International journal of basic and applied research
www.pragatipublication.com
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

Text to voice translator application for blind persons assistant system

Y. HARATHI, Department Of IT, SICET, Hyderabad
JETTY KUSHAL KUMAR, VAJRAPU UTHAM, BILLAKANTI NIHARIKA, VANGALA ESHWAR REDDY
UG Students, Department Of IT, SICET, Hyderabad

ABSTRACT

Viewing text-based material may be challenging for people with partial disabilities, and they need assistance from a third party. A text-to-speech (TTS) system that enables the partially sighted to hear a sound read-back of any text from an image can be created to address the issue. The programming that offers voice assistance to the partially impaired so they can access the internet is included in this audit document. Additionally, with the development of the Internet, everyone now uses a mobile phone. The structural framework for using speech as an input for mail, alerts, URLs, as well as for reading out text messages, is also introduced in this application. The STT and TTS methods are both applied using Natural Language processing and deployed using the flutter application.

KEYWORDS: *Natural language processing, Text to speech synthesis, Easyocr*

INTRODUCTION

A person is considered partially sighted if they can see partially with one or both eyes. When the better-seeing eye's visual acuity is 20/70 or worse and does not improve with the use of contacts or glasses, a severe visual impairment known as low vision arises. Partially impaired people experience more issues in daily life and are less prone to rely on others, therefore advances in technology like text-to-speech and NLP can help these people. Amazingly, voice assistance has been around for more than 100 years despite Alexa being the first system we are aware of being released only ten years prior. Artificial intelligence, or AI, is frequently used nowadays as voice partner. Text-to-speech functionality is essentially built into every modern gadget, including PCs, cellphones, and tablets. Some TTS devices also provide optical person recognition (OCR), which is a function. TTS devices can vocally read text from photos using the technology of OCR and NLP.

Creating the functionality is the first step; the tool will handle the rest. It works with web browsers, desktop operating systems (Windows, Linux, and MacOS), and mobile platforms (iOS and Android).A mobile-based application is provided by this system for those who are partially impaired. Therefore, it would be extremely helpful to develop an app that allows someone with vision impairment to send emails, set alarms, browse websites, and read text messages all with just getting voice as an input and it also includes translation,

International journal of basic and applied research
www.pragatipublication.com
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

summarization of text images. Speech-to-text and text-to-speech are used. As a result, there will be less need for the partially crippled to rely on others for their daily activities.

LITERATURE REVIEW

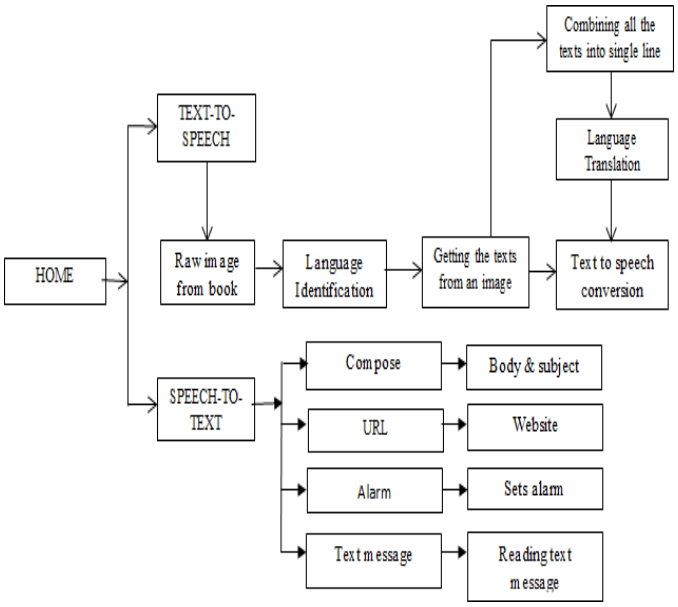
The author [1] develops a text to speech synthesizer in the form of application which converts the processed input text into speech with the Natural language processing tool and reads out to the user in the form of mp3 files using the Digital signal processing. It is based on attractive graphical user interface contains two modules, one module will be functioned based on SWT and DJNativeSwing and second module will implement the freeTTS. Jay Bagrecha et al [2], introduce a lightweight Android GUI programme that helps vision impaired users better perceive their surroundings. Some features of the application include the ability to recognise objects and distances as well as Indian currency. Using NLP and deep learning principles, they build optical character recognition to extract the text from the sheet of paper. The author [3] make review about the various techniques which comes under Speech-to-Text (STT) and Text-to-Speech conversions (TTS). They have concluded that the Hidden Markov Model (HMM) works better for the STT conversion with some computational complexity and parallel or cascade synthesis works well for the TTS conversion. Athira M Nair et al [4], proposed a Deep neural network model for handwritten recognition. The recognized feature will be printed and convert into speech for the end user. They develop the front-end GUI application with the flutter open source tool and the back end will be processed with the Convolutional Neural Network (CNN) modelling and training with text converted into speech conversion. The model is trained and tested with 4,00,000 handwritten images from the kaggle repository and attains above 90% of accuracy. The author [5] developed an application named “EasyReader” a text reader for visually challenged peoples. They desingsed a mobile application with firebase ML vision to scan the text image and it is converted into text. Tessaract OCR is used to extract the text information from the scanned image. Sakshi Bhajikhaye et al [6], proposed a standalone language translator comprises of speech recognition, language translation and speech synthesis. Speech-to-text conversion works with the microphone inputs and converted into text with the following techniques sucha as Natural language processing and Google Nueral machine translation. The TTS module contains speak encoder, synthesizer, and vocoder framework for the conversions. The final GUI applicatio will be developed with flutter applications. Partha Mukerj et al [7] developed a simple and attractive GUI appliclaion which allows user to give their own inputs in the text field. They generate a text to speech engine for the american english named TTS Gramaty. The user can browse a document from the local device and uploaded. The uploaded text will be converted as a audio file and it can be downloaded in their local folder for the future purpose. The techniques include NLP, DSP and flutter framework. TTS for Tamil using Hidden Markov Model proposed by the author [8] for the classification and recognition of Tamil letters based on phonetics available in the Indic word recognition system. Based on the complexity of language, understanding the phonetics, corpus creation and syllable level concatenation were the highest level of challenges were faced during the development of the

