IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.047 ⁜ Vol. 10, Issue 5, May 2022

DOI: 10.17148/IJIREEICE.2022.10558

REAL- TIME CLOUD COMPUTING BASED FACE AND SPEECH RECOGNITION FOR ACCESS CONTROL DEVICES

Manjula BB[1], Guru kiran G[2], MridulPandey[3], Maruthi[4] R , Sandesh[5]

Associate Professor, Dept. of ECE, East West Institute of Technology, Bangalore, India[1]

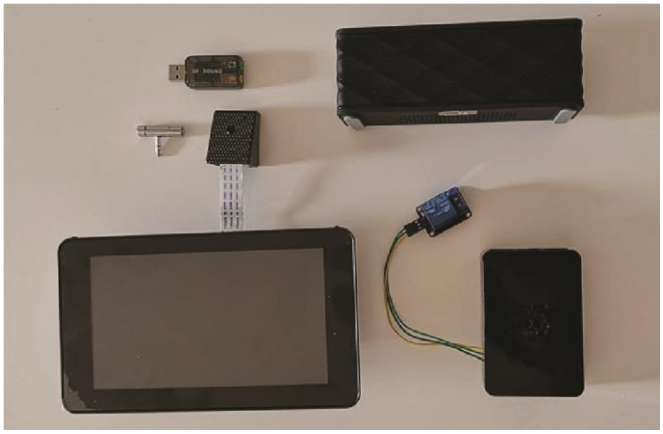
B.E Student, Dept. of ECE, East West Institute of Technology, Bangalore, India [2,3,4,5]

Abstract: The system helps employees to enter the office by unlocking the door without the aid of their id cards. It also helps to prevent spoofing attacks by generating a random code to the employees which they have to type it on the screen, this increases the security. The visitors or delivery persons who wants to meet the employees are given speech recognition services where they tell the name of the employee and the device will send notification to the correct employee respectively.

Keywords: Face and speech, Spoofing attack, Key-tags, Employee.

I. INTRODUCTION

In order to enter the office, employees need id cards. Those who do not have id cards like visitors or delivery persons, or if employees forget their id cards, they must inform the watchmen and take the help of the people inside the office. Then, someone inside the office has to open the door manually, usually those at the desks located closer to the entrance. This does not only consume time and focus of the workday for these staff members but also disturbs others with desks located nearby.



The goal of this project is to develop a system that uses face recognition to manage the lock system of the doorway for workers, and speech-to-text services to notify the proper employee member that a visitor or delivery person is waiting outside. The project will help the office and its staff in two ways.

II. PROBLEM STATEMENT

Disturbances are caused when an individual inside the office must interrupt the work to open the door for an employee, a guest, or to receive delivery. Biometric systems may be attacked via spoofing attacks. The device works well in an office environment, but if it's connected to a slower internet connection, the face recognition and speech-to-text will take longer.

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.047 ⁜ Vol. 10, Issue 5, May 2022

DOI: 10.17148/IJIREEICE.2022.10558

The system is cloud based, which implies that it always requires internet connection to control.

III. METHODOLOGY

The project is split into two phases, the experimentation phase and therefore the implementation phase.

EXPERIMENTATION PHASE: Aims to search out the foremost suitable setup for this project. It lays the inspiration of how the system are going to be implemented.

IMPLEMENTATION PHASE: Seeks to explain how the software and hardware are integrated to make the intended system. It makes use of two Raspberry Pi with a camera, a microphone, and a speaker. Face recognition and speech-to-text conversion are finished with the cloud-based solutions provided by Amazon Web Services and Google Speech - to-Text respectively. The system is complemented with slightly screen display, and a graphical program (GUI) which presents the detected classes to the user. The worker will receive a personal message on Slack with the notification, so others at the office won't be disturbed.

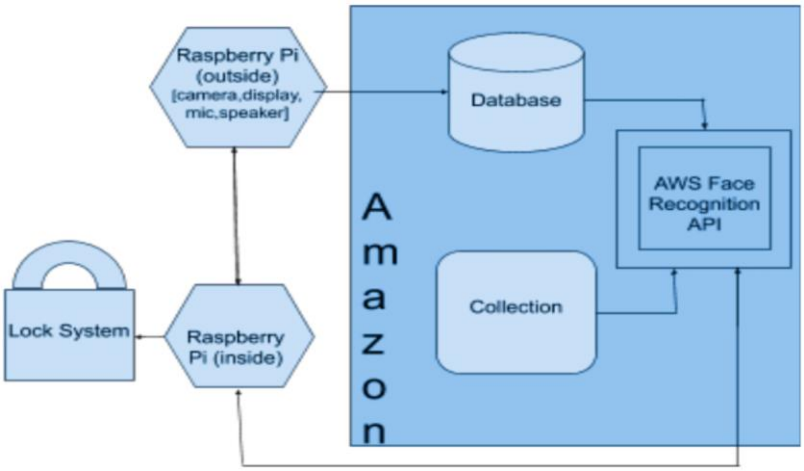


Fig. 1 BLOCK DIAGRAM

Python v3 is employed is because the artificial language. We will consider Java, but Python offered more APIs and libraries. Python offers a simplicity when it involves libraries and APIs for calculations and recognitions, plus it's a well-known artificial language used for image and data analysis

For face recognition Amazon Recognition and Amazon Simple Storage Service (Amazon S3) is employed, both from Amazon Web Services (AWS). The system is split in three functions for *1)* employees, *2*) guests and *3*) deliveries. The worker function has two authentication steps, the face recognition and a random generated code that must be confirmed to guard against spoofing. The guest function includes the speech-to-text service to state an employee's name that the guest wants to fulfill, and therefore the employee is then notified. The delivery function informs the particular person within the office that are to blame for the deliveries by sending a notification.

