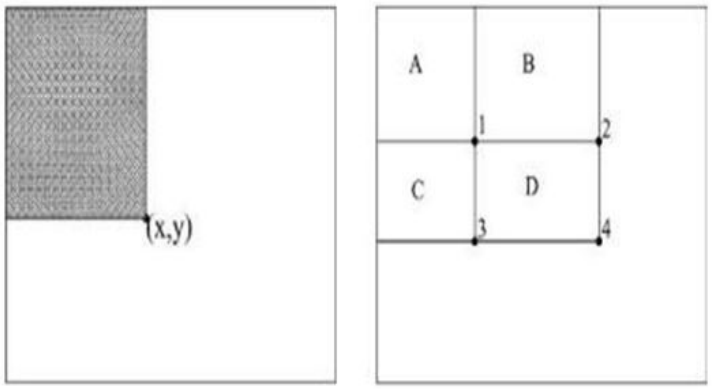

Figure 4.3: Integral Image Calculation
[Source: https://www.youtube.com/watch?v=$_{Q}ZLbR67fUU/Viola - Jones_{e}xplained$]

The entirety of pixels in the rectangle A is the esteem at area 1. So also aggregate of pixels A+B is the esteem at area 2, for area 3 is the total of pixels A+C, and for area 4 is A+B+C+D. So aggregate at area D can be ascertained as 4+1 (2+3). As illustration is demonstrated to higher industry standards figuring the summed territory table. Pixel esteem at one specific spot is the aggregate of pixel estimations of all the pixels in the section earlier or it yet not underneath it and every one of the rows above it on the left-hand side and over it's not the right side.

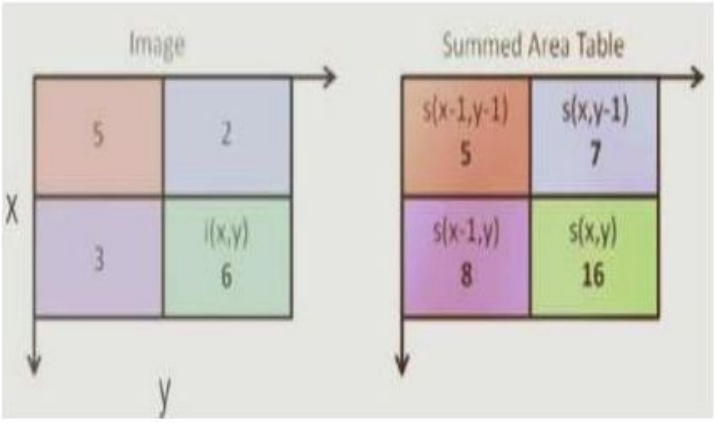

Figure 4.4: Formula for obtaining Summed Area Image
[Source: https://www.youtube.com/watch?v=$_{Q}ZLbR67fUU/Viola - Jones_{e}xplained$]

This summed picture helps in figuring the different Haar feature based control quicker by an extensive sum. Presently we simply need to consider four-pixel values. These four qualities incorporate A, B, C and D as represented in the figure 4.3. These estimations are done in gray scale pictures so at max the estimation of the pixel is 255 so the calculations turn out to be simple. To shape a compelling classifier not very many components is consolidated.The components should be chosen for an effective arrangement so a variant of AdaBoost calculation is utilized.

4.1.3 AdaBoost Algorithm

AdaBoost insinuates to Adaptive Boosting. It tries out various weak classifier more than a few rounds, selecting the best weak qualifier in each round and brushing the best weak classifier to make strong classifier. AdaBoost algorithm definitely realizes what the an-

swer is so, it is administrator sort of calculation and for a situation, it computes different mixes and chooses which suits best for the given situation [9].AdaBoost can utilize classifiers that are reliably wrong by turning around their choice. The weight overhaul step can be intended to just redesign the weight on misses. It can take weeks of handling to decide the last cascade. Along these lines, it makes solid classifier as a direct mix of weighted specimen weak classifiers.

4.1.4 Cascaded Classifier

The negative sub-windows dismissal and detection of every positive window can make a little, solid and effective classifier. Most importantly negative sub-windows are evacuated by basic classifiers, then to keep low false positive rate complex classifiers are utilized.

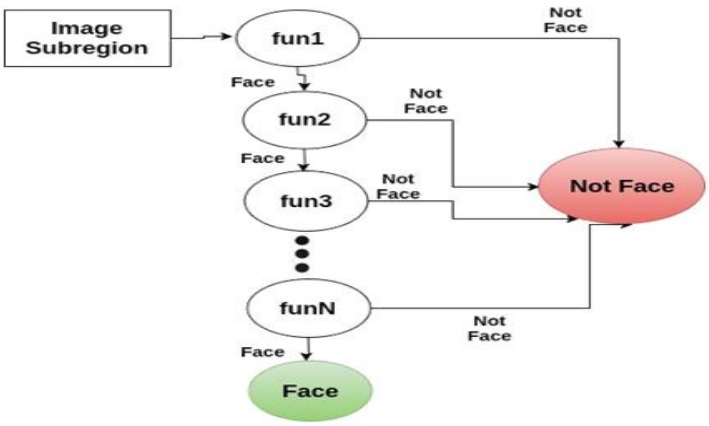

Figure 4.5: Cascade Classifier

Only 1 out of 100 sub-windows contain a face and equal computation time is spent on all windows. For this, we use cascading of classifiers. It has 38 layers cascaded. The rest layer consists of a simple 2 feature classifier, the second layer has 10 features, third has 25 and so on. The window which returns negative for the rest layer is discarded and not tested further. Thus the simple-to-complex cascade makes the computation even more effecient and it reduces the computation speed significantly [19]. As shown in the figure 4.5, we have used N (N = 38) layers cascade and we test our image through these layers and finally after Nth layer we found face detected from an image.

4.2 Eyeglass Removal

4.2.1 Motivation

There would be a crucial task for both face detection and face recognition process while people wears eyeglasses. The appearance of the eyeglass frame is so diverse due to various material properties such as metal or plastic. The respective property of eyeglasses differs significantly from that of human skin. Faces are always well separated with the background whereas the glasses are stuck to the face mixed with eyebrows. For this problem we have two solutions: we can create our dataset as strong as possible i.e. add images of faces which includes all types of eyeglass frames as well as considers the minimum to maximum amount of glass thickness. But for real time FRS, this is not possible to create a dataset because if we try to create a dataset of 50 people the size of dataset would be increased from 13 MB to 85 MB. The second solution is to detect eyeglasses and remove from face images.

