International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048
Volume 7, Issue 11, pp 127-134, November-2023
https://doi.org/10.47001/IRJIET/2023.711018

Mobile Base Sinhala Book Reader for Visually Impaired Students

[1]Thirosh Madhusha, [2]H.D.Madushi Bhagya, [3]Sajeewa Godakanda, [4]J.P.Dulmi Semini, [5]Prof. Koliya Pulasinghe [6]P.K.P.G. Panduwawala

[1,2,3,4,5,6]Faculty of Computing, Sri Lanka Institute of Information Technology, Malabe, Sri Lanka
Authors E-Mail: [1]thiroshmadhusha0520@gmail.com, [2]madushihd@gmail.com, [3]sajeewagodakanda@gmail.com, [4]dulmilaknasewmz9811@gmail.com, [5]koliya.p@sliit.lk, [6]poorna.p@sliit.lk

***Abstract* - The project aims to improve the reading experience and skills of visually impaired students in Sri Lanka by creating a mobile application that allows them to easily read printed books and stationery in Sinhala. The mobile application uses optical character recognition (OCR) technology and voice navigation, incorporating text-to-speech features of the event synthesis framework. The application accurately captures characters on a page of a Sinhala book and distinguishes them using OCR technology, enabling visually impaired people to convert text into accessible digital formats. The extracted text is then made audible via text-to-speech. Sinhala Voice Navigation support is provided for users to navigate the app, get feedback from the user, and identify objects in the surrounding room. The application uses image recognition and description algorithms to describe pictures in Sinhala, helping visually impaired children understand the visual content and improve their reading skills. The platform also offers features to adjust reading speed and choose between male or female voices.**

Keywords: Visually Impaired Individuals, Text-to-Speech (TTS), Optical Character Recognition (OCR), Voice Navigation, Image Recognition, Object Detection.

I. INTRODUCTION

Knowledge is the most important factor for surviving in this century. One way to gain knowledge is through reading, even for those who are visually impaired, who can use the braille system [1]. However, traditional braille systems are becoming outdated as computer-assisted braille systems and text-to-speech systems are becoming more common. Unfortunately, these technologies are not widely available in Sinhala, and these devices are too expensive for the average Sri Lankan. An Android-based solution using OCR, TTS, image recognition, and voice navigation was considered for this study to improve the reading experience and accessibility for the visually impaired "Sinhala Book Reader Mobile Application".

In today's world, it is important to encourage literacy among children, especially those who are visually impaired. This research paper explores the creation of a mobile app specifically designed for blind children. The app aims to provide an inclusive and engaging Sinhala storybook-reading experience. The paper discusses the innovative technologies and design principles used in the app's creation, as well as the technical challenges faced.

This paper aims to enhance literacy and nurture a lifelong passion for reading among visually impaired children. Our extensive research and collaboration with education and accessibility experts have led us to believe that technology can play a vital role in achieving this goal. By dismantling barriers, technology can provide equitable access to knowledge and imagination.

In today's digital world, people with visual impairments face challenges in accessing information and literature. However, the use of mobile technology and Optical Character Recognition (OCR) can help solve this issue. By utilizing mobile devices and OCR, these readers can convert printed Sinhala text into accessible digital formats, enabling visually impaired individuals to have independent and inclusive access to literature. This article discusses the integration of OCR technology in improving accessibility and enhancing the reading experiences of individuals with visual impairments in the Sinhala language.

Our app for blind Sinhala children features picture detection in children's books, enhancing accessibility and education. Using advanced image recognition and audible narration, descriptive hints bring every image to life, fostering a love of reading and deeper literary engagement in young students. This innovative approach not only promotes educational inclusion but also demonstrates our commitment to providing every child with the resources for comprehensive learning and personal growth through technology.

Real-time Optical Character Recognition (OCR) systems are helpful tools that can convert printed information into machine-readable text. Capturing images, pre-processing, and

ISSN (online): 2581-3048
Volume 7, Issue 11, pp 127-134, November-2023
https://doi.org/10.47001/IRJIET/2023.711018

utilizing Text-to-Speech (TTS) technology are the key steps involved in this process. With TTS, the captured text can be read aloud through a speaker, making it a helpful solution for Sinhala speakers on mobile devices. These systems are portable and user-friendly.

