iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** V **Month of publication:** May 2022

DOI: https://doi.org/10.22214/ijraset.2022.43538

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

Voice Based Email for Visually Challenged

Abhiram J[1], Amrutha S[2], Aneetta Susan John[3], Jinshu Maria John[4], Midhun V Nair[5], Alpha Mathew[6]
[1, 2, 3, 4, 5]*Department of Computer Science, College of Engineering Kidangoor, Kottayam, India*
[6]*Department of Computer Science and Engineering, College of Engineering Kidangoor, Kottayam, India*

Abstract: ***The inception of Internet has caused a dramatic revolution in many fields. Internet, being a global computer network, has made life of people easier, as they could access any information they want, more efficiently. Communication is one of the main fields that Internet has revolutionized. Communication has become so easy due to the integration of communication technologies with the internet. E-mails are considered to be the most reliable way of Internet communication, for sending or receiving some important information. Visually challenged person feels difficulty in using these technologies as it requires visual perception. Around 250 million people in this world, are unaware about the usage of Internet or E-mail. The only way by which a visually impaired person can use the current email application is that, they require the help of a third person who would send mail on the behalf of the them.But this does not guarantee privacy and security for that person. This gave the idea of developing a voice-based email which requires only less training. It makes use of mouse operations and speech recognition. It could be used by both visually impaired and also by a normal person***

Index Terms: Feature extraction, MFCC, GMM, Speech recognition, Google API

I. INTRODUCTION

The Voice based Email for visually challenged, however, is a technology which has greater significance that could lead to growing digital world. We will be developing a voice based email system which provides an aid to the visually impaired folks that are na¨ıve to computer systems to use email facilities more securely and efficiently. This e-mail system can be accessed by any user of any age bracket easily. It provides the feature of speech to text and also text to speech with speech reader which makes designed system to be handled by visually impaired person with more ease. It would be a web-based application for visually impaired persons that make use of IVR- Interactive voice response, thus enabling everyone to regulate their mail accounts using their voice only and also to read, send, and perform all other useful activities. The system will prompt the user with voice commands to perform certain action and then, the user will respond accurately to the same. The main advantage of this system is that the use of keyboard is eliminated. The user will have to respond through voice and mouse click only. Also the user needn't worry about which mouse click operation he/she must perform so as to avail a given service as the system itself would be prompting them on which click will provide them with what operations.

II. OBJECTIVES AND SCOPES

This system would be a better aid for visually challenged people to access the mail services without the help of a third person.One of the main objective of this system is that it provides more privacy.Also the system does not require the use of keyboard.Instead, it works only on mouse operations and speech conversions to text. This project is proposed for the betterment of the society.
One of the major issue faced by visually impaired people while using the current mail system is that, they lacks privacy as they requires the support of a third person to use this system. An ideal solution for this problem is to develop a voice based email system that could be accessible by visually impaired people without a third person help. The proposed system make use of Google API and Gaussian Mixture Model(GMM) for feature extraction and speech recognition.

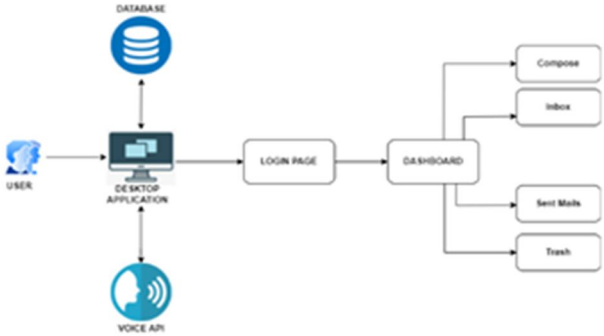
III. PROPOSED METHOD

The task of the proposed system is that, it completely eliminates the use of keyboard and is based on mouse clicks and speech recognition. The user is first asked to login by entering the login credentials. The validity of the details are checked and are encrypted and if valid, we are redirected to the dashboard. It is the main page where the system provides services like Compose, Inbox, Trash etc. The system will prompt the user with voice commands to perform a certain action and the user will respond to the same. To compose a mail, this voice command given by the user is converted to text and is send to the recipient. Similarly, for all the other services, the user is prompted via voice commands.

