



SMART ASSISTIVE TECHNOLOGY FOR ENHANCING THE INDEPENDENCE OF THE VISUALLY IMPAIRED

C.N Okwor¹, M.O Aghware, P.C Mbaoji and D.O Oyeka
Department of Electronic Engineering,
University of Nigeria Nsukka

Abstract: This paper presents the development of a mobile application utilizing a Digital Assistant to bridge the accessibility gap for visually impaired individuals. By leveraging advancements in Artificial Intelligence (AI) and Natural Language Processing (NLP), the assistant will empower users with greater independence by making daily tasks more manageable, reducing reliance on external support. It also leverages on text-to-speech and object recognition features to provide real-time information access and environmental awareness. The target audience comprises people with moderate to severe visual impairments who use internet-connected smartphones. This work fosters independence and autonomy for the visually impaired population, breaks down barriers to written content, opening doors to education and a more informed life and encourages the development of technology that caters to diverse user needs. Tests show a high degree of accuracy in performance of the features of the system

I. INTRODUCTION

In an increasingly modern environment, the digital gap for those with visual impairments deepens. Even seemingly easy chores like checking the time and making phone calls become overwhelming challenges. In an era of fast technology growth, the opportunity for innovation to improve the lives of people with visual impairments has never been greater. Despite great advances in many areas, the accessibility gap for the visually impaired remains large, particularly in terms of digital engagement and information access. People who are visually impaired have several difficulties in their everyday life, from finding information to getting around on their own. Conventional approaches frequently fall short of fully resolving these issues, leaving people with visual impairments in need of outside help or unable to access digital information on an equal basis with their sighted counterparts.

Recognizing these problems and the critical need for specialized solutions, our work aims to create an outstanding digital assistant that is particularly built to empower people with visual impairments. This work aims to close the accessibility gap by developing a digital assistant

specifically designed to meet the requirements of the visually impaired population. Users may safely sign up and log in to the application by incorporating biometric authentication features, which guarantees a smooth and user-friendly experience. By giving those with visual impairments an accessible login process, this solution not only improves security but also fosters inclusion.

Our main goal is to create a smart digital assistant that can comprehend speech requests and instantly respond with audio or other appropriate action. Users will be able to make inquiries on a variety of subjects and get voice responses, giving them the independence and efficiency to obtain knowledge. Users' access to timely and relevant data will be further improved by the digital assistant's ability to request news reports about particular locations and subjects.

Flutter[1] application was used for this work. We chose Flutter as our application's development framework since it enables the development of cross-platform apps with a native user interface and flawless operation. Because of its many accessibility capabilities, such as its ability to support screen readers and customizable widgets, Flutter is a great platform for developing inclusive apps that put accessibility first.

II. RELATED WORKS

Bose et al [2] presented the implementation of a voice-controlled system for the blind. The system uses speech recognition to convert speech input to text and allows for various tasks. The system has modules for speech recognition, text to speech, news feed retrieval, and Google search. It has the potential for future enhancements and is built using open-source software modules. The system is being ported onto a Raspberry Pi Model B+.

Tahoun N, Awad A and Bonny T [3] proposes a Smart Assistant device for blind and visually impaired people. The device uses computer vision techniques to help users navigate their surroundings. It consists of hardware components such as a Raspberry Pi, camera, ultrasonic sensor, bone conduction headset, and power bank. Software components include OpenCV and Python. The device aims to assist users in recognizing objects and colors. The device is portable, does not require constant WIFI connection, and



is powered by a 10000 mAh power bank and Raspberry Pi. It uses Pyttsx3 for text-to-speech conversion and Bluetooth for connectivity. MobileNetSSD is used for object detection and classification. The prototype consists of a 3D model with slots for buttons, camera, ultrasonic sensor, and power bank. The device is affordable, works in real-time, and has a power consumption of 2.5137 watts.

