Group 2: SMART Visitor Recognition

Manjyot Singh Nanra Sharanya Saha Shiv Kumar Yadav Utkarsh Srivastava
21111038, 21111056, 21111057, 21111063
{manjyots21, sharanya21, shivky21, utkarshs21}@iitk.ac.in
Indian Institute of Technology Kanpur (IIT Kanpur)

Abstract

The Internet of Things (IoT) is a concept that describes how real-world devices (things) are connected and monitored remotely over the Internet. When this concept is applied in conjunction with AI techniques, the number of use cases increase exponentially. The application of this concept in our home can make it smarter, safer, and more automated. This project focuses on creating a smart visitor recognition system that provides notifications to the owner through the Internet whenever a person comes at the door. When a person comes at the door, camera sensor captures his/her image and sends it to the ML face recognition module. It recognises the face of the user with any of the allowed members and matches the entry timing. Furthermore, we have also provided a time-based heuristic which allows entry only under desired timings. The advantage gained by choosing this system over similar current systems is that it provides automatic entry in case the person is recognised and is trying to enter within desired timings. In all other cases, the user receives a notification to allow/deny. The developed system provides accurate judgements under different test scenarios. The time bound entry heuristic makes it unique when compared with similar systems developed so far.

1 Introduction

The concept of IoT has entered every sphere of human life. With the advancements in the field of Internet Of Things, the applications are increasing exponentially. One such application is the security of our homes and offices. The security of our home and office spaces is vital in the present day scenario. Traditionally, security was dependent on locks and keys. In such cases, the theft and loss of keys was a major issue. Now, we have shifted to a security system which depends on technology. In this project, we have developed a visitor recognition system that can help to provide security from the touch of a smartphone device. Here, facial recognition is the key to the door. Facial recognition combined with IoT will be much more secure than any other biometric systems [1].

The visitor recognition system consists of a laptop camera acting as an IoT camera sensor and a face recognition algorithm deployed on the local machine. Whenever a person comes in front of the camera sensor, it captures an image of the person. Thereafter, the ML model executes a facial recognition on the captured image. After the algorithm finishes, the following 3 cases can occur:

- **Case 1:** The person is identified successfully and the time of entry is within allowed timings.
- **Case 2:** The person is identified successfully and the time of entry is outside allowed timings.
- **Case 3:** The person is not identified i.e. an unknown person has arrived.

In case 1, entry is granted to the person. However, in cases 2 and 3, entry is put on hold at the moment and a notification is sent to the owner on Telegram. The entry can be allowed or rejected by the owner at real time using the Telegram App on smartphone. Hence, no physical interaction required and security of home becomes intact.

Now, we will briefly discuss the structure of this report in the coming sections. Section 2, i.e. Related Work tells about the different methodologies developed to automate visitor management and entry system. Numerous strategies and facial algorithms that have been used in those systems are discussed in brief. In section 3, we have covered the working strategy of our system. The different

cases and the behaviour of the system under different conditions are covered. We also discuss the facial recognition algorithms that are used and the advantages over conventional methods by using a particular strategy. Section 4 covers the actual implementation of our system namely the dataset used, the pre-processing techniques employed, the ML algorithms deployed and the time bound heuristic used in the system. We also present a comparative analysis among different facial recognition algorithms we encountered during the course of development of this project. In section 5, the outcome of the project is discussed which consists of snapshots and a brief discussion. Furthermore, in section 6, we discuss about the potential future improvements that can be made in this system to make it more robust and dynamic in nature. Towards the end of the report, section 7 gives the conclusion in which we tell the objectives that have been met after the completion of this project. In the end, we mention the individual contributions of the members during the course of development of this project followed by references in the end.

