X-Ray Image Processing and Large Language Models for Multi-Disease Detection and Automated Reporting

Kartik Ranjan (23MDT0065), Mathew Kevin John (23MDT0046), Vishnu Udaiyar (23MDT0027)

Introduction

Introduction

Overview

- Chest X-rays are essential for diagnosing thoracic diseases but are challenging to interpret due to variability and time constraints.
- This project integrates advanced X-ray image processing with large language models (LLMs) to enhance detection and automate reporting.
- Using a dataset of 112,120 annotated images across 14 disease categories, the framework detects diseases and generates clinically relevant diagnostic reports.
- Preliminary results demonstrate improved accuracy, efficiency, and standardization, highlighting AI's potential to transform radiology practices.

Introduction

Objective of Research

Developing a Deep Learning Framework for Multi-Disease Detection
Design a robust deep learning model to detect and classify multiple thoracic diseases from chest X-ray images, enhancing diagnostic accuracy and efficiency.

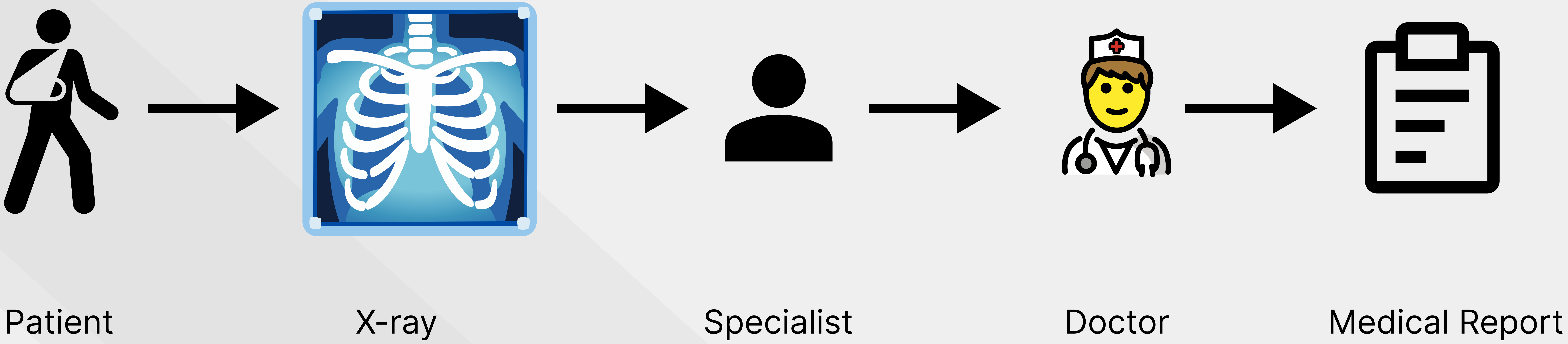
Automated Diagnostic Report Generation with Large Language Models (LLMs)
Integrate findings from the image processing model into an LLM to automatically produce coherent and clinically relevant diagnostic reports, aiming to standardize and streamline the reporting process.

Standardizing Reporting to Reduce Variability and Minimize Errors
Address the variability among radiologists in interpreting X-rays by generating consistent, automated reports that reduce human error and standardize diagnostic practices.

Transforming Radiological Workflows and Improving Patient Outcomes
By improving the speed and consistency of disease detection and reporting, this project has the potential to revolutionize radiology practices, lighten radiologists' workload, and enhance patient care.