In [4], the researchers presented a project on "Virtual Assistance for Visually Impaired" developed by students and faculty at Rao Bahadur Y Mahabaleswarappa Engineering College in. The project aims to create a system/assistant that will assist visually impaired individuals by communicating with them through an earpiece, helping them recognize people, add new faces, and detect objects. They hope this innovative project will have a positive impact on the lives of visually impaired individuals. This work in [5] discusses the development of a voice-controlled desktop application for visually impaired individuals. The application uses speech recognition and synthesis technologies to enable users to control installed applications, send and receive emails, browse the internet, and access home appliances through voice commands. The system utilizes Microsoft SAPI and Bing Speech API for speech recognition and synthesis. The application also includes modules for email, news reading, web surfing, and home automation. The overall goal is to provide a user-friendly and affordable assistive technology solution for visually impaired individuals.

This paper [6] present a smart glass for visually impaired people to overcome the traveling difficulties. It uses an ultrasonic sensor, a microcontroller and the internet to send SMS to a guardian when the subject falls into any danger.

The authors of [7] presented the creation of a smart assistance system for visually impaired people, utilizing technologies such as object detection, speech recognition, and text-to-speech conversion. The system has features including object identification, Wikipedia searching, and email sending. It is built with the Python programming language and the YOLO object identification technique. The system returns audio results. The project employs YOLO v3, a real-time object identification system, and includes functions such as email sending, Wikipedia searching, and object recognition via voice help. The results are in audio format, which is beneficial for visually impaired users. The project can be improved by including other features, such as image captioning, which provides a brief description of the recognized objects.

In [8], the researchers discuss the implementation of a Personal Digital Assistant (PDA) system designed to assist both normal individuals and blind/visually impaired users in accessing online facilities. The PDA is described as inexpensive, lightweight, and easy to use, equipped with desktop-class operating system features such as music, web browsing, and smart calculations. It allows input in voice or text format, which is then parsed by the system to generate

desired output in text or speech format using Google voice API and SAPI5. The main objectives of the PDA system are to provide information to normal and visually impaired users through speech or text input. The system comprises modules for speech to text, speech recognition, and text to speech, enabling functionalities such as news feed, weather forecast, reminders, alarms, and creating to-do lists. The document also highlights the future scope of the PDA system, including the integration of home automation using microcontrollers and invoice recognition.

The Voice Assistant for the Visually Impaired in [9] is a system that functions as a voice assistant for visually impaired people. The system provides configurable messaging, a call record, a dialer, notes, battery level monitoring, and reminders. It allows users to read messages, notes, and dialer numbers aloud. The system is phone-dependent and does not require an internet connection. Text-to-speech (TTS) technology is used to provide vocal feedback and facilitate user interactions. The suggested method intends to make smartphones more accessible to visually challenged users. Existing solutions, such as Talkback, have limits in terms of power use and functionality. The system architecture diagram depicts the relationships between hardware and software components. Language options and other functionality for people with physical disabilities may be added in the future.

The study in [10] focused on how online higher education students with visual impairments used intelligent virtual assistants (IVAs). Participants discussed their ideas and experiences with IVAs, emphasizing benefits such as higher confidence and less limits in online education. The study emphasized the benefits of IVAs over traditional assistive technology and their favorable influence on students with impairments. Participants had positive experiences with IVAs, but also acknowledged accuracy concerns and frustrations.

The research presented in [11] hope to develop a wearable gadget with a virtual assistant system for visually impaired people. It comprises Arduino UNO, RF module, IR sensors, speech module, and earphones to help with independent life. The gadget comprises of a sensor-equipped headgear and a processor unit. Hardware components such as RF modules, encoders, decoders, and the CD4046 micro power phase locked loop are used. Technology such as Node MCU and Arduino boards with shields are used. The project's goal is to aid visually impaired people with daily duties using modern technologies.

