

# Voice Automation Mail System for Visually Impaired

Malathi D

Department of Computing Technologies  
SRM Institute of Science and Technology  
Kattankulathur – 603203

India  
malathid@srmist.edu.in

Dev Awasthi

Department of Computing Technologies  
SRM Institute of Science and Technology  
Kattankulathur – 603203

India

Gopika S

Department of Computer Science,  
Kristu Jayanti College,  
Bangalore

Karnataka - 560077  
[gopikas@krstujayanti.com](mailto:gopikas@krstujayanti.com)

Dorath Jayaseel

Department of Computing Technologies  
SRM Institute of Science and Technology  
Kattankulathur – 603203

India

**Abstract**— The internet has emerged as one of the most crucial elements of modern life. Every single person uses the internet to access knowledge, information, and all communication tools available to them. However, many who are visually impaired find it difficult to use those features and need outside aid to complete their tasks. People with visual impairments all around the world now have a wide range of new opportunities because of the invention of computers. Screen readers, audio-based environments, and other assistive features have made it easier for blind persons to utilize the workspace. Today, email is required to send confidential information. Email is a type of technology that facilitates business correspondence and lets users transmit messages to other people. The main objective of this work is to develop a voice-based email system that will enable people who are blind or visually impaired to send and receive emails using computers. It will make advantage of modern features to create a working environment that enables persons with visual impairments to do their jobs independently.

**Keywords**— Audio Based environment, Text Recognition and Extraction, Text to Speech Synthesizer, Voice Based Technology.

## I. INTRODUCTION

One of the desirable or useful things for daily living is now the Internet. It has improved people's lives by increasing access to information, facilitating interpersonal relationships, and expanding businesses and associations. The first luxury of a 24-hour existence is now the Internet and anyone accesses or uses online data and information. When people utilize the Internet for communication, their lives become easier. Internet has greatly changed the ways in which people communicate. In 2021, there were four billion email users worldwide; by 2025, that number is projected to increase to 4.6 billion. Globally, 306 billion emails were sent and received every day in 2021. Electronic correspondence communication has shown to be the most secure and safest technique for exchanging and entering sensitive or private

information. The requirement of being able to see is a necessity that must be met in order for a person to use the Internet. Internet browsing is a nightmare for persons who are blind because it relies on visual perception. According to a survey, more than 250 million people are unable to communicate through email or Internet infrastructure. An eyeless person can only make use of all these online capabilities if they narrate the entire content to a third party who is not blind. The visually challenged person can then ask for the third party to send, receive, and read out the communications. A variety of dependable and effective technologies have made it possible for the blind to access the Internet. Hence, in this research work, in order to promote society and give such especially suitable persons an equal status and respect, a model is designed for a voice email system that is email-friendly and can be used by visually impaired. This enables voice instructions to be used by blind individuals to send and receive messages. For this system, additional visual aid or a keyboard is not required. This proposed model is based on the practice of using Speech-To-Text (STT) and Text-To-Speech (TTS) transformers, to convert spoken words into written and audible forms respectively. The individual using this method would not be familiar with the layout of the keyboard, the positioning of the letters, or the locations of the keys. The user will answer once the system automatically feeds voice commands for carrying out specific actions. The user only needs to correctly carry out the system's instructions. The recommended technology and voice commands could enable visually challenged people to view emails independently.

## II. RESEARCH CHALLENGES

For those who have the gift of sight, emailing does not provide a significant issue, but for those who do not, because it interferes with so many professional duties, it does. Only people without eyes prefer emails with audio attachments. The oral commands are simple for them to respond. Therefore, there are fewer chances of making this audio-based email accessible to blind people. This voice-based email system is useful for persons without vision because it allows them to comprehend the location. On the other hand, users must keep in mind keyboard shortcuts and

mouse clicks. This approach places more emphasis on user friendliness for all users, including those who are generally blind. This system can only be used on desktop PCs, and sometimes it has trouble correctly decoding the data. New functionality could be added to the system's user interface in the future.

