**Title:** AI Vision Care Assist **Author:** Pavankumar Kurapati **Internship:** Data Science and Generative AI

Innomatics Research Labs

**Abstract**

The "AI Vision Care Assist" application leverages cutting-edge Artificial Intelligence to aid visually impaired individuals in understanding and interacting with their environment. This project integrates scene understanding, text extraction, object detection, and task-specific assistance, making it a versatile and practical tool. With multilingual support and text-to-speech capabilities, it provides a highly accessible user experience.

**Introduction**

* **Problem Statement**:

Visually impaired individuals face significant challenges in navigating their environment and understanding visual information. Existing solutions often lack multi-functionality or regional language support.

* **Objective**:

To develop an AI-powered application that assists visually impaired users by providing scene descriptions, text extraction, object detection, and task-specific guidance in multiple languages.

**Features**

**Scene Description**

* **What it does**: Generates a textual description of the uploaded image using Google Generative AI.
* **How it helps**: Provides an overview of the scene to users who cannot see it.

**Text Extraction**

* **What it does**: Extracts text from images using Tesseract OCR.
* **How it helps**: Allows users to read printed text from documents, labels, and signage.

**Object Detection**

* **What it does**: Detects objects in images using a pre-trained Faster R-CNN model.
* **How it helps**: Assists users in identifying obstacles or items in their environment.

**Personalized Assistance**

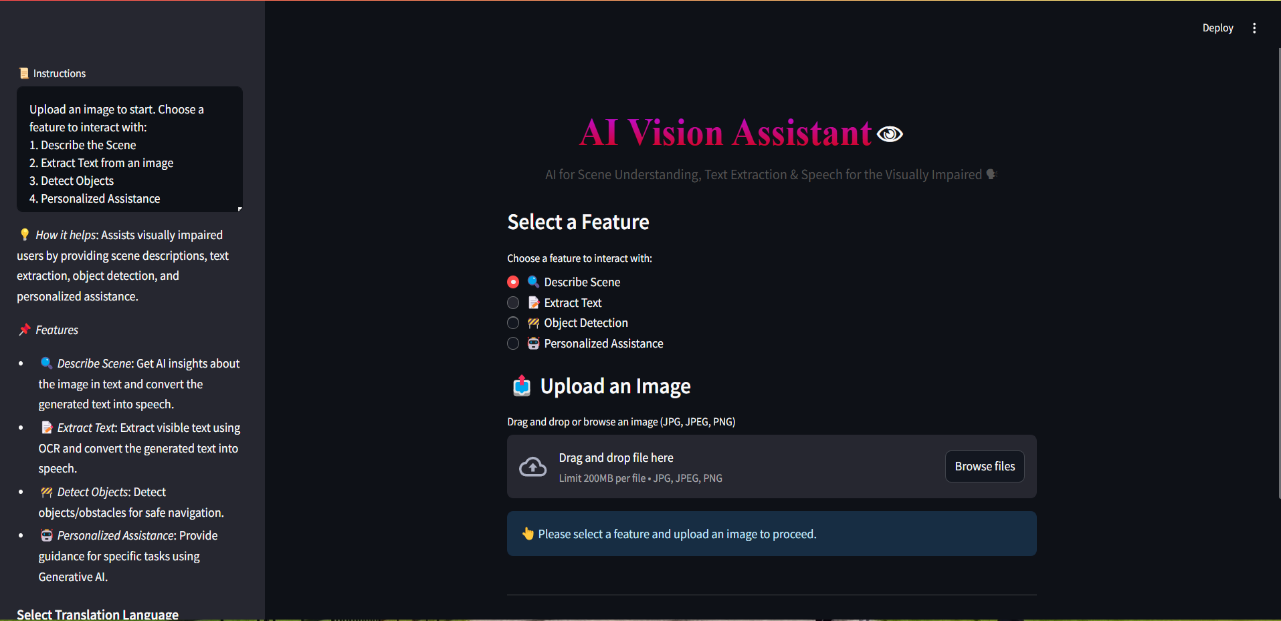
* **What it does**: Uses Generative AI to provide task-specific guidance based on the image content.
* **How it helps**: Offers detailed assistance, such as reading labels or recognizing items.

**Text-to-Speech**

* **What it does**: Converts the generated text into speech using pyttsx3.
* **How it helps**: Enables users to listen to the content in English.

**Multilingual Support**

* **What it does**: Translates text into regional languages (Telugu, Hindi, Kannada, and Malayalam) using Google Translate API.
* **How it helps**: Improves accessibility for non-English-speaking users.



