IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 1, January 2024

DOI: 10.17148/IJARCCE.2024.13125

# VOICE-OPERATED APPLICATIONS LIKE GMAIL, FOR VISUALLY IMPAIRED PERSON

**Mr. Vishnu Parsewar[1], Mr. karan kakade[2], Mr. Aditya Ubale[3], Ms.Radhika Malpani[4]**

Student, MMCOE, Pune[1-3]

Assistant professor, MMCOE, Pune[4]

**Abstract:** In today's interconnected world, communication technologies play a pivotal role in enabling social interactions and fostering connections. The integration of these technologies with the internet has revolutionized communication, expanding its reach and capabilities. However, individuals with physical disabilities often encounter challenges in utilizing these technologies due to visual and motor impairments. Despite technological advancements, many solutions remain inaccessible to a significant portion of the population. Our system aims to bridge this gap by simplifying communication for novice and physically disabled users, eliminating the need for prior training. The system's innovative approach utilizes voice conversion as the sole input method, eliminating the need for keyboard interaction. This feature caters to individuals with limited digital literacy, enabling them to send emails with ease and confidence. The system's user-centric design is further enhanced by its responsive voice interaction approach, ensuring a seamless and intuitive user experience. The system encompasses all the functionalities required for email composition and management, providing a comprehensive solution for effective communication. Upon successful login, the system engages in an interactive dialogue with the user, prompting them to select from available actions such as composing, reading, or deleting emails. For email composition, the system guides the user through the process, requesting the recipient's email address and the desired subject line.

**Keywords:** Internet, Voice, Speech recognition, read email, delete email, compose email.

## I. INTRODUCTION

Communication is most important these days because the way we write letters has completely changed. What followed was the historic invention of mankind so far - telephone communication, which greatly reduced the time required to communicate between two people. When we communicate over the phone, we not only share words, we also share our feelings with others. Subsequently, humanity witnessed a groundbreaking and innovative digital system. Researchers are rapidly updating and developing use cases for digital systems. One way is to communicate via text messages, emails, video calls, etc.

Today, email systems are a means of communication that are likely to be used as official means of communication for large industrial enterprises and other governmental and non-governmental agencies. These systems are much easier for us to use than it is for people with visual impairments. Based on intent and target group, i.e.

This product is designed to serve healthy individuals. For people with physical impairments, using this system can cause headaches. Taking into account all these disadvantages or problems, we have developed a solution to this problem: a "Voice Operated Email System" that allows for easier operation and privacy while using it.

The system consists of multiple modules, one of which is responsible for ensuring the secure and successful login and registration of the users. Privacy is a major concern for this type of system, so we have added a face lock feature to the system that verifies the identity of the users while logging in. The system also alerts the users about any login attempts made In addition to the primary modules, there is an auxiliary module dedicated to managing email communications.

This module empowers users to seamlessly interact with their email accounts, encompassing a range of functionalities. It facilitates reading incoming emails, composing new messages, and even deleting unwanted emails, ensuring efficient email management. This module serves as a valuable tool for streamlining email tasks and enhancing user productivity.

  

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 1, January 2024
DOI: 10.17148/IJARCCE.2024.13125

## II. LITERATURE REVIEW

**a.** The proposed email system is designed with user accessibility and ease of use as its top priorities, making it suitable for both able-bodied and disabled individuals. This innovative system addresses a broader user base than existing solutions, including illiterate individuals and first-time users. To facilitate natural and intuitive interaction, the system utilizes IVR (interactive voice response) technology. Upon user interaction, the system generates corresponding voice prompts to guide users through the desired actions. This step-by-step approach eliminates the need for keyword input and minimizes mouse usage. A key advantage of the system lies in its straightforward operation, requiring only a single mouse click to initiate the system. Thereafter, every step is communicated through voice prompts, allowing users to focus on responding to desired actions without the burden of mouse navigation. Each action performed is accompanied by voice feedback, providing users with clear status updates and explanations in case of successful or unsuccessful execution.