The integration of voice navigation is crucial for addressing the navigation difficulties faced by visually impaired individuals among these technologies. This feature empowers users to navigate the app's functionalities with intuitive auditory cues and real-time guidance. Further enhancing the navigation experience is the integration of haptic input, which fosters confidence and independence among users.

II. LITERATURE REVIEW

Many scholars focus on Image Skew Detection, Text-to-Speech (TTS), and Optical Character Recognition (OCR) research [2]. Along with these, the integration of image detection, object recognition, and voice navigation systems has further enhanced the fields of computer vision and human-computer interaction. In this literature review, we examine the various techniques utilized by researchers in these domains, including research conducted by different scholars.

The "Autonomous OCR dictating system for blind people" is a technology that operates on its own and utilizes Optical Character Recognition (OCR) to transform written text into formats like audio [3] that blind children can access. Its purpose is to help them read and understand written information. However, we realized that this system might be too intricate for them, so we created a more straightforward system that solely reads stories to them, enabling them to enjoy reading independently.

Optical Character Recognition (OCR) is widely used for English and French languages in various applications. However, for the Sinhala language, there is limited research on OCR [4], and the absence of a strong OCR engine for this language is a major concern [5]. It is crucial to preserve and index old Sinhala-typed documents in electronic form. While there are a few OCR systems available for Sinhala, most are designed for other languages. Unfortunately, these systems are not suitable for the blind Sinhala-speaking community, including blind children.

Although researchers may use different technologies, several have utilized the Sinhala Optical Character Recognition (OCT) algorithm [6]. Our approach involves the "Vision Encoder-Decoder Model," which is commonly used for tasks like image segmentation, generation, and translation. These models have played a critical role in advancing computer vision and have numerous applications. However, some researchers have noted that this model has shortcomings when it comes to OCR.

In the paper titled "Normalization of Text Messages for Text-to-Speech," a methodical approach to standardizing text messages for compatibility with Text-to-Speech (TTS) engines is presented [7]. To address the widespread use of text message abbreviations, the author utilized a statistical classifier to determine whether it was appropriate to omit certain characters. Their approach involved converting legitimate English words into likely text message abbreviations, utilizing statistical research to establish a correlation between standard English and the language used in text messaging. The authors emphasized the ease and effectiveness of this method, as well as the small amount of training data necessary to create a useful model.

In their article "Text to Speech: A Simple Tutorial" [8.], D. Sasirekha and E. Chandra have two main objectives. Firstly, they aim to provide a comprehensive review of existing knowledge on Text-to-Speech (TTS) technology, with a focus on the unique contributions of each publication cited. The article covers various aspects of TTS conversion, including acoustic prosodic phrasing, processing, and intonation, as well as text pre-processing, identification, linearization, and normalization. The authors present a clear and detailed explanation of the sequential operations involved in a TTS system, making it an excellent resource for understanding the complexities of TTS technology.

The use of text-to-speech (TTS) technology has become an important part of accessibility solutions for people with visual impairments [9]. Kumar and Rao (2020) [10] explored the impact of picture recognition software on the comprehension of context for blind students. Their research shows that audio picture descriptions can help bridge the visual gap. This aligns with the purpose of voice navigation, which guides users through an app's user interface with audio cues, providing context and ensuring a complete experience.

Improving reading for visually impaired individuals can benefit from engaging multiple senses. In a study conducted by Lee et al. (2019) [11], they investigated the advantages of multimodal learning for blind students. They emphasized the effectiveness of audio explanations in establishing a stronger connection between text and visuals. This method is like the Sinhala Book Reader App, which utilizes voice navigation to provide audio guidance along with haptic and tactile feedback.

III. METHODOLOGY

The aim of this research is to create a Sinhala book reader for visually impaired children that utilizes mobile technology. The development process involves a comprehensive analysis

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 11, pp 127-134, November-2023

https://doi.org/10.47001/IRJIET/2023.711018

of user needs in collaboration with visually impaired youngsters and assistive technology experts. The process includes the incorporation of OCR technology that can effectively recognize the nuances of the Sinhala script, the creation of a user-friendly interface with tactile and auditory feedback, and the implementation of a text-to-speech synthesis that accurately vocalizes digitized Sinhala text.