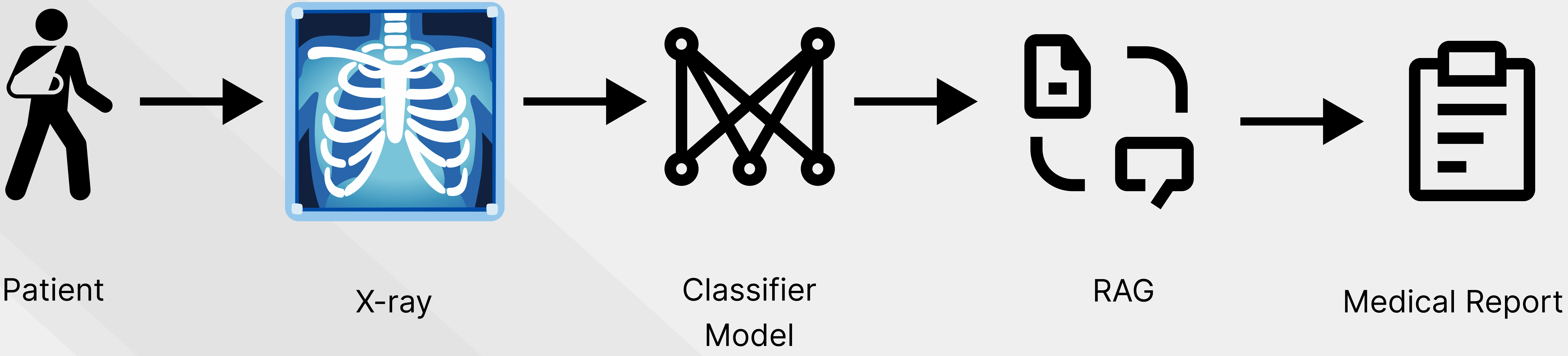
Traditional Approach

Patient - Doctor Interaction



Our Approach

Patient - Diagnosis



Data Preparation

Data Preparation

Data Collection

1. Data Collection Process

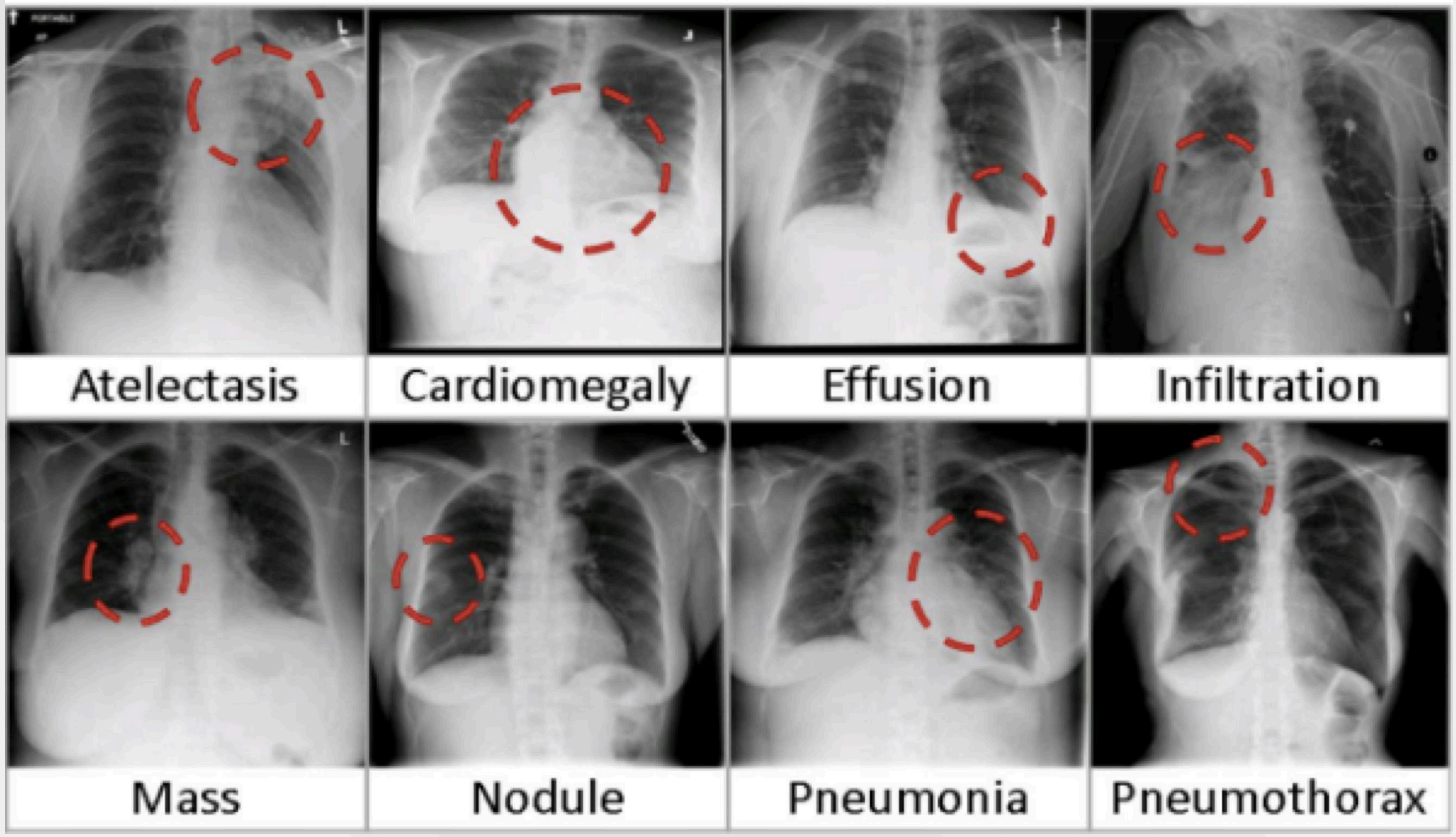
Dataset Source: NIH Clinical Center’s public ChestX-ray dataset, obtained from this link.
https://nihcc.app.box.com/

