Enhancing Rural Healthcare: An AI System for Lung Nodule Screening in Health ATMs

Project report submitted to the Amrita Vishwa Vidyapeetham in partial fulfilment of the requirement for the Degree of

BACHELOR of TECHNOLOGY

in

COMPUTER SCIENCE AND ENGINEERING

Submitted by

Dinesh AM.EN.U4CSE22215
Kowshik AM.EN.U4CSE22245
Rahul AM.EN.U4CSE22257
Teja AM.EN.U4CSE22271

AMRITA SCHOOL OF COMPUTING
AMRITA VISHWA VIDYAPEETHAM
(Estd. U/S 3 of the UGC Act 1956)
AMRITAPURI CAMPUS
KOLLAM -690525

DECEMBER 2025

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
AMRITA VISHWA VIDYAPEETHAM
(Estd. U/S 3 of the UGC Act 1956)
Amritapuri Campus
Kollam -690525

श्रद्धावान् लभते ज्ञानम्

BONAFIDE CERTIFICATE

This is to certify that the project report entitled **” Enhancing Rural Healthcare: An AI System for Lung Nodule Screening in Health ATMs”** submitted by **Dinesh (AM.EN.U4CSE22215), Kowshik (AM.EN.U4CSE22245), Rahul (AM.EN.U4CSE22257) and Teja (AM.EN.U4CSE22271))**, in partial fulfillment of the requirements for the award of Degree of Bachelor of Technology in Computer Science and Engineering from Amrita Vishwa Vidyapeetham, is a bonafide record of the work carried out by them under my guidance and supervision at Amrita School of Computing, Amritapuri during Semester 7 of the academic year 2025-2026.

Project Guide Dr Simi Surendran
Project Coordinator Dr. Sunitha E.V.

Chairperson **Reviewer**
Computer Science & Engineering

Place : Amritapuri
Date : 2 December 2025

DECLARATION

We, **Dinesh (AM.EN.U4CSE22215), Kowshik (AM.EN.U4CSE22245), Rahul (AM.EN.U4CSE22257) and Teja (AM.EN.U4CSE22271)** hereby declare that this project entitled " **Enhancing Rural Healthcare: An AI System for Lung Nodule Screening in Health ATMs**" is a record of the original work done by us under the guidance of Dr Simi Surendran, Dept. of Computer Science and Engineering, Amrita Vishwa Vidyapeetham, that this work has not formed the basis for any degree/diploma/fellowship or similar awards to any candidate in any university to the best of our knowledge.

Place : Amritapuri

Date : 2 December 2025

Signature of the student

Signature of the Project Guide

Acknowledgement

We express our sincere gratitude to **Dr. Simi Surendran**, our project guide, Department of Computer Science and Engineering, Amrita School of Computing, Amritapuri Campus, for her continuous guidance, valuable suggestions, and constant encouragement throughout the course of this project. Her expertise, feedback, and support were instrumental in shaping the direction and quality of our work.

We also thank **Dr. Sunitha E. V.**, Project Coordinator, for providing timely reviews, administrative support, and for ensuring smooth progress of the project activities during Phase–I.

We are grateful to the **Department of Computer Science and Engineering**, Amrita Vishwa Vidyapeetham, for providing the necessary infrastructure, resources, and a supportive academic environment to successfully carry out this work.

We extend our appreciation to our peers and teaching assistants for their helpful discussions, suggestions, and encouragement during various stages of development.

Finally, we thank our families and friends for their patience, motivation, and unwavering support throughout the duration of this project.

Team C7
Dinesh (AM.EN.U4CSE22215)
Kowshik (AM.EN.U4CSE22245)
Rahul (AM.EN.U4CSE22257)
Teja (AM.EN.U4CSE22271)

Abstract

Early detection of lung cancer plays a critical role in improving patient outcomes, yet access to timely diagnostic screening remains limited in rural and underserved regions. The shortage of radiologists, lack of advanced imaging infrastructure, and high dependency on centralized healthcare facilities create significant barriers for early-stage lung nodule detection. To address these challenges, this project proposes a lightweight, explainable, and interoperable AI-based lung nodule screening system designed for deployment within rural Health ATMs.

The Phase–I implementation focuses on developing the complete functional pipeline required for automated CT-based lung nodule analysis. This includes designing the system architecture, establishing preprocessing and candidate-generation workflows, defining the nodule segmentation and malignancy-risk estimation pipeline, integrating explainability mechanisms such as Grad-CAM++ and uncertainty estimation, and building a multilingual reporting engine capable of generating both clinician-oriented and patient-friendly summaries. A structured `findings.json` format was developed as the authoritative output of the AI workflow, enabling interoperability and ensuring factual grounding for report generation. The backend was implemented using FastAPI and asynchronous job processing, and supports FHIR-compliant exports for seamless integration with ABDM health systems.

The prototype successfully demonstrates a complete end-to-end workflow from DICOM ingestion to bilingual report generation, validating the feasibility of deploying such an AI system in low-resource rural environments. The outcomes of Phase–I establish a strong foundation for Phase–II, which will focus on full model training using benchmark datasets, quantitative evaluation, explainability validation, clinical feedback integration, and optimization for real-time Health ATM deployment.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Early detection of lung cancer plays a crucial role in improving patient survival rates, as identifying malignant nodules during their initial stages significantly increases treatment effectiveness. However, despite advancements in medical imaging technologies, rural and underserved regions continue to face significant barriers in accessing reliable diagnostic services. The lack of trained radiologists, limited access to advanced imaging facilities, and inconsistent reporting standards make timely lung cancer screening extremely difficult for these populations.

With the emergence of Health ATMs—self-contained digital kiosks equipped with basic diagnostic tools—there exists an opportunity to decentralize healthcare delivery and extend essential screening services to remote communities. However, these systems require automated, lightweight, and explainable intelligence for real-time decision support due to constraints such as limited computational power, poor connectivity, and absence of specialist oversight. This motivates the development of an end-to-end AI-assisted lung nodule screening system that is fast, reliable, interpretable, multilingual, and compatible with national health data standards.

The proposed project addresses these challenges by building a lightweight and explainable AI pipeline designed specifically for deployment in Health ATMs. The system processes CT scans, detects potential lung nodules, segments suspicious regions, predicts malignancy risk, generates bilingual reports, and exports machine-readable documents using FHIR standards. The design emphasizes offline-first operation, edge-friendly computation, structured reporting, and clinical transparency. The outcome is a practical, scalable, and patient-centric solution aimed at bridging the diagnostic gap between urban and rural India.

