e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:04/April-2022 Impact Factor- 6.752 www.irjmets.com

VOICE-BASED ANDROID ASSISTANT FOR THE VISUALLY-IMPAIRED

Rutika Shetye*1, Komal Rane*2, Tanaya Desai*3, Ajitkumar Khachane*4

*1,2,3Student, Department Of Information Technology, Vidyalankar Institute Of Technology, Mumbai, Maharashtra, India.

*4Professor, Department Of Information Technology, Vidyalankar Institute Of Technology, Mumbai, Maharashtra, India.

ABSTRACT

With the growing demand for technology and gadgets, mobile phones have become an essential part of human life. The use of smartphones has tremendously increased over the past few decades. With the modernization and growth of new technologies, the evolution of smartphones has helped mankind in making their lives easy and fast. But still, there are certain areas where technology has made limited wonders. One of such kind is assistance devices for the blind. There are assistant technologies for the blind but they have limited capabilities. The aim of this paper is to enhance the possibility of developing an integrated solution and develop an android app that will serve as a blind assistant. The motive is to develop an app that will have features like object detection, currency detection, date and time, taking and reading out notes, and setting an alarm. The app will be a voice-based assistant with multilingual support. All the instructions will be given through voice commands. Also, there is a provision for authentication using biometrics.

Keywords: Voice Assistant, Blind Assistant, Text To Speech, Speech To Text, Object Detection, Currency Detection.

I. INTRODUCTION

The purpose of this project is to discuss the development of an Android-based Intelligent Software Assistant application for visually challenged or blind people. The application is an effort to extend help to the blind community and it is intended to help people with visual limitations and blind people to access Android-based devices so that they can use library resources by using android devices. Blind people face many different challenges daily in communicating with the world, traveling and many more. They have to depend on their sighted colleagues for making phone calls and accessing other mobile functionalities. Our project is a voice-recognizing application for mobile phones which helps to access most of the functionalities of the phone and will make it possible for visually impaired people to connect with society. The sighted user's people with limited reading ability can also use this application if they are involved in activities that prevent reading (e.g.: driving or other eyes-occupied situations). Due to a lack of necessary information in the surrounding environment, visually impaired people face problems and are at a disadvantage since visual information is what they require but it is not sufficient. With the help of advanced technology, the visually impaired can be supported. The idea is implemented through an Android mobile app that focuses on a voice assistant, object recognition, currency recognition, calls, etc.

II. SYSTEM DESIGN

The system is designed to work as a blind assistant. The system majorly works on the voice commands of the user. The user is given a welcome window with bilingual support in two languages including Hindi and English. Then the user is asked about the choices of the services he/she wants from the application. There are seven services the app provides. All the inputs are taken from the users through voice commands and then SST and TTS are used to display the output on the screen and pronounce the output via voice. The features like object detection, and money detection all can be accessed via voice commands. The output will be voice-based and text-based. It will be displayed on the screen as well as it would be read. TTS is the process that stands for text to speech, and it converts speech from text. TTS is important for voice output for voice feedback for the user. TTS is implemented in the software in which an audio capacity is required. When the user enters the voice command, TTS converts that voice into text format and performs a specific action. SPEECH TO TEXT(STT) Android has an inbuilt feature that is speech-to-text through which the user can provide speech input to the software. In the background, speech input will be converted to text and perform the action in the form of TTS.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:04/April-2022 Impact Factor- 6.752 www.irjmets.com

