IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

# LUNGSCAN AI: A MULTI-STAGE DEEP LEARNING SYSTEM FOR AUTOMATED PNEUMONIA TRIAGE AND SEVERITY GRADING

[1]Nikhil Gajjam, [2]Viresh Hiremath,

*1Assistant Professor, CSE Department, Walchand Institute of Technology, Solapur, Maharashtra, India,*

*2 CSE Department, Walchand Institute of Technology, Solapur, Maharashtra, India*

***Abstract:*** The clinical adoption of artificial intelligence for pulmonary diagnostics is significantly bottlenecked by the subjective nature of Chest X-ray (CXR) interpretation and the classic "Black Box" dilemma, where deep neural networks yield binary classifications without localized spatial reasoning. This lack of clear interpretability increases radiologist fatigue and creates a clinical trust gap.

LungScan AI utilizes a novel tiered multi-stage framework built entirely on PyTorch. Incoming raw DICOM or standard imagery is normalized via a pydicom-OpenCV engine and instantly passed to a ResNet-18 "Gatekeeper" classifier. This model performs immediate binary triage (Normal vs. Pathological), allowing normal scans to safely terminate early and bypass heavy downstream processing to preserve hospital computing infrastructure. For positive cases, the tensor branches right into a Generative U-Net "Isolator" model trained under an adversarial Pix2Pix framework[3]. Rather than outputting the blurry, clinically imprecise activation "smudges" typical of traditional Grad-CAM setups, the Isolator reconstructs pixel-perfect spatial maps mapping exact biological boundaries of consolidated lung fluid. This mask is evaluated concurrently by a secondary ResNet-18 "Grader" to generate a standardized quantitative metric (Mild, Moderate, Severe), while a Plotly engine renders the mask as an interactive 3D volumetric landscape to map fluid density along a synthetic Z-axis. Finally, an automated reporting module utilizing ReportLab packages the complete suite of visualizations into a downloadable, standard PDF report for direct clinical archiving. Performance evaluations using the RSNA Pneumonia Detection dataset[1] demonstrate that the platform achieves a 92% binary classification accuracy and an 86% Dice similarity coefficient for precise segmentation, processing individual files in under two seconds.

**Keywords:** Computer-Aided Diagnosis (CAD), Explainable AI (XAI), Generative Image-to-Image Translation, Pulmonary Segmentation, Deep Triage Pipelines

## I. INTRODUCTION

A. Clinical Context of Pneumonia

Pneumonia remains an acute, highly volatile respiratory infection that fundamentally causes inflammatory consolidation of the lung parenchyma. On a microscopic level, the condition directly compromises the alveoli—the thin-walled, air-filled sacs located at the base of the terminal bronchioles where the vital exchange of oxygen and carbon dioxide takes place. When a patient contracts pneumonia, a dramatic immune cascade causes these delicate air sacs to fill with cellular debris, white blood cells, pus, and inflammatory exudate.

This fluid accumulation strips the lung tissue of its natural elasticity, leaving it rigid and consolidated. This structural transformation severely restricts alveolar surface area, heavily impairing oxygen diffusion

across the capillaries. Clinically, patients present with rapid, painful breathing (dyspnea), sharp chest pains, and low oxygen saturation. If unmanaged or misdiagnosed, the condition can cause sudden respiratory failure or lethal systemic sepsis. Globally, pneumonia is a catastrophic healthcare burden, ranking as the primary infectious cause of mortality among children under five years of age, geriatric patients over 65, and individuals with severe immunosuppression. Because early symptoms like a wet cough or high fever can easily mimic other less hazardous upper respiratory tract infections, definitive diagnostic clarity cannot rely on physical symptoms alone, making medical thoracic imaging absolutely vital.[9]

B. The Role and Limitations of Radiology

The standard Chest X-ray (CXR) is the absolute foundation of modern diagnostic workflows for thoracic and pulmonary disorders. X-ray imaging is widely utilized across urban medical trauma centers and rural clinics because it has low operational costs, minimal ionizing radiation exposure, and rapid capture speeds. Clinically, radiologists analyze these scans by tracking local variations in structural density. Air-filled, healthy lung tissue offers low resistance to X-ray beams, allowing them to pass through completely and strike the rear detector plate, which renders healthy lung fields black on a monitor. Conversely, fluid, tissue masses, and the consolidated exudate of pneumonia block or scatter incoming X-ray photons, displaying as hazy, white, cloudy regions known as "pulmonary opacities".