If it's an employee, the system goes to ask the person visually through the GUI to square before of the camera to require an image. When the image has been captured, it's sent to the S3 bucket and deleted after the comparison. The image is encrypted on the client-side before it's uploaded to the cloud. The image is decrypted before the comparison, so compared against the collections using the Amazon Face Recognition API. To enhance protection against spoofing attacks, a two-factor authentication is implemented, by combination of face recognition and a code. When the comparison is finished, it returns the similarity score. If the similarity score is above a predefined threshold, the system generates a random 4-digit code and sends it to the worker as a personal message on Slack. This prevents unauthorized persons to enter the office.

If the guest and delivery person option is chosen, the person is asked visually through the GUI to pronounce the name of the worker that wishes to fulfill. The system records the input with the microphone and sends it to Google's Speech-to-Text service. Since it is difficult to use speech-to-text on names, a string comparison algorithm is employed to match the input with all names within the database. If the foremost similar string includes a similarity score over 80, the system sends a non- public message to the worker directly on Slack with the data that a guest is waiting outside the doorway

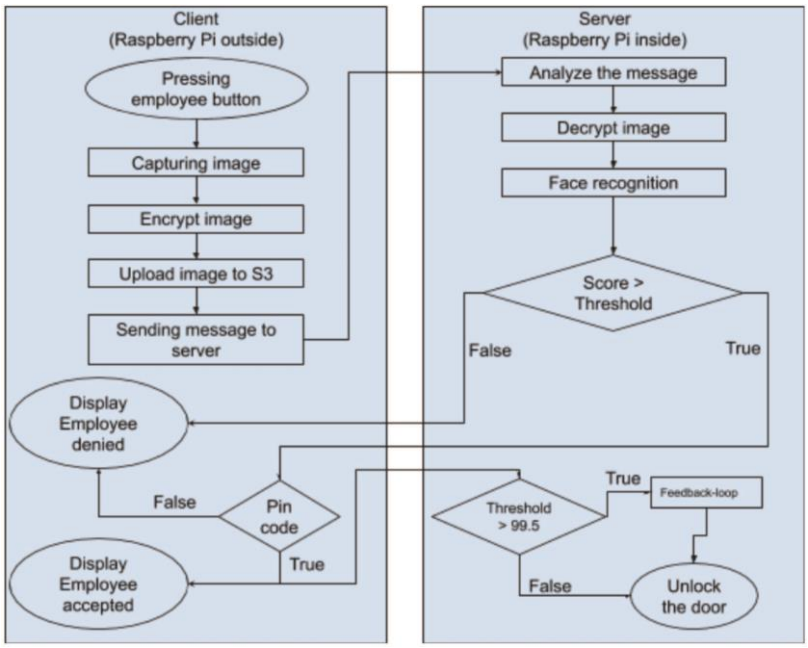
IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

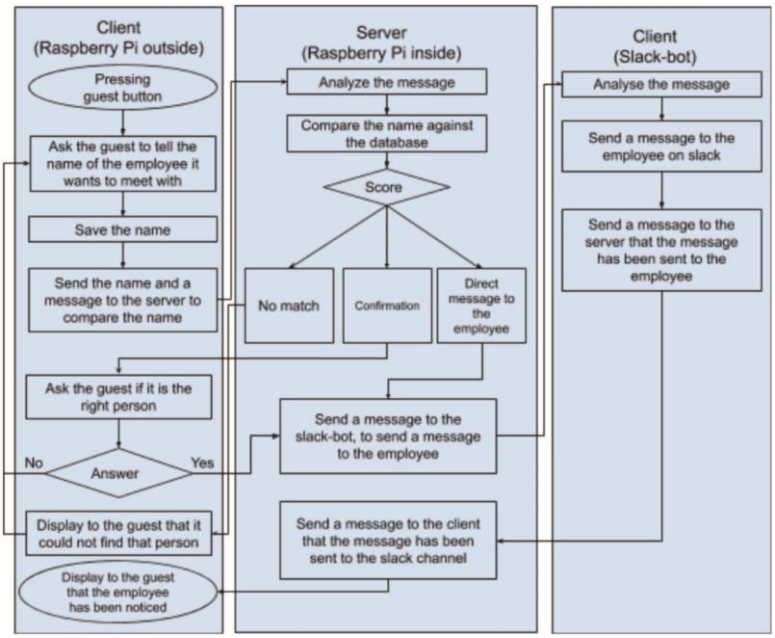
International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.047 ⁜ Vol. 10, Issue 5, May 2022

DOI: 10.17148/IJIREEICE.2022.10558

door. The guest also gets information on the GUI that the worker has been notified. If the score is below 30, it then means that the similarity score is just too low and also the system will ask the user to undertake again. If the score is between these values, the guest must confirm that it is the right individual that they need to satisfy, with help of "Yes" and "No" buttons which can be displayed on the GUI. If the system returns incorrect employees name and also the guest presses "No", the guest will be asked again who wishes to fulfill.