International journal of basic and applied research
www.pragatipublication.com
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

tool. A speech synthesis system is designed by the author [9] and implemented for Tamil language. The corpus is created from various domain words and syllabified using syllable segmentation rule. From the text present in the corpus, speech output for Tamil phrase will be produced with the help of concatenated synthetic speech techniques.

DESIGN ARCHITECTURE AND METHODOLOGY

The architecture in fig.1 includes two module one is Text-To-Speech (TTS) and another one is Speech- To-Text (STT). The TTS module provide text recognition, identification, translation, and summarization. The STT module translates the user commands to the corresponding email composition, URL redirection, setting the alarm and reading the text message received in their mobile.



A. Text recognition

OCR is an abbreviation for "Optical Character Recognition." It is a technological advancement that can recognize text in digital photographs. Using this method, text in scanned documents and photographs is regularly detected. The fact that this library will give us detected text in the form of a list is its strongest feature. The list output will make it simple to automate the text recognition operation. Calculate the text's box bounds accuracy in the image.

B. Text Detection

The process of locating text in an image and containing it in a rectangular bounding box is known as text detection. Both frequency-based and image-based techniques can be used to detect text. Images are segmented into a number of segments using image-based techniques. The pixels that make up each segment

International journal of basic and applied research
www.pragatipublication.com
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

are related and share similar properties. The text is categorised and shaped using the statistical characteristics of connected components. According to the requirements of a machine learning purpose, a bounding box is a fictitious rectangle that is used to outline the item in a box. These are the model's primary results for object detection. the fictitious rectangle border that encloses an object in a picture.

C. Text Translation

Various websites that cater to an international audience frequently translate their text from one language into another. This is made possible using the translate Python package. We sometimes find ourselves in situations where we have to translate a significant section of a text or when we are unable to understand it. While there are online resources for translation, they are not always available and it may be preferable to give our clients a static translated page. So this translation can help partially sighted people very well

