AI First Aid Instructor (Voice + Vision)

Project Overview

The AI First Aid Instructor is a voice and vision-based assistant built to provide immediate first aid guidance using speech and images. Users can speak their symptoms and upload a photo. The system uses Groq-hosted LLMs to understand the problem and return step-by-step visual and spoken first aid instructions.
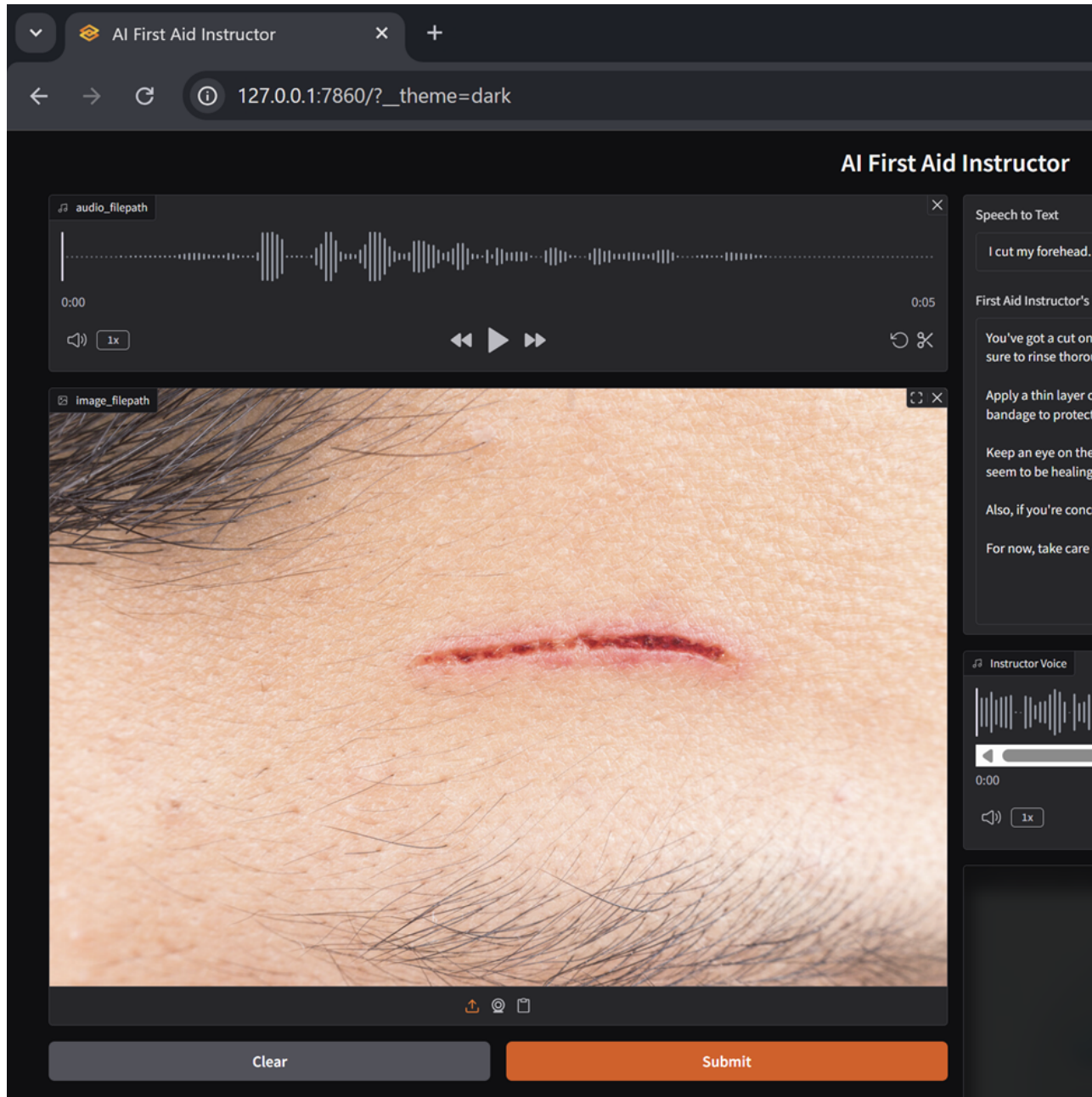
Tech Stack

- LLM: Groq (Meta LLaMA-3)

- STT: Whisper (Groq API)

