Project Proposal: Chest X-ray Report Generation and Interactive Radiology Assistant

Project Members:

- Anand Raj
- Saniya Shinde
- Shanun Randev
- Abishek Chiffon Muthu Raja

Project Overview

The goal of this project is to develop an AI-driven system capable of automatically generating clinically accurate radiology reports from chest X-ray images and serving as an interactive assistant for radiologists. This tool is intended to support radiologists in generating diagnostic reports, reducing their workload, and facilitating better patient outcomes through improved report quality and real-time diagnostic assistance.

Problem Statement

Radiologists face a significant workload due to the high demand for imaging studies, especially chest X-rays, which are one of the most common imaging modalities. This project addresses the need for a system that can:

- **Generate** structured and coherent radiology reports based on X-ray images.
- **Assist** radiologists interactively in answering diagnostic questions, thereby acting as an intelligent radiology assistant. (Optional)

The solution will aim to enhance the efficiency of the diagnostic process and enable radiologists to spend more time on complex cases, ultimately improving clinical workflow.

Dataset and Resources

- **MIMIC-CXR Dataset**: A comprehensive, publicly available collection of chest X-ray images with corresponding radiology reports, providing a robust foundation for both image processing and report generation.
- **CheXpert Labels**: To structure findings, the CheXpert dataset's labels will be used for multi-label classification of common pathologies, facilitating model training on real-world radiological findings.

Data Access Links:

- Chest X-ray Images (MIMIC-CXR)
- Radiology Reports (MIMIC-CXR)

Technical Approach and Model Architecture

Core Components

1. **Image Feature Extraction Module**:
   - A multimodal vision-language transformer model (e.g., MedViLL or VisualBERT) will be used to extract and encode visual features from X-ray images.
   - Alternatively, a pre-trained domain-specific encoder like BioViL-T may be fine-tuned for improved performance on chest X-rays, generating robust image embeddings aligned with clinical standards.
2. **Prompt Construction Module**:
   - Visual embeddings and CheXpert-labeled findings will be combined to form prompts for report generation, allowing the model to produce contextually relevant and clinically accurate reports.
3. **Report Generation Model**:
   - A Large Language Model (LLM), fine-tuned on radiology-specific text data, will be used to generate free-text reports. By leveraging domain adaptation techniques, we aim to align the generated text with radiologist reporting style, ensuring consistency and accuracy in diagnostics.
4. **Interactive QA Model (Optional)**:
   - Using RaDialog, a radiology-specific question-answering model, an interactive component will be integrated, enabling radiologists to ask detailed questions about specific pathologies, findings, or report sections. This will add a conversational element to the system, aiding radiologists in more precise decision-making.

NLP Tasks

The project will address multiple NLP tasks, ensuring that the system meets a wide range of clinical needs:

- **Radiology Report Generation**: Produces comprehensive, structured radiology reports, saved as PDFs or Word documents.
- **Summarization**: Condenses generated reports into concise summaries for quick reference.
- **Simplified Language Output**: Reformulates reports into simpler language, enhancing accessibility for patients and non-specialist staff.
- **Findings Question-Answering (Optional):** Answers specific queries related to pathologies in the image, leveraging the interactive QA model.

Software and Tools

The project will utilize a combination of machine learning frameworks and NLP packages:

- **PyTorch**: For model training and development, offering flexibility for custom architecture design.
- **Hugging Face Transformers**: For integrating and fine-tuning large language models.
- **scikit-learn**: For evaluation metrics and performance analysis.
- **NLTK and BERTScore**: For natural language processing tasks, including text similarity measurement.
- **Streamlit**: To deploy the model and create an interactive interface for users.
- **OpenCV**: For preprocessing and manipulating X-ray images.

Evaluation Metrics

To ensure the accuracy and clinical validity of generated reports, the following metrics will be employed:

- **Clinical Efficacy (CE)**: Measures the correctness of generated pathology findings compared to ground truth labels.
- **BERTScore (BS)**: Evaluates semantic similarity between generated and reference reports, capturing meaning beyond exact word matching.
- **NLG Metrics**: BLEU, ROUGE, and METEOR will be used to assess linguistic similarity and fluency. While these may not fully capture clinical correctness, they provide insight into report coherence and language quality.
- **Human Evaluation**: Domain experts (e.g., radiologists) will review a sample of generated reports to assess clinical relevance, accuracy, and usability.

Development Timeline

Timeline	Tasks
Nov 1st Week	Data exploration, preprocessing, and setup of the image feature extraction module.
Nov 2nd-3rd Weeks	Implement and fine-tune the multimodal transformer model and language model for radiology reports.
Nov 4th Week	Develop and test additional NLP tasks: summarization, language simplification, and interactive QA.
Dec 1st Week	Model refinement based on clinical feedback, final adjustments, and deployment.
Dec 2nd Week	Project wrap-up and documentation.

References

- MedViLL GitHub Repository
- Multimodal in Medical Imaging
- RaDialog Interactive QA Model
- Research Paper on Vision-Language Transformers
- CheXpert Labeling Approach