iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VII **Month of publication:** July 2021

DOI: https://doi.org/10.22214/ijraset.2021.36408

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VII July 2021- Available at www.ijraset.com

Audio and Visual Model for Real World Interaction

Akarshak Bose[1], Shruti Jaiswal[2], Anush Mishra[3]
[1, 2, 3]*Department of Computer Science and Engineering, SRM Institute of Science and Technology Ramapuram, Chennai, India*

Abstract: Communication with the proper information can be helpful for any person to carry out conversations. The proposed system is to help people to interact freely with full information about the past conversations with the person they are meeting. The device UPAL will identify the face and voice of the person and will store necessary details about the meeting, by recording the conversations or by taking inputs from the user. Next time when the user meets the same person, the device will fetch the information from the storage that can be used for a comfortable conversation. UPAL is made up of a Camera and microphone that will use the Face recognition technique and voice recognition system to collect the data. A mobile-based application will be provided to the user for viewing, editing the stored information. UPAL will ensure smart conversation by guiding and reminding the user.

Keywords: Communication, Face Recognition, Voice recognition, UPAL, Conversation.

I. INTRODUCTION

Recent advancements in communications have created the necessary conditions for the networking of a wide variety of devices. This has led to the inserting of devices into everyday life and has enabled new forms of communication between people and objects. The concept of this system is to provide a guidance system to the user which can store data and use it in the future. Often it is found that we forget information about our past meetings with someone. We may find ourselves in an embarrassing situation at times, so it is preferable to have an app that can remind us of previous meetings and guide us on what we should say. The main function of this system is to save information about the meetings that the user attends and synthesize it for future use, as well as to guide different actions based on the information gathered. Intelligent information systems enable the processing of data collected by the sensors. The system will also remind the user about the incomplete task. The system will have a face detection device and a voice recognition device that will identify a particular person and will provide available data. It will record the meeting and identify and store certain keywords. Following the meeting, the system will ask the user basic questions about the meeting, which the user may or may not provide. Next time when they meet again, the system will provide the user with the details about the past meeting. The following project can also assist security personnel in identifying criminals. It may not be possible for the police to know about the background of a person who comes to their state. As a result, if that person commits an offence while using this device, the security personnel will be able to learn about their background by extracting data from a central database.

II. METHODOLOGY

We come across many people in our day to day life for a business purpose and other meetings are casual meetup with a friend or family. It's not always easy to keep in mind the details of the meeting. If the meetings are important, and we try to remember the details because noting them down is not always possible, it is more likely that they will forget in the next three to four days, and if they need these details in a week or a month, they will not have access to them. we can clearly understand that we can never depend on human memory. Things become more complicated for people with Alzheimer's disease because they are not able to remember their recent activities. Sometimes security personnel may come across wanted criminals, but they don't know about them because they might be from other areas or even a different state or country.

The device will be helpful in solving these types of problems where one person will be able to keep notes of important or casual meetups with real time recordings, making everything very transparent and easy, especially when it comes to helping cops. They can easily find out what they have discussed and what was the conclusion of the last meeting on their mobile phones. They can also be able to keep reminders for the next meeting/meetups and what they are going to talk about and also what they need to know from their last meeting if they have already met. They will not have to rely solely on notes; instead, they will have recordings to which they can refer at any time, ensuring transparency and clarity.

©IJRASET: All Rights are Reserved

It will require a mobile application, camera and microphone to be connected with the application when the user installs it. The application will record the face and voice of the discussion/meeting that is happening. The application will be in connection with a central database which includes information about people the user met and also criminal information which will be maintained by the department of police. Every time we ask for details of the particular user's last meetings, the application will automatically search for the user's name and if the name matches, then will provide the details to the user. Once the user install the application and connects the camera and microphone, then the device is ready to use the system.