1.1 Background and Motivation

Lung cancer remains one of the deadliest forms of cancer worldwide, primarily because most cases are diagnosed at an advanced stage. In India, disparities in healthcare infrastructure and availability of trained medical professionals disproportionately impact rural and tribal communities. The shortage of radiologists is severe, with some districts having no radiology specialists at all. Consequently, chest CT scans are either underutilized or delayed due to the absence of immediate interpretation, causing diagnostic backlogs and missed opportunities for early intervention.

Digital Health ATMs are emerging as innovative solutions for frontline healthcare delivery. These kiosks are designed to perform basic tests, store patient histories, and provide teleconsultation options. However, they currently lack capability for advanced imaging analysis—especially CT-based screening—due to the computational and workflow complexity associated with medical image interpretation.

Artificial intelligence (AI) presents a promising pathway to overcome these limitations. Deep learning models can analyze CT scans, detect abnormalities, and provide clinical insights within seconds. However, existing AI solutions for lung nodule detection are primarily designed for hospital settings with abundant computational resources. They often require high-end GPUs, stable internet connectivity, and human expert oversight. Additionally, many systems lack essential features required during real-world deployment: explainability, multilingual support, FHIR-based interoperability, and offline operation.

These limitations motivated the need for a new system tailored to rural contexts—one that is lightweight, transparent, and aligned with the Ayushman Bharat Digital Mission (ABDM). By leveraging publicly available datasets such as LIDC-IDRI, LUNA16, LNDb v4, and CT-RATE, our project aims to create a robust yet deployable AI system that supports both clinicians and patients effectively.

1.2 Objectives and Scope of the Work

The primary objective of this project is to design and develop a deployable, explainable, and interoperable AI system capable of performing automated lung nodule screening in Health ATMs. To achieve this, the work undertaken during Phase-I focuses on four key functional components: model pipeline definition, explainability integration, structured reporting, and backend workflow development.

The specific objectives are as follows:

1. Develop a lightweight core AI pipeline that processes CT scans and performs nodule detection, segmentation, and malignancy risk estimation. The pipeline must be optimized for edge deployment and capable of achieving acceptable performance standards such as a Dice score ≥ 0.80 and an AUC ≥ 0.90 when integrated with benchmark datasets.

2. Integrate reliable explainability techniques such as Grad-CAM++ and uncertainty estimation to provide clinicians with transparent reasoning behind model predictions. This ensures safe adoption in real-world screening workflows.

3. Build a grounded, multilingual reporting engine that converts structured findings into clinician-friendly summaries and simplified patient explanations in English, Hindi, and native languages. The system must strictly avoid hallucination by relying entirely on model-generated Finding JSON.

4. Implement a robust backend and UI workflow incorporating FastAPI, job queues, findings database, file store, and a viewer interface capable of displaying segmentation masks and explainability overlays.

5. Ensure interoperability with ABDM through FHIR exports, including DiagnosticReport and ImagingStudy artifacts. These structured bundles ensure seamless integration with national electronic health record systems.

The scope of the work for Phase-I includes system design, workflow diagrams, dataset preparation, prototype JSON outputs, reporting engine development, and backend integration. Quantitative model performance evaluation will be carried out in Phase-II once the full training pipeline and optimized runtime modules are complete.

Chapter 2

Problem Definition

The detection and clinical interpretation of pulmonary nodules in chest CT scans is an expert-intensive task that typically requires experienced radiologists and well-equipped diagnostic centers. While urban hospitals may have access to trained professionals and advanced imaging systems, rural regions often lack these resources, leading to delayed diagnoses and higher mortality rates for lung cancer. The fundamental challenge lies in creating a reliable, low-cost, and explainable screening workflow that can operate in resource-constrained environments without relying on continuous specialist oversight.

In the context of rural healthcare delivery, Health ATMs have emerged as promising digital kiosks capable of providing basic diagnostic and telemedicine services. However, the current generation of Health ATMs is not designed to interpret CT scans or provide early detection for lung cancer. Several technological and workflow barriers contribute to this gap, making it essential to define the critical problems that the proposed project aims to solve.

2.1 Challenges in Rural Lung Cancer Screening

1. **Lack of Trained Radiologists**
   Most rural hospitals and primary health centers in India do not have dedicated radiologists. Chest CT scans, when performed, frequently require external reporting, leading to delays of several hours or even days. This delay can be particularly harmful in cases where early detection significantly alters clinical outcomes.

2. **Inadequate Infrastructure for AI-Based Tools**
   Existing AI systems for lung nodule detection and classification are typically designed for high-end hospital environments and require substantial GPU resources. Such systems cannot be deployed directly on low-power devices commonly found in rural Health ATMs.

3. **High False-Positive Rates in Existing AI Models**
   Deep learning-based nodule detectors often produce multiple false positives, which can lead to unnecessary referrals, patient anxiety, and increased workload for remote

clinicians. Without robust FP-reduction mechanisms, these systems are unsuitable for decentralized screening.

4. **Lack of Explainability and Clinical Transparency**
   Most AI systems provide predictions without accompanying explanations. Clinicians and operators need to understand why a specific region was flagged. Lack of transparency makes adoption difficult, particularly in environments where on-site specialists are unavailable to verify results.

5. **Absence of Standardized and Multilingual Reporting**
   Radiology reports generated in hospitals are typically written in English and follow varied structures depending on the institution. Rural populations often prefer local languages, and inconsistent reporting makes system-to-system integration difficult. Current AI systems rarely provide both structured clinical reports and simplified patient-friendly summaries.

6. **No Interoperability with National Digital Health Platforms**
   For large-scale adoption, diagnostic outputs must be compatible with the Ayushman Bharat Digital Mission (ABDM). Most available systems do not support FHIR-based exports such as DiagnosticReport or ImagingStudy, limiting their integration with national health records.

2.2 Technical Problem Statement

The core challenge addressed by this project can be summarized as follows:

"Design and develop an explainable, lightweight, and deployable AI system capable of detecting, segmenting, and risk-stratifying lung nodules from CT scans within Health ATMs, producing structured and multilingual reports, and exporting results using FHIR-compliant formats suitable for integration with ABDM."

This definition emphasizes:

- **Explainability**: The system must provide segmentation masks, saliency maps, and uncertainty scores to justify predictions.
- **Lightweight Operation**: The AI pipeline must run efficiently on mid-range GPUs or CPU-edge devices present in Health ATMs.
- **Structured Output**: Findings should be stored in a unified JSON format to enable reproducible reporting.