2 Related Work

Several works were reported in the literature based on home security using Internet of Things. In this section, a brief summary of the works are presented. The authors in [2] have used Raspberry Pi as the central processing module. Further, they talk about some image augmentation techniques to enhance their dataset coverage. For facial recognition, they have used Deep Learning CNN and on the front end, they have used Blynk IoT App for handling notifications. The authors in [1] have developed a novel system using Haar Cascade Face Classifier. The major drawback is the lack of facility for adding new faces to the model during runtime. In an ideal home scenario, such situations can arise when we have guests visiting or during festivals. The authors in [3, 4] dicuss about facial recognition techniques in depth. The techniques explored include ANN (Artificial Neural Network), Eigen Faces, SVM, CNN. The model in [5] was built only upon 4 output classes and failed generalize to cases when the number of students need to increase during run time. Whereas during prediction in [3], faces in the image are misidentified, and most of them have interference such as angle, illumination, and occlusion. And it offers no info on how to adapt to unknown people. However, we were motivated to use idea of face encodings and SVM for classification from [4]. The authors in [6, 7, 5] talk about developing a visitor system for office spaces and home. Particularly, [6] uses user authentication instead of facial recognition for granting access. And it does not provide decision making capabilities to the system. For use in office spaces, authors in [7] follow a different approach for entry. They prioritise entry of different persons based on situation. Our system can easily be extended to priority based entry when deployed to office spaces. In the end, [8] developed an interrupt driven system which sends email notification. However, the authors failed to provide functionality of adding new members on the go. After comparing the functionalities and drawback, we analysed that our visitor recognition system provides the following advantages over the systems discussed above:

1. Rule based entry to the users.
2. Retraining the facial recognition model in real time.
3. Facility of adding new members from smartphone in real time.
4. Automatic decision making capabilities to the system based on user's input.

3 Proposed Idea

The project focuses on automating the visitor recognition and corresponding door opening. The project minimises the interaction with the owner and smartly allows the people inside the house based on certain heuristics. The idea is based on notion of people belonging to different classes of relation with respect to the owner. These include people who are part of owners family, some people who are part of owner's daily routine but not in blood relation to the owner and some people who randomly visit at times and need to be brought to the owner's attention when they visit. There are two heuristic rules, one that allows the people belonging to family to enter the house at any point of time and second that allows only time bound entry to the people that are not part of family but have almost daily visits.

All the people visiting the vicinity of door will be captured by the camera and a face recognition algorithm will start processing it to identify the person. In case of a recognition the corresponding heuristic based on person's role will be applied and decision on allowing the person or not will be taken automatically. The notification of such visit will be sent to the user to inform about the opening of door. In case the person is not identified then the camera will capture another photo and notify the user about the presence of someone at the door. Based on owner's choice whether to allow or deny the entry the system will ask the owner whether it should save this person's face for future visits. If the owner opts for saving, then the system will retrain the machine learning model to accommodate the prediction of new user as well. The system allows the user to set the new person's name, role and expected time of visit in case of one time visitors. The working of the system is shown in Figure 1.

Figure 1: A flow chart of the proposed system

4 Methodology

The built smart visitor recognition system works completely online and performs real-time face recognition. Depending upon the recognised entity and its corresponding heuristics it allows or rejects the entry. We have analysed a few ML models like KNN and SVM and have finally used KNN for real time face recognition. The smart visitor system sends a notification to the user about the person recognised via Telegram. The system also holds the capability to add a new person to the data base and retrain the model in real-time.

4.1 Data set used

Face detection and recognition being a major portion of our project, we have explored a few data sets like:

1. The Extended Yale face database B
2. Yale Face data set
3. A data set by Robotics Laboratory of National Cheng Kung University

The data set was required to determine the ML model which would work the best for us and their corresponding evaluation scores. The **Extended Yale face database B** was selected for analysis of the ML algorithms, mostly because of the format in which the data was available and its characteristics matched the best with our requirements.
The characteristics of the data set are as follows :

1. Total number of images: 16128
2. Number of human subjects: 28
3. Each human subject has a total of 576 images in 9 poses under 64 illumination conditions.

However, as the analysis was entirely done in our personal laptops and processing such a huge data set was a challenge, so we have only used a part of the data set for our analysis. We have randomly selected 10 subjects each with 100 images from the data set and have used them for training and testing of our models.

4.2 Data pre-processing

As mentioned in the above section, we have used a part of The Yale face Extended B data set for analysis of the ML model that works the best for us. We performed sheering and rotation in the images captured by our webcam while adding a new person, to increase the robustness of the model Figure 2 shows the image augmentation steps that we have used on our own dataset.

(a) Original Image (b) Rotated Image (c) Sheered Image

Figure 2: Image Pre-processing and Augmentation

The final system implemented uses a database built in real time using the web-cams of our laptops. The images captured by the web-cam on detection of a face and are further processed. The face encoding of such captured images are used for training the ML model.

