Grenze International Journal of Engineering and Technology, Jan Issue

Designing a Secure Interactive ATM for Visually Impaired People

Kargil Thakur[1], Aditya Tayde[2], Aditya Wanjari[3], Amit Tembhurkar[4] and Shilpa Sonawani[5]
[1-4]Department of Computer Science Engineering MITWPU, Pune, India
Email: tkargil0@gmail.com, adityatayade52@ gmail.com, wanjari.aditya@ gmail.com, amit.18.ynrs@ gmail.com
[5]Asst. professor, Department of Computer science Engineering, MITWPU, Pune, Maharashtra, India
Email: shilpa.sonawani@mitwpu.edu.in

***Abstract*—ATM(Automated Teller Machine) is among the most important tools for financial transactions. Nevertheless, for years, the visually impaired have been denied various banking facilities, especially ATM cards. The most common arguments presented by banks are 'someone might misuse your card or cheat you, you might be stalked by miscreants and robbed'. Therefore, an ATM interface should be made which allows the visually impaired to do transactions safely, independently, and efficiently. This work aims to implement an ATM interface that has all these qualities. In this interface, the user will hear audio instructions from the interface about various facilities in the ATM, implemented using text-to-speech conversion. Moreover, the ATM will also notify the breakup of dispensed cash and directions for the cash dispenser via audio instructions. The user will also be able to give voice instructions to the interface for speeding up the transactions, implemented using NLP(Natural Language Processing). For avoiding the snooping of ATM PIN(Personal Identification Number), a distance detection algorithm will be used, to ensure that there is no one else present in the ATM.**

***Index Terms*— computer vision, human computer interaction, speech recognition, speech processing, text-to-speech, visually impaired.**

I. INTRODUCTION

Globally the number of people of all ages visually impaired is estimated to be 285 million, of whom 39 million are blind. People 50 years and older are 82% of all blind. The major causes of visual impairment are uncorrected refractive errors (43%) and cataract (33%); the first cause of blindness is cataract (51%), according to WHO(World Health Organization) [1]. Looking at these numbers, there is a need for technologies to be developed to assist visually impaired people in performing day-to-day activities. These technologies are known as assistive technology. ATM(Automated Teller Machine) is a technology which enables users to do many financial transactions, like cash withdrawals, cash deposit, balance check, etc. The accessibility and availability of these ATMs make it a very essential device. However, the lack of assistive technology in ATMs discourages many blind users from using ATMs. In India, for years, the visually impaired have been denied various banking facilities, especially ATM cards. The most common arguments presented by banks are 'someone might misuse your card or cheat you, you might be stalked by miscreants and robbed. Although there are some advancements made to make ATMs more accessible to visually impaired people, they are not enough. Some of these advancements are:-

Grenze ID: 01.GIJET.8.1.14

1. The keypad for entering a PIN(personal identification number) has braille text[2].
2. The ATMs give voice instructions to the users[3]

To address all limitations of the current ATM system, we have developed an ATM interface with the following objectives- create an ATM interface that gives audio instructions, use NLP for identifying the instructions given by the user and executing them, use the video feed of atm to prevent snooping of atm pins by creating a distance detection algorithm, give details about the breakdown of cash dispensed by the ATM, give voice instructions about the location of the cash dispenser and card slot The flow of the research paper would be as follows: Existing methodology talks about various techniques already facilitating the visually impaired to do transactions in ATM and their drawbacks. The next section, Setup talks about various hardware and software requirements of the system. And it also talks about how these all components are placed and used. The next section, Design talks about the general architecture of the system, and how different users will interact in the system. The next section, our methodology talks about the different modules we have used in the system and their working methodology. The next section talks about the conclusion we draw from our work and the future scope of the system. Finally, this paper ends with listing all the references used to design this system.