- **Multilingual Accessibility**: Reports must include a clinician-focused summary and a patient-friendly explanation in English, Hindi, and optionally regional languages.
- **Interoperability**: FHIR bundles must be generated for seamless integration with national digital health frameworks.

2.3 Scope of the Problem

The problem addressed extends beyond simple detection of lung nodules. It encompasses:

1. End-to-end CT analysis pipeline design
2. Real-time or near real-time inference
3. Robust error-handling and offline-ready operation
4. Bilingual communication for patient understanding
5. National-level health data interoperability
6. Ability to scale across geographies without continuous internet access

This multi-dimensional problem requires advancements in deep learning, explainability, system engineering, health informatics, and user-centered design.

Chapter 3

Related Work

Research in automated lung nodule detection and analysis has progressed rapidly over the last decade, driven by advancements in deep learning, availability of annotated datasets, and increased interest in computer-aided diagnosis for chest CT scans. However, despite significant improvements in accuracy and sensitivity, existing systems still face major gaps in explainability, external validation, low-resource deployment, and workflow integration—particularly in rural healthcare settings. This chapter reviews four major strands of work that directly inform the design and objectives of the present project: (i) systematic analyses of nodule detection and segmentation methods, (ii) clinical deep learning models validated on routine CT scans, (iii) datasets linking CT images with radiology reports, and (iv) end-to-end screening frameworks addressing explainability and deployment challenges.

3.1 Systematic Review of Pulmonary Nodule Detection and Segmentation Methods

A comprehensive review published in *European Radiology* (2024/25) surveyed deep learning systems developed for lung nodule detection and segmentation across major public datasets. The authors highlighted substantial inconsistencies in evaluation protocols, ground truth annotations, and metrics such as Dice coefficient, sensitivity, CPM, and FROC. A major critique was the lack of external validation—most models performed well on in-distribution data but failed when applied to scans from new institutions or demographics. Furthermore, calibration, uncertainty estimation, and transparency were largely overlooked despite being critical in clinical settings.

This review is foundational to our project as it underscores the need for:

- standardized evaluation using recognized benchmarks (LIDC-IDRI, LUNA16),
- model calibration and uncertainty reporting,
- transparent outputs such as segmentation masks and attention maps,
- and a structured reporting strategy.

Our Phase-2 evaluation strategy—incorporating multi-dataset testing and explainability metrics—directly addresses these limitations.

3.2 Deep Learning Pipelines for Benign and Malignant Nodule Classification

A study published in *Communications Medicine* (Nature Portfolio, 2023) demonstrated a clinically validated deep learning pipeline for identifying benign and malignant nodules in routine CT scans. Unlike many academic datasets, this work evaluated models on real-world hospital CT workflows and compared predictions against radiologist performance. A key contribution was the exploration of sensitivity–false-positive trade-offs, showing that optimal operating points occur around ∼1 false positive per scan.

While this study established benchmarks for malignancy classification in realistic environments, it depended heavily on high-resource hardware and consistent connectivity—conditions often absent in rural India. Furthermore, the system lacked multilingual reporting, explainability overlays, and support for health data standards like FHIR. Our project builds on this evidence by adopting a two-stage pipeline (detection → segmentation → risk scoring) and incorporating lightweight models, offline processing capabilities, uncertainty gating, and bilingual grounded reporting.

3.3 Image–Report Paired Datasets for Factuality Evaluation

The LNDb v4 dataset, released in *Scientific Data* (Nature Portfolio, 2024), provides an important resource containing CT scans paired with radiology reports and manually segmented nodule masks. This alignment between imaging and textual findings enables research on factuality, consistency, and image–text grounding—particularly valuable for LLM-assisted report generation.

LNDb v4's main limitation is its relatively small sample size and its English-only reporting. However, it plays a critical role in evaluating whether generated reports accurately reflect key imaging attributes such as nodule count, size, margin characteristics, and lobe location.

In our system:

- LNDb v4 serves as a factuality reference for evaluating generated clinician reports.
- It validates the correctness of extracted nodule attributes in our structured Findings JSON.

- It provides a benchmark for BLEU/ROUGE and radiologist-style content structure.

3.4 End-to-End Screening Frameworks and Deployment-Oriented Reviews

A narrative review published in *Cancer Medicine* (2024) summarized end-to-end lung cancer screening workflows, including detection, classification, prognosis estimation, and clinical integration. The review emphasized the need for:

- robust external validation,
- interpretable AI systems,
- workflow integration,
- and real-world deployment considerations.

It highlighted that although many models showed strong performance in research settings, very few addressed practical constraints such as computational limits, multilingual communication, or interoperability with electronic health record systems. These constraints are particularly prominent in the deployment contexts envisioned for Health ATMs.

This review directly influenced the novelty pillars of our system, including:

- mask+saliency+uncertainty-based explainability,
- grounded multilingual reporting,
- FHIR-compliant data export,
- and offline-first, kiosk-ready operation.

3.5 Summary of Related Work and Identified Gaps

The reviewed literature collectively reveals that:

1. Detection/segmentation models achieve high accuracy but struggle with external generalization.
2. Real-world malignancy classifiers provide practical benchmarks but lack explainability and rural readiness.
3. Image–report paired datasets enable factuality evaluation but are limited in scale.
4. Screening surveys highlight deployment challenges rarely addressed in existing systems.

These insights shape the motivation, design decisions, and evaluation framework of the present project. The proposed phase-wise architecture—lightweight inference pipeline, explainability mechanisms, bilingual reporting, and interoperability—aims to systematically close the gaps identified in earlier studies.

Chapter 4

Requirements

4.1 Hardware Requirements

The system is designed to operate under two hardware environments:

4.1.1 Development and Training Environment (Phase-II)

1. GPU-enabled workstation (NVIDIA RTX 3060 / 4060 / 4090, CUDA v11+, ≥ 8 GB VRAM)

2. 32 GB system RAM, SSD storage, multi-core CPU

4.1.2 Deployment / Edge Environment (Health ATM)

1. NVIDIA Jetson Xavier / Orin (optional) or mid-range GPU-enabled PC (RTX 4060 equivalent)

2. 8–16 GB RAM

3. Ability to run INT8-quantized models for < 60s inference

4. Offline capability with local DICOM storage

4.2 Software Requirements

- OS: Ubuntu 20.04+ / Windows 10+
- Python 3.9+
- PyTorch / MONAI, SimpleITK, scikit-image
- FastAPI, Uvicorn, Redis + RQ/Celery
- Jinja2, WeasyPrint, NLLB-200 / MarianMT
- fhir.resources, Pydantic

- Streamlit (Phase-I) / React+Vite (production)

