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

Group Code (224616)

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

Abstract

Chest X-rays are a fundamental diagnostic tool for identifying a wide range of thoracic diseases. However, the interpretation of these images is often time-consuming and subject to variability among radiologists. This project aims to enhance multi-disease detection and automate the reporting process by integrating advanced X-ray image processing techniques with large language models (LLMs). Utilizing a comprehensive Chest X-ray dataset comprising 112,120 high-resolution (1024×1024) frontal-view PNG images across 14 common thoracic disease categories, along with detailed metadata and bounding box annotations, we develop a robust deep learning framework for accurate disease identification and localization. The dataset is strategically divided into training, validation, and testing sets on a patient-level to ensure unbiased evaluation. After training the image processing model to detect and classify multiple pathologies, the identified findings are input into a large language model to generate coherent and clinically relevant diagnostic reports automatically. This integrated approach not only aims to improve the accuracy and efficiency of disease detection but also seeks to standardize reporting, thereby reducing the workload on radiologists and minimizing human error. Preliminary results indicate promising performance in both multi-disease detection and automated report generation, demonstrating the potential of combining image processing and natural language processing technologies in medical imaging applications. This project underscores the potential of artificial intelligence to transform radiological practices, enhance diagnostic workflows, and ultimately improve patient outcomes.

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/

   Data Acquisition: The dataset includes 112,120 frontal-view X-ray images of 30,805 unique patients.
   Labeling Process: Disease labels were text-mined using natural language processing (NLP) from associated radiological reports. The accuracy of this method is >90%.

2. **Dataset Overview**
   Image Details: Each X-ray is in PNG format with a resolution of 1024x1024.
   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 Collection

3. **Why This Dataset is Valid for Our Project**

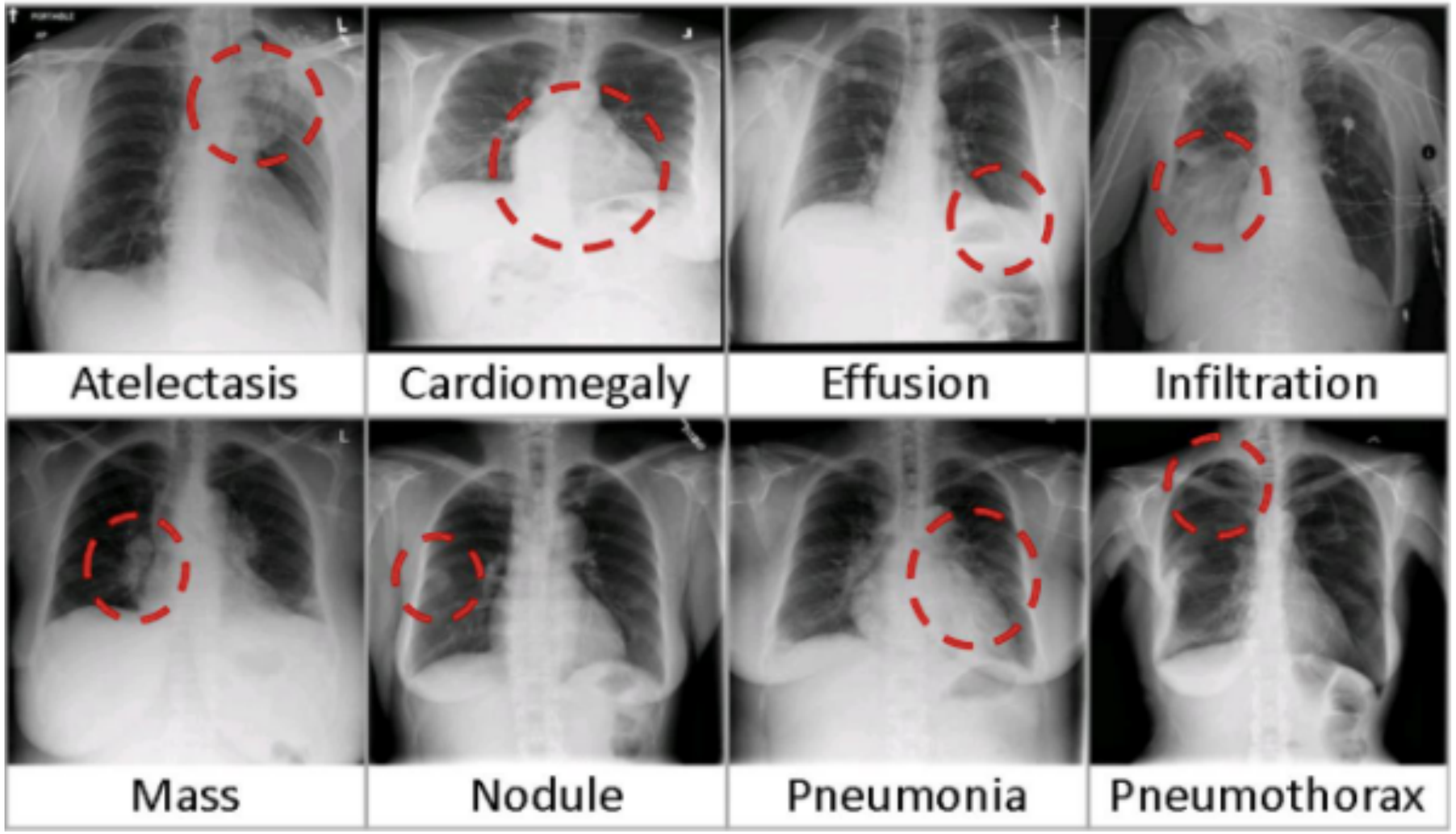
   Scale & Variety: With over 112,000 images and 30,805 patients, the dataset offers extensive diversity, which can improve model generalization.