2. Dataset Overview

Image Details: Each X-ray is in PNG format with a resolution of 1024×1024.
Pathologies: There are 14 thoracic diseases labeled, including common conditions like Atelectasis, Pneumonia, and Cardiomegaly.
Metadata: The dataset also includes patient demographics (e.g., age, gender), view positions, image size, and pixel spacing.
Bounding Boxes: ~1000 images have bounding box annotations indicating areas of abnormalities.
Data Split: Train/validation and test sets are provided at the patient level to ensure no overlap in patient data between the sets.

Data Preparation

Data Collection



	A	B	C	D	E
1	Image Index	Finding Labels	Follow-up #	Patient ID	Patient Age
2	00000001_000.png	Cardiomegaly	0	1	57
3	00000001_001.png	Cardiomegaly|Emphysema	1	1	58
4	00000001_002.png	Cardiomegaly|Effusion	2	1	58
5	00000002_000.png	No Finding	0	2	80
6	00000003_001.png	Hernia	0	3	74

A	B	C	D	E	F
Image Index	Finding Label	Bbox [x	y	w	h]
00013118_008.png	Atelectasis	225.0847458	547.0192168	86.77966102	79.18644
00014716_007.png	Atelectasis	686.1016949	131.5434984	185.4915254	313.4915
00029817_009.png	Atelectasis	221.8305085	317.0531151	155.1186441	216.9491
00014687_001.png	Atelectasis	726.2372881	494.9514202	141.0169492	55.3220
00017877_001.png	Atelectasis	660.0677966	569.7807865	200.6779661	78.1016
00003148_004.png	Atelectasis	596.0677966	505.7807865	56.40677966	180.0677
00012515_002.png	Atelectasis	289.0847458	638.1378608	83.52542373	56.4067
00022098_006.png	Atelectasis	494.1016949	577.3920981	271.1864407	154.0338
00014198_000.png	Atelectasis	676.3389831	512.3073524	98.71186441	193.0847
00021007_000.png	Atelectasis	344.4067797	468.9175218	105.220339	101.9661

Data Preparation

Data Collection

1. **Comprehensive Research Collection:** Conducted an extensive literature review, gathering research papers from reliable academic sources, and categorized them by disease type (e.g., infectious, chronic, genetic) to create an organized knowledge base.