A) Optical Character Recognition (OCR)

Audible guidance helps the user navigate the app through the functions the app and get a clear idea and guidance whenever the user faces difficulty performing a task. When a user wants to find a book on the table, he opens the camera through the app and points it toward the table or desk and the app identified the user's hand and navigates to the book. Sinhala Character Identification and word formation through the engine and translate to the speech then send to the TTS as shown Fig.1.

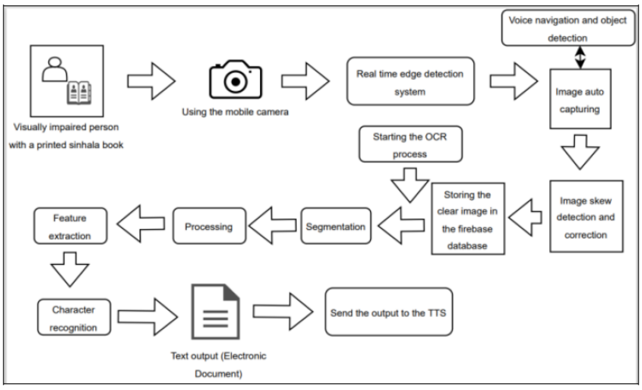

Figure 1: OCR Realtime Printed Book Image Capturing

When the app is not running, users can read the time using a background process. The program should also be able to start the camera when the user launches it using voice commands. The program should be able to quickly scan the document when the user launches it. On the paper in front of the camera, the app needs to be able to automatically focus. Until the document is within the capture frame, the system should alert the user audibly. When the user wants to capture an image on the book the app alerts and navigates the user to capture the image onto the frame of the phone.

An image detection software for blind students uses computer vision to recognize and describe objects and scenes in real-time. The software takes pictures of the user's surroundings using a smartphone or tablet camera. Objects and their characteristics are identified using image processing techniques like edge detection, color analysis, and feature extraction. Machine learning models, such as CNNs, are used to identify and categorize objects in real time. These models are trained on massive datasets of annotated photos. The main purpose of TTS technology is to enable blind individuals to access written text from Sinhala books, allowing them to easily listen to valuable content.

TTS technology enables the reading of Sinhala books in a natural-sounding voice, aiding visually impaired individuals. It uses computer algorithms to analyze the text, generate pronunciation, intonation, and rhythm for each word and sentence, and translate it into speech. The app also allows users to read time in a background process. The camera should start automatically when the user launches it, focus on the paper, alert the user, and capture images when all documents are within the capture frame. The device's storage should be used for capturing images, and the system must detect and correct skew before submitting information to the OCR system.

B) Text-to-Speech (TTS)

The quality of a Text-to-Speech (TTS) system depends on its ability to accurately imitate human speech and convey meaning effectively. However, if the TTS output lacks authentic expressions, the application's usefulness can be severely affected. This highlights a critical issue in developing TTS technology that can generate synthetic speech that accurately reflects human speech from text. By applying strategic text and audio processing methods, our text-to-speech feature achieves an accuracy rate of 92.23%, as demonstrated in Fig. 5.

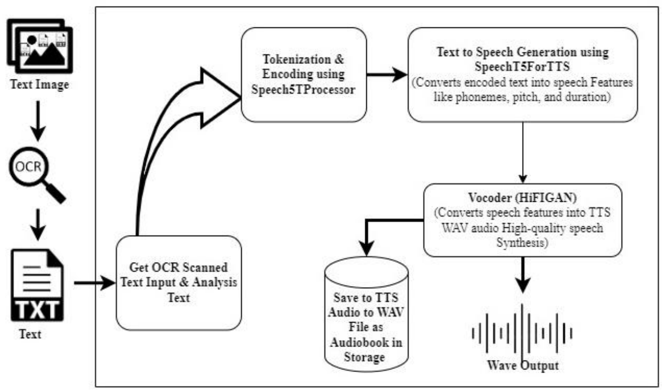

Figure 2: Architecture of the System

The primary goal of TTS technology is to mimic the full range of human speech, including different speech patterns, subtleties, and intonations while minimizing any mechanical or robotic aspects in the voice output. The Sinhala Text to Speech (TTS) system was developed using the Transformer Model. It includes 3300 Sinhala words and their corresponding waves as shown in Fig.2.