   Multi-label Capability: Each image can have multiple disease labels, allowing for nuanced learning about thoracic pathologies.

   High Label Accuracy: Disease labels were mined with >90% accuracy, enhancing reliability.

   Benchmark Availability: Pre-existing models and performance benchmarks are available in literature, providing a reference point for evaluation.

Dataset Preview



Data Preprocessing

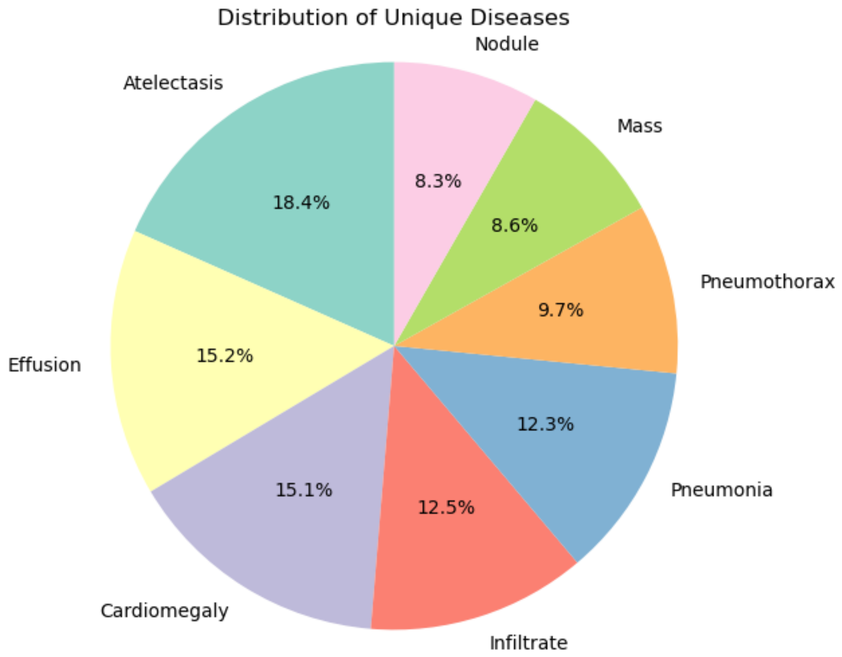
- **Preprocessing Steps:**
  - Normalization of Bounding Box Labels:
  - Adjusts bounding boxes for consistency across different images.
- **Vectorization:**
  - Input Vector: Pixel values from 1024x1024 grey-scaled images reshaped to (1024*1024, m).
  - Output Vector: Bounding box labels reshaped to (15*5, m).

- **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.

