Project Documentation:

Smart Attendance System

Submitted by

Harsh Kamleshbhai Shah
Kavish Shah

TABLE OF CONTENT

ABSTRACT

In the traditional system, it is hard to handle the attendance of huge employees in a company. As it is time-consuming and has a high probability of error during the process of inputting data into the computer. Real-Time Face Recognition is a real-world solution which comes with day to day activities of handling a bulk of student's attendance. Face Recognition is a process of recognizing the employees face for taking attendance by using face biometrics. In this project, a computer system will be able to find and recognize human faces fast that are being captured through a surveillance camera. Numerous algorithms and techniques have been developed for improving the performance of face recognition but our proposed system uses Haar cascade classifier to find the positive and negative of the face and LBPH (Local binary pattern histogram) algorithm for face recognition by using python programming and OpenCV library. Here we use the tkinter GUI for user interface purposes.

List Of Figures

List of Tables

ABBREVIATIONS

Abbreviations used throughout this whole document for Survey Application are:

ABBREVIATION	FULL FORM
LBPH	LOCAL BINARY PATTERN HISTOGRAM
UI	USER INTERFACE
UX	USER EXPERIENCE
OpenCV	OPEN SOURCE COMPUTER VISION LIBRARY
SEO	SEARCH ENGINE OPTIMISATION
AI	ARTIFICIAL INTELLIGENCE
GUI	GRAPHICAL USER INTERFACE
IOT	INTERNET OF THINGS
ML	MACHINE LEARNING
DL	DEEP LEARNING
I/O	INPUT AND OUTPUT
RFID	RADIO-FREQUENCY IDENTIFICATION
LBP	LINEAR BINARY PATTERN
LBPH	LINEAR BINARY PATTERN HISTOGRAM
CSV	COMMA SEPARATED VALUES
RAM	RANDOM ACCESS MEMORY
CPU	CENTRAL PROCESSING UNIT
IDLE	Integrated Development and Learning Environment
OS	OPERATING SYSTEM
RGB	RED GREEN BLUE

CHAPTER 1

INTRODUCTION

- **PROJECT SUMMARY**
- **PROJECT PURPOSE**
- **PROJECT SCOPE**
- **OBJECTIVES**
- **TECHNOLOGY AND LITERATURE OVERVIEW**
- **SYNOPSIS**

1.1 PROJECT SUMMARY

Our Smart Attendance System being a biometric technique implies determination if the image of the face of any particular person matches any of the face images that are stored in a database. This difficulty is tough to resolve automatically because of the changes that several factors, like facial expression, aging and even lighting can affect the image. Our Smart Attendance System is natural, feasible and does not require assistance. Our expected system engages the face recognition approach for automating the attendance procedure of employees without their involvement. A web cam is used for capturing the images of employees. The faces in the captured images are detected and compared with the images in the database and the attendance is marked.

1.2 PROJECT PURPOSE

Attendance maintenance is a significant function in all the institutions to monitor the performance of the employees. Face recognition for the intent of marking attendance is a resourceful application of attendance system. It is widely used in security systems and it can be compared with other biometrics such as fingerprint or eye iris recognition systems. As the number of employees at an organization increases, the needs for lecturers or to the organization also increase the complication of attendance control. This project may be helpful for the explanation of these types of problems.

1.3 PROJECT SCOPE

Our project simplifies time tracking, and there is no need to have personnel to monitor the system 24 hours a day. A time and attendance system using facial recognition technology will accurately report attendance, absence, and overtime with an identification process that is fast as well as accurate. It will accurately track time and attendance without any human error. Facial biometric time tracking allows you to not only track employees but also add visitors to the system so they can be tracked throughout the worksite. Our system has high security that won't allow anyone to invade the privacy or fraudulent usage as well as it is not time consuming.

1.4 OBJECTIVES

1.4.1 Main Objectives

- Our main objective is to provide an attendance system for the employees to ease the burden of the institute

1.4.2 Secondary Objectives

- To accurately track time and attendance without any human error.
- We want to provide a fast and efficient attendance system.

1.5 TECHNOLOGY OVERVIEW:

❖ **Python**

Python is an interpreted, object-oriented, high-level, general-purpose programming language. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse. The Python interpreter and the extensive standard libraries are available in source or binary form without charge for all major platforms and can be freely distributed. Some of the features of Python are Interactive, Interpreted, Modular, Dynamic, Object-oriented, Portable, High level, Extensible in C++ & C, Extensive Libraries, Simple & Easy, Readable, Free and Open source, Less Coding, etc. Table 1.1 shows some of the advantages and disadvantages of using Python.

Table 1.1 Advantages and Disadvantages of Python

Advantages	**Disadvantages**
Vast Libraries Support	Runtime Error
Improved Productivity	Design Restrictions
IOT opportunities, Embeddable	Database Access
Portable, Free and Open Source	Weak in Mobile Computing

Python is used to almost all disciplines and in wide areas. Below mention are some of the important areas where Python is popularly used.

- ➢ Web and Internet Development
- ➢ Education
- ➢ Scientific and Numeric
- ➢ Software Development
- ➢ Business applications
- ➢ Network programming
- ➢ Games and 3D graphics
- ➢ Desktop GUI applications
- ➢ Database access

❖ Machine learning

Machine Learning is one of the most influential and powerful technologies in today's world where a huge amount of data is available to train the computer to do defined tasks. It will continue to be making headlines for the foreseeable future. Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it to learn for themselves. The primary goal is to allow the computers to learn automatically without human intervention or assistance and adjust actions accordingly. Machine learning algorithms can be categorized as supervised, unsupervised, semi-supervised, and reinforcement. Fig. 1.1 shows the difference between Machine Learning, Artificial Intelligence, and Deep Learning.

Some of the applications of machine learning in important sectors are mentioned below.

- ➢ Virtual Personal Assistants
- ➢ Image recognition
- ➢ Stock marketing
- ➢ Healthcare
- ➢ Online fraud detection
- ➢ Advertise recommendation
- ➢ Traffic prediction
- ➢ Email spam filtering

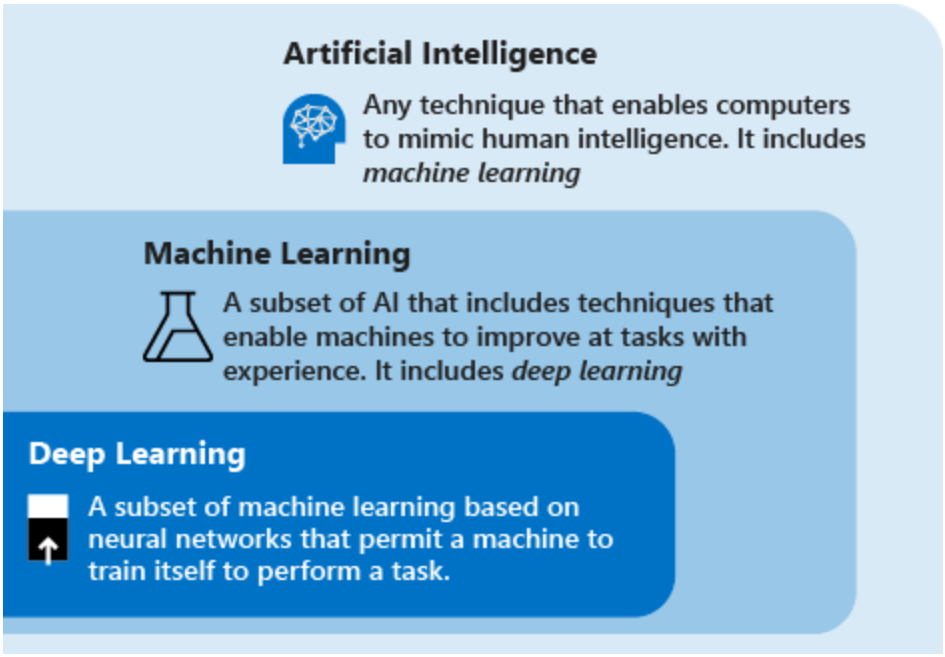


Fig. 1.1 Difference between AI, ML and DL [6]

❖ **Tkinter**

Tkinter is an open source, portable graphical user interface (GUI) library designed for use in Python scripts. Tkinter relies on the Tk library, the GUI library used by Tcl/Tk and Perl, which is in turn implemented in C. Therefore, Tkinter can be said to be implemented using multiple layers. Several competing GUI toolkits are available to use with the Python language.

❖ **Numpy**

NumPy is the fundamental package for scientific computing in Python. It is a Python library that provides a multidimensional array object, various derived objects (such as masked arrays and matrices), and an assortment of routines for fast operations on arrays, including mathematical, logical, shape manipulation, sorting, selecting, I/O, discrete Fourier transforms, basic linear algebra, basic statistical operations, random simulation and much more.

❖ **Pandas**

Pandas is an open source Python package that caters diverse tools for data analysis. The package contains various data structures that can be used for many diverse data manipulation tasks. It also includes a range of methods that can be invoked for data analysis, which becomes feasible when working on data science and machine learning problems in Python.

❖ **Open CV**

Open CV (Open Source Computer Vision Library) is a open source computer vision software library for the purpose of machine learning. Open CV was developed to serve the purpose of computer vision applications and to stimulate the usage of machine perception in the commercially viable products. Open CV is a BSD- licensed product which is easy for the utilization and modification of the code. The library contains more than 2500 advanced algorithms including an extensive set of both typical and state-of-the-art computer vision and machine learning algorithms. These algorithms can be employed for the detection and recognition of faces.

1.5 SYNOPSIS

This report summarizes the work we have done at Maxgen Technologies Pvt. Ltd., as a part of my B.Tech final semester project. Chapter 1 includes the overview, purpose, objectives, and introduction of our project. We also briefed project significance, and tools, technologies opted for this project in Chapter 1. We understood the project, analysed all the previous literature that we have covered in Chapter 2. In Chapter 3, we mentioned the project schedule, work breakdown, timeline chart and risk management of the project. In Chapter 4, we have discussed all the hardware and software requirements of the project. The feasibility of the project and the accuracy is mentioned in Chapter 5. In Chapter 6, we have explained the methodology of the Machine Learning Algorithm that we have used in the project. Testing and validation of applied techniques are available in Chapter 7. Some screenshots of

the implementation of our project are in Chapter 9. We discussed limitations and future advancement that can be made based on this work in Chapter 10. The conclusion part is covered in Chapter 11.

CHAPTER 2

LITERATURE SURVEY

- **LITERATURE SURVEY**
- **DESCRIPTION OF FACE DETECTION AND FACE RECOGNITION**
- **DESCRIPTION ON APPLIED MACHINE LEARNING ALGORITHM**

2.1 LITERATURE SURVEY

We have gone through different published literature to find out the gap between our system and existing systems. There are different types of systems existing in the market till now for attendance of the employees. Some of the approaches are :-

In [1] the authors proposed an attendance System using RFID cards. In this approach, The proposed RFID tag uses energy from the tag reader. The problem with this approach is that an unknown person can make use of a valid ID card and enter the office.