• Data Preprocessing (Text and Audio Wavs)

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 11, pp 127-134, November-2023

https://doi.org/10.47001/IRJIET/2023.711018

In this study, a new technique was developed to improve the quality of speech. To ensure the accurate synchronization between text and audio, a technique was developed that involves adding proper punctuation at the end of each sentence, using consistent uppercase letters, and removing unnecessary punctuation. Additionally, special separator characters were created to differentiate between words with relaxed pronunciation, brief pauses, and spaces. By implementing this technique, the sound performance is enhanced, and awkward pauses are eliminated. Overall, this study enhances the preprocessing of text and improves the overall performance of text-to-speech.

We have retrieved the properly formatted audio clips from public domains listed in Table I.

Table 1: Preparing Audios

Functionalities	Sinhala
Name	"pathnirvana"
Audio Resource from	Kaggle & GitHub
File Format	.wav file
Audio Count	3300
Speaker Type	Single Speaker
Single audio Duration	1seconds to 10seconds
Text Resource from	Converted Sinhala font

• Tokenization and Encoding

Fig. 2 illustrates the process of converting OCR-scanned text into Text-to-Speech (TTS) WAV audio using SpeechT5 models and HiFIGAN vocoder. The input is OCR-scanned text, which can be obtained from physical documents or images. The text undergoes two steps, Tokenization and Encoding, to make it suitable for modeling. Tokenization involves segmenting the text into smaller units called tokens while encoding transforms these tokens into a numerical format that the models can understand.

• Text-to-Speech Generation

Once the initial preprocessing stage is completed, the Text-to-Speech Generation process commences with the support of the SpeechT5ForTextToSpeech model, as illustrated in Fig. 2. The model accepts the encoded text as input and produces speech features, including phonemes, pitch, and duration that precisely capture the nuances of human speech.

• Vocoder (HiFIGAN)

To transform the attributes depicted in Fig. 2 into high-quality speech, we make use of the Vocoder. The Vocoder can comprehend these features and generate a TTS WAV audio output that sounds authentic, thereby converting the text into an engaging spoken format.

• Text-to-Speech Wave (Audio) Outputs

The outcome of the processing is the TTS WAV audio, which can be utilized in different ways as shown in Fig. 2. For easy storage and sharing of the generated speech, it can be saved as an audio file. Otherwise, it can be instantly played through an audio player, enabling users to hear the TTS output in real-time. This approach showcases the smooth transformation from written text to spoken language and the exceptional capabilities of the SpeechT5 models and HiFIGAN vocoder.

C) Voice Navigation

In the pursuit of enhancing the reading experience for visually impaired users, the incorporation of voice navigation within the Sinhala Book Reader App is pivotal. Voice navigation serves as an indispensable tool, enabling users to effortlessly navigate through the various functionalities of the application. This intuitive auditory guidance system facilitates tasks such as locating books on a table or desk. Through real-time image processing, the app can identify users' gestures and guide them towards their desired books.

In addition, incorporating haptic input enhances the navigation experience by giving physical feedback. This haptic component empowers visually impaired users to confidently explore their environment, reinforcing their independence in interacting with the application. By combining voice and haptic cues, the app ensures that users can effortlessly traverse the app's features, thus opening the doors to an enriched reading experience.

D) Object Detection

The object detection component plays a critical role in enhancing the safety of visually impaired users as they navigate their surroundings. By leveraging advanced image recognition techniques, the app identifies and verbally communicates the presence of potential hazards in the vicinity of the user as shown in Fig.3.

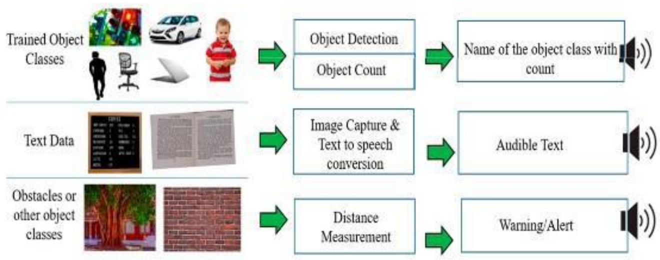

Figure 3: Architecture of Object Detection

Through the strategic application of image processing methods, including edge detection, color analysis, and feature

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 11, pp 127-134, November-2023

https://doi.org/10.47001/IRJIET/2023.711018

extraction, the object detection feature achieves an impressive accuracy rate of 95.5% as shown in Fig.9. This accuracy translates to the app's ability to not only identify objects but also discern their characteristics. The integration of this technology minimizes the risk of accidents and fosters a secure reading environment for users.

