Voice Based Mail System for Visually Impaired Using AI/ML

Dr. P. D. Halle[1], Anish Mavkar[2], Omkar Dalvi[3], Pranav Gonte[4], Vivek Yewale[5].

[*1] Asso. Prof, Dept. Of Information Technology Engg, SKN Sinhgad Institute of Technology & Science Lonavala, India

[*2,3,4,5] Student, Dept. Of Information Technology Engg, SKN Sinhgad Institute of Technology & Science Lonavala, India.

ABSTRACT: In today's digital world, communication tools are essential for maintaining both social and professional connections. While these technologies continue to evolve, they often remain inaccessible to individuals with visual or physical impairments. This project proposes a voice-driven email system tailored specifically for visually impaired and illiterate users, eliminating the need for keyboard usage or prior technical training.The system functions solely on voice input and output, utilizing speech recognition to understand user commands and text-to-speech for system responses. Users can perform standard email operations through intuitive spoken instructions and minimal physical interaction, creating a user-friendly and inclusive communication experience.

I. INTRODUCTION

This system aims to simplify communication for individuals who are blind or visually impaired by using advanced voice technologies such as Speech-to-Text (STT) and Text-to-Speech (TTS). By replacing the need for visual cues or keyboard inputs, users can send and receive emails using spoken commands and audio prompts.

Acting as a voice-activated email assistant, the system is particularly beneficial in regions like India, where many individuals face challenges due to visual disabilities or illiteracy. It allows users to interact with digital content independently, thereby enhancing their ability to connect with others and participate more fully in the digital world.

System Overview

The proposed system is designed to help visually impaired individuals use email services using voice interaction alone. By replacing traditional graphical interfaces with speech recognition and audio feedback, the system promotes accessibility and independence in digital communication. Users can perform all email-related tasks hands-free, making it ideal for those with visual limitations.

1.1 Components

1.Voice Input Module: Captures spoken commands through a microphone and processes them using Google's Speech Recognition API to convert voice to text.

Email Management Module: Handles

2.login, inbox viewing, composing, and sending emails. It interprets user instructions through Natural Language Processing (NLP) to determine the intended action.

3.Text-to-Speech Module: Converts system responses and received messages into spoken output using the pyttsx3 engine, enabling users to interact without visual support.

1.2 Functional Workflow

The stepwise functioning of the system is as follows:

1.The user launches the application and completes voice-based login.

2.The system provides voice prompts such as "Compose Email", "Read Inbox", or "Check Unread Messages".

3.The selected module is activated based on the user's spoken command.

4.While composing emails, users dictate messages which are confirmed via voice before sending.

5.The system retrieves new emails and reads them aloud to the user.

1.3 Technologies Used

The system is built with the following core technologies:

1.Python: Used to implement the system's logic and control structure.

2.Google Speech Recognition API: Converts audio input into textual commands.

3.Pyttsx3: Offers offline support for text-to-speech conversion. .

4.NLP Techniques: Help in interpreting user commands more effectively.

II.LITERATURE SURVEY

A literature survey helps to understand the existing systems and technologies used to assist visually impaired users in accessing email or

digital communication services. Several studies and projects have explored voice-based systems, AI/ML techniques, and accessibility improvements for such users.

1. Prakash et al. (2020) proposed a Voice-Based Email System for the Visually Challenged that relied on speech recognition to allow users to compose and read emails using voice commands. The system improved accessibility but lacked multi-language and emotion recognition support.

2. Kumar et al. (2021) developed a Voice Email Based on SMTP for Physically Handicapped, emphasizing efficient communication via speech input and audio feedback. However, their system required manual setup, limiting real-time usability.

3. Tiwari et al. (2022) reviewed several Voice-Based Email Systems for the Blind, highlighting the benefits of integrating Natural Language Processing (NLP) and smartphone-based accessibility. The study suggested that AI-driven adaptive learning could further enhance performance.

4. Baptista et al. (2019) designed a Telephone Interface for Email Services using Interactive Voice Response (IVR), allowing users to access emails through telephonic communication. Though functional, it lacked personalization and contextual understanding.

5. Malik et al. (2020) presented an enhanced Speech Recognition Email System incorporating text-to-speech and facial recognition. This improved authentication security but still faced challenges with noisy environments.

From the survey, it is evident that while previous systems successfully enabled basic email access for visually impaired users, most lacked deep AI/ML integration, emotion-aware interaction, and strong security features. The proposed system in this project addresses these gaps by combining speech recognition, NLP, and facial authentication, thereby.

III. SYSTEM ARCHITECTURE AND FUNCTIONALITY

2.1 Speech-Based Email Operations

The system allows users to carry out email tasks entirely through spoken language. Commands for composing, reading, or deleting emails are processed by a speech recognition engine, while system guidance and feedback are delivered using audio responses. This hands-free approach ensures that users can independently manage their emails without relying on a screen or keyboard.