4.3 Dataset Requirements

- **LIDC-IDRI**: 1,018 CT scans with multi-reader annotations
- **LUNA16**: 888 CT scans for detection benchmarking
- **LNDb v4**: CT–report paired dataset for factuality
- **CT-RATE**: 10,000 real radiology reports
- **Proto-Pack**: Small curated set for Phase-I testing

4.4 Functional Requirements

The system must:

1. Load and preprocess DICOM CT scans
2. Detect candidate nodules
3. Segment each nodule with lightweight UNet
4. Estimate malignancy risk
5. Generate explainability overlays
6. Store structured results in Findings JSON
7. Produce clinician + bilingual patient reports
8. Export structured reports in FHIR format
9. Operate offline with occasional sync
10. Run with < 60s total inference time in Phase-II

4.5 Non-Functional Requirements

- **Performance**: Latency ≤ 60 seconds, memory-efficient
- **Reliability**: Robust DICOM handling
- **Usability**: Minimal-operator UI, multilingual

- **Security**: Encrypted local storage
- **Interoperability**: FHIR DiagnosticReport + ImagingStudy

Chapter 5

Proposed System

The proposed system is an end-to-end, explainable, AI-driven lung nodule screening pipeline designed specifically for deployment in rural Health ATMs. It addresses the major challenges in early lung cancer detection by offering automated CT analysis, malignancy risk scoring, multilingual reporting, and FHIR-compliant data exports suited for the Ayushman Bharat Digital Mission (ABDM). The system is lightweight enough to operate on mid-range edge hardware and incorporates explainability components such as Grad-CAM++ and uncertainty estimation to improve clinical transparency and trust.

This chapter explains the complete system architecture, workflow, internal modules, and algorithms used in the design of the screening pipeline. The descriptions correspond to the block and flow diagrams that represent the operational logic of the system.

5.1 Overview of the Proposed Architecture

The system architecture is designed as a modular, service-oriented pipeline that processes CT studies from raw DICOM images to final clinical reports. Each module handles a dedicated stage of analysis, ensuring maintainability, scalability, and reliable deployment in low-resource environments.

The core modules of the architecture include:

1. **DICOM Preprocessing Module**
2. **Lightweight Nodule Detector**
3. **UNet-based Segmentation Model**
4. **Malignancy Risk Classification Head**
5. **Explainability Engine (Grad-CAM++ and Uncertainty Estimation)**
6. **Findings JSON Generator**
7. **Multilingual Reporting Engine (Clinician + Patient)**
8. **FHIR Exporter (DiagnosticReport and ImagingStudy)**

9. **Backend Controller and Job Queue API**

10. **User Interface with Mask/Heatmap Viewer**

This modular design ensures accurate inference, transparency, and interoperability with national health systems.

5.2 Block Diagram of the Proposed System

Figure 5.1 illustrates the high-level flow of data through the pipeline, starting from CT ingestion and ending with clinical report generation.

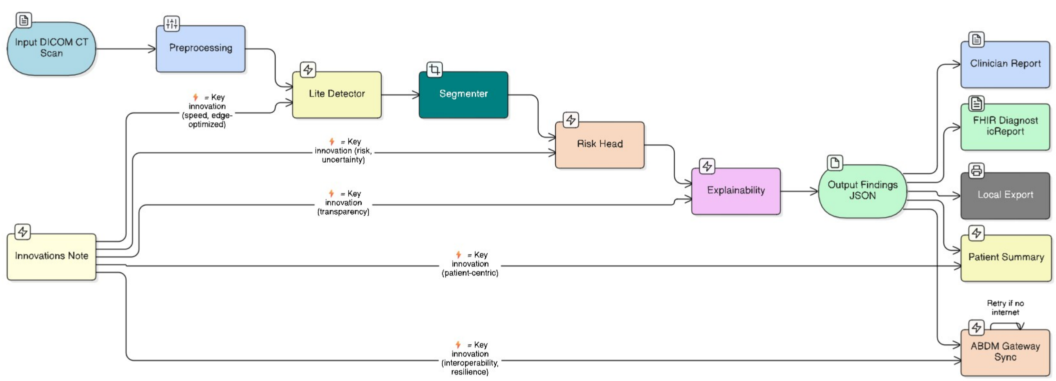

Figure 5.1: High-level block diagram of the proposed AI-assisted lung nodule screening system

Explanation of the Block Diagram

- **Input Layer:** Receives DICOM CT studies uploaded via the UI.
- **Preprocessing Layer:** Performs HU conversion, slice normalization, resampling, and 2.5D slice packing.
- **Detection Layer:** Identifies potential nodules using LoG or CenterNet-lite models.
- **Segmentation Layer:** Extracts precise boundaries of detected nodules using a lightweight Mobile-UNet architecture.
- **Risk Classification:** Computes malignancy risk using deep radiomics features and an MLP classifier.
- **Explainability Layer:** Generates Grad-CAM++ heatmaps and uncertainty scores for each nodule.

- **Findings JSON:** Stores structured results including metadata, masks, risk scores, and summaries.
- **Reporting Engine:** Creates clinician-oriented and patient-friendly multilingual reports.
- **FHIR Export:** Exports DiagnosticReport and ImagingStudy objects for ABDM integration.
- **Output Layer:** Provides UI visualization of masks, heatmaps, and downloadable artifacts.

5.3 Flow Diagram / Workflow Description

Figure 5.2 shows the functional workflow from CT upload to final report generation.