II. Existing Methodology

The first step towards making ATMs more accessible to visually impaired people was introducing "talking ATMs". The world's first talking ATM for the visually challenged was an NCR machine unveiled by the Royal Bank of Canada on October 22, 1997 [4]. These ATMs give audio instructions to visually impaired people, about what operations they can perform on the ATM. Once the user inserts the earphone in the 3.5mm audio jack, the ATM detects that the user is visually impaired and starts to give audio instructions. This was augmented by the presence of braille text on the keypads, and near the card slot and cash dispenser. These ATMs are used even today. Unfortunately, these ATMs cannot solve all the problems visually impaired people face while using the ATM. The next advancement in the talking ATM was the introduction of tactile keys around the ATM screen, to facilitate some common transaction types. This adds only a set of convex tactile symbols around a touch screen and an earphone jack as hardware to normal touch screen ATMs with no physical keys[5]. For example, "-" tactical key was given for withdrawal, "+" was given for deposit, "x" was given for cancellation, etc. However, the problem with this technique was blind users had to search for these keys, which increased the transaction time a lot. Moreover, there were some minor problems as well, like users were finding it difficult to locate the headphone and insert the headphone.

One major problem in designing ATM interfaces for visually impaired people is creating a secure and usable PIN entering environment. One solution proposed for this was using hand gestures for entering ATM PIN. Such a solution was presented in [6], where hand gestures were used to identify digits of PIN. The problem with this is that it was tested in a specific environment, i.e., the hand was placed inside a box, and its image was taken from there to identify the PIN. Hence, there could be a possibility that accuracy is less for this algorithm when tested in normal light conditions. A similar approach was also used in [7], but this algorithm had better accuracy and speed. In this approach, the user could also enter its username using hand gestures. But still, it was not able to counter the risks of snooping ATM PINs.

The next approach uses various sensors for enabling a user to do a transaction. The approach[8] is that visually impaired users can give various instructions like ATM PIN, withdrawal amount, etc. using flex sensors. These sensors are fitted in hand gloves, which users can wear to give instructions to the ATM. This approach also uses other sensors like IR(infrared) sensors and RFID(Radio-frequency identification). The problem with this approach is that it requires a lot of additional components to implement, like flex sensor, RFID, etc. Also, it is still unable to tackle the issues like preventing snooping of PINs or making the transactions quick. A similar approach of using flex sensors is also found in [9] demonstrated using the Arduino Mega board. The next approach is to implement ATM processing using voice command and fingerprint[10]. In this approach, only a single person is strictly permitted at a time. The door will be locked when the sensors detect human presence inside the cabin. The user can use voice commands to choose between normal mode and blind mode. After inserting the smart card, which already has the fingerprint template number of the visually challenged person encrypted on it, the user can use the fingerprint scanner which compares the one in the smart card with the fresh one for authentication. The user can even give voice commands to withdraw cash or check balance. This approach can be further improved, and new features can be added to it for making it more safe and secure. In continuation to this approach, we have developed our model.

III. SETUP

As we know that in the real world, the ATM's processing unit is a small microcontroller board, which does not have high processing capability. Hence, to create a real-like environment for testing the ATM interface, we are using a Raspberry Pi microcontroller board. The various components we have used to create this system are- Raspberry Pi 3B+, Display monitor, Mic, Headphones, HDMI connector , Power source, USB audio adapter, Raspberry Pi camera module V2, REES52 Flex Cable 2 meters, Mouse , Keyboard. Apart from these hardware requirements, the setup needs some software requirements as well. The software requirements of the systems are- Raspberry Pi OS(Operating System), Thonny Python IDE(Integrated Development Environment), Python 3, Guizero - Python library, Pyttsx3 - Python library, SpeechRecognition - Python library, Pyaudio - Python library, Google speech to text API(Application Programming Interface), cv2 - Python library for OpenCV, HOG(Histogram of Oriented Gradients) descriptor.

