ABSTRACT

Every living creature in this world has a unique face it's a biological law or one may even call it law of nature. Even if they are identical twins still they have some dissimilarities. The face is the most important representation of a person's identification. In this project we aim to build a model which can recognise the face of a person and validate their identity and attendance. Attendance is an important part of every organisation be it school, college, universities and school. The conventional methods used by pronouncing name or roll number of the person and then the attendance is recorded. There are plenty of drawbacks in this methodology such as time consumption, proxy record and many more. To remove these drawbacks we are going to build a fully automated attendance recorder which focuses primarily on image processing and recording attendance. Face detection is used to identify the location of the face in the image and then Face recognition is used to match the image captured by the webcam and then it is matched with the images stored in the database and when their is a match then the attendance is recorded in with live time.

INTRODUCTION

Attendance is a prime and basic factor for proper functioning of any organisation. The problem arises when we try to take attendance by traditionally calling each name. The technology of automatic attendance will come and act as a time and energy saver. It can also act as an important factor in flight transportation where rather than carrying a boarding pass and waisting a lot of paper irrespective we can install this system and our face will act as a boarding pass.

Variety of other automatic attendance marking systems and technologies do exist which are presently being used at various places. The most popular of these methods is biometric technique and RFID system. Although it is automatic and a tiny bit advanced of traditional system yet it fails to save our resources of time and energy constraint. A person has to wait in queue for giving attendance. Also in this time of COVID the touch factor has also evolved and during this period even this methods implementation was also not possible.

This project works upon the most popular and demanding computer vision technology of facial recognition and its ability to function in various aspects of our life. The system can also be used in various other surroundings such as offices, sports complexes, golf courses, airports and even railway station. The only criteria is that a clear image of the person should be updated in the data base and the source of the image should be authenticated such as Aadhar or passport another way could be when ever a student enrols in a university, his/her image must be uploaded at that same time by officials.

BACKGROUND

Today in this world led by computer science and artificial intelligence continues research and development has been done in the field. Computer vision is a vital part of this research and development. Tremendous work has been done in development of computer vision because it is a very important part of Robotics, Machine Learning and many other sub fields. All this research and development has enabled our today to have a face recognition system in our phones. This describes the essentialness of this topic.

Now this technology has found applications in various other fields for example in Air travel , there are a lot of models being made to use facial recognition technology for the passengers so that they are not required to carry their boarding passes. This tech has also found its application in attendance system which rules out proxy attendance.

The most descriptive and unique feature of a living creature is FACE. This is because of the fact that it is most visible and clear physical characteristic. The working of face recognition a clear image of a person's face is uploaded in the database and then the computer vision software works upon the image and learns about its characteristics by making a histogram. This histogram is saved linked with a unique ID to the name and original photo of that person. Now the trained model is set and made to work it scans a image form the webcam via image capturing or live video analysis now this image is converted into a new histogram and then it is matched with the records in the database and when match is found the recognition task is complete.

Now, this technology has been evolving ever since it had been discovered from mere face recognition we have moved on to autonomous driving cars where they enhanced the vision to an

extent that a car is able to drive on its own. Today these technologies are going to work wonders but in this project we are working on basics of computer vision i.e face recognition.

