Exclusive Premises Access Monitoring System

Mahesh Kumar Gupta
Assistant Professor, Department of Mechanical Engineering
SRM Institute of Science and Technology, Delhi-NCR Campus, Modinagar, Uttar Pradesh, India

Siddham Sharma
Department of Mechanical Engineering
SRM Institute of Science and Technology, Delhi-NCR Campus, Modinagar, Uttar Pradesh, India

Ayushman Thakur
Department of Mechanical Engineering
SRM Institute of Science and Technology, Delhi-NCR Campus, Modinagar, Uttar Pradesh, India

Utkarsh Sen
Department of Mechanical Engineering
SRM Institute of Science and Technology, Delhi-NCR Campus, Modinagar, Uttar Pradesh, India

Melvin Emmanuel
Department of Mechanical Engineering
SRM Institute of Science and Technology, Delhi-NCR Campus, Modinagar, Uttar Pradesh, India

Shahnawaz Alam
Assistant Professor, Department of Mechanical Engineering
SRM Institute of Science and Technology, Delhi-NCR Campus, Modinagar, Uttar Pradesh, India

Abstract- Security of personal properties or exclusive events is a vital, furthermore a hectic task to perform and observe. Plenty of human resources as well as valuable company time is spent on performing it manually. The resources can marginally be saved in these technologically advanced days. The idea of the uniqueness of a human face is being explored here for identification as well as verification purposes. Such levels of automation options are researched with the help microcontroller device, namely Raspberry Pi. OpenCV framework has been brought into application here. The entire suggested system can be broken down to five functional parts namely- Detection, Processing, Training, Recognition and Buzzer/Notification Module. Here the usage of Raspberry Pi maintains the viability of the project with sufficient automation requirements being fulfilled. The system stores database in order to acknowledge the features for desired and allowed guests, residents as well as owners of the premises. Similarly, authentication of the same is also being performed using the same technology. In this paper, utilization and application of Viola-Jones' face detection concept with Haar's Cascades algorithm is also done. Local Binary Pattern technology is also exploited in face recognition while Service from Amazon have been utilized (namely S3) for cloud upload of images. Twilio Notifier has been used for communicating uploaded pictures to a pre-defined user smartphone. Hence, alerting the owner, resident or the responsible body with a notification or a buzzer, about any possible breach in the secured premises by any intruder.

Keywords – Adaboost classifier, HaarCascade algorithm, Facial database, Access control system

I. Introduction

Face recognition is one of the main implementations of the biometric-based authentication method over the past few decades. Face recognition is fairly recognition activity pattern, where a face is labelled as either recognized or unknown by comparing it to a recognized person's photographs stored within the database. Face recognition may be a difficulty due to some variability in knowledge due to random variation among different populations, including systemic variations from different factors such as lightening and posing conditions. Many problems must be dealt with by statistical methods of face recognition. Such similar patterns of difficulties arise because faces must be interpreted in such a simplest way that the available face knowledge is best used to distinguish a selected face from all the opposite faces within the database.

While figuring out this aspect, poses of the human face can be a particularly difficult problem just because all faces appear similar; particularly, all faces consist of two eyes, mouth, nose, and other features within the same area. The exterior part of the body is a highly complex and dynamic system with characteristics which may change in time dramatically and rapidly. Scientists seek to understand the design of the external part of the body while constructing or designing face recognition systems with the help of machine learning. The expertise and experience in external body component recognition system techniques may enable researchers to understand the fundamental structure. External body part recognition system uses any or all of the senses, such as visual, auditory, and tactile data. Each of those data is used to memorize and store facets either individually or together.

Conditions around the individual are critical in an overly face recognition system in many instances. The management and combining of substantial data is complicated for a computer recognition program. However, it is often difficult to memorize several names. A computer system's main benefit is that it has the memory space. Human features that can be used to recognize the face are being constantly researched, and arguably so. Local as well as global features are required to recognize the face.

