

AI Powered Solution for Assisting Visually Impaired Individuals

An AI-powered application to assist visually impaired individuals through real-time scene understanding, text-to-speech, object detection, and more.

Shivam Dubey
intern ID = IN9240906

Problem Statement

- This project aims to leverage Generative AI to assist visually impaired individuals in perceiving and interacting with their surroundings.
- Visually impaired individuals often face challenges in understanding their environment, reading visual content, and performing tasks that rely on sight.
- There is a need for an intelligent, adaptable, and user-friendly solution that provides:
 - - Real-time scene understanding.
 - - Text-to-speech conversion for reading visual content.
 - - Object and obstacle detection for safe navigation.
 - - Personalized assistance for daily tasks.

Task

- Develop an AI-powered application using Streamlit that provides assistive functionalities through image analysis. The application should allow users to upload an image and implement at least two of the following features:
 - - Real-Time Scene Understanding
 - - Text-to-Speech Conversion for Visual Content
 - - Object and Obstacle Detection for Safe Navigation
 - - Personalized Assistance for Daily Tasks

Features Overview

- The application includes the following functionalities:
- - Real-Time Scene Understanding: Describes the content of the uploaded image.
- - Text-to-Speech Conversion: Extracts and reads text from images using OCR.
- - Object and Obstacle Detection: Identifies and highlights objects/obstacles.
- - Personalized Assistance: Provides task-specific guidance and information.

Technologies Used

- - Langchain: For building language-based applications to assist with scene understanding and text conversion.
- - Streamlit: For creating a user-friendly interface for image upload and result display.
- - Google Generative AI: To leverage models for text generation, scene understanding, and speech conversion.
- - Tesseract OCR: For text extraction from images and conversion to speech.
- - OpenCV: For image processing and object detection.
- - TensorFlow: For deep learning models used in object detection and scene understanding.
- - YOLO: For real-time object detection.
- - Python: The programming language used to implement the application.

Demo of Features

- 1. Scene Understanding Demo: Upload an image and generate a descriptive text output.
- 2. Text-to-Speech Demo: Upload an image with text and hear the content being read aloud.
- 3. Object Detection Demo: Detect objects or obstacles in an uploaded image and highlight them for safer navigation.
- 4. Task Assistance Demo: Upload an image and get task-specific guidance, such as reading labels or recognizing items.



Upload Image

Choose an image (jpg, jpeg, png)

Drag and drop file here

Limit 200MB per file • JPG, JPEG, PNG

Browse files

00 (176).jpg
73.2KB

AI Assistive Tool for Visually Impaired

This AI-powered tool assists visually impaired individuals by leveraging image analysis. It provides the following features:

- **Scene Understanding:** Describes the content of uploaded images.
- **Text-to-Speech Conversion:** Extracts and reads aloud text from images using OCR.
- **Object & Obstacle Detection:** Identifies objects or obstacles for safe navigation.
- **Personalized Assistance:** Offers task-specific guidance based on image content, like reading labels or recognizing items.

Upload an image to get started and let AI help you understand and interact with your environment!



 Describe Scene

AI

Object recognition

Image generation

Text reading

Object tracking

Text reading

Object detection

Upload Image

Choose an image (jpg, jpeg, png)

Drag and drop file here


Limit 200MB per file • JPG, JPEG, PNG

Browse files

00 (176).jpg

73.2KB

0:00 / 3:31



A night-time street scene in a city, likely New York City, showing heavy traffic with yellow taxis and cars. Pedestrians are visible on the sidewalks. In the foreground, a black sign for 'JERSEY BOYS' is visible. Two red bounding boxes are drawn on the image: one around a traffic light and another around a 'No Left Turn' sign. The background features illuminated buildings and streetlights.

Highlighted Image with Detected Objects

✕

📁 Upload Image

Choose an image (jpg, jpeg, png)

Drag and drop file here

Limit 200MB per file • JPG, JPEG, PNG

Browse files

📄 00 (102).jpg

310.8KB

✕

🔑 Instructions

1. Upload an image.

2. Choose an option below:

- 🖼️ Describe Scene: Get a description of the image.
- 📄 Extract Text: Extract text from the image.
- 🚧 Detect Objects & Obstacles: Identify obstacles and highlight them.

AI Assistive Tool for Visually Impaired

This AI-powered tool assists visually impaired individuals by leveraging image analysis. It provides the following features:

- **Scene Understanding:** Describes the content of uploaded images.
- **Text-to-Speech Conversion:** Extracts and reads aloud text from images using OCR.
- **Object & Obstacle Detection:** Identifies objects or obstacles for safe navigation.
- **Personalized Assistance:** Offers task-specific guidance based on image content, like reading labels or recognizing items.

Upload an image to get started and let AI help you understand and interact with your environment!

A photograph of a city street scene. In the foreground, a pedestrian is walking across a crosswalk marked with a blue and white zebra crossing sign. Several cars are parked or moving along the street. In the background, there are buildings and a tram. The sky is overcast.

Share ☆ ✎ 🔄 ☰

< Manage app

Upload Image

Choose an image (jpg, jpeg, png)

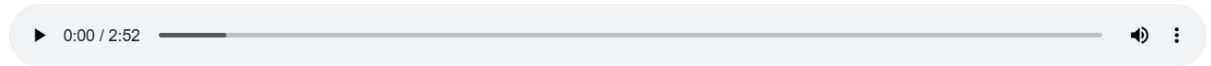
Drag and drop file here
Limit 200MB per file • JPG, JPEG, PNG

Browse files

00 (102).jpg
310.8KB

Instructions

1. Upload an image.
2. Choose an option below:
 - Describe Scene: Get a description of the image.
 - Extract Text: Extract text from the image.
 - Detect Objects & Obstacles: Identify obstacles and highlight them.



Highlighted Image with Detected Objects

Manage app

Future Improvements

- - Real-Time Image Analysis: Implement live camera feed for real-time analysis.
- - Enhanced Object Detection: Detect a wider range of objects for better assistance.
- - Multi-language Support: Provide text-to-speech in multiple languages.
- - User Preferences: Allow users to personalize the AI's assistance based on their needs.