**Technologies Used**

* **Streamlit**: For building the user interface.
* **Google Generative AI (Gemini API)**: For generating scene descriptions and personalized assistance.
* **Tesseract OCR**: For extracting text from images.
* **PyTorch (Faster R-CNN)**: For object detection.
* **Google Translate API**: For multilingual text translation.
* **pyttsx3**: For text-to-speech functionality.
* **Python Libraries**: PIL, torchvision, dotenv, and others.

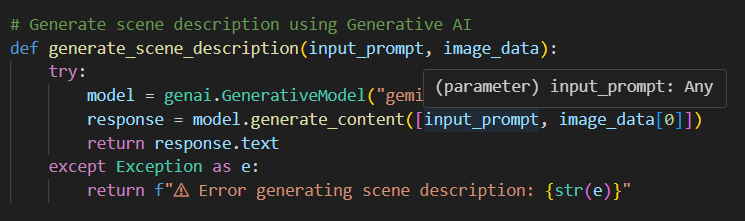
**Implementation :**

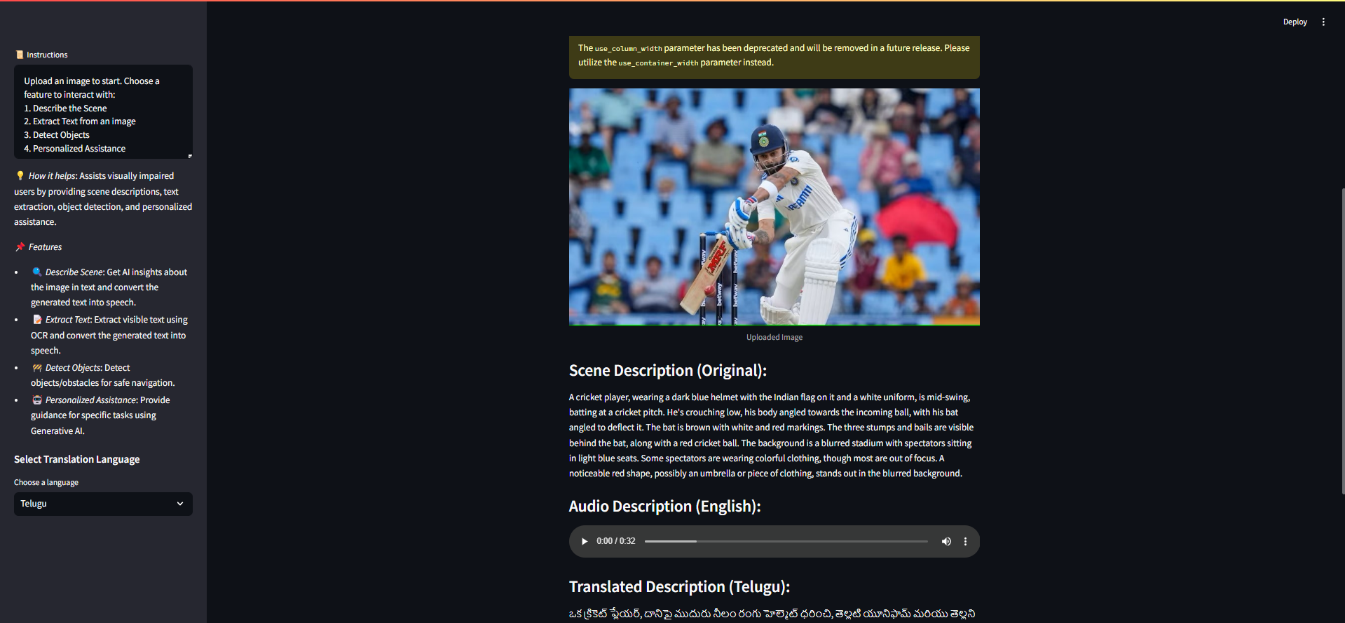
**Scene Description**

Provides an overview of the scene to users who cannot see it.