The substantial growth is due to the availability of real-time equipment, the increasing need for surveillance applications, a rising focus on commercial civilian research projects as well as studies on natural network classifiers, which emphasize real-time computing and adaptation. Two classifications fall under the face recognition system: verification and identification. Face verification (one-to-one matching) which compares the image of the face against the images of a prototype whose identity is claimed. Face identification (one-to-many matching) that compares a face image of a question against all image templates in a database of the very face.

Although the face recognition (one-to-many) system works to identify the unknown face and mark them as intruders, in these types of hardware monitoring system the response system is a requirement because alerting the user about an intruder is much more important than just spotting him / her on the live stream. There are plenty of things that can be done to warn the user depending on the user 's requirement and the situation in which this device is being used.

The approach we use in our "Exclusive Premises Access Monitoring System" is that the Face Identification (one-to-many matching) software that compares a face picture query, i.e. an intruder's face or one that tries to access the premises, against all image models in a very face database that is additionally created by us providing data of all user-related facets.

The first and foremost thing that should be done is to attach a buzzer or alarm device inside the premises and that should be done in case there is no need for remote access to the system and either the customer or his / her associate is present to tackle the situation if any intruder is identified .

The second choice usually considered and more sophisticated and likely to use is a Thing Internet system, where the computational device defining the faces is linked to the internet. When an intruder is detected the device will upload his / her pictures to a cloud-based storage device with the time and other important information and from there this information will be sent to the user on his / her mobile with the help of a Notifier API.

Our aim is to warn the user in the most likely way when an intruder is spotted, making the device that recognizes face, as described above.

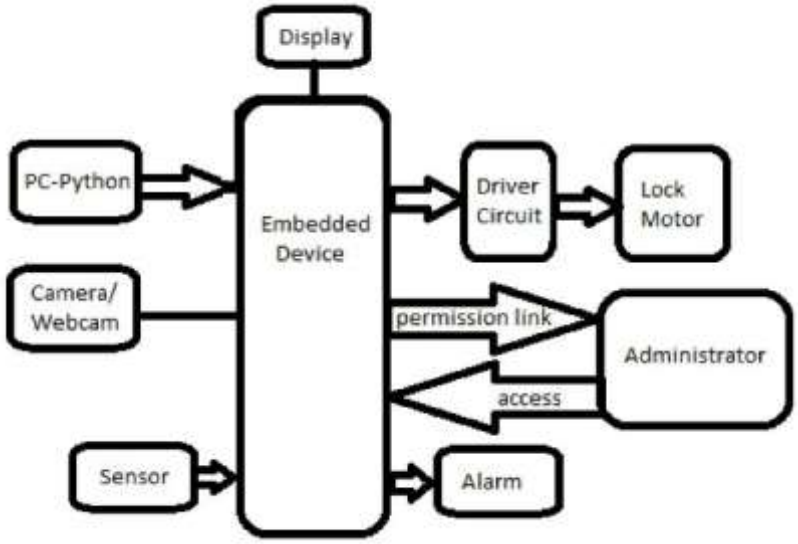

Figure 1. Schematic representation of simplified control access system with proposed design

II. Materials

Our system required a high-performance processor, and for that specific reason, Raspberry Pi 3 Model B module was selected as the viable choice. This Raspberry Pi module runs on a Linux operating system, based on Debian, known as Raspbian. Other modules such as the Module 3B+, Module 4, etc can also be considered based on availability, compatibility, size limitation and viability.

The system has sequential functions that are laid out as follows:

- Firstly, a Passive Infrared Sensor, also known as a PIR sensor is employed for sensing the presence of a human. It successfully detects if a human comes in the proximity of the door, or the entry to which the system is installed.
- The PIR sensor captures a set of images of the human standing in the proximity of the door/entrance, in front of the camera and in its field of view.
- The image is sent to the remote user or the system administrator via mail (with password protection available).
- Here, Principle Component Analysis (PCA) [8] and Neural network algorithms [12] are used to process the captured image with the help of Open Computer Vision (Open CV) in the Raspbian processor, in order to compare the present image with the image already saved in the database of the system.
- If the face gets recognized, i.e. if the current image matches the one in the database, the door is unlocked and the image is sent to the system administrator with specific date and time of entry. In all other cases, same details, i.e. image with date and time of approach, are sent to the administrator in the password protected link. The administrator then has the option of either granting access to the visitor or denying it.