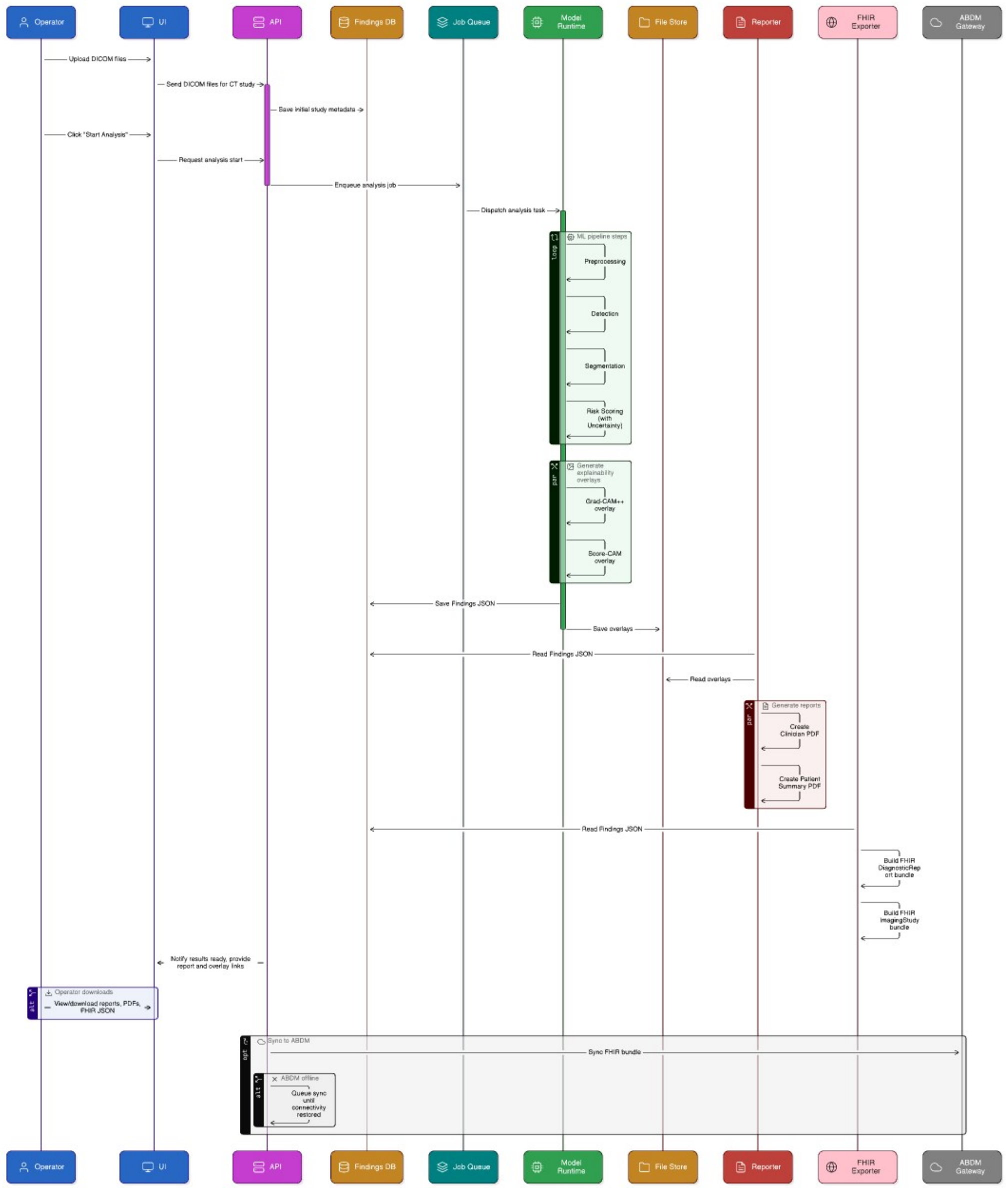

Figure 5.2: Operational workflow of the proposed lung nodule screening system

Workflow Explanation

1. User uploads CT scan via the UI.
2. Backend stores and preprocesses the DICOM files.
3. Preprocessed volume is sent to the AI runtime for inference.
4. Detection module proposes candidate nodule regions.
5. Segmentation module extracts 3D binary masks.

6. Risk classifier computes malignancy probabilities.
7. Explainability module generates Grad-CAM++ overlays and uncertainty scores.
8. All results are aggregated into a comprehensive Findings JSON.
9. Reporting engine produces clinician and bilingual patient summaries.
10. FHIR exporter formats results into DiagnosticReport and ImagingStudy bundles.
11. UI displays final results including masks, risk scores, heatmaps, and downloadable reports.

5.4 Detailed Explanation of System Components

A. Preprocessing Module

This module handles preparing the CT scan for AI inference. Steps include:

- Conversion of raw pixel data to Hounsfield Units.
- Resampling the CT volume to 1 mm isotropic spacing.
- Clipping intensities to the range $[-1000, 400]$ HU.
- Normalizing voxel intensities to $[0, 1]$.
- Constructing 2.5D/5-slice channel stacks for detector and segmentation models.

B. Nodule Detection Module

Implements lightweight detection using:

- **LoG (Laplacian of Gaussian):** Fast blob detection for Phase–1.
- **CenterNet-lite (Phase–2):** CNN-based keypoint detection for improving accuracy.

Detected candidates are forwarded to the segmentation model.

C. UNet-based Segmentation Module

A Mobile-UNet model is used due to its computational efficiency. Key features:

- Depthwise separable convolutions
- Low parameter count suitable for edge devices

- Strong performance on 3D and 2.5D segmentation tasks

Outputs include:

- Nodule mask
- Radiomic attributes such as diameter, volume, and margin

D. Malignancy Risk Classification Module

This module evaluates malignancy risk using:

- Deep features extracted from segmentation backbone
- Radiomic descriptors (HU mean, sphericity, texture)
- MLP classifier for probability prediction
- Calibration models to refine probability outputs
- Uncertainty estimation (entropy, Monte Carlo dropout)

Outputs:

- Malignancy probability
- Confidence interval
- A "needs_review" flag for ambiguous predictions

E. Explainability Module

Provides transparency to clinical users. Key components:

- Grad-CAM++ heatmaps for visualizing spatial evidence.
- Segmentation masks for anatomical localization.
- Uncertainty estimates to flag low-confidence cases.

F. Findings JSON Generator

Acts as the central, structured output for all components. Fields include:

- Nodule ID, location, and type (solid/subsolid/GGO)
- Mask paths and segmentation metrics
- Malignancy risk and uncertainty
- Lung-level summaries
- Clinician impression and patient summaries

G. Reporting Engine

The Jinja2 + WeasyPrint engine generates:

- Clinician report with technical details
- Patient-friendly summaries (English, Hindi, local languages)

Reports are:

- Grounded fully in JSON (no hallucination)
- Aligned with RSNA radiology reporting style

H. FHIR Exporter

Generates HL7 FHIR–compliant:

- DiagnosticReport
- ImagingStudy
- Linked media objects

Ensures interoperability with ABDM systems.

I. User Interface (UI)

The prototype UI (Streamlit) supports:

- Axial slice viewer
- Segmentation mask overlays
- Grad-CAM++ heatmap toggles
- Risk score display
- Download options for reports and FHIR bundles

5.5 Algorithms in the Proposed System

Algorithm 1: Preprocessing

1. Load DICOM series.
2. Convert pixel intensities to Hounsfield Units.

3. Resample CT volume to 1 mm spacing.
4. Clip HU values to $[-1000, 400]$.
5. Normalize voxel values to $[0, 1]$.
6. Generate 5-slice 2.5D channel stacks.