### III. OBJECTIVES

The main objective of this research work is to develop a voice-based emailing system that will make it possible for those who are blind or illiterate to use commonplace technology like sending and receiving emails and using the Internet. This technique makes it simple for blind users to sign in by speaking. The main objective of creating this kind of system described in the study is to give persons who are blind or visually impaired a stronger sense of community. Our email system is easily accessible to users of any age. Future uses for this technology could include in addition to email, enhanced and utilized for further services like messaging and using voice commands to operate other programmes.

### IV. SCOPE OF THE PROJECT

The creation of computer-based accessible solutions has opened up several new options for blind and visually impaired people anywhere in the world. However, visually impaired people have a very difficult time using this equipment because using it requires visual sight. The proposed system corrects the flaws in the existing one. The suggested method is voice command anchored, in contrast to the current systems. The main consideration that has been made is whether the suggested system is available. Because the system in place today and the one we are developing are fundamentally dissimilar.

The requested task is implemented using Python. Ideas focuses on offering fundamental features including voice-based guidance and the ability to compose, read, transfer, and receive emails. Unlike the current approach, which prioritizes compatibility and user usability. With this method, the user is prompted to give the system specific instructions before reading the subject and content of a given email ID, such as "compose the mail," "show the email id of the previous three unread messages," and "choose one mail id." There is direct user-system communication. While the existing system focuses more on normal people, our approach is beneficial to all types of people, whether they are normal, visually impaired, or illiterate.

### V. MODULES

#### A. The STT (Speech-To-Text):

This module interprets speech that has been detected with the aid of using a microphone into textual content and stores the written layout textual content in a variable. If not, it raises an exception.

#### B. The TTS (Text To Speech):

This module converts the given text to the audio, in .mp3 file format and plays the file created.

#### C. The Python's GTTS (Google Text To Speech) library

This is cross-platform, so it works with the speech libraries on Windows, Linux, and MacOS. This module accepts text-based input together with language and speed parameters and automatically creates an audio file. Additionally, read out emails' messages to the user.

D. *SMTP* : The "smtplib" keyword is used to import the SMTP (Simple Mail Transfer Protocol) module, required to connect to Gmail servers. This protocol is required for composing emails and connecting to Gmail, and helps you perform other tasks.

### VI. RELATED WORK

#### A. Voice Based Technology for Blind

Management of speech automation systems [1] use three different forms of technology, including Text-to-Speech (TTS) and Speech-to-Text (STT), which convert speech to text, and Interactive Voice (IVR). We built an application using response. IVR is a term that explains how a user can communicate with technology using a keyboard or voice message. The user can also communicate with the mail system. The use of highly sensitive microphones, which are typically not accessible to all users, is the biggest drawback.

This research proposes a voice-based e-mail system that allows visually impaired people to easily access their email [19]. Jan.V. et al. [2] makes it easier for blind people to send and receive voicemail. The focus of this research is improving text-to-speech email delivery for blind or partially sighted people. This study provides text-to-speech and speech-to-speech e-mail access technologies for the visually impaired. This makes it possible for blind people to send mail using voice control rather than a keyboard.

A voice-based system with multilingual capabilities was proposed by Diveshh Jimmy in [3]. This system offers a beautiful GUI for all user types. Figure 2 shows the GUI. User has access to the e-mail system and can send, receive, read and delete e-mail. The main drawback, however, is the need to click the mouse in certain areas of the proposed program.

To be able to read and record symbolic linguistic expressions such as phonetic transcriptions, the author Britha presented a system with TTS and STT [4]. The system design consists of two modules: *Choice of interfaces and mailing options*. The first module selects whether the user is blind or sighted. Additionally, the second module provides a simple mailing alternative to complete all jobs. Using additional voice recognition technologies in addition to all other converters and IVR, Parkh Bhardwaj et al. [5] have designed a system which offers more functions than the current GUI does. The primary programming language utilized was Java. All forms of handicapped individuals, even those who are illiterate, can utilize the programme.