TTS AND STT

A class of technology known as Assistive technology, which is also referred as assistive, adaptive and rehabilitative devices for people with disabilities were designed to read digital text aloud. Text-to-speech (TTS) conversion is a sophisticated feature found in smart devices including text scanners, ATMs, and online translators. Due to its appropriate accessibility, text-to-speech technology use in the programme improves the user experience. Making books audio is now a common practice. Even a well-known website called Audible uses TTS technology to provide hundreds of books in audio format. Many APIs in python allow you to convert text into speech. Among them Google Text to Speech API or gTTS API is one of these . Typed text can be converted into audio and saved as an mp3 file with the help of an easy-to-use program called gTTS. May languages are supported by the gTTS API, such as English, Hindi, Tamil, French and many more. Any one of the two accessible audio speeds, rapid or slow, may be used to deliver the speech. The voice of the created audio, however, cannot be changed as of the most recent version. **Pytsx,** An open- source text-to-speech wrapper is called Pytsx.In accordance with your operating system, it employs various speech engines. **NSSpeechSynthesizer** It's for Mac OS X 10.5 and later (nsss)Windows XP, Windows Vista, and (untested) Windows 7 all support SAPI5. espeak - eSpeak on any operating system or platform (like Ubuntu or Fedora Linux) that can host the shared library. sapi5- An open-source text-to-speech wrapper is called Pytsx. Software that converts speech to text listens to audio and outputs an editable, verbatim transcript on a specific device. Through voice recognition, the software accomplishes this.

ANALYSIS OF RESULT AND DEPLOYMENT

Text-to-speech Module : The image or pdf file of the Tamil book page of fig.2 is uploaded with the read command available in python. Text recognition is accomplished through the use of an optical character recognition (OCR) technique that converts document photographs into text format by extracting text. This type of information retrieval is known as recognition-based retrieval, allowing for more searching.

International journal of basic and applied research
www.pragatipublication.com
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

2 அறத்துப்பால்

1. கடவுள் வாழ்த்து

1. அகர முதல எழுத்தெல்லாம்; ஆதி
பகவன் முதற்றே உலகு.

2. கற்றதனால் ஆய பயனென்கொல் வாலறிவன்
நற்றாள் தொழாஅர் எனின்.

3. மலர்மிசை ஏகினான் மாணடி சேர்ந்தார்
நிலமிசை நீடுவாழ் வார்.

4. வேண்டுதல்வேண் டாமை இலான்அடி சேர்ந்தார்க்கு
யாண்டும் இடும்பை இல.

5. இருள்சேர் இருவினையும் சேரா இறைவன்
பொருள்சேர் புகழ்புரிந்தார் மாட்டு.

6. பொறிவாயில் ஐந்தவித்தான் பொய்தீர் ஒழுக்க
நெறிநின்றார் நீடுவாழ் வார்.

7. தனக்குவமை இல்லாதான் தாள்சேர்ந்தார்க்கு அல்லால்
மனக்கவலை மாற்றல் அரிது.

8. அறஆழி அந்தணன் தாள்சேர்ந்தார்க்கு அல்லால்
பிறஆழி நீந்தல் அரிது.

9. கோளில் பொறியின் குணமிலவே எண்குணத்தான்
தாளை வணங்காத் தலை.

10. பிறவிப் பெருங்கடல் நீந்துவர்; நீந்தார்
இறைவன் அடிசேரா தார்.

Figure 2 TEXT IMAGE

Machine learning models based on AI are created using the image annotation method known as bounding boxes. Several uses for them include robots, drones, autonomous vehicles, security cameras, and other machine vision systems, it is used for target recognition and object detection. Using the NLTK tool, the detected Tamil words are tokenized with the corresponding index values and a corpus is created with the individual unique values as show in the fig.3 . With the Tamil word corpus, the summarization and translation can be done based on the selected target language.

International journal of basic and applied research
www.pragatipublication.com
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

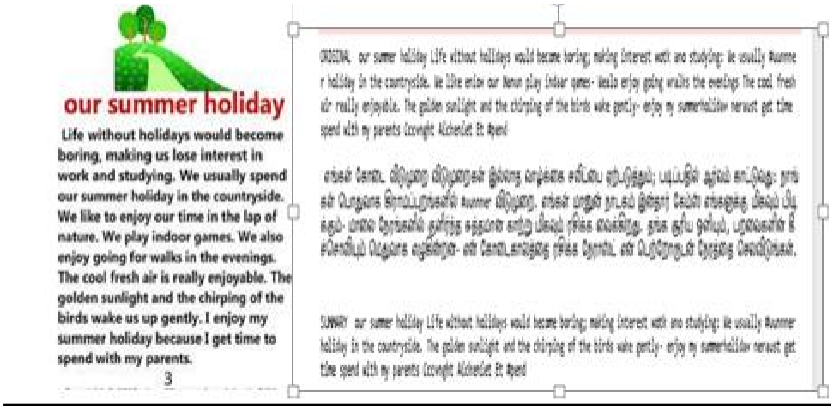


Figure 4 LANGUAGE TRANSLATION

The flutter framework with dart language is used to develop the user interface in the way of "Look and Feel" structure. If the original text is in English language, after using translation package in easyocr the english text in will be translated in Tamil language. Summarization can also be done in original text which can be converted into speech. The fig.4 & 5 shows the user can upload text file to get an output as audio format, translation and summarization. The playsound() function of concatenative speech synthesizer which combine the words available in the corpus in to mp3 or wav file and it can be downloadable format for future use.