4.2.1 Face Recognition

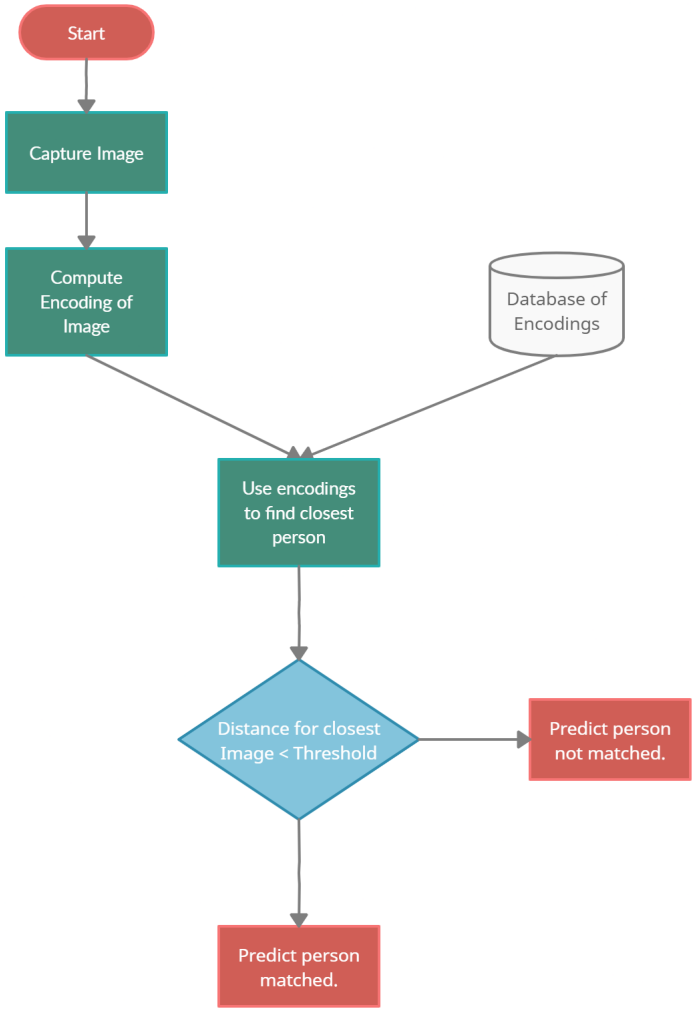


Figure 3: Original Image

For face recognition step, first we need compute the encoding of the face. This process consists of computing **128 dimensional vector** which is essentially extracting features of face. We will also have the database of stored encodings of already registered users. These encodings will be compared and we will use classification algorithm to classify the encoding to the closest user the encoding matches with. We will also get distance to the closest image. If this distance is less than or equal to the **threshold value of 0.5 units**, we will conclude that our image matches with the closes person our model found. Else we will predict that person is not found.

4.3 Analysis of different ML models used for face recognition

We have implemented and analysed two different ML algorithms for performing face recognition. The algorithms were trained and tested on the customized data set created by us from The Yale Extended Data set B.

4.3.1 Face Encoding

Step one of the face recognition system is to get encoding of the face. Encoding algorithm accepts face image as input and performs feature extraction to get a representation array 128-d long used to

represent the face.

Inception CNN Model We have done analysis on CNN Inception model. Inception-ResNet-v2 is a convolutional neural architecture that builds on the Inception family of architectures but incorporates residual connections. This model helps in generating face encodings. Face encoding computation using inception model was taking around 0.2 secs which was significantly high.

ResNet Model Face-recognition module in python uses state of the art dlib model ResNet for computing encodings of an image. A residual neural network (ResNet) is an artificial neural network (ANN) of a kind that builds on constructs known from pyramidal cells in the cerebral cortex. Residual neural networks do this by utilizing skip connections, or shortcuts to jump over some layers. Typical ResNet models are implemented with double- or triple- layer skips that contain nonlinearities (ReLU) and batch normalization in between.

4.3.2 Encoding Classification

After obtaining the encoding, for image, we will also have stored encodings in the database. To perform encoding classification to get predictions on which person does the encodign belong to, we explore following methods.

K-Nearest Neighbours K-Nearest Neighbours calculates the euclidean distance of the test face encoding from the known face encodings. If the calculated distance is less than a threshold value then it considers it as a match and predicts the name accordingly. If no matches are found then the algorithm predicts the person as unknown. The already implemented knn model available with sklearn library was used for training and testing of out dataset.

