International Journal of Aquatic Science
ISSN: 2008-8019
Vol 12, Issue 02, 2021

Mobile application development aimed at assisting automated services to the visually impaired people using IOT

Dr.S.Hemalatha[1], Mrs.M.VIDHYA[2], Mrs. E. Archana[3], Mrs. M. Abirami[4], Mrs. A. Porselvi[5], Mrs. S. Annie Shery[6]

[6]Professo, Department of Computer Science and Engineering, Panimalar Institute of Technology, Chennai - Tamilnadu
[1,2,3,4,5]Assistant Professor, Department of Computer Science and Engineering, Panimalar Institute of Technology, Chennai - Tamilnadu

Email: [1]pithemalatha@gmail.com, [2]Mecse.vidhya@gmail.com

Abstract- The Internet of Things (IOT) is a concept that includes linking physical things to the internet in order to create smart systems or applications. The IoT paradigm implies a large number of devices connected via a traditional intent network. Because these devices typically have limited capabilities, shifting a portion of the service execution to a cloud infrastructure is a viable option. In this project, we suggest that human voice be used as a potential interface for one or more devices in the IoT ecosystem, allowing instructions to be issued and information to be received. Web crawling is also used to improve the system by incorporating trending data. An automated ticketing system will be created as a proof of concept. The suggested system is conceptualised as an Android application in which the user issues voice instructions. The voice is recorded and transferred to the STT cloud service using the HTTP POST mechanism. It receives a JSON object with many alternative transcripts of the studied speech sequence as a response. The Google STT service's automatic speech recognition (ASR) for giving voice commands is then used to match transcript alternatives against a set of supported instructions. When the ticketing procedure is finished, the cloud sends a ticket verification request to the IOT board mounted on the vehicle.

Keywords*: Visually impaired people, Mobile platforms, Automatic Speech Recognition(ASR),Cloud platforms, Internet of Things*.

1. INTRODUCTION

Reading, driving, and walking are just a few of the challenges that visually impaired persons experience in their daily lives. i)Loss of visual acuity ii)Loss of visual field- incapacity of a person to see as wide as a normal person iii)Photophobia-inability of a person to stare at light iv)Diplobia-double vision v)Visual distortion vi)Visual perceptual difficulties. The term "visual acuity" refers to the sharpness of one's vision. The sharpness of the retinal focus within the eye, the health and functioning of the retina, and the sensitivity of the brain's interpretative faculty are all elements that affect visual acuity.
Visual impairment occurs when the eyes are damaged or when the brain fails to receive and

interpret the visual cues supplied by the eyes. Diabetic retinopathy, age-related macular degeneration, cataract development, and increased pressure within the eyes, which causes glaucoma, are all common causes of visual impairment. Visual impairment comes in a variety of shapes and sizes, with varied degrees of severity.
Even in familiar surroundings, blind people have difficulty self-navigating. The inability of blind individuals to use technology, such as reading or looking for information on the internet, is a major issue. We can utilize screen readers to solve this difficulty, however it takes longer to read the material.
A screen reader converts the text on the computer into a tactile or auditory format that visually impaired persons may understand. The majority of screen readers do not assist blind persons in navigating between online pages, but with the inclusion of some additional functions, the user can work freely on a computer. Some screen readers use a synthetic voice to read out loud the words on the screen, while others use a braile display. These screen readers make use of crystals that grow when exposed to specific voltage levels, allowing blind persons to read text with their fingers. Hardware for reading from a computer screen is quite pricey.
We will use speech recognition to book tickets in this paper. The method of developing technologies for converting speech to text is known as speech recognition. Approaches like the Hidden Markov model and artificial neural networks dominated speech recognition. Many components of voice recognition are now handled by a deep learning technique known as long short-term memory (LSTM). The identity activation function for the memory cell allows LSTM to store values for both long and short periods of time.. The LSTM algorithm is used to determine the time series, time lag, and length of a significant event.