There are also papers involving using biometrics for attendance. In [2] fingerprint is used for marking the attendance of employees. There is a biometric sensor that will take the fingerprint, feature extraction done on that data. If it is for enrollment then that data is stored in the database else if it is for authentication then that data is started matching with the data in the database. The problem with this method is that for attendance students should go to the place where this hardware device is located or pass the hardware device around the students during class which can be a distraction to the students. And in [3] and [4] Iris based attendance system is used. The problem with this approach is that it is sensitive to environmental factors.

In [5] authors proposed a face recognition based attendance system based on Eigenface recognition. Images are converted into eigenfaces, Recognition is performed by comparing eigenface got from the input image and eigenfaces in the database.The problem with this approach is that this method is very sensitive to face background, head orientations and it doesn’t recognize the face of a person if the person is wearing glasses or a grown beard, etc.

But in the approach proposed in our system is not sensitive to face background, head orientations and it recognizes a person's face even if he grows a beard or wears glasses, etc.

A critical survey is very important to find out the gap between the different existing system and our proposed system. Following table shows the types of attendance system and their drawbacks.

Table 2.1: Drawbacks of different attendance system as per the survey

Types of Attendance System	**Drawback**
Manual	Time Consuming Error in entry Low Security Duplicate Data
RFID-based	Fraudulent usage
Fingerprint-based	Time consuming for students to wait and give their attendance
Iris-based	Invades the privacy of the user

Our Smart Attendance System proposed an attendance system that overcomes the problem of the manual method of existing system. It is face recognition method to take the attendance.

2.2 DESCRIPTION OF FACE DETECTION AND FACE RECOGNITION

2.2.1 Face Detection

Face Detection is a method of detecting faces in the images. Its consist of two outcomes i.e. positive image (face) and negative image (no-face). Positive images are the images which contain the object which you want to detect i.e. face in our case. Negative images are the images which doesn't contain object that you want to detect. It is the first and essential step needed for face recognition. It mainly comes under object detection like for example a car in an image or any face in an image and can use in many areas such as security, bio-metrics, law enforcement, entertainment, personal safety, etc.

2.2.2 Face Recognition

Face Recognition is a method of identifying or verifying a person from images and videos that are captured through a camera. Its Key role is to identify people in photos, video, or in real-time.

2.3 DESCRIPTION OF APPLIED MACHINE LEARNING ALGORITHM

2.3.1 Haar Cascade Classifier

It is an Object Detection Algorithm used to identify faces in an image or a real time video. The algorithm uses edge or line detection features proposed by Viola and Jones in their research paper "Rapid Object Detection using a Boosted Cascade of Simple Features" published in 2001. The algorithm is given a lot of positive images consisting of faces, and a lot of negative images not consisting of any face to train on them.

2.3.2 Local Binary Patterns Histograms (LBPH)

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

It was first described in 1994 (LBP) and has since been found to be a powerful feature for texture classification. It has further been determined that when LBP is combined with histograms of oriented gradients (HOG) descriptor, it improves the detection performance considerably on some datasets.
Using the LBP combined with histograms we can represent the face images with a simple data vector.

As LBP is a visual descriptor it can also be used for face recognition tasks

Following are the advantages of LBPH:

- LBPH is one of the easiest face recognition algorithms.
- It can represent local features in the images.
- It is possible to get great results (mainly in a controlled environment).
- It is robust against monotonic gray scale transformations.
- It is provided by the OpenCV library (Open Source Computer Vision Library).

CHAPTER 3

PROJECT MANAGEMENT

- **PROJECT PLANNING OBJECTIVES**
- **PROJECT SCHEDULING**
- **RISK MANAGEMENT**

3.1 PROJECT PLANNING OBJECTIVES

Project planning is an organized and integrated management process, which focuses on activities required for successful completion of the project.

Objectives of Project Planning:

- It defines the roles and responsibilities of the project management team members.
- It ensures that the project management team works according to the business objectives.
- It checks the feasibility of the schedule and user requirements.
- It determines project constraints.

The Project development was planned and divided into various activities like:

- Initially, requirement gathering and research work.

- Based upon the research, learning and understanding of tools and libraries required for development.
- After understanding, start the actual implementation of the project.

3.1.1 Software Scope:

Software scope is the part of project planning that involves determining and documenting a list of specific project goals, tasks, costs and deadlines. Software scope describes the data and control to be processed, function, performance, constraints, interfaces, and reliability. Functions described in the statement of scope are evaluated and in some cases refined to provide more detail prior to the beginning of estimation.

3.1.2 Resource:

A resource is a necessary asset whose main role is to help carry out a certain task or project. A resource can be a person, a team, a tool, finances, and time. Most projects require many different resources to be completed. Resources should be assessed and allocated before a

project begins. Poor resource planning can result in running out of resources midway through a project or delaying deadlines in delivering the final product or service.

3.1.2.1 Human Resource:

The process of identifying and documenting project roles, responsibilities, required skills, reporting relationships, and creating a staffing management plan.

3.1.2.2 Reusable Software Resources:

Software reuse is the process of implementing or updating software systems using existing software assets.

- The systematic development of reusable components.
- The systematic reuse of these components as building blocks to create new system.

3.1.2.3 Environment Resource:

The essential aspects of environmental resource management are ethical, economical, social, and technological. These underlie principles and help make decisions in the project.

3.1.3 Project Development Approach:

The activities followed for this project is listed below:

- Requirement Gathering
- Analysis and Survey

- Organizing and distributing work
- Back end code development
- GUI development
- Integration of GUI and backend
- Testing the system
- Process model used: Incremental Process model

3.2 PROJECT SCHEDULING

3.2.1 Basic Principle

Software project scheduling can be defined as an activity that distributes the estimated effort across the planned project duration by allocating the effort to specific software engineering tasks. Simply one can say that project schedule is a tool which communicates:

1) What works has to be performed.

2) Who will perform the work.

3) Time duration during which the works need to be completed.

3.2.2 Compartmentalization:

The development of the project is divided into the following set of activities/tasks:

1. Implement code for making a Graphical user interface (GUI).
2. Implement code for storing user Id and Name.
3. Implement code to clear the text area of user Id and Name.
4. Implement code for opening the camera in the system.
5. Implement code for detecting the face of the user using Haar Cascade Classifier.
6. Implement code for showing the notification after completion of each task.
7. Implement code for taking pictures of the user and storing it in the system.
8. Implement code for training the pictures of the user.
9. Implement code for detecting the face of the user using LBPH.
10. Implement code for merging the data to a CSV file and marking the attendance and updating.
11. Implement the code to quit the system.

3.2.3 Work Breakdown Structure:

Total work is divided into five parts:

PART I: Gathering information regarding the past and current attendance system and thoroughly reading different Libraries NumPy and Pandas which we are going to use in this project and going through Machine Learning Algorithms.

PART II: Making the Graphical User Interface using Tkinter.

PART III: Implementing Haar Cascade Classifier used for face detection and merging it's functionality to GUI.

PART IV: Implementing LBPH used for face recognition and merging it's functionality to GUI.

PART V: Updating the CSV file by recording the attendance. Training and Testing of the project.

3.2.4 Project Organization:

The project organization is the structure of the project. It's created separately, with specialists and workers from various departments. These personnel work under the project manager.

Project organization is a process. It provides the arrangement for decisions on how to realize a project. It decides the project's process: planning how its costs, deadlines, personnel, and tools will be implemented.

3.2.5 Timeline Chart:

Following figure illustrates the timeline chart of the project. The timeline chart is divided into different tasks undergone in a designated time period.

Fig. 3.1 Timeline Chart

3.2.6 RISK MANAGEMENT

- Facial recognition technology required a highly controlled environment to get high accuracy which can be solved by having proper lightning and enough data set for recognition
- This technology also has a high privacy risk which can only be solved by using trusted facial recognition software
- This technology is new and not so easy so it maybe a little too complicated for people and hard to understand

CHAPTER 4

SYSTEM REQUIREMENTS

- **USER CHARACTERISTICS**
- **FUNCTIONAL REQUIREMENT**
- **NON-FUNCTIONAL REQUIREMENT**

- **HARDWARE REQUIREMENT**
- **SOFTWARE REQUIREMENT**

4.1 USER CHARACTERISTICS

User can use our system for marking their own attendance hasslefree and in an efficient way without wasting any time.

4.2 FUNCTIONAL REQUIREMENTS

Functional requirements define the internal working of the software: that is, the face detection, face recognition, training and processing of the images and other specific functionality that show how the use cases are to be satisfied.

The functional requirements of the system are mentioned as follows:

- ➢ Taking and tracking student attendance by facial recognition in specific time.
- ➢ Generating the Attendance sheet of that particular day.
- ➢ Updating the names of the present employees directly to the attendance sheet
- ➢ Permitting the admin to modify the attendance sheet.

4.3 NON FUNCTIONAL REQUIREMENT

Nonfunctional Requirements are characteristics or attributes of the system that can judge its operation. The following points clarify them:

- **Accuracy and Precision**

The system should perform its process in accuracy and Precision to avoid problems. As we were developing the system, we must make the system that is very accurate in its functions. All the functions should keep working properly, keep getting perfect input, process accurately and produce the perfect output. Accuracy is the most important non-functional characteristic or requirement of the system.

- **Modifiability**

The system should be easy to modify, any wrong should be correct.

- **Security**

The system should be secure and save student's privacy.

- **Usability**

The system should be easy to deal with and simple to understand.

- **Maintainability**

The maintenance group should be able to fix any problem that occurs suddenly.

❖ **Speed and Responsiveness**

Execution of operations should be fast.

❖ **Reliability**

Error handling mechanism must be robust to avoid failure of operation and in case of failure the system reports it to the admin without any due harm.

4.4 HARDWARE REQUIREMENT

The total amount of data that is processed through the company's hardware will be approximately 30 Gigabytes. Table 4.1 denotes the hardware required.

Table 4.1 Hardware Requirements

Requirements	Specification
Processor	i3 or equivalent
RAM	4GB
Memory	> 250GB
OS	Windows 7 or equivalent
CPU	Quad-Core

4.5 SOFTWARE REQUIREMENTS

The project is developed using the python programming language. Table 4.2 lists the software specifications used to process the project.

Table 4.2 Software Requirements

Requirement	Specification
Platform	Python 3.8
Text editor	IDLE Used
Technology	Image processing, Machine learning, Statistics
Libraries	Tkinter ,Numpy ,Pandas ,OpenCv , Shutil, Win-os , Sys , PIL

The main advantage of using Python as a platform is that its vast library support.The table 4.3 below shows the description of the libraries that we used in our project.

Table 4.3 Libraries Used

Library	Description

Tkinter	Tkinter is portable graphical user interface (GUI) library designed for use in Python scripts
Numpy	Numpy is mainly used for mathematical functions as it provides multidimensional array object
Pandas	Pandas is mainly used for data analysis and data manipulation
OpenCv	OpenCv is mainly used for real time Computer Vision
Shutil	Shutil provides functions of high-level operations on files and collections of files that helps in automating the process of copying and removal of files and directories.
PIL	Python Imaging Library adds support for opening, manipulating, and saving many different image file formats.
Sys	The sys module provides functions and variables used to manipulate different parts of the Python runtime environment.
OS	The OS module in python provides functions for interacting with the operating system.