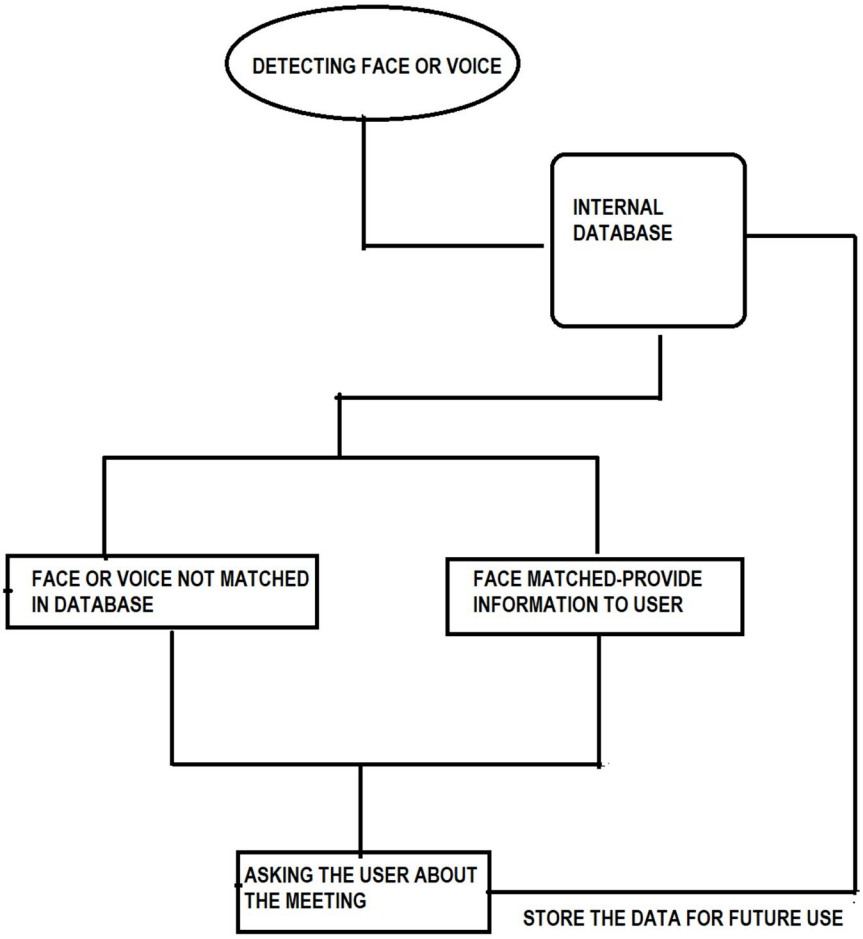


Diagram 1

Now, as soon as an individual person or group of people approaches the user, then the application will detect the face of the person and then the application will search for information available in the database or not. If the user data is not available, then the application will inform the user and will record the meeting proceedings. The application will also ask the user about the meeting after the end of the conversation. If the faces or the voice match, the application provides the user details of the past meeting and background.

With the help of face recognition, the application will search its storage, if there is no particular information found, then it will simply return/show a message on the application (on the user's phone) that "NO INFORMATION FOUND". But if it did find some information, then it will simply show it on the user's phone. Now this whole process needs to be done before the meeting starts, but if the meeting starts and the application is still searching, then recording will start side by side. During the whole process of recording, the application doesn't need to do anything except record it.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VII July 2021- Available at www.ijraset.com

Once the app detects that the person has moved/left from the sight of the camera (or when the user says "meeting ended"), the application will know that the meeting ended. Following that, the application will ask the user whether they want to keep this recording or not (yes/no). If they press "no," the recording will be deleted immediately, and if they press "yes," the application will ask a few questions about the meeting to keep a quick review and reminders if needed for future to ask for second meeting means what review the user will give here will be saved. The questions could be like:

1) How was the meeting? Sad, OK or happy. This will be helpful for judging the mood of the person from the user's perspective
2) Will this be their second meeting? Yes or no. If the user selects "no," it will proceed to the next question; if the user selects "yes," it will ask what topic needs to be discussed in the next meeting and whether the user needs to set a reminder for it, similar to an alarm clock.
3) Next it might ask what topics they discussed in today's meeting?
4) Next it would ask what was the conclusion of the meeting. Users can write all these in bullet points just like short notes.
5) At the end, it might ask whether the user needs to keep any reminder of anything they remembered from a meeting, like to buy something important or to wish a birthday to someone. The application can ask other questions too, like it will end with saying "do you need to add anything else important" and the user can write what they need. After this, the application will store the details in the database in the user's folder (with their name, the other person's name).

