

# [9] nandini rawool.docx

*by* Dr. Sunil L. Bangare

---

**Submission date:** 24-Oct-2024 12:06AM (UTC-0700)

**Submission ID:** 2399487529

**File name:** \_9\_nandini\_rawool.docx (224.13K)

**Word count:** 2158

**Character count:** 12833

# Vocal Mail: An Accessible Email System for the Visually Impaired

## 1. Abstract

Despite the widespread use of email, visually impaired people encounter major barriers to accessing standard email systems since they rely on visual contact. In order to overcome these difficulties, a voice-based email system is presented in this work. The system makes the experience of visually impaired users smooth, accessible, and easy to use by leveraging face recognition for secure authentication, text-to-speech (TTS) processing, and speech-to-text (STT) using custom-trained datasets. By doing with external APIs, the system prioritizes increased privacy and autonomy. According to experimental evaluations, the method greatly increases accessibility and provides visually impaired people with a useful alternative for managing emails through voice commands.

<sup>3</sup>  
Keywords: Text-to-speech (TTS), Speech-to-text (STT), Face recognition, Voice commands

## 2. Introduction

Email is becoming a necessary tool for both business and personal communication. However, graphical user interfaces (GUIs) and input methods in traditional email systems necessitate visual interaction from users, which poses serious obstacles for those with visual impairments. While they may provide only partial answers, current assistive technologies—such as screen readers and Braille keyboards—are frequently unwieldy and ineffective. This study offers a voice-based email system that uses a mix of facial recognition, text-to-speech conversion, speech recognition, and natural language processing (NLP) for safe login, therefore overcoming these restrictions. The technology improves security and usability by allowing visually impaired users to connect with their email only through voice commands.

The datasets used in the development of the suggested system are particularly trained to recognize face cues. With the help of this innovative method, users may safely access their email accounts and manage their inboxes without requiring visual assistance.

<sup>1</sup>  
TABLE: TABLE BELOW SHOWS THE COMPARISON BETWEEN  
TRADITIONAL SYSTEM AND PROGRESSIVE SYSTEM OF EMAILING

S. No	Traditional System	Proposed System
1.	Compared to the progressive system, it is unstable.	It is more trustworthy because of its high level of security.
2.	This system demands the use of the keyboard.	A keyboard is not necessary since the system relies on IVR or Interactive Voice Response.
3	lengthy procedure (Slow execution).	It is quicker and more effective than a normal system.
4	The system can only be used by those who are disability free.	Users of the system can be both able-bodied and impaired.
5	People with visual impairments cannot communicate with others or use online email services.	The web application based email system will enable interaction for blind persons (using voice commands).

## Literature survey:

Voice-based email systems have become much more accessible for people with visual impairments due to new technology. Researchers have tried out different ways to make systems where people can send and receive emails just by talking. These systems use Speech-to-Text (STT), Text-to-Speech (TTS), and Natural Language Processing (NLP) to make it easier for everyone to use email, no matter what they can see.

Researchers (Rajput et al.) [2] have developed a user-friendly voice-based email system for people with visual impairments using AI and PyAudio. The system enables users to compose and send emails hands-free, using speech-to-text conversion. It also provides access to other services. AI-powered interfaces and NLP technologies allow the system to understand user commands and provide appropriate responses. This approach prioritizes ease of use, making the system accessible to individuals with technical limitations or vision issues.

Awasthi [3] and colleagues have developed a voice-activated email system aimed at making email accessible to visually impaired individuals. This system is designed to enhance the user experience by utilizing speech-to-text (STT) and text-to-speech (TTS) transformers, enabling users to manage their email using voice commands. The system eliminates the reliance on visual cues and keyboard navigation, addressing the challenges faced by visually impaired users, such as memorizing keyboard shortcuts and using mouse clicks. Instead, the voice-based system provides audio feedback and intuitive command responses, enhancing accessibility and ease of use.

Jain [1] and Suresh's [3] research aims to enhance the accuracy of voice-based email systems. They highlight the significance of Google's Text-to-Speech (gTTS) and speech recognition engines, enabling visually impaired individuals to access email effectively. Jain's system includes a user-friendly interface that accommodates both regular and visually impaired users, utilizing Python's speech recognition capabilities for voice input and standard email protocols (IMAP/SMTP) for email management.