CHAPTER 5

SYSTEM ANALYSIS

- **FEASIBILITY STUDY**
- **ACCURACY**

5.1 FEASIBILITY STUDY

5.1.1 Economical Feasibility

Economic feasibility determines whether the required software is able to reap the financial benefits of an organization. It includes costs incurred by the software development team, estimated hardware and software costs, feasibility studies, and so on. In this regard, it is important to consider the costs incurred in the purchase and the activities required to perform the software development.

5.1.2 Technical Feasibility

Technical feasibility assesses current resources and technologies, which are required to meet the user's needs in software within the allotted time and budget. With this in mind, the software development team ensures that existing resources and technologies can be upgraded or added to the software to meet specific user needs.

5.1.3 Operational Feasibility

Operational feasibility assesses how much does the required software performs in a series of steps to solve business problems and user needs. This may depend on human resources and

includes seeing if the software will work after it has been upgraded and will work once it is installed.

5.1.4 Schedule Feasibility

We planned keeping all the development phase in our mind so our system can be ready in time without any error, after keeping this in mind we planned our deadline one week before actual submission date so we can complete system without error and we found out that we will require minimum 4 months to implementation of the complete project with all the features implemented. This also includes the testing and debugging phase.

5.2 ACCURACY OF PROPOSED SYSTEM:

We considered 3 feet as the distance of an object for recognition. As shown in Figure 5.1, the Face recognition rate of students is 77% . This system is recognizing students even when students are wearing glasses or grown a beard.

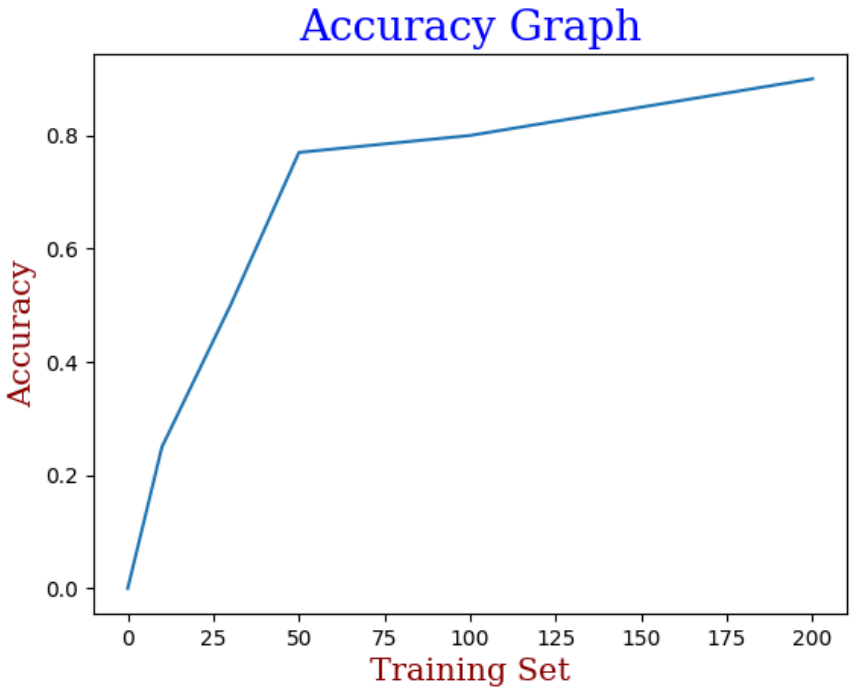


Figure 5.1 Accuracy - Training Set Graph

CHAPTER 6

METHODOLOGY

- **METHODOLOGY FLOW**
- **FACE DETECTION**
- **FACE RECOGNITION**

6.1 METHODOLOGY FLOW

The approach performs face recognition based employee attendance systems. The methodology flow begins with the capture of image by using a camera as the face is detected, followed by pre-processing of the captured facial images, then feature extraction from the facial images, subjective selection and lastly classification of the facial images to be recognized.

6.2 FACE DETECTION

6.2.1 Pre-Processing

Testing set and training set images are captured using a camera. There are unwanted noise and uneven lighting exists in the images. Therefore, several pre-processing steps are necessary before proceeding to feature extraction. Pre-processing steps that would be carried out include scaling of image, median filtering and conversion of colour images to grayscale images. The details of these steps would be discussed in the later sections

6.2.1.1 Scaling of Image

Scaling of images is one of the frequent tasks in image processing. The size of the images has to be carefully manipulated to prevent loss of spatial information. In order to perform face recognition, the size of the image has to be equalized. This has become crucial, especially in the feature extraction process, the test images and training images have to be in the same size and dimension to ensure the precise outcome. Thus, in this proposed approach test images and train images are standardized at size 225 × 225 pixels.

6.2.1.2 Median Filtering

Median filtering is a robust noise reduction method. It is widely used in various applications due to its capability to remove unwanted noise as well as retaining useful detail in images. Since the colour images captured by using a camera are RGB images, median filtering is done on three different channels of the image. Figure 6.1 shows the image before and after noise removal by median filtering in three channels. If the input image is a grayscale image, then the median filtering can be performed directly without separating the channels.

Figure 6.1 Median Filtering

6.2.1.3 Conversion to Grayscale Image

Color images have to be converted to grayscale images before proceeding to the later steps. By converting color images to grayscale images, the complexity of the computation can be

reduced resulting in higher speed of computation. Figure 6.2 and Figure 6.3 shows the conversion of images to grayscale images.

Figure 6.2 Conversion of Image into Grayscale Image

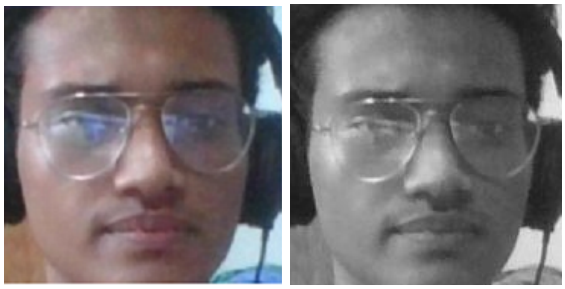


Figure 6.3 Conversion of Image into Grayscale Image

6.2.2 Haar Cascade Classifier Methodology:

Detecting objects with the help of Haar cascade classifiers is an effective method proposed by Paul Viola and Michael Jones in their paper, "Rapid Object Detection using a Boosted Cascade of Simple Features" in 2001. Object Detection comes under a machine learning based approach where a cascade function is trained from lots of positive and negative images.

Now what are these positive and negative images?

A classifier (namely cascade of boosted classifiers working with haar like features) which is trained with many samples of a specific object (i.e., a face or a car), called positive example. So, whatever you want to detect if you train your classifier with those kinds of values. For example, if you want to detect faces then you need to train your classifier with a number of

images which contain faces. So, these are called positive images which contain the object which you want to detect.

Similarly, we want to train the classifier with negative images that means the images which don't contain objects that you want to detect. For example, if we want to detect the face then the image which doesn't contain the face is called a negative image. In the same way if the image contains a face or number of faces then it is called positive images.

After a classifier is trained it can be applied to the region of interest in an input image and classifier outputs 1 if the region is likely to show the object or 0 otherwise.

Here we will work with face detection. Initially, in order to train the classifier, the cascade function needs a lot of positive images (images which contain faces) and negative images (images without faces). Then we need to extract features from it. For this, we use Haar features shown in the below image. They are just like our convolutional kernel. Each feature is claimed to be one value which is obtained by subtracting the sum of pixels under the white rectangle from the sum of pixels under the black rectangle.

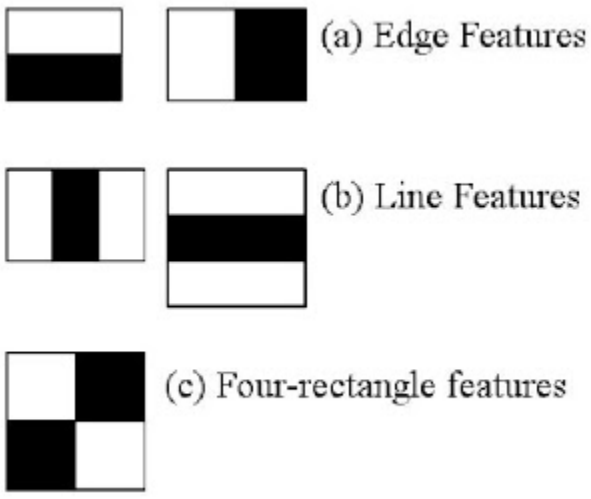


Figure 6.4 Types of Haar-Feature

Now to calculate lots of features, all possible sizes and locations of each kernel are used. (Just imagine how much computation it needs? Even a 24x24 window results over 160000 features). In order to calculate each feature, we need to find the sum of the pixels under white and black rectangles. To get over from it, they introduced the integral image. Calculation depends upon the size of the image if how large your image, it reduces the calculations for a given pixel to an operation involving just four pixels.

But among all these features most of them are irrelevant that we calculated. For example, consider the image below. The top row shows two good features. In the first feature it focuses on the region of the eyes which is commonly darker than the region of the nose and cheeks. When it comes to the second feature it focuses on the property that the eyes are often darker than the bridge of the nose. But if it is applied to cheeks or any other place is irrelevant that you can observe in the image. By using Adaboost we select the best features out of 160000+ features.

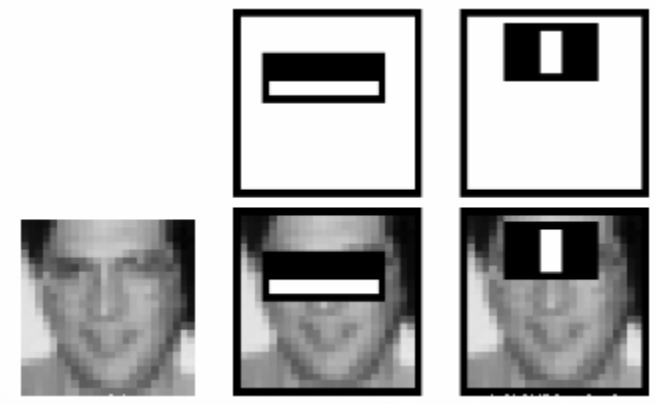


Figure 6.5 Example of Extracting Haar-Features from Facial Image

In the same way, we have to apply each and every feature on all the training images. It finds the best threshold for each and every feature which will classify the faces to positive and negative. Obviously, there will be errors or misclassifications. We only select the features with minimum error rate because they are the features that most accurately classify the face and non-face images. (The process is not as simple as this. Each and every image is given an equal weight in the beginning. After each classification, there will be a change in weights in which weights of misclassified images are increased. Then the same process is done again. New error rates and new weights are calculated. The process will be continued until the required accuracy or error rate is achieved or the required number of features is found).

The final classifier is obtained by a weighted sum of these weak classifiers. It is then called a weak classifier because it alone can't classify the image, but together with others forms a strong classifier.

6.3 FACE RECOGNITION

6.3.1 Local Binary Patterns Histogram Methodology:

Local Binary Patterns Histogram algorithm (LBPH) is for face recognition. It is based on a local binary operator, and it is one of the best performing textures descriptors. The need for facial recognition systems is increasing day by day as per today's busy schedule. They are being used in entrance control, surveillance systems, smartphone unlocking etc. In this article, we will use LBPH to extract features from an input test image and match them with the faces in the system's database.