Detailed description of all the major functioning elements that are employed in this IOT Access Control System are given as follows:

RASPBERRY Pi MODULE: Raspberry Pi is a series of single board computers, termed as microcontrollers. The module we are using here specifically is the Raspberry Pi 3 Model B (as shown in Figure 2). It consists of a 1.2 GHz 64-bit quad core processor, with Bluetooth, Wi-Fi, and USB boot functionalities available. Raspberry Pi does not have a built-in real-time clock in any of the modules. The time has to be set either manually while booting, or from previously saved configurations.

Figure 2. Raspberry Pi Module 3 Model B

ESP32 – CAMERA MODULE: ESP32-CAM (as shown in Figure 3) is a small size camera module with sufficiently low power consumption. It consists of an OV2640 camera and an onboard TF card slot for data storage. It also has Bluetooth and Wi-Fi capabilities. The camera employed here is equipped with flash, with total power consumption 310mA at 5V (with flash) and 180mA at 5V (without flash). It also has a working temperature range of -20 ℃ to 85℃. This camera can be used in multiple applications ranging from QR scanning, wireless video monitoring to image capturing.

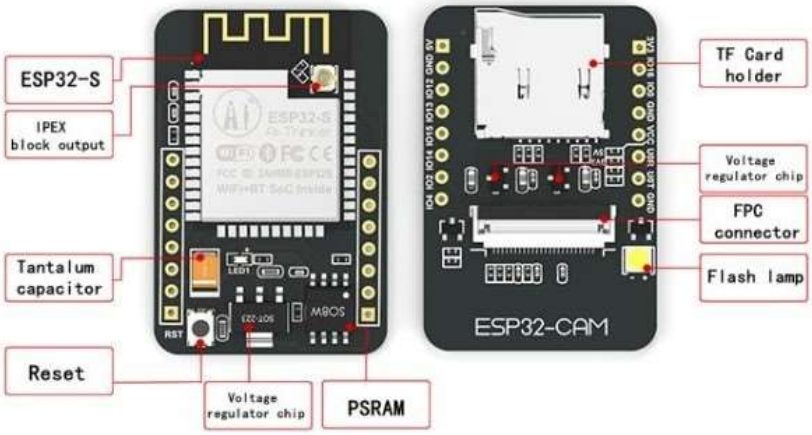

Figure 3. ESP32 Camera Module with component representation

SERVO MOTOR: Servomotor is a type of rotary or linear actuator that provides precise control over linear or angular position, acceleration and velocity. It requires a controller module, in our case, the Raspberry Pi 3 Model B, which can be used to program it accurately. A basic general servo motor can be programmed in many ways to provide specific degrees of rotation depending upon usage and application and hence is considered a versatile device.

OPEN CV: Open Computer Vision or Open CV is a real-time computer vision tool which is a library of programming functions. Open CV is free for use under open-source BSD license and can be accessed for a variety of programs and functions as required by the user.

GSM MODULE: GSM module is a device which employs mobile telephone technology and in-turn provides a data link to a given remote network. It stands for Global System for Mobile. These GSM modules (as shown in Figure 4) are used as a part of embedded systems to provide a wide array of functionality options.

Figure 4. GSM Module with dedicated SIM Slot

ARTIFICIAL NEURAL NETWORKS: Neural Networks or Artificial Neural Networks (ANN) are systems consisting of series of algorithms that perform tasks already programmed by using pattern recognition and classification. They often use multi-layer perception method to obtain a usable data map. Normally, predefined model (or test image/object) is used which could be used by the artificial neural network for comparison and classification.