The "Insight" smartphone app [12] is intended to help visually impaired people by offering a variety of services such as learning modules, messaging, calculator, games, scheduling, and navigation support. It employs a braille-based input system and seeks to improve accessibility for visually impaired individuals. The program received a lot of great comments and users were eager to use it. Future goals include adding more functionality and language support.



The research [13] is on a system that employs Deep Learning to aid visually impaired people. It includes modules for face and object detection. The system is a real-time web application that can handle both images and speech. It is implemented as a web application with facial recognition capabilities. Future ideas include voice calls and a text reader for visually challenged people. While the works mentioned here have solved the issue of accessibility for people with visual impairments to various degrees, our work aims to provide a cheaper alternative which would be accessible to people of lower economic capabilities.

III. SYSTEM DESIGN AND IMPLEMENTATION

This section outlines the methods and techniques utilized to achieve the objectives of developing a mobile app for the visually impaired. It describes the particular operations, techniques, and tools utilize to meet our project's objectives.

Figure 1 shows the flowchart illustrating the process flow for the voice-controlled application. It highlights the critical steps involved in processing voice commands, performing actions, and delivering feedback to the user. These steps are shown in the flowchart:

1. **Authentication:** This step verifies the user's credentials before proceeding.
2. **Voice Command:** Following successful authentication, the user interacts with the application through voice commands.
3. **Speech Recognition:** The application utilizes a speech-to-text conversion process to translate the spoken command into a machine-readable format.
4. **Assistant Command Check:** The processed command is then evaluated to determine if it falls within the scope of the virtual assistant's functionalities. If not, the interaction terminates.
5. **AI Processing:** Identified assistant commands are directed to the AI request module for further processing.