©IJRASET: All Rights are Reserved

IV. SYSTEM DESCRIPTION

A. *Architecture*



The system begins with registration of new user by entering his/her details like name, mail id, training model etc. It is done by an admin. Entered informations are stored to the database. These informations are then fetched from the database when-ever needed. Already registered users can directly login to the system by entering the email id via voice commands. We use Google API for speech recognition. If the system detects the user as valid, he/she will be directed to the dashboard where the email services can be accessed.

B. *System modules*

The System mainly consists of Registration module, Login module and Dashboard module.

1) *Implementation of Registration Module:* New User Registration is done in this module. It is done by an admin. Registration is done by entering details like name, mail id, gender etc. These entered informations are stored in the database which can be monitored by the admin. Along with registration, feature extraction and training is also done on the data set.

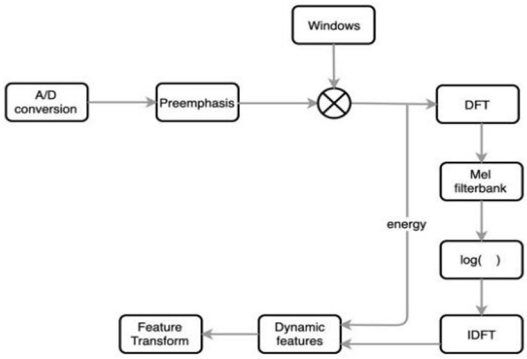


Fig. 1. Feature extraction using MFCC

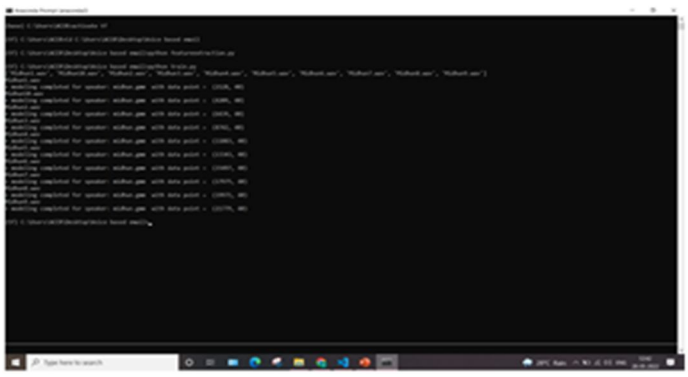


Fig. 2. Training Phase

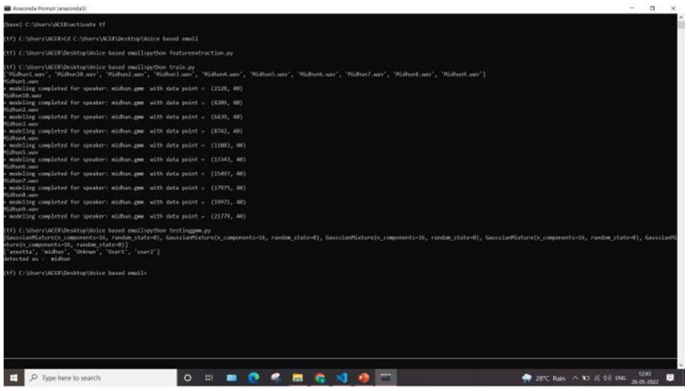


Fig. 3. Testing Phase

2) *Implementation of Login Module:* After Registration, the user can login to the system via Login module. Here, the user is prompted to enter mail id as voice command.I f the mail id is valid, then the user is asked for confirmation. After confirmation, system requests the user to say ”password”. If the voice matches with the trained dataset, the user will be directed to the Dashboard.



Fig. 4. Registration module



Fig. 5. Login module

3) *Implementation of Dashboard Module:* After successful login, we enter the dashboard module.There are mainly 5 services and it can be accessed either by mouse clicks or voice commands.

a) *Compose:* User is asked to speak the recipient mail id,mail subject and content to be composed via voice.After each entry, system asks for confirmation.After getting confir-mation, the mail is sent to the desired recipient.

b) *Inbox:* User can check all unread mails and recently received mails

c) *Read Messages from an Email Id:* User is able to search a specific mail from inbox.User will be asked to speak the mail id and thus, mails from that particular mail id can be accessed via voice.

d) *Delete Mails:* User can delete unnecessary mails from inbox.Mails from specific users can be searched by saying the mail id and then it can be deleted.

e) *Logout:* User can logout from the system by selecting the logout option.