- TTS: gTTS and ElevenLabs

- Image Generator: Hugging Face FLUX.1-dev

- UI: Gradio

- File Handling: Python, Pydub, FFmpeg

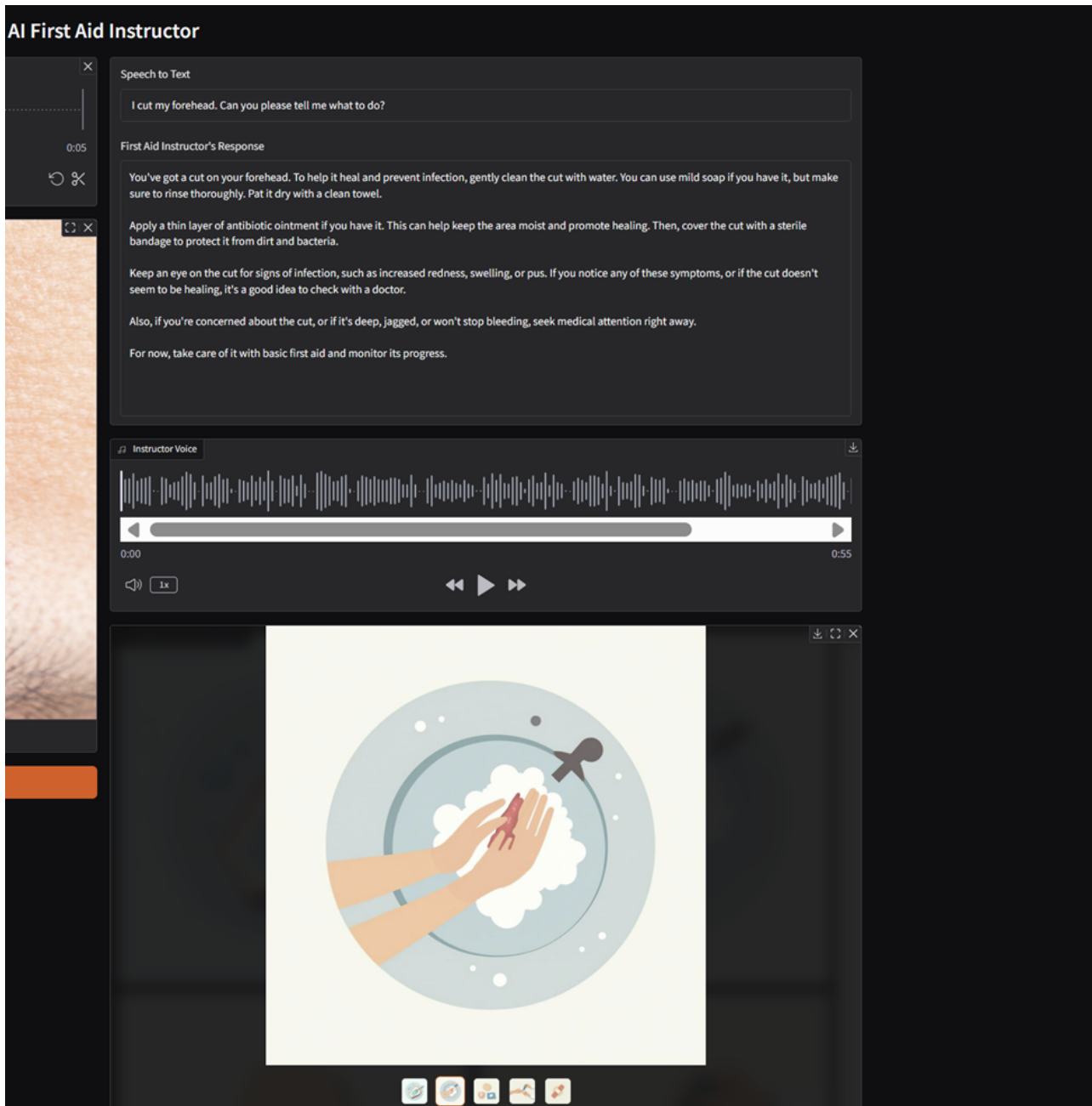# AI First Aid Instructor (Voice + Vision)

User Input: Voice + Image

The user uploads a photo of the injury and records a voice query explaining the situation.

AI First Aid Instructor (Voice + Vision)

Output: LLM Response + Voice + Visual Instruction

The system returns simplified medical guidance via text, converts it to audio using ElevenLabs, and generates a matching illustration using FLUX.

AI First Aid Instructor (Voice + Vision)

LLM Intermediate Steps

Raw first aid instructions are extracted and then simplified into prompts suitable for image generation.

```
🧠 LLM RAW OUTPUT (extracted steps):
[
  "Step 1: Gently clean the cut with water.",
  "Step 2: Use mild soap if available, but make sure to rinse thoroughly.",
  "Step 3: Pat the cut dry with a clean towel.",
  "Step 4: Apply a thin layer of antibiotic ointment if available.",
  "Step 5: Cover the cut with a sterile bandage to protect it from dirt and bacteria."
]
```

```
🎯 Simplified Visual Prompts:
 ["Clean a cut with water.", "Clean a cut with mild soap and water.", "Pat a cut dry with a towel.", "Apply antibiotic ointment to a cut.", "Cove
r a cut with a bandage."]
✅ Image generated: step_image_1.png
✅ Image generated: step_image_2.png
⚠️Attempt 1 failed for step 3: 504
⚠️Attempt 2 failed for step 3: 504
✅ Image generated: step_image_3.png
✅ Image generated: step_image_4.png
⚠️Attempt 1 failed for step 5: 504
✅ Image generated: step_image_5.png
```

LLM Intermediate Steps

AI First Aid Instructor (Voice + Vision)

Technical Architecture

The flowchart below outlines the step-by-step pipeline of how audio, image, and AI models interact to deliver voice and visual first aid outputs.