WORKING MECHANISM CHART

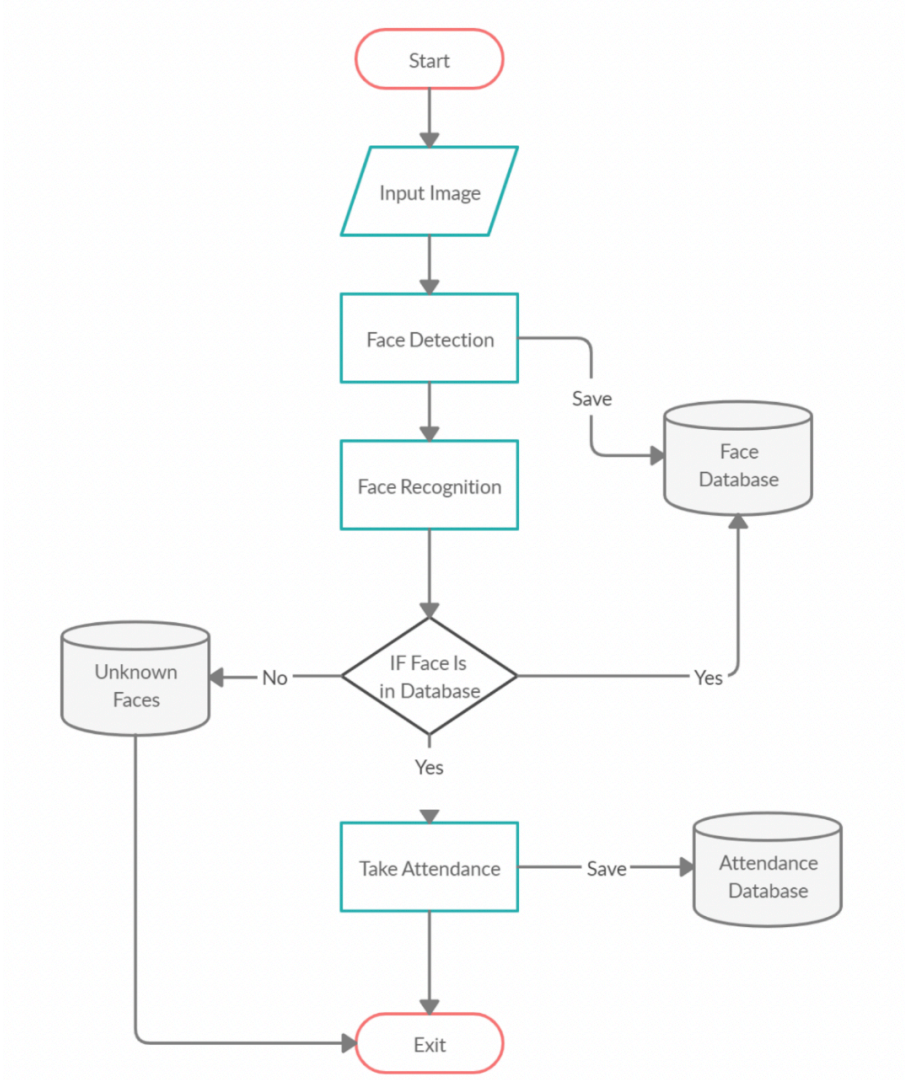


Fig: Flow Chart

Various Attendance System and Their Drawbacks

Radio Frequency Identification card system is one of the most popular and easy to use system. However, there are many drawbacks and one of most common one is proxy cards. To overcome this drawback finger print verification system was introduced but it had another drawback of a lot of time consumption. Instead for face recognition system, the face of a person is always exposed and it contains less personal information when compared to iris. There is no doubt that iris recognition might me more effective but it contains some detail that may hamper the privacy of a user. There is another recognition system also known as voice recognition system but it's very difficult to operate in open surroundings and also its efficiency is poor when compared to the above methods. Therefore the face recognition system is one of the most apt and quick system in terms of biological recognitions system.

SYSTEM TYPE	MERITS	DEMERITS
RFID CARD SYSTEM	Simple	Fraudulent usage
FINGERPRINT SYSTEM	Accurate	Time-consuming
VOICE RECOGNITION SYSTEM		Less accurate compared to others
IRIS RECOGNITION SYSTEM	Accurate	Privacy Invasion

Table 2.1: Merits and Demerits of Different Biometric System

Digital Image Processing

The method of processing and working on images which are digital in nature by a digital computer is called Image Processing. These process techniques are motivated by three major applications :

• Improvement of pictorial information for human understanding
• Image processing for machine learning
• More efficient storage and transmission.