IV. DESIGN

The fig 1 is a visual representation of how different users in the system will interact with each other. It also displays what functions the ATM system will perform. The three main components of this design are:- the visually impaired user, the ATM interface, and the bank database.
The visually impaired user cannot directly communicate with the bank database to perform various actions. Hence, the ATM interface helps him/her to provide a bridge for communicating with the database. The user performs various actions in the ATM. The functions are inserting headphones, inserting ATM cards, entering ATM PINs, and giving audio instructions. These functions are interpreted by the ATM interface, and then necessary steps are taken by the interface. For example, when a user inserts the headphones in an audio jack, he/she is prompted with a welcome message and then given directions to the card slot. This can be seen in fig.1.
Now, for the actions which need a bank's database, an ATM interface is used. For example, when the user enters the ATM PIN, the ATM interface needs to verify that PIN to enable further operations. For this, the ATM interface compares the ATM PIN entered by the user to the PIN available in the bank database. If the PINs match, the user is given further options, for different types of transactions. The flow of different operations can be understood using the flowchart of fig. 2.

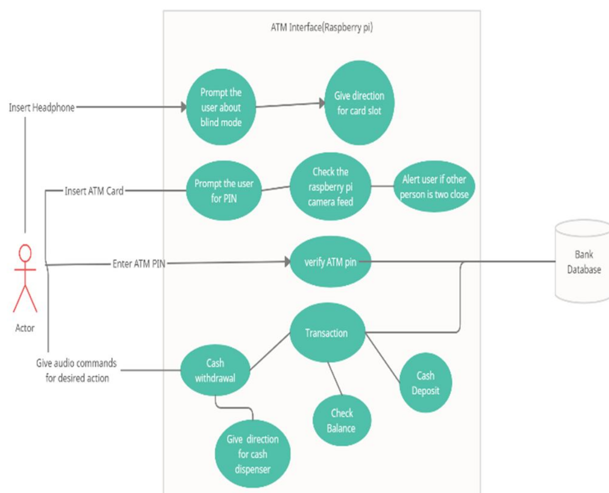

Fig. 1 Use Case Diagram

V. OUR METHODOLOGY

We have divided the system into four main modules. These modules are:-

1. User interface module.
2. Text-to-speech module.
3. Speech recognition module.
4. Distance detection module.

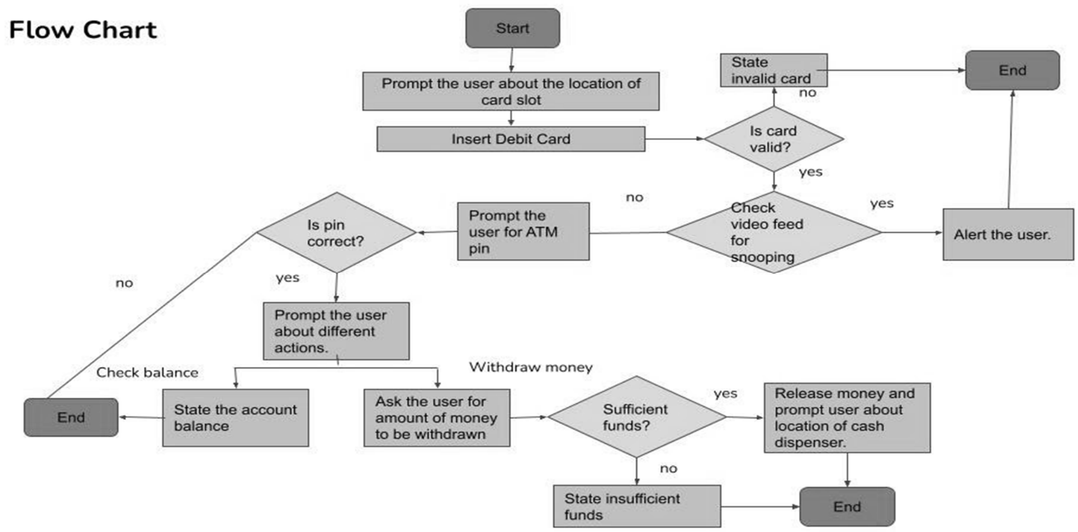

Fig. 2 Flowchart of the system

A. User Interface Module

This module deals with the designing of a user interface for the ATM system. This interface is built in the Python programming language. Our main aim was to keep the interface as simple as possible considering the limited memory in Raspberry Pi. As we know, Raspberry Pi has low processing power, hence the UI(User Interface) should be light to make the transactions quickly and efficiently. We have used the Guizero library in Python for creating UI[11].

The UI contains various windows for different functions in the ATM system. The different windows are- Welcome window, , Enter PIN window, choose a type of transaction window, Cash withdrawal window, Cash deposit window, Balance check window, Change PIN window

The welcome window contains a welcome message from the bank, and this window automatically redirects to enter the PIN screen once the card is detected. As of now, we do not have a system to read the ATM card, hence the window is automatically redirected to the enter PIN window after 10 seconds.

The next screen is the enter PIN window. This window contains a text box to enter the ATM PIN. While the user enters the PIN using the Braille keypad, the distance detection module is started, and it checks whether there is anyone else snooping on the ATM PIN of the visually impaired user. If someone is snooping on the ATM PIN then the user is prompted with an error message, and the user is given a choice whether to continue the transaction or cancel the transaction.