4.2.2 Check whether Person wears Eyeglass or not

There are various techniques to check whether the person wears eyeglasses or not from the face image.

- Find out the pixel intensity of detected face and then compare that value with the normal facial image(i.e. without eyeglass face image), if intensity value of detected face is greater than the normal face image then that person from face image wears eyeglasses otherwise not [20].

- Find out pixel intensity of detected face and focus on only that area of the face which covers eyes and nose and find out intensity of the pixels which covers only nose-tips and compare that value with the threshold value(i.e. the intensity of the nose-tip pixels of without eyeglass image)[20]

- At the time of face detection, if we put LED lights nearer to the camera, and us-

ing the reflection of eyeglasses [20], figure out that if a person wears eyeglasses or not.[20] have used 4 LEDs at four corners and in the middle they have put web-cam and rest of all they have turned on all lights, then turn o one or more lights, and then turn o all lights and taking result in all conditions and after that comparing then they concluded that if person wears eyeglasses and LEDs are on then due to reflection of eyeglasses, they can distinguish that person wears eyeglasses or not.

4.2.3 Eyeglass frame detection and removal

There is a difficulty to remove eyeglasses entirely, researchers firstly remove only eyeglass frame by replacing skin color to the detected glass frame. There are various methods for glass frame detection.

Table 4.1: Eyeglass frame detection approaches

Author	Approach	Working Cases	Issue
Saito et al[21]	Active contour model	Saito et al[21] have used snakes for eyeglass frame extraction. They have taken one assumption that the eyeglass frame is symmetric and exists in upper part of the facial image.[21] have implemented Generic Algorithm for searching optimum parameters of the snake	Have to define initially snakes for all type of eyeglass frame
Jiang rt.al[22]	Edge Detection	Jiang rt.al.[22] have proposed a method using six measures of edge information within several regions near eyes by which can decide either input face wears glasses or not.	Can not detect position and shape of the glasses.
Jing and Mariani[23]	Deformable contour	Jing and Mariani[23] has used a common and useful facial feature nosepiece for checking the presence of eyeglasses. In [2], firstly edge map was produced using Canny edge detector and then unnecessary edge points will be removed by the filtering process	If eyeglass frame is large and overlapped with eyebrows then edge detector will not separate eyebrow and eyeglass frame.
Wu et al[24]	3D Hough Transform	Wu et al[24] have presented a method to identify glass frames and separate it from the facial features using stereo facial images. For this purpose [24] has used the 3D shape of various kinds of glasses and identified that the rims lie on the same plane approximately while the other facial features	The rim of defined eyeglasses should be in the same plane and even for an image for which we want to check

4.2.4 Proposed Method for Eyeglass Frame Detection and Removal

Following steps are used for eyeglass frame detection and removal in our system. First of all we will convert detected face into gray scale image and as discussed earlier, using eyeglass detection method we will find out if the person wears eyeglass or not. If the person wears eyeglass, then we will select ROI(Region Of Interest) which covers only eyes and then we will apply thersholding on that region with maximum threshold value=40. Then using canny edge detector, we will find out the region which contains only eyeglass frames and then we dilate an image. Finally we apply skin color inpainting on dilated image.

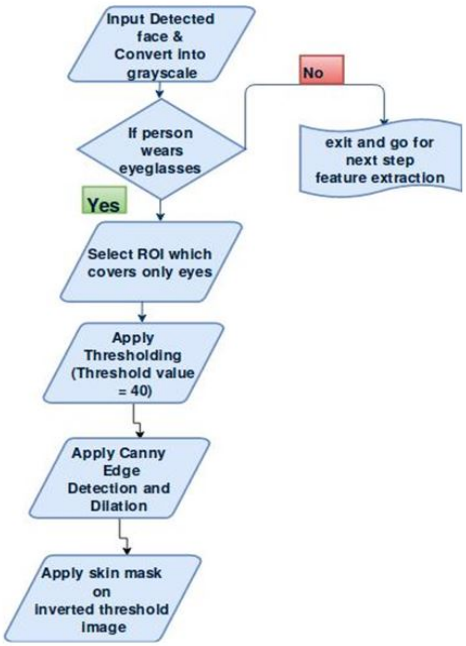

Figure 4.6: Eyeglass Frame Detection and Removal

4.2.5 Remove eyeglasses and reconstruct glass-less image

Eyeglass Removal based on a set of training images:

After extracting eyeglass region from the detected face, we have to remove eyeglasses from extracted region. Saito et al[14] have developed PCA Reconstruction Method. As [13] uses PCA for reconstruction, it depends on the training set and [13] has created a

training set of a person for whom we want to create a glass-less image with glass-less images so that after using PCA reconstruction we can get a glass-less image.

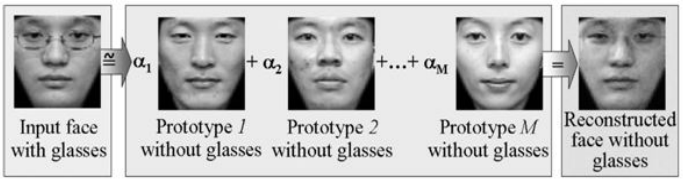

Figure 4.7: Eyeglass PCA Reconstruction

Inpainting:

Inpainting is used to reduce the gap between glasses and non-glasses faces for recognition. After extracting eyeglasses from an image, we apply inpainting approach on the image with the help of mask which will be created as the same size of input image where non-zero pixels corresponds to the area which will be inpainted. In our case, we will use skin color to inpaint eyeglass area and reconstruct the glass-less image.

4.3 Feature Extraction using PCA

We cannot analyze complex data using a single image of face.So, we have to use large amount of images to analyze complex data but for this purpose we need a huge amount of memory and it may occur over-tting problems and generalize poorly to new samples. To overcome these problems, we can use dimensionality reduction techniques.

Principal Component Analysis(PCA) is commonly used technique for dimensionality reduction in the various areas like for feature extraction in pattern recognition, data representation in computer vision etc. Kirby and Sirovich [14] applied PCA for representing faces and Turk and Pentland [14] applied PCA for recognizing faces.PCA is commonly referred to as the use of Eigenfaces [13]. In this technique, each face image would be presented as a weighted sum of the eigenfaces and PCA eigenfaces.