Researchers Rajole and Phursule [1] identified that regular email systems often rely on visual displays, making them challenging for those with visual impairments. To address this issue, they developed a system that combines facial recognition for secure login and a voice-to-text feature that translates spoken commands into written text. This voice-activated system eliminates the need for manual interaction, allowing users to navigate email functions like inbox, sent items, and outbox solely through spoken commands.

In 2020, Sherly Noel developed a voice-controlled email application that uses human-computer interaction (HCI) to make it easier for visually impaired people to send and receive emails. This system gives users the ability to interact with their emails using voice commands, making it accessible to those who cannot use a graphical interface. Noel's system is a significant advancement in making technology more accessible to the visually impaired, highlighting the value of voice-based interactions in creating a positive user experience for this community.

While technology has made great strides, there are still challenges that need to be addressed. These include finding solutions for email attachments, improving the accuracy of speech recognition systems, and ensuring that these systems work with all assistive technologies. Additionally, even though these systems have shown to be beneficial for visually impaired users, ongoing testing, updates, and user training are essential to maintain and improve their effectiveness.

These systems promote inclusivity, but improvements are needed. Expanding language support and including features like file attachments and secure communication would enhance their accessibility. Wang et al. [2] suggest increasing security through hashing algorithms and better error-handling systems, leading to reduced errors and improved user experience.

### 3.Related Work

A number of methods have been put forth to improve visually impaired people's digital accessibility. Early systems made extensive use of keyboard shortcuts and screen readers, which, although useful, were not at all natural for non-technical users. Speech recognition software is one of the more modern methods, enabling voice commands to operate email operations. The majority of systems, however, still need some kind of visual interaction and frequently rely on third-party APIs, which poses privacy issues and restricts offline operation.

Dasgupta et al.'s work suggested voice-assisted systems that used simple text-to-speech modules, but it did not integrate security protocols—which are essential for private interactions. Similarly, Iglesias et al. built a haptic interface for visually handicapped people but did not focus on email systems. By combining voice commands, face recognition, and bespoke datasets, this research expands on the solutions that already exist and develops a more complete, safe, and usable system.

## 4.Methodology

### ➤ System Design

The suggested system's architecture is organized into four major modules: face recognition for authentication, speech-to-text (STT) for input, natural language processing (NLP) for command comprehension, and text-to-speech (TTS) for email output. Each module is created utilizing pre-trained models and specific datasets, resulting in excellent accuracy and efficiency.

