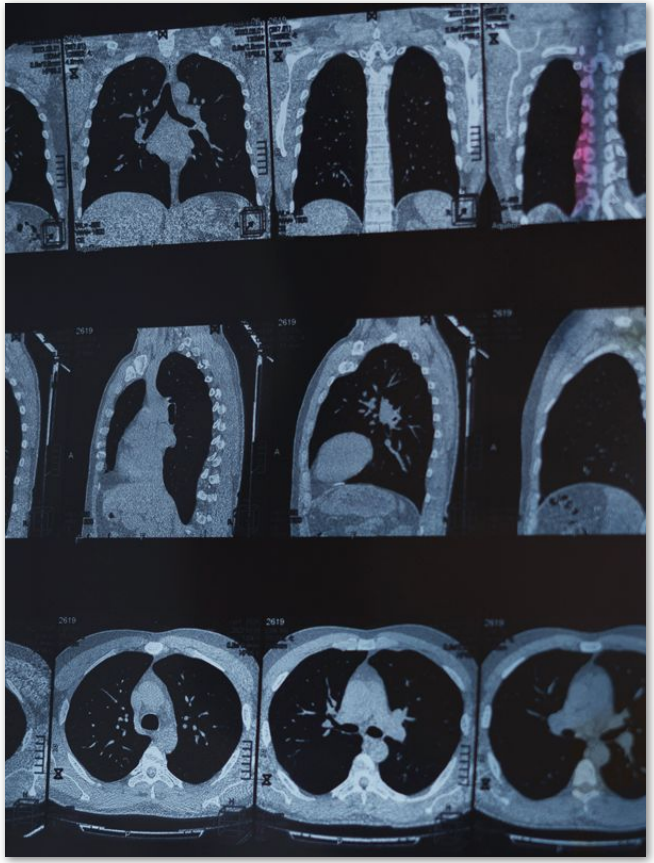
EECS E6895 Spring 2026
Agentic AI
Midterm Project Plan

Radiologist 1:
“RAV”

Rithika Devarakonda
Wei **A**lexander Xin
Vikas Chelur



We will build an agentic AI workflow that imitates the basic diagnostic tasks of a professional radiologist.

Given medical images such as x-ray, MRI, or CT scans, our agent will analyze, segment, and extract relevant information using computer vision.

Then, the agent will output key diagnostic information such as the presence of a fracture or a tumor.

Additionally, our agent will provide qualitative remarks on the diagnosis. Is it a compound or simple fracture? Is the tumor benign or malignant?

Goal

For model training, we will use an open source dataset of relevant medical images, available from various online repositories.

As an example, Stanford AIMI has a selection of public datasets, one of which is the CheXpert Plus dataset. This dataset includes "223,462 unique pairs of radiology reports and chest X-rays across 187,711 studies from 64,725 patients."

Training images must be linked to reports that describe the image and provide diagnostic information. We will extract key information as text labels to use in supervised training.

https://aimi.stanford.edu/shared-datasets

Dataset

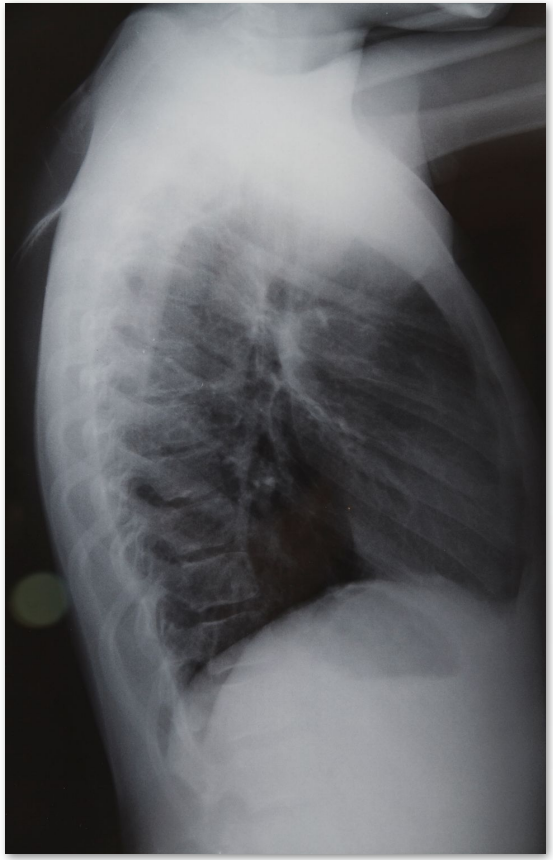
We plan to use a pre-trained *computer vision* model as a foundation model. We will then fine-tune the model to tailor it our specific medical imaging task(s). Choosing what specific task(s) we want to address will be key to driving this additional layer of training.

In our preliminary analysis, we've already identified a few pre-trained models, such as RF-DETR and YOLOv12. However we will conduct additional research to determine the one most ideal for radiologist context.

Additionally, we'll use Gemini 1.5 Flash to enable natural language input and output, which will also require fine-tuning.

https://blog.roboflow.com/pre-trained-models/

Models

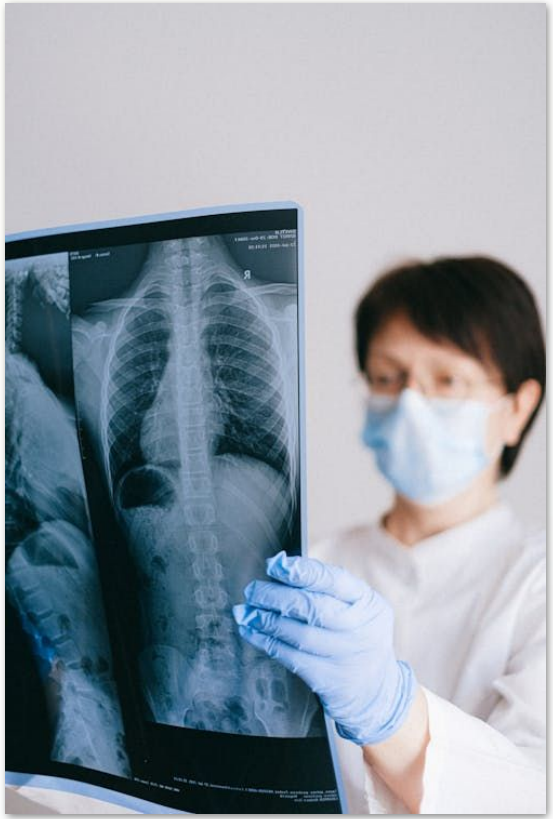


We will leverage Model Context Protocol (**MCP**), developed by Anthropic, to enable our radiologist agent “RAV” to interact with external resources.

The use context will be limited to diagnostic medical imaging. Decision making and text generation will be autonomous. There will be a simple I/O interface.

RAV will accept a medical image from the user via file upload, then apply our fine-tuned computer vision model to locate and identify visual areas of interest. These areas will be classified accordingly, and then a summary of findings will be generated.

Architecture



We will evaluate the agent using a multi-pronged approach.

The MCP integration will be evaluated using **MedAgentBench** from Stanford, which tests the efficiency and accuracy of an agent's interactions with external medical resources.

As a whole, we will test the agent *quantitatively* by prompting it with questions with fixed answers based on the computer vision model's output, to reduce hallucinations. We will also test it *qualitatively* by having a more powerful or specialized LLM (such as **MedGemma**) evaluate the agent's outputs, to improve relevance

Evaluation

Thank you.

Any questions? Visit us at:
radiologist-rav.com

Rithika Devarakonda
Wei **A**lexander Xin
Vikas Chelur