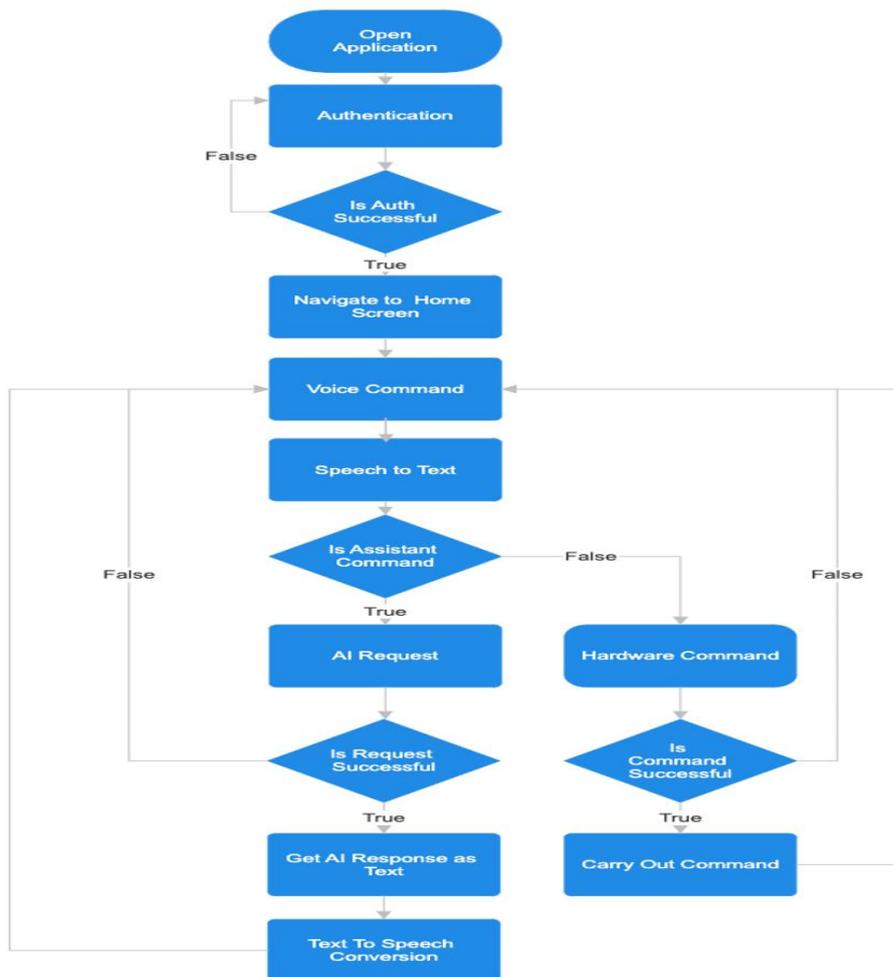


Fig. 1. System Flowchart



6. **Request Success Evaluation:** The outcome of the AI processing is assessed. If unsuccessful, the interaction concludes.
7. **Hardware Interaction Path:** For successful requests, the process diverges into two potential paths. The first path caters to hardware interactions. Here, the application generates a hardware command.
8. **Hardware Command Execution:** The generated hardware command undergoes evaluation to verify its success. If unsuccessful, the interaction terminates.
9. **Hardware Command Success:** Following successful hardware command execution, the application carries out the intended action.
10. **Text Response Path:** The second path from the successful request evaluation pertains to text-based responses. The application retrieves a textual response relevant to the user's query.
11. **Text-to-Speech Conversion:** The retrieved text response is then converted into an audible format using a text-to-speech module for user feedback.
12. **Home Screen Return:** The interaction concludes with the application returning to the home screen.

Among the countless biometric authentication solutions available, the `local_auth` package on `pub.dev`[14] emerges as a versatile and popular choice for integrating biometric authentication into Flutter applications. Some of its features are cross-platform compatibility, multiple biometric methods, user friendly integration and enhanced user experience.

STT is built on complex algorithms that have been trained on large amounts of speech data. These algorithms evaluate acoustic variables such as pitch, timing, and pronunciation to translate spoken language into text. Advancements in AI and ML have enhanced accuracy, making STT more trustworthy and user-friendly. Voice-activated features are now available in Flutter apps thanks to the speech-to-text package [15], which translates spoken words into text. Its features include: basic recognition, text handling, error handling and platform specific options.

Text-to-speech (TTS) is an assistive technology that converts any text into a speech [16]. This report delves into the fundamental principles, functionalities, applications, and future possibilities of TTS, exploring its impact on various sectors. At the heart of TTS lies a complex process involving natural language processing (NLP) and machine learning (ML). NLP segments text, identifies grammatical structures, and understands pronunciations. ML models predict speech patterns and intonation based on vast audio datasets, generating realistic and expressive voices. Features include: voice variety, customization, emotional expression, real-time conversation and integration with platforms. The text-to-speech package of flutter utilized in this work is `flutter_tts`[17].

Dart [18], an object-oriented programming language developed by Google is used for this work. Positioned as a language for building web, mobile, and desktop applications, Dart offers a comprehensive toolkit for developers to create high-quality software solutions across various platforms.

To create intelligent and customized virtual assistants (VAs), for the visually impaired, this paper integrates Gemini, a family of large language models (LLMs), with Flutter.

IV. SYSTEM FEATURE ACCURACY TEST AND ANALYSIS

4.1 Features

Time and Date

When a user says "time please," the voice assistant app activates its time retrieval function. This feature utilizes speech recognition to understand the user's request and then translates it into a call for the requested data. The app retrieves the current time and date, either from the device itself or a reliable network source, and converts that data into spoken language through a text-to-speech engine. Finally, the app vocalizes the time and potentially the date for the user.

Flashlight

- **User Commands:** The user can activate the flashlight by saying "Lumos" and deactivate it by saying "no Lumos." These commands offer a unique and potentially memorable user experience.
- **Technical Functionality:** Behind the scenes, the app leverages speech recognition technology to understand these spoken commands. It then employs natural language processing (NLP) to interpret the user's intent, which in this case is controlling the flashlight. Finally, the app interacts with the device's hardware to turn the flashlight on or off based on the user's command.