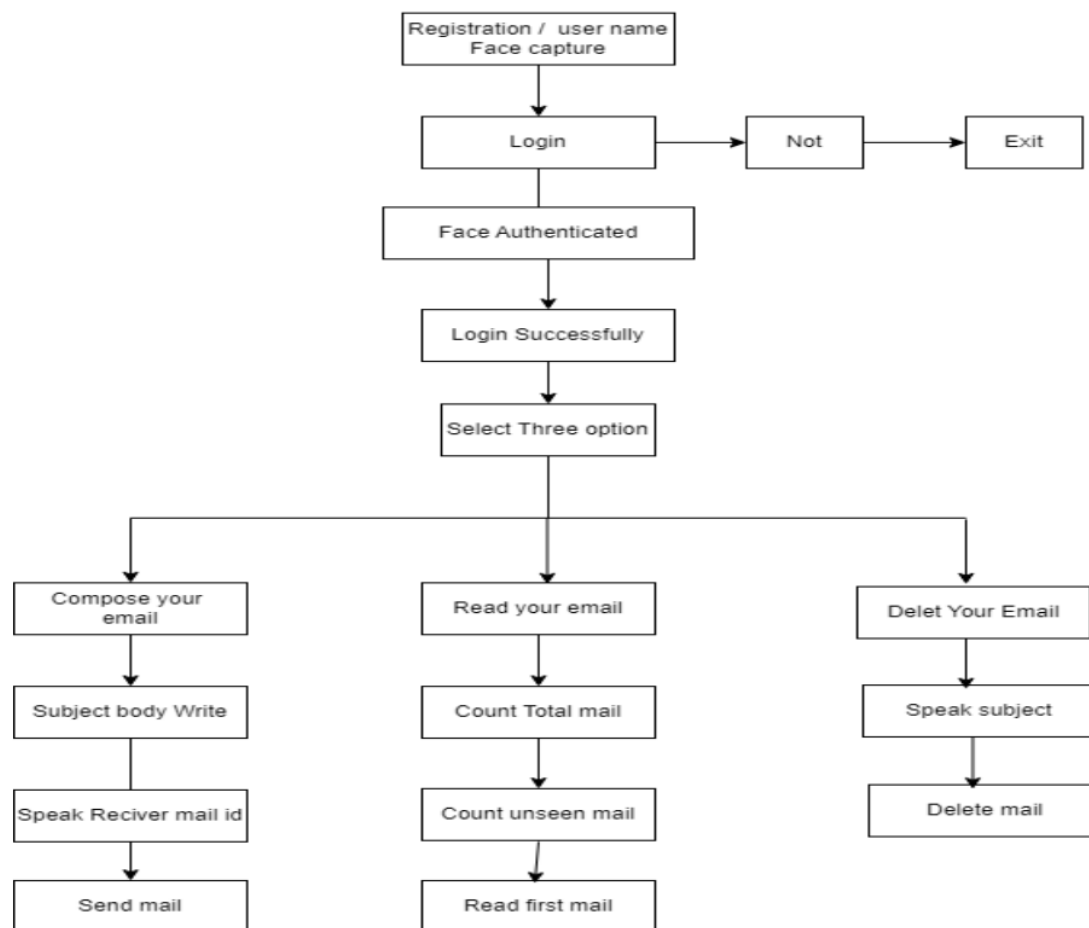


Fig. System Architecture diagram.

## ➤ Face Recognition for Secure Login

To ensure user privacy and security, the system incorporates **face recognition** as the primary authentication method. This eliminates the need for password-based or voice-based logins, which can be insecure or prone to eavesdropping. The face recognition module is built using the **OpenCV** library and **Haar Cascade Classifiers** for detecting and recognizing user faces.

The process involves the following steps:

1. **Data Collection:** The system collects facial images during the registration phase. These images are analyzed and stored in a database, with features extracted for facial recognition.
2. **Face Detection and Matching:** During login, a real-time image of the user is captured through the webcam. The system uses **Haar Cascades** to detect facial landmarks, followed by **Euclidean distance** matching to verify the identity of the user against the stored images.
3. **Authentication:** If the match confidence exceeds a set threshold, the user is granted access to their email account. This approach minimizes the risk of unauthorized access.