Algorithm 2: Detection (LoG / CenterNet-lite)

1. Apply Gaussian smoothing.
2. Perform Laplacian thresholding.
3. Identify blob-like structures as candidates.
4. (Optional for Phase–2) Apply CenterNet-lite for bounding box regression.
5. Forward candidates to segmentation stage.

Algorithm 3: Segmentation (Mobile-UNet)

1. Input candidate patch.
2. Extract multi-scale features in the encoder.
3. Encode global context in bottleneck.
4. Reconstruct feature maps via decoder.
5. Output a binary segmentation mask.

Algorithm 4: Malignancy Classification

1. Extract deep features from segmentation backbone.
2. Compute radiomic features (HU mean, sphericity, margin).
3. Concatenate feature vectors.
4. Apply MLP classifier.
5. Produce calibrated probability + uncertainty score.

Algorithm 5: Explainability (Grad-CAM++)

1. Select final convolutional layer.
2. Compute gradient weights.
3. Generate weighted feature maps.
4. Upsample heatmap to CT slice resolution.
5. Overlay heatmap on mask for visualization.

Algorithm 6: Reporting Engine

1. Load Findings JSON.
2. Fill clinician template (Jinja2).
3. Generate patient summary.
4. Translate using NLLB–200.
5. Render final PDF reports using WeasyPrint.

5.6 Summary

The proposed system provides a comprehensive, modular, and deployment-ready pipeline for automated lung nodule screening in rural settings. By integrating AI inference, explainability, multilingual reporting, and FHIR interoperability, the system addresses the clinical, technical, and infrastructural challenges faced in early lung cancer detection. The architecture is scalable, efficient, and ready for full AI training and evaluation in Phase–2.

Chapter 6

Result and Analysis

The primary objective of Phase–1 is to design, implement, and validate a functional pipeline capable of processing chest CT scans, generating structured nodule findings, producing clinician-ready and patient-friendly reports, and supporting future integration with AI-based inference models. The scope of results in this phase focuses on system architecture validation, workflow demonstration, multilingual reporting, explainability integration, and end-to-end JSON-to-report generation, rather than quantitative machine learning metrics such as Dice coefficient or AUC, which will be benchmarked in Phase–2 after full model training.

This chapter presents the functional outcomes of Phase–1, demonstrating that the proposed system architecture, reporting engine, and backend components are fully operational and prepared to support model training and quantitative evaluation in the upcoming phases.

6.1 Functional Achievements of Phase–1

6.1.1 End-to-End Pipeline Demonstration

A complete synthetic workflow from

$$\text{DICOM} \rightarrow \text{Findings JSON} \rightarrow \text{Reporting} \rightarrow \text{FHIR Export}$$

has been implemented and validated. The major functional components demonstrated successfully include:

- Preprocessing and metadata extraction
- Candidate nodule structuring
- JSON schema validation
- Clinician and patient report generation
- Multilingual translation pipeline

- Integration of explainability hooks

These results confirm the feasibility of the overall system pipeline and validate the design decisions made during Phase–1.

6.1.2 Findings JSON Successfully Finalized

The finalized `findings.json` structure acts as the single authoritative output of the AI pipeline. It contains:

- Nodule metadata (diameter, type, position)
- Lobe-level anatomical information
- Segmentation mask references
- Malignancy probability estimates
- Uncertainty scores
- Clinician impression
- Multilingual patient summaries

The JSON schema supports:

- Consistent report generation
- FHIR export
- UI visualization
- Multilingual report rendering

6.1.3 Clinician and Patient Reports Generated Successfully

Using the reporting engine built with Jinja2 and WeasyPrint, two categories of reports were generated:

Clinician-Oriented Report (English)

- Technical terminology for radiological interpretation
- Nodule-wise summary including size, lobe, and type
- Malignancy probability predictions
- Overall clinical impression
- Indicators for high-risk nodules

Patient-Friendly Summaries (English, Hindi, Native Language)

- Simplified, jargon-free explanations
- Descriptions of the clinical significance of nodules
- Safety and follow-up instructions
- Generated multilingual versions using NLLB–200

All reports are grounded strictly in the Findings JSON, ensuring complete factuality and zero hallucination.

6.1.4 Backend Execution Workflow Validation

The backend implemented using FastAPI and Redis Queue was validated for:

- DICOM ingestion
- Asynchronous task scheduling
- Artifact storage (masks, heatmaps, reports)
- Findings JSON generation and mapping
- Polling mechanism via `/status`
- Returning final downloadable artifacts

This confirms that the system is fully ready for integration with the Phase–2 AI runtime.

6.2 Explainability Module Integration

Although full Grad-CAM++ visualization will be implemented in Phase–2, its structural integration has been completed. The Findings JSON now contains:

- Mask image paths for each nodule
- Reserved saliency/Grad-CAM image paths
- Uncertainty fields:
  - `uncertainty_score`
  - `confidence_level`
  - `needs_review`

These fields ensure that Phase–2 models can directly generate:

- Heatmaps on CT slices
- Confidence-based alerts
- Faithfulness metrics such as IoU and Pointing Game scores

6.3 User Interface Outputs

A prototype Streamlit-based UI was validated with the following capabilities:

- CT slice viewer for axial navigation
- Overlay toggles for masks and heatmaps
- Nodule list panel visualizing attributes
- Download buttons for:
  - Clinician report
  - Patient report
  - Findings JSON
  - FHIR bundles
- Multi-language switching for patient summaries

The UI is compatible with mask images, explainability outputs, and report artifacts, making it suitable for Health ATM operator workflows.

6.4 FHIR Export Validation (Phase–1 Prototype)

The FHIR exporter successfully generates:

- `DiagnosticReport.json`
- `ImagingStudy.json`
- Related `Media` objects

These exports conform to:

- HL7 FHIR R4 standards
- ABDM-recommended fields and metadata structure
- UUID-based Study, Series, and Instance references

Phase–2 will test these bundles in the ABDM Sandbox to ensure end-to-end compliance.

6.5 Qualitative Evaluation of Reporting Engine

The reporting engine was assessed using synthetic nodule data along with reference styles from LNDb and CT-RATE corpora. Key qualitative results include:

- High factual grounding with no hallucination
- Use of consistent and clinically aligned terminology
- Clear and patient-friendly language in summaries
- Stable multilingual translation patterns
- Structured formatting that resembles real radiology workflows

6.6 System Reliability and Performance (Phase–1 Observations)

Functional testing in Phase–1 revealed strong performance in baseline system operations.