Read From

The app offers users two convenient options:

- **Reading from Existing Photos:** By saying "read from storage," users can access their photo gallery and select an image containing text.
- **Reading from Captured Images:** Saying "read from camera" activates the device's camera, allowing users to take a picture of an object with text they want to hear read aloud.

The app provides access to the device's storage or camera based on the user's chosen method ("read from storage" or "read from camera"). Once an image is selected or captured, the app employs Optical Character Recognition (OCR) technology to identify and extract text within the image. The extracted text is converted into an audible format using the



app's TTS engine. The app then reads the text aloud to the user.

4.2 Accuracy Test

To guarantee a smooth user experience in the world of voice assistant applications, accuracy testing is essential. Two basic elements are included in this evaluation process:

1. **Speech Recognition Accuracy:** This evaluates how well the app understands spoken commands. It involves assessing:
 - Ability to correctly transcribe what you say, even with background noise or different accents?
 - Ability to understand specific keywords and phrases relevant to its functionalities?

2. **Response Accuracy:** This assesses how well the app responds to your requests. It involves assessing:
 - Ability to answer questions correctly using reliable information sources
 - Ability to understand complex questions and provide helpful answers

Accuracy testing essentially verifies the voice assistant's capacity to reliably understand the user's request and take appropriate action. This thorough evaluation is essential to guaranteeing a trustworthy and easy-to-use interface between people and voice-activated technologies.

Table 1 shows the results of the speech-to-text recognition accuracy test

Table 1: STT Recognition Accuracy

Accuracy of STT	Frequency	Relative Frequency	% Frequency
TRUE	37	0.95	95
FALSE	2	0.05	5
Total	39	1	100

The STT system achieved near-perfect performance, scoring 37 correct transcripts out of 39 tests. This translates to a

95% success rate, meaning the system accurately converted spoken language to text in almost all cases.

Table 2 shows the results of the response accuracy test

Table 2: STT Response Accuracy

Response Accuracy	Frequency	Relative Frequency	% Frequency
TRUE	36	0.93	92.30
FALSE	3	0.07	7.70
Total	39	1	100

The report demonstrates the response system's effectiveness. It achieved an impressive 92% accuracy, correctly classifying 36 out of 39 responses as TRUE. With only 3 FALSE classifications (representing 8%), this performance indicates the response system is functioning very well.

Table 3 shows the accuracy test of flashlight activation and deactivation. . A total of 50 tests were conducted. In all 50

tests, the Lumos command successfully turned on the flashlight and there were zero instances where the Lumos command did not turn on the flashlight. The Lumos command achieved a perfect accuracy of 100% in controlling the flashlight during testing.

Table 3: Flashlight Activation Accuracy Test

Activation Accuracy	Frequency	Relative Frequency	% Frequency
TRUE	50	1	100
FALSE	0	0	0
Total	50	1	100

Table 4 shows performance of the "No Lumos" command in turning off the flashlight on the app under development. 50 tests were conducted to assess its effectiveness. In all 50 tests, the "No Lumos" command successfully turned off the previously activated flashlight and there were zero instances

where the "No Lumos" command did not deactivate the flashlight. The "No Lumos" command achieved a perfect accuracy of 100% in turning off the flashlight during testing. This indicates that the app functioned flawlessly in all test cases.



Table 4: Flashlight Deactivation Accuracy Test

Deactivation Accuracy	Frequency	Relative Frequency	% Frequency
TRUE	50	1	100
FALSE	0	0	0
Total	50	1	100

Table 5 shows the results of the accuracy test of the time check feature. The test assesses the accuracy of the "Time Please" command in providing the current time and date within the developed app. A total of 50 tests were conducted. In all 50 tests, the "Time Please" command successfully displayed the current time and date accurately and there were zero instances where the "Time Please" command provided incorrect information about the time or

date. The "Time Please" command achieved a perfect accuracy of 100% during testing. This indicates that the app functioned as intended in all test cases.