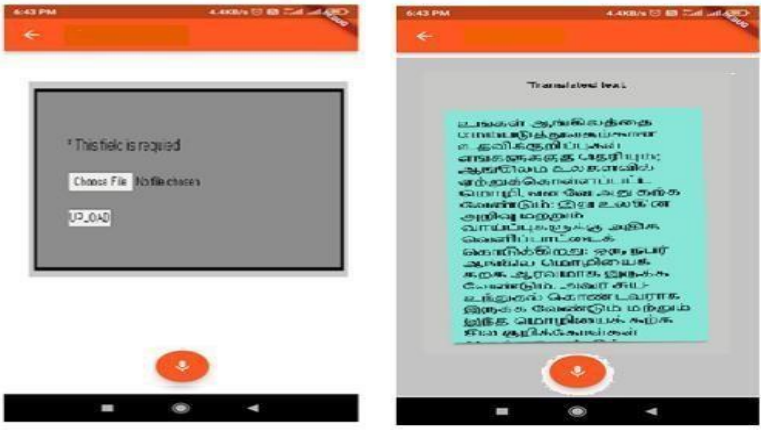


Figure 5 GUI APPLICATION

SPEECH-TO-TEXT MODULE

Input speech command will be processed to Gmail which were already logged int the mobile application, it transfers from send mail page to compose mail. In Compose mail Module the user gives the input as speech for body of the message. In URL module, after getting the Website URL as an speech input the app moves to the corresponding web page shown in fig ().

International journal of basic and applied research
www.pragatipublication.com
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

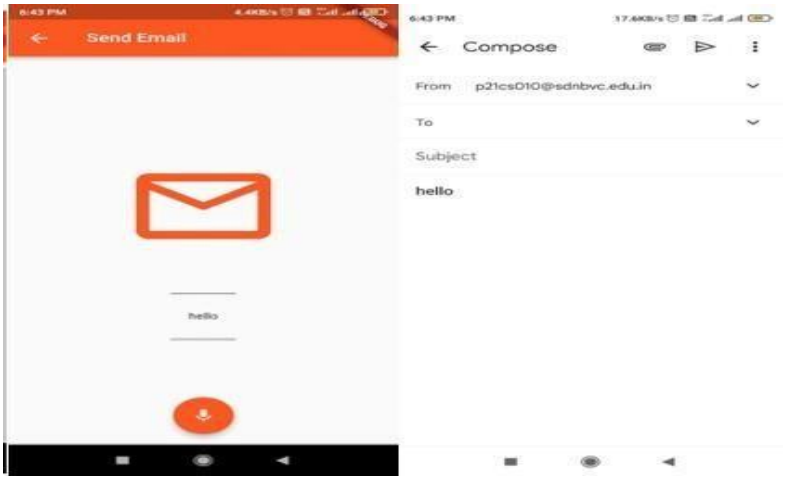
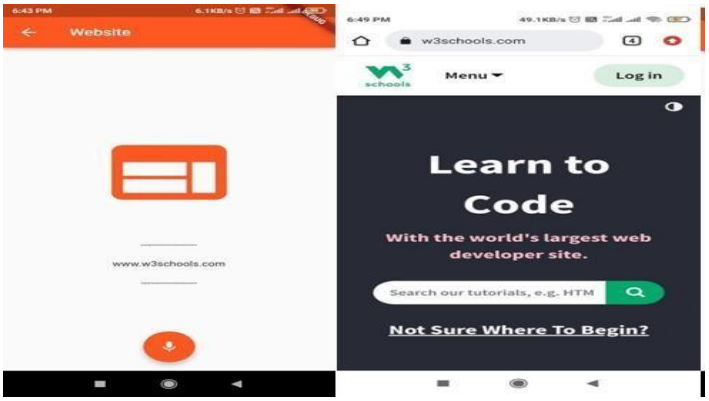
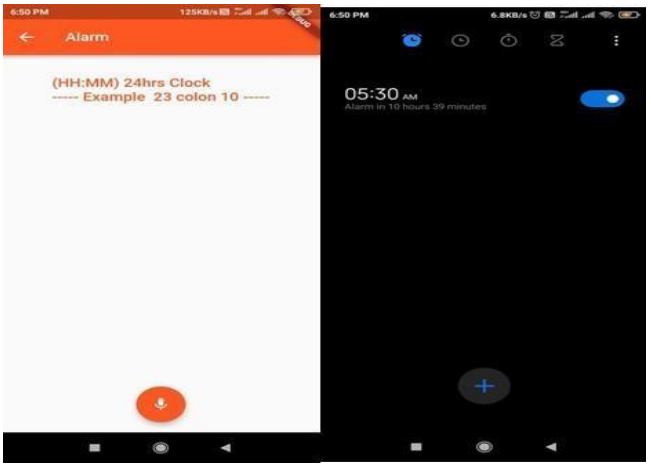