The next window is the Choose type of transaction window. In this window, the user can choose various actions he/she wants to perform. The four types of transactions that the user can do are cash withdrawal, cash deposit, balance check, and change the PIN. This window also communicates with the bank database to perform various transactions. For example, in the balance check inquiry, the interface sends a request to the bank database for fetching the user balance.

All the windows after this will perform tasks according to the title given to the window. For example, a cash withdrawal window facilitates the cash withdrawal from the ATM. It will ask for the amount to be withdrawn, then it will send a request to the bank database for checking the available balance, and if the balance is greater than the amount requested for withdrawal, the money is withdrawn. Similarly, other transaction types also function.

B. Text-to-speech Module

This module deals with converting the textual instructions to audio instructions. We have used the Pyttsx3 library to convert the text instructions to audio instructions. Pyttsx3 is a text-to-speech conversion library in Python. Unlike alternative libraries, it works offline and is compatible with both Python 2 and 3. An application invokes the pyttsx3.init() factory function to get a reference to pyttsx3 engine instance. It is a very easy-to-use tool that converts the entered text into speech. The engine object has the following variables that can be fine-tuned:[12].

1. Rate - Integer speech rate in words per minute. The default rate is 200.
2. Voice - String identifier of active voice.
3. Volume - Floating point volume between 0.0 and 1.0.
4. Voices - List of pyttsx 3.voice.Voice descriptor objects.

This module is activated once the user inserts headphones in the audio jack. The system starts with giving an audio welcome message to the user. Once the user swipes the card, the audio prompt to enter the ATM PIN is given. Furthermore, all the types of transactions which a user can perform at the ATM are given as audio instructions. Accordingly, all the instructions regarding the types of transactions are given vocally. Audio instructions are also given about the location of the card slot and cash dispenser. For example, the audio instruction that "card slot is located at the bottom left corner of the system" is given for helping the user find the location of the card slot.

Furthermore, the breakup of the cash dispensed is also given to the user using voice instructions. For example, if the user withdraws 10000 from the account, then he/she is told about the number of 500 notes and the number of 2000 notes dispensed by the ATM. These instructions vary according to the types of notes dispensed by the ATM.

C. Speech Recognition Module

The Speech recognition module is the part where the system takes audio inputs from users and converts them into text. We have used the NLP(Natural Language Processing) API, namely, Google speech to text API to convert the speech from a visually impaired person to text outputs. The python library for using this API is the SpeechRecognition library.

