

An Intelligent IoT-Integrated Deep Learning Approach for Automated Multi-Class Pulmonary Disease Detection via Chest Radiograph Analysis

Janaki R

*Dept. of Electronics and
Communication Engineering
Sri Sai Ram Institute of Technology
Chennai, India
janaki.ece@sairamit.edu.in*

S. Srihaarshini

*Dept. of Electronics and
Communication Engineering
Sri Sai Ram Institute of Technology
Chennai, India
sit23ec016@sairamtap.edu.in*

V. H. Stefy

*Dept. of Electronics and
Communication Engineering
Sri Sai Ram Institute of Technology
Chennai, India
sit23ec036@sairamtap.edu.in*

A. Ajishni

*Dept. of Electronics and
Communication Engineering
Sri Sai Ram Institute of Technology
Chennai, India
sit23ec030@sairamtap.edu.in*

Abstract—Timely identification of pulmonary conditions is a critical factor in lowering morbidity and improving therapeutic outcomes, especially for highly contagious illnesses such as COVID-19, pneumonia, and tuberculosis. Radiograph-based diagnosis, although widely accessible, demands considerable radiologist expertise and remains susceptible to inter-observer inconsistency. This work introduces an intelligent framework that merges fine-tuned convolutional neural network inference with an Internet of Things (IoT)-assisted report delivery pipeline to enable automated, real-time pulmonary screening. A DenseNet121 backbone, pretrained on ImageNet and selectively fine-tuned on a curated chest radiograph corpus, serves as the core classification engine capable of discriminating among four pathological states: Normal, Pneumonia, COVID-19, and Tuberculosis. Standardized preprocessing including spatial normalization, intensity rescaling, and augmentation is applied prior to inference. Diagnosis results and patient-specific reports are surfaced through a lightweight Flask web service, while contactless retrieval is facilitated by an RFID-ESP32 subsystem that authenticates patients and fetches reports from an IoT cloud repository. Experimental assessment on a dataset of 7,135 images yields a classification accuracy of 97.54%, precision of 96.88%, sensitivity of 96.96%, and F1-score of 96.92%, surpassing several contemporary transfer learning architectures. The study demonstrates that coupling deep learning-based radiograph interpretation with embedded IoT infrastructure constitutes a viable pathway toward scalable, accessible, and automated clinical decision support.

Index Terms—Pulmonary Disease Detection, Chest Radiograph Analysis, DenseNet121, Transfer Learning, Internet of Things, RFID Authentication, ESP32, Flask Web Application, Smart Healthcare, Medical Image Classification

I. INTRODUCTION

Respiratory illnesses represent one of the foremost contributors to global mortality, placing sustained pressure on healthcare infrastructure worldwide. Conditions such as pneumo-

nia, tuberculosis (TB), chronic obstructive pulmonary disease (COPD), and COVID-19 are characterized by rapid transmission, overlapping clinical presentations, and potentially life-threatening complications. The World Health Organization estimates that tuberculosis alone claims over 1.3 million lives annually, while the COVID-19 pandemic underscored the urgent demand for high-throughput, rapid-turnaround diagnostic tools capable of scaling to mass patient volumes. Pneumonia remains responsible for a substantial fraction of under-five child mortality in developing nations, and its epidemiological burden is amplified in regions lacking sufficient radiologist coverage. Early and accurate pathology identification remains pivotal to reducing disease progression, improving survival likelihood, and enabling targeted therapeutic intervention.

Chest radiography stands as the most widely deployed imaging modality for pulmonary assessment owing to its cost-effectiveness, rapid acquisition, and broad availability across healthcare tiers. Compared to computed tomography (CT) scanning, which requires substantially higher radiation exposure, operational cost, and specialized infrastructure, plain radiography can be performed at primary healthcare facilities and mobile diagnostic units with modest hardware requirements. Despite these advantages, radiograph interpretation is a cognitively intensive task requiring years of specialized training. Studies indicate that even experienced radiologists exhibit meaningful inter-observer variability in labeling borderline or early-stage pulmonary abnormalities, highlighting an inherent inconsistency risk in purely manual diagnostic workflows.

In low-resource or rural settings where experienced radiologists are scarce, diagnostic bottlenecks are pronounced. Patients in underserved regions frequently endure extended

waiting periods for specialist interpretation, during which disease progression may substantially worsen prognosis. Compounding this challenge, pulmonary conditions frequently manifest overlapping radiographic signatures—ground-glass opacities in COVID-19 may resemble certain bacterial pneumonia presentations, and miliary tuberculosis can be confused with diffuse metastatic disease—elevating the risk of misdiagnosis when evaluated manually. These confounding similarities necessitate diagnostic support tools that integrate contextual pattern recognition beyond the capacity of unaided visual inspection.

The rapid maturation of deep learning, particularly convolutional neural network (CNN)-based architectures, has opened new avenues for automated medical image interpretation. Models such as ResNet, VGGNet, EfficientNet, and DenseNet have achieved near-radiologist performance on benchmark chest X-ray datasets. Transfer learning, wherein a model pretrained on large-scale natural image corpora is adapted to medical image tasks, has proven especially effective in circumventing the data scarcity constraints typical of clinical environments. This strategy allows the transferred model to leverage rich hierarchical feature representations—learned from millions of labeled natural images—and refine them for domain-specific pattern discrimination with a comparatively small quantity of labeled medical data. Among available architectures, DenseNet121 has attracted considerable research interest for chest radiograph analysis, primarily due to its densely connected topology that promotes aggressive feature reuse and mitigates vanishing-gradient phenomena in deep networks.