Figure 6 EMAIL ASSISTANCE



ALARM: It gives voice as an input. The user can give the Speech input as 24-hrs timing to set an alarm with colon(:)Format, It transfers to the page where the alarm can set. The following timing will be set as alarm app.



Index in Cosmos
March 2024, Volume 14, ISSUE 1
UGC Approved Journal

International journal of basic and applied research
www.pragatipublication.com
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

TEXT MESSAGE:

It gets the text messages from your mobile into this app and reads out the messages. And also it separates the sent messages and received messages. Following the user's voice input of a message, the application receives and incorporates the text messages from the mobile device. There it reads the messages to the user.

CONCLUSION

TTS combination remains quickly developing part of PC innovation and is progressively assuming more significant part in the manner of communication with framework and connection points over various stages. In future many new designs are implemented for an attempt to the area of execution of text-to- discourse framework on different stages. For example, communication framework ATMs, computer games whatever other stages where TTS innovation could be an additional benefit and increment usefulness.The application has modules for object detection, object distance from the user, object captioning, Indian bank note deduction, which warns the user if the note is a fake, and optical character recognition, which enables the user to quickly and easily read whatever is written on a piece of paper, in an image, or anywhere else around them.

REFERENCES

1. Itunuoluwa Isewon, Jelili Oyelade, Olufunke Oladipupe, " Design and Implementation of Text to Speech Conversion for visually impaired people", International Journal of Applied Information Systems, pp. 25-30, vol.7(2), APril 2014.
2. Jay Bagreeha, Tanay Shah, Karan shah, Tanvi Gandhi, Sushila Phalwe, "Virtual Eye: Android Application for Visually Impaired", Recent Trends in Intensive computing, doi:10.3233/APC210204, 2021.
3. Ayushi Trivedi, Navya Pant, Pinal Shah, Simran Sonik, Supriya Agrawal, "Speech to Text and Text to Speech recognition system -A review", IOSR Journal of Computer Engineering (IOSR-JCE), pp. 36-43, vol.20(2), Apr 2018.
4. Athira M Nair, Chrissie Aldo, Blessil Bose, Alex Joseph, Praseetha V.M, Elizabeth M J, " HandWritten Character Recognition using Deep Learning in Android Phones", International Research journal of Engineering and Technology (IRJET), pp. 61-66, vol.8(7), July 2021.
5. Amirul Abdullah, Mohamed Amir Fahmi Mohd Rosli, Norsaremah Salleh, "EasyReader: Text Reader for Visually Impaired People", International Journal on Perceptive and Cognitive Computing (IJPCC), Vol.8(1), 2022.
6. Sakshi Bhajikhaye, Dr.Sonali Ridhorkar, Vidhi Gautam, Mamta Soni, Mayank Badole, Adrash Kant, Pranjali Rewatkar, " Language Translation by Stand-Alone Voice Cloning: Amultipseaker Text-to-Speech Synthesis Approach based on Transfer Learning.

7. Partha Mukherjee, Soumen Santra, Subhajit Bhowmick, Ananya Paul, Pubali Chatterjee, Arpan Deyasi, "Development of GUI for Text-to-Speech Recognition, Using Natural Language Processing", International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech). IEEE, 2018.
8. A.Femina Jalin, J.Jayakumari, " Text to Speech Synthesis System for Tamil using HMM", Proceedings of 2017 IEEE International Conference on Circuits and Systems (ICCS2017).
9. J.Sangeetha, S.Jothilakshmi, S.Sindhuja, V.Ramalingam, " Text to Speech Synthesis system for Tamil",International Journal of Emerging Technology and Advanced Engineering, pp.170-75, Vol.3(1), Jan 2013.