III. REQUIREMENTS

A. *General Requirements*

1) Must have mobile phones with a good internet connection.
2) Need to carry a phone everywhere all the time for this application/system to work.
3) UPAL

B. *UPAL*

External Device

1) Night vision lens Camera Resolution : 720P microphone
2) 2000mAh Battery
3) 32GB storage

Mobile Application

```
image = cv2.imread(imagePath)
gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
faces = upalCascade.detectMultiScale(
   gray,
   scaleFactor=1.1,
   minNeighbors=5,
   minSize=(30, 30),
   flags = cv2.cv.CV_HAAR_SCALE_IMAGE
)
```

This is the primary surce code for the face recognition part. The above code will be implemented in oder to get the functional detection of the Face.

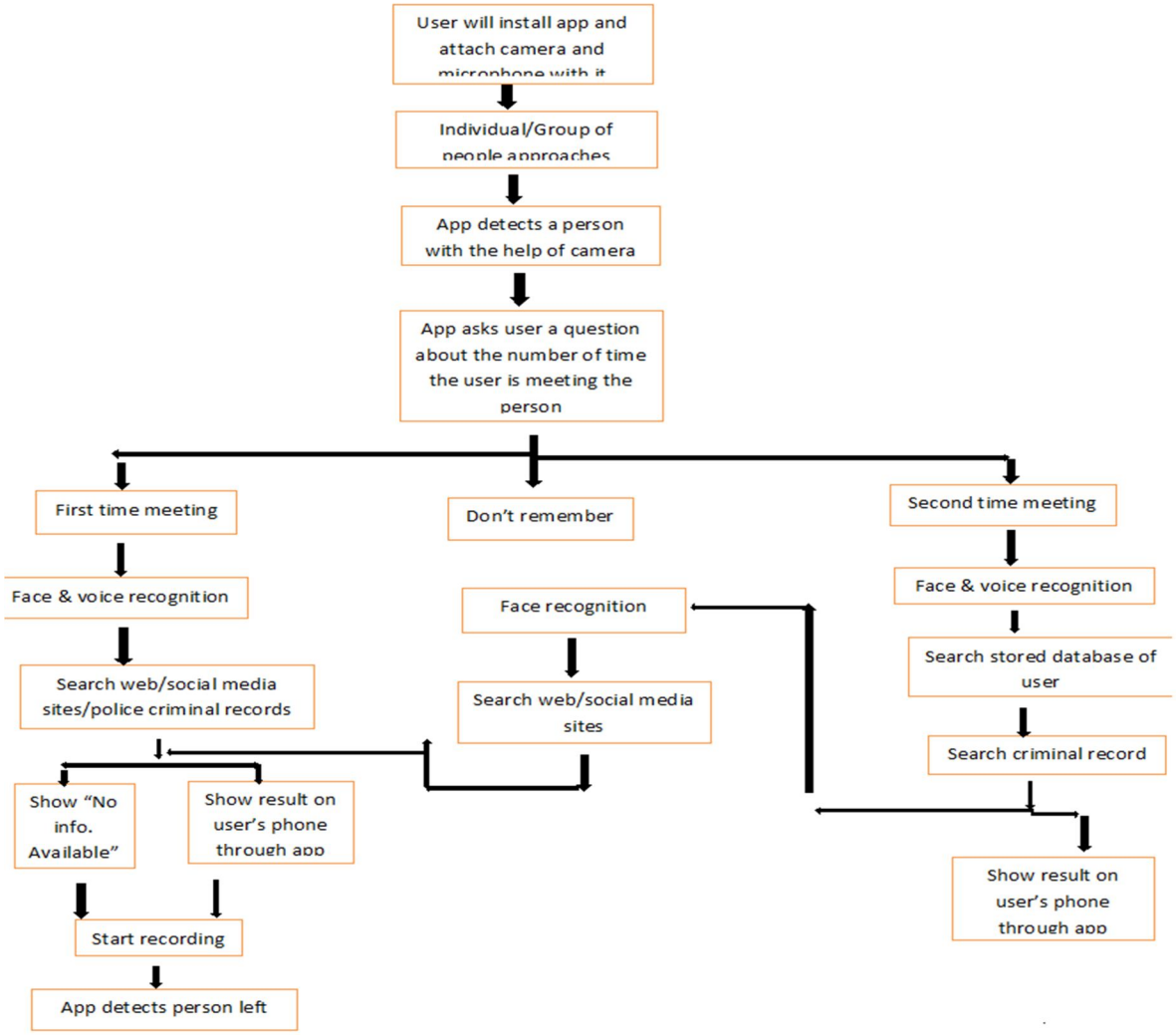
Similarly for voice recognition following code will be used.

```
p = pyaudio.PyAudio()
```

The application will be installed in the users phone and all the information will be available in text format about the past meetings. The user can access the information anytime.