2. LITERATURE SURVEY

[1] Karl Pichotta and Raymond J. Mooney show how to use LSTM Recurrent Neural Networks to learn statistical scripts. Scripts are used to store knowledge of archetypal event sequences in this paper. It explains a Recurrent Neural Network model for statistical script learning with Long Short-Term Memory, an architecture that has been shown to operate effectively in a variety of Artificial Intelligence applications**.** They put the system to the test on two tasks: inferring held-out events from text and inferring new events from text, and it outperformed earlier systems in both. They compared their results to a variety of baselines, including the previous best-published method, and found considerable improvements in inferring both held-out and original events.

[2] A review on voice recognition was presented by Santosh K.Gaikwad, Bharti W.Gawali, and Pravin Yannawar. Speech is the most common and fundamental means of communication among humans, according to them. Speech has the potential to be an essential way of computer interaction. This paper provides an overview of significant technological perspectives and an understanding of the underlying progress of speech recognition, as well as an outline of the techniques established at each stage of the process. This document aids in the selection of a technique, as well as its relative merits and drawbacks. A stage-by-stage comparison of several techniques is carried out. The selection on feature direction for developing techniques in human computer interaction systems concludes this study. MFCC is commonly used for speech feature extraction and GHM and HMM are the best among all modeling techniques, according to this review.

[3] The Visually Impaired People's Common Problems were explained by Anu Arora and Anjali Shetty. Visual impairment has traditionally been considered a barrier to proper human functioning, particularly in the areas of participation and economic production. Uneven street surfaces, open manholes, parked cars, sellers, and other obstacles made ambulation difficult for the visually impaired. Make a formal request to the government to enhance road infrastructure. The width of the sidewalk or surface should be sufficient. Railings should be installed between the roads and the pavements. Short films about the issues experienced by blind people, posters, and other methods can be used to raise awareness, particularly in rural areas where religious myths about blindness exist.

[4] Automatic Speech Recognition, according to Preeti Saini and Parneet Kaur, is a beneficial technique. As explained in this paper, the accuracy of automatic speech recognition (ASR) remains one of the most important research challenges, along with speaker and language variability, vocabulary size and domain, and noise. Speech recognition system design necessitates paying close attention to issues such as different types of speech classes, speech representation, feature extraction methodologies, database, and performance evaluation. The results of this investigation of basic approaches to voice recognition suggest that they are more accurate. The research that has been done to address the issue of ASR is also discussed in this study.

[5] Speech Recognition Using Artificial Neural Network was presented by Nidhi Srivastava. For speech recognition, this article employs Neural Networks (NN) and Mel Frequency Cepstrum Coefficients (MFCC). Mel Frequency Cepstrum Coefficients (MFCC) is a tool for extracting speech features since it generates training vectors by converting speech signals into frequency domain. Artificial Neural Network (ANN) is a self-learning computer model for information processing that integrates artificial neurons to analyse data and solve issues. The weights of an artificial neuron are adjusted to produce a given output from a specific input. In MATLAB, 120 samples were recorded at a sampling frequency of 44100 Hz, with 60 samples used for training and 60 samples used for testing. The data train algorithm is used for training. The MFCC coefficients are calculated using the VOICEBOX speech processing toolbox function melcepst. Finally, the networks were created, trained, and simulated using the MATLAB Neural Network Toolbox, with mean square error used to assess their performance.

[6] Mrs. Omprakash Yadav, Ryan Fernandes, Rohit Tiwari, and Sheenam Kaul gave a presentation titled Online Reservation System Using QR Code Based Android Application System. This study introduces a new Seat Allocation method based on a QR code image that provides information about the ticket and the passenger in the form of a two-dimensional image, reducing scanning time. The major goal of this study paper is to make waiting list passengers' journeys in Indian Railways more convenient. Wireless standards are used to connect HHT to the DSA server, which provides authentication to each ticket. The DSA server's automatic upgradation method allows reservations to be made while the train is moving, as well as providing transparency in berth or seat bookings made online or at the counter. HHT devices scan the QR code, which encodes the URL as part of the check-in procedure, redirects to the PRS server, and retrieves the stored data to validate the passenger. The check-in process updates all passengers' information and allows the DSA server to make a seat reservation or vacant. The seats of absent passengers are assigned to the queue by the DSA server. The check-out process allows the passenger to take a break from

his journey. The booking interface allows you to book tickets for passengers on board.