Image Representation in a Digital Computer

An Image is a 2-dimensional light intensity function
$\mathbf{f}(\mathbf{x},\mathbf{y}) = \mathbf{r}(\mathbf{x},\mathbf{y}) \times \mathbf{i}(\mathbf{x},\mathbf{y})$ - (2.0)

Where,
r(x,y) represents the reflectivity of the surface of the corresponding image point.
i(x,y) represents the intensity of incident light

A digital image f(x,y) is discretised both in spatial coordinates by grids and in brightness by quantisation. Hence, the image can be represented effectively as a matrix whose row, column indices specify a point in the image and the element value identifies the grey level value at that point. These elements are referred to as pixels and pels.

The image size used for image conversion to matrix's applications is 256 * 256 elements, 640 * 480 pels for 1024 * 1024 pixels. Quantisation of these matrix pixels is done at 8 bits for black and white image sand 24 bits for coloured images (because of three colour planes Red, Green and Blue each at 8 bits).

Digital Image Processing Technique

The technique in digital image processing :

- Image Acquisition - An image sensitive device and the ability to digitise the signal produced by that device.
- Preprocessing - This enhances the image quality, contrast enhancement, filtering etc.
- Segmentation - Partitions an input image into constituent parts of objects
- Description/ Feature selection - This extracts the image objects description which are further suitable for further computer processing.
- Recognition and Interpretation - The object is given a label based on the information provided by its descriptor. Whereas interpretation assigns a meaning to a set labelled objects.
- Knowledge Base - This is useful for inter module cooperation as well as for efficient processing.

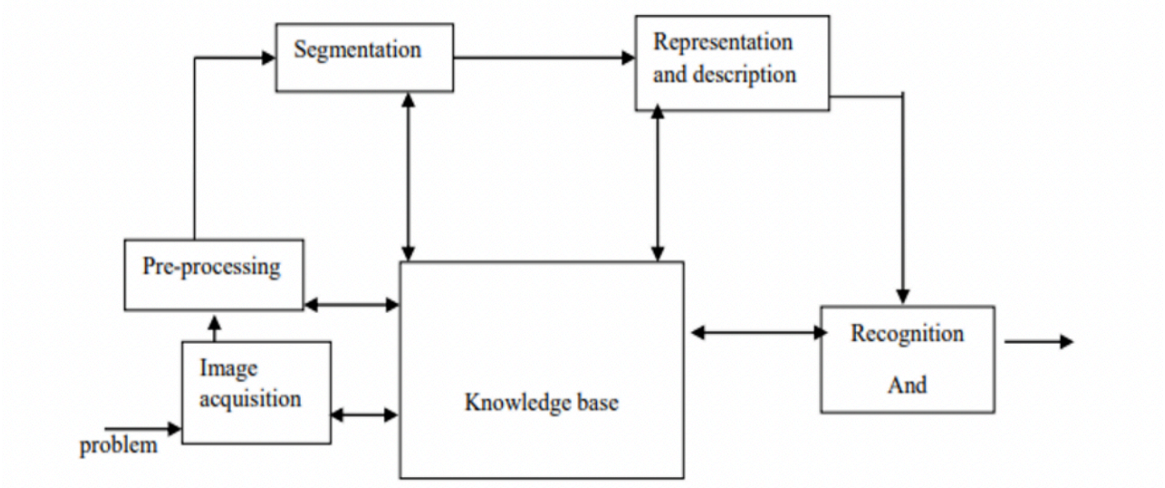


Fig: A diagram showing the digital image processing technique

FACE DETECTION

Face detection is the technique and process of searching, identifying and locating all the faces that are found in single image or video irrespective of their scale, position, orientation, age and expression. Also, this detection is regardless of extraneous illumination conditions and the content of the image or video.

FACE RECOGNITION

Face Recognition is visual pattern identification problem, where the face, is represented as a three dimensional object that is subject to varying illumination, pose and various other factors, needs to be identified based on its learning from other images.

Face Recognition has a simple task of finding an already identified face from database as a known face or unknown face. In some cases it can even tell the name of the person linked with the ID that matches i.e. whose face it is.