Support Vector Machines Support Vector Machine (SVM) is a supervised machine learning model used for classification problems. After giving a SVM model a set of labeled training data for each category,the model is able to categorize new test data. The model also gives the probability with which the category is predicted. We have used a probability threshold of 0.85, i.e. if the model predicts a catergory with a probability greater than 0.85 then we accept the prediction otherwise it is predicted as Unknown. We have trained and tested the SVM model using our customised dataset. The already implemented SVM model available with sklearn library was used for training and testing of out dataset.

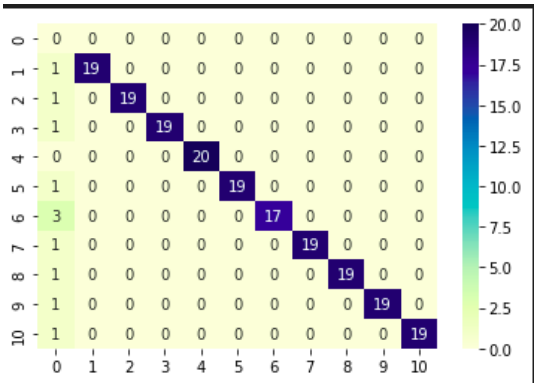
4.3.3 Comparison of the Models

K-NN and SVM, both the algorithms performed well with our training and testing data-sets.We have compared both the models on basis of their accuracy scores, time required for training and testing on the same data sets. Our proposed system being completely robust and dynamic, demanded an ML model with less training and prediction time along with good accuracy scores.

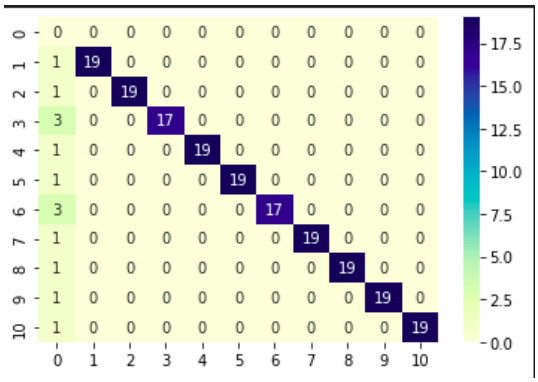
Ml Model	**Accuracy**	**Training Time (For 9 subjects) approx.**	**Prediction Time (For 9 subjects) approx.**
KNN	94.5%	2 minutes 30 seconds	39.4 seconds
SVM	93%	2 minutes 32 seconds	35.4 seconds

Table 1: Comparison of the Models

Note: The training and prediction times shown in Table 1 are local system dependent.



(a) KNN Model



(b) SVM Model

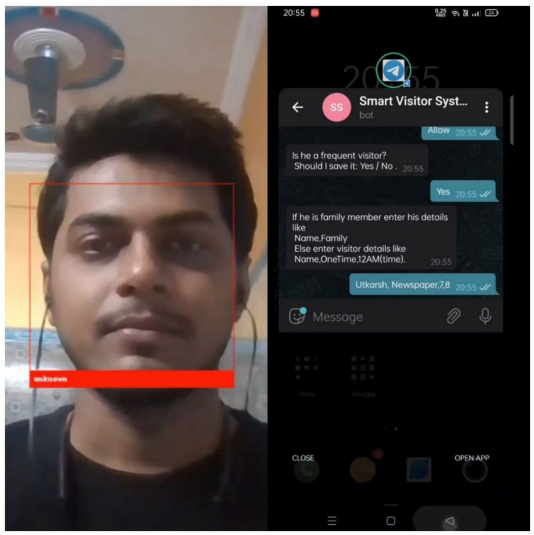
Figure 4: Confusion Matrix for the ML models

K-NN gives better accuracy scores in less training time. The prediction time for SVM model is however lesser for the same data set. Considering all the above results we chose to deploy K-NN model for real time face-recognition in our smart visitor management system.