Step 1: Create a Training set and Convert it to Face Vectors

- A training set is a set of face images($I_1, I_2, I_3, ...I_M$) in which size of each image ($N \times N$) must be same. For an example, there are 100 images with same dimension 50×50 for the training dataset. So, it would have 50×50 i.e. 2500 dimensions.

- PCA does not work on images directly but works in vector format.So, we have to convert $N \times N$ sized image into a $N^2 \times 1$ format.

- For an example, consider the following fig: which has an image size of 50×50, after converting this image into vector we can get 2500×1 sized vector of each image.

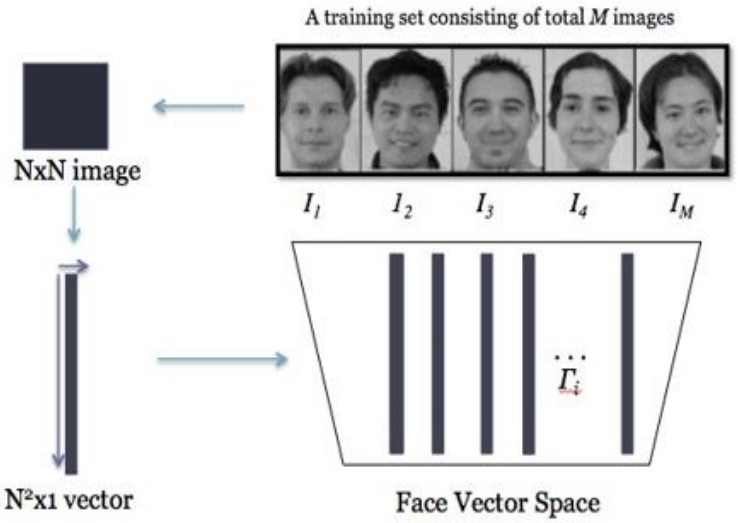

Figure 4.8: Dataset converted to Face Vectors
[Source: https://www.youtube.com/watch?v=SaEmG4wcFfg]

Step 2: Normalize the Face Vectors We are going to remove features which are common to all the images of the training dataset. So that each face in the dataset will remain with its unique features.

- **Calculate average face vector**

  Average face vector contains facial features which are common in all images of the dataset and it will be calculated by the following equation:

$$\theta = \frac{1}{M} \sum_{n=1}^{M} \Gamma_n \tag{4.2}$$

- **Subtract average face vector**

  The average vector θ will be subtracted from the original faces Γ_i and the result will be stored in variable Φ_i.

$$\Phi_i = \Gamma_i - \theta \tag{4.3}$$

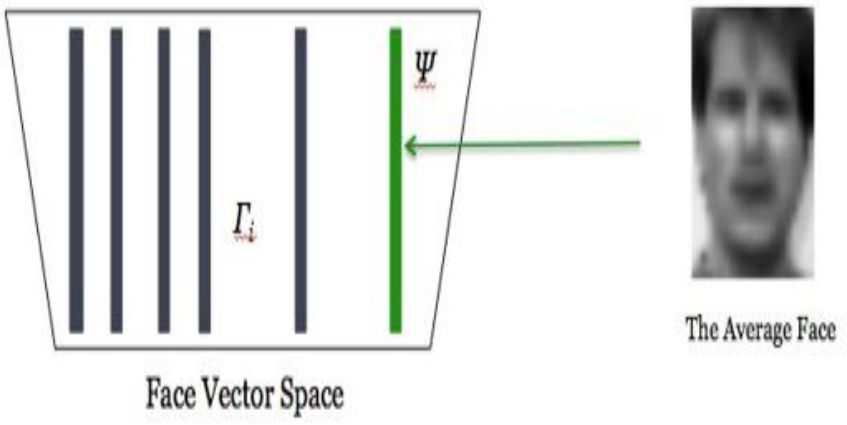

Figure 4.9: Average face
[Source: https://www.youtube.com/watch?v=SaEmG4wcFfg]

Step 3: Calculate co-variance matrix

- To find out the eigenvectors, we will calculate co-variance matrix using the equation $C = AA^T$ where A is the matrix containing all the normalized face vectors i.e. A = $[Q_1, Q_2, ...Q_M]$. So, we will get co-variance matrix C with dimension of $N^2 \times N^2$ and for our case C will be of dimension 2500×2500 and will generate 2500 eigenvectors which will increase the computations required by the system and slow down the system. To overcome this problem, we use dimensionality reduction, in which we

will calculate eigenvectors from a co-variance with reduced dimensionality. We will use C = $A^T A$ to find out co-variance matrix by which we can generate co-variance matrix of the dimension of $M \times M$ which will return M eigenvectors.

- Selected K eigenfaces should be in its original dimensions i.e. $N^2 \times 1$. So, we will multiply each calculated eigenvectors by A since it must be in the original dimensionality of the face vector space: $U_i = AV_i$

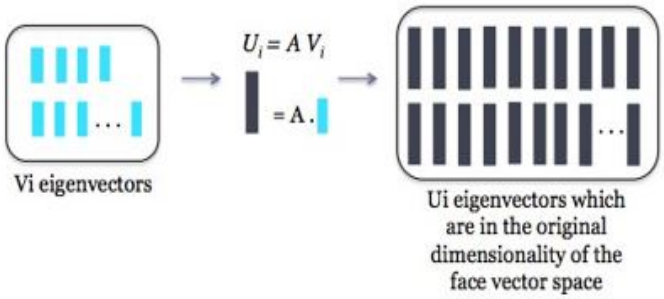

Figure 4.10: Finding co-variance matrix
[Source: https://www.youtube.com/watch?v=SaEmG4wcFfg]

Step 4: Select K best eigenfaces, such that KM and can represent the whole training set

- Using co-variance matrix C, we will obtain eigenvalues and eigenvectors and then from the set of eigenvectors, we will select only K eigenvectors on the basis of K largest eigenvalues.We will represent each face image as a linear combination of all K eigenvectors.