PASSIVE INFRARED SENSOR: Passive Infrared Sensors (or PIR Sensors) are hardware sensor units which detect foreign infrared radiations reflected or radiated by different objects (as shown in figure 5). They do not generate radiation by themselves, hence the name passive infrared sensor. They are generally considered for usage in form of motion detectors or even proximity detectors due to their capability to sense or detect radiations in their given field of view.

Figure 5. Passive Infrared Sensor Pmod PIR

III. Methodology

3.1 Dataset Creation -

We want to remember faces that don't seem to be part of the existing dataset, and identify faces of ourselves, friends, family members, colleagues, etc. To do this, we would like to gather images of faces that we would like to remember and then measure in some way. Typically, this process is taken up as the Face Enrolment Recognition Process. We call it "enrolment" because in our dataset and application, we "enrol" and "register" the user as an example person. For this, we built our own custom face recognition dataset, we want to have a selected individual physical access to capture example images of their face. Such a program will be common for businesses, schools or other organizations where the process requires people to be physically present.

For the sake of gathering examples, image sets of those people, whom we may escort to a specially assigned room where a setup with pre-set video camera is present to firstly, detect the x and y coordinates corresponding to their face in a live and continuously streaming video and secondly, writing the frames that contain their face to the disk. We performed this process over several days or weeks to collect samples of their face in specific luminaires, day times, the moods and feelings; creating a more diverse collection of photographs representative of the identity of that individual person.

3.2 Face Detection -

Face detection is the opening for biometrics of the entire face and its accuracy significantly affects the efficiency of sequential operations. A cascade of Adaboost classifier with a haar like feature is intended for face detection in their process. Learning Adaboost helps to pick the most useful features from an outsized feature pool to make a robust classification. With the cascade structure, most non-face samples are easily rejected in the preceding stages to accelerate the process of face detection. Nowadays, a related method employs not only frontal face detection but also multi-view (or multi-pose) face detection algorithms.

The face detection technique can also be extended to localization of landmarks (e.g., eyes, nose tip, and mouth), which can then be used to face geometrical normalization. Common approaches to locating facial landmarks are also the Active Form Model (ASM) and Active Appearance Model (AAM) (Coote set al., 1995). In ASM, facial landmarks are first identified by searching local regions, and then these landmarks are jointly optimized by constraints of global form. AAM employs not only shape constraint but also texture constraint to improve localization accuracy.

Within the two important characteristics for Face Detection, the horizontal and also the vertical characteristics explain what the eyebrows and also the nose look as if to the computer respectively which eventually helps the computer to understand what an image really is. Also, each feature contains a value of its own when the images are checked.

The general concept for face detection is to combine the HaarCascade face detection and the methods of Local Binary Pattern Histogram (LBPH). Because LBPH needs grayscale cropped face from the detection of HaarCascade, we can use this detection to take pairs of user pictures so that they can be stored in our face recognition dataset and we can recognize the user later. The source code for this part, which we will shortly clarify.

3.3 Face Recognition -

Facial recognition is the job of making a correct identification of a face against a pre-existing facial database in a picture or video image [16]. It starts with detection, which we discussed in the previous section, which distinguishes human faces from other objects in the image and then works on identifying those detected faces. It uses the basic biometric identification scheme, but on an increasing, algorithmic scale. Recognition technology looks at data where we see a picture. It stores and access the stored information.

Below, we discuss how the biometric identification works; the technologies differ but here are the basic phases:

- Phase 1: You catch a picture of your face from a photo or video. Your face can appear alone or in a crowd of excesses. Your picture could show you looking straight ahead, or almost in profile.
- Phase 2: The FERET authentication program reads your facial geometry to recognize and remember you. Primary variables include the area between your eyes, and the distance between your forehead and chin.
- Phase 3: The mathematical formula for the facial signature is compared to a database of recognized faces.
- Phase 4: Whether or not the face fits within the stored faces in the dataset will form a determination.

There are two simple kinds of errors we observed:

- False Negative Error: A "false negative" is when the face recognition device does not fit the face of a person with an image stored in the database.
- False Positive Error: A "false positive" is when, in a very database, the face recognition system matches a person's face to a picture, but that match is really wrong.