©IJRASET: All Rights are Reserved

IV. ALGORITHM



V. LIMITATIONS

A. Camera and microphone can get affected on windy, dusty day. Microphone will get affected because of crowd and noisy places too (to catch the exact/particular voice).

B. Time complexity: for finding details of a person since we are searching here multiple sites.

C. Time of showing results of the person to user might clash with the starting time of the meeting.

D. If the person is a criminal but cannot speak and also gone through a plastic surgery, then, it becomes difficult to identify them even if they are around us and we are not meeting with them.

E. Needs to detect only human and not other's as it would be really annoying to tell application to not record. This could be solved by giving user the option to switch application on or off according to their needs. But it cause problem for Alzheimer patient than.

F. People do need to file any crime they encountered to police and the cops are needed to investigate the case taking it seriously and catch the criminal. Otherwise the person who has done crime will be treated as normal people and who knows what he/she capable of doing.

©IJRASET: All Rights are Reserved

VI. CONCLUSION

The proposed device- UPAL will be helpful not only for user based interacion but also for the security staffs to identify the criminals. UPAL recognises the face and identifies the voice with an accuracy of 90%. The UPAL is preferably a genuine option for remembering the keypoints during a Conversation or storing the Minutes of the meeting. We developed this device integrated with all applications of facial and voice recognition to help the peoplewho forget the things which may lead to embrassing situations .Hence, we tried to make it as friendly, convenient, and feasible as possible along with cheap so that it can be helpful for a much larger cause. Choosing materials which were easily available locally and user friendly mobile based application will be also helpful for quick congfiguration. The user have to install the Application then connecting UPAL device with it using wifi. And as it is just a prototype, real device can be ready by within few hours.

VII. FUTURE SCOPE

A. In future we can develop a system which will be all connected with a spectacles like everything that we are showing on phone now can be seen on the specs the user is wearing which will ultimately reduce much attention on phone and nobody will find about it which is kind of cool. But this type of system requires a lot of high technology the AI and it is also costly too so for now phones are good to use.

B. We can provide an option to the user that he/she can put bookmark on the recording like from where they need to see the video next time if they need it. It will just save the time by skipping all intro, greeting and random talks.

C. We can detect a person's mood while searching for their details and it will ultimately will help user to have a friendly talk means they can know what he/she should avoid saying just in case not to hurt the person.

D. This mood detection can be used to identify from their recent posts on social media site which will help analyze how much they are active on it and what mood they are currently going through, sometimes people do mention a bit about their problem in their social media accounts.