### B. Inferences from survey

A literature review examined various techniques for creating voice-based e-mail systems[6]. Common systems, including screen readers and other tools, also have some drawbacks. The work proposed in this model takes full advantage of technology by using features that have the greatest accuracy in audio capture and content presentation required by visually impaired people.

## VII. PROPOSED SYSTEM ARCHITECTURE

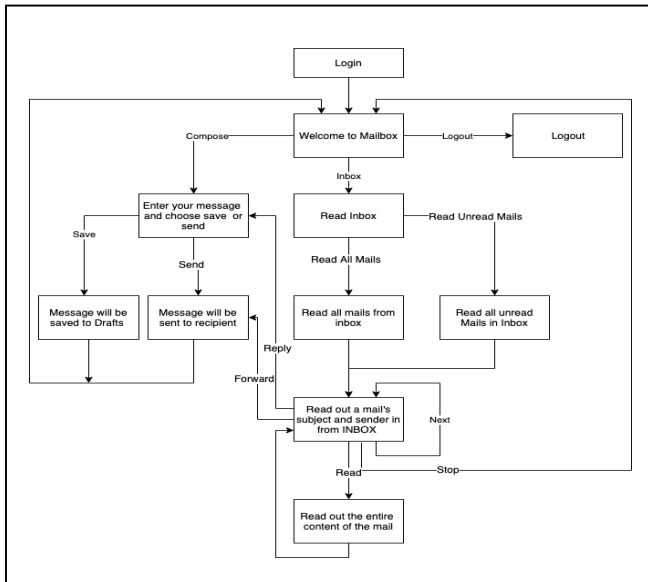


Fig. 1. Architecture for Voice based email system for visually impaired

### A. User Login

The login module in the proposed model is entirely voice-based, requiring the user to enter an email address and password. After the user enters the email address and password, the login system notifies them and asks if they are correct. If they say yes, they are allowed to proceed with the operation. If the user replies that something is wrong, the login system prompts them again for the email address and password.

Implemented in the Python language, it recognizes speech using a speech recognition engine and text-to-speech library functions[7] that help implement text-to-speech code. The login page frontend was developed using the tkinter module of the Python language.

### B. Option Choosing

After successfully logging in, the email system prompts the user with options to compose a message or check the user's mailbox. The user can press "Enter" to start composing the email. Use the spacebar to indicate that mail is checked in your inbox.

This module is built using a menu-driven Python application. The software prompts the user to press "Enter" to compose a message or "Space" to check the inbox of the user's logged-in email address. To start composing an email

when the user presses the "Enter" key, call `compose_mail()`. When the user presses the "space" key, call `check_inbox()` to check the user's logged in inbox.

### C. Checking the Inbox

The user can view their logged-in mail's inbox by hitting the "Space Bar" while selecting an option. The user is initially prompted to identify the precise user whose communication has to be read. The machine then reads the message's subject and content. The Google Text-to-Speech (GTTS) reading method is employed. The user's vocal instructions were used to obtain each of these inputs. The GUI part of the front end of this page is implemented with "tkinter".