It is important to have a careful look at the rate of "false positive" and also the rate of "false negative," since there is almost always a trade-off.

If a face is inserted into the view of our surveillance camera and the face cannot be detected, then a buzzer is activated and if the user needs remote access to the monitoring system then a text message notification containing an intruder's snapshot can be sent to his / her smartphone. We ensured that the "intruder" was labelled as "Unknown" for a sufficient number of frames before activating the buzzer or sending this text message to guard against false-positive detections. Whenever the main code requiring the processing of a video stream and the derivation of information from the frames

is implemented, it is good practice to ensure that an event, activity or identification takes place for a pre-set number of frames before sending an alert or taking further action.

3.4 The Reaction Setup -

Research has shown that humans respond to the reliable or inaccurate alarm system with predictability. Our purpose was not just to match faces on live stream with our face dataset to recognize whether the individual is recognized or an intruder; the main objective was to let the user know whether an intruder is trying to enter the property / location.

The most important thing in the entire system to accomplish this function is the aspect in which the user is warned or notified, whether with the aid of a buzzer or an instant message on his smartphone.

We addressed two ways of alerting the user in such a situation-

- By adding an alarm buzzer to the device, which is activated if any intruders occur or try to enter the premises. This option is best suited in a condition where remote accessibility to the system or entry is not necessary and the user or his / her partner is there to handle the situation in person.
- Having the system so that when the intruder or someone attempting to enter the premises attempts to do the same thing; their picture will be uploaded to a cloud service (Amazon S3) and then forwarded to the user with a connection, with the aid of some online notifier (Twilio API). So, he / she can remotely grant that person access or deny that person access.

Amazon(S3) is Web warehouse service. It's designed to make computing on a web scale easier for developers. It features a simple web services interface that can be used at anytime from anywhere on the net to store and retrieve any amount of information. It gives any developer access to the same highly scalable, secure, fast, and inexpensive data storage infrastructure that Amazon uses to operate its own global website network.

Twilio Notification enables us to send notification messages from the same unified API to multiple users across different communication channels. We can connect to the user with a single API request via SMS, mobile apps, Facebook Messenger and more. Also, we can identify preferred channels to meet certain users and tag them for more granular sending capabilities.

IV. Result & Discussion

For the successful implementation of our Facial recognition system, it was made to run through a number of theoretical cases and scenarios which in turn helps in making the Access Control System as fool-proof as possible.

The tests have been taken in multiple situations that range from controlled conditions like ample light, proper face orientation, stable face, etc to uncontrolled conditions like low light, unstable face, improper orientation to cases of no light or stability. Analysis is also done for the number of frames captured for a single visitor, for which the minimum has been set to 10 frames.

- Case of a visitor is taken, visiting at nearly 12.30 p.m. in the day. As soon as approach is made to the entrance where the apparatus is installed, the Passive Infrared Sensor comes into function and senses the visitor in proximity.
- ESP32-CAM Module comes into function. The camera module feeds information to the embedded system where face detection step is initiated. Principle Component Analysis is performed and with the help of ANN, a neural network is built and the facial features are stored in the database.
- The system recognises the facial characteristics with the defined test picture in the database. So, the visitor is given the access.
- ON01 is the resulting command after which driver circuit and lock motor operate to grant access to this visitor.

