Curabot

A Multimodal AI-Powered Medical Assistant

Introduction

Curabot is a lightweight, multimodal AI-powered medical assistant designed to support healthcare professionals through natural, accessible interaction. By integrating text, audio, and image understanding, Curabot assists in answering clinical questions, analyzing medical images, and generating visual content to aid communication and education.

Built with efficiency in mind, Curabot targets deployment in low-resource environments—such as clinics, field hospitals, and mobile devices—where computing power is limited. The system integrates several state-of-the-art models into a unified pipeline, each specialized for a key task within the healthcare domain.

---

1. Medical Question Answering via Fine-Tuned Language Models

Dataset and Preprocessing

The heart of Curabot's chatbot functionality is a fine-tuned language model trained on the *"lavita/ChatDoctor-HealthCareMagic-100k"* dataset. This dataset includes 100,000 real-world doctor-patient Q&A pairs, offering structured, clean medical dialogue ideal for training. Compared to alternatives like MedDialog or HealthCareMagic-Raw, it provides a lower noise-to-signal ratio, resulting in improved model performance.

Model Selection and Justification

We evaluated **Qwen/Qwen2.5-3B-Instruct** against **LLaMA 2 8B**. While LLaMA 2 achieved slightly better raw performance, it demanded significantly more computational resources. Qwen2.5-3B-Instruct, with fewer parameters, offered comparable accuracy and greatly improved inference speed—especially after quantization—making it better suited for real-time, resource-constrained applications.

Optimization: Quantization and LoRA

To reduce memory usage and improve runtime efficiency, the Qwen model was quantized to **4-bit precision** using *bitsandbytes*. We fine-tuned the model using **LoRA (Low-Rank Adaptation)** with rank *r=8*, allowing knowledge injection without retraining the full model. This parameter-efficient technique retained the model's core language abilities while minimizing compute requirements.

Use Case: Multimodal Q&A

Users can ask medical questions by typing or speaking. Spoken queries are transcribed via the Whisper model and processed by the Qwen model—optionally enhanced by a RAG (Retrieval-Augmented Generation) pipeline for better accuracy.

---

2. Speech-to-Text with Whisper

Audio Processing

To support voice interaction, Curabot uses OpenAI's **Whisper-base model**. Audio inputs are resampled to 16 kHz and transcribed with high accuracy, even in noisy or accented environments. We opted for the base model to balance performance with speed and hardware efficiency.

Use Case: Voice-Based Accessibility

This module allows hands-free interaction, enabling patients or doctors to use voice for medical queries. It's particularly helpful in low-literacy settings or for users with limited typing ability, supporting broader accessibility.

---

3. Retrieval-Augmented Generation (RAG)

Motivation and Dataset

To improve factual grounding and reduce hallucinations in responses, we implemented a RAG pipeline. We used the *"DSWF/ai_medical_chatbot_train"* dataset with 270,000 Q&A entries. Each was embedded using **all-MiniLM-L6-v2** and indexed with **FAISS** for fast retrieval.

Retrieval Process

When a query is submitted, it is embedded and matched with the top 10 similar Q&A pairs. These are appended as context for the language model, providing a dynamic knowledge base without increasing model size.

Use Case: Context-Rich Answers

RAG enhances the chatbot's accuracy in complex queries by grounding its responses in relevant historical data—simulating an "open-book" environment while maintaining efficiency.

---

4. Stable Diffusion for Medical Image Generation

Dataset Synthesis

To enable medical text-to-image generation, we created a synthetic dataset by converting radiology reports into descriptive captions using a large language model. These were paired with corresponding medical images for training.

Model and Training

We fine-tuned **CompVis/stable-diffusion-v1-4** on the synthetic dataset. This allowed the model to generate medical images—e.g., "MRI of brain showing pituitary adenoma"—from descriptive input.

Use Case: Medical Visualization

Doctors can describe conditions to generate illustrative visuals for education or simulation. This is especially valuable in training scenarios and for communicating complex diagnoses to patients.

---

5. YOLO for Brain Tumor Detection

Dataset and Annotation

We trained the detection module using the *brain-tumor-yzzav* dataset, which includes labeled MRI scans for three tumor types: glioma, meningioma, and pituitary adenoma.

Model Architecture and Results

A **YOLOv5** model was fine-tuned using data augmentation, input normalization, and anchor optimization. The model achieved:

- **mAP@50**: 0.992
- **mAP@50–95**: 0.798

These metrics indicate high detection accuracy and reliable segmentation performance.

Use Case: Automated Diagnosis Support

Curabot's image analysis module enables automated tumor detection and classification from brain MRI scans. It supports radiologists by highlighting potential tumors and reducing diagnostic time, particularly in areas with limited expertise.

---

6. System Integration and Architecture

Modular and Maintainable Design

Each core function—language modeling, speech recognition, image generation, retrieval, and analysis—is developed as an independent module. This modularity ensures flexibility, allowing easy upgrades or replacement of components without disrupting the system.

Designed for Efficiency

Every component is optimized for performance on limited hardware: quantized models, low-rank adapters, compact embedding networks, and lightweight vision models keep resource usage minimal while maintaining robust functionality.

---

7. Use Case Summary

- **Medical Q&A**: Type or speak questions to receive medically informed answers.
- **RAG Support**: Enhances chatbot accuracy through context retrieval.
- **Voice Interaction**: Enables hands-free communication with the assistant.
- **Text-to-Image Generation**: Creates medical visuals from descriptions.
- **Tumor Detection**: Automatically identifies brain tumors in MRI scans.

🎬 Demonstration Video Script and Structure for Curabot:

https://drive.google.com/file/d/1v1-9qFTXpQsp0niiGLwCMucBpiy-YGAB/view?usp=sharing

---

Conclusion

Curabot illustrates how modern AI components can be strategically combined to create a practical, efficient, and intelligent healthcare assistant. Its lightweight architecture, multimodal capabilities, and focus on real-world constraints make it a valuable tool for accessible, AI-driven medical support—especially in under-resourced settings.

As technology advances, Curabot offers a strong foundation for future systems that blend AI with frontline healthcare delivery.