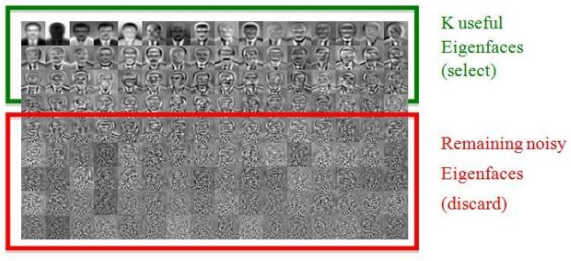

Figure 4.11: Eigenfaces
[Source: https://www.youtube.com/watch?v=SaEmG4wcFfg]

- As shown in the fig, we can see the most major features of a dataset by the first eigenface and the next component shows the next most possible features.

- Here, principle components start from major features of the training set to lesser features, so we will select only a few first principle components which can have major facial features and rest of the last components will be discarded. The selected K principle components represent the whole original dataset as they contain the major facial features that create the dataset

4.4 Training using Convolution Neural Network

Artificial Neural Network with multiple hidden layers has an excellent ability of features learning. The features obtained from learning have an essential description to data, then facilitate visualization or classification.

The difficulties of deep neural network in training can overcome by layer-wise pre-training. In this report, the implementation of layer-wise pre-training is achieved through unsupervised learning. Feed forward neural network or Multilayer Perceptron with multiple hidden layers in artificial neural networks is usually known as Deep Neural Networks (DNNs).[16] Convolutional Neural Networks (CNN) is one kind of feedforward neural network

4.4.1 An Overview of CNN

CNN is an efficient recognition algorithm, widely used in pattern recognition and image processing. It has many features such as simple structure, less training parameters and adaptability. It has become a hot topic in voice analysis and image recognition. Its weight shared network structure make it more similar to biological neural networks. It reduces the complexity of the network model and the number of weights.

Generally, the structure of CNN includes two layers one is feature extraction layer, the input of each neuron is connected to the local receptive fields of the previous layer, and

extracts the local feature. Once the local features is extracted, the positional relationship between it and other features also will be determined. The other is feature map layer, each computing layer of the network is composed of a plurality of feature map. [16]Every feature map is a plane, the weight of the neurons in the plane are equal. The structure of feature map uses the sigmoid function as activation function of the convolution network, which makes the feature map have shift invariance. Besides, since the neurons in the same mapping plane share weight, the number of free parameters of the network is reduced.

Each convolution layer in the convolution neural network is followed by a computing layer which is used to calculate the local average and the second extract, this unique two feature extraction structure reduces the resolution. CNN is mainly used to identify displacement, zoom and other forms of distorting invariance of two-dimensional graphics. Since the feature detection layer of CNN learns by training data, it avoids explicit feature extraction and implicitly learns from the training data when we use CNN.[16] Furthermore, the neurons in the same feature map plane have the identical weight, so the network can study concurrently. This is a major advantage of the convolution network with respect to the neuronal network connected to each other. Because of the special structure of the CNNs local shared weights makes it have a unique advantage in speech recognition and image processing. Its layout is closer to the actual biological neural network. Shared weights reduces the complexity of the network. In particular multi-dimensional input vector image can directly enter the network, which avoids the complexity of data reconstruction in feature extraction and classification process.

4.4.2 Principle of CNN

Convolution neural network algorithm is a multilayer perceptron that identifies two-dimensional image information.It has four major layers: input layer, convolution layer, sample layer and output layer. But,in a deep network architecture the convolution layer and sample layer can have multiple layers. CNN is not as restricted boltzmann machine,in cnn as each neuron just feel the local area of the image rather than global image. In addition, each neuron parameter is set to the same, namely, the sharing of weights , namely each neuron with the same convolution kernels to deconvolution image.[16]

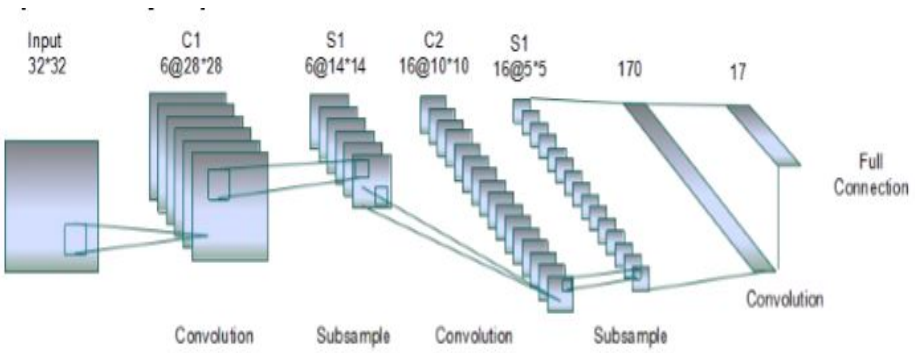

Figure 4.12: Architecture of CNN
[Source: https://www.youtube.com/watch?v=Desfjw5Wvbfg]

4.4.3 Main Process of CNN

CNN algorithm has two main processes:convolution and sampling.

- Convolution process: use a trainable filter Fx, deconvolution of the input image (the first stage is the input image, the input of the after convolution is the feature image of each layer,namely Feature Map), then add a bias bx, we can get convolution layer Cx.

- A sampling process: n pixels of each neighborhood through pooling steps, become a pixel, and then by scalar weighting Wx + 1 weighted, add bias bx + 1, and then by an activation function,produce a narrow n times feature map Sx+1.

To extract feature and reduce the size of the training parameters,the key factors of CNN are the local receptive field, sharing of weights , sub sampling by time or space.The advantage of CNN algorithm is that to avoid the explicit feature extraction, and implicitly learn from the training data[16];The same neuron weights on the surface of the feature mapping, thus reducing the complexity of the network and parallel learning ;Adopting sub sampling structure by time or space, can achieve some degree of robustness, scale and deformation displacement; Input information and network topology can be a very good match, It has unique advantages in speech recognition and image processing.