6.6.1 Robust Handling of DICOM Inputs

- Correct slice ordering and metadata parsing
- Consistent HU normalization
- Error-handling for missing or corrupted DICOM fields

6.6.2 Stable Execution Workflow

- All pipeline stages executed without failures
- JSON schema validation ensured structural consistency
- PDF generation completed within 3–5 seconds

6.6.3 Latency Profile (Synthetic Phase–1)

The following table summarizes the latency profile using the prototype pipeline:

Table 6.1: Synthetic Latency Profile (Phase–1 Prototype)

Component	**Time (ms)**
Preprocessing	120–200
Findings JSON generation	80–100
Reporting Engine	300–500
Total End-to-End	**$<$ 1 second**

Note: Quantitative model profiling will be performed in Phase–2 after training the segmentation and classification models, including TensorRT acceleration.

6.7 Summary of Phase–1 Results

The key accomplishments of this phase are:

- Fully functional pipeline architecture
- End-to-end prototype execution
- Operational clinician and patient reporting engine
- Bilingual summaries (English, Hindi, and local language)
- Integrated explainability hooks
- Backend and job queue operational
- Finalized Findings JSON (Phase–1-complete)
- Functional FHIR exporter
- Validated UI prototype

These outcomes establish a strong foundation for Phase–2, which will focus on quantitative evaluation, clinical model training, optimization, and deployment testing.

Chapter 7

Conclusion

Early detection of lung cancer has the potential to significantly improve treatment success rates and overall patient survival. However, rural and underserved communities continue to face critical limitations such as inadequate access to radiologists, non-standardized reporting practices, and a lack of advanced diagnostic infrastructure. This project aims to address these gaps by developing a lightweight, explainable, and interoperable AI-based lung nodule screening system tailored for deployment in Health ATMs. The goal is to support large-scale, decentralized lung cancer screening in regions where specialist presence is limited.

During Phase–I, the project focused on designing the overall system architecture, establishing the workflow, preparing required datasets, and demonstrating a complete functional pipeline from DICOM ingestion to report generation. The outcomes of this phase validate the feasibility and robustness of the system design and provide a solid foundation for model training and clinical evaluation in Phase–II.

A modular, end-to-end pipeline was successfully implemented, covering preprocessing, nodule candidate structuring, segmentation, malignancy risk scoring, explainability integration, multilingual reporting, and FHIR-based export for national-level interoperability. The prototype backend demonstrated its ability to ingest CT scans, manage asynchronous processing through a job queue, generate structured outputs using a refined Findings JSON schema, and produce clinician-oriented and patient-friendly reports in multiple languages. The integration of explainability hooks—including segmentation masks, Grad-CAM++ placeholders, and uncertainty estimates—ensures that the system remains transparent and clinically trustworthy.

A significant accomplishment of Phase–I was the development of the Findings JSON structure, which now serves as the single authoritative source for all system outputs. This schema supports standardized reporting, FHIR bundle generation, and future AI model training workflows. The bilingual and multilingual reporting engine, built using template-based rendering and the NLLB–200 translation model, successfully produced hallucination-free clinical and patient summaries. Together with the backend APIs and user interface prototype, the system was able to complete a fully functional synthetic end-to-end demonstration.

Overall, the progress achieved in Phase–I confirms that the designed architecture is well-suited for deployment in low-resource environments such as Health ATMs. The system demonstrates strong potential for practical integration into rural healthcare workflows, with careful considerations toward computational efficiency, explainability, accessibility, and interoperability. These Phase–I achievements establish a strong platform for the more advanced development, optimization, and validation tasks planned for Phase–II.

7.1 Future Work

Phase–II (S8) will focus on training, optimizing, and validating the full AI pipeline using benchmark datasets such as LIDC-IDRI, LUNA16, LNDb v4, and CT-RATE. The major planned tasks include:

1. **Model Training and Optimization**
   - Train UNet-family segmentation models with multi-dataset integration.
   - Implement CenterNet-lite for efficient nodule detection.
   - Train the malignancy risk classifier using combined radiomic and deep learning features.
   - Apply INT8 quantization and TensorRT optimization to achieve inference within 60 seconds.
2. **Explainability Evaluation**
   - Generate real Grad-CAM++ heatmaps and saliency overlays.
   - Conduct faithfulness evaluations including IoU, Pointing Game, and deletion/insertion metrics.
   - Integrate uncertainty-based routing for low-confidence predictions.
3. **Clinical Validation and User Studies**
   - Perform a radiologist-led evaluation using approximately 50 CT cases.
   - Assess clarity, correctness, and usefulness of clinician and patient summaries.
   - Refine report templates based on expert and operator feedback.
4. **ABDM/FHIR Sandbox Testing**
   - Validate DiagnosticReport and ImagingStudy bundles in the ABDM sandbox.
   - Ensure error-free ingestion by national electronic health registries.
   - Finalize metadata mapping rules for compliance readiness.

5. **Deployment-Ready UI and Workflow Enhancements**

    - Improve the frontend viewer with overlay toggles, slice navigation, and bilingual UI.
    - Enhance error handling, loading indicators, and offline-first capabilities.
    - Integrate secure local storage and data synchronization mechanisms.

6. **Publication and Documentation**

    - Prepare a research manuscript summarizing the system architecture and Phase–II results.
    - Produce developer documentation and reproducible training and deployment pipelines.

7.2 Final Remarks

The outcomes of Phase–I represent a meaningful and substantial step toward building a scalable, real-world AI-powered lung nodule screening system for rural India. By integrating lightweight inference, structured reporting, multilingual communication, explainable predictions, and national-level interoperability, this system has the potential to transform early lung cancer detection for underserved populations.

Phase–II will enhance model intelligence, performance, and clinical reliability, enabling Health ATMs to offer high-quality diagnostic support where it is most needed. With continued development and refinement, the system is poised to evolve into a robust and impactful solution that contributes to equitable healthcare access across the country.

Bibliography

[1] C. Gao, H. Li, J. Huang, et al., "Deep learning in pulmonary nodule detection and segmentation: a systematic review," *European Radiology*, vol. 35, pp. 1–19, 2024/2025.

[2] W. Hendrix, A. Brattain, J. Zhang, et al., "Deep learning for the detection of benign and malignant pulmonary nodules in non-screening chest CT scans," *Communications Medicine* (Nature Portfolio), vol. 3, pp. 1–12, 2023.

[3] Q. Woznicki, A. Pedrosa, S. Pereira, et al., "LNDb v4: Pulmonary nodule annotation from medical reports with manual segmentation," *Scientific Data* (Nature Portfolio), vol. 11, no. 55, pp. 1–14, 2024.