**b.** This paper introduces a user-friendly audio-based tool designed to help visually impaired individuals determine whether a room's light is on or off. Despite the prevalence of Information and Communication Technology (ICT)-based solutions, these can be challenging for visually impaired individuals due to the requirement for digital device familiarity and the complexity of repetitive tasks. The proposed tool, which requires no specific skills or knowledge from the user, addresses these issues. It utilizes a electronic board and gives response user about the light's status. Two different models have been developed for different types of installation purposes, with the second prototype have three versions for enhanced support in identifying the light status. The new design features improved response and dimension modifications. Evaluations involving various end-user groups have demonstrated the tool's effectiveness. A survey of 100 visually impaired individuals highlighted the challenges they face using existing devices and showed that 94% of participants were interested in a new basic tool that could be integrated with the existing lighting system. This study contributes to the field of ambient intelligence by demonstrating how an auditory-based tool can autonomously and simply assist visually impaired people in checking lights. It proposes an idea that could be applied to other cases using light feedback.

**c.** Three Catering specifically to the needs of visually impaired individuals, this methodology presents a voice-based email application. Conventional email applications typically rely on a Graphical User Interface , which proves challenging not only for visually impaired users but also for physically disabled to navigate. The proposed application tackles this accessibility hurdle by adopting voice commands as the primary interaction method. The application's architecture comprises five core modules: user voice input processing, voice comparison against a pre-recorded sample, text generation, output text verification, and email sending. Voice commands serve as the primary means of user input. Upon user sign-in or registration, the application verifies the user's identity by comparing their voice against a reference sample securely stored in a database. If unauthorized access is detected, the application reverts to the sign-in stage. The application maintains an active listening state, awaiting user voice commands such as "compose email" or "read email." Upon receiving the "compose email" command, the application initiates the email composer module. Users can then dictate the email content using voice commands such as "edit to," "edit subject," and "edit body." The Speech-To-Text (STT) interface seamlessly converts the user's speech into text and inputs the content into the corresponding fields. After composing the email, users can have the email read back to them using the "read email" command. The composed email is then sent to the intended recipient upon receiving the "send email" command. Conversely, if the user issues the "read email" command, the application retrieves the relevant email and utilizes Text-To-Speech (TTS) technology to read the email content aloud to the user. Users can sign out of the application at any time.

**d.** The widespread use of the Internet has led to the development of various communication applications, with email being the most prevalent and reliable method. However, visually impaired individuals face challenges using traditional email systems due to their reliance on visual cues. To address this issue, a voice-controlled email system has been developed to provide visually impaired individuals with easy and accessible email communication. This framework can also benefit individuals with other disabilities.The voice-controlled email system utilizes Interactive Voice Response (IVR) technology, which enables interaction between humans and computers using voice and Dual-Tone Multi-Frequency (DTMF) tones. IVR systems are commonly used in telecommunications, allowing customers to interact with a company's system via voice or keypad input.The system employs Text-To-Speech (TTS) technology to convert text into spoken language, providing visually impaired users with auditory feedback. TTS has evolved beyond the robotic voices typically associated with computers, now utilizing real voices synthesized in segments and smoothed before playback.Speech recognition (SR) technology enables the system to interpret human speech and convert it into text, allowing users to compose and send emails using voice commands. SR systems employ acoustic modeling to analyze the acoustic properties of speech, such as pitch, frequency, and duration, to identify individual phonemes.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 1, January 2024
DOI: 10.17148/IJARCCE.2024.13125

Additionally, language modeling considers the statistical relationships between words and phrases to determine the most likely sequence of words given a particular acoustic input.The development of big data and deep learning algorithms has significantly enhanced the accuracy and robustness of SR systems, making them more practical for real-world applications. However, SR still faces challenges, including the variability of human speech, which can be affected by factors such as accent, pronunciation, and background noise. Additionally, SR systems are typically trained for specific languages or domains, limiting their applicability to other contexts.Despite these challenges, the voice-controlled email system represents a significant advancement in making email communication accessible to visually impaired individuals and those with other disabilities. As technology continues to evolve, we can expect to see even more innovative and user-friendly solutions in the future.The research studies examined provide a means for visually impaired individuals to access email, a crucial communication tool in today's world. The proposed framework aims to reduce the obstacles encountered by visually impaired individuals, such as the need to memorize and employ mouse clicks and keyboard shortcuts for email access. Consequently, we propose a voice-based authentication system instead of the conventional username and password approach