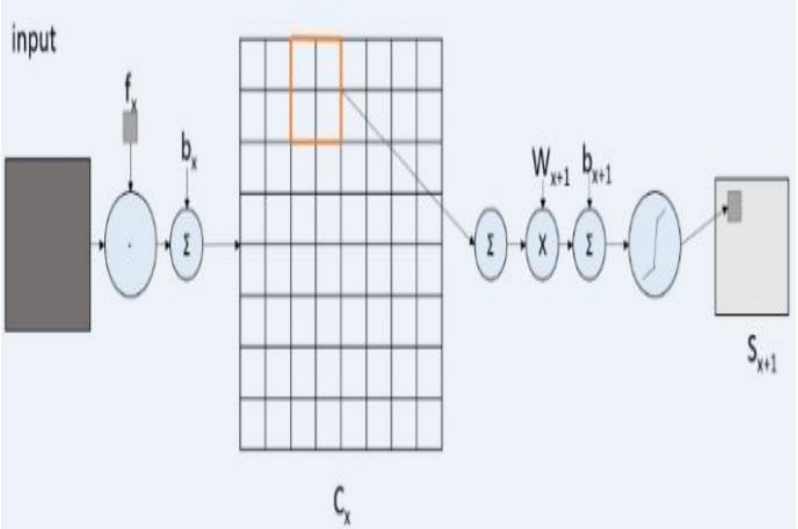

Figure 4.13: Main process of CNN
[Source: https://www.youtube.com/watch?v=FeEmG10wcSvbfg]

4.5 Face Recognition

We have training dataset of faces which is represented as a linear combination of weighted combination of the weighted sum of K eigenfaces added to mean face. Each image will be made of all the eigenfaces in proportions that can be presented as weight vector that contains all the weights and then this face vectors will be trained using CNN with RBF kernel. At the time of training of face vectors, we can find out the probability of each face vectors using RBF. After classification of face vectors, the next step is face recognition. Now, when a new image will be passed for the testing, then first of all that image will be converted to face vector and then using CNN trainer we can find out the margin of the face vector of the test image and then can find out probability of test image using RBF and after comparing that value with the calculated probability of each face vectors probability we can get the recognized face from dataset.

4.6 Recognizing an unauthorized face

We have created a dataset of authorized person for face recognition system. Now, if any person who is unauthorized comes for the face recognition, then that person should not be recognized. Consider the following fig, to distinguish unauthorized people from authorized person at the time of face recognition.

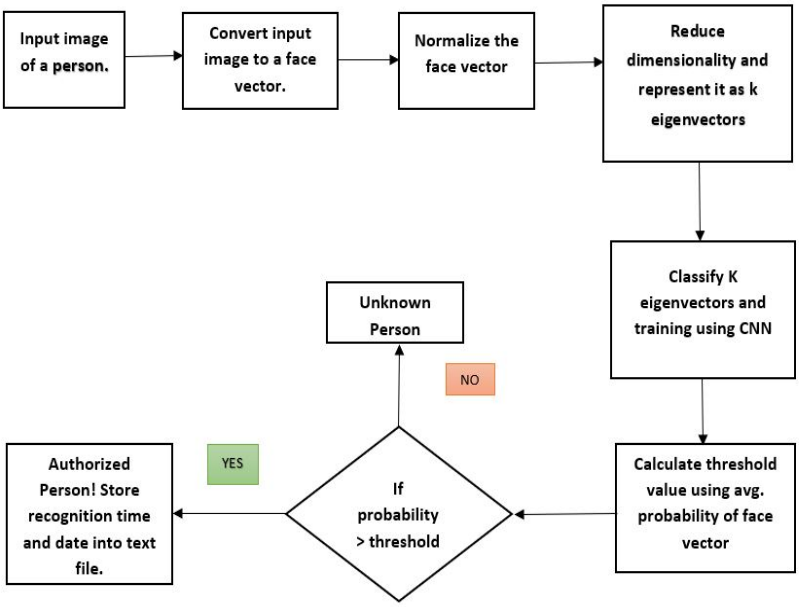

Figure 4.14: Flow Diagram of recognizing unknown person

Chapter 5

Experiments And Results

5.1 Experimental setup

As shown in the fig 5.1, for the testing purpose, we have used one PC with external webcam. The software and hardware specifications for our system is explained below.

Figure 5.1: Experimental Setup

- Software Requirements

  In our system, we have done implementation, training of data and testing of project using OpenCV. We require OpenCV 2.4.9 or upper version with python and Ubuntu as on Operating System.

- Hardware Requirements

  For IFRS, we use one computer system which has RAM 4 GB, internal hard-disk is 500 GB and processor is i3.

- Face Datasets

  We have different face datasets available which can be used for testing purpose. Here, we use Yale [25], Yale B database [25] and AT T [25] for testing of PCA, LBPH, Fisherfaces and proposed face recognition method. The main purpose of selecting these face datasets is to test the robustness of the proposed method for recognizing authorized person. The Yale database consists of 15 persons individual images where each person will have its own 11 images in different conditions like center light, left light, right light, with glasses, without glasses, normal face, happy, sad, surprised etc. The Yale database B consists of 10 persons individual images where each person will have its own 576 images in different conditions like center light, left light, right light, with glasses, without glasses, normal face, happy, sad, surprised etc. For reducing the computational complexity, we reduce the size of image samples of each face datasets.

5.2 Results

5.2.1 Face Detection

An effective face detection in a picture with a frontal perspective of a human face is as shown in the Fig.5.2 Face will be detected from an image using Viola Jones algorithm [9] by running a window over an image. As shown in the fig 5.2, our system takes 226.2 ms to detect real-time face from an image.

Figure 5.2: Face Detection

5.2.2 Dataset Creation

- Check whether person wears eyeglasses or not:

  To check whether a person wears eyeglasses or not, we will firstly find out the region of eyes from the detected face and then find the intensity value of each pixels and calculate the total pixel value. If the intensity of pixels is greater than threshold value(without eyeglass intensity level, here threshold value=160) [23].