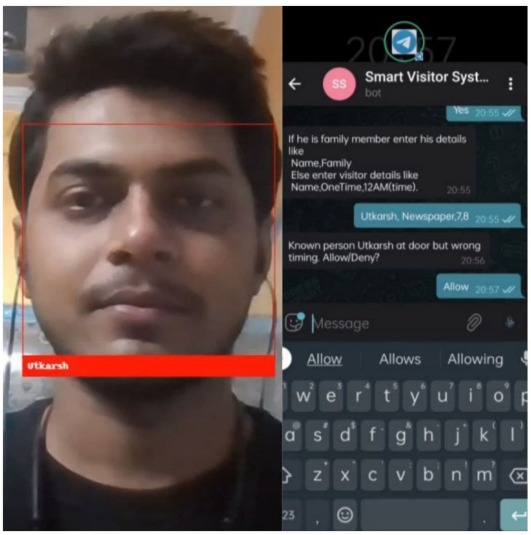
5 Results

We have deployed a smart visitor recognition system in the home scenario. It covers the following cases of entry as shown in the following figures.

- Case 1: Unknown person addition to the system in Figure 5a.
- Case 2: Unknown person arriving at the wrong time in Figure 5b.
- Case 3: A family member arriving at any time in Figure 6a and Figure 6b.
- Case 4: Adding a new member in Figure 6c.

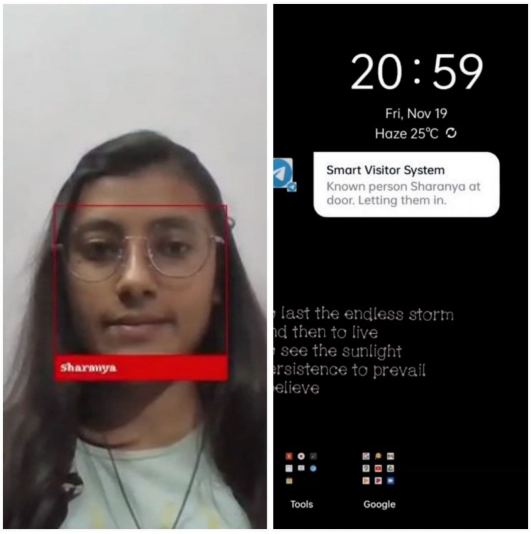


(a) Case 1: Unknown person entry.

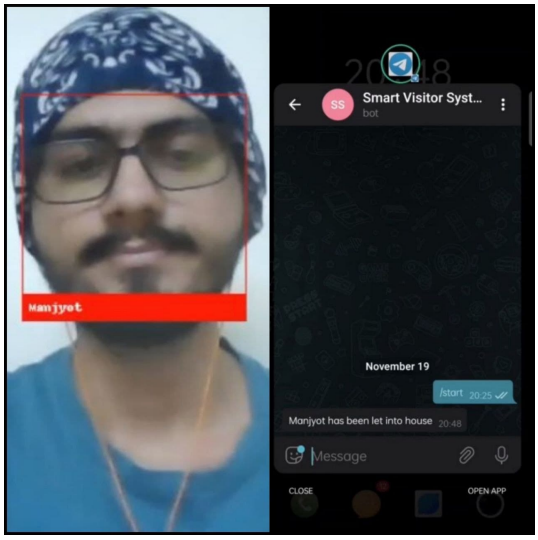


(b) Case 2: Known person at wrong time.