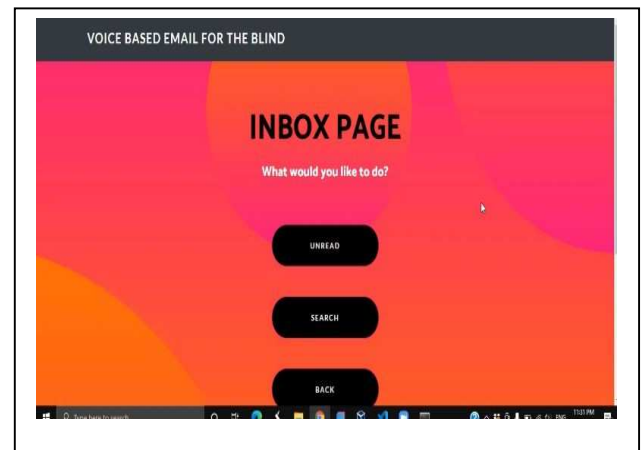


Fig 2: GUI Design

Users can view their subscribed email inbox by pressing the "spacebar" while selecting an option. The email IDs of users' first 3 unread messages are first stored in a dictionary key. The subject and content are then pulled into a Python list and mapped as values to the corresponding key values in the dictionary. An email ID from the list of saved email IDs must be entered by the user. A list containing the subject and content of received messages is read after the user enters the selected email ID value into the dictionary. "Tkinter" is used for front-end development of this site.

### A. Text to Speech Approach(gTTS)

In this implementation, this module serves multiple areas by converting the text entered to an .mp3 file, as in Fig 4, and then plays the created .mp3 file.

This function is implemented by importing the gTTS (Google text-to-speech) module.

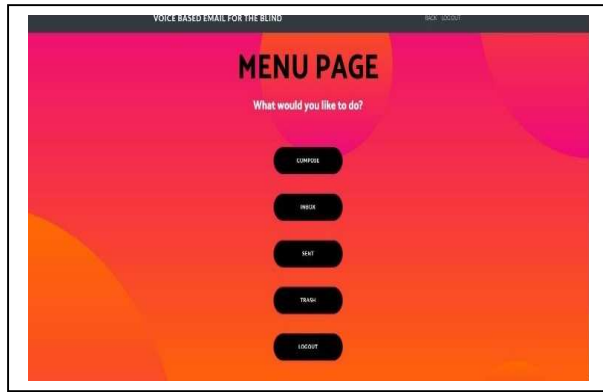


Fig 3 –menupage

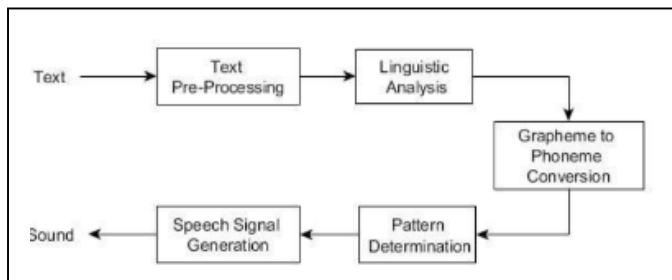


Fig 4: Various stages in text to speech Conversion

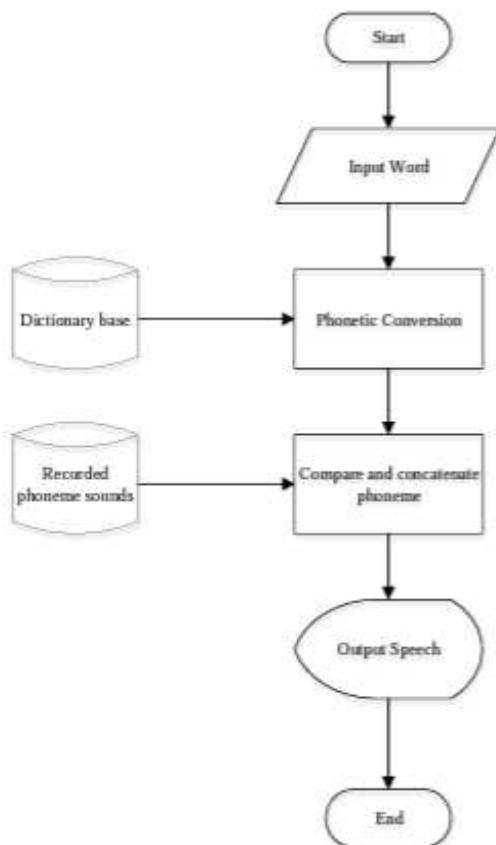


Fig 5 : Flow chart on Text to Speech Conversion

#### A. Speech to Text Approach

The method of turning vocal speech signals into text using transcripts is known as speech recognition. In the current era of the computer revolution, the speech recognition plays a significant part in improving the user experience while interacting with the computers in a natural way. This sub module is used in all other modules. This function converts the speech recognized through system microphone to the text and stores it in the variable. Our implementation raises an exception if the entered audio data is unrecognizable[9].