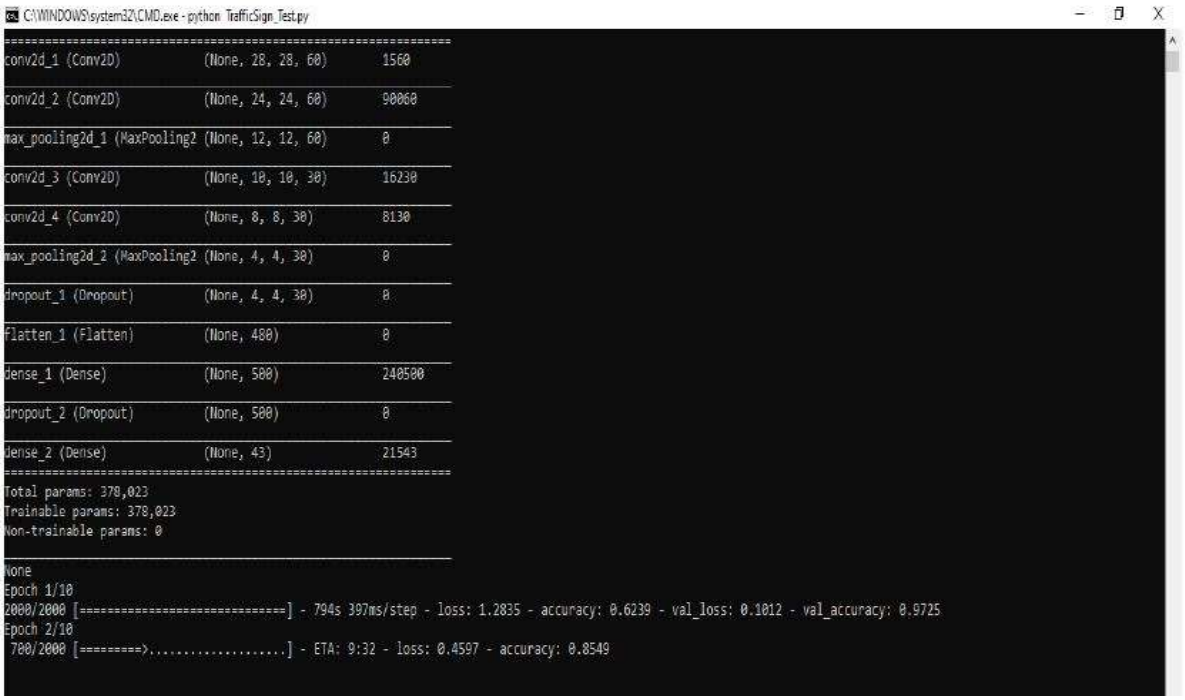

Figure 6. Compilation after execution has started.