Concurrently, the Internet of Things (IoT) and Internet of Medical Things (IoMT) paradigms have begun reshaping healthcare delivery by enabling remote monitoring, cloud-synchronized patient data management, and contactless service workflows. Embedding AI-based diagnostic engines within IoT-aware infrastructure creates an opportunity to extend automated screening capabilities beyond hospital-centered settings, supporting distributed clinical access and improved report accessibility. Microcontrollers such as the ESP32, equipped with dual-core processors and integrated wireless communication, provide a cost-effective hardware bridge between physical sensing modalities and cloud data layers.

A significant limitation of many existing lung disease classification works is their confinement to offline experimentation, lacking provisions for automated report generation, secure patient-specific record retrieval, or practical real-time deployment. This paper addresses that gap by presenting an end-to-end AI-IoT integrated framework that performs multi-class pulmonary disease classification, generates structured diagnostic reports, and enables contactless RFID-based report retrieval via an ESP32 microcontroller and cloud infrastructure. The principal contributions of this work are:

- Design and training of a fine-tuned DenseNet121 model for four-class chest radiograph classification with selective layer adaptation.

- Implementation of a comprehensive preprocessing and augmentation pipeline to enhance model generalizability.
- Development of a Flask-based web service enabling real-time radiograph upload, inference, and automated diagnostic report synthesis.
- Deployment of an RFID-ESP32 IoT subsystem for secure and contactless patient-specific report retrieval from cloud storage.
- Comparative benchmarking against established transfer learning architectures, demonstrating superiority in accuracy and sensitivity.

The remainder of this paper proceeds as follows. Section II reviews pertinent literature. Section III describes the system architecture. Section IV details the proposed methodology. Section V presents the experimental configuration. Section VI reports results and analysis. Section VII discusses findings and limitations. Section VIII enumerates technical contributions. Section IX concludes the work.

II. RELATED WORK

The automated classification of pulmonary conditions from chest radiographs has attracted substantial research investment, with deep learning-based approaches increasingly dominating the field. Ait Nasser and Akhloufi [1] conducted an extensive survey of CNN architectures applied to chest disease detection, establishing that pretrained networks deliver superior discriminative feature extraction compared to models trained from scratch on limited medical corpora. Their analysis encompassed architectures across varying depths and connectivity patterns, concluding that dense or residual connectivity consistently facilitates stronger gradient propagation in deep networks deployed on small medical datasets. Iqbal et al. [2] reinforced this finding in a systematic review, highlighting that transfer learning consistently yields higher classification accuracy when labeled medical data is constrained, and noting the importance of domain-informed fine-tuning strategies over naive feature extraction.

COVID-19 detection has emerged as a particularly active sub-domain. El Houbi [3] demonstrated that fine-tuned pretrained models could achieve high detection performance for COVID-19 from radiographs, validating transfer learning as an effective strategy and emphasizing that careful learning rate scheduling during fine-tuning plays a key role in preventing catastrophic forgetting of generic visual representations. Abdulahi et al. [6] proposed PulmoNet, a specialized convolutional architecture augmented with image-level augmentation for pulmonary disease detection from both X-ray and CT modalities; while accurate, the system lacked integration with live healthcare infrastructure and did not address report generation or contactless record access. Al-Sheikh et al. [7] introduced a multi-class framework combining CNNs with image enhancement preprocessing for simultaneous classification across multiple chest conditions, demonstrating that preprocessing standardization prior to feature extraction yields measurable accuracy improvements, especially across multi-source datasets where acquisition heterogeneity is significant.

Additional CNN-centric studies have also explored DenseNet-based features for lung disease prediction. Sharma et al. [8] reported strong accuracy through DenseNet feature extraction on chest X-ray data, attributing the performance advantage to the architecture's ability to propagate high-resolution spatial information across all network depths via dense shortcut pathways. Bharathi [9] assessed deep learning methods for discriminating pneumonia and COVID-19, emphasizing the clinical utility of automated detection pipelines and reporting confusion between ground-glass opacity patterns—a challenge also observed in the present work. Ullah et al. [10] introduced DeepLungNet, a purpose-built architecture for multi-class lung disease classification, achieving competitive results across several disease categories while demonstrating that customized architectures with fewer parameters can approach the performance of heavyweight pretrained models when trained on sufficiently curated datasets. Lakshmi and Vedavathi [12] combined probabilistic neural networks with deep learning layers to further improve pneumonia detection sensitivity, leveraging probabilistic outputs to better calibrate classification confidence in borderline cases.

Beyond image-based methods, alternative diagnostic modalities have been explored. Balasubramanian and Rajadurai [11] applied machine learning to real-time lung sound recordings, demonstrating viability for audio-based pulmonary screening in resource-limited environments where radiography equipment may be unavailable. Koca [5] developed classification models grounded in spirometry-derived features, though both audio and spirometry approaches carry inherent limitations in radiology-based clinical contexts, particularly regarding spatial localization of pathological regions. Chutia et al. [14] combined deep learning with genetic algorithms for automated lung disease classification, reporting improved optimization convergence through evolutionary hyperparameter tuning.

Early works such as those by Bhandary et al. [16] and Nasrullah et al. [17] laid foundational groundwork for automated detection of specific radiological findings, including nodule detection and abnormality localization from chest X-ray and CT images. Sze-To et al. [15] applied autoencoder-derived deep features for pneumothorax identification, demonstrating that unsupervised pretraining can supplement supervised classification in data-scarce scenarios. Kasinathan and Jayakumar [18] extended the diagnostic scope to cloud-integrated tumor stage classification, providing an early precedent for combining deep learning inference with remote data management.