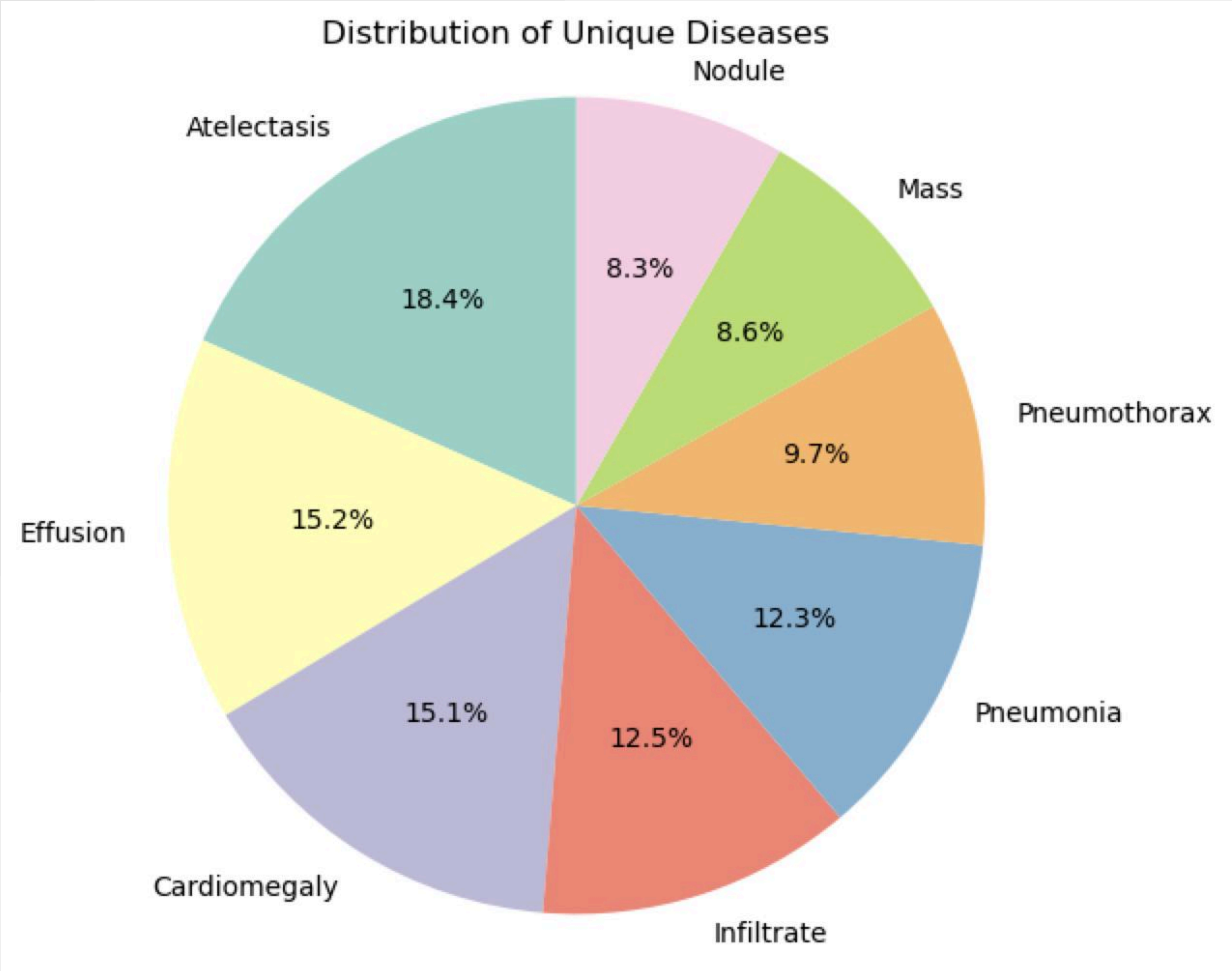
2. **Detailed Data Extraction:** Essential information, such as abstracts, methodologies, findings, and conclusions, was extracted from each paper, and each document was tagged with metadata (author, publication year, journal, keywords) for enhanced searchability.

3. **Data Preparation for Model Compatibility:** The dataset was cleaned, standardized, and formatted to align with the input requirements of RAG models, including tokenization and text segmentation, ensuring the data was structured for precise retrieval.

4. **Model Integration and Testing:** After building the knowledge base, we integrated it with the RAG models, performing testing and fine-tuning to ensure the models produced contextually relevant and accurate disease-specific responses for clinical, research, and educational applications.

Data Preparation

Data Visualisation



Atelectasis: 18.4%
Effusion: 15.2%
Cardiomegaly: 15.1%
Infiltrate: 12.5%
Pneumonia: 12.3%
Pneumothorax: 9.7%
Mass: 8.6%
Nodule: 8.3%

Methodology

Data Preprocessing

Preprocessing Steps:

Normalization of Bounding Box Labels

Adjusts bounding boxes for consistency across different images.

Vectorization:

Input Vector: Pixel values from 1024×1024 grey-scaled images reshaped to (224*224, 3).

Output Vector: Classiication labels reshaped to (15, 1).