Local Binary Patterns Histogram algorithm was proposed in 2006. It is based on a local binary operator. It is widely used in facial recognition due to its computational simplicity and discriminating power.

The steps involved to achieve this are:

- ➢ Creating datasets
- ➢ Face acquisition
- ➢ Feature extraction
- ➢ Classification

6.3.1.1 Steps involved in LBPH:

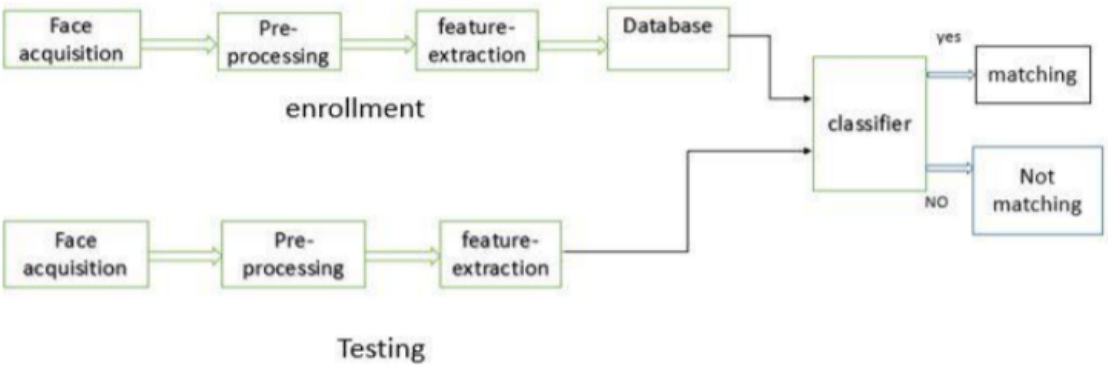


Figure 6.6 Steps Involved in LBPH

- Suppose we have an image having dimensions N x M.
- We divide it into regions of the same height and width resulting in m x m dimension for every region.
- Local binary operators are used for every region. The LBP operator is defined in window size of 3x3

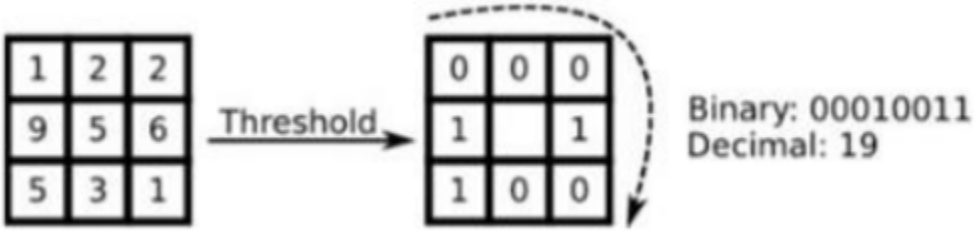
$$LBP(x_c, y_c) = \sum_{p=0}^{P-1} 2^p s(i_p - i_c)$$

Here '(Xc,Yc)' is considered as the central pixel with intensity 'Ic'. And 'In' being considered as the intensity of the neighbor pixel

- It compares a pixel to its 8 closest pixels using this function.

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

- If the value of neighbor is greater than or equal to the central value it is set as 1 otherwise it is set as 0.
- Thus, we obtain a total of 8 binary values from the 8 neighbors.
- After combining these values, we get an 8 bit binary number which is translated to decimal number for our convenience.
- The obtained decimal number is said to be the pixel LBP value and its range is 0-255.



- After the generation of LBP value histogram for each region of the image is created by counting the number of similar LBP values in the region.
- After creation of histogram for each region all the histograms are merged to form a single histogram and this is known as the feature vector of the image.
- Now we compare the histograms of the test image and the images in the database and then we return the image with the closest histogram.
- We can use various kinds of approaches to compare the histograms (calculate the distance between two histograms), for example: Euclidean distance, chi-square, absolute value, etc.

- The Euclidean distance is calculated by comparing the test image features with features stored within the dataset. The minimum distance between test and original image gives the matching rate.

$$d(a,b) = \sqrt{\sum_{i=1}^{n} |a_i - b_i|^2}$$

- As an output we get an ID of the image from the database if the test image is recognized.

Figure 6.7 Output of Face Recognition

- LBPH can recognize both side and front faces and it is not affected by illumination a variation which means that it is more flexible

❖ Let us consider an example:-

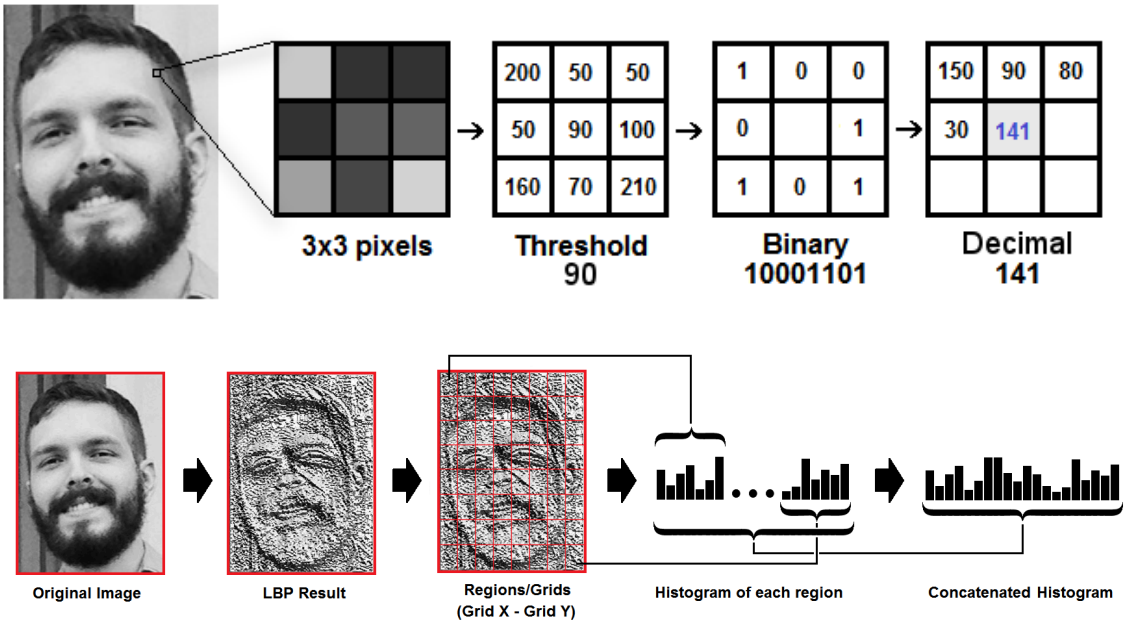


Fig. 6.8 Example of LBPH Methodology

CHAPTER 7

TESTING

- **BLACK BOX TESTING**
- **WHITE BOX TESTING**
- **TEST CASES**

7.1 BLACK BOX TESTING

The technique of testing without having any knowledge of the interior workings of the application is called black-box testing. The tester is oblivious to the system architecture and does not have access to the source code. Typically, while performing a black-box test, a tester will interact with the system's user interface by providing inputs and examining outputs without knowing how and where the inputs are worked upon.

This method of testing is named as black box because the software program, in the eyes of the tester, is like a black box; inside which one cannot see. This method attempts to find errors in the following categories:

- ➢ Incorrect or missing functions
- ➢ Interface errors
- ➢ Errors in data structures or external database access
- ➢ Behaviour or performance errors
- ➢ Initialization and termination errors

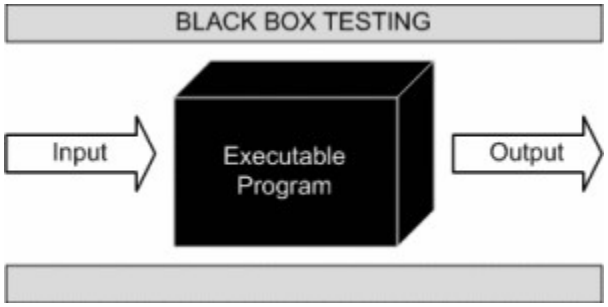


FIGURE 7.1 BLACK BOX TESTING

7.2 WHITE BOX TESTING

White-box testing is the detailed investigation of internal logic and structure of the code. White-box testing is also called glass testing or open-box testing. In order to perform white-box testing on an application, a tester needs to know the internal workings of the code.

This method of testing is named as white box because the software program, in the eyes of the tester, is like a white/transparent box; inside which one clearly sees. The tester chooses inputs to exercise paths through the code and determines the appropriate output.

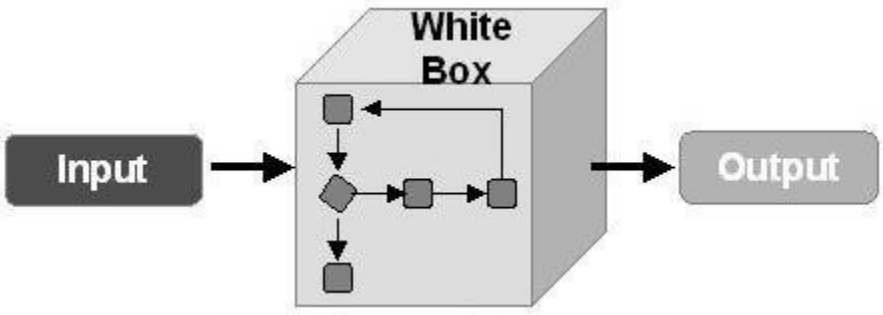


Figure 7.2 White-Box Testing

7.3 TEST CASES

Table 7.1 Test Cases

Test Id	Objective	Prerequisite	Expected Output	Actual Output	Status
1	Take Image	Camera Employee details should already be given	Pictures should be saved in the given folder to train	As expected	Pass
2	Image taken Status	Webcam to be closed after performing “Take Image”	Notification “Images saved for ID: and Name:” should be shown in the notification bar	As expected	Pass
3	Train Image/ Trained image status	Pictures of the employee should be taken before	Pictures should be trained to recognise the employee and should show notification”Images Trained”	As expected	Pass
4	Track Image	Camera Images should already be trained by the system	Employee details to be shown in the camera	As expected	Pass
5	Attendance	Camera Employee should have already passed through the camera	Attendance should be filled in the new csv file which changes daily	As expected	Pass
6	Attendance Status	Webcam to be closed after filling attendance	Notification “ Name: ID: Date: Time: “ should be shown in the attendance bar	As expected	Pass
7	Quit	Face recognition software should be	Close the software	As expected	Pass