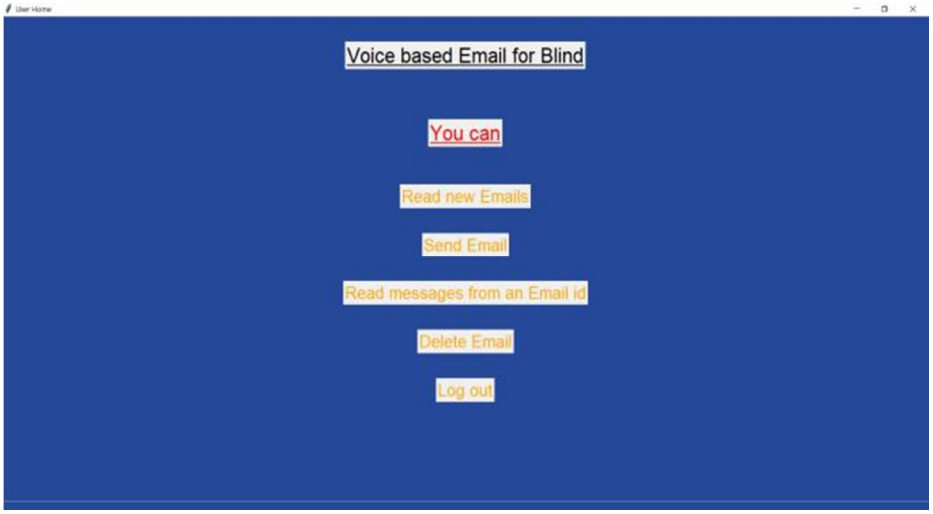


Fig. 6. Dashboard module

V. MODELS USED

A. *Mel Frequency Cepstral Coefficient*

Mel Frequency Cepstral Coefficient(MFCC) are coefficients that collectively make up an MFC. They're derived from a sort of cepstral representation of the audio clip (a nonlinear "spectrum-of-a-spectrum"). The difference between the cep-strum and Mel-frequency cepstrum is that, within the MFC, the frequency bands are equally spaced on the Mel scale, which approximates the human auditory system's response more closely than the linearly-spaced frequency bands that are used in normal spectrum. This frequency warping allows better representation of sound, for example, in audio compression.

MFCCs are commonly derived as follows:

1) Fourier Transform: Take the Fourier transform of (a windowed excerpt of) the given signal.
2) Mapping: Map the powers of spectrum obtained above onto Mel scale, using triangular overlapping windows or cosine overlapping windows.
3) Applying log: Take the logs of powers at each Mel frequencies.
4) Applying DCT: Take the discrete cosine transform of the list of Mel log powers, as if it was a signal.
5) Computing MFCCs: The obtained MFCCs are the am-plitudes of resulting spectrum.

B. *Gaussian Mixture Model*

Gaussian Mixture Model(GMM) is a type of machine learning algorithm used for data clustering.It classifies data into different categories based on the frequency or pitch of the user's voice. GMM mainly uses Unsupervised Learning and is more robust. GMM is a probabilistic model that assumes all the data points are generated from a mixture of a finite number of Gaussian distributions with unknown parameters. One can think of mixture models as generalizing k-means clustering to incorporate information about the covariance structure of the data as well as the centers of the latent Gaussians.It make use of maximum-likelihood estimation.