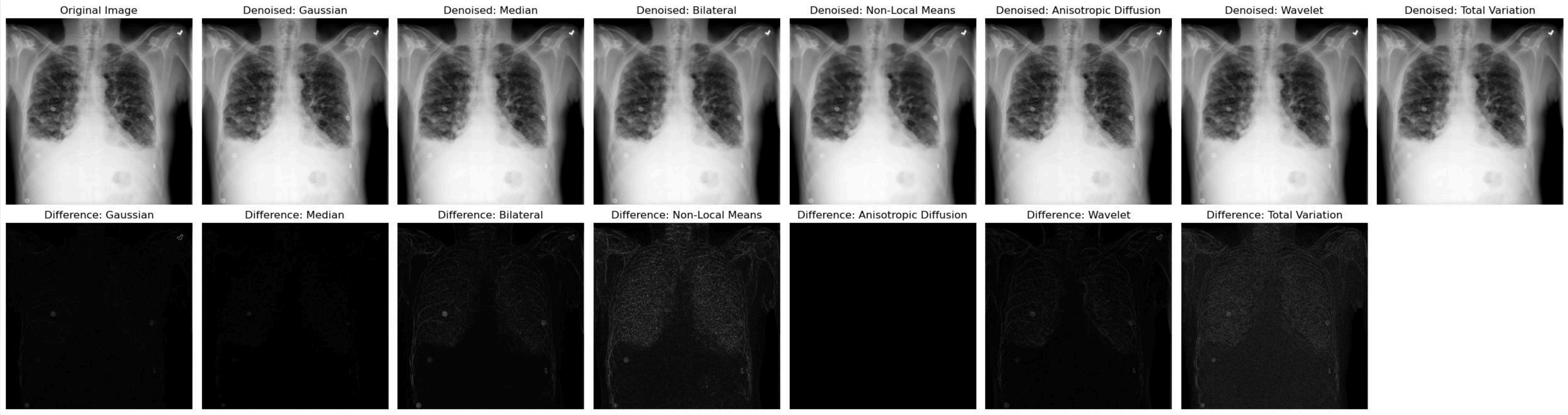
Filtering Techniques Applied:

- Histogram Equalization: Enhances image contrast.
- Gaussian Blur: Reduces noise and smoothens images.
- Median Blur: Removes noise while preserving edges.

Additional Preprocessing:

- Data Augmentation: Random rotations, flips, and translations.
- CLAHE: Local contrast improvement.
- Edge Detection (Canny/Sobel): Highlights anatomical features.

Data Preprocessing



Models Used

X-ray Image processing

1. **EfficientNetB0:**
   input shape: (224, 224, 3) output shape: (1, 15) optimizer: Adam loss: binary_crossentropy metrics: accuracy

2. **NASNetLarge:**
   input shape: (224, 224, 3) output shape: (1, 15) optimizer: Adam loss: binary_crossentropy metrics: accuracy