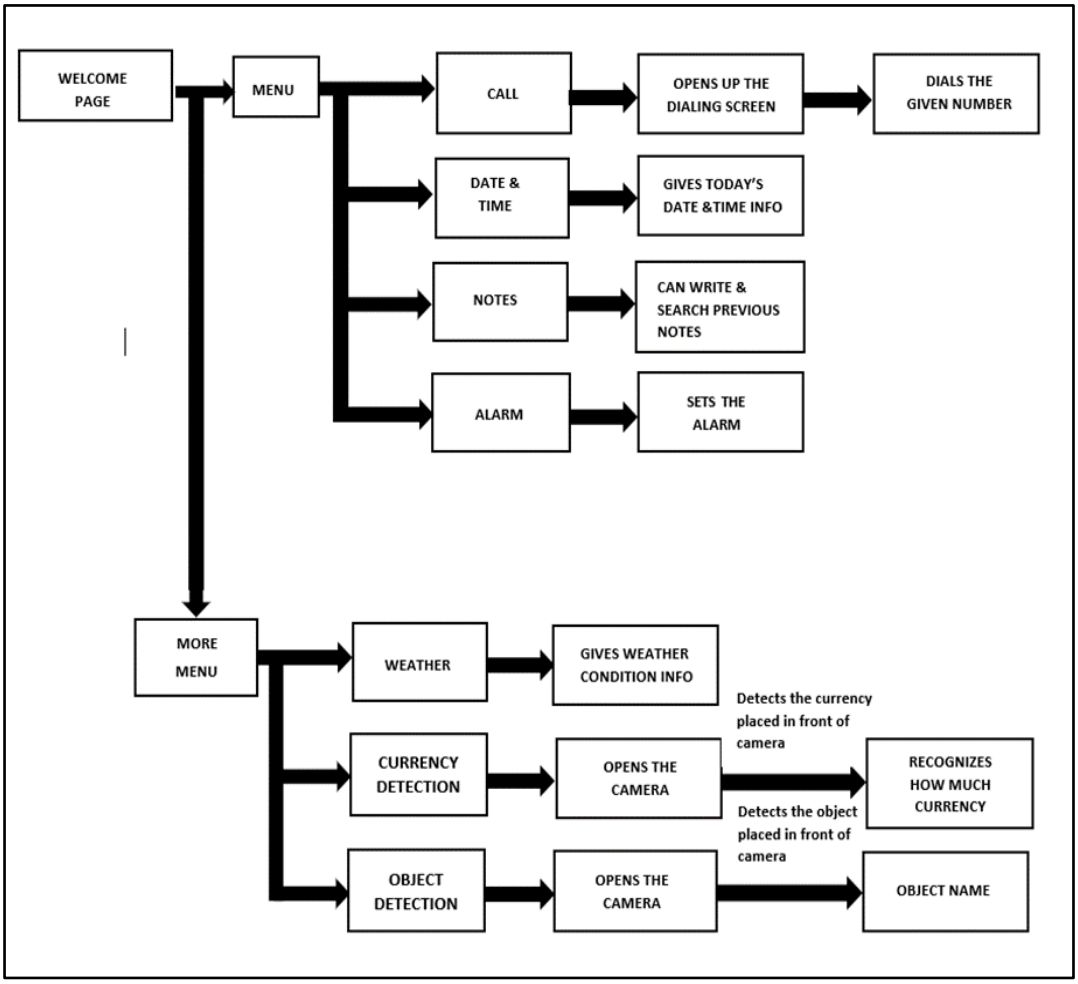

Figure 1: Data Flow Diagram

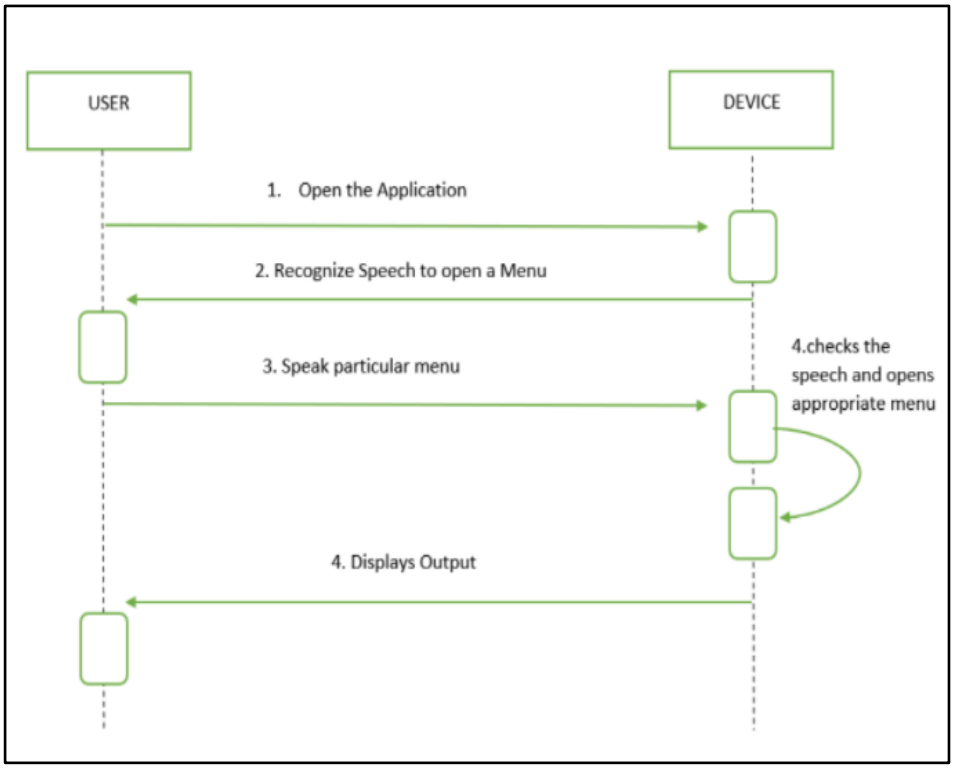

Figure 2: Sequence Diagram

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:04/April-2022 Impact Factor- 6.752 www.irjmets.com

