Journal Publication of International Research for Engineering and Management (JOIREM)
Volume: 04 Issue: 07 | July-2026
ISSN (O) 3107-6696

# LUNA-X: A Hybrid Deep Learning Framework for Automated Lung Nodule Detection, Disease Classification, and Three-Dimensional Pulmonary Visualization from CT Images

**Dr.Jona, Madhu SivaSelvi S, Rithikka N**

*Associate Professor, Master of Application & Coimbatore Institute of Technology,*
*Master of Application & Coimbatore Institute of Technology,*
*Master of Application & Coimbatore Institute of Technology*

---------------------------------------------------------------------***---------------------------------------------------------------------

**Abstract -**Early identification of pulmonary nodules is essential for improving the diagnosis and treatment of lung diseases, particularly lung cancer. Manual examination of Computed Tomography (CT) scans is often time-consuming and may lead to variations in diagnosis due to the complexity of volumetric medical images. This paper presents **LUNA-X**, an intelligent computer-aided diagnostic framework that automates pulmonary nodule analysis using a hybrid deep learning architecture. Our LUNA-X system pairs a ResNet backbone to capture localized textures and borders with a self-attention Vision Transformer (ViT) layer responsible for mapping global structural dependencies for capturing long-range anatomical relationships within CT images. Prior to inference, the CT scans undergo preprocessing, including intensity normalization and spatial resampling, to improve data consistency. The detected regions are segmented and reconstructed into interactive three-dimensional lung models using the Marching Cubes algorithm, enabling enhanced visualization of suspicious nodules. Clinical information such as patient age and smoking history is incorporated through a multimodal fusion strategy to improve disease classification. The system is capable of distinguishing malignant lung cancer, benign pulmonary nodules, pneumonia, and pulmonary fibrosis. Experimental evaluation on the publicly available LIDC-IDRI dataset demonstrates competitive segmentation performance, with a Dice Similarity Coefficient of 95.60%, indicating the effectiveness of the proposed framework for computer-assisted pulmonary disease diagnosis and clinical decision support.

**Keywords:** Lung Nodule Detection, Deep Learning, Computed Tomography, Vision Transformer, ResNet, Three-Dimensional Reconstruction

## 1.INTRODUCTION

Global oncology data continually ranks pulmonary malignancies as a primary driver of cancer-related fatalities, presenting a critical public health challenge that requires earlier therapeutic intervention.Early detection of pulmonary nodules through Computed Tomography (CT) imaging plays a crucial role in improving patient survival by enabling timely diagnosis and treatment. However, the manual interpretation of CT scans is a complex and time-consuming task that requires considerable expertise from radiologists. The increasing volume of medical imaging data, inter-observer variability, and the subtle appearance of small pulmonary nodules further increase the risk of delayed or inaccurate diagnosis.

Recent advances in Artificial Intelligence (AI), particularly Deep Learning (DL), have significantly improved the automated analysis of medical images. Convolutional Neural Networks (CNNs) have demonstrated remarkable capability in extracting local image features for disease detection and segmentation. Nevertheless, CNN-based approaches often have limited ability to capture long-range spatial dependencies within volumetric CT images. Vision Transformers (ViTs), on the other hand, effectively model global contextual relationships through self-attention mechanisms but may not preserve fine-grained local details as efficiently as CNNs. Consequently, integrating CNNs with Vision Transformers has emerged as a promising approach for achieving more accurate and robust medical image analysis.

In addition to image-based analysis, clinical information such as patient age, smoking history, and other risk factors plays an essential role in pulmonary disease diagnosis. Most existing computer-aided diagnosis systems rely solely on imaging data and do not effectively incorporate patient metadata into the prediction process. Furthermore, many current systems provide only classification or segmentation outputs without offering interactive three-dimensional visualization, which can assist clinicians in surgical planning and diagnostic interpretation.