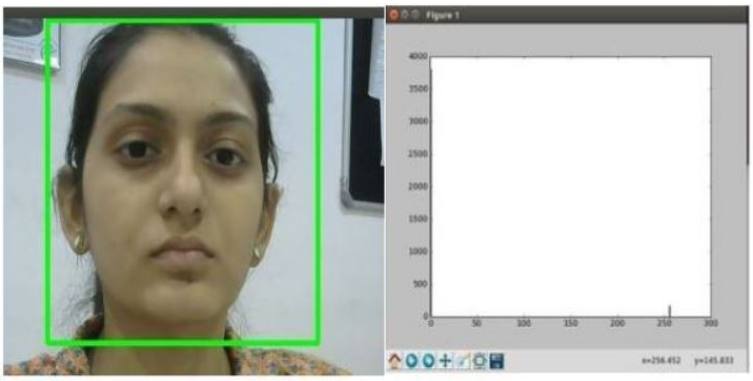

Figure 5.3: Eyeglass Detection

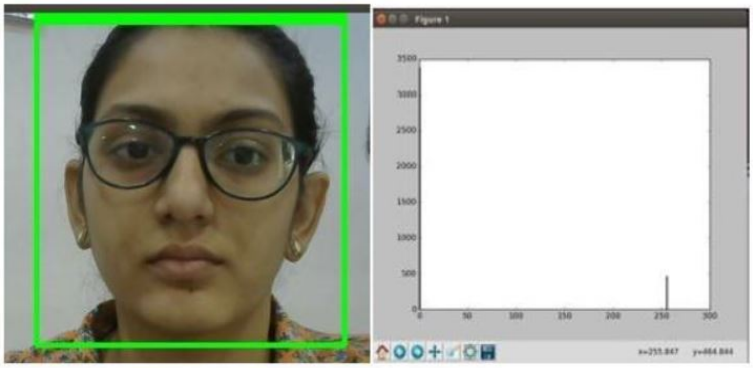

Figure 5.4: Eyeglass Detection

As shown in the fig 5.3, the intensity level of a face without glasses is 145 pixels and with eyeglasses the intensity level is 464 pixels. Here,intensity level¿threshold value. So person has wore eyeglasses.

- Dataset without Eyeglasses:

  As explained previously, if the threshold value of pixels would be less than 160 pixels then the person do not wear eyeglasses. As shown in the fig, the person do not have wear eyeglasses, so we will create dataset of that person by storing images in one specific folder.

Figure 5.5: Dataset Creation

- Dataset with Eyeglasses:

  If the intensity level of detected face is greater than the threshold value(= 160 pixels) then it means that person wears eyeglasses. Now for the time of dataset creation, if the person wears eyeglasses, we will insert 50 % images with eyeglasses and 50 % images without eyeglasses. eyeglass removal will be done using edge detection and skin color inpainting as discussed earlier.

5.3 Face Recognition

In this section, we will discuss about face recognition with IP camera and WebCam.

5.3.1 Face Recognition with IP Cam

We have used HikVision IP camera for FRS, which has frame rate 25 fps. With the help of this IP camera, we can successfully recognize people in a smaller amount of time. If we use IP camera, then our system can recognize person within 2 to 3 seconds.

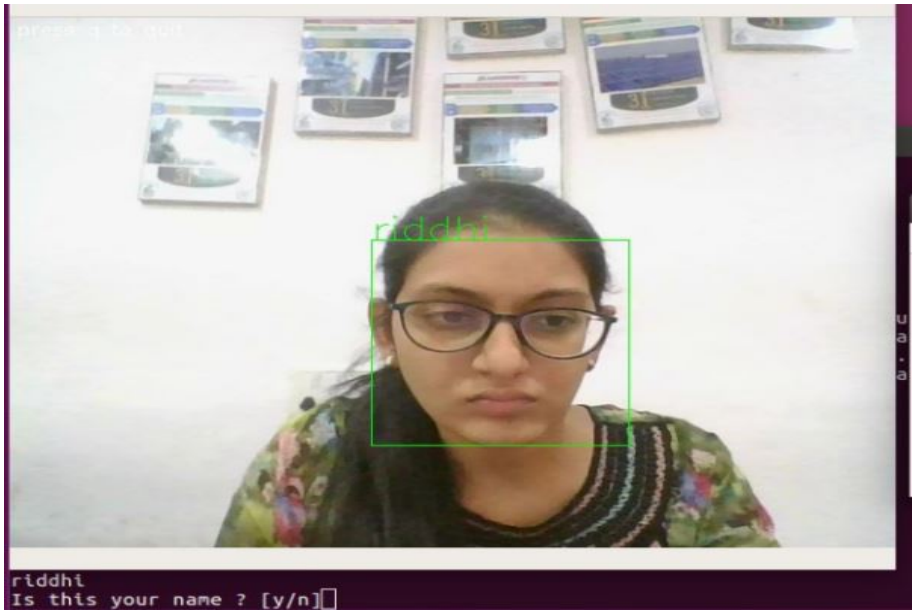

Figure 5.6: Face Recognition using IP-camera

Figure 5.7: Correct Face Recognition within 1 sec

Now, if we do not have good internet connectivity, then we may get delay between two frames. Here, due to less network connectivity, sometimes we get frames will be totally lost or delay between two frames i.e. if one person is recognized and the next person comes we get delay of 10 to 15 frames. So, for realtime, if we have good internet connection, then and then only we can use IP camera for face recognition.

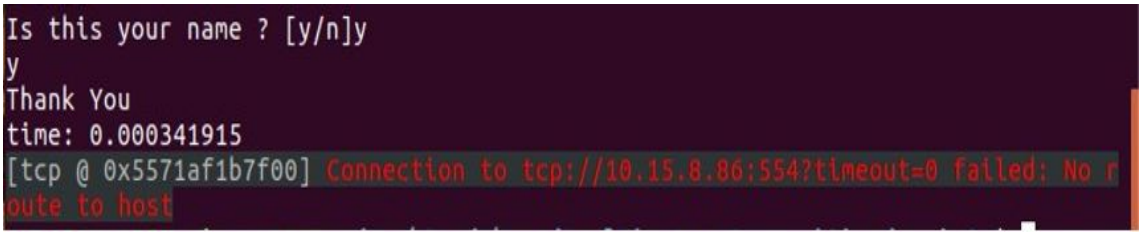

Figure 5.8: Problem at IP cam

5.3.2 Face Recognition with Web-cam

Due to lesser network connectivity, we cannot use IP camera for real-time face recognition.So, we have used web cam for this purpose. In our system, we have used Logitech camera which is USB camera with frame-rate of 25 fps. The frame delaying problem with IP camera will be solved here. We get 2 to 4 frames delay with Web cam in worst case otherwise there will not be any delay between two frames. We can also recognize multiple faces at a time.