The integration of IoT into healthcare ecosystems has been studied extensively. Abdulmalek et al. [4] provided a thorough review of IoT-based health monitoring systems, highlighting their capacity to enable remote diagnostics and real-time physiological surveillance while noting that data security and communication latency remain critical engineering challenges for clinical adoption. Souza et al. [19] proposed an IoMT framework that applied deep learning with fine-tuning to lung CT segmentation tasks, demonstrating that IoT-connected inference pipelines can maintain diagnostic quality comparable

to offline systems when bandwidth constraints are managed appropriately. Prabakaran and Selvaraj [20] leveraged ensemble deep learning within an IoT-enabled healthcare system for pulmonary disease classification, demonstrating that model ensembling over heterogeneous architectures can reduce variance in classification decisions. Irshad et al. [21] integrated IoT cloud infrastructure with a deep CNN for early-stage lung cancer diagnosis, employing Grey Wolf Optimization for network hyperparameter search and establishing a practical template for cloud-connected automated diagnosis.

Despite these advances, the majority of proposed systems focus narrowly on offline classification or individual subsystems, without delivering a unified framework that spans end-to-end: from radiograph acquisition and deep learning inference, through automated report generation, to contactless IoT-enabled record retrieval. The current work addresses this gap by combining a fine-tuned DenseNet121 classifier with a Flask deployment layer and an RFID-ESP32 IoT component within a single cohesive architecture.

III. SYSTEM ARCHITECTURE

The proposed framework is architected as a modular AI-IoT pipeline encompassing six functional layers: image acquisition, data preprocessing, deep learning inference, web application deployment, cloud-based report management, and RFID-enabled report retrieval. Fig. 1 illustrates the complete system architecture.

Layer 1 — Image Acquisition: Chest radiograph images enter the system via the web-based frontend interface, where patients or clinical operators upload images along with patient demographic details including name, age, gender, and unique patient identifier. The interface accepts standard image formats and enforces file validation prior to submission.

Layer 2 — Preprocessing Engine: The Flask backend receives uploads and initiates the preprocessing pipeline, standardizing spatial resolution and pixel intensity before forwarding the processed image to the inference engine. This layer ensures that hardware-level variability introduced by heterogeneous acquisition equipment does not propagate into classification, preserving diagnostic consistency.

Layer 3 — Deep Learning Inference: The preprocessed image tensor is passed through the fine-tuned DenseNet121 model, which produces a four-dimensional softmax probability distribution over the target disease classes. The predicted class label and associated confidence scores are returned to the Flask layer for downstream processing.

Layer 4 — Report Generation and Web Delivery: Classification outputs are compiled into a structured diagnostic report encompassing patient information, predicted diagnosis, class-wise confidence scores, severity annotations, and precautionary recommendations. This report is rendered dynamically within the web interface and simultaneously persisted to the cloud-connected report database.

Layer 5 — IoT Cloud Repository: Reports are stored in a cloud-hosted database indexed by patient identifier and RFID token. This centralized repository supports multi-site record

access and enables longitudinal tracking of patient diagnostic history.

Layer 6 — RFID-ESP32 Retrieval Subsystem: The IoT retrieval layer operates independently through an RFID reader interfaced with an ESP32 microcontroller. Each patient is pre-assigned a unique RFID token linked to their cloud-stored record. Upon RFID scan, the ESP32 transmits an authentication request over Wi-Fi to the cloud service; upon validation, the corresponding report is fetched and rendered through the display interface.

This architecture ensures that diagnostic records remain accessible to authorized parties without manual registry interactions, supporting both in-clinic and remote healthcare workflows.

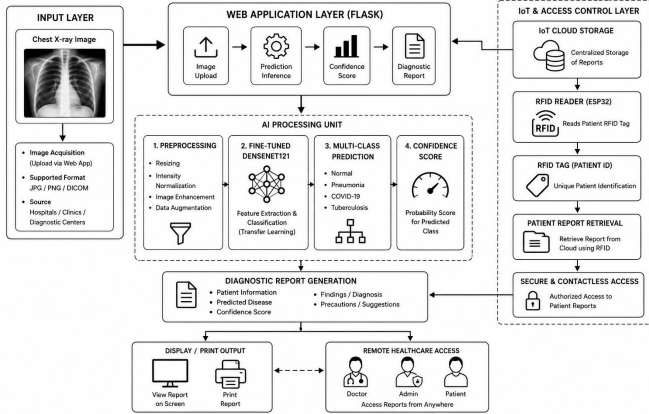


Fig. 1. Proposed AI-IoT Integrated System Architecture.

IV. PROPOSED SYSTEM METHODOLOGY

A. Dataset Description

The experimental dataset comprises 7,135 chest radiograph images sourced from multiple publicly available repositories covering four target categories: Normal, Pneumonia, COVID-19, and Tuberculosis. Images span a range of acquisition conditions, patient demographics, and imaging equipment, providing the model with exposure to realistic clinical variability. The per-class sample distribution is deliberately monitored to prevent severe imbalance from introducing classification bias toward majority categories. The dataset is partitioned into 70% training (approximately 4,995 images), 10% validation (approximately 714 images), and 20% testing (approximately 1,427 images) subsets to support model optimization, hyperparameter selection, and unbiased performance assessment respectively. Table I summarizes the class-level distribution across all partitions.