2.2 Facial Authentication Module

To strengthen security, the system uses OpenCV and Haar Cascade classifiers for real-time facial recognition. Once a user's face is validated through the webcam, a voice confirmation is requested to complete the login. This dual-factor authentication—face and voice—adds an extra layer of protection while maintaining ease of use.

2.3 Registration and Training Process

User registration is conducted entirely through speech. Personal details are captured by voice, and facial data is collected via the webcam. This information is trained and stored to enable future logins. This process simplifies onboarding for visually impaired users.

2.4 Graphical Interface Integration

Although the core interaction is voice-driven, a simple graphical interface is included for testing and support purposes. It features accessible design elements like animations and video backgrounds to accommodate sighted users or developers during testing.

2.5 Error Handling and Feedback

The system is equipped with voice-based feedback for all actions. If a command is unclear, the system asks for clarification. Audio prompts such as "Command not recognized" or "Please repeat" maintain smooth communication. Status confirmations like "Email sent successfully" or "No new unread messages" enhance user experience.

2.6 Database and Training Storage

User Data Storage: The system maintains encrypted user credentials, facial images, and voice samples in a secure local database. Training Data: Facial data and voice patterns are trained to recognize individual users and reduce false login attempts

2.7 Multilingual Support (Optional Expansion)

The architecture supports adding multiple language models using Google's multilingual speech recognition API. This allows users to interact in their preferred language, enhancing accessibility for non-English speakers.

2.8 Hardware and Software Requirements

Hardware: Microphone, Webcam, and Speaker for input/output interaction. Software: Python 3.x, OpenCV, Pyttsx3, Google Speech Recognition API, smtplib, imaplib, and Tkinter for GUI testing.

IV. SYSTEM DESIGN

The system begins with facial authentication. Once access is granted, the user gives voice commands to perform actions like composing emails or reading inbox messages. The system uses speech processing and NLP to interpret the command and provides spoken feedback accordingly.

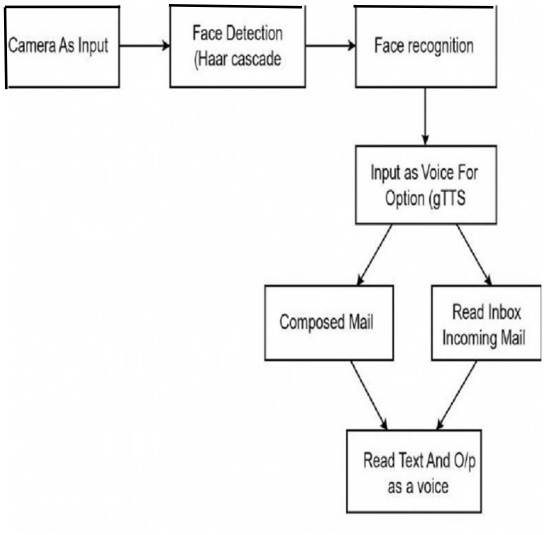

FIGURE 1: SYSTEM ARCHITECTURE

V. COMMUNICATION INTERFACE

The communication interface acts as the bridge between the user and the system, enabling seamless voice-based interaction. For visually impaired users, it replaces traditional graphical elements with auditory and tactile feedback. The design ensures accessibility, simplicity, and real-time response.

3.1 Voice-Based Communication

Voice interaction forms the primary mode of communication in the system. It allows users to perform all major operations like composing, sending, and reading emails using spoken commands.

(a) Speech-to-Text (STT)

Converts the user's speech into text using Google Speech Recognition API. Detects speech patterns and adjusts to the speaker's accent and pronunciation. Includes a noise-filtering mechanism to improve accuracy in noisy environments. The converted text is processed by NLP to extract meaning and intent.

(b) Text-to-Speech (TTS)

Uses the pyttsx3 engine to convert text responses into clear, human-like speech output. Provides customizable voice parameters (speed, tone, pitch, and gender). Offers offline functionality, ensuring uninterrupted service without internet dependency. Supports multiple languages for regional accessibility.

(c) Voice Command Interpreter

Interprets user instructions like "Compose new mail", "Read latest inbox", or "Delete message".

Utilizes NLP to understand the intent and context of commands. Includes confirmation prompts

3.2 Braille Interface

Although primarily voice-driven, the system can also integrate Braille-based interfaces to support users familiar with tactile communication

(a) Braille Displays

Converts text output into tactile Braille dots for real-time reading. Can display email text or menu navigation information. Connects via USB or Bluetooth for portability.