Individual Significance of Face Detection and Face Recognition

When ever we see a person's photo or an video we have two primary tasks/questions

A. Where is the person in the photo/video?

B. Who is this?

Therefore to answer these to primary and basic questions we have been given two techniques of Face Detection and Face Recognition.

Face Detection answers the question A. It identifies the person's face in the input photo/video. Whereas on the other hand we have Face Recognition, this technique identifies who is this? It decides that if the input face is a recognised or unrecognised face based on

the model's learning from the database from where it validates this input image.

Hence it is clearly understandable that face detection provides the input for face recogniser and then the recogniser provides the final output.

FACE DETECTION

The task performed by the face detector is to tell that wether the image given be it any arbitrary size contains a human face or not and if it contains then detector provides the recogniser with the exact coordinates of the Face detected. Face detection can be performed based on several cues:

1. Skin colour (for faces in colour images and videos)
2. Motion (for faces in videos)
3. Facial/head shape
4. Facial appearance

Or a combination of these parameters.

Majority of facial recognition algorithms are based on these cues only i.e. they do not use any other cue. An input image is scanned at any of the locations and scales by a given sub window. Face detection works by classifying the pattern in sub-window as a face or non-face. The face/non-face classifier uses statistical learning with the help of training examples. Most of the modern algorithms use Viola Jones object detection framework, which uses Haar Cascades.

FACE DETECTION METHODS	BENEFITS	DRAWBACKS
Viola Jones Algorithm	1. High Speed detection 2. High Accuracy.	1. Slow/Long training Time. 2. Limited Head Pose. 3. Limited Functioning in low brightness
Local Binary Pattern Histogram	1. Simple and Easy Computation 2. High tolerance against the monotonic illumination changes.	1. Only possible for binary and grey images 2. Overall performance is inaccurate compared to Viola-Jones.
Ada Boost Algorithm	It does not need to have any prior knowledge about face structure.	The result is majorly depended on the training data and affected by weak classifiers
SMQT Features and SNOW Classifier Method	1. It has the ability to deal with the lighting problem in object detection. 2. It is extremely efficient in computation.	The only drawback is that the region which are similar to grey value region will be misidentified as a face.

FACE DETECTION METHODS	BENEFITS	DRAWBACKS
Neural Network	It provides high accuracy only if large size of image were trained.	1. Detection process is slow and computation is complex. 2. Overall performance is lower than Viola-Jones algorithm

Table: Benefits & Drawbacks of Face Detection Methods Viola-Jones Algorithm

It was in 2001 when P. Viola and M.J. Jones designed todays most popular algorithm for extracting face parts from a static image or video frame. This algorithm is known as Viola-Jones algorithm. The implementation of this algorithm has been divided in into 4 sub parts:

1.) Haar Feature
2.) Creating Integral Image
3.) Applying Adaboost
4.) Finally Cascading

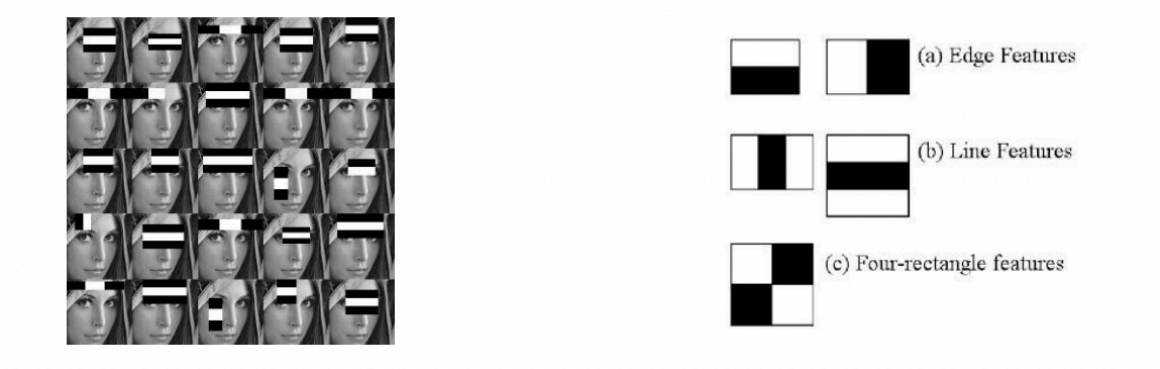


Fig: Haar Feature

This algorithm analyses the input image using Haar features consisting of multiple rectangles.
In the fig shows different types of Haar features.The feature performs the task of window function mapping of the image. Now a single value result is to be computed which is going to represent each feature. This result can be calculated by subtracting sum of white rectangles from sum of black rectangles.