Accuracy of 'Time Please'	Frequency	Relative Frequency	% Frequency
TRUE	50	1	100
FALSE	0	0	0
Total	50	1	100

Table 5: Time Check Feature Accuracy

Table 6 shows the result of the call feature accuracy test. This works by saying "call x", where 'x' is a pre-saved contact on the mobile phone. A total of 50 tests were conducted. In all 50 tests, where the user had sufficient

airtime, the "Call Francis" command successfully initiated a call to the pre-programmed number and there were zero instances where the "Call Francis" command malfunctioned or initiated a call to an incorrect number.

Table 6: Accuracy of 'Call x' Feature

Accuracy of 'Call x'	Frequency	Relative Frequency	% Frequency
TRUE	50	1	100
FALSE	0	0	0
Total	50	1	100

Table 7 shows the test of accuracy or the 'Read from Camera' feature. The performance of this feature was analyzed in capturing and processing text from images within the app. A total of 50 tests were conducted. In 48 out of 50 tests, the "Read from Camera" command successfully captured an image, processed it, and accurately read aloud typed text present within the image. There were two instances (4%) where the app malfunctioned due to

limitations in text recognition. This was because of the system's inability to recognize a specific font style. The "Read from Camera" command achieved a success rate of 96% in accurately reading typed text from captured images, with font style being a potential factor contributing to some errors. This indicates that the app functions well in most cases, but there's room for improvement.

Table 7: 'Read from Camera' Feature Accuracy Test

Accuracy of 'Read from Camera'	Frequency	Relative Frequency	% Frequency
TRUE	48	0.96	96
FALSE	2	0.04	4
Total	50	1	100



The test results for the accuracy of the ‘Read from Storage’ feature is shown in Table 8. This test assesses the performance of the "Read from Storage" command in accessing the user's gallery, selecting an image, and processing text within the app. A total of 50 tests were conducted. In 47 out of 50 tests, the "Read from Storage" command successfully:

- Opened the user's gallery.
- Allowed image selection.

- Processed the selected image.
 - Accurately read aloud typed text present within the image.
- There were three instances (6%) where the app malfunctioned:
- Font Style: Similar to the "Read from Camera" function, specific font styles in the image hindered the text recognition.
 - In rare instances, the app misread the text.

Table 8: ‘Read from Storage’ Feature Accuracy Test

Accuracy of ‘Read from Storage’	Frequency	Relative Frequency	% Frequency
TRUE	47	0.94	94
FALSE	3	0.06	6
Total	50	1	100

V. CONCLUSION

This paper has presented a development of a voice assistant application with a specific emphasis on innovative user interactions and functionalities. The core features explored were speech recognition for accurate understanding of user commands, time and date retrieval upon user request, voice-controlled flashlight activation and deactivation using unique commands (“Lumos” and “no Lumos”), text-to-speech functionality that allows users to convert text from existing photos or captured images into spoken words (“read from” feature). The work involved the implementation of various technologies, including speech recognition engines, Natural Language Processing (NLP), Optical Character Recognition (OCR), and Text-to-Speech (TTS) engines. Test results show a high level of accuracy in the performance of the features with little errors which can be associated with systemic glitches and limitations.

VI. REFERENCES

[1]. Flutter, “Flutter - Build apps for any screen.” Accessed: May 15, 2023. [Online]. Available: <https://flutter.dev/>

[2]. P. Bose, A. Malphak, U. Bansal, and A. Harsola, “Digital assistant for the blind,” in 2017 2nd International Conference for Convergence in Technology, I2CT 2017, Institute of Electrical and Electronics Engineers Inc., Dec. 2017, pp. 1250–1253. doi: 10.1109/I2CT.2017.8226327.