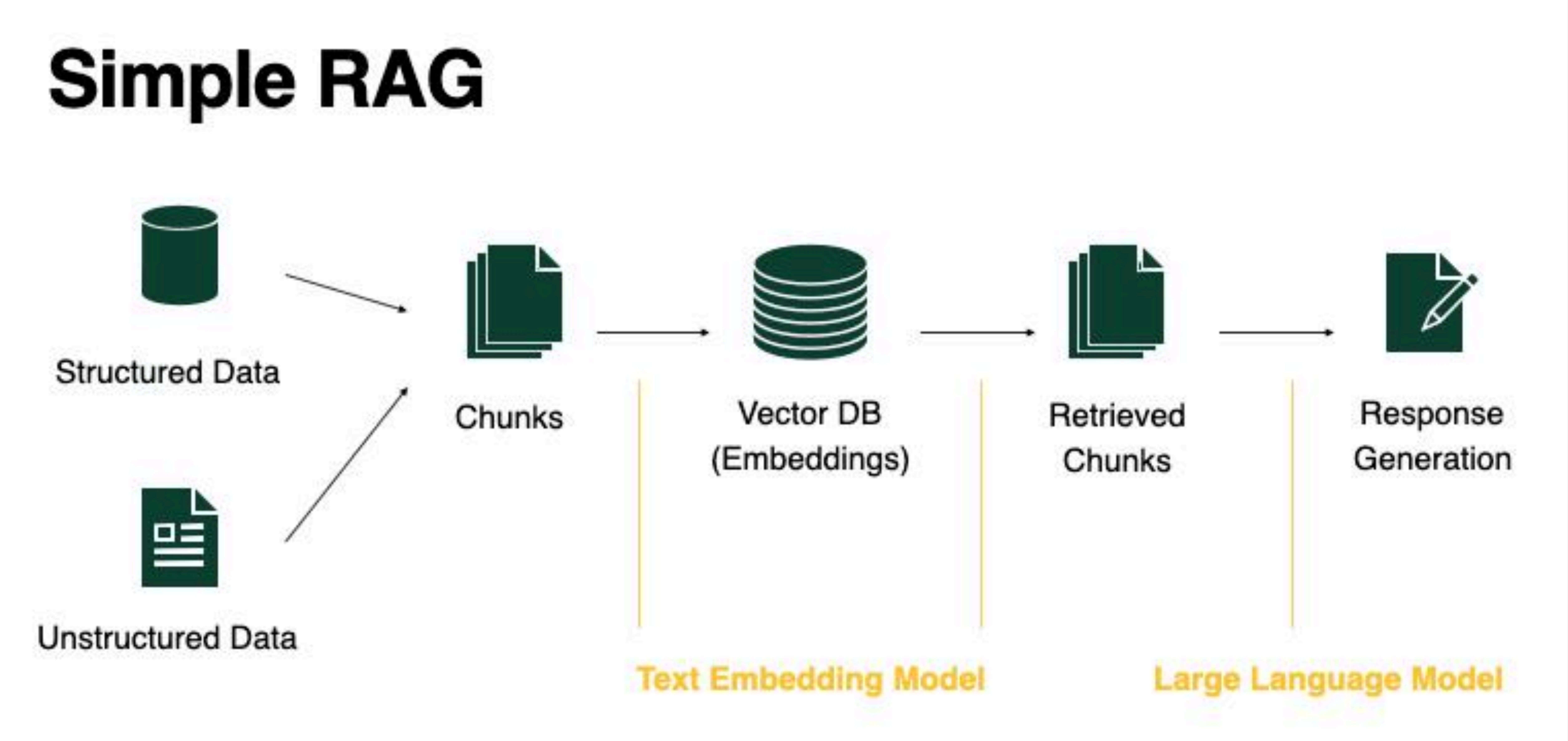
3. **ResNet50**
   input shape: (224, 224, 3) output shape: (1, 15) optimizer: Adam loss: binary_crossentropy metrics: accuracy

4. **VGG16**

5. **MobileNetV3Large**

Models Used

RAG



Models Used

RAG

Textual Data Preprocessing:

- Converting .docx to .pdf using PyPDF to load read pdf documents in a structured format.
- Converting the retrieved textual data to embeddings using OllamaEmbeddings for model “Llama3”

Vector Store:

- FAISS

LLM:

- ChatGroq model ‘mixtral-8×7b-32768’ along with Ollama’s Llama3 for determining context

Client app:

- Streamlit (Python)

Models Used

RAG

```
if "vectors" not in st.session_state:
    st.session_state.embeddings = OllamaEmbeddings(model="llama3")

    if os.path.exists(VECTOR_STORE_PATH):
        st.session_state.vectors = FAISS.load_local(VECTOR_STORE_PATH, st.session_state.embeddings, allow_dangerous_deserializatio
        print("\nLoaded Vector Store from disk-----------\n")
    else:
        st.session_state.docs = load_documents(DATA_FOLDER)
        print("\nRead Documents ------------\n")

        st.session_state.text_splitter = RecursiveCharacterTextSplitter(chunk_size=1000, chunk_overlap=200)
        st.session_state.final_documents = st.session_state.text_splitter.split_documents(st.session_state.docs)

        st.session_state.vectors = FAISS.from_documents(st.session_state.final_documents, st.session_state.embeddings)

        # Save vector store to disk
        st.session_state.vectors.save_local(VECTOR_STORE_PATH)
        print("\nSaved Vector Store to disk-----------\n")
```

```python
llm = ChatGroq(
    model="mixtral-8x7b-32768",
    temperature=0,
    max_tokens=None,
    timeout=None,
    max_retries=2,
    # other params...
)


print("\nInitialised LLM -----------------\n")

prompt = ChatPromptTemplate.from_template(
"""
You are given a patient's data of Age, Gender and Diagnosis from X-ray images.
You are an experienced clinical assistant specializing in medical diagnostics and treatment recommendations. You can

<context>
{context}
<context>

Patient Data: {input}
"""
)

document_chain = create_stuff_documents_chain(llm, prompt)
retriever = st.session_state.vectors.as_retriever()
retrieval_chain = create_retrieval_chain(retriever, document_chain)
```