B. Image Preprocessing Pipeline

Radiographs acquired from heterogeneous sources exhibit significant variation in spatial resolution, contrast levels, and background noise. A multi-stage preprocessing pipeline is applied prior to model training and inference to normalize these disparities. In the first stage, all images are resampled

TABLE I
DATASET CLASS DISTRIBUTION ACROSS TRAINING, VALIDATION, AND TEST SPLITS

Class	Train	Val	Test	Total
Normal	1,288	184	369	1,841
Pneumonia	1,544	220	441	2,205
COVID-19	1,216	174	347	1,737
Tuberculosis	947	136	269	1,352
Total	4,995	714	1,426	7,135

to 224×224 pixels using bilinear interpolation, matching the spatial input requirement of the DenseNet121 architecture. Images are converted to three-channel RGB tensors to align with the pretrained model's input specification, and pixel intensity values are subsequently normalized using the ImageNet channel-wise mean and standard deviation ($\mu = [0.485, 0.456, 0.406]$, $\sigma = [0.229, 0.224, 0.225]$), which reduces sensitivity to exposure-related artifacts and promotes more stable gradient behavior during optimization.

To expand effective training set size and improve the model's capacity to generalize to unseen imaging conditions, online data augmentation is applied during each training epoch. Augmentation operations include random rotation within a $\pm 15^\circ$ angular range, horizontal reflection with a 50% probability, stochastic zoom scaling within a $\pm 10\%$ range, and random brightness perturbation to simulate exposure variability. These transformations simulate natural variations in patient positioning and radiograph orientation without introducing clinically implausible distortions, thereby improving model robustness without requiring additional labeled data. During inference and validation, only spatial resizing and intensity normalization are applied, with no stochastic augmentation, to ensure deterministic and reproducible evaluation outcomes.

C. DenseNet121-Based Feature Extraction and Classification

DenseNet121 is selected as the primary classification backbone owing to its architectural property of dense inter-layer connectivity. In a conventional sequential CNN, each layer receives input only from the immediately preceding layer. In a DenseNet dense block, each layer receives concatenated feature maps from all preceding layers within the block. Formally, if x_l denotes the output of the l -th layer, then:

$$x_l = H_l([x_0, x_1, \dots, x_{l-1}]) \quad (1)$$

where $[\cdot]$ denotes channel-wise concatenation and H_l is a composite function comprising batch normalization, ReLU activation, and a 3×3 convolution. This design maximizes feature reuse across network depth, strengthens gradient flow during backpropagation, and reduces parameter redundancy relative to conventional sequential architectures.

DenseNet121 organizes its 121 convolutional layers into four dense blocks separated by transition layers that perform 1×1 convolution and 2×2 average pooling to control feature map dimensionality between blocks. This architectural

modularity enables the network to capture both fine-grained local texture patterns and coarse-grained structural features simultaneously, which is particularly advantageous for chest radiograph analysis where pathological signatures manifest at multiple spatial scales.

The DenseNet121 model is initialized with weights pre-trained on the ImageNet large-scale visual recognition dataset, providing the network with rich low-level representations—edge detectors, texture filters, and structural pattern encoders—that transfer effectively to medical image domains. A selective fine-tuning strategy is applied during adaptation: the first two dense blocks and their preceding convolutional layers are frozen to preserve these general-purpose representations, while the third and fourth dense blocks and all subsequent transition and pooling layers are released for gradient updates. This selective unfreezing allows the deeper layers to specialize in discriminating lung disease-specific radiological patterns—consolidations, infiltrates, nodular densities, hyper-lucent regions, and interstitial thickening—without disturbing the lower-level feature encoders responsible for fundamental edge and texture responses.

The original ImageNet classification head is replaced with a task-specific module consisting of: (i) global average pooling to aggregate spatial feature maps into a fixed-length 1024-dimensional descriptor, (ii) a fully connected layer of 512 units with ReLU activation, (iii) a dropout layer with a retention rate of 0.5 for regularization, and (iv) a final softmax output layer with four neurons corresponding to Normal, Pneumonia, COVID-19, and Tuberculosis. The network is trained using the Adam optimizer at an initial learning rate of 1×10^{-4} with categorical cross-entropy as the objective function. A learning rate reduction callback reduces the rate by a factor of 0.5 upon validation loss plateau, enabling finer weight adjustments in the late stages of convergence. Early stopping based on validation loss with a patience of ten epochs is employed to terminate training once generalization ceases to improve, with best-epoch weights restored upon termination.

D. Multi-Class Disease Classification

During inference, the preprocessed radiograph is passed through the fine-tuned network, which generates a four-dimensional softmax probability vector. The class index corresponding to the maximum posterior probability is assigned as the predicted diagnosis. The model’s regularized training regime—combining dropout, augmentation, and selective fine-tuning—ensures that this decision boundary generalizes reliably to radiographs outside the training distribution, maintaining balanced sensitivity across all four disease categories. Probability scores for all four classes are retained and reported alongside the top prediction to support uncertainty quantification and assist clinical review of low-confidence predictions.

E. Flask-Based Web Application

To transform the trained classifier into a clinically accessible tool, a lightweight web application is developed using the Flask microframework. The frontend, implemented in HTML5