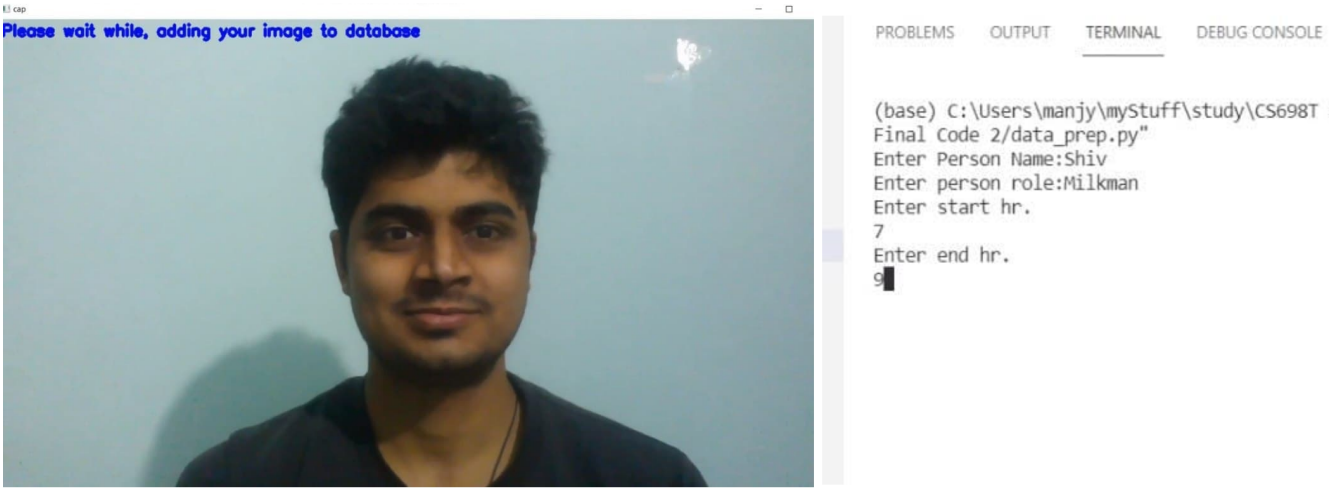
Figure 5: Entry cases 1 and 2



(a) Case 3: Family member entry.



(b) Case 3: Family member entry.



(c) Case 4: Adding new member.

Figure 6: Entry cases 3 and 4

6 Discussion and Future Work

The primary aim to develop this project was to provide security to our homes. However, the number of use cases can be multiple if the system is slightly modified. If we add the facility of adding and removing members on the go, it can be deployed to Banks where only authorised personnel can be allowed to enter vaults. In office spaces, the same model can be configured to give access to meeting rooms. This will avoid unnecessary interference during important meetings and conferences. If the facial recognition model is replaced with a car License Plate identification model, it can be deployed in Toll Plazas.

7 Conclusion

The aim of the project was to design and develop a visitor recognition system based on face recognition using Machine Learning and apply a time based heuristic to facilitate automatic decision making capabilities to the system. During the course of the work, we have been able to successfully complete the stated work. Moreover, we added a couple of unique features to make the project more ideal for a real world use case. The highlights of the system are listed as

- Real time addition of new people to the database.
- Rule based Entry.
- Decision making Ability.
- Smartphone based control system.

8 Individual Contributions

Work	Manjyot	Sharanya	Shiv	Utkarsh
Total Contribution	**29.17%**	**29.17%**	**12.5%**	**29.17%**
Literature Survey	✓	✓	✓	✓
Preprocessing of Image Integration	✓	✓	-	✓
ML Model RnD	✓	✓	-	✓
ML Model Implementation	✓	✓	-	-
Telegram Notification Send Message and Image	-	-	✓	✓
Telegram Notification receive reply	-	-	✓	✓
Integration	✓	✓	-	-
Analysis of various models	✓	✓	-	-
Report	✓	✓	✓	✓
Final Presentation	✓	✓	✓	✓

References

[1] S. Ghafoor, K. B. Khan, M. R. Tahir, and M. Mustafa, "Home automation security system based on face detection and recognition using iot," *Communications in Computer and Information Science*, p. 67–78, 2020.

[2] S. A. Radzi, M. M. Alif, Y. N. Athirah, A. S. Jaafar, A. H. Norihan, and M. S. Saleha, "Iot based facial recognition door access control home security system using raspberry pi," *International Journal of Power Electronics and Drive Systems (IJPEDS)*, vol. 11, no. 1, p. 417, 2020.

[3] Y. Zhou, H. Ni, F. Ren, and X. Kang, "Face and gender recognition system based on convolutional neural networks," in *2019 IEEE International Conference on Mechatronics and Automation (ICMA)*, pp. 1091–1095, 2019.

[4] M. Lal, K. Kumar, R. Hussain, A. Maitlo, S. Ali, and H. Shaikh, "Study of face recognition techniques: A survey," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 6, 2018.

[5] S. Kakarla, P. Gangula, M. Rahul, C. S. Singh, and T. H. Sarma, "Smart attendance management system based on face recognition using cnn," *2020 IEEE-HYDCON*, 2020.

[6] M. Oktaviandri and K. K. Foong, "Design and development of visitor management system," 2019.

[7] S. Rao, S. A, and K. Kulkarni, "Smart phone based cost effective visitor management system for smart offices," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 12, no. 6, pp. 112–123, 2018.

[8] R. K. Gupta, S. Balamurugan, K. Aroul, and R. Marimuthu, "Iot based door entry system," *Indian Journal of Science and Technology*, vol. 9, no. 37, 2016.