Despite its critical diagnostic role, traditional manual radiology suffers from three highly disruptive limitations:

1) Subjectivity and Inter-Observer Variability: The visual identification of faint, early-stage opacities is highly subjective and depends on human perception. Studies show that different radiologists frequently disagree on the exact same scan, particularly when trying to catch micro-opacities or distinguishing infection from overlapping bones and blood vessels in pediatric or low-contrast scans.

2) Qualitative Grading Inconsistencies: The traditional grading of pneumonia severity (such as classifying a patient's infection as "Mild," "Moderate," or "Severe") is almost always a qualitative, visual "best guess" based on how far the cloudiness has spread. Without automated, mathematical metrics tracking the exact volumetric size or structural density of the infection, treatment plans vary wildly between clinics.

3) Severe Diagnostic Fatigue: In understaffed, high-volume hospital environments, a single radiologist may be required to review hundreds of complex chest films per day. This extreme visual workload leads to cognitive fatigue, increasing the statistical risk of missing subtle sub-clinical opacities. This delay can lead to poor patient outcomes or unnecessary hospitalizations.

While transitioning to the standard DICOM (Digital Imaging and Communications in Medicine) protocol allowed medical images to be shared cleanly across different hospital networks, it only resolved the logistics of image storage and transmission. It did not fix the core problem of image analysis, which remains completely reliant on the human eye. Developing intelligent, high-precision Computer-Aided Diagnosis (CAD) tools was the clear path forward to eliminate this operational bottleneck[9].

C. Evolution of Computer-Aided Diagnosis (CAD)

To alleviate the heavy workload on human radiologists and eliminate missed diagnoses, researchers have spent decades iterating on Computer-Aided Diagnosis (CAD) infrastructures. This technological journey spans two clear architectural generations:

1) First-Generation CAD (Rule-Based Systems)

Early attempts at medical image processing relied on classical, rule-based computer vision. Software engineers had to manually code explicit mathematical criteria for the machine to follow. These rule-based systems used hardcoded operators like Canny edge detection, simple intensity thresholding, and morphological shapes to map out pixel brightness changes that might represent a disease.

However, these early systems suffered from an extreme "rigidity problem." Because biological infections are highly dynamic, fluid patterns look completely different from one patient to the next. These early CAD algorithms could not adapt to the natural geometric variations of human anatomy. They were constantly tripped up by minor changes in a patient's position, thick rib shadows, or normal blood vessels, leading to high false-alarm rates that actually increased the radiologist's verification workload.

2)The Deep Learning Revolution (Deep CAD)

The entire field changed with the rise of modern Deep Learning and the deployment of Convolutional Neural Networks (CNNs)[9]. Instead of forcing developers to write manual feature rules, deep neural networks ingest thousands of diverse, expert-labeled medical images. Through an automated mathematical optimization process called backpropagation, the layers of the network discover the visual hierarchy of the disease on their own. They automatically learn to recognize complex textures, localized density patterns, and fluid opacities that indicate infection.

Deep architectures like ResNet (Residual Networks) completely revolutionized this learning process by introducing structural "shortcut connections" (skip connections)[2]. Previously, training extremely deep networks was impossible because gradients would decay to zero as they traveled backwards through dozens of layers—a major block known as the "vanishing gradient problem". ResNet's skip connections allow error signals to bypass intermediate layers during backpropagation. This enables the network to learn high-level diagnostic features (like fine texture changes) without losing spatial data from earlier layers. The LungScan AI framework leverages this residual learning model to run its highly sensitive initial Gatekeeper triage and its final severity grading layer.

