INNOMATICS®
RESEARCH LABS

INNOVATION. AUTOMATION. ANALYTICS

PROJECT ON

AI Assistant for Visually Impaired

By : Prabin Basyal

Introduction

The AI Assistant for the Visually Impaired is a web-based tool designed to assist individuals with visual disabilities by offering real-time guidance and information through advanced technologies like machine learning, OCR (Optical Character Recognition), and Text-to-Speech. By leveraging cutting-edge models from LangChain and Google Vision, this assistant provides useful features such as describing scenes, identifying objects, detecting obstacles, and guiding users through tasks in an intuitive, accessible format.

Objective

The main goal of this project is to provide a robust solution that enables visually impaired individuals to navigate and interact with their surroundings more independently. The application's core objectives are:

- To offer detailed descriptions of scenes and images.
- To detect objects and obstacles, assisting with navigation.
- To provide personalized task-specific help, such as identifying items or reading labels.
- To convert text into speech, making the information audible and more accessible.

Problem

Visually impaired individuals face significant challenges in everyday tasks, especially when it comes to interpreting visual information. Whether it's reading signs, navigating unfamiliar spaces, or identifying objects, there is a clear gap in the tools available to help them in real-time. Existing solutions often fall short by providing insufficient context or being limited in scope, which hampers their independence and mobility.

Solution

This AI-powered assistant aims to bridge this gap by offering a solution that not only provides a description of the surroundings but also identifies objects, detects obstacles, and helps with specific tasks. Through seamless integration of OCR and AI models, the system analyzes images and provides detailed outputs in both text and audio formats. These features allow users to make informed decisions and carry out their daily activities more confidently and independently.

Code Breakdown

Here's a breakdown of the main components of the application:

- **handle_error**: This function is responsible for catching and logging any errors that occur during the execution of the application. It ensures that any issues are properly recorded and an appropriate message is displayed to the user.
- **scene_understanding**: This function takes an image as input and provides a comprehensive description of the scene. It includes details about the layout, colors, objects, and activities within the image. This feature aims to help users understand what is happening in the image by providing them with a clear, structured response.
- **detect_objects_and_obstacles**: This function identifies potential hazards or obstacles in the image, offering guidance on how to navigate the environment safely. It calculates distances and suggests safe paths for movement.
- **provide_task_assistance**: This function provides specialized assistance depending on the selected task. Whether it's identifying items, reading labels, or offering navigation tips, this function tailors its response to suit the specific needs of the user.

Code Breakdown

- **text_to_speech**: This feature converts text-based outputs into audio, using Google's Text-to-Speech technology. It enables visually impaired users to listen to the assistant's responses, providing an auditory feedback loop for better interaction.
- **main**: The core of the application, this function sets up the user interface with Streamlit. It handles user interactions such as uploading images and selecting different features. It also displays the results from the other functions in an accessible format.

How to Use

Using the AI Assistant is simple and intuitive. Here's how to get started:

- **Upload an Image**: Upload an picture or image using the file uploader. This could be any image that contains visual information you want to analyze, such as a scene, a label, or a document.

How to Use

- **Choose a Feature**: Select a feature based on the type of assistance you need. You can choose from:
  - Scene Description: Get a detailed description of the image.
  - Object Detection: Detect objects and obstacles within the image.
  - Task Assistance: Receive specific help with tasks like identifying items or navigating spaces.
- **Get Results**: After selecting a feature, the assistant will analyze the image and provide the results in text format. Additionally, audio feedback will be provided, allowing you to listen to the results.

Conclusion

The AI Assistant for the Visually Impaired is a powerful tool that leverages modern technologies to provide real-time, detailed assistance for visually impaired individuals. It offers a blend of image recognition, object detection, scene understanding, and task-specific guidance, all designed to make the world more accessible. This application empowers users to interact with their environment more confidently and independently, offering a new level of autonomy.

Future Enhancement

While this AI Assistant is already a significant step forward, there are several opportunities to further enhance its capabilities:

1. **Multilingual Support**: Expanding the language options for both image descriptions and speech output would make the tool accessible to a broader audience.
2. **Improved Object Detection**: The object detection model can be enhanced to recognize a greater variety of objects and hazards, further improving navigation and safety.
3. **Real-Time Video Assistance**: Integrating real-time video analysis could allow the assistant to provide continuous support as the user moves through their environment, offering immediate guidance and feedback.
4. **Voice Command Integration**: Adding voice command functionality would allow users to interact with the assistant more naturally, without needing to manually select options.
5. **Enhanced OCR Accuracy**: Improving the accuracy of OCR, especially for different fonts and handwritten text, would ensure more reliable text extraction.

THANK YOU

INNOMATICS RESEARCH LABS