ISSN: 2395-7852

International Journal of Advanced Research in Arts, Science, Engineering & Management

Volume 10, Issue 3, May 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 6.551

+91 9940572462 +91 9940572462 ijarasem@gmail.com www.ijarasem.com

International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 6.551 |Bimonthly, Peer Reviewed & Referred Journal|

| Volume 10, Issue 3, May 2023 |

Speech Recognition Solution with Commands for Visually Impaired

[1]Avani Hegde, [2]Shreya R J, [3]Sohan Mohandas Shetty,[4] Mr Rajesh N Kamath

Student, Dept. of Information Science & Engineering, Mangalore Institute of Technology & Engineering, Moodbidri, Karnataka, India[1,2,3]

Dept. of Information science & Engineering, Mangalore Institute of Technology & Engineering, Moodbidri, Karnataka, India[4]

ABSTRACT: Speech is a crucial aspect of communication, but individuals with dysarthria face page in controlling PCs. To address this, we propose developing a speech recognition application to decode acoustic signals and convert them into syllables or words. However, no commercial or open-source product currently allows for unrestricted text input through automatic speech recognition. The fast-growing technology of speech recognition has numerous applications in different areas and can significantly benefit the nearly 20% of people globally with the disabilities, including those are visually impaired or have limited hand function. With our project, users can share information and operate computers through voice input. This is especially helpful for thousands of individuals who cannot type or see, ,those are simply too lazy to do so. Our application recognizes speech and converts audio input into text while also enabling users to perform operations like opening, closing, exiting, and reading program applications and files through voice commands**.**

KEYWORDS: Speech Recognition, unrestricted text input, automatic speech recognition

I.INTRODUCTION

Speech Recognitiontechnology is a groundbreaking development in the domain of computer science, linguistics and computer engineering. The ability of a machine or program to recognize spoken words and convert them into written text. While early versions of the software had limited vocabularies and the ability to recognize words and phrases spoken clearly, modern systems are advanced, able to handle natural speech, different accents, and various languages.
It is becoming increasingly prevalent in many modern devices and text-focused programs, enabling users to interact with their devices more easily or even hands-free. For individuals with disabilities like blindness or limited mobility, speech recognition technology can particularly be transformative, allowing them to use computers and other devices with greater independence.

To achieve accurate speech recognition, computer algorithms are used to process and interpret spoken words, converting them into text that computers and humans can understand. This involves capturing sound via a microphone, converting the electrical signal from the microphone into a digital signal using an Analog to Digital (ADC) converter, and recording the digitized sample into computer memory. The computer then compares and attempts to match the input voice with a stored digitized voice sample, ultimately identifying the speaker.

Given the potential benefits, speech recognition technology, it has variety of applications across various fields. For example, it is used to improve accessibility for the same, enhance customer service experiences, or streamline transcription services in healthcare or legal settings.
In today's world, where approximately 20% people suffer from various disabilities, speech recognition technology provides a critical means of improving quality of life by promoting greater inclusion. While there is a lot of work to be done in refining and advancing speech recognition technology, its potential to change the way we interact with technology is undeniable.

II.LITERATURE SURVEY

In [1]Yogesh Kumar, Navdeep Singh, "A comprehensive view of automatic speech recognition system-a systematic literature review ",[2019], Human beings always try to communicate with objects using natural language, and communication will always be a vital aspect of human life for sharing and building information across generations. The

most difficult speech processing problems is converting recorded speech signals into text using automatic speech recognition mechanisms.

In [2] Automatic speech recognition :Dong Yu, Li Deng Springer, 2016,Automatic Speech Recognition (ASR), which is aimed to enable natural human–machine interaction, has been an intensive research area for decades. Many core technologies, such as Gaussian mixture models (GMMs), hidden Markov models (HMMs), mel-frequency cepstral coefficients (MFCCs) and their derivatives, ngram language models (LMs), discriminative training, and several adaptation techniques developed along the way, mostly prior to the new millenium.

In [3] Speech Recognition Technology: JianliangMeng, Junwei Zhang, Haoquan Zhao2012 Speech recognition is a cross-disciplinary field that involves the identification and understanding of speech signals to convert them into text or commands, enabling natural voice communication between humans and machines. It encompasses physiology, psychology, linguistics, computer science, signal processing, and even body language, aiming to achieve human-like communication with machines.

