How to Use Vision.ai

Overview

Vision.ai is an AI-powered image captioning system designed to assist visually impaired individuals. It allows users to upload an image, generates a meaningful caption, and reads the caption aloud using Text-to-Speech (TTS). Additionally, users can translate the caption into their preferred language.

System Requirements

- Python 3.x

- Required Libraries: Install all dependencies via the requirements.txt file.

Installation

Step 1: Clone the repository

Clone the Vision.ai repository using the command:

git clone https://github.com/nikhilitz/Vision.ai.git

Step 2: Install dependencies

Navigate to the project directory and install dependencies:

cd Vision.ai

pip install -r requirements.txt

Step 3: Start the backend (FastAPI)

Run the FastAPI backend:

uvicorn backend:app --reload

This starts the server at http://127.0.0.1:8000

Step 4: Start the frontend (Streamlit)

Launch the user interface using Streamlit:

streamlit run frontend/app.py

Usage Instructions

1. Open the frontend URL in your browser (typically http://localhost:8501).

2. Upload an image using the upload interface.

3. Wait for the caption to be generated and displayed.

4. Click the play button to hear the caption read aloud.

5. (Optional) Choose a different language to see the translated caption.