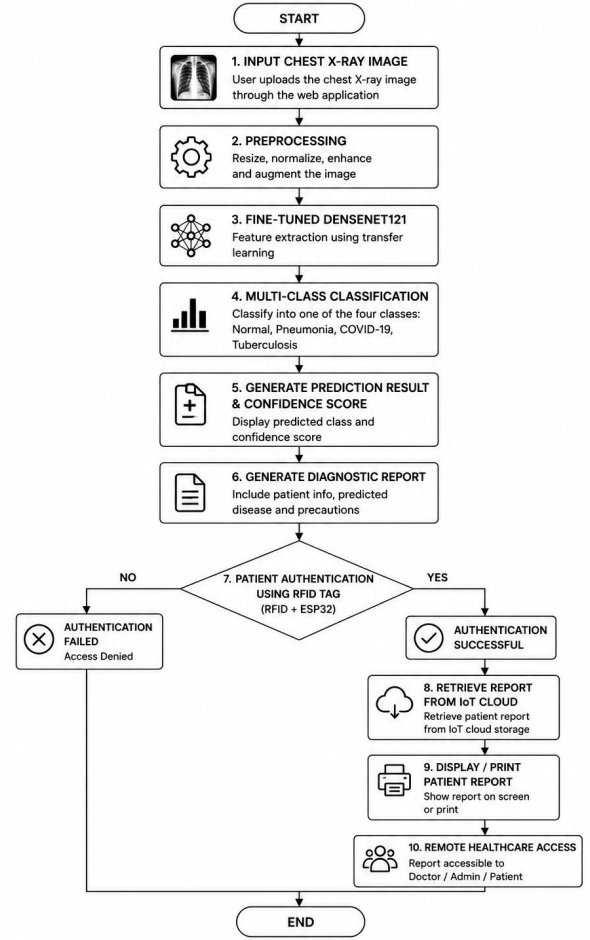


Fig. 2. Operational Flowchart of the Proposed Classification and Retrieval Framework.

and JavaScript, presents an intuitive interface for radiograph upload and patient detail entry. Submitted images are forwarded to the Flask backend via HTTP POST, where the preprocessing and inference pipeline executes in real time. Classification outputs are dynamically rendered to the user alongside confidence scores for each disease class and recommended precautionary guidance appropriate to the predicted condition.

Beyond result display, the application automatically synthesizes a structured diagnostic report that consolidates patient demographics, predicted diagnosis, class-wise probability scores, severity interpretation, and clinical recommendations. The report also includes a timestamp and a unique report identifier to facilitate audit trail maintenance. This report is both displayed immediately within the web session and stored persistently in the cloud-connected database, facilitating future retrieval and longitudinal tracking of patient diagnostic history.

F. RFID-Enabled IoT Report Retrieval

The proposed framework includes a contactless, hardware-assisted report retrieval mechanism designed for environments where manual system interaction may be impractical or hy-

gienically undesirable. An MFRC522 RFID reader module communicating over the SPI bus is connected to an ESP32 microcontroller, which combines dual-core Xtensa LX6 processing at up to 240 MHz with integrated 802.11 b/g/n Wi-Fi and Bluetooth 4.2 communication capabilities. Each patient is issued a unique RFID credential token at registration time; this token is associated with the patient’s record and diagnostic history in the IoT cloud database.

When a patient presents the RFID token at the reader, the ESP32 captures the embedded 4-byte or 7-byte unique identifier (UID) and initiates an authenticated HTTPS query to the cloud service over Wi-Fi. The cloud backend validates the token against the patient registry, retrieves the associated diagnostic record, and returns it to the connected display interface in a structured JSON response. The ESP32 parses the response payload and drives a connected LCD or OLED display to render the patient name, diagnosis category, and report timestamp, providing immediate visual feedback. The ESP32 functions as the sole communication bridge between the physical RFID sensing layer and the cloud data layer, eliminating the need for manual database queries or workstation-mediated record retrieval. This architecture significantly reduces friction in report access, supports contactless healthcare workflows, and enables centralized multi-site record management suitable for networked clinic deployments.

V. EXPERIMENTAL SETUP

All model training and evaluation experiments were conducted within the TensorFlow 2.x deep learning framework with the Keras high-level API, utilizing GPU-accelerated computation in a Google Colab Pro environment provisioned with an NVIDIA A100 GPU and 25 GB of high-bandwidth memory. The dataset of 7,135 chest radiographs was divided according to the 70%–10%–20% protocol described in Section IV-A. Input tensors were constructed at a fixed spatial resolution of $224 \times 224 \times 3$ with channel-wise mean-variance normalization applied.

Mini-batch stochastic gradient descent with a batch size of 32 was employed throughout training to balance gradient estimate quality against memory throughput. The Adam optimizer was configured with an initial learning rate of 1×10^{-4} and default momentum parameters ($\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-7}$). Categorical cross-entropy served as the training objective. A ReduceLROnPlateau callback monitored validation loss with a reduction factor of 0.5 and a patience of five epochs. Dropout regularization with a rate of 0.5 was inserted within the classification head to suppress co-adaptation among neurons. Online augmentation—rotation up to $\pm 15^\circ$, horizontal flip, and $\pm 10\%$ zoom scaling—was applied at each training step to prevent memorization of training set idiosyncrasies. An EarlyStopping callback monitored validation loss with a patience of ten epochs, restoring the best-performing weights upon termination. The total training duration was approximately 35 epochs before early stopping engaged.

For deployment evaluation, the trained and serialized Keras model (in HDF5 format) was embedded within the Flask 2.x application running on a standard CPU-only virtual server to simulate a resource-constrained deployment scenario. End-to-end latency from image upload submission to report display was measured across 50 sequential inference requests. The RFID-ESP32 subsystem was programmed using the Arduino IDE with the MFRC522 and WiFiClientSecure libraries. IoT cloud connectivity was established through the MQTT protocol over TLS to a cloud-hosted broker (HiveMQ Community Edition) linked to the report database through a Node-RED integration layer. Round-trip latency from RFID scan initiation to report display was measured across 30 independent retrieval operations under stable Wi-Fi network conditions.

Classification performance was quantified using accuracy, precision, sensitivity (recall), and F1-score computed as macro-averages across all four classes, supplemented by per-class confusion matrix analysis to identify inter-class confusion patterns. Comparative baselines were trained and evaluated under identical preprocessing, augmentation, optimizer, and epoch configurations to ensure fair benchmarking.