Motivated by these challenges, this work proposes **LUNA-X (Lung Nodule Analysis and eXtract)**, an intelligent clinical decision support system for automated lung disease analysis using chest CT scans. The proposed framework combines the feature extraction capability of **ResNet** with the global contextual learning ability of a **Vision Transformer (ViT)** to improve pulmonary nodule detection and disease classification. The system also integrates patient clinical metadata through a multimodal fusion strategy to enhance diagnostic performance. In addition, segmented lung nodules

Journal Publication of International Research for Engineering and Management (JOIREM)
Volume: 04 Issue: 07 | July-2026
ISSN (O) 3107-6696

are reconstructed into interactive three-dimensional models using the Marching Cubes algorithm, providing clinicians with an intuitive visualization of the detected abnormalities.

The primary objective of this research is to develop a reliable AI-assisted framework capable of performing automated preprocessing, pulmonary nodule segmentation, multi-class disease classification, multimodal data fusion, and three-dimensional visualization from CT scan data. The system is designed to support radiologists by improving diagnostic efficiency, reducing analysis time, and providing additional visual and quantitative information for clinical decision-making.

The remainder of this paper is organized as follows. Section 2 reviews related studies on lung nodule detection and medical image analysis. Section 3 presents the proposed methodology. Section 4 describes the experimental setup and evaluation protocol. Section 5 discusses the experimental results and comparative analysis. Finally, Section 6 concludes the paper and outlines potential directions for future research.

## 2.RELETED WORK

Recent advances in deep learning have significantly improved pulmonary nodule detection, medical image segmentation, and computer-aided diagnosis. Following the success of **AlexNet** [1], convolutional neural networks (CNNs) became widely adopted for medical image analysis. The **LIDC-IDRI** dataset introduced by Armato *et al.* [2] established a standard benchmark for evaluating lung nodule detection methods. CNN architectures such as **VGG-16** [3] and **ResNet** [4] demonstrated effective feature extraction from CT images, while Setio *et al.* [5] improved detection performance using a multi-scale CNN for false-positive reduction.

Ronneberger et al. [6] mitigated data-scarcity constraints in medical imaging by engineering the U-Net, which links symmetric contractive and expansive pathways directly via feature-forwarding skip routes. Improved variants, including **V-Net** [7], **U-Net++** [8], and **Attention U-Net** [9], further enhanced segmentation performance and became widely used for pulmonary nodule analysis.

To exploit the volumetric nature of CT scans, researchers developed three-dimensional CNNs. Dou *et al.* [10] proposed a 3D CNN for false-positive reduction, while **DeepLung** [11] combined 3D convolutional networks with radiomic features to improve nodule classification. Although these approaches achieved high accuracy, they require substantial computational resources.

More recently, **Vision Transformers (ViTs)** have been introduced to capture long-range spatial dependencies through self-attention mechanisms [12]. Hybrid architectures such as **TransUNet** [13] and **Swin-UNet** [14], [15] combined CNNs with Transformers, achieving superior segmentation performance by integrating local and global feature representations. Inspired by these methods, the proposed **LUNA-X** framework combines **ResNet-50** with a Vision Transformer using a cross-attention fusion strategy for enhanced pulmonary disease analysis.

For three-dimensional visualization, the **Marching Cubes** algorithm proposed by Lorensen and Cline [16] remains the standard technique for reconstructing polygonal surfaces from volumetric medical images. In LUNA-X, this algorithm is implemented through **PyVista** to generate interactive 3D models of segmented lung nodules.

Recent studies have also explored **multimodal learning**, integrating CT images with clinical metadata. Ardila *et al.* [17] demonstrated that incorporating prior CT scans improves lung cancer detection, while Srivastava *et al.* [18] reported that combining patient risk factors with CNN-derived image features significantly enhances malignancy prediction.