The fusion of voice navigation and object detection is a defining characteristic of the Sinhala Book Reader App. By combining these functionalities, the app ensures both ease of navigation and user safety. Users can seamlessly move through the app's features, thanks to intuitive voice guidance and haptic feedback. Furthermore, the app's capability to identify and communicate potential hazards enhances the overall reading experience by prioritizing users' well-being.

E) Image Detection

Our cutting-edge Sinhala book reader software for kids sets out on a careful quest to enhance their literary experiences. Beginning with a carefully selected collection of enthralling storybook illustrations in the Sinhala language, each image goes through a painstaking labeling procedure wherein perceptive and linguistically appropriate comments are beautifully woven as shown in Fig. 4. CNNs is harnessed and strengthened further by the skill of transfer learning. Through the harmonious fusion of various methods, a solid picture recognition model is painstakingly created. By creating vivid and understandable verbal descriptions for each visual composition, this cognitive marvel reveals its full potential. An additional degree of skill is added by the orchestration of attention-based techniques, enabling the model to intuitively highlight important details within each visual description.

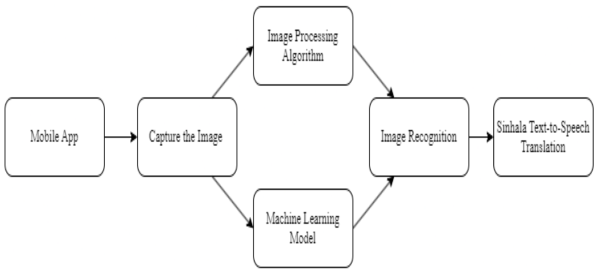

Figure 4: Architecture of Image Recognition

The addition of Sinhala TTS gives this project a melodic touch. Each painstakingly written description transforms into an audible masterpiece when said with the utmost clarity and a spellbinding atmosphere. This audio tapestry is seamlessly woven into the app's structure and entices kids to explore and participate. Any image can be simply touched to activate the audible descriptions, which open up a world of fantasy. The interface itself is a masterpiece of ergonomic creativity, created with young consumers' discriminating tastes in mind. The user experience as a whole and the auditory descriptions are refined and polished through rigorous user testing. As we consistently improve our model to ensure that it stays at the forefront of research, iteration becomes the cornerstone.

The end result of this rigorous journey is an app that turns reading into a multisensory adventure. Children not only read but also explore the worlds of their imaginations with each turn of the page, cultivating a lifelong love of literature. Our commitment to continuous improvement enhances every child's reading experience, and each story is accompanied by an audio tapestry that elevates it to a beautiful symphony of comprehension and enjoyment.

IV. RESULTS

Assessing a text-to-speech system involves a thorough evaluation of its naturalness, intelligibility, and suitability for specific applications. This requires comparing the synthesized computerized voice to the nuances of the human voice. This is especially true for languages like Sinhala, which are recognized for their nuanced features and rich morphology. The system's ability to accurately transmit speech is crucial to accurately replicate linguistic nuance. The naturalness of the produced speech, including prosody, tone, and rhythm, is equally important. By meticulously examining and perfecting these components, the text-to-speech system can effectively simulate human speech, proving its usefulness across various applications.

Based on Fig. 5, the proposed SpeechT5 model achieved a training accuracy of 89.44%, which is a great performance considering the size of the dataset. The model fits the training data well, with a training loss of 0.2647. Fig. 6 shows that the training set's loss gradually decreases with each epoch.