		running			

CHAPTER 8

SYSTEM DESIGN

- **SYSTEM FLOW CHART**
- **TRAINING FLOW CHART**
- **FACE DETECTION AND RECOGNITION FLOW CHART**

8.1 SYSTEM FLOW DIAGRAM

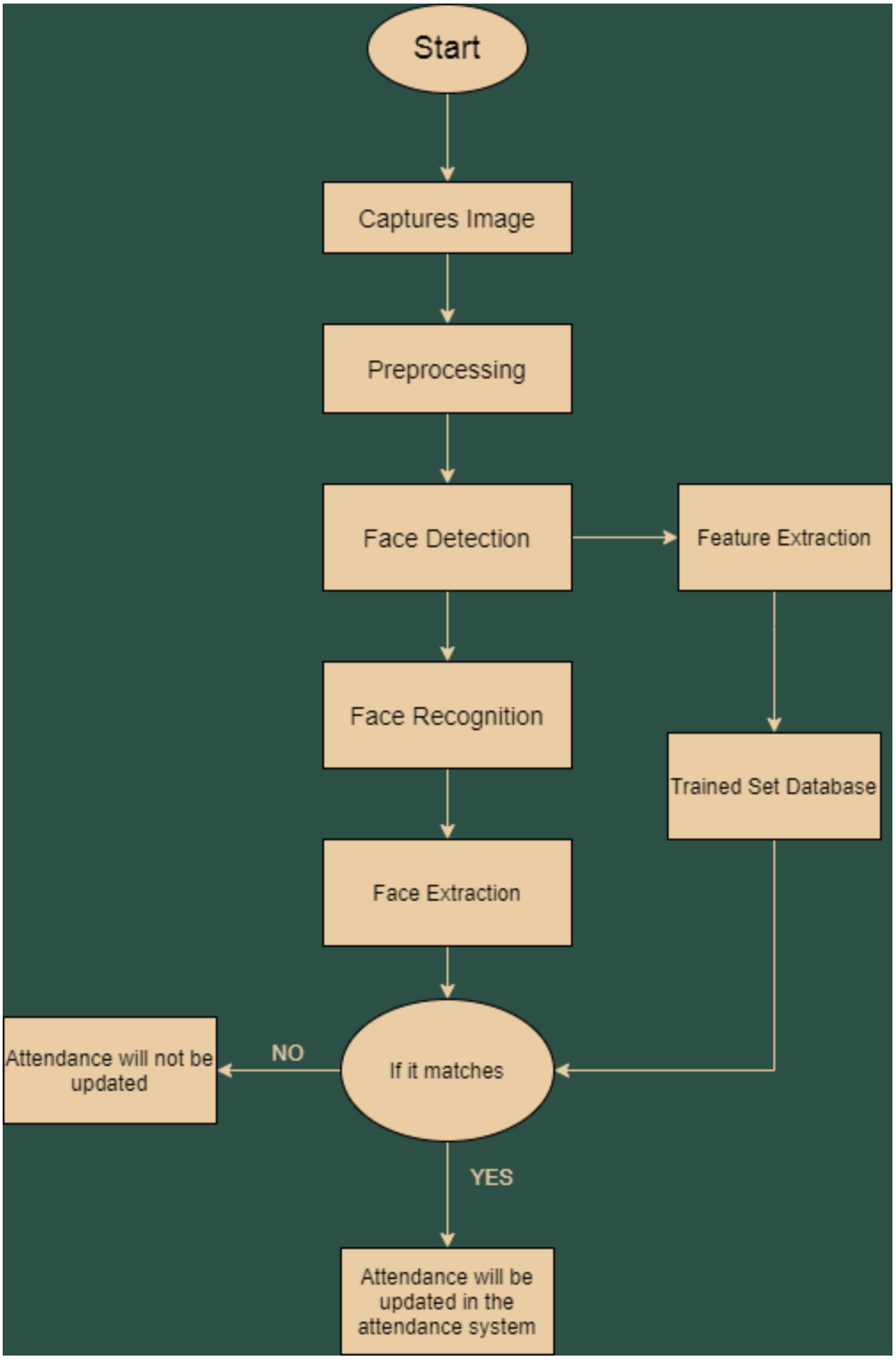


Figure 8.1 SYSTEM FLOW DIAGRAM

8.2 TRAINING FLOW DIAGRAM

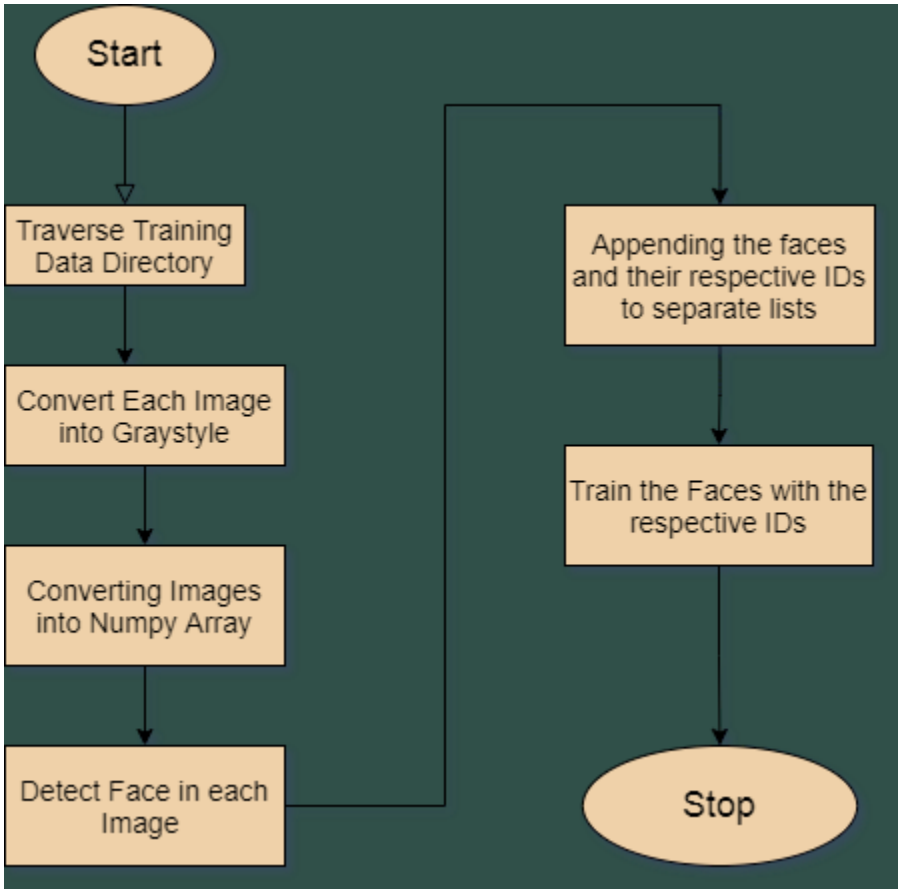


Figure 8.2 TRAINING FLOW DIAGRAM

8.3 FACE DETECTION AND RECOGNITION FLOW CHART

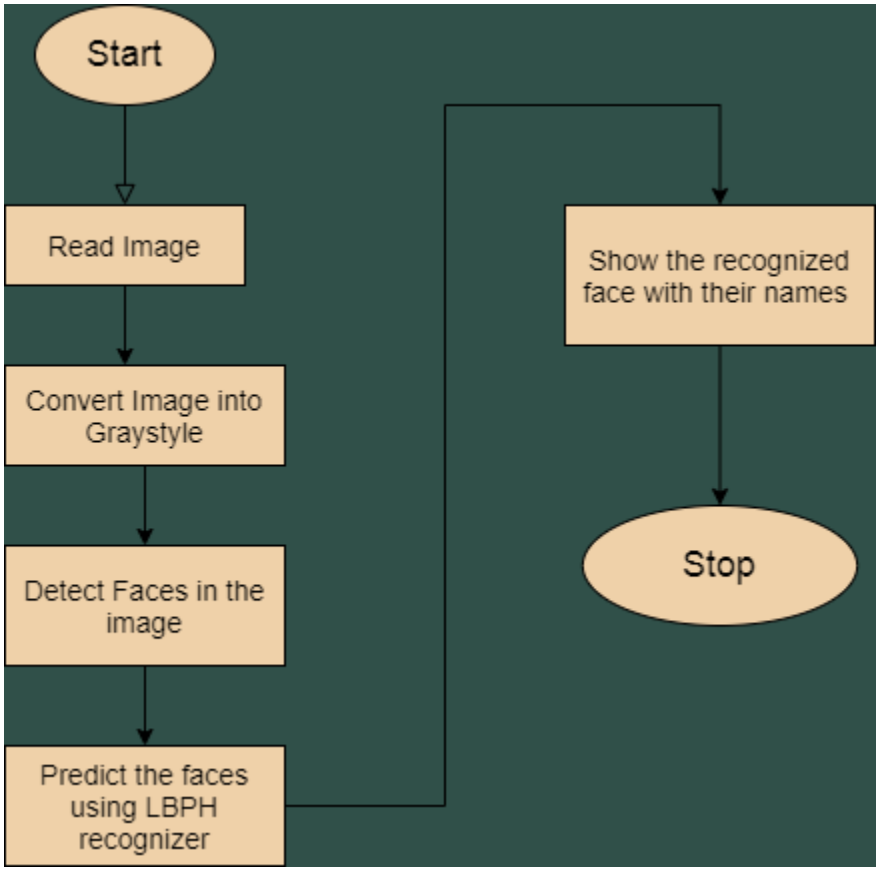


Figure 8.3 FACE DETECTION AND RECOGNITION FLOW CHART

CHAPTER 9

IMPLEMENTATION

- **WORKING OF OUR SYSTEM**
- **SCREENSHOTS**

9.1 WORKING OF OUR SYSTEM

When we run the program, a window will pop out and asks for Enter Id and Enter Name. After entering respective name and id fields then we have to click the Take Images button. By clicking the Take Images button, a camera of the running computer is opened and it starts taking image samples of the person. This Id and Name is stored in the Student Details folder and file name is saved as Student Details.csv. It takes 60 images as samples and stores them in the Training Image folder. After completion it notifies that images are saved. After taking image samples in order to train the image samples we have to click the “Train Image” button. Now it takes a few seconds to train the machine for the images and creates a Trainner.yml file and stores them in the TrainingImageLabel folder. Now all initial setups are done. After completion of take images and Train images we have to click the Track images button which is used to track the faces. If the face of a particular student is recognized by the camera then Id and Name of person is shown on Image. After coming out of it, attendance of a particular person will be stored in the Attendance folder as a csv file with name, id, date and time and it is also available in the window.