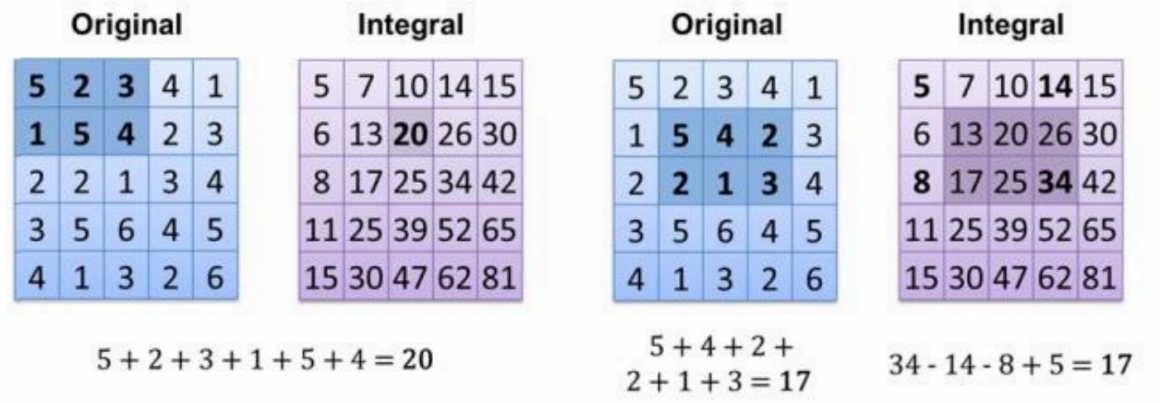


Fig: Integral of Image

The value of integration of image at a particular location is represented by the sum of pixels on the left and top of that particular location. In order to be descriptive and clear, the value of the integral of the image at location 1 is represented as the sum

of pixels in rectangle A. The value of all other integral images at different locations is cumulative.

In order to illustrate clearly, the value of the integral image at location 1 is the sum of the pixels in rectangle A. The values of integral image at the rest of the locations are cumulative. For example, the result at location 2 his calculated by sum of A and B, (A+B). In the the same way at position 3 the result is calculated by sum of A and C, (A+C).

Local Binary Pattern Histogram

Local Binary Pattern (LBP) is a very essential and very efficient texture operator , it marks the pixels of the image by thresholding the neighbourhood of all individual pixel and calculate the result as a binary number.

It has provided a feature for texture classification since 1994. It has been discovered that to improve detection performance considerably on some datasets we use to combine LBP with histograms of oriented gradients (HOG) descriptor. This combination of LBP with histograms enables us to show the face images as a simple data vector.

LBH Algorithm work step by step:

This Algorithm works in 5 steps: -

1. **Parameters:** This algorithm has 4 parameters:

   (.) Radius: This builds a circular LBP and returns the radius surrounding the central pixel and it is often close to 1.

   (.) Neighbours: This is the number of points used to build the circular local binary pattern. It is also to be kept in mind that the more points we include the more is the computational time. It is usually set to 8.

   (.) Grid X: This is the count of cells in the horizontal direction in the grid. It is to be kept in mind that the more the cells more finer is the grid and this results in higher dimensionality of the feature vector. It is also set to 8.

   (.) Grid Y: This is same as GRID X but just the cells are taken in vertical direction.

2. Training the Algorithm: This is one of the earliest step to train the algorithm. This is achieved by creating a database with all the images of people we want to recognise in the future. We would need to assign a unique ID to each image to link the image to the name of the person and then this ID will act as a key to full fill the job.