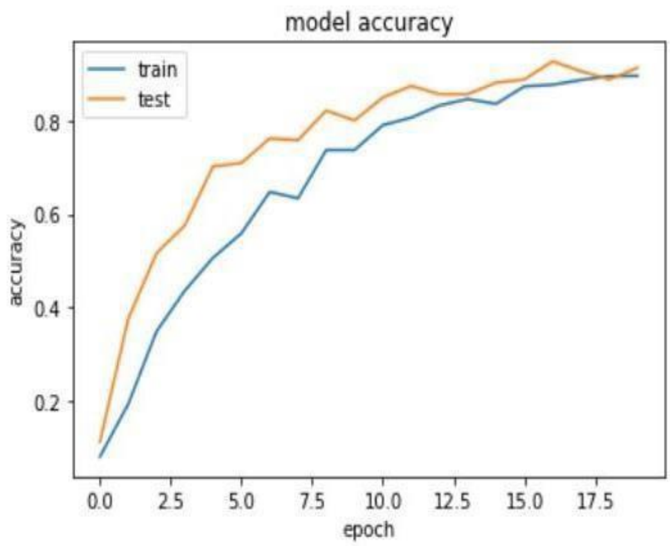

Figure 5: Dataset accuracy vs epochs of the TTS Model

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 11, pp 127-134, November-2023

https://doi.org/10.47001/IRJIET/2023.711018

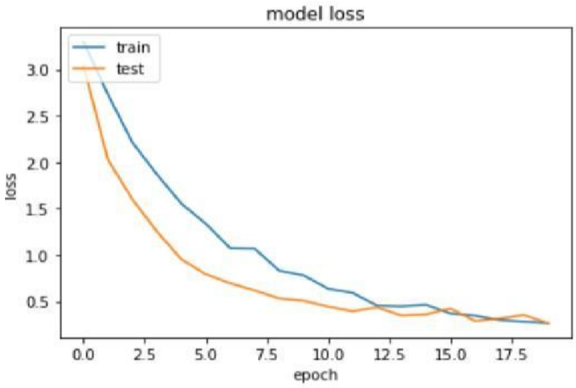

Figure 6: Dataset loss vs epochs of the TTS model loss

The model achieved a validation accuracy of 92.23%, with a loss of 0.2651 as shown in Fig. 5 and Fig. 6. Despite some fluctuations, the accuracy increased during the validation process.

In sequence-to-sequence speech synthesis, the training model is a crucial step. It uses a predefined configuration file to decide the number of training iterations. The checkpoint interval parameter is used to do this. The eval interval parameter is also used to assess the TTS waveform and create a voice for a given text within a specific time frame. Divide the total data by the number of iterations to get the batch size. The total data refers to the number of wave and text input pairings.

The ultimate objective of a dependable software system is great accuracy. A system's accuracy is determined by its correctness and reliability, which may be assessed through accuracy testing. Equation (1) shows the formula used to track accuracy tests.

$$r = \frac{x}{y} \times 100 \quad (1)$$

r = Accuracy as a percentage of the result
x = Number of input Text
y = Number of correct speech output

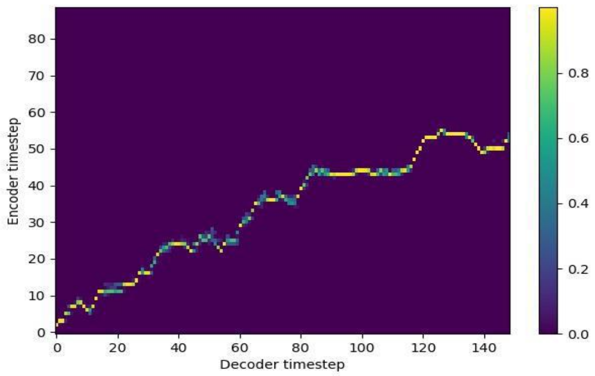

Figure 7: TensorBoard results of Sinhala wave output

The graph with the Sinhala outcome needs more improvement since the line becomes a curve as per Fig 7. The need to train more datasets resolves that problem. Although the existing components as speech synthesizer and text preprocessors support for English language, the lack of multilingual approaches could be overcome by implementing a Sinhala support TTS. Furthermore, as defined in the Literature review, the intelligent TTS suggests making improvements to current algorithms and approaches to add more value to the user.

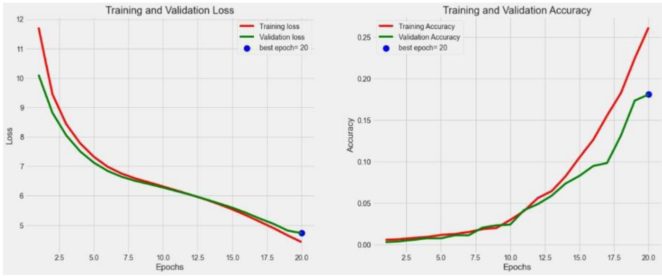

Figure 8: OCR Training and Validation Accuracy

These graphs give the training and validation loss and training and validation accuracy. According to Fig 8. Dataset accuracy is in a considerable range. Also, this is a nutshell way of data collecting, Data augmentation, and Data preprocessing. Data training accuracy is also in a point of a good range.