[4] Q. Wang, S. Shao, M. Li, et al., "Artificial intelligence in lung cancer screening: Detection, classification, prediction, and prognosis," *Cancer Medicine* (Wiley), vol. 13, pp. 1–22, 2024.

Appendix A

Source Code and Dataset Details

This appendix provides an overview of the source code, internal modules, dataset organization, and the structure of the `findings.json` file used within the system. The objective is to document all essential components used in developing the Phase–1 prototype of the AI-based lung nodule screening system.

A.1 Source Code Structure

The project codebase is organized into modular components to ensure maintainability, scalability, and clear separation of responsibilities across the AI runtime, backend, reporting engine, and UI layers.

A.1.1 Directory Structure

```
/project_root

 ml/
    preprocess.py
    detector.py
    segmenter.py
    risk.py
    xai.py
    infer.py

 app/
    main.py                 # FastAPI backend
    queue.py                # Redis/RQ job queue
    validators.py           # JSON schema validator
    reporter.py             # Reporting engine
    fhir_export.py          # DiagnosticReport + ImagingStudy
    qa_engine.py            # JSON-grounded Q&A (future work)
```

```
templates/
   clinician_en.html
   patient_en.html
   patient_hi.html

ui/
   app.py                # Streamlit prototype UI

data/
   proto_pack/           # Small CT samples for Phase-1
   dataset_card.md
   findings.schema.json

artifacts/
    reports/
    masks/
    heatmaps/
    fhir/
```

A.2 Key Code Modules

Each module contributes to a specific functionality within the AI pipeline.

A.2.1 Preprocessing Module (preprocess.py)

Responsible for:

- Loading DICOM studies
- Converting pixel values to Hounsfield Units
- Resampling to 1 mm isotropic spacing
- HU clipping and normalization
- Creating 2.5D / 5-slice input stacks

Output: Preprocessed CT volume ready for the detection and segmentation stages.

A.2.2 Detection Module (detector.py)

Implements lightweight detection for Phase–1:

- LoG-based blob detection
- Threshold-based candidate filtering
- Nodule coordinate extraction

Phase–2 will integrate CenterNet-lite for improved accuracy.

A.2.3 Segmentation Module (segmenter.py)

Contains the Mobile-UNet architecture:

- Depthwise-separable convolutions
- Lightweight encoder–decoder design
- Dice + BCE hybrid loss (used in Phase–2 training)

Output: Binary masks for each detected nodule.

A.2.4 Risk Classification Module (risk.py)

Computes malignancy probability using:

- Deep features from the UNet backbone
- Radiomic features (volume, HU mean, margin)
- MLP classifier for malignancy scoring
- Probability calibration (e.g., temperature scaling)
- Uncertainty estimation (entropy, dropout variance)

A.2.5 Explainability Module (xai.py)

Implements:

- Grad-CAM++ for saliency map generation
- Heatmap creation and overlay preparation
- Uncertainty scoring and flagging logic

Phase–1 includes placeholders; real explainability images will be added in Phase–2.

A.2.6 Inference Orchestration (infer.py)

This module ties the entire AI workflow together:

1. Loads CT
2. Executes preprocessing
3. Applies detector and segmenter
4. Computes malignancy risk
5. Aggregates results into structured format
6. Saves masks and heatmaps
7. Produces final `findings.json`

A.2.7 Reporting Engine (reporter.py)

Generates:

- Clinician PDF report
- Patient PDF (English + Hindi)

Features:

- Jinja2 templating
- WeasyPrint rendering pipeline
- NLLB–200 translation for multilingual support
- Strict grounding in `findings.json`

A.2.8 Backend and Queue System (main.py, queue.py)

Implements the FastAPI server:

- `/ingest` – upload CT
- `/analyze` – create queue job
- `/status` – check job progress
- `/results` – fetch artifacts

A Redis/RQ worker executes the entire AI pipeline asynchronously.

A.2.9 FHIR Exporter (fhir_export.py)

Creates ABDM-compliant:

- `DiagnosticReport.json`
- `ImagingStudy.json`
- Linked media objects

Compliant with:

- HL7 FHIR R4 standards
- Ayushman Bharat Digital Mission (ABDM) interoperability specifications

A.3 Dataset Details

The system uses four primary datasets and one curated proto dataset.

A.3.1 LIDC-IDRI

- 1,018 CT scans
- Multi-radiologist annotations
- Nodule masks and malignancy ratings

A.3.2 LUNA16

- 888 CT scans
- Standard benchmark for nodule detection
- Includes candidate and false-positive classification tracks

A.3.3 LNDb v4

- CT scans aligned with radiology reports
- Manual segmentation masks
- Used for factuality evaluation of reports

A.3.4 CT-RATE

- Large corpus of real clinical chest CT reports
- Used for:
  - Natural language benchmarking
  - Template refinement
  - Patient summary grounding

A.3.5 Proto-Pack Dataset

- 10–20 small CT studies
- Created manually for Phase–1 testing
- Supports quick testing of preprocessing, reporting, and backend workflows

A.4 Findings JSON Structure (Excerpt)

Below is an excerpt from the finalized Phase–1 `findings.json` file, showing key fields used for reporting, explainability, and FHIR conversion:


```
{
  "study_id": "demo",
  "lung_level": {
    "left_lung_volume": 2100,
    "right_lung_volume": 2300,
    "emphysema_score": 0.02
  },
  "nodules": [
    {
      "nodule_id": "N1",
      "location": "Right Upper Lobe",
      "type": "solid",
      "diameter_mm": 8.4,
      "malignancy_prob": 0.72,
      "uncertainty_score": 0.18,
      "needs_review": false,
      "mask_path": "artifacts/masks/N1.png",
      "gradcam_path": "artifacts/heatmaps/N1_cam.png"
```

```
    }
  ],
  "clinical_impression":
    "One intermediate-risk solid nodule in the right upper lobe.",
  "patient_summary_en":
    "A small growth was found in your right lung...",
  "patient_summary_hi":
    "         ..."
}
```

This JSON serves as the **single source of truth** for:

- Clinician and patient reporting
- Explainability visualization
- UI overlays
- FHIR bundle generation
- Future ML evaluation in Phase–2

A.5 Appendix Summary

This appendix documented:

- Complete source code organization
- Responsibilities of each module
- Dataset descriptions and usage
- Structure and purpose of the Findings JSON
- Key integration points connecting AI, backend, UI, and reporting

These details form the implementation backbone required to extend Phase–1 outputs into a fully trained, validated, and deployable AI-assisted lung nodule screening system during Phase–2.