3. Using the LBP operation: In the beginning of LBPH is to generate a new image which is driven from the original image and has the ability to describe the original image much more descriptively by highlighting the facial characteristics. To achieve this the algorithm applies the idea/concept of sliding window which has the parameters of radius and neighbours.

This image shows this procedure:

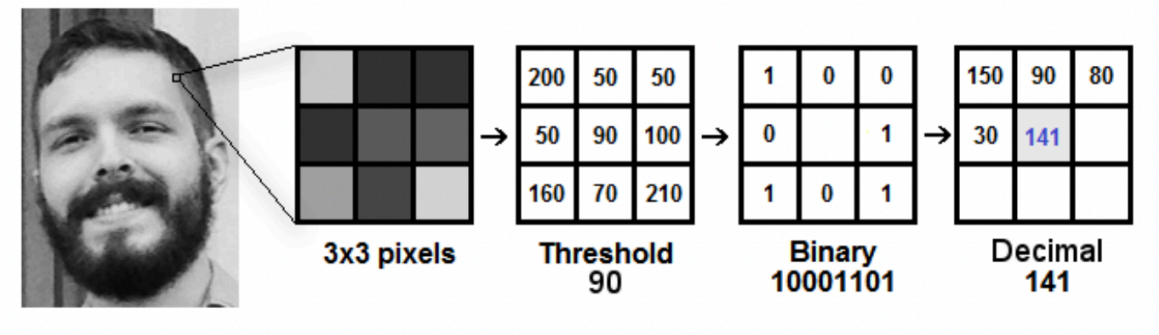
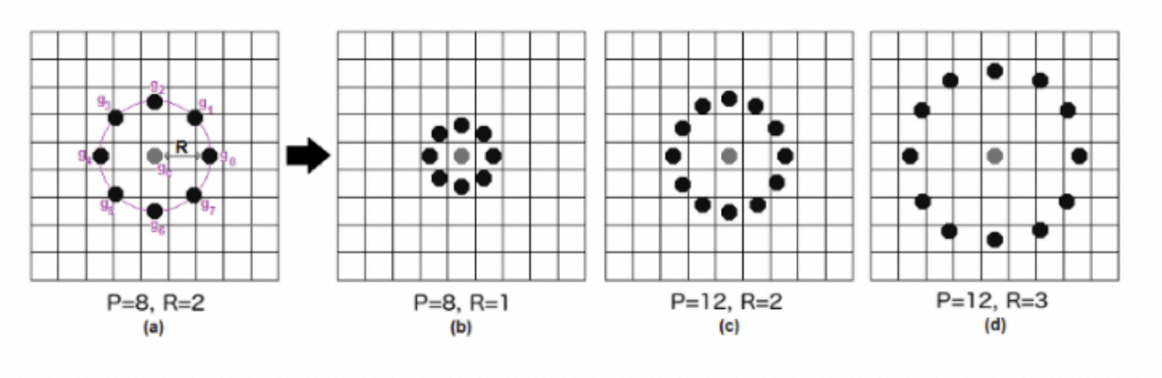


Fig: LBP Operation

Lets break the application of LBP in small steps:

- Input is an image describing a face in greyscale.



- We distribute this image into a window of 3 X 3 pixels.
- Then we represent it in a 3 X 3 matrix representing the impact of each pixel(0~255).
- Then we take the value described by the centre of the matrix as threshold.
- This value is use to define new values from 8 neighbours.
- For each neighbour of the threshold value, we have to set a new binary value. We have a rule for this function.
  - We set 1 for values >= threshold
  - and 0 for values < threshold value.
- The matrix has been redescribed and each position contains only binary number by the previous step just except the central value. Now we concatenate each binary number from all the positions of the matrix line by line into a new binary value. There is also another technique where we concatenate binary values in a clockwise direction.
- Then we convert the binary value into a decimal value and then it is set to the central element of the matrix value, which originally is a pixel of the image.
- Now after all of this process we have generated a new image which describes the image much more descriptively than the original image.