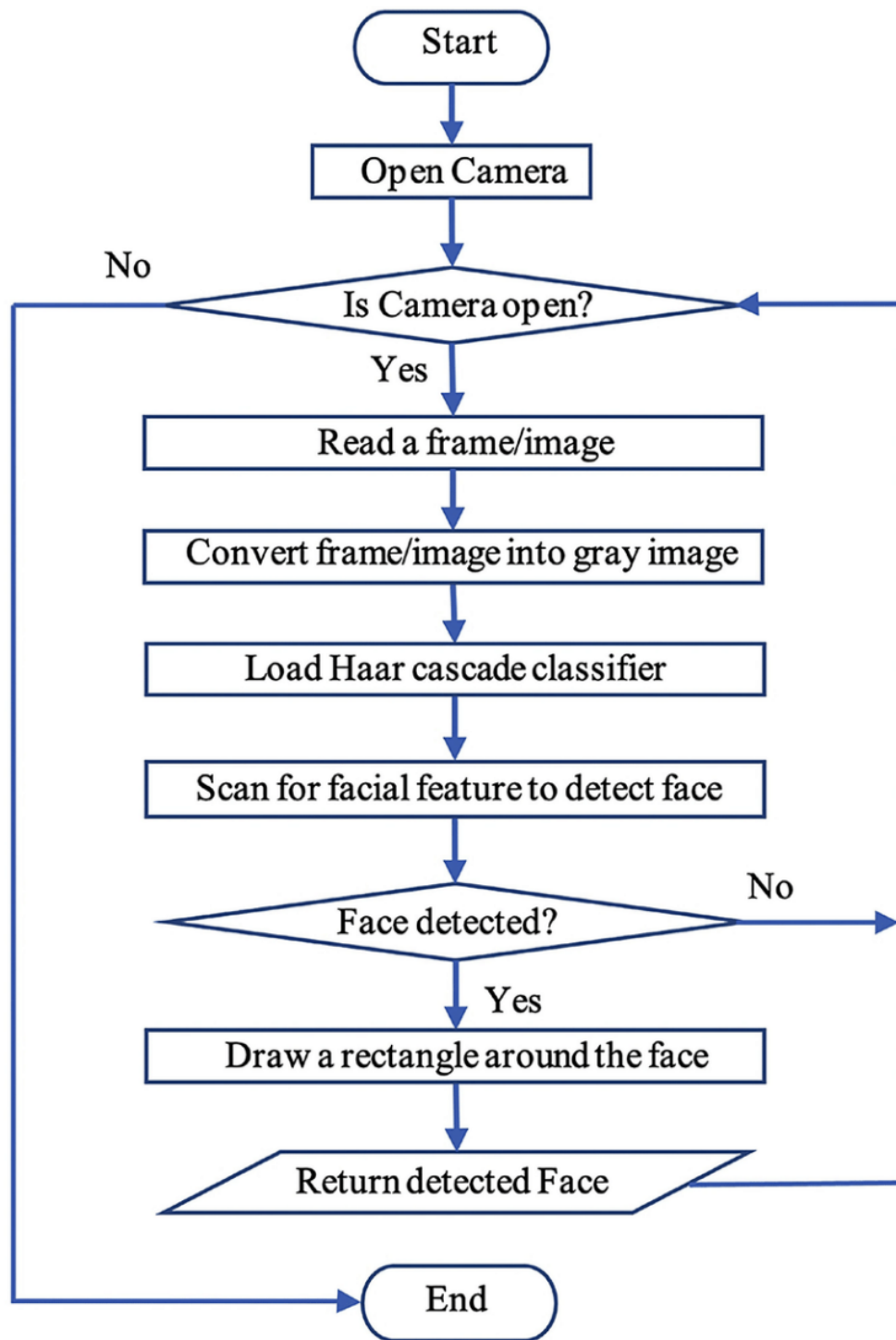


Fig. Face Detection

## ➤ <sup>5</sup> Speech-to-Text (STT) Module

The **Speech-to-Text (STT)** module allows users to input commands verbally. The system is built using datasets such as **LibriSpeech** and **Common Voice**, which are trained to recognize email-specific commands like "compose email," "send email," or "read inbox."

Key components of the STT module include:

1. **Preprocessing:** The audio data is filtered to remove noise and enhance clarity, ensuring that the system can recognize spoken commands even in environments with moderate background noise.
2. **Model Training:** The system uses **deep neural networks (DNNs)** to map audio signals to text. This is accomplished by training on a wide variety of speech samples, ensuring robust recognition of different accents and speaking styles.
3. **Command Parsing:** After converting speech to text, the system parses the command and determines the action to take. For example, saying "read inbox" will prompt the system to retrieve and read out the most recent emails.

## ➤ Natural Language Processing (NLP) for Command Execution

The **NLP module** plays a crucial role in <sup>7</sup>interpreting the user's spoken commands and converting them into actionable instructions. Built using **Recurrent Neural Networks (RNNs)** and **Long Short-Term Memory (LSTM)** layers, the NLP engine is trained on a custom dataset of email-related commands.

1. **Command Recognition:** The system breaks down the user's input into syntactic elements and identifies key actions <sup>1</sup>such as composing or replying to emails. It also processes additional details, such as **the recipient's email address or message content**.
2. **Action Mapping:** <sup>6</sup>The recognized command is then mapped to a specific action within the email system. If the user asks to "compose email," the system will initiate the email composition process

## ➤ <sup>8</sup> Text-to-Speech (TTS) Module

The **Text-to-Speech (TTS)** module allows users to listen to their emails. Built using **Tacotron 2**, the TTS model converts email text into human-like speech. The system reads out the sender, subject, and body of incoming emails, providing users with an easy way to access their inbox without visual interaction.

The module operates as follows:

1. **Text Preprocessing:** The email content is broken down into manageable segments, ensuring smooth and clear speech synthesis.



2. **Speech Generation:** The TTS model generates voice output, ensuring natural intonation and rhythm in spoken emails. Users can also reply or take action based on the content by providing additional voice commands.