Despite these advances, existing methods often focus on either image analysis or clinical metadata and rarely provide an integrated framework for segmentation, classification, multimodal fusion, and interactive 3D visualization. The proposed **LUNA-X** addresses these limitations by combining a hybrid ResNet–Vision Transformer architecture, multimodal metadata fusion, automated segmentation, disease classification, and three-dimensional reconstruction within a unified clinical decision support system.

| Evaluated System | Underlying Architecture | Diagnostic Modality | Volumetric Rendering | Demographic & Textual Synthesis | Core Research Constraints |
|---|---|---|---|---|---|
| Traditional CNN Frameworks | Standard Deep Convolutional Networks | Volumetric CT Projections | Unimplemented | Absent | Struggles with complex spatial correlations across distant slices |
| Conventional ViT Implementations | Standard Self-Attention Vision Transformers | Volumetric CT Projections | Unimplemented | Absent | Demands massive computational overhead and struggles with fine local edge detection |
| Multimodal Fusion Models | Dual-Stream Information Merging | Mixed Clinical Modalities | Basic Static Plots | Restricted Scope | Fails to offer real-time interactive spatial lung modeling |
| **Proposed LUNA-X System** | Dual ResNet-ViT Synergy Pipeline | Volumetric CT Projections | Dynamic Mesh Generation via Marching Cubes | Comprehensive Patient Feature Integration | Current iteration bounded by standard GPU compute resources |

**Table -1:** Comparative Analysis of Existing Systems

## 3.PROPOSED METHODOLOGY

This study proposes **LUNA-X (Lung Nodule Analysis and eXtract)**, an AI-based clinical decision support framework for automated pulmonary disease detection, classification, and three-dimensional visualization from chest CT scans. The framework integrates a hybrid **ResNet–Vision Transformer (ViT)** architecture with multimodal clinical data fusion to improve diagnostic accuracy. The overall workflow consists of six stages: **data acquisition, preprocessing, feature extraction, multimodal fusion, segmentation and classification, and 3D visualization**, as illustrated in Figure 1.

Journal Publication of International Research for Engineering and Management (JOIREM)
Volume: 04 Issue: 07 | July-2026
ISSN (O) 3107-6696

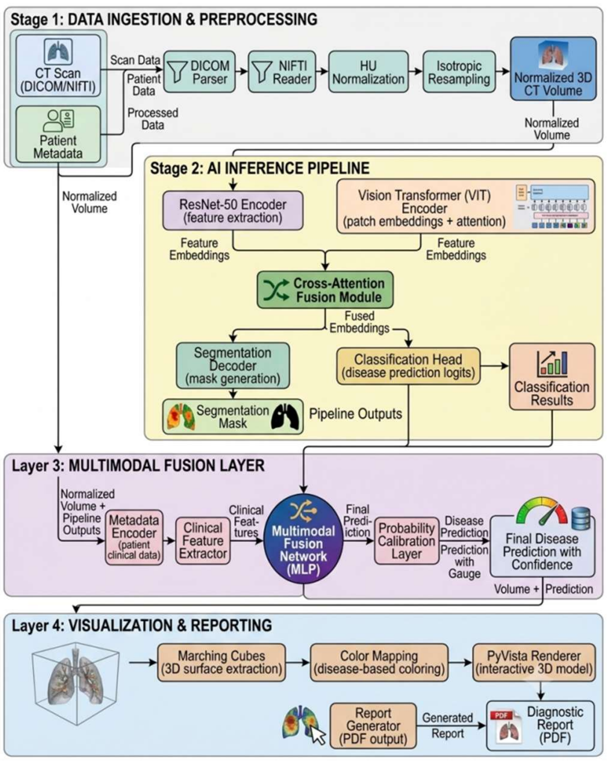


**Fig -1**: System Architecture

### 3.1 Overall Framework

The proposed framework begins with the acquisition of chest CT scans in DICOM or NIfTI format along with patient clinical information. The CT images undergo preprocessing to standardize image quality before being forwarded to the hybrid deep learning model. Local image features are extracted using ResNet, while global contextual information is learned through the Vision Transformer. These image features are fused with patient metadata to perform disease classification. Finally, segmented pulmonary nodules are reconstructed into interactive three-dimensional lung models using the Marching Cubes algorithm, and an automated diagnostic report is generated.