For audio input, we use a microphone with a USB audio adapter. Pyaudio python library is used to access the microphone input which is further processed using Google API. SpeechRecognition is a Python library for performing speech recognition with support for various APIs, both online and offline. Some examples of APIs supported by SpeechRecognition are CMU sphinx(offline), IBM speech to text, Google speech to text API, etc.[13]. Google has a great Speech Recognition API. This API converts spoken text (microphone) into written text (Python strings), briefly Speech to Text[14].

The module is used when the user is asked to select any option or provide any input to the system. For example, users can give the amount to be withdrawn through voice instruction. Like, when asked for the amount, the user can say "5000" which will be converted into a machine-understandable format and can further be processed for the desired action. We have also used audio inputs for selecting the type of instructions i.e., users can say 1 for cash withdrawal, 2 for a cash deposit, etc. Using such a speech recognition module enables the system to be fast and efficient.

D. Distance Detection Module

This module is very essential for ensuring the safety of users. By safety, I mean that it can help prevent snooping of ATM PINs, and further it can help to reduce ATM frauds on visually impaired people. Python libraries like cv2 (OpenCV) and imutils will be used to implement this module. We have developed this module taking inspiration from [15], which is a social distancing detector. In [15], the YOLO(You Only Look Once) detector is used to segregate humans from the background and then find the pairwise distance between these users. However, OpenCV already has a pre-trained HOG(Histogram of Oriented Gradients) + Linear SVM(Support Vector Machine) model to detect people in images and video streams. So, we use the HOG descriptor to find out humans in a video feed. After detecting the people, the pixel distance between the people can be used to determine whether snooping of PIN is possible or not, and if the snooping is possible, then an alert is given to the user.

Before implementing the module, we must enable the camera in Raspberry Pi configuration. Once enabled, we can use the video feed from the Pi camera to perform further operations. The video feed can be accessed in the Python program using the cv2 library. The Pi camera must be placed at the top of the room so that it can cover the whole area of the ATM room.

The main idea behind the HOG(histogram of oriented gradients) is that the local appearance and shape of an object in an image can be described by the intensity distribution of gradients or the direction of the contours. The implementation of these descriptors can be obtained by dividing the image into small, connected regions, called cells. Then, for each cell, we compute a histogram of gradient directions or edge orientations for all pixels of the cell. The combination of these histograms is the descriptor[16].

The module is activated when the user is prompted to enter the ATM PIN. The Pi camera video feed is accessed in the program and the HOG descriptor identifies the humans in the video feed. Then we draw rectangles on the humans identified using the imutils library. After this, we find the centroids of all the rectangles drawn for identified humans. After this, we find the pixel distance between all the identified centroid pairs, and if any of the pixel distance is less than the threshold, the visually impaired user is given an alert stating that “someone may be snooping your PIN, do you want to go forward”. After this, the user can take the necessary steps to either make the ATM room secure or cancel the transaction.

This image shows how the HOG descriptor identifies the humans in the video feed while ignoring the background details. These rectangles drawn around the humans can be used to detect the distance between the two humans and further to alert the user if the other user is too close.

Finally, all the modules are integrated to create one system. We could not test the system with visually impaired people; hence the efficiency of the system cannot be calculated. But still, we have tested all the modules individually and can get accurate outcomes.

VI. CONCLUSION

In this paper, we have designed an ATM interface for visually impaired people which overcomes the shortcomings of the existing system.

1. Our system provides audio instructions about various functions of the ATM to the blind user.
2. It also enables the user to give audio instructions to the ATM system for faster transactions.
3. It provides a safer transaction environment, by alerting the user if someone may be snooping on their ATM PIN.
4. It also tells the user about the location of the card slot and cash dispenser.
5. It also prompts the user about the breakdown of cash dispensed.

In the future, some functionalities can be added to this system.