III. IMPLEMENTATION

Modules Implemented

Object Detection

The object acquisition model is trained to determine the presence and location of multiple object classes. For example, a model might be trained with pictures that contain pieces of various fruits, as well as a label that specifies the category of the fruit represented (e.g. apple, banana, or strawberry), and data that specifies where each item came from the images. When a photo is given to a model, it will release a list of items it detected, a location for the binding box containing each item, and a score that shows confidence in detecting the right output.

Currency Detection

The function of identifying what an image represents is called image classification. The image classification model is trained to identify different categories of images. For example, you can train a model to see pictures representing three species of animals: rabbits, hamsters, and dogs. TensorFlow Lite provides advanced models that you can use in your mobile applications. During training, the image classification model is provided with images and related labels. Each label is a name for a different concept, or category, model that you will learn to identify. If sufficient training data is provided (usually hundreds or thousands of images per label), the image classification model can learn to predict that new images belong to any of its trained classes. The entire process for prediction mentioned above is known as inference.

Biometric Authentication For Login

Biometrics is the technical calculations and body measurements of individuals that are related to human characteristics. Biometrics are further separated into different categories. The first one is Physical and the second one is behavioural. Physical biometrics are our human features such as our fingerprints, iris, and many others. Behavioural biometrics include our actions. For example, this may include smiling, hand gestures, and others. In today's world biometrics has brought up an enormous change and an easy and more secure means to identify a specific user by making use of their physical traits. Biometrics, time and time again, have proved to be efficient and less prone to cyber-attacks.

Weather Detection

The Open Weather Map is a service that provides weather data, including the current data about the weather, forecasts about the weather, and historical data. It provides this data to web services and mobile app developers. It provides an API with JSON, XML, and HTML endpoints and a limited free usage tier. Making more than 60 calls per minute requires a paid subscription starting at USD 40 per month. This API provides an API key which is nothing but an URL that is appended with the latitude-longitude parameters to obtain JSON content. We then read the JSON content using Json Object Request class of the volley library. Then the result string is manipulated and we get the output string as voice. We need to add INTERNET permission to get results without the internet the result will not be obtained and ACCESS_FINE_LOCATION for latitude and longitude.

Alarm

In Alarm, we get the input from the user as a specified string and then separate the hours and minutes from the string obtained. Then using the alarm broadcast receiver class we generate a call to the alarm notification class at the specified time, on that call alarm notification class is intended which has a ringtone and alarm display for the user to cancel on the swipe moment on the screen.

Time and Date

In android, there is a Calendar class in the util library which gives the output as the current date and time. The date and time are then formatted in a string along with the day and then the output is given by voice.

Notes

In notes, we are using the SQLite database to store the notes in the phone database of the user. Here we are taking the notes input from the user on-screen press and the input is then stored in the database on swipe we get the activity where all the notes are displayed to hear the notes user needs to press on them so that the notes are read out to the user. To delete any specified note user needs to Long press on the screen. Notes Activity uses

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:04/April-2022 Impact Factor- 6.752 www.irjmets.com

READ_EXTERNAL_STORAGE permission.

Call

In Call, we are taking two inputs one is a mobile number and another in contact. The input is then checked for numbers and characters to identify if it is a number or contact name. According to the input string if it is a number, it is directly sent to the call intent of android to make a call and if it is a contact name then a cursor is created to go through the contacts of our phone which then matched if matched then the number of that contact is extracted and sent to call intent to place a call. Call Activity requires two permissions one is READ_CONTACTS and the other is CALL_PHONE.