In [4] Sven L Mattys, Matthew H Davis, Ann R Bradlow, Sophie K Scott,"Speech recognition in adverse conditions: A review" ,[2012],This article presents a review of effects of adverse conditions (ACs) on the perceptual, linguistic, cognitive, and neurophysiological mechanisms underlying speech recognition. The review starts with a classification of ACs based on their origin: Degradation at the source (production of a noncanonical signal), degradation during signal transmission (interfering signal or medium-induced impoverishment of the target signal), receiver limitations (peripheral, linguistic, cognitive).

In [5] Maree Johnson, Samuel Lapkin, Vanessa Long, Paula Sanchez, Hanna Suominen, Jim Basilakis, Linda Dawson," A systematic overview of speech recognition technology in health care",[2014],This study involved a systematic review of literature from 2000 to examine the application of speech recognition technology in healthcare settings used by healthcare professionals, with a focus on evaluation, patient outcomes, and staff outcomes.

In [6]TaabishGulzar, Anand Singh , Dinesh Kumar Rajoriya and Najma Farooq, "A Systematic Analysis of Automatic Speech Recognition: An Overview",[2014], This paper assesses advancements in modern speech recognition systems compared to human performance. Sources of knowledge and the role of hypotheses in verification are outlined.

III.SCOPE AND METHODOLOGY

Aim of the project

The aim of our project is to work on a speech recognition system that is capable of accurately interpreting spoken words and executing tasks accordingly. This is achieved through the help computer algorithms that process and interpret the sound captured by a microphone, converting it into written language that can be understood by both humans and computers.

Existing system

The recent speech recognition systems have come a long way since their outgrowth, with Siri, Apple's digital assistant, becoming a household name. Despite the advancements made in speech recognition technology, there is room for improvement, particularly in applications for individuals with disabilities.

Proposed system

The proposed system for speech recognition involves encoding and parameterizing speech signals in a 2-D time matrix with four parameters. The mean and variance of each pattern are then optimized using genetic algorithms to generate the rule base of a Mamdani fuzzy inference system. This system has been designed to help individuals with visual impairments by allowing them to give instructions through speech recognition, enabling tasks such as making phone calls, sending text messages, accessing text files, and writing notes to be executed seamlessly. The app works by converting electrical signals captured by the microphone into digital signals, with no additional prompts or calls required. By improving communication and task execution for those with disabilities, our speech recognition system has the potential to enhance their quality of life and promote greater inclusivity. Continued research and development of this technology holds limitless possibilities

IV.DESIGN AND IMPLEMENTATION

The Design phase is an important step in software development, involving creating a system architecture that meets both functional and non-functional requirements. This involves breaking downof large systems into sub-systems that offer a various set of related services. The end result of this is a detailed description of the software structure

Speech Recognition System Architecture

A sequence diagram is a graphical representation of an interaction between objects in a software system. It is used to describe the flow of messages that occur between objects during a particular interaction. These diagrams are part of the Unified Modeling Language (UML) and are commonly used in software design and development.In a sequence diagram, objects are represented by vertical lines called lifelines. These lifelines represent the lifespan of an object and are arranged from top to bottom focused on the time of their creation. The messages between objects are shown as horizontal arrows between the lifelines.