[3]. N. Tahoun, A. Awad, and T. Bonny, “Smart assistant for blind and visually impaired people,” in ACM International Conference Proceeding Series, Association for Computing Machinery, Oct. 2019, pp. 227–231. doi: 10.1145/3369114.3369139.

[4]. Abhijna UJ, Dharani K, Kavya K, Roja G, and Prof. Kavitha Juliet, “Virtual Assistance for Visually

Impaired,” International Journal of Advanced Research in Science, Communication and Technology, pp. 649–655, Jun. 2022, doi: 10.48175/ijarsct-5087.

[5]. Southeast University (Bangladesh), Southeast University (Bangladesh). Department of Computer Science and Engineering, Southeast University (Bangladesh). Department of Electrical and Electronic Engineering, and Institute of Electrical and Electronics Engineers. Bangladesh Section, ICCIT 2019: 22nd International Conference on Computer and Information Technology : December 18-20, 2019.

[6]. M. Razu Miah and M. Sanwar Hussain, “A Unique Smart Eye Glass for Visually Impaired People.”

[7]. R. Reddy Appidi and U. Avinash, “A Smart Assistance for Visually Impaired,” International Research Journal of Engineering and Technology, 2023, [Online]. Available: www.irjet.net

[8]. R. Sarikaya, “The technology behind personal digital assistants: An overview of the system architecture and key components,” Jan. 01, 2017, Institute of Electrical and Electronics Engineers Inc. doi: 10.1109/MSP.2016.2617341.

[9]. S. Upadhay, S. Upadhay, A. Ghosh, and D. Jha, “Issue 3 www.jetir.org (ISSN-2349-5162),” JETIR, 2019. [Online]. Available: www.jetir.org94

[10]. S. Florida USF Tampa Graduate Theses et al., “Experiences of Using Intelligent Virtual Assistants by Visually Experiences of Using Intelligent Virtual Assistants by Visually Impaired Students in Online Higher Education Impaired Students in Online Higher Education Experiences of Using Intelligent Virtual Assistants by Visually Impaired Students in Online Higher Education,” 2019. [Online]. Available: <https://digitalcommons.usf.edu/etd>



- [11]. N. KumarV, “The Virtual Assistant: A Wearable Device for Independent Living of the Visually Impaired,” 2023. [Online]. Available: www.ijrti.org
- [12]. N. D. K. Guruge, R. A. G. P. Ranasinghe, S. H. P. D. L. Abeywardana, N. Siriwardhana, and D. H. De Silva, “INSIGHT: PERSONAL ASSISTANT MOBILE APPLICATION FOR THE VISUALLY IMPAIRED IN SRI LANKA.” [Online]. Available: www.apiar.org.au
- [13]. V. Sharma, V. Mahendra Singh, S. Thanneeru, and A. Kharat, “Virtual Assistant for Visually Impaired.” [Online]. Available: <https://ssrn.com/abstract=3580035>
- [14]. “local_auth | Flutter package.” Accessed: Mar. 22, 2023. [Online]. Available: https://pub.dev/packages/local_auth
- [15]. “speech_to_text | Flutter package.” Accessed: Mar. 22, 2023. [Online]. Available: https://pub.dev/packages/speech_to_text
- [16]. “Introduction to Text to Speech.” Accessed: Mar. 22, 2023. [Online]. Available: https://www.cs.cmu.edu/~srallaba/Learn_Synthesis/intro.html
- [17]. “flutter_tts | Flutter package.” Accessed: Mar. 22, 2023. [Online]. Available: https://pub.dev/packages/flutter_tts
- [18]. “A Bit About Dart - Learn Dart: First Step to Flutter,” Educative: Interactive Courses for Software Developers, Accessed: Mar. 22, 2023. [Online]. Available: <https://www.educative.io/courses/learn-dart-first-step-to-flutter/g7knn5r74ok>