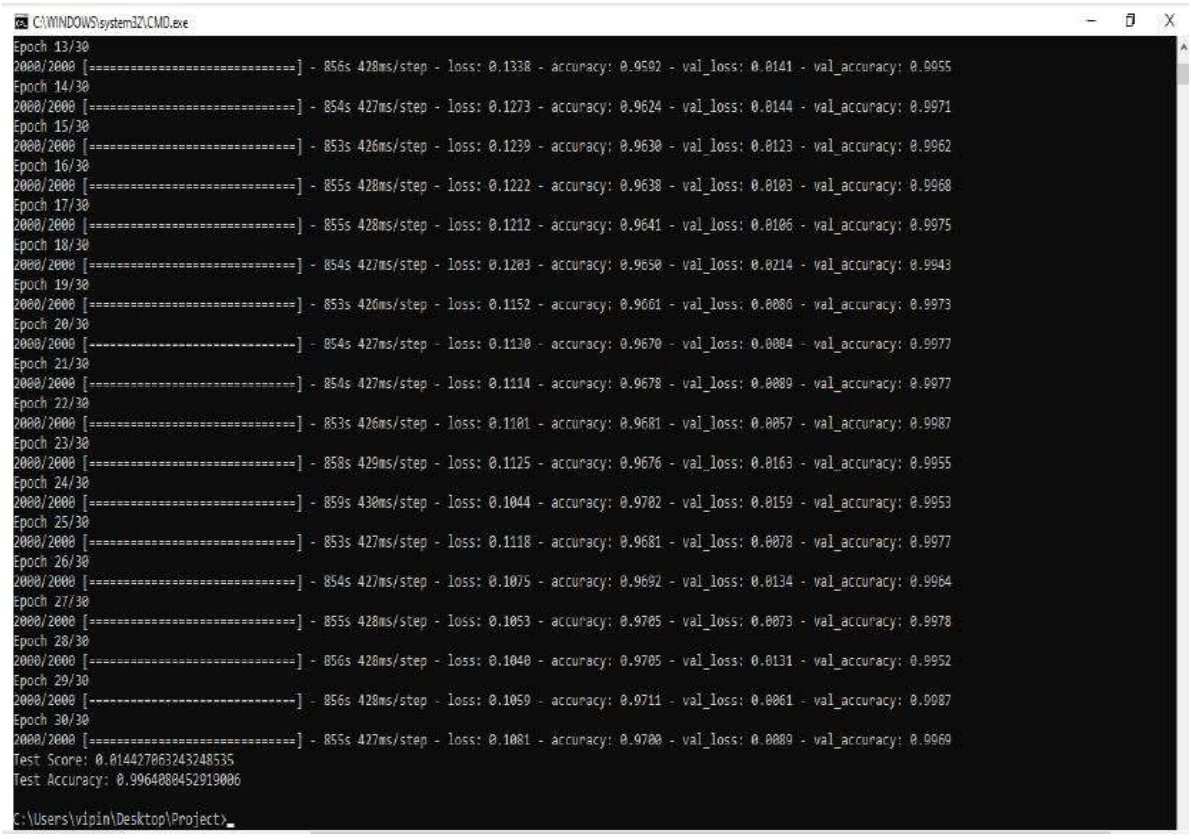

Figure 7. After Final Compilation

In the final output we can clearly see that the accuracy is 99.64 which is better than where we started. The test score is 0.0144.

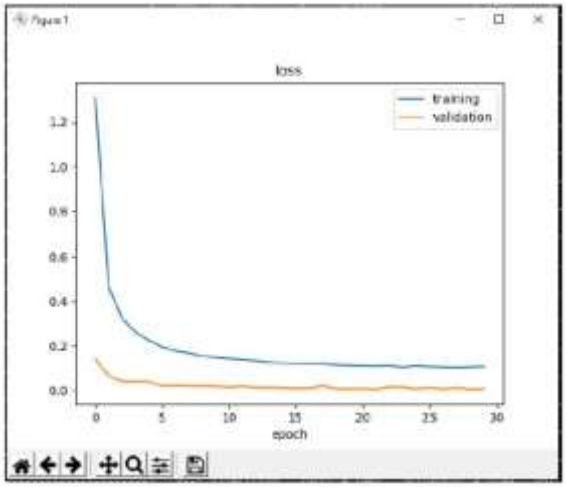

Figure 8. Graph of loss vs epochs

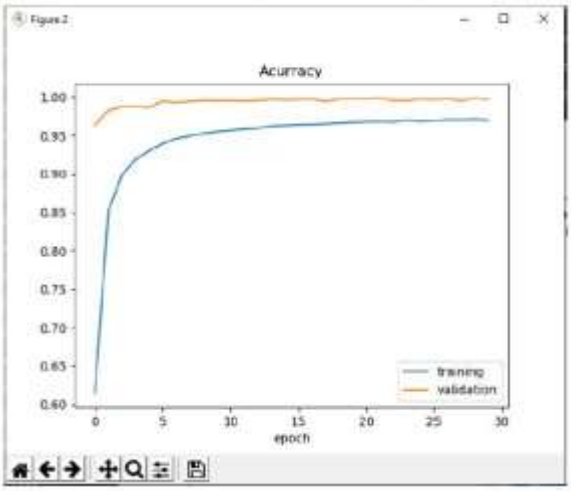

Figure 9. Graph of accuracy vs epochs

The graphs clearly depict a decrease in loss with increase in numbers of epochs. Also, the model's accuracy increases with increase in epoch numbers. Since the model is trained with data set improvement, the accuracy is increased and the loss is reduced with every epoch.

With the help of a variety of test cases, we have observed the versatility of an FRI Access Control System. We have also observed their dependence on external factors as well as efficiency of algorithms used for programming them. These tests have also been able show the superiority of such Facial Recognition Systems over other primitive access control systems, which were stated at the beginning.