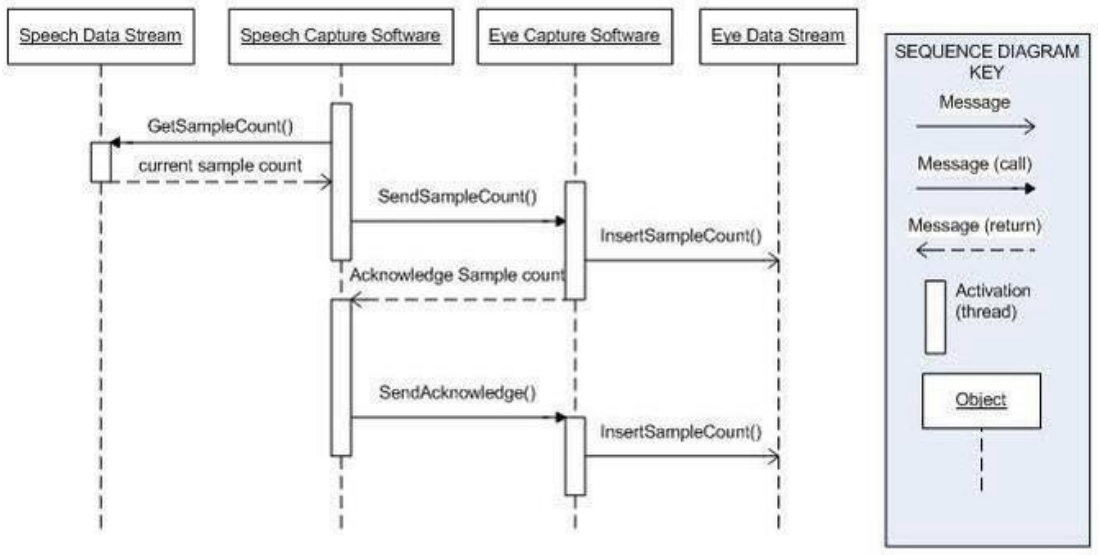

Fig 1.Speech Recognition System Sequence Diagram

A speech recognition system includes several essential components that work together to convert spoken language into text. The first component is the acoustic front-end, which converts incoming speech signals from a microphone into a sequence of fixed-size acoustic vectors using a process called feature extraction. During the training phase, the system estimates the parameters of word or phone models based on the acoustic vectors of the training data.The acoustic model maps the acoustic features to phonetic units, such as phones or triphones, while the lexicon provides information about the pronunciations of individual words. The language model captures the statistical properties of the language being recognized, providing context to help the system choose the most likely sequence of words.Finally, the decoder uses all of this information to produce the final recognition output. The acoustic front-end, acoustic model, lexicon, language model, and decoder work together seamlessly to ensure the accurate transcription of spoken language into text.In summary, the success of a speech recognition system relies on the effectiveness of its individual components and the synergy between them. By combining the expertise of acoustic signal processing, machine learning, and natural language processing, speech recognition systems have become increasingly powerful and accurate, enabling new applications and opportunities for human-machine interaction.

V.CONCLUSIONS

The purpose of the paper is to assess the efforts made to alleviate accessibility barriers faced by individuals who are visually impaired when accessing printed, written, or visual information using the speech channel.

REFERENCES

[1] Nirmal Varghese Babu, E Kanaga, “Sentimental analysis in social media data for depression detection” [2022].

[2] Sonam Gupta, Arjun Singh, LipikaGoel, Ajay Prasad, Mohammad AmanUllar, “Psychological Analysis for Depression Detection from Social Networking sites” [2022].

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 6.551 |Bimonthly, Peer Reviewed & Referred Journal|

| Volume 10, Issue 3, May 2023 |

[3] HessaAlmutairi, KholoundAlharbi, “Postpartum Depression Detection: Concept Analysis” [2022].

[4] Xiaowei Li, Xin Zhang, Jing Zhu, Wandeng Mao, Shuting Sun, Zihan Wang, Chen Xia, Bin Hu, “Depression recognition using machine learning methods with different feature generation strategies”, [2019].

[5] AastikMalviya, Rahul Meharkure, Rohan Narsinghani, VirajSheth, “Depression Detection Through Speech Analysis: A Survey” [2019].

[6] Qingxiang Wang, Huanxin Yang, Yanhong Yu, “Facial expression video analysis for depression in Chinese patient”, [2018].

[7] J Koo LB Marangell, M Nakamura, A Armstrong, C Jeon, T Bhuthani, JJ Wu, “Depression and suicidality in psoriasis review of literature including the cytokine theory of depression”, [2017].

[8] Paula Lopez-Otero, Laura DocíoFernández, Alberto Abad, Carmen Garcia-Mateo, “Depression Detection Using Automatic Transcriptions of De-Identified Speech”,[2017].

[9] VikramjitMitra, Elizabeth Shriberg, “Effects of type, learning algorithm and speaking style for depression detection from speech”,[2015].

International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

| Mobile No: +91-9940572462 | Whatsapp: +91-9940572462 | ijarasem@gmail.com |

www.ijarasem.com