Fig: The LBP operation Radius Change

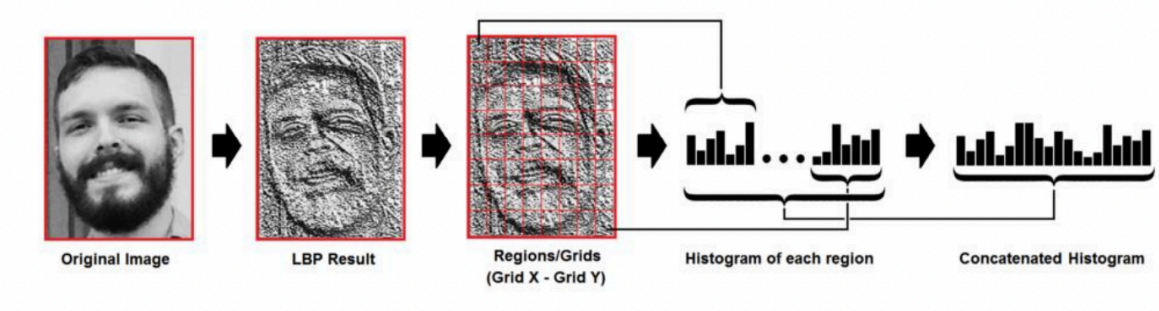


Fig: Extracting The Histogram

4. Extracting the Histograms: From the new image created by the algorithm, we use the two parameters Grid X and Grid Y to divide the image into a new multiple grids.

Now, the following method is used to extract the histogram of each region as follows;

- We consist the image in greyscale, each histogram will contain two hundred fifty-six positions (0~255) depicting the occurrences of each variable intensity.
- Now we have to combine each histogram to form a new and bigger histogram. Suppose we have a 8X8 grid then we will have 8 X 8X 256 = 16384 points in the new generated final histograms. The new histogram shows the features of the original image.

5. Face Recognition: Now we have already trained the algorithm. Now we have a histogram created for each image in the dataset. Now when we read a new image for recognition we again create a histogram for the given input image.

- Now we compare the input histogram with all of database histogram and return the histogram which matches closest to the input histogram and from the ID in database linked to the image we can return the Name of the person.
- We have number of approaches to perform comparison between the two histograms like Euclidean distance, chi-square, absolute value, etc. The most popular one is Euclidean distance and the formula used is,

$$D = \sqrt{\sum_{i=1}^{n}(hist1_i - hist2_i)^2}$$

- The program has to return the calculated distance. This distance is stored as confidence distance. We use the threshold and the confidence distance to automatically estimate the wether the program has correctly functioned and identified the image or not. The criteria for successful recognition is that the confidence is less than the threshold.