(b) Braille Keyboards

Allow users to compose or reply to emails using Braille input. Supports Grade 1 and Grade 2 Braille encoding, enabling both beginners and advanced users to operate the system.

3.3 Audio Navigation and Feedback

The system provides audio cues and navigation sounds for each action (like opening inbox, sending email, or logging in). Error sounds or verbal messages alert the user to incorrect input or failed actions. Provides real-time progress updates such as "Fetching your inbox…" or "Message sent successfully."

3.4 Accessibility Features

Adjustable Speech Rate and Volume: Allows users to personalize the speed and clarity of audio output.

Context-Aware Feedback: The system provides feedback relevant to the user's last command.

Help Mode: A special mode guides first-time users by describing available commands.

Auto Command Suggestion: If a command is incomplete, the system suggests possible actions (e.g., "Did you mean compose a new email?").

3.5 Multilingual and Regional Support

Integrates multilingual models for Hindi, Marathi, English, and other regional languages. Supports accent adaptation

through dynamic voice profiling. Enables language switching during runtime without restarting the system.

3.6 Future Integration Possibilities

Gesture or Head-Movement Commands: Integration with camera sensors for limited physical interaction.

Emotion Detection via Voice Tone: To enhance personalization and user comfort.

Haptic Feedback Devices: Vibrational alerts for incoming emails or system errors.

REFERENCES

[1] Prakash, S., Agarwal, K., Agarwal, G., & Singh, S. (n.d.). Voice Based Email System for the Visually Challenged. The study highlights improved accessibility and independence for visually impaired users by enabling voice-based email interactions.

[2] Kumar, S., Yogitha, & Aishwarya, R. B. (n.d.). Voice Email Based on SMTP for Physically Handicapped. This paper discusses a system enabling email operations via voice commands, enhancing accessibility and global appeal.

[3] Kumar, S., Malik, S., & Sharma, S. (n.d.). Voice-Based E-Mail System for Visually Challenged People.
The system boosts accessibility for users with visual impairments by simplifying email operations.

[4] Tiwari, P. A., Zodewar, P., Nimkar, H. P., Rokde, T., Wanjari, P. G., & Samarth, U. (n.d.). A Review on Voicebased E-Mail System for Blind. The paper discusses voice-based email systems with potential for integration into smartphones and support for multiple languages.

[5] Baptista, N., Prior, E., & Correia, M. E. (n.d.). Telephone Interface for the Email Service. This work introduces an
IVR system for email access via telephone, enhancing accessibility through features like automatic language detection

[6] Halle, P. and Shiyamala, S. (2019) "Architectural Integration for Wireless Communication Security in terms of integrity for Advanced Metering Infrastructure-Survey Paper", Asian Journal For Convergence In Technology (AJCT) ISSN -2350-1146. Available https://asianssr.org/index.php/ajct/article/view/771

[7] Halle, P. and Shiyamala, S. (2019) "Secure Routing through Refining Reliability for WSN against DoS Attacks using AODSD2V2 Algorithm for AMI," International Journal of Innovative Technology and Exploring Engineering. Blue Eyes Intelligence Engineering and Sciences Engineering and Sciences Publication - BEIESP. https://doi.org/10.35940/ijitee.I8178.0881019 (Scopus)

[8]Halle, P.D., Shiyamala, S.and Rohokale, Dr.V.M. (2020) "Secure Direction-finding Protocols and QoS for WSN for Diverse Applications-A Review," International Journal of Future Generation Communication and Networking, Vol. 13 No. 3 (2020) Available at: https://sersc.org/journals/index.php/IJFGCN/article/view/26983. (Web of Science)

[9]Halle, P.D. and Shiyamala, S. (2020) "Trust and Cryptography Centered Privileged Routing Providing Reliability for WSN Considering Dos Attack Designed for AMI of Smart Grid," International Journal of Innovative Technology and Exploring Engineering. Blue Eyes Intelligence Engineering and Sciences Engineering and Sciences Publication-BEIESP. doi:10.35940/ijitee.b7449.019320.

[10]Halle, P.D. and Shiyamala, S. (2021) Ami and its wireless communication security aspects with QOS: A Review, SpringerLink. Springer Singapore. Available at: https://link.springer.com/chapter/10.1007/978-981-15-5029-4_1 (Scopus: Conference Proceeding Book Chapter)

[11]Halle, P.D. and Shiyamala, S. (2022) "Secure advance metering infrastructure protocol for smart grid power system enabled by the internet of things," Microprocessors and Microsystems, 95, p. 104708. Available at: https://doi.org/10.1016/j.micpro.2022.104708 (SCI)

[12]Halle, P.D. and Shiyamala, S. (2022) "Internet of things enabled secure advanced metering infrastructure protocol for smart grid power system" SN Computer Science. (Scopus under review)