**e.** This approach introduces an Android application that addresses the limitations of existing mobile apps, which are often designed with sighted users in mind and disregard accessibility standards. These apps fail to accommodate the unique needs of blind individuals, leading to various challenges in single-device and multi-device interactions, user interface synchronization, consistent output formats, output redirection, and maneuvering user interface components such as buttons, layouts, and navigation menus. The introduction of new features further complicates the learning process for blind users, requiring them to continuously adapt their experiences to each new application and device. This study presents an effective solution in the form of a semantically enriched, context-aware, and simplified user interface design that seamlessly integrates with multiple devices, including smartphones and smartwatches. The study's findings reveal that this redesigned interface significantly reduces cognitive strain and enhances the user experience for blind individuals. Initial results indicate that the proposed solution improves the user experience for blind people when performing common tasks on smartphones and smartwatches.

## III. PROPOSED SYSTEM

Visually impaired users (VIU) often face challenges in using email applications that rely heavily on graphical user interface (GUI). Most of the GUI features are designed for normal users and do not provide adequate accessibility or usability for VIU. Therefore, there is a need for technological innovation that can address the specific needs and preferences of VIU in email communication.

We present a novel approach to develop an application for sending emails that can accommodate the needs of visually impaired users. Our application leverages speech-to-text and text-to-speech technologies to enable voice-based communication via email. This application is not limited to visually impaired users, but can also benefit normal users who prefer voice input and output. The application operates through simple and user-friendly voice commands that can be customized according to the user's preferences.

The proposed system consists of several modules which work together to serve the visually impaired people.

**a.** The login module, Traditional user account security measures often prioritized user experience over user protection, leading to vulnerabilities and unauthorized access. To address this, face lock technology has been introduced as an additional layer of security, complementing traditional password or PIN authentication. During the registration process, the system captures the user's face and prompts the user to enter personal information. This captured face image serves as a unique identifier and is essential for subsequent login attempts. To capture and recognize user faces accurately, the system employs the MTCNN algorithm, which effectively detects and aligns facial features, even under challenging conditions. Face lock offers several significant benefits for user security, including enhanced protection against unauthorized access, convenience, and accuracy. Overall, face lock represents a significant improvement in user account security, providing a convenient and effective way to prevent unauthorized access and protect user data.

**b.** Upon successful login, the voice-based email assistant greets the user and initiates a conversation to understand their desired actions. The assistant can assist with various email tasks, including reading emails, composing new emails, and deleting unwanted messages. These operations are performed entirely through voice commands, eliminating the need for keyboard or mouse interaction. This voice-based interface caters to users who may face challenges with traditional input methods, enhancing accessibility and providing a user-friendly email experience. The assistant's ability to recognize and respond to voice commands provides a convenient and efficient way to manage emails. Users can simply speak their instructions, and the assistant will execute the corresponding actions, making it easier to navigate their inbox and perform various email tasks.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 1, January 2024

DOI: 10.17148/IJARCCE.2024.13125

This hands-free interaction also allows users to multitask while managing their emails, further enhancing productivity. Overall, the voice-based email assistant represents a significant advancement in email accessibility, enabling users to communicate and manage their emails effectively using voice commands. This technology has the potential to revolutionize the way people interact with email, particularly for those with disabilities or who prefer a hands-free approach.

## IV. ALGORITHMS & LIBRARIES

**i.**MTCNN is a deep learning-based framework that utilizes a cascade of convolutional neural networks (CNNs) to detect and align faces in digital images and videos. It is capable of accurately detecting faces of different scales and orientations, even in challenging lighting conditions and with partial occlusions. The MTCNN framework consists of three main stages:

1. Proposal Network (P-Net): The P-Net is a shallow CNN that generates a set of candidates bounding boxes that may contain faces
2. Refine Network (R-Net): The R-Net takes the candidate bounding boxes from the P-Net and refines them using a more complex CNN. This stage further reduces the number of candidates and improves the accuracy of the face detection.
3. Output Network (O-Net): The O-Net takes the refined bounding boxes from the R-Net and outputs the final face detection results, including the bounding box coordinates and five facial landmarks (left eye, right eye, nose, left mouth corner, and right mouth corner).