[7] Ms.Vrinda and Mr.Chander Shekhar delivered a paper on an English language speech recognition system. The purpose of this study is to provide an overview of speech recognition technologies. It explains how voice recognition systems function as well as the level of accuracy to be expected. The major goal of this study is to create a speech recognition system for physically challenged people who can't use a keyboard or mouse to operate a computer. The HMM (hidden Markov model) is utilised to recognise speech samples in this paper, and the results for single words are outstanding. This research investigates how to capture human speech in a digital computer and decode it into text, i.e., converting speech to text. This project can be scaled up to a very big size with very few changes. A medium-sized vocabulary system was implemented during the trial. Using the HMM Technique or other developing techniques like Artificial Neural Network, the system can be extended to continuous word recognition with a vast vocabulary based on a phone acoustic model.

[8] An Android Application for Ticket Booking and Ticket Checking in Suburban Railways was given by Subarnarekha Ghosal, Shalini Chaturvedi, Akshay Taywade, and N. Jaisankar. The creation and execution of a smartphone application that is more effective than the current ticketing system is the subject of this study. The "Android Suburban Ticket (ASR)" can be purchased at any time and from any location, and the ticket will be stored in the customer's phone as a "Quick Response Code." The ticket is validated at the source and deleted at the destination using GPS technology. which is not available in the present suburban railway system. Also included with the ticket checker is an application that searches the cloud for the user's ticket using the ticket number. The implementation is done in the platform-independent language java. Additionally, for user and ticket information, SQLite and Cloud Database are employed as databases. PHP is also utilised as a framework for web development.

[9] A Survey on Speech Recognition was delivered by Harpreet Singh and Asok Kumar Bathla. This document provides an overview of the voice recognition system as well as its most recent developments. The primary goal of this work is to compare and summarise some of the commonly used approaches in voice recognition systems at various stages. Humans' most common means of communication is speech. Human computer interface refers to the communication between humans and computers. The process of a computer recognising human voice in order to generate a string of words or commands is known as speech recognition. Speech recognition systems' output can be used in a variety of fields. For the development of Automatic Speech Recognition (ASR), there are a variety of artificial intelligence techniques accessible. This research compares the performance of the ASR system based on the adopted feature extraction technique and the speech recognition approach for the specific language.

[10] M.A.Anusuya and S.K.Katti described Automatic Speech Recognition briefly (ASR). The accuracy of automatic speech recognition remains one of the major research challenges after years of research and development.
The following issues must be carefully considered while designing a Speech Recognition system: Different sorts of speech classes are defined, as well as speech representation and feature extraction. The challenges that exist in ASR, as well as the numerous strategies

developed by various research workers to overcome these problems, have been given in chronological sequence. The goal of this paper is to outline and compare some of the commonly used approaches in various phases of speech recognition systems, as well as to identify research topics and applications that are at the cutting edge of this interesting and challenging field. Scientists are interested in speech recognition because it is a critical sector with a technical impact on society and is expected to increase in this area of human-machine interaction.

3. PROPOSED MODEL

Continuous Data Integration and Data Retrieval in a Cloud-based Automatic Speech Recognition Engine is a multi-faceted problem. It's difficult to improve the performance of a Language Model for Live Data without impacting the total performance. We look at the quantity of data that needs to be kept and processed on a device and implement Android App Voice Data Processing with low memory footprint and energy usage. A similar amount of data must be sent and sent to the cloud server as well. We don't compress recorded voice because Google STT only supports FLAC as a compressed audio file. The suggested solution overcomes the limitations of the existing system by scouring the internet for relevant information and properly weighing it. The main concept behind our solution is that it does intense voice recognition and synthesis in the cloud. The technology can detect the speaker's voice and take the necessary steps to book tickets. The voice that has been recognized is sent to the Google STT service, where it is matched. The appropriate operation is done out if the voice matches. With the help of this technology, the user (visually impaired persons) can be self-sufficient and accomplish all necessary tasks. To book tickets, automatic speech recognition (ASR) is used (ASR).