IV. RESULTS

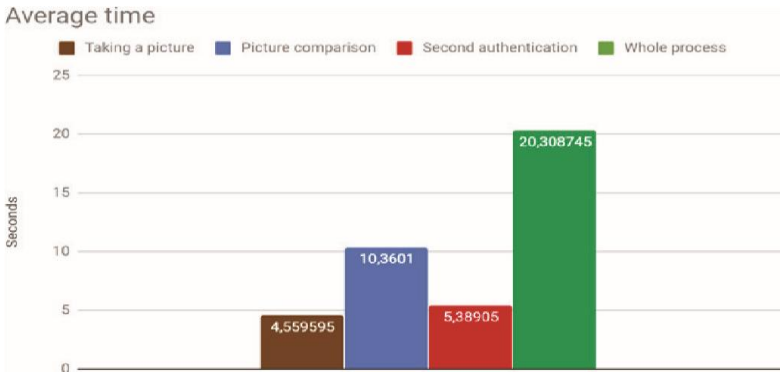
Amazon Recognition provided perfect separation between genuine and impostors. The smallest genuine score is 94.25%, while the very best impostor score is 73.1%. The system will be therefore considered reliable to permit employees to access the office, and to deny access to unauthorized people, a minimum of size evaluated here and here within the imaging environment where the corporate operates.

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

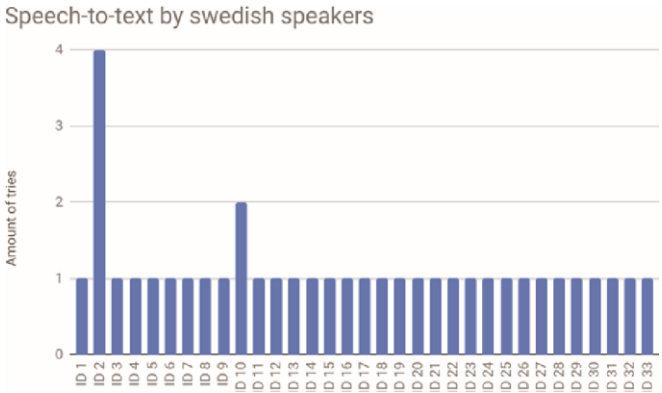
International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.047 ⁜ Vol. 10, Issue 5, May 2022

DOI: 10.17148/IJIREEICE.2022.10558



Each employee name has been spoken by the testers to work out what percentage attempts was needed to match with the proper employee. The results are shown in Figure. speakers managed to match the right employee within the majority of cases (average amount of tries equal to 37/33= 1.12). On the opposite hand,speakers had to do over once with approximately 27% of the employees (9 out of 33 IDs), resulting in a mean amount of tries adequate to 50/33=1.51.



V. CONCLUSION

The goal of this work is to develop a system that may control the most entrance of an office by using face recognition and speech-to-text. It makes use of two Raspberry Pi with a camera, a microphone, and a speaker. Face recognition and speech-to-text conversion are through with the cloud-based solutions provided by Amazon Web Services and Google Speech-to-Text, respectively The system is complemented with a slightly screen display, and a graphical program (GUI) which presents the detected classes to the user.

ACKNOWLEDGMENT

Authors K. H.-D. and F. A.-F. thank the Swedish Research Council for funding their research. Authors also acknowledge the CAISR program of the Swedish Knowledge Foundation.

REFERENCES

[1] A. F. Abate, M. Nappi, D. Riccio, and G. Sabatino, "2d and 3d face recognition: A survey," *Pattern Recognition Letters*, vol. 28, no. 14, pp. 1885 – 1906, 2007.

[2] S. Singh and S. Prasad, "Techniques and challenges of face recognition: A critical review," *Procedia Computer Science*, vol. 143, pp. 536 – 543, 2018.

[3] A. A. Albahdal and T. E. Boult, "Problems and promises of using the cloud and biometrics," in *11th International Conference on Information Technology: New Generations*, April 2014, pp. 293–300.

[4] M. Ayad, M. Taher, and A. Salem, "Real-time mobile cloud computing: A case study in face recognition," in *28th International Conference on Advanced Information Networking and Applications Workshops*, May 2014, pp. 73–78.

[5] D. von Sohsten and S. Murilo, "Multiple face recognition in real-time¨ using cloud computing, emgu cv and windows azure," in *13th International Conference on Intelligent Systems Design and Applications*, Dec 2013, pp. 137–140.

[6] R. Kesava, "Cloud software as a service with iris authentication," *Journal of GlobalResearch in Computer Science*, 2010.