Building AI-Powered Solution for Assisting Visually Impaired Individuals

Project Features:

1. Real-Time Scene Understanding:

- Users upload an image.
- The application generates descriptive textual output using Google Generative AI to interpret the scene.

2. Text-to-Speech Conversion:

- Users upload an image containing text.
- The application extracts the text using Tesseract OCR.
- The extracted text is converted into audible speech using the gTTS library.

Technologies Used:

- Streamlit: For building the user-friendly web interface.
- Google Generative AI: For scene interpretation.
- Tesseract OCR: For text extraction from images.
- gTTS (Google Text-to-Speech): For converting text into speech.

Repository Contents:

1. Code files
2. Outputs:

- Image description: Description of uploaded image is generated in the first project.
- Audio File: Output speech generated from the text in the second project.

3. Requirements:

- Install dependencies using `pip install -r requirements.txt`.

Run the Project:

1. Real-Time Scene Understanding:

code: *streamlit run realtimescene.py*

2. Text-to-Speech Conversion:

code: *streamlit run texttovoice.py*

3. Upload an image and test both functionalities.

Evaluation Criteria:

1. Successful implementation of two key functionalities:

- Real-Time Scene Understanding.
- Text-to-Speech Conversion.

2. Technical accuracy and usability.
3. Clear documentation and ease of execution.

Demo:

1. Real-Time Scene Understanding: Upload an image of a scene.
2. Text-to-Speech Conversion: Upload an image with text, and listen to the audio output.