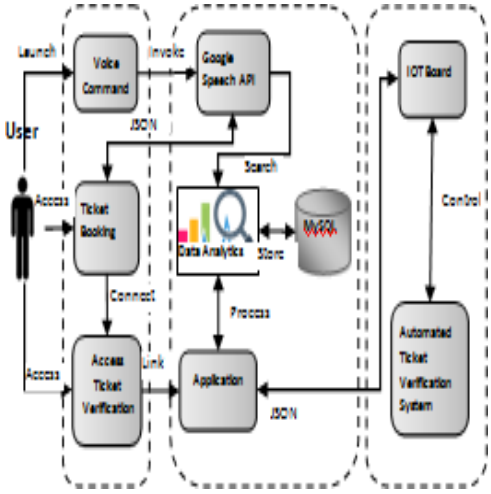

Figure 1.Architecture diagram

MODULES USED

1.Android Based Voice Enabled Ticketing System using Google Speech API2.Data Analytics through LSTM using Python and Tensorflow

3. IoT Board Interface with PHP-based Cloud Communication

MODULE DESCRIPTION

5.1 Android based voice enabled ticketing system using Google Speech API

Developers can utilise the Google Cloud Speech API to convert audio to text using strong neural network models in an easy-to-use API. To serve global users, the API detects over 110 languages and variants. You may transcribe text dictated to an application's microphone, enable voice command-and-control, and transcribe audio files, among many other things. Recognize audio in the request and combine it with your Google Cloud Storage audio storage using the same technique Google uses to power its own products. The voice API can stream text results, returning partial recognition results as they become available and displaying the recognised text as you speak. Alternatively, the Speech API may extract text from audio files and return it as text that has been recognised. Deep learning neural networks enable Automatic Audio Recognition (ASR) to power your voice search and speech transcription applications. In our project, this module is used for the following purposes. When a user talks into the created Android app, the app sends the request to the speech recognition process, which recognises the user's voice. The voice is recognised and transferred to the cloud as training logs, which are subsequently returned back to speech recognition as forecast logs. It delivers the response back to the Android application after successful recognition. The available menu is sent to the screen reader, which performs a text to speech operation and reads the contents on the screen. The voice I/O response is sent to the user by the android app. The cloud is also communicated with by the screen readers. The created ticket is saved in the device data store after a successful ticket booking. The ticket can be retrieved from the device's data store by the user.

5.2 Data Analytics through LSTM using Python and Tensorflow

Data analytics (DA) is the act of analysing data collections in order to derive conclusions about the information they contain, with the use of specialised systems and software becoming more common. Data analytics tools and techniques are widely utilised in the commercial sector to help businesses make better business decisions, as well as by scientists and academics to validate or refute scientific claims. TensorFlow is a graph-based computation framework. It's a different way of thinking about mathematical calculations. A recurrent neural network's hidden layer output is fed back into itself after passing through a conceptual delay block. Recurrent neural networks are extremely adaptable, however the vanishing gradient problem is their fundamental drawback. Lengthy memory are required for recurrent neural networks so that the network can connect data relationships across long periods of time. Long-short term memory (LSTM) networks are the most prevalent method for dealing with this problem. It accomplishes this by generating an internal memory state that is simply added to the processed data, considerably reducing the multiplicative effect of minor gradations. A fascinating notion known as a forget gate controls the temporal dependence and effects of earlier inputs, determining which states are remembered or forgotten. The input gate and output gate are two more gates found in LSTM cells. This module performs the following tasks in this project. The Google speech recognition mechanism collects the audio stream (voice) from the user and stores the obtained text in the training data logs to test and train the datasets. The LSTM algorithm extracts texts from training data logs and stores them in a results database with which the Google speech recognition process interacts directly. The text is provided to the word error rate calculator and the loss functions from the outcome database. The process input output operations are sent from the LSTM neural network to the tensor flow implementation process, which gives the LSTM with graphical display.