Conclusions

Results

This research highlights the transformative potential of AI in medical imaging, particularly in automating the diagnostic process and reducing human error. The combination of deep learning models for image processing and LLMs for report generation has shown significant promise in improving diagnostic workflows, ultimately contributing to better patient outcomes.

Preliminary results indicate that the integration of X-ray image processing with LLM-based report generation shows promising outcomes. The multi-disease detection accuracy has reached competitive levels compared to existing models in the literature, and the generated diagnostic reports were found to be coherent and clinically relevant.

MODEL	ACCURACY
EfficientNet	67.02%
NasNet	57%
MobileNetV3	57.81%
ResNet50	57.63%

Results

RAG For X-ray images using ChatGroq

Enter Patient Data of Age, gender and Diagnosis

Age: 22, Gender: Male, Diagnosis: Effusion & Cardiomelagy

Based on the patient data provided, the patient is a 22-year-old male who has been diagnosed with effusion and cardiomegaly.

Effusion refers to the abnormal accumulation of fluid in the body, most commonly in the pleural space surrounding the lungs. Cardiomegaly, on the other hand, refers to an enlarged heart, which can be caused by various conditions such as high blood pressure, heart valve problems, or cardiomyopathy.

In this patient's case, the presence of effusion and cardiomegaly could suggest a number of possible etiologies. One possibility is that the effusion is causing the cardiomegaly, as the fluid accumulation can put pressure on the heart and cause it to enlarge. Alternatively, the cardiomegaly could be due to an underlying cardiac condition, such as alcoholic or hypertrophic cardiomyopathy, which can lead to heart enlargement and fluid accumulation as a result of heart failure.

The patient's young age and the absence of any mention of alcohol abuse or hypertension in the context provided suggest that alcoholic cardiomyopathy is less likely in this case. However, hypertrophic cardiomyopathy cannot be ruled out based on the information provided.

In addition, the presence of effusion and cardiomegaly could increase the risk of complications such as trapped lung and pneumothorax, particularly in the context of hernia repair and the presence of

The patient's young age and the absence of any mention of alcohol abuse or hypertension in the contex provided suggest that alcoholic cardiomyopathy is less likely in this case. However, hypertrophic cardiomyopathy cannot be ruled out based on the information provided.

In addition, the presence of effusion and cardiomegaly could increase the risk of complications such as trapped lung and pneumothorax, particularly in the context of hernia repair and the presence of subcutaneous emphysema. Trapped lung occurs when air becomes trapped in the lung and cannot be exhaled, while pneumothorax refers to the presence of air in the pleural space outside the lung. Both conditions can cause respiratory distress and require immediate medical attention.

In the presence of pleural effusion, the risk of trapped lung and pneumothorax may be higher due to the presence of fluid in the pleural space. In the presence of pneumothorax, the effusion may worsen by allowing air to accumulate in the pleural space, further compromising respiratory function.

Therefore, it is important to closely monitor this patient's respiratory status and consider further diagnostic testing, such as chest imaging, to evaluate the extent of the effusion and cardiomegaly and t assess for any complications such as trapped lung or pneumothorax. Treatment recommendations may include diuretics to reduce fluid accumulation, medications to manage any underlying cardiac conditio and respiratory support as needed.

Limitations and Future Scope

Limitations:
Computational resource requirements.
Dependency on high-quality bounding box annotations.

Future Scope:
Extend framework to other modalities like CT and MRI.
Incorporate federated learning for privacy-preserving training.
Continuous model improvement using external datasets.

Acknowledgments

Thanks to ICMR-NIRT Chennai, Dr. Ponnuraja, and VIT management for their support. Platforms like Radiopaedia and ScienceDirect provided invaluable resources.

INDIAN COUNCIL OF MEDICAL RESEARCH | NATIONAL INSTITUTE FOR RESEARCH IN TUBERCULOSIS

Thank You