The database containing voice samples has been trained and tested using Convolutional Neural Networks (CNN) as shown in Fig 5, A thorough training of the system yields higher testing performance, which is the key characteristic of weight connectivity, local connectivity, and pooling results[10]. Wideband speech signals are used in experiments, which significantly improves the system's performance over conventional methods.

### VIII. RESULTS AND DISCUSSION

#### A. Libraries used by the system

The speech recognition package from the Python library is used by the proposed system and is installed using the pip command. Speech recognition is a crucial element used in many artificial intelligence applications and home applications. Another crucial programme is Yagma, which is used to send email from the sender's address to the recipient's mail id. The Google-text-to-speech gTTS module is used to convert text to speech. The imaplib client side library of Python is used to access emails through the IMAP protocol. It enables the client application to modify the email message. Tkinter and Pyglet, which offer utilities for designing interfaces for multimedia applications, are imported to create a nice interactive user interface

#### B. Speech to Text in Proposed approach

Speech recognition uses linguistics and related technology to recognise the user's spoken words and translate them into text. It aids in the computer's comprehension of human language. Python's voice recognizer is utilized to respond to them with the necessary functions. Fig 6 illustrates the suggested system makes use of this capability to receive voice instructions from visually impaired people in order to collect inputs, checks the condition specified in the code, and performs the necessary function as a result[15].

The external microphone is used to input speech. Voice input can be collected through blue tooth devices and it is also utilized to collect feedback from users because the external microphone is set up for that purpose. The software is permitted to change one second in accordance with the threshold value of the external noise adjustments since the user's immediate environment may be noisy[11,12].

Google Speech Recognition is used to translate speech to text. However, some offline recognition systems involve complex installation procedures and a large number of dependencies[9]. One quick and effective method for understanding user commands and correctly translating them to text is Google Speech to Text[13].

The user's voice commands are entered through the microphone in the system. The user's voice commands are collected using the listening command, which accepts input from a specific source, which is the system's microphone[14]. Convert audio to text by recognizing Google commands to convert audio files to text. These functions can be used with the "reco" function of Python's speech recognition engine.

### C. Text to Speech in Proposed approach

The suggested model transforms text written in human language into human-like speech. For the same purpose, Python offers a variety of application interface. One of the most effective methods is Google Text to Speech, and this paper takes advantage of this API. The program helps the user with an efficient effort less and time advantageous solution to use and it has a lot of built-in features that are used to store text[8].

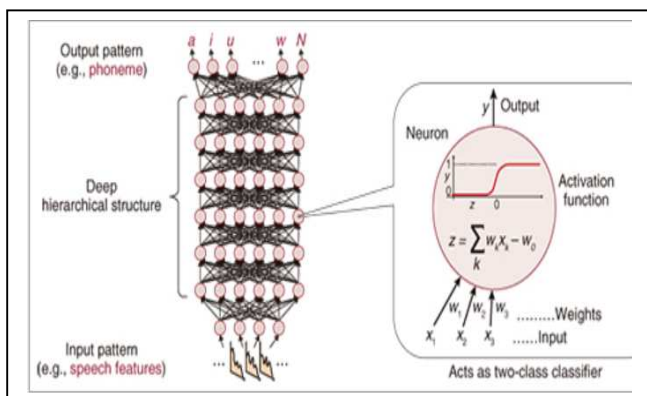


Fig 6 : Speech to Text Conversion

The convolutional neural network matches parts of the signal rather than the entire signal of pixels because it is difficult for a computer to identify the signal when the entire set of pixels is considered. Filtering is the mathematics in matching these. [10] [11]

The language of the text is supplied as one of the function parameters in the gtts() function, and in our system, the text is converted into an mp3 file. Direct saving of the mp3 file occurs in the same directory as the project file.