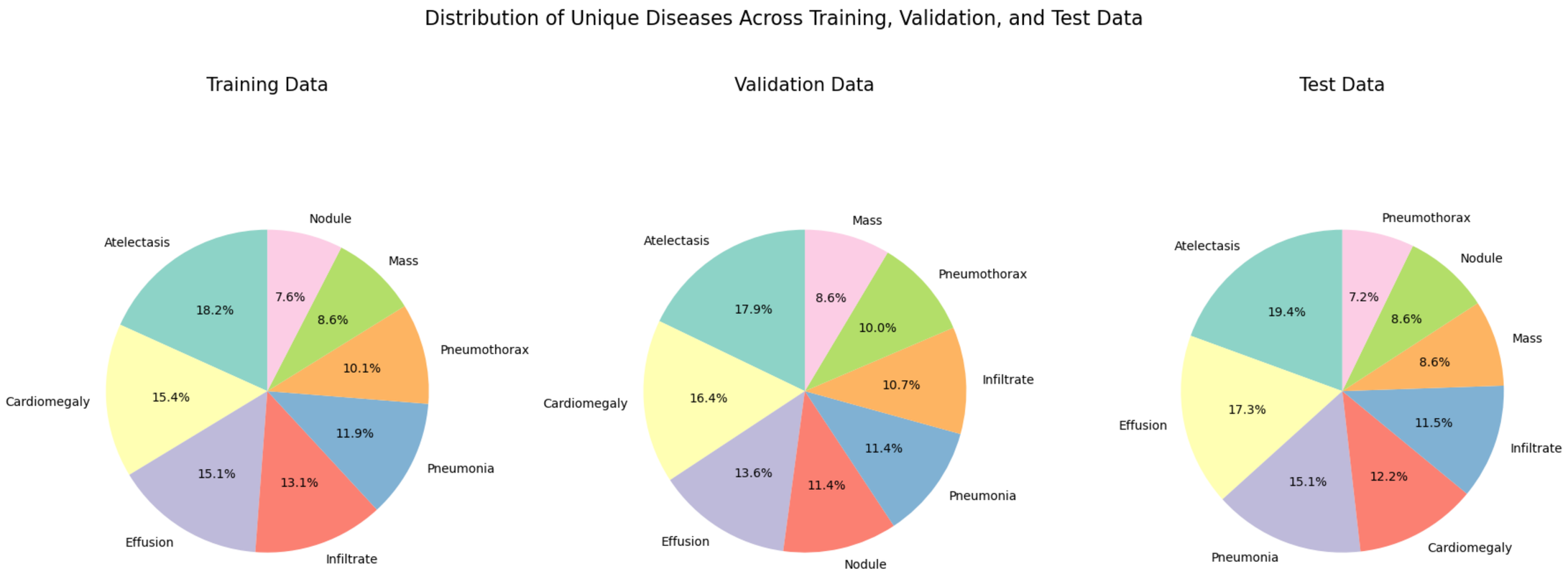
- **Model Preferences:**
  - NasNet and EfficientNet for high efficiency.
  - YOLO for real-time object detection.
  - ResNet/DenseNet for strong feature extraction.

EDA



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

EDA


Distribution of Unique Diseases Across Training, Validation, and Test Data
Training Data
Nodule
Atelectasis
Mass
7.6%
18.2%
8.6%
Pneumothorax
10.1%
Cardiomegaly
15.4%
11.9%
Pneumonia
15.1%
13.1%
Effusion
Infiltrate
Validation Data
Mass
Atelectasis
Pneumothorax
8.6%
17.9%
10.0%
Infiltrate
10.7%
Cardiomegaly
16.4%
11.4%
Pneumonia
13.6%
11.4%
Effusion
Nodule
Test Data
Pneumothorax
Atelectasis
Nodule
7.2%
19.4%
8.6%
Mass
8.6%
Effusion
17.3%
11.5%
Infiltrate
15.1%
12.2%
Pneumonia
Cardiomegaly

Methodology

Stage 1:

- Train and test object detection models on a diverse range of X-ray images (with and without anomalies).
- Aim to achieve maximum accuracy in detecting anomalies.

Stage 2:

- Use the trained model to derive inferences from the detected anomalies (e.g., position, size).
- Analyze these inferences to assess the severity of diseases or infections.

Methodology

Stage 3:

- Collect a large corpus of textual data related to all labels and conditions.
- Fine-tune pretrained Large Language Models (LLMs) using Retrieval-Augmented Generation (RAG) models with tools like LangChain.
- Generate clinical reports based on the detected anomalies, integrating additional patient medical data via prompts.

Acknowledgement

We sincerely thank ICMR-NIRT, Chennai for the insightful visit, especially Dr. C. Ponnuraja for his session on tuberculosis and intern Krishna for his talk on X-ray diagnosis, which will aid our project team.

Special thanks to Dr. Sathya Narayana Sharma K. for organizing this visit and to VIT management for their support in making this enriching experience possible.

Thank You!