IV. RESULTS

Fig 3: Home Page

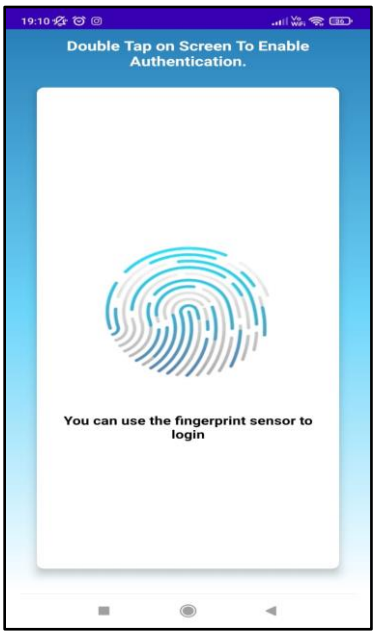

Fig 4: Login Page

Fig 5: Language Setup

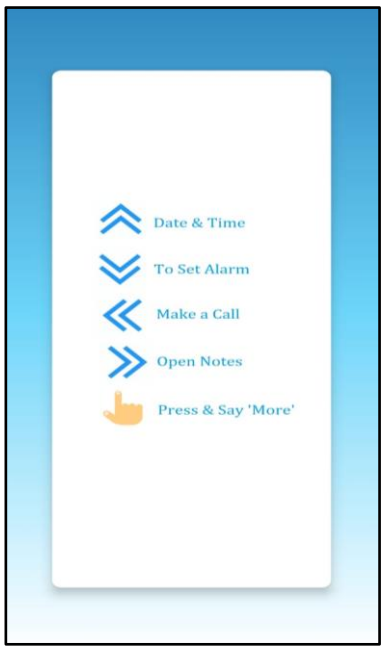

Fig 6: Main Menu

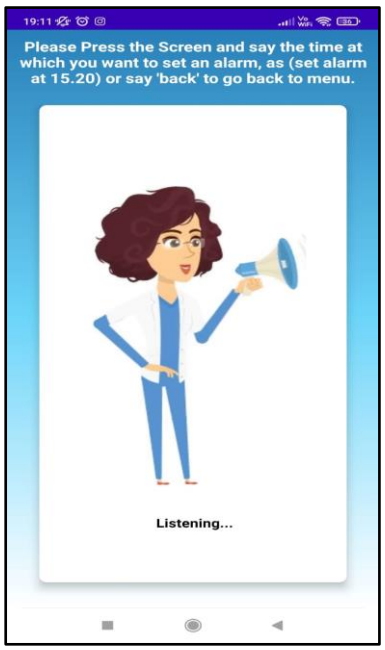

Fig 7: Set Alarm

Fig 8: Advanced Menu

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:04/April-2022 Impact Factor- 6.752 www.irjmets.com

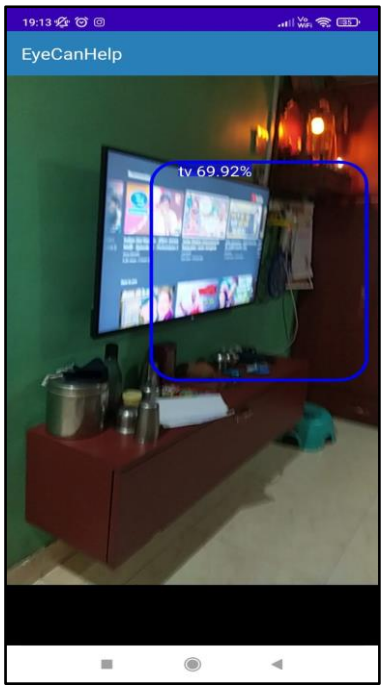

Fig 9: Object Detection

Fig 10: Currency Detection

Fig 11: Read Notes

V. CONCLUSION

Our project is an idea to solve the day-to-day issues faced by blind people and people with limited reading abilities. The application is an android solution that includes object detection, currency detection, calling features, weather statistics, and much more all accessible through the voice commands of the user. The aim is to provide an enhanced and integrated solution for blind people.

VI. FUTURE SCOPE

For future scope, we can add extra security login methods like login via OTP. By this, the information would be very much secure and the authentication system will be very much strong. Also, the addition of multi-lingual support will enhance the features of the system. The addition of other Indian languages will also boost user interaction. A voice-based system finder can be implemented. Where the user will be able to find the lost device by calling out certain commands. The device will be trained based on the user voice itself and then if it is under a detectable range where the voice can be reached easily the system will respond back to the user and get itself detected. Money detection and object detection features can be enhanced more by using larger datasets to train the model but the app size should be kept as minimal as possible. Also dynamic and fast object detection can be implemented to be used on roads .OCR technique can also be implemented where the image will be translated to text and then it will be read out to the user.

VII. REFERENCES

[1] Kwang B.Lee(2006, Jul), "The Design and Development of User Interfaces for Voice

[2] Jagtap Nilesh, Bendre M.R(2014, Feb)"COEP Voice-Based System in Desktop and Mobile Devices for blind people

[3] J.A.Larson (2003,Oct),VoiceXML:Introduction to Developing Speech Applications.

[4] Google Open Source Project. Talkback: An Open Source Screen reader for android.