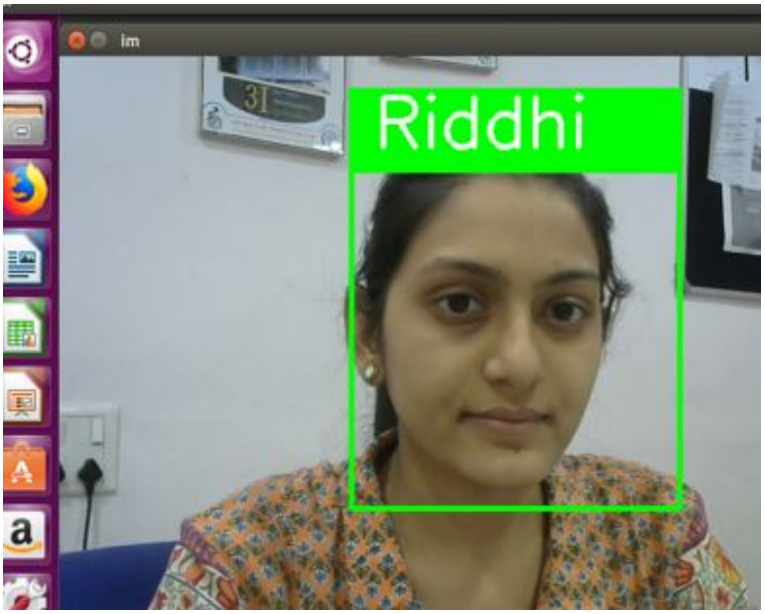

Figure 5.9: Face Recognition using Web-camera

5.3.3 Face Recognition: GUI

In this section, we will show face recognition using Graphical User Interface. As shown in the fig, if we press Run button, the face recognition process will be started in back-end. First of all, dataset will be trained and then the image which is matched with dataset and name of the person will be displayed.

Figure 5.10: Graphical User Interface

As shown in the fig 5.10, after clicking Run button, we get an image which is matched with the dataset and the name of that person as well as we can see one pop up window which shows one instruction that Choose Yes or No which will tell user to choose appropriate option. Now, if a person is unknown i.e. if we do not have dataset of that person then that person will not be recognized and it will show unknown.

5.4 Result set Comparisons

5.4.1 Databases used for Testing purpose

We have used AT T, Yale Face Database and Yale Face Database B for testing purpose. We have shown comparison between these datasets. AT T database contains total 400 images of 40 persons, Yale face database contains 165 images of 15 persons and Yale face database B contains 5760 images of 10 persons.

5.4.2 Face Recognition using various algorithms

We have compared face recognition algorithms with standard database. As shown in fig, we can get higher accuracy with PCA + CNN algorithm using Yale face database B.

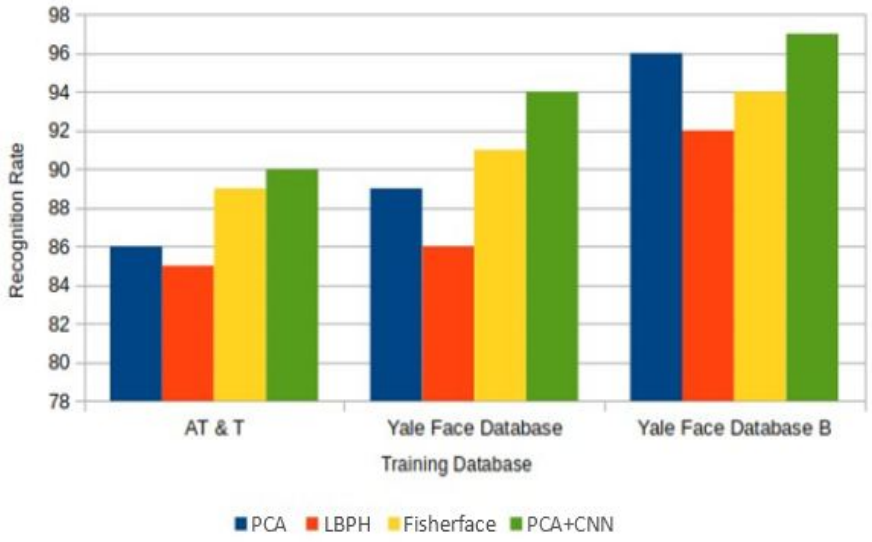

Figure 5.11: Face Recognition Using Standard Dataset

We have compared face recognition algorithms with standard database and our own dataset shown as face profiles. We have added faces within lighning conditions. We have added frontal face, oriented face and face with facial expressions.

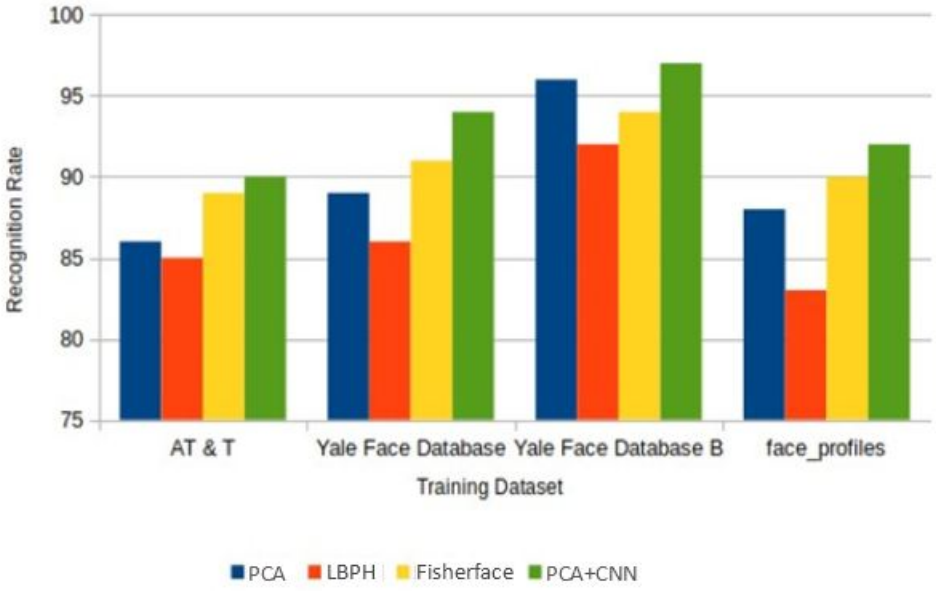

Figure 5.12: Face Recognition Using Standard Dataset and own dataset

We have compared standard dataset and own dataset with face recognition algorithms and proposed method and we can get more accuracy than other algorithms.