V. Conclusion & Summary

Through this, we have successfully reviewed the current "state-of-art" face recognition algorithms and systems as well as different software architectures and main subsystems required for implementation in a fully functional System. Here, we also established specific embedded systems with multiple modules programmable through computer language of our selection, i.e. Python.

Also, we successfully analysed the implementation of Principle Component Analysis for its speed and storage efficiency. Apart from PCA, analysis of Artificial Neural Networks algorithm made accuracy of face recognition and realisation much more accurate as compared to that of PCA, as it occasionally fails when facial orientation changes in front of the camera.

The proposed system is tested for positive tests, i.e. cases where access is granted as well as deliberately performed negative tests for denying access. Uncontrolled and controlled conditions have also been tested for with different levels of ambient lighting, stability, etc. The system is also observed to increase visitor data and make it feature rich in cases of variation from original facial data.

This reflects that the project gives a definite proposed set of functionalities that work best for an efficient Facial Recognition IOT Access Control System.

VI. Acknowledgement

We express our deep sense of indebtedness to **Mr. Mahesh Kumar Gupta** and also to **Mr. Shahnawaz Alam**, Assistant Professors, Department of Mechanical Engineering, SRM Institute of Science and Technology, Delhi-NCR Campus, Modinagar for their excellence, guidance and encouragement towards the completion of this paper.

REFERENCES

[1] P. J. Phillips, H. Moon, S. Rizvi, and P. J. Rauss. The FERET evaluation methodology for face-recognition algorithms. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(10):1090–1104, 2000.

[2] P. J. Phillips, H. Wechsler, and P. Rauss. The FERET database and evaluation procedure for face-recognition algorithms. Image and Vision Computing, 16(5):295–306, 1998.

[3] D. Blackburn, M. Bone, and P. J. Phillips. Facial recognition vendor test 2000: evaluation report, 2000.

[4] D. Bolme, R. Beveridge, M. Teixeira, and B. Draper. The CSU face identification evaluation system: its purpose, features and structure. In International Conference on Vision Systems, pages 304–311, 2003.

[5] A. Georghiades, D. Kriegman, and P. Belhumeur. From few to many: generative models for recognition under variable pose and illumination. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(6):643–660, 2001

[6] M. Turk and A. Pentland. Face recognition using eigenfaces. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1991

[7] P. Hallinan. A Deformable Model for Face Recognition under Arbitrary Lighting Conditions. PhD thesis, Harvard University, 1995.

[8] P. Hallinan, G. Gordon, A. Yuille, P. Giblin, and D. Mumford. Two- and Three-dimensional Patterns of the face. A.K. Peters, Wellesley, MA, 1999

[9] A. R. Martinez and R. Benavente. The AR face database. Technical Report 24, Computer Vision Center (CVC) Technical Report, Barcelona, 1998

[10] M.-H. Yang, D. Kriegman, and N. Ahuja. Detecting faces in images: a survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(1):34–58, 2002.

[11] H. Rowley, S. Baluja, and T. Kanade. Neural network-based face detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(1):23–38, 1998.

[12] K.-K. Sung and T. Poggio. Example-based learning for view-based human face detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(1):39–51, 1999.

[13] H. Schneiderman and T. Kanade. A statistical method for 3D object detection applied to faces and cars. In Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, pages 746–751, 2000.

[14] Canming Ma; Taizhe Tan; Qunsheng Yang; "Cascade boosting LBP feature based classifiers for face recognition," Intelligent System and Knowledge Engineering, 2008. ISKE 2008. 3rd International Conference on, vol.1, no., pp.1100-1104, 17-19 Nov. 2008 doi: 10.1109/ISKE.2008.4731094

[15] Viola, P.; Jones, M.; "Rapid object recognition using a boosted cascade of simple features," Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, vol.1, no., pp. I-511- I-518 vol.1, 2001 doi: 10.1109/CVPR.2001.990517

[16] Froba, B.; Ernst, A.; "Face recognition with the modified census transform," Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on, vol., no., pp. 91- 96, 17-19 May 2004 doi: 10.1109/AFGR.2004.1301514