©IJRASET: All Rights are Reserved

VI. SYSTEM REQUIREMENTS

A. Software Requirements

1) *Tensorflow:* Currently, the most famous deep learning library in the world is Google's TensorFlow. Google product uses machine learning in all of its products to improve the search engine, translation, image captioning or recommendations. It has several wrappers in several languages like Python, C++ or Java. The model can be trained and used on GPUs as well as CPUs. GPUs were initially designed for video games. In late 2010, Stanford researchers found that GPU was also very good at matrix operations and algebra so that it makes them very fast for doing these kinds of calculations. Deep learning relies on a lot of matrix multiplication. TensorFlow is very fast at computing the matrix multiplication because it is written in C++. Although it is implemented in C++, TensorFlow can be accessed and controlled by other languages mainly, Python.
2) *Tkinter:* Tkinter is the standard GUI library for Python. Python when combined with Tkinter provides a fast and easy way to create GUI applications. Tkinter pro-vides a powerful object-oriented interface to the Tk GUI toolkit.Tkinter provides various controls, such as buttons, labels and text boxes used in a GUI application. These controls are commonly called widgets. There are cur-rently 15 types of widgets in Tkinter.All Tkinter widgets have access to specific geometry management methods, which have the purpose of organizing widgets throughout the parent widget area.
3) *Xampp:* XAMPP is one of the widely used cross-platform web servers, which helps developers to create and test their programs on a local webserver.It consists of Apache HTTP Server, MariaDB, and interpreter for the different programming languages like PHP and Perl. It is available in 11 languages.XAMPP is an abbreviation where X stands for Cross-Platform, A stands for Apache, M stands for MYSQL, and the Ps stand for PHP and Perl, respectively. It is an open-source package of web solutions that includes Apache distribution for many servers and command-line executables along with modules such as Apache server, MariaDB, PHP, and Perl.
4) *Scikit-learn:* Scikit-learn (formerly scikits.learn and also known as sklearn) is a free software machine learning library for the Python programming language. It features various classification, regression and clustering algorithms including support-vector machines, random forests, gradient boosting, k-means and DBSCAN, and is designed to interoperate with the Python numerical and scientific libraries NumPy and SciPy. Scikit-learn is a NumFOCUS fiscally sponsored project
5) *Python IDLE:* IDLE (Integrated Development and Learning Environment) is an integrated development environment (IDE) for Python. The Python installer for Windows contains the IDLE module by default.IDLE can be used to execute a single statement just like Python Shell and also to create, modify, and execute Python scripts. IDLE provides a fully-featured text editor to create Python script that includes features like syntax highlighting, autocompletion, and smart indent. It also has a debugger with stepping and breakpoints features.In IDLE, we write code line by line.
6) *Google API:* Google APIs are application programming interfaces (APIs) developed by Google which allow communication with Google Services and their integration to other services. Examples of these include Search, Gmail, Translate or Google Maps. Third-party apps can use these APIs to take advantage of or extend the functionality of the existing services.The APIs provide functionality like analytics, machine learning as a service (the Prediction API) or access to user data (when permission to read the data is given). Another important example is an embedded Google map on a website, which can be achieved using the Static Maps API, Places API or Google Earth API.
7) *NumPy:* It is a library consisting of multidimensional array objects and a collection of routines for processing of array.Using NumPy, a developer can perform the following operations such as Mathematical and logical operations on arrays,Fourier transforms and routines for shape manipulation,Operations related to linear algebra etc. NumPy has in-built functions for linear algebra and random number generation.

VII. RELEVANCE

A. It makes the life of differently abled people more easier.
B. This system makes disabled people feel like a normal person.
C. The use of keyboard is eliminated as, in this application, the user need to respond only through voices and mouse clicks.
D. It provides more privacy.

VIII. FUTURE SCOPE

1) Used for the betterment of society.
2) Helps the visually impaired people to be a part of growing digital India by using internet and also aims to make life of such people quite easy.
3) The success of this project will also encourage developers to build something more useful for visually impaired or illiterate people, who also deserves an equal standard in society.