MTCNN has been shown to be very effective for face detection and alignment, and it is widely used in a variety of applications, such as facial recognition, face tracking, and augmented reality

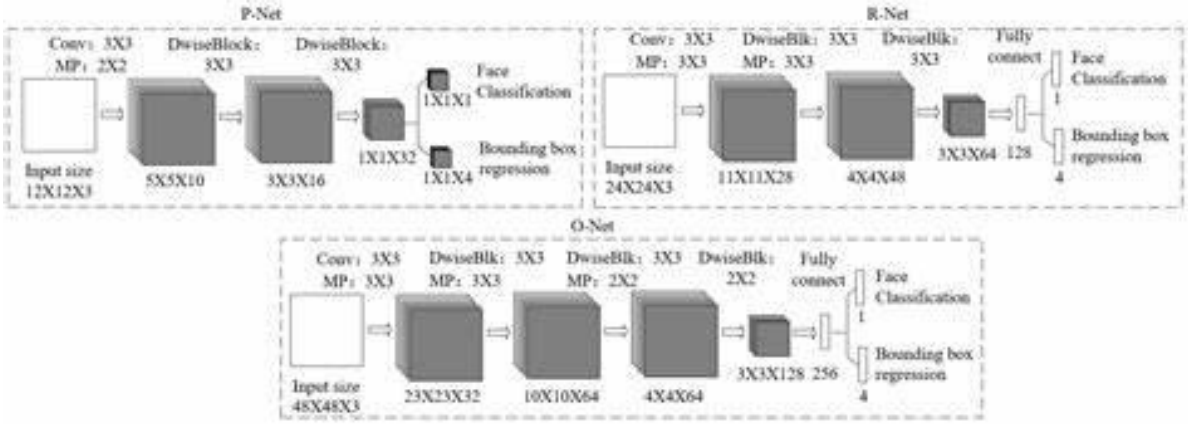


Figure 1: MTCNN Process Flow

**ii.NLP:** Natural Language Processing: - NLP algorithms often employ machine learning techniques. Rather than manually crafting extensive rule sets, NLP can harness machine learning to automatically derive these rules by analyzing a collection of examples and drawing statistical inferences. Natural Language Processing (NLP) endeavors to develop machines capable of comprehending and responding to text or voice data, producing text or speech in response.

1. **Speech Recognition**: This library enables speech recognition tasks, offering compatibility with different engines and APIs, for both online and offline use. It encompasses a wide range of speech recognition engines/APIs, including Google Speech Recognition, Google Cloud Speech API, Microsoft Azure Speech, IBM Speech to Text, and others.

2. **Google Text-to-Speech:** Harnessing the power of cloud computing, Google Text-to-Speech seamlessly transforms written text into natural-sounding speech. Leveraging advanced technologies, it generates voices that closely resemble human speech, providing a diverse selection of voice options across multiple languages and their variants. To enhance the speech output, you can utilize SSML tags to introduce pauses, numerals, date and time formatting, and other pronunciation instructions. Moreover, the GTTS (Google Text-to-Speech) Python library acts as an interface to Google Cloud's text-to-speech API. This library empowers you to generate spoken mp3 data directly to a file, a file-like object (byte string) for further audio manipulation, or stdout (standard output).

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 1, January 2024

DOI: 10.17148/IJARCCE.2024.13125

## V. SYSTEM ARCHITECTURE

The software begins with a user login process, where it captures the user's face for authentication. If the face is recognized, the user is logged into their account. If the authentication fails, the system retries the authentication process. Once logged in, the system interacts with the user and offers actions such as reading emails, deleting emails, and composing emails.

If the user chooses to compose an email, the system interacts with the user to obtain the recipient's email address and the content to be sent. If the user opts to read an email, the system starts by reading the first unread email from the list. If the user selects the delete email option, the system asks the user which email they want to delete obtains the address and subject of the email, and proceeds with the deletion. This process ensures a user-friendly and efficient email management system for the user.

