

# **Voice Based Email System for Blind with Image Captioner**

Name: Nandana M USN No.: 1KG21CS070 Email id: nandanamurthy6203@gmail.com

Name:Dhanush U S USN No.: 1KG21CS032 Email id: dhanushupallapaate@gmail.com Guide 1 Name:Nagaveni B. Nimbal Email id:nagaveni@kssem.edu.in Guide 2

Name: Prof. Jayashubha J Email id: jayashubha@kssem.edu.in KS School of Engineering and Management Department of Computer Science and Engineering

#### **Keywords:**

Voice based email, Optical character recognition, image caption generator, Text to speech, speech to text, flask, inception v3, Tesseract OCR.

#### Introduction:

The Voice-Based Email System is a specially designed application aimed at supporting visually impaired individuals by providing a user-friendly way to access and understand email content. This innovative system focuses on enhancing accessibility by enabling users to listen to their email messages through voice output. The main feature of the system is its ability to read out the first email message received by the user. This ensures simplicity and focus, allowing blind users to access the most recent communication without navigating through multiple emails. A unique and powerful component of this application is its image captioning functionality. When an email includes an image, the system uses a pre-trained image captioning model to analyze the content of the image and generate a descriptive caption. This caption is then converted to speech, allowing the blind user to understand the visual content without needing to see it. The use of AI-driven image captioning bridges the gap between visual and audio information, creating a more inclusive digital communication experience. This solution is particularly valuable as it combines natural language processing, text-to-speech technology, and computer vision to provide a seamless experience. The system helps blind users recognize and interpret both textual and visual information from emails, improving their independence and communication. By focusing on accessibility and ease of use, the Voice-Based Email System becomes a practical tool that empowers the visually impaired to stay connected in the digital world.

L



Ť					
User	Application	Openc	Flickr30kData	iset inceptionV3	LSTM
Upload Ima	Prepr Retric	ocess Image ave Image Caption ge Caption ct Image Features ge Features rate Caption tion	<b>&gt;</b>	<b>&gt;</b>	
User	Application	Openc	Flickr30kData	set InceptionV3	LSTM

Figure 1: Sequence diagram for voiced based email system with image captioner

## **Objectives:**

- Design a UI/UX design front page which recognize voice.
- Integrate Text-to-Speech (TTS) and Speech-to-Text (STT) for seamless email interaction. Make the system usable for both blind and other disabled users.
- Enabled email functionalities: send, read, search, and spam.
- Enable visually impaired users to send, receive, read, delete and identify the image captions emails independently. Provide a multi- lingual voice-based interface for better accessibility.

## Methodology:

The project can be built as a standalone application either for desktop or mobile devices, making it convenient for users across platforms. For desktop development, the framework Electron can be used, which allows developers to create powerful desktop applications using web technologies like HTML, CSS, and JavaScript. On the other hand, for mobile platforms, React Native stands out as a suitable framework due to its cross-platform nature and support for native mobile performance. Another practical and flexible approach is to make the system a web-based solution, hosted on a cloud platform that ensures seamless access and scalability. In this model, serverless functions play a crucial role by handling real-time tasks such as image captioning and email generation, eliminating the need for managing traditional servers. For better performance and scalability, the machine learning models can be deployed using cloud solutions like AWS, Azure, or Google Cloud, which offer robust infrastructure and advanced ML support. The main programming languages to be used in this project are Python and JavaScript, where Python handles ML operations and backend logic, while JavaScript takes care of the user interface and frontend interaction. To build APIs and handle ML processing, efficient frameworks such as PyTorch, TensorFlow, and FastAPI will be utilized. For tasks like image understanding, pre-trained models like CLIP, BLIP, and those from Hugging Face will be highly effective, enabling the system to analyze and describe images intelligently. To further support backend services and ensure smooth execution, cloud tools such as AWS Lambda, Azure AI, and Google Cloud Vision API will be integrated, helping process and deliver features efficiently without overloading the user's device. This combination of tools and technologies makes the system scalable, responsive, and future-ready.

## **Result and Conclusion**:

In conclusion, this research is centered on assisting visually impaired individuals in accessing and understanding content within their emails, particularly focusing on images that are often overlooked in existing systems. Despite the advancement of technology, email remains one of the most widely used communication methods globally. However,

L

traditional email platforms fall short when it comes to helping visually impaired users identify or interpret images embedded in emails. This gap creates a communication barrier and adds unnecessary complexity and frustration to the user experience. The proposed framework in this project is designed specifically to eliminate these challenges and provide a more inclusive, intuitive solution. It introduces a voice-based system that not only reads the first email message aloud but also generates and speaks a caption for the image attached, allowing users to understand visual content through audio. The system goes a step further by associating images

with user IDs, making recognition more structured and simplified for the user. This approach enables image identification within emails to become a user-friendly and accessible experience. The aim is not just to improve accessibility, but to empower visually impaired users with greater independence and confidence while interacting with digital communication. Overall, the system reflects a meaningful step toward a more inclusive technology landscape where everyone can communicate effectively, regardless of visual ability.

# **Future Scope:**

The future scope of this project includes:

- Research focuses on enabling visually impaired individuals to access emails and identify the images.
- Email remains one communication globally. of the most widely used forms of communication globally.
- Current methods for visually impaired users can't able to identify the images which can be problematic and cumbersome.
- The proposed framework reduces barriers to email access.
- Develops a platform in which the visually impaired individuals can be able to recognise the image by knowing the ID.
- Designed to simplify image identification in email for visually impaired users.

L