The user's own PCs' speakers are used to play the saved mp3 file after it has been loaded using the Pyglet module. The voice speed can be set using the time package in the Python library, and the system can also wait a while after conversion. Additionally, the remove

command, which is included in the OS package in the Python library, can subsequently be used to remove the saved MP3 file from the directory[12].

### D. Mail Box Filtering

The first three unread messages in the user's inbox are read by the proposed algorithm. A data structure called a dictionary provided by the Python library implements the filtering process as an additional feature.[17] The content of the message is defined as the key of a specific key in the dictionary, which holds the email ID of the message as the key. When the user selects an email ID, the specific message is read out from the dictionary. This system has an innovative feature that allows users to easily select the messages they want without having to listen to all the messages in the mailbox. This incredibly convenient and efficient feature greatly simplifies the work process for the visually impaired.

### E. To send and receive Mails

The imaplib library is used for sending and receiving mail. Python built-in SMTP library used for sending email. Email packages serve HTML content and attachments using available email packages. Additionally, the system uses the Yagmail package to send emails about specific users with Gmail accounts.

IMAP is an e-mail retrieval protocol that saves disk space by reading and displaying e-mail instead of downloading it to your system. It also provides various other functions such as viewing, replying, deleting, etc. I used the read function in my module. The IMAP login feature allows users to specify an email username and password. You can use the get\_payload function to get the message out of the HTML junk. This happens even if there are no values left in the inbox.

### F. Graphical User Interface

Python offers a variety of ways to develop graphical user interfaces. Tkinter is one of the interfaces for the tk GUI toolkit that comes with Python. It provides a quick and easy way to create GUI applications. Before creating a GUI application with a main window, we need to import the Tkinter module. Additionally, you can develop widgets that perform the functions you want.

The suggested employs straightforward methods to create the interface because sending mail via the visual interface doesn't require graphical content. Consequently, the straightforward tool was put into use to create a setting where visually challenged people could send and receive letters.



### G. Analyzing performance and results

When compared to certain pre-existing models[18], the suggested approach for a voice-based email system has very few drawbacks. The usage of relatively few keyboard clicks is the only drawback. With this interface, only two moves are required to operate. However, the system has a lot of benefits, including the ability for users to filter the mail's contents and select the mail ID from which to view them. It works well with the standard microphone found in personal computers.

### IX. CONCLUSION AND FUTURE SCOPE

This application was developed for the visually impaired to help them use the free-mail service. Screen readers have been removed. A major goal in creating the system described in this study is to increase a sense of community among visually impaired people in confined spaces. Each operation has a unique outcome, so the decision tree follows a specific path. This greatly improves system compatibility. With the help of this system, disabled people will be able to interact with the outside world. Anyone can easily use the mail system regardless of age. There are audio-to-content capabilities and content-to-audio capabilities[16]. Blind people can now easily send and receive emails using only voice commands, with little or no keyboard or mouse. It helped remove the difficulties faced by blind people and make them more sociable people. A screen reader could be used in addition to console shortcuts to reduce the cognitive load associated with remembering console shortcuts, but that idea was abandoned.

In future, this model can be extended for functionality other than email by adding chat, Notes, and other voice-activated programs. In future work, this model could be extended for use with a variety of non-email services, such as messaging. Other features found in standard email, such as image attachment and indentation and font selection, may be added to the language features in the future.