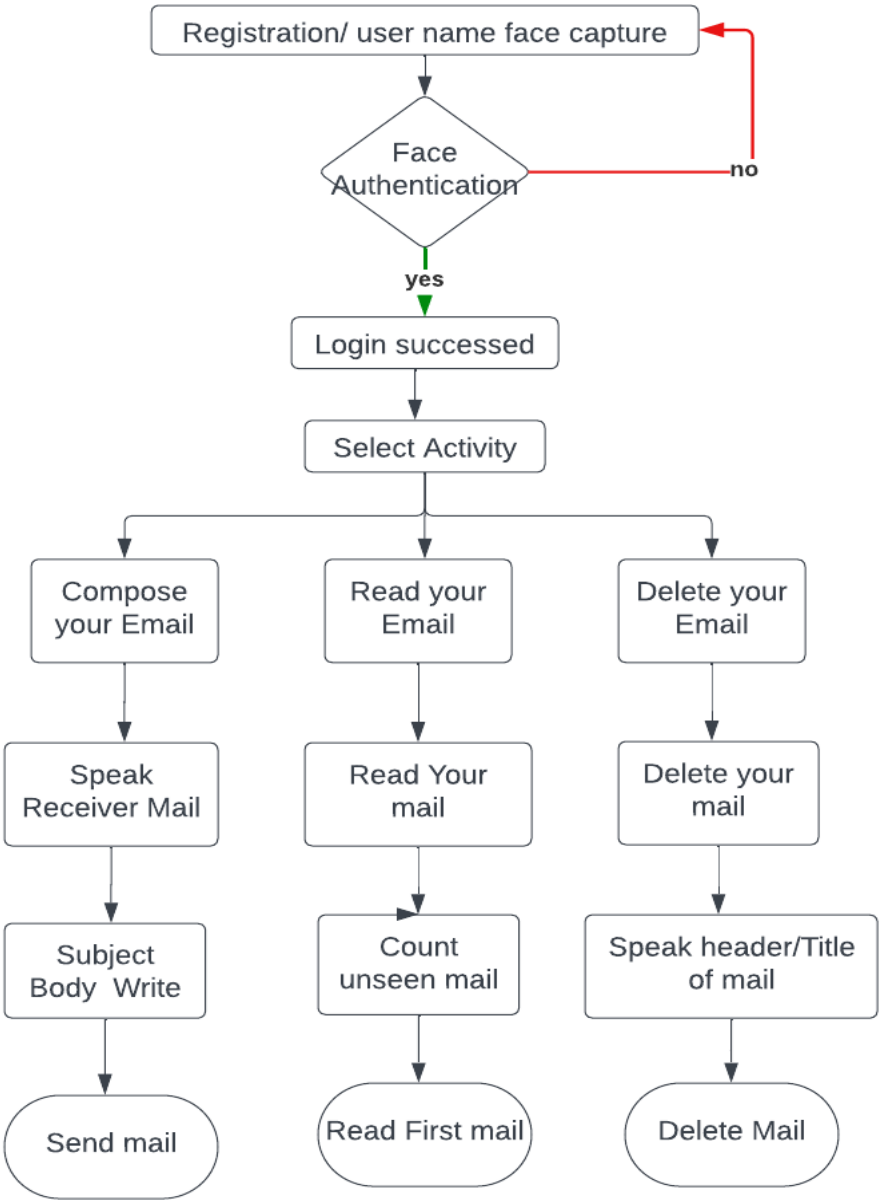


Figure 2 System Architecture

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 1, January 2024

DOI: 10.17148/IJARCCE.2024.13125

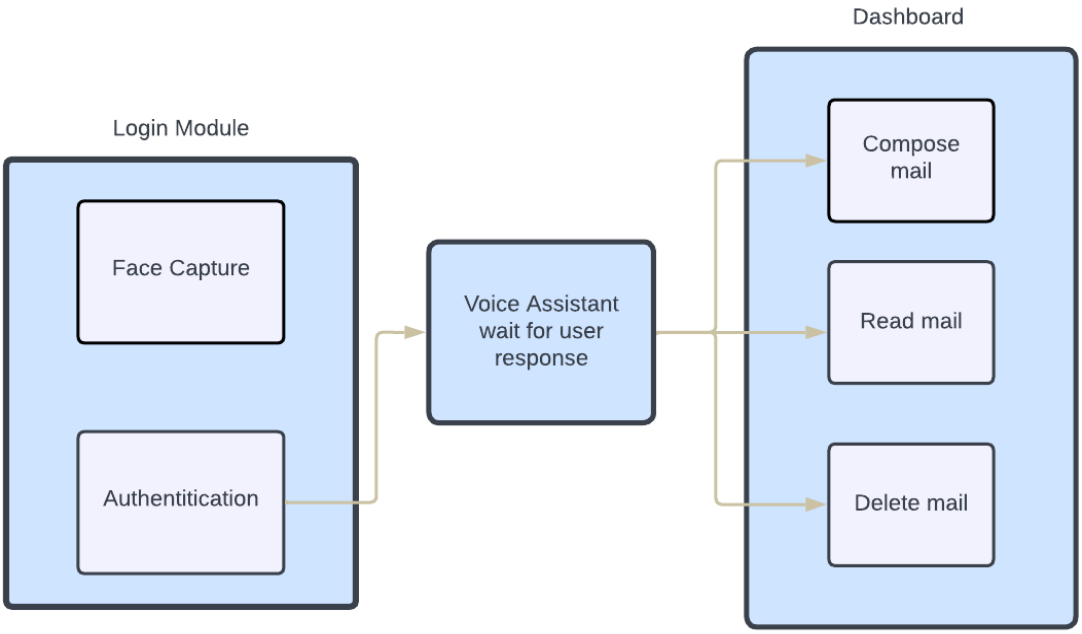


Figure 3: Block Diagram

## VI. OBJECTIVE

**a.** Implement voice-based email composition, sending, and reception to empower users with hands-free email management.

**b.** Provide feedback to users about their actions, such as confirming the recipient's email address or reading the contents of an email aloud.

**c.** Be easy to use and navigate, with clear and concise voice prompts.

**d.** Be affordable and accessible to visually impaired people of all income levels.

## VII. CONCLUSION

In summary, the implementation of a voice-activated email system for visually impaired individuals can revolutionize their communication and social engagement. This system, powered by cutting-edge speech recognition and text-to-speech technology, empowers users to compose, send, and receive emails solely through voice commands, eliminating the need for conventional reading and typing. By fostering independence and self-sufficiency among visually impaired individuals, this system promotes inclusivity and accessibility in the digital era.

Future advancements may include multilingual support and integration with other communication platforms to expand its usability further. However, it is crucial to continuously address potential challenges, such as privacy concerns and speech recognition accuracy, to guarantee a smooth user experience. Overall, the development and adoption of this system highlight the transformative power of technology in bridging accessibility divides.