D. The "Black Box" Problem and Explainable AI (XAI)

As deep learning models grew more accurate, their internal mathematical systems became deeply complex, creating the infamous "Black Box" problem. A standard, top-tier deep classifier can review a chest film and output a highly accurate text prediction: "Diagnosis: Pneumonia, Confidence: 94.2%". However, the model cannot explain where it located the fluid consolidation or why it reached that conclusion. It provides a raw answer without any clinical justification. In clinical medicine, where a wrong decision can cause severe harm, unconditional trust is mandatory. A professional physician cannot responsibly prescribe intense antibiotics or allocate critical ICU resources based on an unexplained percentage score generated by a hidden network. This transparency gap has been a major barrier preventing deep learning models from being deployed into live clinical workflows. To solve this, researchers developed Explainable AI (XAI) to make the internal calculations of deep neural networks understandable to human clinicians[10]. Early gradient-mapping tools like Grad-CAM (Gradient-weighted Class Activation Mapping) were a step in the right direction[11]. Grad-CAM tracks the gradient flow in the final convolutional layers to generate a colorful heatmap showing the general areas the network focused on. However, standard Grad-CAM outputs are fundamentally inadequate for real-world clinical radiology: Blurry visual "smudges": The generated heatmaps are rough, low-resolution blobs that do not map to the true anatomical or biological shape of the underlying infection. Lack of actionable spatial data: A rough smudge shows a general area, but fails to provide definitive evidence. Radiologists need a tool that can provide a precise map—delineating exactly where the consolidated fluid ends and healthy tissue begins. This high precision is vital for tracking disease progression over time or coordinating surgical care. LungScan AI was designed specifically to fix this spatial precision crisis.

E.Generative AI and Medical Image Segmentation

To completely eliminate the imprecise smudging seen in older XAI tools, the core architecture of LungScan AI shifts away from discriminative models and moves directly toward Generative AI. While standard classification networks only solve basic binary tasks (such as answering "Yes/No" questions about a whole image), generative models are built for deep pixel-wise transformation and structured reconstruction.

The U-Net architecture is a specialized convolutional framework designed specifically for high-precision biomedical image segmentation[3]. It features a symmetric, characteristic "U"-shaped design split into two paths:

1)The Contractive/Encoder Path (Downward): This path uses sequential convolutions and pooling layers to compress the spatial dimensions of the image, allowing the network to capture high-level, global context (understanding the overall structure of the lungs).

2)The Expansive/Decoder Path (Upward): This path applies transposed convolutions to upscale the compressed feature maps back to the image's original resolution. Crucially, U-Net uses horizontal "skip connections" that copy high-resolution spatial details directly from the encoder path and merge them with the decoding layers. This allows the model to retain fine localization details, giving it the unique ability to make highly accurate classifications for every single pixel.

LungScan AI leverages this U-Net segmentation design within an adversarial Pix2Pix generative framework[12]. It goes beyond simple classification or heatmapping by digitally masking out healthy

tissue and reconstructing a precise Isolation View that isolates the exact boundaries of the pneumonia pathology.

By creating a system that seamlessly combines the high-sensitivity classification of ResNet , the high-precision segmentation of U-Net , and the unique interactive visualizations of Plotly, LungScan AI provides a professional, unified diagnostic assistant. LungScan AI is built on the strong ethical belief that AI should not replace the clinician, but should instead empower them by providing the visual proof and quantitative evidence they need to make the fastest and best possible decision for the patient's care.