### REFERENCES

- [1] Guillermo Arturo Hernández Tapia, Ana Lilia Reyes Herrera. "Email management system for blind people in Spanish language International Conference on Human Computer Interaction Cancun Mexico September, 2020.
- [2] Jagtap Nilesh, Pawan Ala, Chavhan Swapnil, and Bendre M.R. "Voice Based System in Desktop and Mobile Devices for Blind People". In International Journal of Emerging Technology and Advanced Engineering (IJETA), 2018.
- [3] Payal Dudhbale J. S.Wankhade, P. S. Narawade. "Voice- Based System in Desktop and Mobile Devices for Blind People", 2020.
- [4] In International Journal of Scientific Research in Science and Technology, 2018.
- [5] Ruchikshedekar, Sonu Gupta, 2019, Voice based email System for Blinds, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY(IJERT), Issue 10 (October 2019).
- [6] G. Shoba, G. Anusha, V.Jeevitha, R.Shanmathi. "AN Interactive Email for Visually Impaired". In International Journal of Advanced Research in Computer and Communication Engineering (IJARCEE), 2019
- [7] Itunuoluwa Isewon, Jelil Oyelade and Olufunke Oladipupo. "Design and Implementation of Text To Speech Conversion for Visually Impaired People", International Journal of Applied Information Systems (IJASIS), Volume 7- No. 2, April 2014.
- [8] S. Venkateswarlu, D. B. K. Kamesh, J. K. R. Sastry and Radhika Ranjan. "Text to Speech Conversion", Indian Journal of Science and Technology, DOI: 10.17485/jst/2016/v9i38/102967, Year: 2016, Volume: 9, Issue: 38, Pages: 1-3
- [9] Teena Varma, Stephen S Madar, Lenila L Monthero, and Rachna S Poojary. "Text Extraction From Image and Text to Speech Conversion", International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181, 2021
- [10] Itunuoluwa Isewon, O. J. Oyelade, O. O. Oladipupo, "Design and Implementation of Text To Speech Conversion for Visually Impaired People" International Journal of Applied Information Systems (IJASIS) – ISSN : 2249-0868
- [11] Teena Varma, Stephen S Madar, Lenila L Monthero, "Text Extraction From Image and Text to Speech Conversion", International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181.
- [12] Hay Mar Htun, Theing Zin, Hla Myo Tun, "Text To Speech Conversion Using Different Speech Synthesis", International Journal Of Scientific & Technology Research Volume 4, Issue 07, July 2015
- [13] Nwakanma, Cosmas & Oluigbo, Ikenna & Izunna, Okpala. (2014). "Text – To – Speech Synthesis (TTS)". project 2012, department of Information Management Technology, Federal University of Technology Owerri. 2. 154-163
- [14] Benkerzaz, Salha & Elmri, Youssef & Denna, Abdeslem. (2019). A Study on Automatic Speech Recognition. 10. 77-85. 10.6025/jitr/2019/10/3/77-85.
- [15] Venkateswarlu, S. & Duvvuri, B K Kamesh & Jammalamadaka, Sastry & Ranjan. (2016). Text to Speech Conversion. Indian Journal of Science and Technology. 9. 10.17485/jst/2016/v9i38/102967.
- [16] Onaolapo, J.O. & Idachaba, Francis & Badejo, Joseph & Odu, Tawalade & Adu, O.I.. (2014). "A Simplified Overview of Text-To-Speech Synthesis". Engineering and Computer Science. 1. 582-584.
- [17] Benkerzaz, Salha & Elmri, Youssef & Denna, Abdeslem. (2019). "A Study on Automatic Speech Recognition", 10. 77-85. 10.6025/jitr/2019/10/3/77-85.
- [18] Ankur Lundia, Sounak Bhattacharya, D. Malathi, J. D. Dorathi Jayaseel (2020). Development of Hand Sign Recognition Model for Differently-abled Persons. International Journal of Advanced Science and Technology, 29(06), 2368 - 2373
- [19] Gopika S, D Malathi (2020). Implementation of a Novel Blind Quality Evaluation Algorithm for Natural Scene Statistics Images. International Journal of Advanced Science and Technology, 29(06), 2777 - 2787