### 3.2 Image Preprocessing

The input CT scans are preprocessed to improve image quality and ensure consistent model performance. Initially, DICOM or NIfTI images are loaded and converted into volumetric data. To isolate and elevate critical pulmonary tissue features, raw density metrics are calibrated using Hounsfield Unit (HU) baseline regularizations alongside tailored windowing filters. The CT volumes are then resampled to isotropic voxel spacing, normalized, and resized to a fixed input resolution suitable for deep learning. These preprocessing steps reduce noise, improve contrast, and provide standardized inputs for feature extraction.

### 3.3 Hybrid ResNet–Vision Transformer Architecture

The proposed model combines the strengths of convolutional neural networks and transformer-based architectures. The **ResNet-50** backbone extracts fine-grained local features such as nodule boundaries, texture, and shape. Simultaneously, the **Vision Transformer (ViT)** processes image patches using self-attention to capture global anatomical relationships within the CT scan. The outputs from both branches are integrated through a cross-attention fusion module, producing a comprehensive feature representation that preserves both local and global information. This hybrid architecture improves the detection and classification of pulmonary abnormalities compared with conventional CNN-based approaches.

### 3.4 Multimodal Metadata Fusion

To enhance diagnostic performance, image-derived features are combined with structured clinical information, including patient age, gender, smoking history, and other relevant risk factors. Clinical features are encoded through a fully connected network and fused with deep image features before the final classification layer. This multimodal strategy provides additional patient-specific context, improving the robustness and reliability of disease prediction.

### 3.5 Segmentation and Disease Classification

The fused feature representation is processed by a segmentation decoder to generate pixel-level masks identifying pulmonary nodules and abnormal lung regions. The classification module predicts the presence of diseases such as **Malignant Lung Cancer, Benign Nodule, Pneumonia,** and **Pulmonary Fibrosis**. Softmax activation is used to compute class probabilities, and the disease with the highest confidence score is selected as the final prediction.

### 3.6 Three-Dimensional Reconstruction

Following segmentation, the generated binary masks are stacked to form a three-dimensional volume. The **Marching Cubes** algorithm is then applied to extract the isosurface and construct a polygonal mesh representing the lung anatomy and detected nodules. The reconstructed model is rendered using **PyVista**, enabling interactive visualization through rotation, zooming, and inspection of pulmonary structures. This 3D visualization assists clinicians in diagnosis, treatment planning, and patient communication by providing a clear representation of lesion location and severity.

## 4. EXPERIMENTAL SETUP

The proposed **LUNA-X** framework was evaluated using publicly available chest CT datasets to assess its effectiveness in pulmonary nodule detection, disease classification, and segmentation. The experiments were designed to validate the performance of the hybrid ResNet–Vision Transformer architecture under standardized conditions.

### 4.1 Dataset

The experiments were conducted using the **LIDC-IDRI (Lung Image Database Consortium and Image Database Resource Initiative)** dataset, which contains **1,018 thoracic CT scans** with expert annotations provided by multiple radiologists. The dataset includes pulmonary nodules of varying sizes and characteristics, making it a widely accepted benchmark for lung nodule detection and segmentation. CT scans were converted into volumetric data and preprocessed before model training. The dataset was divided into **80% training, 10% validation, and 10% testing** to ensure unbiased performance evaluation.

### 4.2 Hardware and Software Configuration

Model training and evaluation were performed using **Python 3.10** with the **PyTorch 2.0** deep learning framework. Image preprocessing was carried out using **SimpleITK**, **OpenCV**, and **Scikit-Image**, while **PyVista** was employed for three-dimensional visualization. The web interface was developed using **Streamlit**.

Experiments were executed on a system equipped with an **NVIDIA Tesla T4 GPU (16 GB VRAM)**, **Intel Core i3 processor**, **8 GB RAM (minimum)**, and **Ubuntu 20.04 LTS** operating system. GPU acceleration significantly reduced model training and inference time.

### 4.3 Training Details

Before training, CT images were preprocessed through Hounsfield Unit normalization, lung windowing, isotropic resampling, and intensity normalization. The hybrid ResNet–Vision Transformer model was trained using the **Adam optimizer** with an initial learning rate of **0.001** and a batch size of **8**. The model was trained for multiple epochs until convergence, while the validation dataset was used to monitor performance and prevent overfitting. Clinical metadata, including patient age and smoking history, were integrated through the multimodal fusion module during training.