**Google Generative AI (Gemini API)**: For generating scene descriptions and personalized assistance.

**pyttsx3**: For text-to-speech functionality.



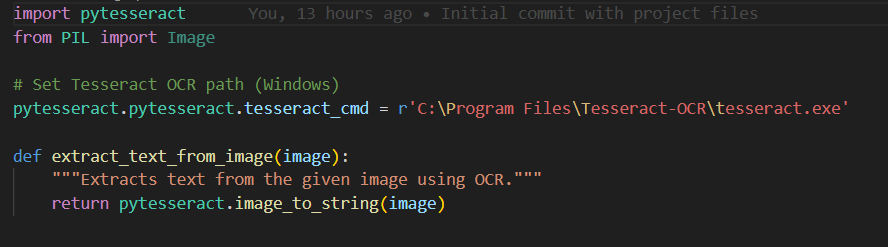


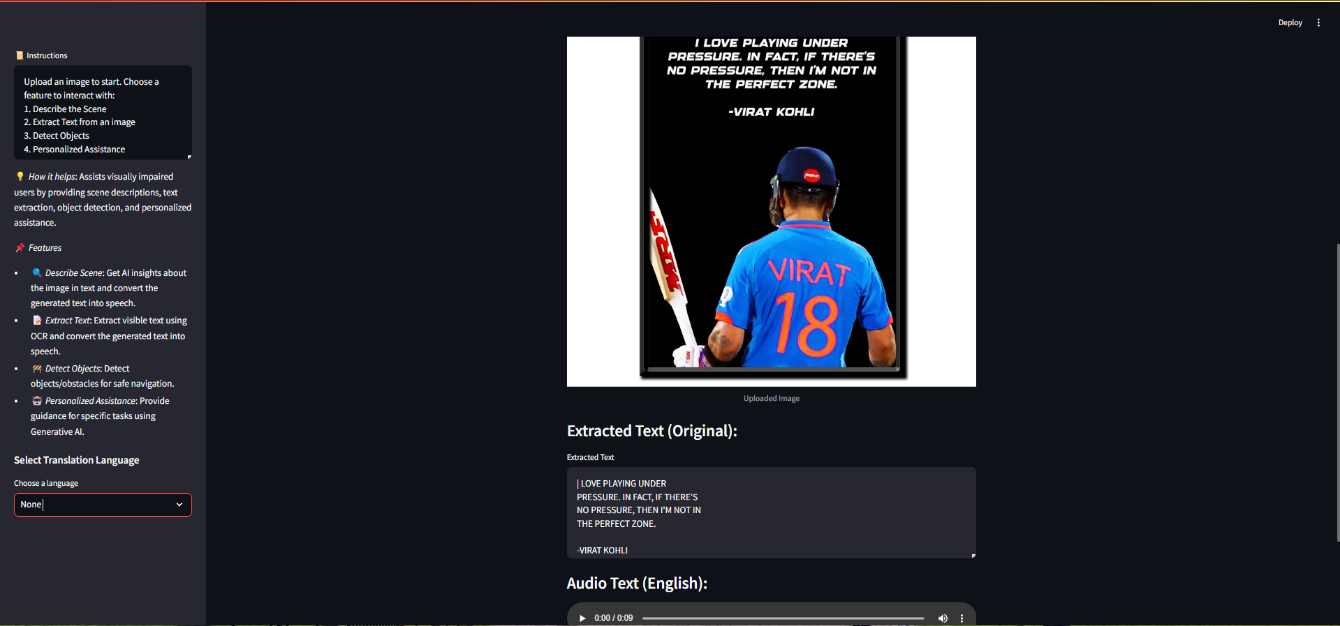
**Text Extraction :**

Allows users to read printed text from documents, labels, and signage.