9.2 SCREENSHOTS

1. Front View:

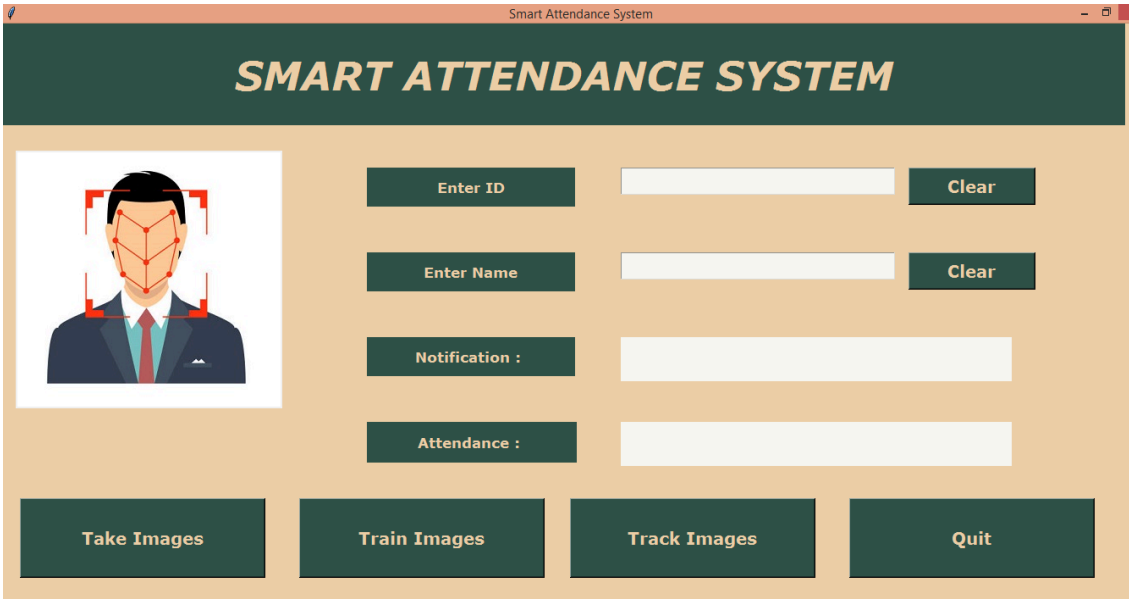


Figure 9.1 Front End of the system

2. Enter name and ID of the employee.

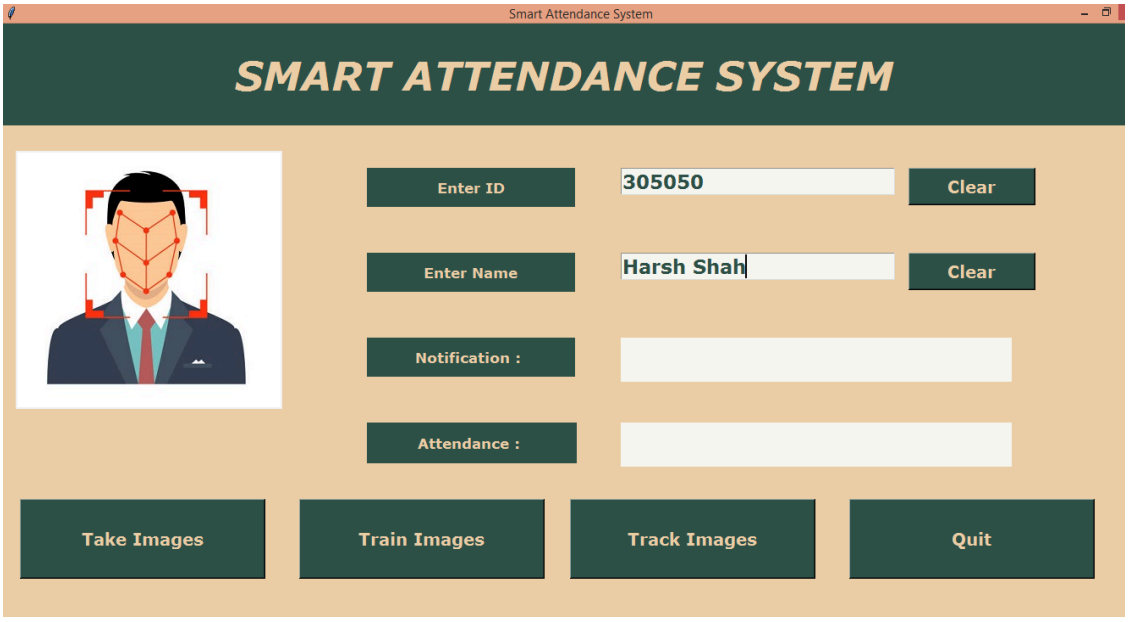


Figure 9.2 Registration

3. Clear Button will clear out the entry of the respective label.

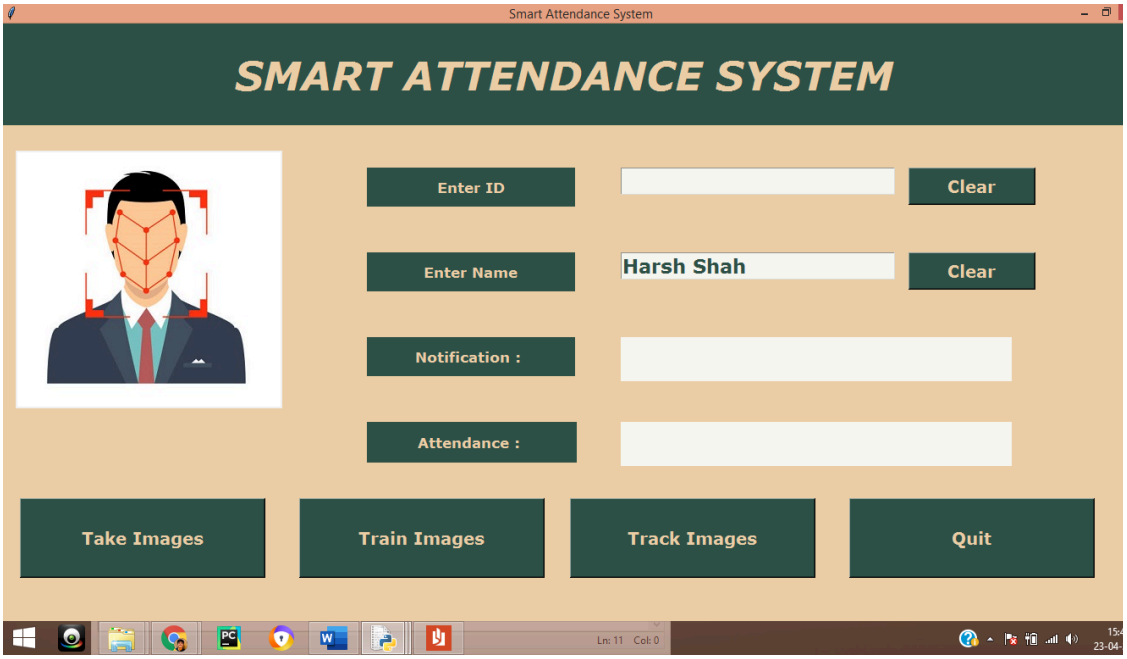


Figure 9.3 Clearing ID

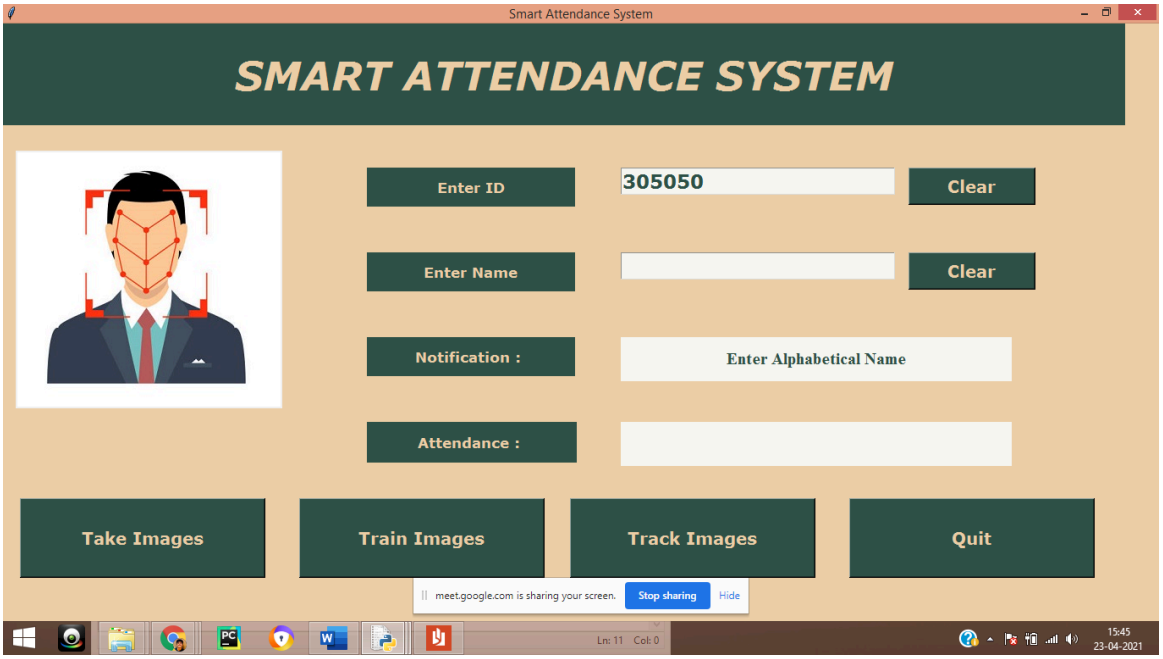


Figure 9.4 Clearing NAME

4. Notification will pop-up if the admin has left out any entry.

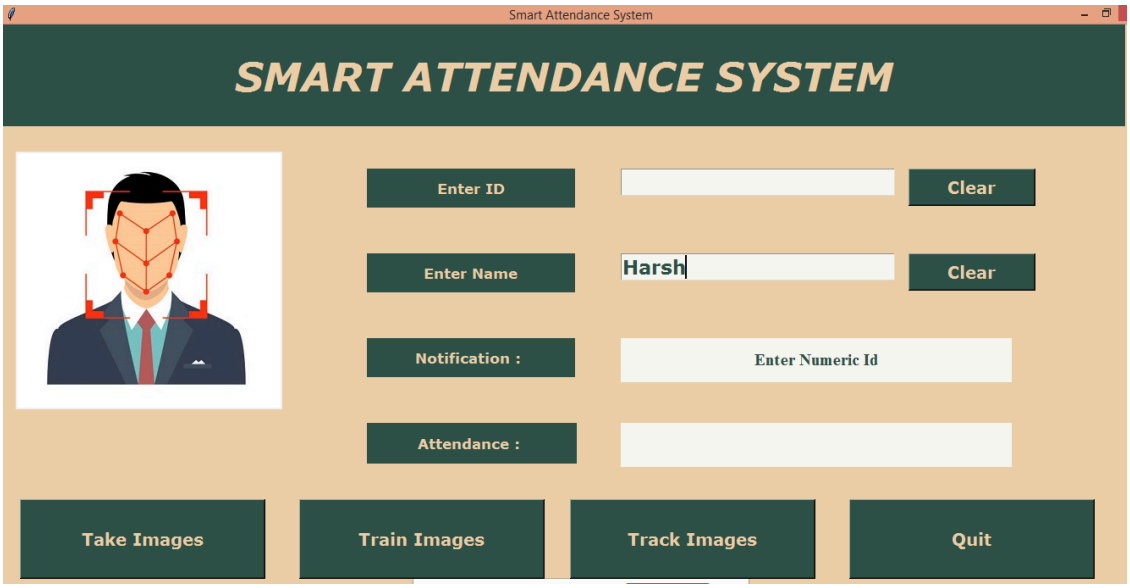


Figure 9.5 USER ID Validation

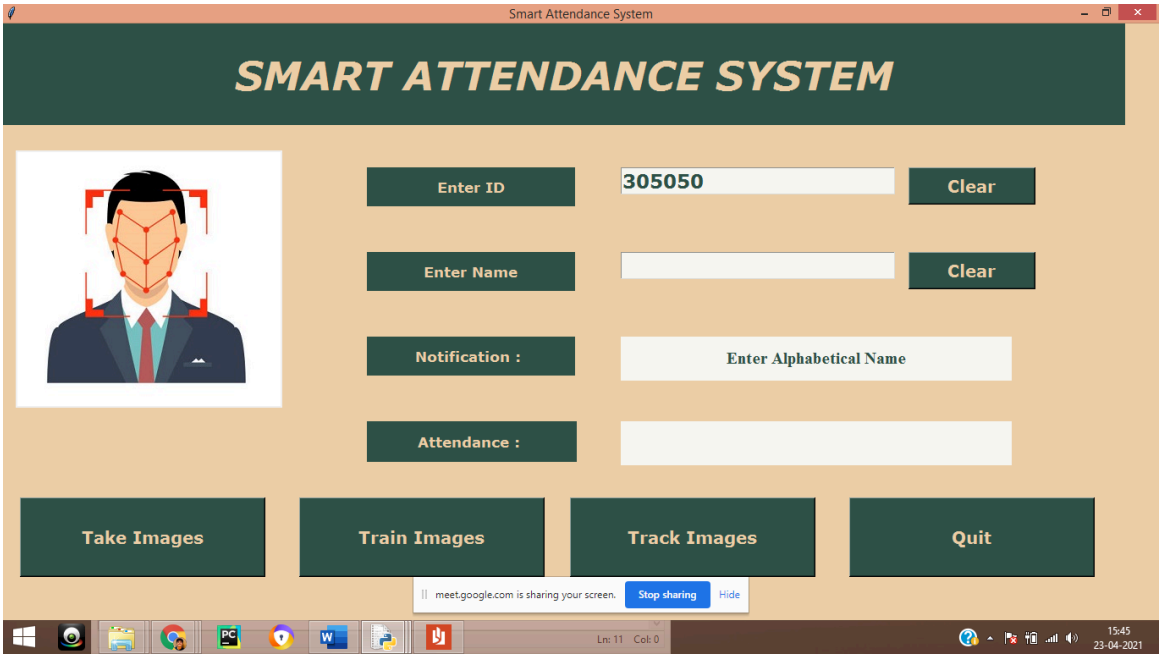


Figure 9.6 User NAME Validation

5. By pressing Take Images, the camera will capture the image of a particular employee.