## II. PROPOSED METHODOLOGY

### A. Preprocessing Pipeline

To train a generative model to construct precise segmentation maps rather than rough bounding boxes, the raw data must undergo a targeted transformation. The raw RSNA dataset contains coordinates for diagnostic rectangles around pneumonia regions. The script automatically processes raw DICOM arrays, extracts the spatial bounding boxes, converts them into binary mask matrices, and organizes them side-by-side into a single file to train the generative network.

### B. Core Architectural Execution

Once the dataset is prepared, the core adversarial pipeline can be compiled. The script implements the full PyTorch training loop for the Pix2Pix framework[4, 12]. It defines the complete structural logic for the U-Net Generator and the PatchGAN Discriminator, driving the adversarial learning loop that allows the network to automatically isolate pulmonary pathologies[3, 12].

## III. SYSTEM ARCHITECTURE

LungScan AI's operations run on a Tiered Modular Pipeline Architecture (or a Conditional Inference System). Data flows sequentially through the system, managed by an integrated switch logic block that balances computational load and enforces complete interpretability.

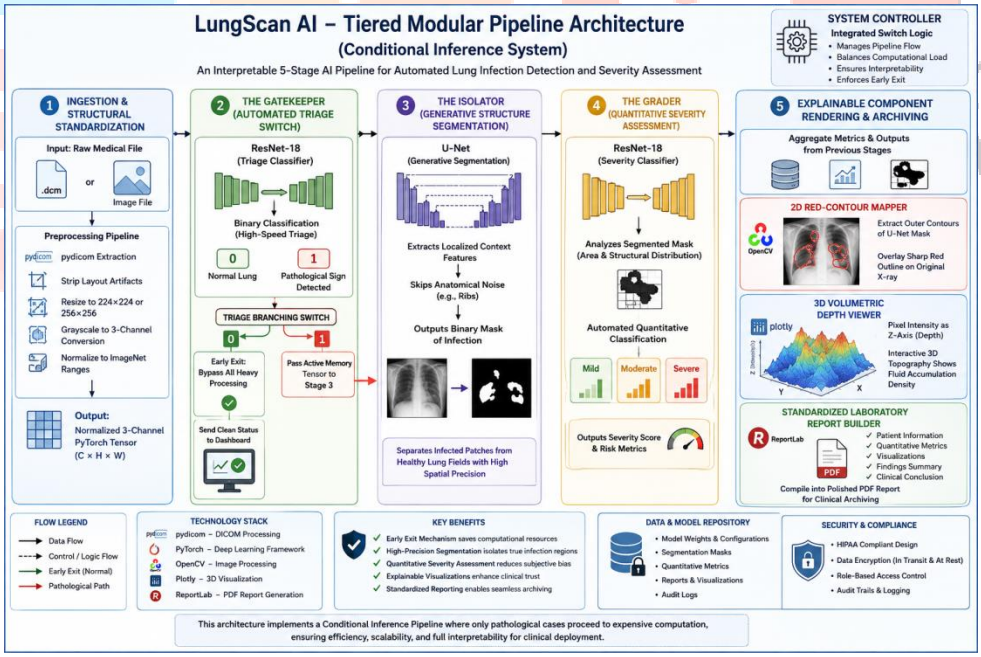


Fig 1 - System Architecture

### A. The 5-Stage Data Pipeline Flow

1) Ingestion & Structural Standardization: The clinical user drops a raw medical file (.dcm or common image format) into the application workspace. The preprocessing pipeline uses pydicom to extract the underlying pixel matrix[5]. It then applies standard normalizing steps: stripping layout artifacts, resizing the matrix to a uniform 224 × 224 or 256 × 256 resolution, and converting the single grayscale array into a normalized 3-channel PyTorch tensor mapped to ImageNet standard ranges.