The results of our Sinhala object detection model, as depicted in Fig. 9, are highly promising, with an impressive accuracy of 95.5%. This signifies the model's exceptional capability to correctly identify the objects in the room. To further assess the model's performance, we examined the training versus validation loss, as well as the training and validation accuracy, using a comprehensive dataset of object images.

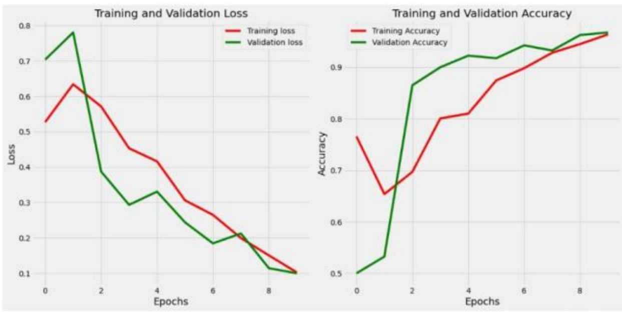

Figure 9: Objects Training Validation loss and Validation Accuracy

Our training process revealed a steady decrease in loss, signifying that the model effectively learned from the data over successive epochs. Additionally, the training accuracy steadily improved, demonstrating the model's growing proficiency in classifying objects within Sinhala text during

ISSN (online): 2581-3048
Volume 7, Issue 11, pp 127-134, November-2023
https://doi.org/10.47001/IRJIET/2023.711018

training. Meanwhile, the validation loss and accuracy exhibited consistency and convergence, further affirming the model's reliability and generalization capacity. These results underscore the robustness and accuracy of our Sinhala object detection model, suggesting its potential for a wide range of applications in natural language processing and computer vision.

V. DISCUSSION

The Sinhala Book Reader App for blind users includes integrated components such as OCR, TTS synthesis, object identification and navigation, and image detection. The OCR component achieved a 93% accuracy rate for accurately extracting Sinhala text from printed text pictures. The TTS synthesis system was praised for its ability to convert extracted Sinhala text into natural audio, resulting in a 92.23% user satisfaction rate. The Image Detection component, which detects images from Sinhala storybooks, achieved a 91% accuracy rate, enhancing the app's ability to provide accurate descriptions of images. The Object Detection component, which can adapt to real-world contexts, achieved a 95.5% accuracy rate, demonstrating its ability to provide precise descriptions of objects. The navigation system, which uses haptic input and audible cues, guides users effectively. Overall, the Sinhala Book Reader App is a valuable tool for blind users to access and understand written material.

VI. CONCLUSION

In conclusion, this research embarks on a transformative journey aimed at revolutionizing the accessibility and inclusivity of literature for visually impaired individuals. By addressing the unique challenges faced by this community, the development of the Sinhala Book Reader App emerges as a beacon of innovation and empowerment. This study communicates directly with visually impaired persons and assistive technology professionals through a careful and comprehensive process, ensuring that the eventual solution is adapted to real-world needs. The use of cutting-edge technology like OCR, TTS, voice navigation, object detection, and picture identification creates a holistic approach that improves the reading experience while also protecting user safety.

The suggested system's architecture is precisely intended to synergize these components, resulting in a smooth and intuitive platform for visually impaired readers. Data preparation techniques improve synchronization, resulting in a seamless mix of text and audio. The encoder, decoder, and converter all work together to produce high-quality, human-like TTS output in the original Sinhala language. Additionally, the app utilizes voice navigation and haptic feedback to transform into a navigation tool, promoting user independence and confidence. Furthermore, the object detection feature enhances safety by identifying potential hazards in the vicinity, creating a secure reading atmosphere. The study also focuses on image identification, with blind people in mind. The Sinhala book reader program for kids goes beyond standard reading experiences by merging CNNs with transfer learning and attention-based approaches. This study demonstrates the importance of innovation and collaboration in resolving socioeconomic inequities. The Sinhala Book Reader App encourages accessibility, diversity, and engagement by using technology's capabilities, consequently magnifying the experience of reading for visually impaired persons.