5.3 IoT Board Interface with PHP-based Cloud Communication

Cloud communications are Internet-based voice and data communications in which telecoms apps, switching, and storage are hosted by a third party outside of the firm using them and available via the public Internet. Cloud communications providers supply voice and data communications apps and services, which they host on servers they own and administer. Voice is a terrific way to communicate these days. This is the notion we employed in our project to make work easier and more beneficial for persons who are visually impaired. Today, just a few commercial mobile apps are self-contained. Most interact with back-end services, some of which are on-premises but many of which are hosted in the cloud. We can talk about how PHP can be used to build mobile apps and services, as well as how Right Scale can help manage and automate mobile app development and deployment in the cloud. 81 percent of websites, according to Web Technology Surveys. PHP was more cloud-friendly than their second most-used language. Because we'll be working on cloud systems, PHP will be the ideal option for mobile apps. After booking tickets, our information is stored in a database as logs in our programme. We can validate our tickets by saying verify ticket. If we say open later, the IOT board checks the data and the motor in it starts working, allowing the vehicle door to open.

4. CONCLUSION

The main aim of this project is to develop a ticket booking application that performs all the operation by speech recognition. Hence, visually impaired people can make use of this application. The main goal of this project is to create a ticket booking application that uses speech recognition to fulfil all tasks. As a result, this tool can be used by those who are blind or visually challenged. A variety of applications are available to help the visually impaired. This programme improves the functionality of currently accessible software. Without the assistance of personal assistants, visually impaired people can utilise this to order tickets

5. REFERENCES

[1] Karl Pichotta and Raymond J. Mooney,” Learning Statistical Scripts with LSTM Recurrent Neural Networks”, In Proceedings of the 30th AAAI Conference on Artificial Intelligence
[2] (AAAI-16), Phoenix, Arizona, February 2016.
[3] Santosh K.Gaikwad,Bharti W.Gawali,Pravin Yannawar,” A Review on Speech Recognition Technique”, International Journal of Computer Applications (0975 – 8887),Volume 10– No.3, November 2010
[4] Anu Arora, Anjali Shetty,” Common Problems Faced By Visually Impaired People”, International Journal of Science and Research (IJSR),Volume 3 Issue 10, October 2014
[5] Preeti Saini, Parneet Kaur,” Automatic Speech Recognition: A Review”,International Journal of Engineering Trends and Technology- Volume4Issue2- 2013
[6] Nidhi Srivastava,”Speech Recognition using Artificial Neural Network”,International Journal of Engineering Science and Innovative Technology (IJESIT) Volume 3, Issue 3, May 2014
[7] Mrs. Omprakash Yadav, Ryan Fernandes, Rohit Tiwari, Sheenam Kaul,” Online Reservation System Using QR Code basedAndroid Application System ”, International Journal of Scientific and Research Publications, Volume 4, Issue 12, December 2014
[8] Ms. Vrinda1 ,Mr. Chander Shekhar,” Speech recognition system for english language”, International Journal of Advanced Research in Computer and Communication Engineering,Vol. 2, Issue 1, January 2013
[9] Subarnarekha Ghosal, Shalini Chaturvedi, Akshay Taywade and N. Jaisankar,” Android

Application for Ticket Booking and TicketChecking in Suburban Railways"Indian Journal of Science and Technology, Vol 8(S2), 171–178, January 2015

[10] Harpreet Singh,Ashok Kumar Bathla," A Survey on Speech Recognition",International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume No. 2, Issue No. 6, June 2013

[11] M.A.Anusuya,S.K.Katti," Speech Recognition by Machine: A Review", (IJCSIS) International Journal of Computer Science and Information Security,Vol. 6, No. 3, 2009