**Tesseract OCR**: For extracting text from images.

**pyttsx3**: For text-to-speech functionality.

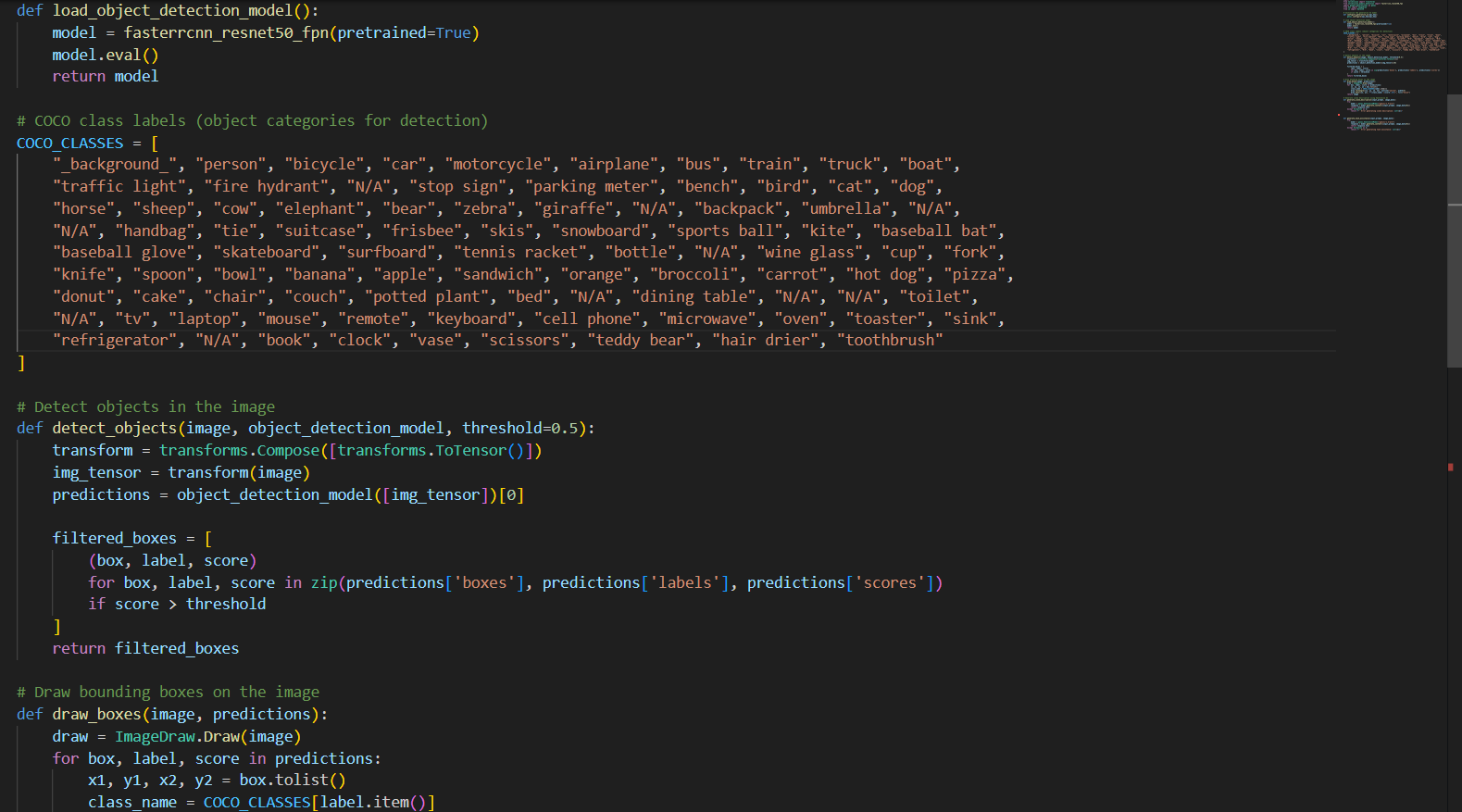


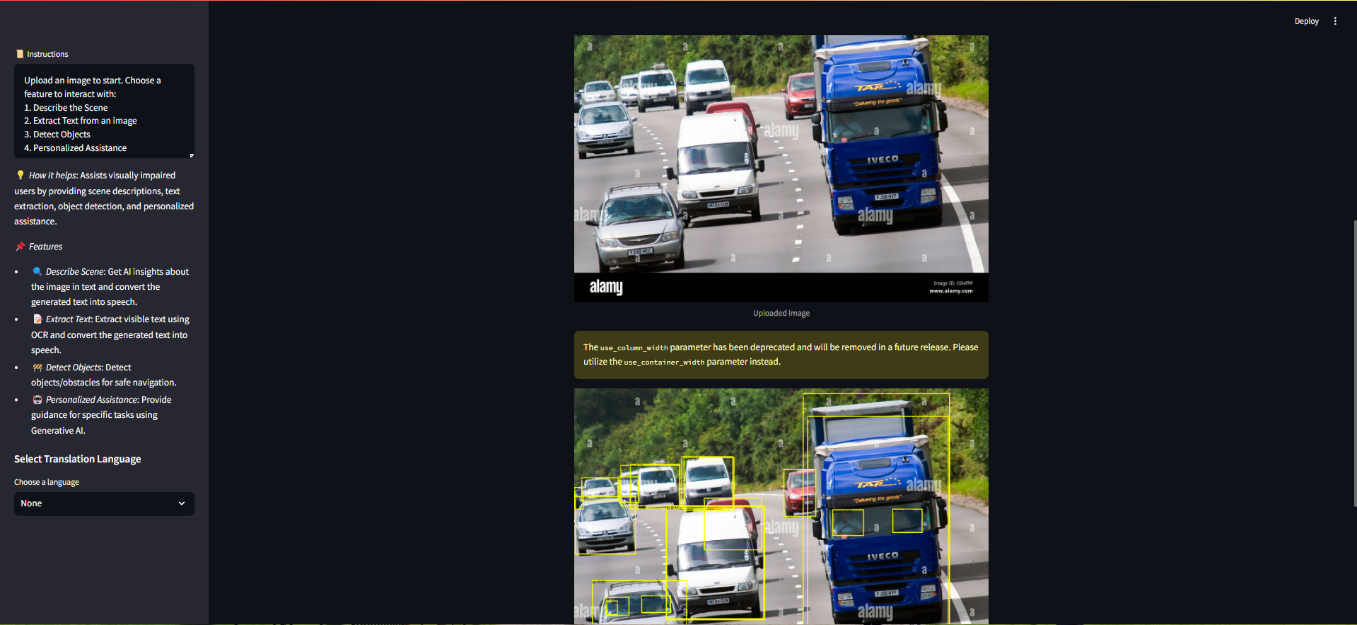


**Object Detection**

Assists users in identifying obstacles or items in their environment.

**PyTorch (Faster R-CNN)**: For object detection.

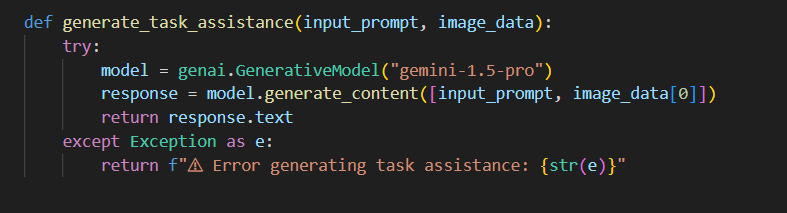
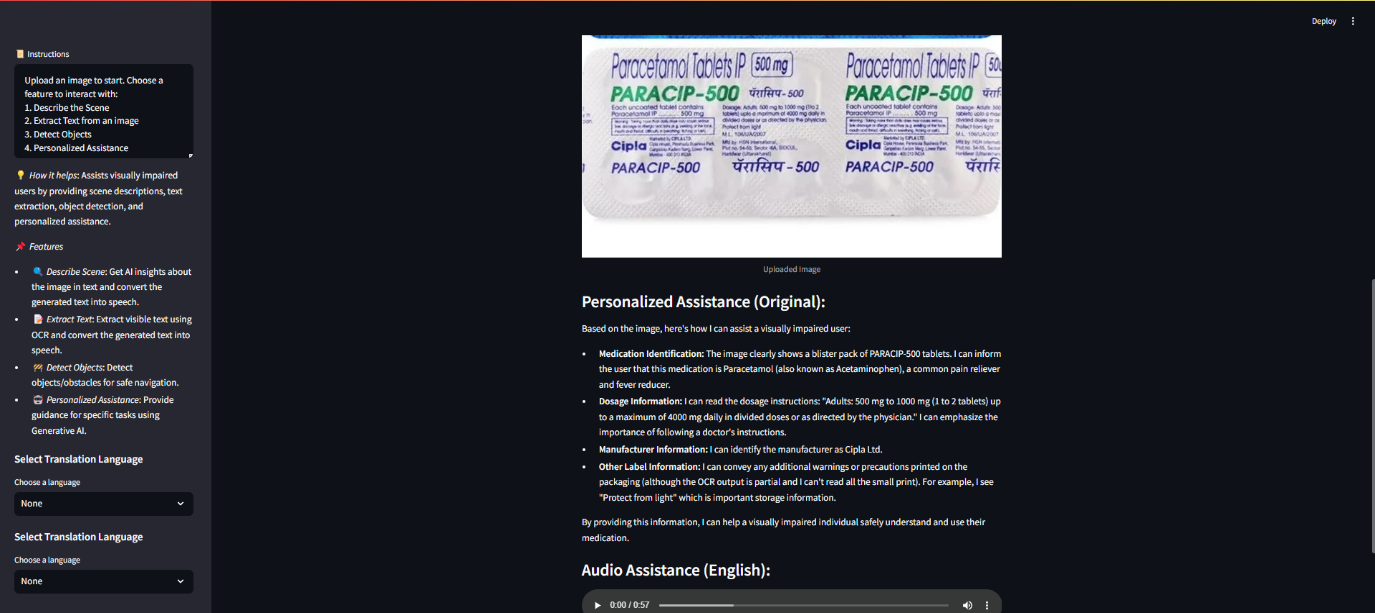




**Personalized Assistance :**

Offers detailed assistance, such as reading labels or recognizing items.

**Google Generative AI (Gemini API)**: For generating scene descriptions and personalized assistance.



**Conclusion :**

* The project provides a versatile tool to improve the independence and quality of life for visually impaired individuals.
* It combines state-of-the-art AI technologies with practical features like translation and audio output to enhance accessibility.