Model Implementation

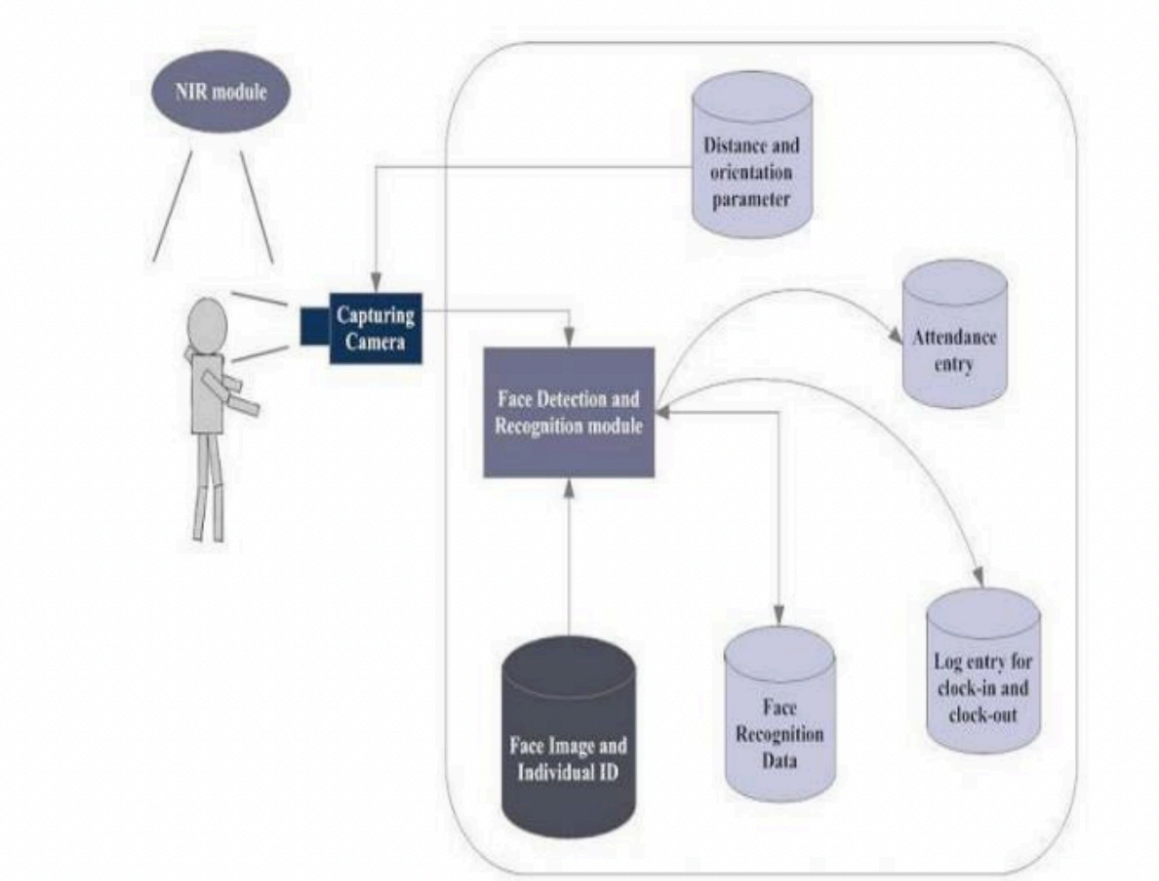


Fig: Model Implement

The backbone of this model of implementation is the open source computer vision library in python (OpenCV). The most important significance of OpenCV is to provide a simple to use computer vision tools that helps people build great computer vision apps.. The working of this model is the administrator uploads the persons image in the database , the image should be taken from at-least 50cm distance. If this is correct then the image is converted into a grayscale and then its Eigen values are stored in a xml file.

When the user request's for recognition the camera scans and generates new Eigen values for each test case and then searches for a match in data base and when a match is found it adds the record having name and time in the excel sheet.

Design Requirements

Software Implementation

1. OpenCV: This is an open source computer vision library in python. This library provides a lot of image processing features and functions. This is a very easy to use library and has brilliant functions which can help create very complex computer vision applications and very quickly. This is open source software under BSD license. Example of some supported functions are:
   - Derivation: Gradient / laplacian computing, contours delimitation
   - Hough transforms: lines, segments, circles, and geometrical shapes detection
   - Histograms: computing, equalisation, and object localisation with back projection algorithm
   - Segmentation: thresholding, distance transform, foreground / background detection, watershed segmentation
   - Filtering: linear and nonlinear filters, morphological operations • Cascade detectors: detection of face, eye, car plates
   - Interest points: detection and matching
   - Video processing: optical flow, background subtraction, camshaft (object tracking)
   - Photography: panoramas realisation, high definition imaging (HDR), image in-painting

2. Python IDE: Today when python is the most popular language for machine learning applications. There are a number of

IDE available for its development such as PyCharm, Thonny, Ninja, Spider etc. I have used PyCharm because it’s free and it’s very easy to use and extremely feature rich.

HARDWARE REQIREMENTS:

Fig: Raspberry Pi 3 Model B+

Fig: Webcam

Fig: Power Source

Fig: Prototype Device