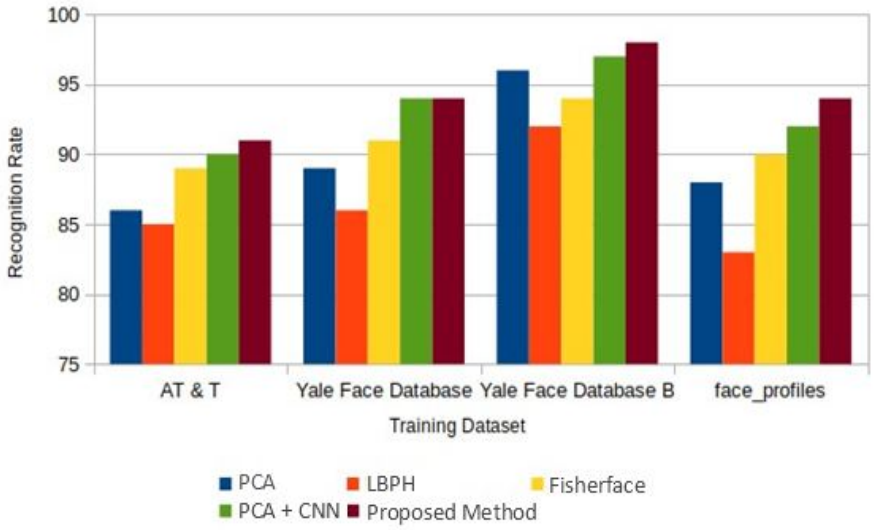

Figure 5.13: Face Recognition with proposed method Using Standard Dataset and own dataset

Chapter 6

Conclusion & Future Scope of Work

6.1 Conclusion

With the advancement of learning algorithms we have to train machines using different algorithm to make intelligent autonomous system to recognize face. The work presented here is exhaustive literature survey of face recognition approaches which includes face detection, feature extraction and face recognition and their implementation. For face detection we have use Viola Jones algorithm, which gives accuracy of 93%.The implemented face detection algorithm also include the variation of face images like occlusion, frontal face, 45°orientation, with or without eyeglasses.For feature extraction and dimensionality reduction PCA algorithm is used. For face recognition CNN is used training of dataset. We keep threshold value with the help of RBF to recognize only known faces. We have tested our proposed system with various datasets and various algorithms and we got 93% accuracy.

6.2 Future Scope of Work

We have observed that due to constraint of space for database, we would like to extend our work by configuring cloud platform with this project.

Bibliography

[1] S.-H. Lin, "An introduction to face recognition technology," *Informing Science*, vol. 3, no. 1, pp. 1–8, 2000.

[2] G. Yang and T. S. Huang, "Human face detection in a complex background," *Pattern recognition*, vol. 27, no. 1, pp. 53–63, 1994.

[3] "Computer vision." https://en.wikipedia.org/wiki/Computer$_v ision$.

[4] "Opencv." https://www.learnopencv.com/.

[5] T. K. Leung, M. C. Burl, and P. Perona, "Finding faces in cluttered scenes using random labeled graph matching," in *Computer Vision, 1995. Proceedings., Fifth International Conference on*, pp. 637–644, IEEE, 1995.

[6] A. Lanitis, C. J. Taylor, and T. F. Cootes, "Automatic face identification system using flexible appearance models," *Image and vision computing*, vol. 13, no. 5, pp. 393–401, 1995.

[7] D. Pissarenko, "Eigenface-based facial recognition," *December 1st*, 2002.

[8] M.-H. Yang, D. J. Kriegman, and N. Ahuja, "Detecting faces in images: A survey," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 1, pp. 34–58, 2002.

[9] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, vol. 1, pp. I–I, IEEE, 2001.

[10] J. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "Face recognition using lda-based algorithms," *IEEE Transactions on Neural networks*, vol. 14, no. 1, pp. 195–200, 2003.

[11] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, "Face recognition by independent component analysis," *IEEE Transactions on neural networks*, vol. 13, no. 6, pp. 1450–1464, 2002.

[12] I. Craw, D. Tock, and A. Bennett, "Finding face features," in *European Conference on Computer Vision*, pp. 92–96, Springer, 1992.

[13] M. Turk and A. P. Pentland, "Face recognition system," Nov. 17 1992. US Patent 5,164,992.

[14] H. Moon and P. J. Phillips, "Computational and performance aspects of pca-based face-recognition algorithms," *Perception*, vol. 30, no. 3, pp. 303–321, 2001.

[15] J. Yang, D. Zhang, A. F. Frangi, and J.-y. Yang, "Two-dimensional pca: a new approach to appearance-based face representation and recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 26, no. 1, pp. 131–137, 2004.

[16] S. H. Kewen Yan1, "Face recognition based on convolution neural network," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 36, no. 4, pp. 4077–4081, 2017.

[17] L. Wiskott, N. Krüger, N. Kuiger, and C. Von Der Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 19, no. 7, pp. 775–779, 1997.

[18] G. J. Edwards, T. F. Cootes, and C. J. Taylor, "Face recognition using active appearance models," in *European conference on computer vision*, pp. 581–595, Springer, 1998.

[19] A. K. Jain and S. Z. Li, *Handbook of face recognition*. Springer, 2011.

[20] J.-S. Park, Y. H. Oh, S. C. Ahn, and S.-W. Lee, "Glasses removal from facial image using recursive error compensation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 27, no. 5, pp. 805–811, 2005.

[21] Y. Saito, Y. Kenmochi, and K. Kotani, "Extraction of a symmetric object for eyeglass face analysis using active contour model," in *Image Processing, 2000. Proceedings. 2000 International Conference on*, vol. 2, pp. 231–234, IEEE, 2000.

[22] X. Jiang, M. Binkert, B. Achermann, and H. Bunke, "Towards detection of glasses in facial images," *Pattern Analysis & Applications*, vol. 3, no. 1, pp. 9–18, 2000.

[23] Z. Jing and R. Mariani, "Glasses detection and extraction by deformable contour," in *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, vol. 2, pp. 933–936, IEEE, 2000.

[24] H. Wu, G. Yoshikawa, T. Shioyama, T. Lao, and T. Kawade, "Glasses frame detection with 3d hough transform," in *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, vol. 2, pp. 346–349, IEEE, 2002.

[25] "The database of faces." http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html.