1. Audio instructions can be given in other languages, like the local language of the area where the ATM is situated.
2. Two-factor authentication can be enabled while entering the ATM PIN, using voice authentication or a fingerprint scanner.
3. The location of the cash dispenser or card slot can be given in real-time, based on the location of the user's hands.

REFERENCES

[1] WORLD HEALTH ORGANISATION (2010). Global data on visual impairment 2010[Online].Available:https://www.who.int/blindness/publications/globaldata/en/

[2] B. R. Sanchez and G. G. Garcia, "Keyboard of Automatic Teller in Braille Code," 16th International Conference on Electronics, Communications and Computers (CONIELECOMP'06), 2006, pp. 39-39, doi: 10.1109/CONIELECOMP.2006.36.

[3] Lucas D. Introna & Louise Whittaker (2006) Power, Cash, and Convenience: Translations in the Political Site of the ATM, The Information Society, 22:5, 325-340, DOI: 10.1080/0197224060090427

[4] Dahiru, Umar & Elango, Sangeetha & Ariyaluran Habeeb, Riyaz Ahamed. (2015). System Design for Talking ATM for The Blind Users.

[5] Miki, Hiroyuki & Hirano, Kazuhiko & Suzuki, Kunikazu & Nomura, Masatoshi. (2005). Another type of talking ATM: ATM with tactile symbols for visually impaired users.

[6] G, Shruthi & Sarayu.K.P, Sarayu.K.P. (2015). Design of ATM Accessing System for Blind using Real-Time Video Processing through Gestures. International Journal of Computer Applications. 119. 47-49. 10.5120/21115-3941. .

[7] S. R. Rupanagudi et al., "A high speed algorithm for identifying hand gestures for an ATM input system for the blind," 2015 IEEE Bombay Section Symposium (IBSS), 2015, pp. 1-6, doi: 10.1109/IBSS.2015.7456642

[8] M. S. Kavitha,Soundariya. P,Vaishnu. J. M,Yamuna. M, ”An Embedded Based ATM for Blind People",IJESC,vol.3,issue 3

[9] Sasirekha, K., Nivetha, M., Indumathi, A., & Renukadevi, D. (2016). ATM Machine for Blind People. international journal of chemical sciences, 911-916

[10] Aneesh Chandran, Aiswarya Paulson, Abee Joe Varghese DeenShifaz ,”ATM for visually challenged people” IRJET , vol.4,issue 3, Mar. 2017 Art. No. V4I372

[11] Martin O'Hanlon, Laura Sach,”Create Graphical User Interfaces with Python”, in 2020 by Raspberry Pi Trading Ltd, Maurice Wilkes Building, St. John's Innovation Park, Cowley Road, Cambridge, CB4 0DS

[12] N.G Foong.”An Introduction to pyttsx3: A Text-To-Speech Converter for Python” .betterprogramming.pub [https://betterprogramming.pub/an-introduction-to-pyttsx3-a-text-to-speech-converter-for-python-4a7e1ce825c3 (accessed May 30, 2021)

[13] SpeechRecognition(v3.8.1),Accessed:May,30,2021[Onlline].Available:https://pypi.org/project/SpeechRecognition

[14] Pythonspot,"Speech Recognition using Google Speech API".[Online].Available:https://pythonspot.com/speech-recognition-using-google-speech-api/#:~:text=Google%20has%20a%20great%20Speech,excellent%20results%20for%20English%20language. (accessed May 30, 2021)

[15] Narinder Singh Punn, Sanjay Kumar Sonbhadra, Sonali Agarwal, & Gaurav Rai. (2021). Monitoring COVID-19 social distancing with person detection and tracking via fine-tuned YOLO v3 and Deepsort techniques.

[16] M. Kachouane, S. Sahki, M. Lakrouf and N. Ouadah, "HOG based fast human detection," 2012 24th International Conference on Microelectronics (ICM), 2012, pp. 1-4, doi: 10.1109/ICM.2012.6471380.