## 5.Results and Discussion

### ➤ System Evaluation

The system was tested with visually impaired participants, who were tasked with logging in, composing emails, and navigating their inboxes through voice commands. Performance was evaluated based on the following criteria:

1. **Face Recognition Accuracy:** The face recognition module achieved a 98% success rate in well-lit environments and a 92% success rate in low-light conditions. Future enhancements could include infrared detection for improved low-light performance.
2. **Speech-to-Text Accuracy:** In quiet environments, the speech recognition system achieved an 86% accuracy rate, dropping to 78% in noisier settings. This demonstrates the need for further optimization in challenging acoustic environments.
3. **User Feedback:** Most users reported that the system was intuitive and easy to use, with the face recognition module providing a convenient and secure alternative to traditional login methods.

### ➤ Challenges

Despite the system's success, many obstacles were identified:

- **Lighting:** Accuracy of face recognition declined in dimly lit areas, despite good performance under ideal lighting. Infrared sensors or sophisticated lighting techniques could be used to address this.
- **Noise and Accents:** The voice recognition module has trouble in noisy surroundings and with some regional accents. Performance may be enhanced by adding more training data and using noise reduction.

## 6.Conclusion

The voice-based email system presented in this study is intended to improve accessibility for users who are blind or visually challenged. The system provides an effective and user-friendly interface for email management by fusing speech-to-text and text-to-speech technologies with face recognition for safe login. Utilizing datasets that have been specially trained guarantees privacy and lessens dependency on other providers. Future research will concentrate on enhancing the accuracy of speech recognition, increasing multilingual support, and incorporating attachment management.

## **8.References:**

- [1] Conferencing, Paging, Voice Mailing via Asterisk PBX – Ale Imran and Mohammed A. Qadeer [2019]
- [2] Finding with the design of a command- based speech interface for a voice mail system - S,Gamm and D.Langmann [2018]
- [3] Design and implementation of voice mail in IP telephony based on Mediastreamer2- Xu Yue and Xu Bing[2021]
- [4] On the integration of MPEG-4 video streams with voice and email data packet traffic over wireless picocell networks- P. Koutsakis, S. Psychic and M. Paterakis[2021]
- [5] Performance analysis of Voice message service in CDMA cellular systems”- Song Liu, Zhisheng Niu and Dawei Huang [2020]
- [6] Harivans Pratap Singh, Aman Pratap Kushwaha, Aayushmaan, Harendra Singh, “Voice-Based Email System”. International Journal of Innovative Science and Research Technology Volume 6, Issue 7, July – 2023
- [7] Mullapudi Harshasri , Manyam Durga Bhavani , and Misra Ravikanth, “Voice Based Email for Blind”. - International Journal of Innovative Research in Computer Science & Technology (IJRCST) ISSN: 2347-5552, Volume-9, Issue-4, July 2021

ORIGINALITY REPORT

---

13%

SIMILARITY INDEX

11%

INTERNET SOURCES

3%

PUBLICATIONS

1%

STUDENT PAPERS

---

PRIMARY SOURCES

---

1

[assets.researchsquare.com](https://assets.researchsquare.com)

Internet Source

7%

---

2

[ijarcce.com](https://ijarcce.com)

Internet Source

1%

---

3

N Kavi Prakash, Tenzin Monlam, Rohan Singh, RP Aravindhan, Pentapuri Vishnuvardhan Reddy, Richa Jain. "Voice based E-Mail with Attachment for Blind", 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI), 2022

Publication

1%

---

4

[www.patenthub.cn](https://www.patenthub.cn)

Internet Source

1%

---

5

[ebin.pub](https://ebin.pub)

Internet Source

1%

---

6

Submitted to October University for Modern Sciences and Arts (MSA)

Student Paper

1%

---

7

[www.researchgate.net](https://www.researchgate.net)

Internet Source

1%

---

Rajesh, Sangeetha, and Lifna C.S. "UI design for Language Translator Module in Swasthya slate (m-Health tool)", 2015 International Conference on Advances in Computing Communications and Informatics (ICACCI), 2015.

Publication

---

<1 %

---

Exclude quotes Off

Exclude matches Off

Exclude bibliography On