©IJRASET: All Rights are Reserved

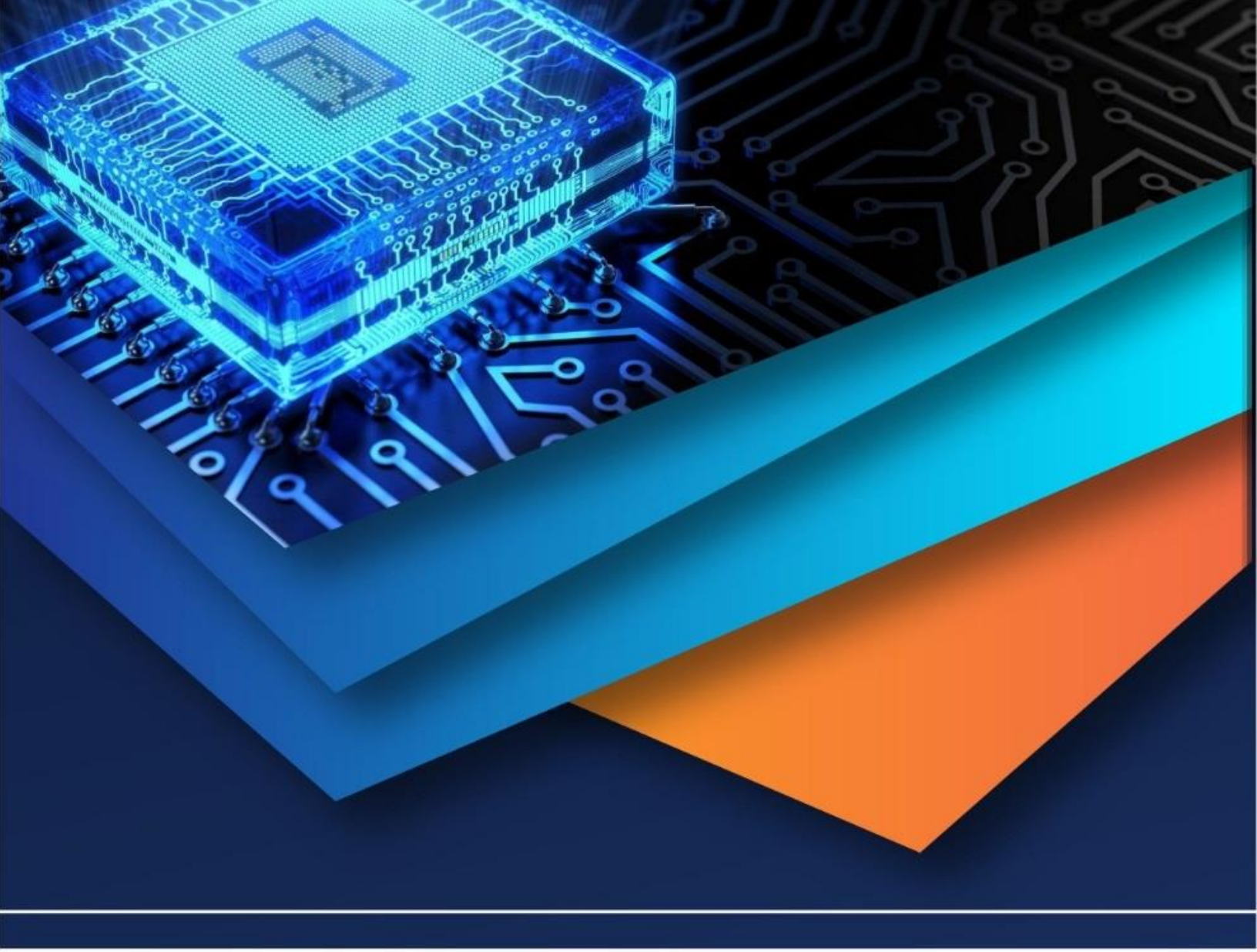
IX. CONCLUSION

Voice-based Email System for visually challenged will make the email system easily accessible to visually challenged people. Privacy is the most important feature that is considered while developing this system. Both fully and partially blind people can use this system. With the help of our system visually challenged people will become independent as they can use email services without the support of a third person. The system makes use of an efficient voice input and mouse click based technology which reduces the burden of accessing email service. As blind people become capable of performing mail services their own they will be able to contribute to the growing digital world.

REFERENCES

[1] Ayisha Zubain Bhandari,Prof.B C.Melinamath ”A Survey on Auto-matic Recognition of Speech via Voice Commands” International Jour-nal of New Innovations in Engineering and Technology,ISSN: 2319-6319,Volume 6 Issue 4-January 2017
[2] S. Usharani,P. Manju Bala,R. Balamurugan ”Voice Based Form Filling System for Visually Challenged People” ISBN 978-1-7281-6202-7,IEEE ICSCAN 2020
[3] Angayarkanni.S.A et al.,”SHOPAIDE: Voice Based AI Assistant for E-Shopping” Interantional Journal of Modern Agriculture,ISSN: 2305-7246,Volume 10 Issue 3,2021
[4] Subhash S et al.,”Artificial Intelligence-based Voice Assistant” Fourth World Conference on Smart Trends in Systems,Security and Sustain-ability(World S4) 2020

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)