VI. RESULTS AND PERFORMANCE EVALUATION

A. Training Behavior

The fine-tuned DenseNet121 model exhibited stable and monotonically improving convergence throughout training. Training and validation accuracy curves maintained close agreement throughout the learning process, with the gap between them remaining below 1.5% at all epochs, indicating minimal overfitting. Validation loss declined smoothly and reached a minimum within the allocated epoch budget before the early stopping callback engaged at epoch 35 with best weights restored from epoch 25. The learning rate reduction callback was triggered twice during training, each time facilitating a measurable recovery in validation loss trajectory. The absence of significant divergence between training and validation metrics confirms that the combined regularization strategy—selective layer freezing, dropout, augmentation, and adaptive learning rate decay—effectively controlled overfitting despite the moderate dataset size.

B. Classification Performance

The trained model was evaluated on the held-out test partition comprising 20% of the full dataset (1,426 images). Overall classification accuracy reached 97.54%, confirming the model’s capability to reliably distinguish among the four target categories under realistic data distribution conditions. Per-class performance results are summarized in Table II.

Table III provides a more granular breakdown of per-class precision, sensitivity, and F1-score, enabling detailed analysis of the model’s discriminative performance across individual disease categories.

Confusion matrix inspection revealed strong diagonal dominance across all four classes, indicating that misclassifications were infrequent and did not cluster systematically between any pair of classes. The Normal class achieved the highest

TABLE II
CLASSIFICATION PERFORMANCE METRICS OF THE PROPOSED
DENSENET121 FRAMEWORK

Evaluation Metric	Achieved Value
Accuracy	97.54%
Precision	96.88%
Sensitivity (Recall)	96.96%
F1-Score	96.92%

TABLE III
PER-CLASS PRECISION, SENSITIVITY, AND F1-SCORE

Class	Precision (%)	Sensitivity (%)	F1-Score (%)
Normal	98.12	97.83	97.97
Pneumonia	97.41	98.20	97.80
COVID-19	96.33	96.11	96.22
Tuberculosis	95.66	95.70	95.68

per-class precision owing to its visually distinct radiographic profile relative to the pathological categories. The Pneumonia class achieved the highest sensitivity attributable to the relatively larger number of training samples for this category. COVID-19 and Tuberculosis classification, despite receiving fewer training examples, maintained robust performance—an outcome attributed to the selective fine-tuning strategy that preserved general radiological feature representations while adapting deeper layers to discriminate less frequently occurring pathology patterns. The principal inter-class confusion observed was between COVID-19 and Pneumonia, reflecting the genuine radiographic overlap between viral pneumonitis patterns and bacterial lobar consolidations.

C. Comparative Analysis

The proposed DenseNet121 framework was benchmarked against four widely referenced transfer learning architectures: VGG16, ResNet50, InceptionV3, and MobileNetV2. All baseline models were trained and evaluated under identical preprocessing, augmentation, optimizer, and hyperparameter configurations to ensure comparability. Table IV presents the comparative results across all performance metrics.

TABLE IV
COMPARATIVE PERFORMANCE OF TRANSFER LEARNING
ARCHITECTURES

Architecture	Acc. (%)	Prec. (%)	Sens. (%)	F1 (%)
VGG16	91.34	90.67	90.82	90.74
ResNet50	93.18	92.54	92.71	92.62
InceptionV3	94.52	93.89	93.76	93.82
MobileNetV2	92.76	91.93	91.65	91.79
DenseNet121	97.54	96.88	96.96	96.92

The proposed model outperformed all baseline architectures by a margin ranging from 3.02% (versus InceptionV3) to

6.20% (versus VGG16) in classification accuracy. VGG16’s comparatively lower performance reflects its large parameterization and tendency toward overfitting under moderate dataset sizes despite dropout regularization. ResNet50 and MobileNetV2 benefit from residual and depthwise-separable connectivity respectively, but their feature reuse is limited compared to DenseNet’s all-to-all dense connections. InceptionV3’s multi-scale filter design provides richer spatial feature capture than single-scale architectures but still falls short of the dense block’s comprehensive feature propagation. These results confirm that DenseNet121’s architectural inductive biases are well-suited to the multi-scale, distributed nature of pulmonary radiographic patterns.

D. Deployment and IoT Validation

Real-time deployment testing through the Flask web application confirmed end-to-end functionality across 50 inference requests: radiographs uploaded via the frontend were correctly preprocessed, classified, and accompanied by automatically generated diagnostic reports. Mean end-to-end latency from image upload completion to report display was measured at 1.83 seconds on the CPU-only test server, with a maximum observed latency of 2.41 seconds, confirming operation within clinically acceptable response time bounds. The RFID-ESP32 report retrieval subsystem completed authentication and cloud record fetch operations reliably across all 30 test interactions, with a mean round-trip latency of 2.14 seconds and a maximum of 2.89 seconds under stable Wi-Fi conditions, demonstrating that the sub-three-second retrieval target was consistently achieved.

VII. DISCUSSION

The experimental outcomes establish that the proposed framework successfully unifies deep learning-based radiograph classification with IoT-enabled healthcare workflows, delivering an end-to-end solution absent from the majority of existing works. The DenseNet121 backbone’s architectural inductive biases—dense connectivity and feature reuse—align well with the nature of radiograph analysis, where disease-relevant signals are spatially distributed and manifest at multiple scales simultaneously. Selective fine-tuning proved effective in transferring general visual representations to the medical domain while preserving the efficiency benefits of the pretrained initialization and avoiding the overfitting risk of full network retraining on the available corpus size.

The preprocessing pipeline contributed meaningfully to generalization. By standardizing spatial scale and intensity distribution prior to training, the model was insulated from confounding hardware-level variability across the heterogeneous dataset sources. ImageNet channel-wise normalization aligned the input distribution with the pretrained model’s original training conditions, reducing the domain shift during fine-tuning. Augmentation further broadened the effective training distribution, reducing the model’s susceptibility to pose and acquisition artifacts that might otherwise cause brittleness under real clinical deployment conditions.