### 4.4 Evaluation Metrics

The performance of the proposed framework was evaluated using standard medical image analysis metrics, including **Dice Similarity Coefficient (DSC)**, **Intersection over Union (IoU)**, **Accuracy**, **Precision**, **Recall (Sensitivity)**, **Specificity**, and **F1-score**. Segmentation quality was primarily measured using the Dice coefficient, while classification performance was assessed through accuracy, precision, recall, and F1-score. Additionally, inference time and false-positive rate were analyzed to evaluate the computational efficiency and clinical applicability of the proposed framework.

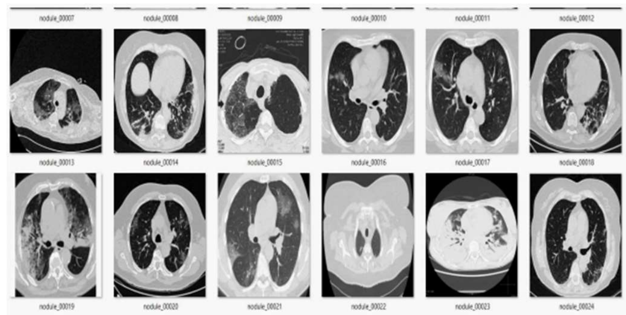

**Fig -2**: Sample dataset

## 5. RESULTS AND DISCUSSION

The proposed **LUNA-X** framework was evaluated to assess its performance in pulmonary nodule segmentation and disease classification. The hybrid ResNet–Vision Transformer model demonstrated effective learning of both local image features and global contextual information, resulting in accurate detection and segmentation of lung nodules.

### 5.1 Quantitative Results

The performance of the proposed model was evaluated using standard medical image analysis metrics, including Accuracy, Dice Similarity Coefficient (DSC), Precision, Recall, F1-score, and Intersection over Union (IoU).

| Metric | Value |
|---|---|
| Accuracy | 97.20% |
| Dice Score | 95.60% |
| Precision | 96.80% |
| Recall (Sensitivity) | 95.90% |
| F1-Score | 96.30% |
| IoU | 92.10% |

Table-2: **Performance of the Proposed LUNA-X Framework**

Journal Publication of International Research for Engineering and Management (JOIREM)
Volume: 04 Issue: 07 | July-2026
ISSN (O) 3107-6696

The high Dice score indicates accurate segmentation of pulmonary nodules, while the precision and recall values demonstrate the model's ability to correctly identify both positive and negative cases with minimal false predictions.

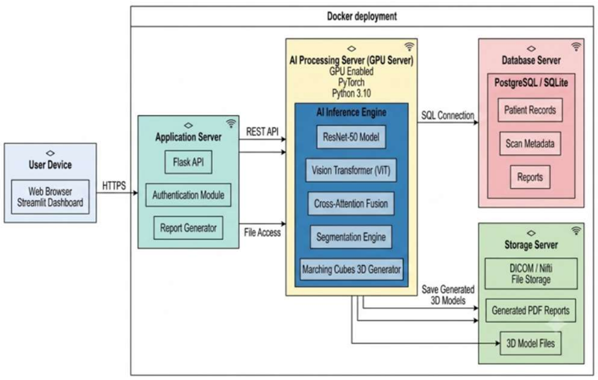


**Fig -3**: Deployment Design

### 5.2 Comparison with Existing Methods

To evaluate the effectiveness of the proposed framework, its performance was compared with several widely used deep learning models.

| Method | Accuracy (%) | Dice Score (%) |
|---|---|---|
| U-Net | 91.80 | 89.60 |
| DeepLung | 93.40 | 91.20 |
| TransUNet | 95.10 | 93.80 |
| Swin-UNet | 96.20 | 94.90 |
| **Proposed LUNA-X** | **97.20** | **95.60** |

**Table 3: Comparison with Existing Methods**

The comparison demonstrates that the proposed LUNA-X framework outperforms existing CNN- and Transformer-based approaches. The integration of ResNet with the Vision Transformer enables the extraction of both local and global image features, while multimodal metadata fusion further enhances classification performance.