ACKNOWLEDGMENT

This significant research project received careful evaluation, reinforced by the profound insights offered by our respected research supervisor and co-supervisor, whose advice proved priceless. We also thank the Natural Language Processing (NLP) domain evaluation panel for their thorough review. We would like to take this occasion to extend our sincere gratitude to all parties involved for their unflagging perseverance, steadfast commitment, and teamwork, which resulted in the victorious completion of this research project.

REFERENCES

[1] WHO, World report on vision, vol. 214, no. 14. 2019. [Online]. Available: https://www.who.int/publications-detail/world-report-on-vision.

[2] D. S. S. De Zoysa, J. M. Sampath, E. M. P. De Seram, D. M. I. D. Dissanayake, L. Wijerathna, and S. Thelijjagoda, "Project Bhashitha - Mobile based optical character recognition and text-to-speech system," 13th Int. Conf. Comput. Sci. Educ. ICCSE 2018, no. Iccse, pp. 623–628, 2018, doi: 10.1109/ICCSE.2018.8468858.

[3] A.Joy* and D. R. Saranya, "A Pilot Research on Android Based Voice Recognition Application," Int. J. Recent Technol. Eng., vol. 8, no. 4, pp. 7272–7277, 2019, doi: 10.35940/ijrte.d5284.118419.

[4] S. Chanda, S. Pal, and U. Pal, "Word-wise Sinhala Tamil and English script identification using Gaussian kernel SVM," Proc. - Int. Conf. Pattern Recognit., no. February 2015, 2008, doi: 10.1109/icpr.2008.4761823.

[5] S. Gallege, "Analysis of Sinhala Using Natural Language Processing Techniques," 2010, [Online]. Available: www.defence.lk/

[6] M. Rimas, R. P. Thilakumara, and P. Koswatta, "Optical character recognition for Sinhala language," c2013 IEEE Glob. Humanit. Technol. Conf. South Asia

ISSN (online): 2581-3048
Volume 7, Issue 11, pp 127-134, November-2023
https://doi.org/10.47001/IRJIET/2023.711018

Satell. GHTC-SAS 2013, no. August, pp. 149–153, 2013, doi: 10.1109/GHTC-SAS.2013.6629906.

[7] D. L. Pennell and Y. Liu, "NORMALIZATION OF TEXT MESSAGES FOR TEXT-TO-SPEECH The University of Texas at Dallas Computer Science Department 800 West Campbell Road," 2010 IEEE Int. Conf. Acoust. Speech Signal Process., pp. 4842–4845, 2010.

[8] S. Ueno, M. Mimura, S. Sakai, and T. Kawahara, "Synthesizing waveform sequence-to-sequence to augment training data for sequence-to-sequence speech recognition," Acoust. Sci. Technol., vol. 42, no. 6, pp. 333–343, 2021, doi: 10.1250/ast.42.333.

[9] P. J. S. P. R. S. -Jasmita Khatal, "IRJET- Real Time Image Captioning and Voice Synthesis using Neural Network," Irjet, vol. 8, no. 1, pp. 415–419, 2021.

[10] H. Miyauchi, "A systematic review on inclusive education of students with visual impairment," Educ. Sci., vol. 10, no. 11, pp. 1–15, 2020, doi: 10.3390/educsci10110346.

[11] T. Dinushika, L. Kavmini, P. Abeyawardhana, U. Thayasivam, and S. Jayasena, "Speech Command Classification System for Sinhala Language based on Automatic Speech Recognition," Proc. 2019 Int. Conf. Asian Lang. Process. IALP 2019, pp. 205–210, 2019, doi: 10.1109/IALP48816.2019.9037648.

Citation of this Article:

Thirosh Madhusha, H.D.Madushi Bhagya, Sajeewa Godakanda, J.P.Dulmi Semini, Prof. Koliya Pulasinghe, P.K.P.G. Panduwawala, "AI-Based Movie Content Rating and Recommendation System" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 7, Issue 11, pp 127-134, November 2023. Article DOI https://doi.org/10.47001/IRJIET/2023.711018

****