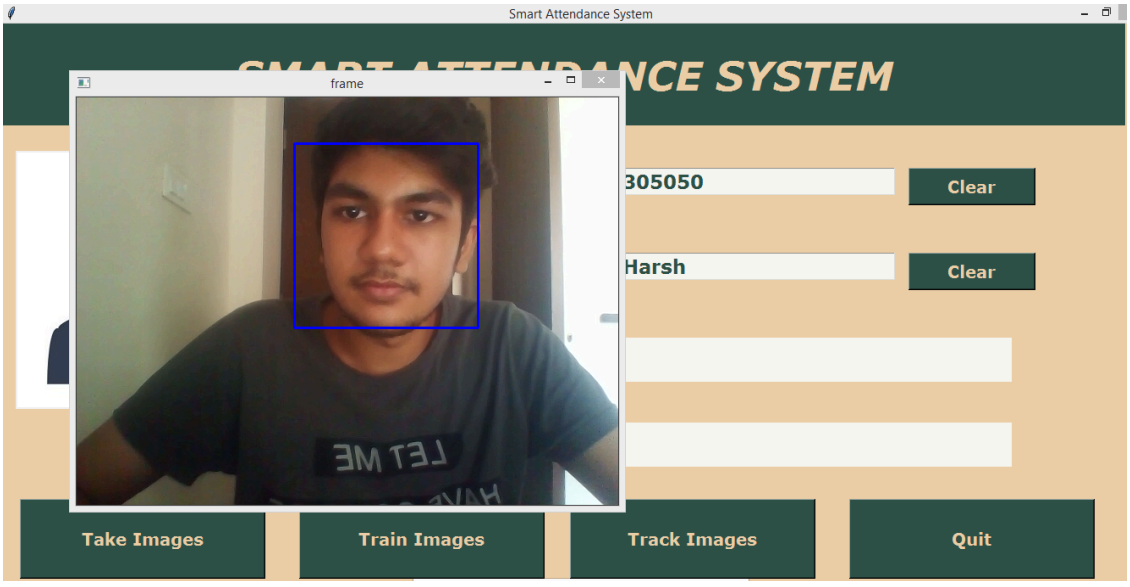


Figure 9.7 Face Detection Case-1

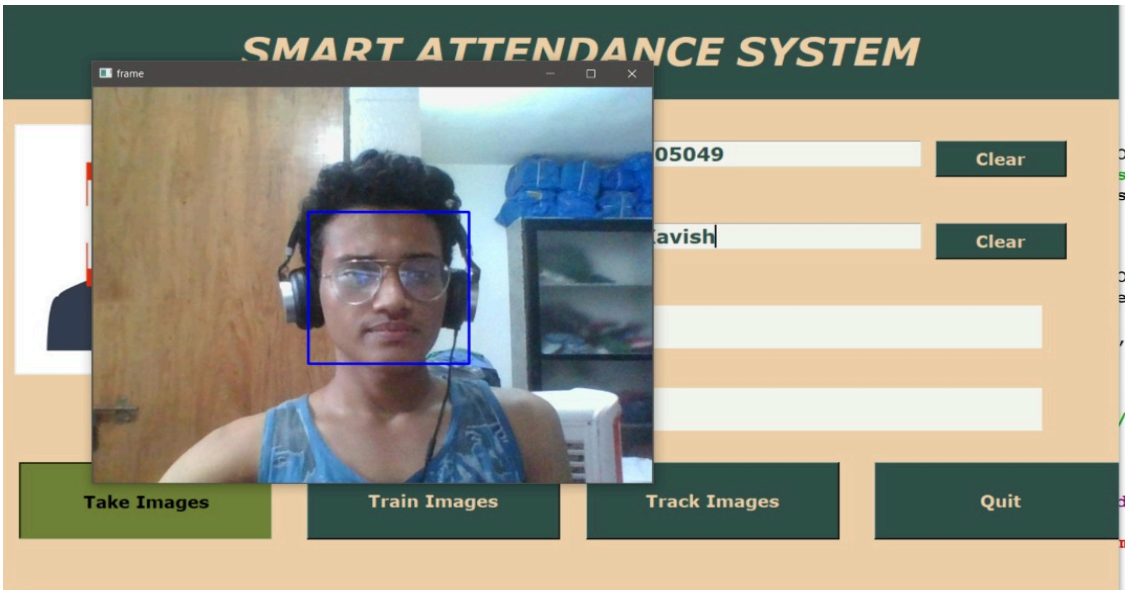


Figure 9.8 Face Detection Case- 2

6. A notification message displayed like an image saved for a particular employee with id and name.

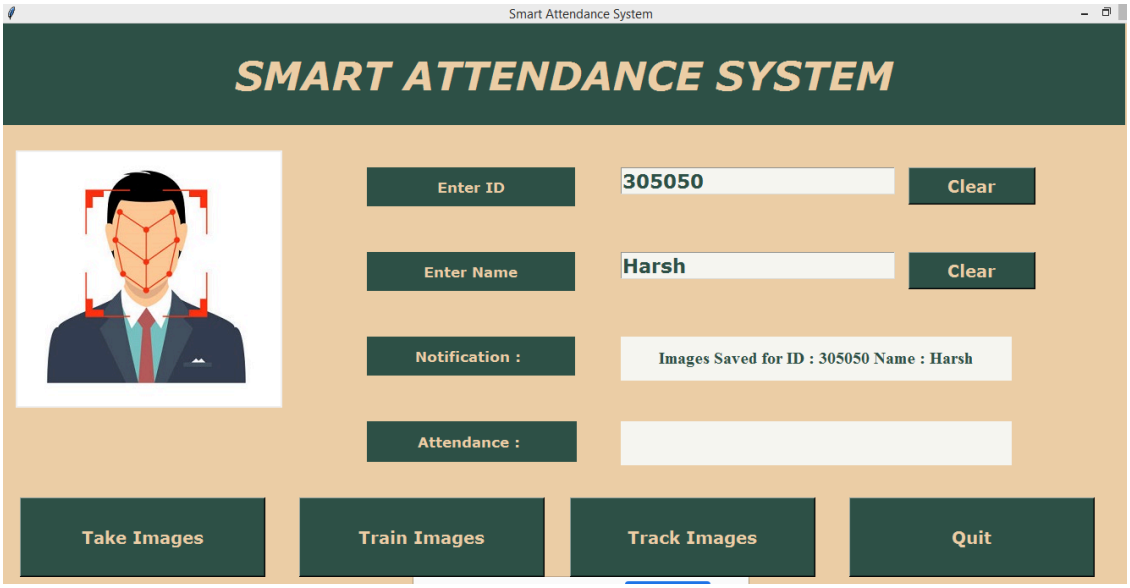


Figure 9.9 Images Saved Notification

7. Clicking on Train image button, it displays a notification message like “Image Trained”

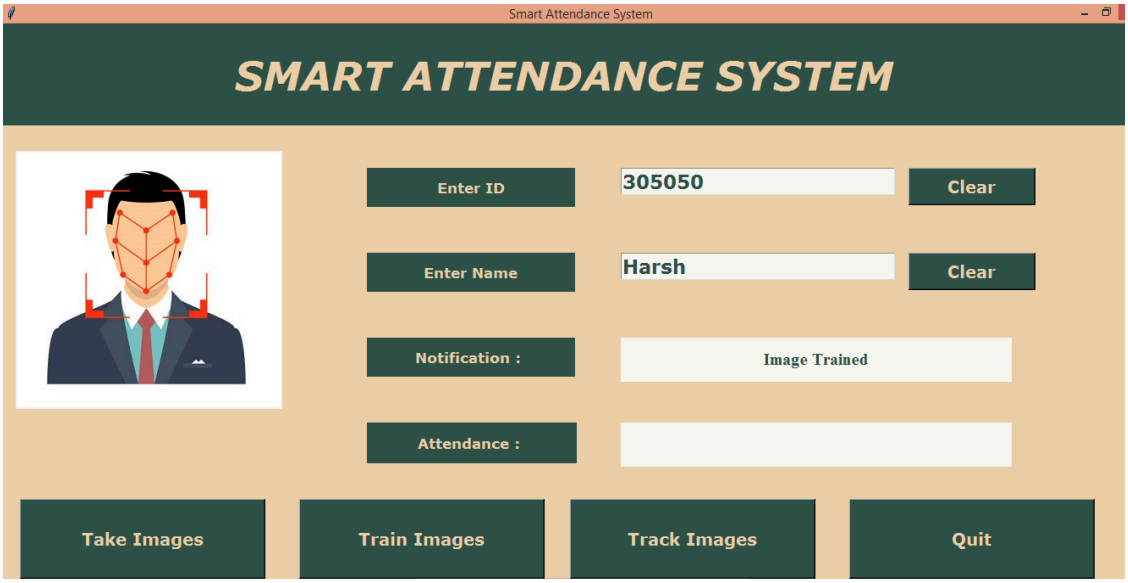


Figure 9.10 Images Trained Notification

8. On clicking the track image button, it recognizes the face (which is already trained) and displays the name and id of the particular person.