E. The background check by searching web portals can be used.

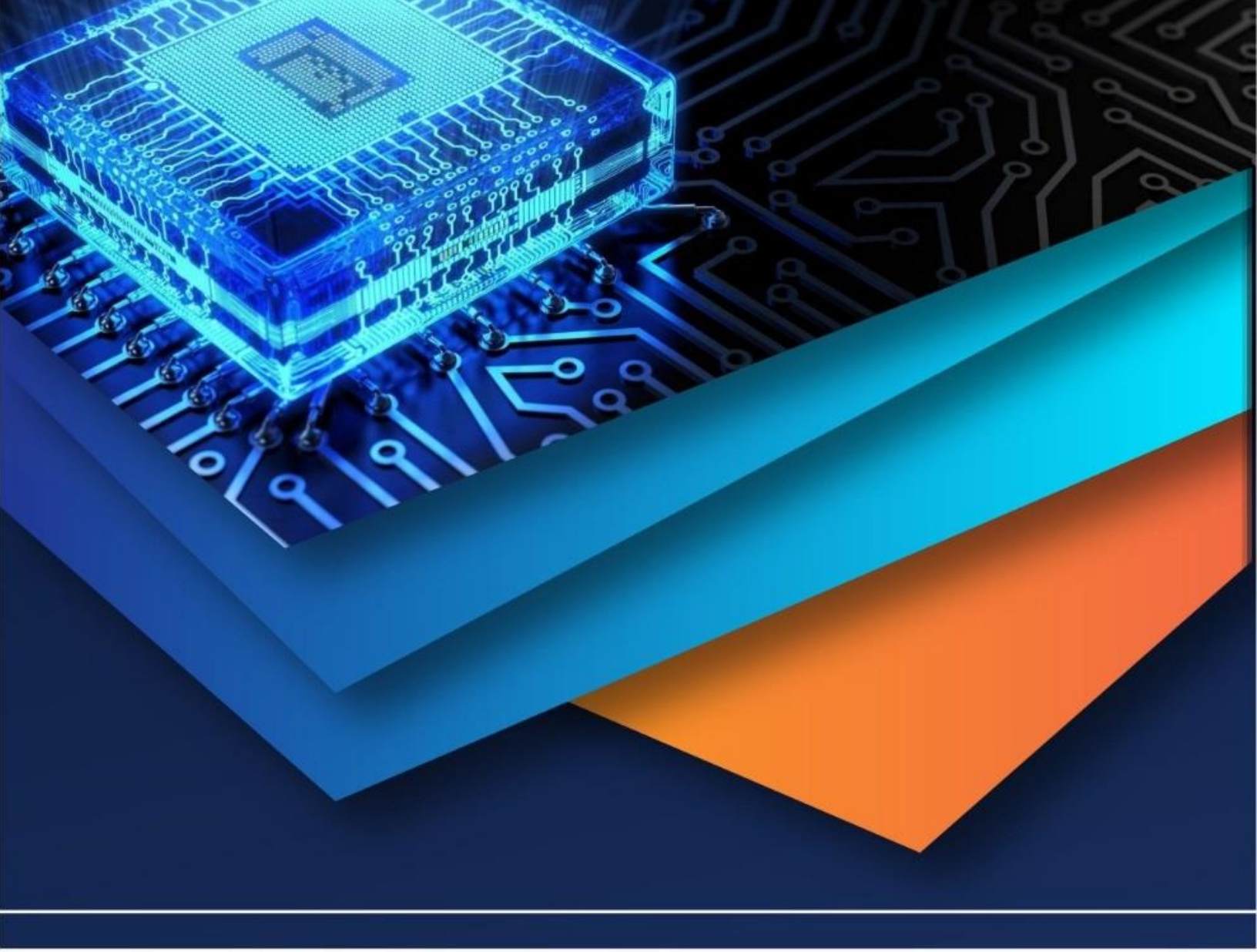
VIII. ACKNOWLEDGEMENT

The authors would like to thank Dr. V Subbiah Bharathi (Director SRM IST RAMAPURAM) for providing the necessary guidance, Abhishek Ghosh for his assistance and faculties for their cooperation.

REFERENCES

[1] S.P. Larcombe,J.M. Stern, P.A. Ivey, N.L. Seed"A low cost, intelligent micro-camera for surveillance" European Convention on Security and Detection, 1995
[2] Rowley, H., Baluja, S., and Kanade, T., "Rotation Invariant Neural Network-Based Face Detection", Computer Vision and Pattern Recognition, pp. 38-44, 1998.
[3] Tolba, Ahmad & El-Baz, Ali & El-Harby, Ahmed. (2005). Face Recognition: A Literature Review. International Journal of Signal Processing. 2. 88-103.
[4] Kortli, Yassin & Jridi, Maher & Falou, & Atri, Mohamed. (2020). Face Recognition Systems: A Survey. Sensors. 20. 342. 10.3390/s20020342.
[5] Benkerzaz, Saliha & Elmir, Youssef & Dennai, Abdeslem. (2019). A Study on Automatic Speech Recognition. 10. 77-85. 10.6025/jitr/2019/10/3/77-85.
[6] Pala, Mahesh. (2016). A New Human Voice Recognition System. AJSAT. 5. 23-30.
[7] Nazeer, Shahrin & Omar, Normah & Khalid, Marzuki. (2007). Face Recognition System using Artificial Neural Networks Approach. Proceedings of ICSCN 2007: International Conference on Signal Processing Communications and Networking. 420 - 425. 10.1109/ICSCN.2007.350774.
[8] Teimoor, Ramyar. (2018). A survey of Face recognition with machine learning. 10.13140/RG.2.2.16820.50563.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)