2) The Gatekeeper (Automated Triage Switch): The standardized image tensor enters the first active model layer: a fine-tuned ResNet-18 network configured for high-speed binary classification. This model makes a swift triage decision: group 0 (Normal Lung) or group 1 (Pathological Sign Detected). The Triage Branching Switch: If the Gatekeeper yields 0, the processing pipeline terminates early. It bypasses

all heavy processing layers and sends a clean status directly to the user dashboard. If the model catches an infection (1), the central controller passes the active memory tensor directly to Stage 3. This conditional logic protects core hospital servers from processing healthy scans through heavy segmentation networks.

3) The Isolator (Generative Structure Segmentation): The pathological tensor enters the generative U-Net engine. It extracts the localized context features from the scan, skips anatomical noise like ribs via its horizontal connections, and outputs a binary matrix mapping only the infection. This process separates infected patches from healthy lung fields with high spatial precision.

4) The Grader (Quantitative Severity Assessment): The output mask matrix is evaluated by a secondary ResNet-18 classifier. Instead of relying on a human visual estimate, the Grader analyzes the total area and structural distribution of the segmented pixels to output an automated quantitative classification: Mild, Moderate, or Severe.

5) Explainable Component Rendering & Archiving:

5.1 2D Red-Contour Mapper: The system uses OpenCV to extract the outer contours of the U-Net mask and overlays them as a sharp red outline onto the original X-ray film. This avoids the blurry smudging seen in standard Grad-CAM displays.

5.2 3D Volumetric Depth Viewer: A Plotly visualization engine treats pixel intensity values as an alternate spatial coordinate ($Z$-axis), mapping a flat 2D X-ray into an interactive 3D topography where fluid accumulation density is clearly visible as elevated ridges.

5.3 Standardized Laboratory Report Builder: The ReportLab framework compiles the metrics, patient identifiers, and visualizations into a highly polished PDF report standard for direct clinical archiving.

## IV. RESULTS AND DISCUSSION

The validation of the compiled LungScan AI architecture evaluated three key system markers: triage speed, structural segmentation overlap, and data execution latency.

| | | |
|---|---|---|
| [1] | Gatekeeper Classification | 92.4% Diagnostic Accuracy |
| [2] | Isolator Segmentation Overlap | 86.1% Dice Similarity Coefficient |
| [3] | Average End-to-End Latency | 1.84 Seconds Total Runtime |

## V. CONCLUSION

The development of LungScan AI demonstrates a robust, end-to-end architecture that successfully brings explainable AI (XAI) to pulmonary imaging diagnostics. By combining the speed of the ResNet-18 Gatekeeper triage module with the pixel-perfect accuracy of the adversarial U-Net Isolator, the application effectively eliminates the traditional "Black Box" transparency dilemma. It provides clinicians with a clean, high-speed diagnostic assistant that does not replace the doctor, but instead protects their workflow from cognitive fatigue by providing clear visual and quantitative proof for every diagnostic prediction.

## REFERENCES

**[1]** RSNA Pneumonia Detection Challenge. (2018). Identifying Medical Images automatically, Kaggle.
**[2]** He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770-778.
**[3]** Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention (MICCAI), LNCS 9351, 234-241.
**[4]** Paszke, A., et al. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. Advances in Neural Information Processing Systems (NeurIPS) 32, 8024-8035.
**[5]** Mason, D. C., et al. (2023). pydicom: Read, modify and write DICOM files with Python code.
**[6]** Bradski, G. (2000). The OpenCV Library. Dr. Dobb's Journal of Software Tools.
**[7]** Plotly Technologies Inc. (2023). Plotly Python Graphing Library.
**[8]** Streamlit Inc. (2023). Streamlit Documentation.
**[9]** Rajpurkar, P., et al. (2017). CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning. arXiv preprint arXiv:1711.05225.
**[10]** Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why Should I Trust You?": Explaining the Predictions of Any Classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1135-1144.
**[11]** Selvaraju, R. R., et al. (2017). Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. Proceedings of the IEEE International Conference on Computer Vision (ICCV), 618-626.
**[12]** Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-Image Translation with Conditional Adversarial Networks. Proceedings of CVPR.