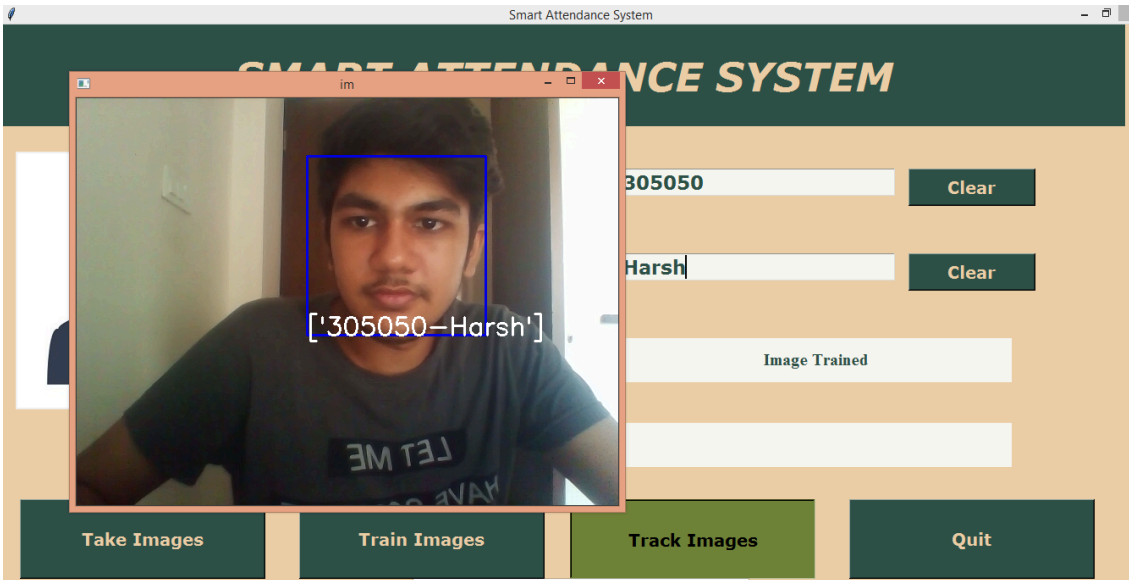


Figure 9.11 Face Recognition Case-1

Figure 9.12 Face Recognition Case - 2

9. Attendance is updated as shown in the attendance bar.

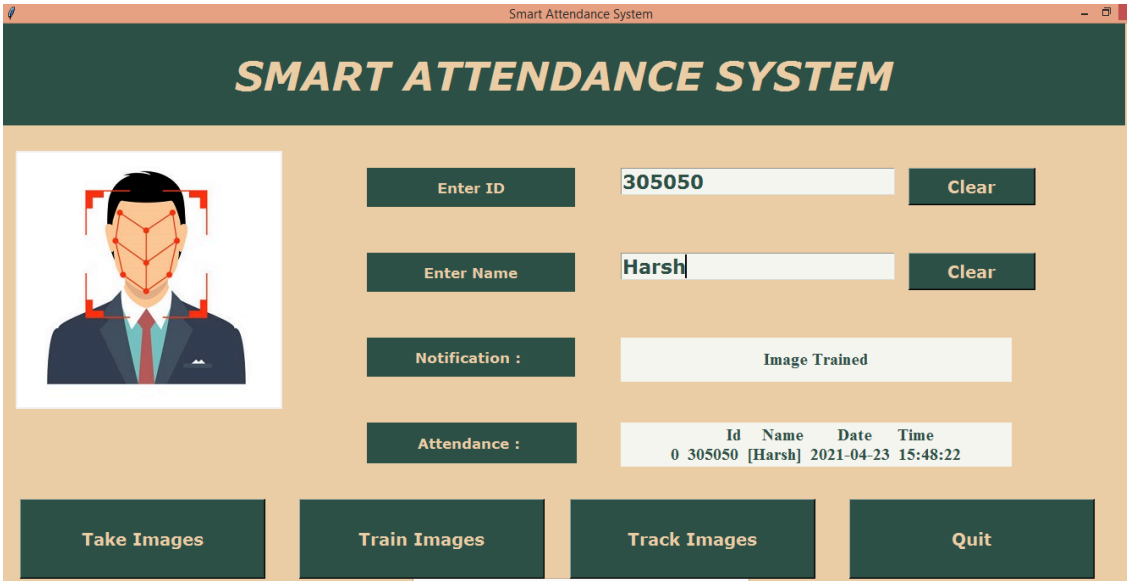


Figure 9.13 Attendance Marked Notification

10. An Attendance file will be created of that particular date.

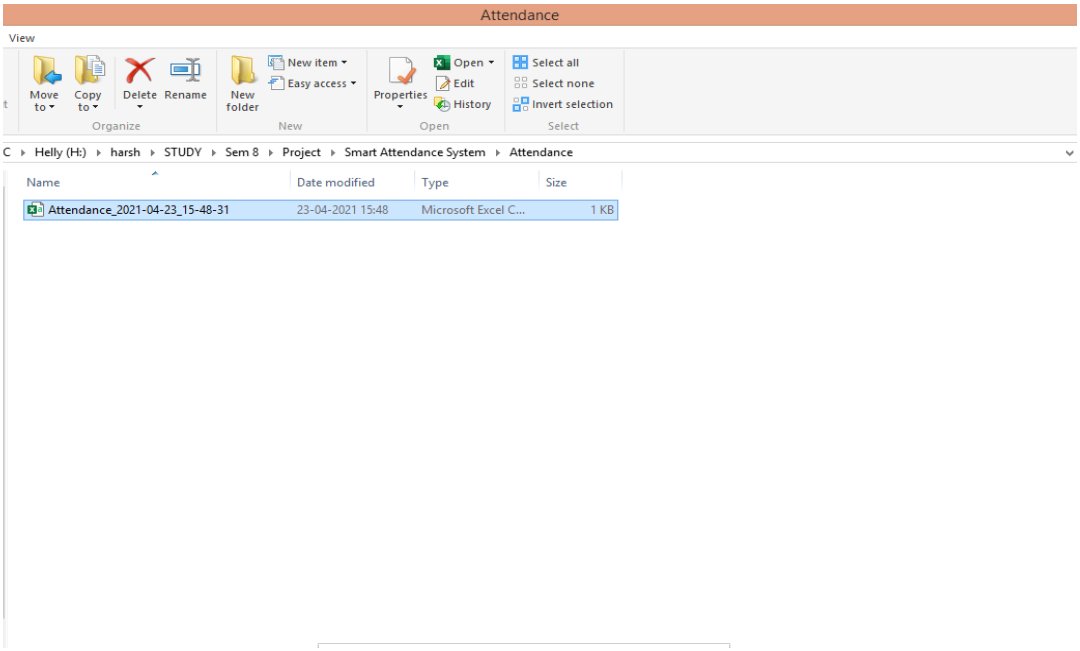


Figure 9.14 Generate Attendance Sheet

11. Attendance of particular epmloyee is updated in the “Attendance file”.

	A	B	C	D	E	F	G	H	I	J
1	Id	Name	Date	Time						
2	305050	['Harsh']	23-04-2021	15:48:22						
3	305049	['Kavish']	23-04-2021	15:57:06						
4										
5										
6										
7										
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										

Figure 9.15 Updation of Attendance

CHAPTER 10

LIMITATIONS AND FUTURE ENHANCEMENT

- **LIMITATIONS**
- **FUTURE ENHANCEMENT**

10.1 LIMITATIONS

- ➢ This system will not work without a computer and a webcam.
- ➢ It threatens individual privacy.
- ➢ It may create data vulnerabilities.
- ➢ Technology is still imperfect.
- ➢ It is costly.

10.2 FUTURE ENHANCEMENT

- ➢ Surveillance and Security can be increased drastically using this technology.
- ➢ School , Universities and Health care can start using face recognition for better management.
- ➢ Fraud can be detected more easily with advancements.

CHAPTER 11

CONCLUSION

- **CONCLUSION**

11.1 CONCLUSION

We have implemented an attendance management system for student's attendance. It helps to reduce time and effort, especially in the case of a large number of students marked attendance. The whole system is implemented in the Python programming language. Facial recognition techniques used for the purpose of the student attendance. And also, this record of student attendance can further be used mainly in exam related issues like who are attending the exams and who are not attending. On this project, there is some further work remaining to do like installing the system in the office. It can be constructed using a camera and computer.

Bibliography

LITERATURE:

[1] Ononiwu G. Chiagozie, Okorafor G. Nwaji, “Radio-frequency identification (RFID)based Attendance System with Automatic Door Unit”, Academic Research International(2012).

[2] O. Shoewu, PhD, O.A. Idowu, B.Sc., “Development of Attendance Management System using Biometrics.”, The Pacific Journal of Science and Technology(2012).

[3] A. Khatun, A. K. M. F. Haque, S. Ahmed and M. M. Rahman, ”Design and implementation of iris recognition based attendance management system”, 2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), Dhaka, 2015, pp.1-6.

[4] Kadry, Seifedine Smaili, Mohamad. (2010). ”Wireless attendance management system based on iris recognition”. Scientific Research and Essays. 5. 1428-1435.

[5] Jomon Joseph1, K. P. Zacharia, “Automatic Attendance Management System Using Face Recognition”, International Journal of Science and Research (IJSR), 2013.

[6] Bharath Tej Chinimilli, Anjali T, Akhil Kotturi, Vihas Reddy Kaipu, Jathin Varma Mandapati, “Face Recognition based Attendance System using Haar Cascade and Local Binary Pattern Histogram Algorithm”, Fourth International Conference on Trends in Electronics and Informatics (ICOEI 2020).

[7] Kritika Shrivastava, Shweta Manda , “Conceptual Model for Proficient Automated Attendance System based on Face Recognition and Gender Classification using Haar-Cascade, LBPH Algorithm along with LDA Model”, International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 10 (2018) pp. 8075-8080.

[8] Dr.B.Kameswara Rao, Anusha Baratam, Gudla Shiridi Venkata Sai, B.Radhika, A.Vineeth, "ATTENDANCE SYSTEM BASED ON FACE RECOGNITION USING LBPH", International Journal of Creative Research Thoughts (IJCRT), Volume 8, Issue 5 May 2020 | ISSN: 2320-2882

[9] Abhishek Jha, "Class room attendance system using facial recognition system", The International Journal of Mathematics, Science, Technology and Management (ISSN : 2319-8125) Volume: 2, Issue: 3, 2014.

[10] Sudhir Bussa, Shruti Bharuka, Ananya Mani, Sakshi Kaushik , "Smart Attendance System using OPENCV based on Facial Recognition", International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 ,Vol. 9 Issue 03, March 2020.

LINKS

https://docs.microsoft.com/en-us/azure/machine-learning/concept-deep-learning-vs-machine-learning

https://docs.opencv.org/3.4/d2/d99/tutorial_js_face_detection.html

https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_objdetect/py_face_detection/py_face_detection.html

https://towardsdatascience.com/face-detection-with-haar-cascade-727f68dafd08

https://towardsdatascience.com/face-recognition-how-lbph-works-90ec258c3d6b

https://iq.opengenus.org/lbph-algorithm-for-face-recognition

https://thesai.org/Downloads/Volume10No5/Paper_35-LBPH_based_Enhanced_Real_Time_Face_Recognition.pdf