## REFERENCES

[1]. "Voice Email Based On SMTP For Physically Handicapped" Sunny Kumar, Ms. Yogitha. R, Ms. R. Aishwarya[2021]

[2]. “Designing assistive technology for getting more independence for blind people when performing everyday tasks: an auditory-based tool as a case study” Barbara Leporini, · Michele Rosellini, · Nicola Forgione[2021]

[3]. “Human computer interaction(HCI) based Smart Voice Email (Vmail) Application - Assistant for Visually Impaired Users (VIU)” Sherly Noel[2020]

[4]. Performance analysis of Voice message service in CDMA cellular systems”- Song Liu, Zhisheng Niu and Dawei Huang. [2020]

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 1, January 2024

DOI: 10.17148/IJARCCE.2024.13125

[5]. A Voice Communication-Augmented Simulation Framework For Aircraft Trajectory Simulation- Yuhao Wang, Michael Thomas and Yutian Pang[2021]
[6]. Design and implementation of voice mail in IP telephony based on Mediastreamer2- Xu Yue and Xu Bing[2021]
[7]. Transfer Learning From Speech Synthesis to Voice Conversion with Non-Parallel Training Data- Mingyang Zhang, Yi Zhou and Li Zhao [2021]
[8]. Discovery of Activities Actor Perspective from Emails based on Speech Acts Detection- Marwa Elleuch, Nassim Laga, Walid Gaaloul and Oumaima Alaoui Ismaili [2020].