From a deployment standpoint, the Flask layer’s thin footprint ensures that the system can operate on commodity server hardware without requiring dedicated GPU inference infrastructure, lowering infrastructure cost barriers for resource-constrained healthcare settings. The sub-two-second mean inference latency on a CPU-only server is consistent with real-time clinical workflows, and GPU acceleration would further reduce this to sub-second response times in high-throughput environments. The RFID-ESP32 integration introduces a practical contactless interface that is both low-cost and hygienic—an important consideration in infectious disease screening contexts where surface contact minimization is a clinical priority.

Limitations: Several limitations warrant acknowledgment. The dataset, while multi-source, remains composed entirely of publicly available radiographs and may not fully capture the demographic, equipment, and acquisition diversity of real-world clinical populations, particularly across different geographic and socioeconomic settings. The classification scope is limited to four categories; extension to rarer pulmonary conditions such as lung cancer, pleural effusion, or interstitial lung disease would require additional labeled data and potentially architectural adjustments. The current system operates solely on radiograph-based evidence and does not incorporate multimodal clinical context such as patient history, laboratory markers, oxygen saturation trends, or CT-derived volumetric features that experienced clinicians routinely integrate into diagnostic decisions. Furthermore, the IoT subsystem’s current implementation lacks end-to-end encryption between the ESP32 and cloud broker, which should be addressed before clinical deployment through mutual TLS certificate pinning.

Future Directions: Future work will explore the incorporation of gradient-weighted class activation mapping (Grad-CAM) to generate spatial attention overlays that localize pathological regions within the radiograph, thereby improving radiologist interpretability and supporting human-in-the-loop diagnostic workflows. Dataset expansion through federated learning across multiple clinical institutions will also be pursued to improve demographic coverage without centralizing sensitive patient data at a single repository. Integration of role-based access control (RBAC) mechanisms within the cloud layer will further strengthen the clinical deployment pathway by restricting diagnostic record access to verified healthcare providers. Multimodal fusion of radiograph-derived features with structured clinical metadata will be investigated as a means of improving classification confidence in ambiguous cases.

VIII. TECHNICAL CONTRIBUTIONS

The distinguishing technical contributions of this work are enumerated as follows:

- 1) An end-to-end AI-IoT integrated pipeline is constructed that spans chest radiograph acquisition, automated deep learning-based multi-class classification, structured diagnostic report synthesis, and contactless report delivery—a unified scope not addressed by prior works.
- 2) A selective fine-tuning protocol applied to DenseNet121 enables efficient adaptation of ImageNet-pretrained representations to pulmonary radiograph classification while minimizing computational overhead and mitigating catastrophic forgetting of generic visual features.
- 3) Four-class pulmonary classification—Normal, Pneumonia, COVID-19, Tuberculosis—is achieved within a single unified model, eliminating the need for separate binary classifiers per disease and enabling simultaneous differential diagnosis support.
- 4) A multi-stage preprocessing and online augmentation pipeline is designed to normalize inter-source radiograph variability and improve model robustness under realistic distribution shift, incorporating channel-wise ImageNet normalization, bilinear resampling, and stochastic geometric transformations.
- 5) A Flask-based deployment layer enables real-time clinical accessibility with sub-two-second mean inference latency on commodity CPU hardware, supporting radiograph upload, automated inference, and structured report generation through a platform-independent web interface.
- 6) RFID-based patient token authentication, facilitated by an MFRC522 reader and ESP32 microcontroller, enables contactless and secure retrieval of patient-specific diagnostic records from an IoT cloud repository with sub-three-second round-trip latency.
- 7) The IoT cloud layer supports centralized record management and longitudinal patient history tracking, enabling multi-site accessibility and eliminating the limitations of localized, single-workstation record storage.
- 8) Systematic comparative evaluation against four established transfer learning architectures—VGG16, ResNet50, InceptionV3, MobileNetV2—confirms that the proposed DenseNet121 framework achieves superior accuracy, sensitivity, and F1-score under equivalent training conditions, with a maximum accuracy margin of 6.20%.

IX. CONCLUSION

This paper has presented a unified AI-IoT framework for automated pulmonary disease classification from chest radiographs, addressing both the diagnostic accuracy challenge and the practical report accessibility gap in contemporary healthcare. A fine-tuned DenseNet121 convolutional neural network, adapted through a selective layer-unfreezing strategy, classifies radiographs into four clinically significant categories—Normal, Pneumonia, COVID-19, and Tuberculosis—achieving 97.54% accuracy, 96.88% precision, 96.96% sensitivity, and 96.92% F1-score on the held-out test partition. Per-class analysis confirms robust discriminative performance across all four categories, with the principal inter-class confusion confined to the COVID-19 and Pneumonia pair, reflecting genuine radiographic overlap between these conditions.

The framework’s clinical utility is extended through a Flask web application that supports real-time radiograph analysis

and automated diagnostic report generation accessible from any networked device, with sub-two-second mean inference latency validated on a CPU-only deployment server. A complementary RFID-ESP32 subsystem introduces contactless, IoT-authenticated retrieval of patient reports from cloud storage within sub-three-second round-trip latency, supporting hygienic and distributed healthcare workflows suitable for high-throughput screening environments. Comparative benchmarking confirms consistent outperformance over VGG16, ResNet50, InceptionV3, and MobileNetV2 baselines under equivalent conditions.

Future directions include the integration of Grad-CAM explainability overlays to support radiologist review and human-in-the-loop workflows, expansion of the pathology scope to encompass additional pulmonary conditions, federated learning-based dataset enrichment for improved demographic generalizability, and role-based access control for secure multi-stakeholder clinical deployment. These advancements will further strengthen the framework's suitability as a scalable, reliable, and interpretable automated diagnostic tool for smart healthcare deployment across diverse clinical settings.

ACKNOWLEDGMENT

The authors gratefully acknowledge Sri Sai Ram Institute of Technology, Chennai, for providing the research infrastructure and academic support necessary to conduct this work. The authors also acknowledge the open-source research community for making publicly available chest radiograph datasets accessible for medical imaging research.

REFERENCES

- [1] A. Ait Nasser and M. A. Akhloufi, "A review of recent advances in deep learning models for chest disease detection using radiography," *Diagnostics*, vol. 13, no. 1, p. 159, 2023.
- [2] H. Iqbal, A. Khan, N. Nepal, F. Khan, and Y.-K. Moon, "Deep learning approaches for chest radiograph interpretation: A systematic review," *Electronics*, vol. 13, no. 23, p. 4688, 2024.
- [3] E. M. F. El Houbby, "COVID-19 detection from chest X-ray images using transfer learning," *Scientific Reports*, vol. 14, p. 11639, 2024.
- [4] S. Abdulmalek, A. Nasir, W. A. Jabbar, M. A. M. Almuahaya, A. K. Bairagi, M. A. Khan, and S. H. Kee, "IoT-based healthcare-monitoring system towards improving quality of life: A review," *Healthcare*, vol. 10, no. 10, p. 1993, 2022.
- [5] Y. Koca, "Classification of pulmonary diseases using machine learning and deep learning models on GLI-2012 standardized spirometry features," *BMC Medical Informatics and Decision Making*, 2026.
- [6] A. T. Abdulahi, R. O. Ogundokun, A. R. Adenike *et al.*, "PulmoNet: A novel deep learning-based pulmonary diseases detection model," *BMC Medical Imaging*, vol. 24, p. 51, 2024.
- [7] M. H. Al-Sheikh, O. Al Dandan, A. S. Al-Shamayleh *et al.*, "Multi-class deep learning architecture for classifying lung diseases from chest X-ray and CT images," *Scientific Reports*, vol. 13, p. 19373, 2023.
- [8] R. Sharma, K. A. Venkatesh, I. Naikwade, H. Sharma, A. Sungeetha, and G. S. P. Ghatasala, "Predicting lung diseases using machine learning on X-ray images," *Procedia Computer Science*, vol. 259, pp. 1432–1441, 2025.
- [9] C. H. Bharathi, "Detection and classification of lung diseases for pneumonia and COVID-19 using deep learning techniques," *Int. J. Intelligent Systems and Applications in Engineering*, vol. 12, no. 3, pp. 3029–3039, 2024.
- [10] N. Ullah, M. Marzougui, I. Ahmad, and S. A. Chelloug, "DeepLungNet: An effective DL-based approach for lung disease classification using CRIs," *Electronics*, vol. 12, no. 8, p. 1860, 2023.
- [11] S. Balasubramanian and P. Rajadurai, "Machine learning-based classification of pulmonary diseases through real-time lung sounds," *Int. J. Engineering and Technology Innovation*, vol. 14, no. 1, pp. 85–102, 2024.
- [12] P. V. N. Lakshmi and K. Vedavathi, "Deep learning and probabilistic neural networks based detection and classification of lung diseases for pneumonia," *Int. J. Intelligent Systems and Applications in Engineering*, vol. 12, no. 12s, pp. 708–713, 2024.
- [13] M. Aykanat, O. Kiliç, B. Kurt, and S. B. Saryal, "Lung disease classification using machine learning algorithms," *Int. J. Applied Mathematics Electronics and Computers*, vol. 8, no. 4, pp. 125–132, 2020.
- [14] U. Chutia, A. S. Tewari, and J. P. Singh, "Lung disease classification using deep learning and genetic algorithm," *Neural Computing and Applications*, vol. 37, no. 25, pp. 20519–20534, 2025.
- [15] A. Sze-To, A. Riasatian, and H. R. Tizhoosh, "Searching for pneumothorax in X-ray images using autoencoded deep features," *Scientific Reports*, vol. 11, no. 1, p. 9817, 2021.
- [16] A. Bhandary, G. A. Prabhu, V. Rajinikanth *et al.*, "Deep-learning framework to detect lung abnormality – A study with chest X-ray and lung CT scan images," *Pattern Recognition Letters*, vol. 129, pp. 271–278, 2020.
- [17] N. Nasrullah, J. Sang, M. S. Alam, M. Mateen, B. Cai, and H. Hu, "Automated lung nodule detection and classification using deep learning combined with multiple strategies," *Sensors*, vol. 19, no. 17, p. 3722, 2019.
- [18] G. Kasinathan and S. Jayakumar, "Cloud-based lung tumor detection and stage classification using deep learning techniques," *Biomed Research International*, vol. 2022, p. 4185835, 2022.
- [19] L. F. F. Souza, I. C. L. Silva, A. G. Marques *et al.*, "Internet of medical things: An effective and fully automatic IoT approach using deep learning and fine-tuning to lung CT segmentation," *Sensors*, vol. 20, no. 23, p. 6711, 2020.
- [20] J. Prabakaran and P. Selvaraj, "Advance IoT intelligent healthcare system for lung disease classification using ensemble techniques," *Computer Systems Science and Engineering*, vol. 46, no. 2, pp. 2141–2157, 2023.
- [21] R. R. Irshad, S. Hussain, S. S. Sohail *et al.*, "A novel IoT-enabled healthcare monitoring framework and improved grey wolf optimization algorithm-based deep convolution neural network model for early diagnosis of lung cancer," *Sensors*, vol. 23, no. 6, p. 2932, 2023.