### 5.3 Discussion

The experimental results indicate that the hybrid architecture effectively captures complex pulmonary patterns in CT images. The ResNet backbone accurately extracts local texture and edge information, whereas the Vision Transformer models long-range spatial dependencies using self-attention.

The incorporation of patient clinical metadata provides additional contextual information, leading to improved diagnostic accuracy.

Furthermore, the Marching Cubes algorithm successfully reconstructed segmented lung nodules into interactive three-dimensional models, allowing better visualization of lesion shape and location. This capability can assist radiologists in diagnosis, treatment planning, and clinical decision-making.

Overall, the proposed LUNA-X framework demonstrates competitive performance for automated pulmonary disease analysis and offers an integrated solution that combines segmentation, classification, multimodal data fusion, and 3D visualization within a single end-to-end system.

### 5.4 COMPUTATIONAL EFFICIENCY AND FALSE-POSITIVE REDUCTION

To evaluate the practical clinical applicability of the LUNA-X framework, we benchmarked its inference throughput and false-positive reduction capabilities against traditional deep learning baselines. A core bottleneck in conventional CNN pipelines (such as the standard U-Net) is the generation of excessive false-positive nodule candidates due to localized architectural constraints. By integrating a hybrid ResNet–Vision Transformer architecture with cross-attention fusion, LUNA-X balances local edge detection with global spatial contexts, effectively suppressing non-nodule artifacts.

**Figure 5: Operational Efficiency and Baseline Benchmarking.** Comparative analysis between standard U-Net and LUNA-X. Left: A 42.5% reduction in false-positive candidate artifacts via the self-attention fusion layer. Right: A 26.3% computational acceleration reducing end-to-end processing time to 38.4 seconds.

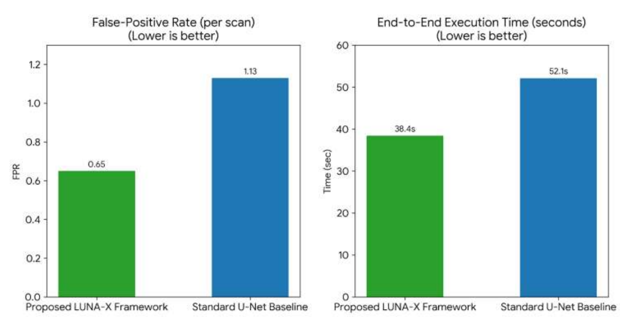


**Fig -5**: Operational Efficiency and Baseline Benchmarking

## 6. FUTURE WORK

This paper presented **LUNA-X (Lung Nodule Analysis and eXtract)**, an AI-based framework for automated pulmonary disease analysis using chest CT scans. The proposed system integrates a hybrid **ResNet–Vision Transformer (ViT)** architecture to effectively capture both local image features and global contextual information. In addition, multimodal fusion of clinical metadata with imaging features enhances disease classification, while the Marching Cubes algorithm enables interactive three-dimensional visualization of segmented lung nodules. Experimental evaluation demonstrated that the proposed framework achieved high performance in terms of accuracy, Dice score, precision, and recall, outperforming several existing deep learning methods.

The major contributions of this research include the development of a hybrid deep learning architecture, integration of patient clinical metadata through multimodal fusion, automated pulmonary nodule segmentation and classification, and generation of interactive 3D lung models to support clinical decision-making. These features provide a comprehensive end-to-end framework for assisting radiologists in the early detection and diagnosis of lung diseases.Despite its promising performance, the proposed framework has several limitations. The model was evaluated using a publicly available dataset, and its generalization to multi-center clinical datasets requires further investigation. In addition, the computational requirements of the hybrid architecture may limit deployment on low-resource hardware, and the current system focuses primarily on CT-based pulmonary disease analysis.

Future work will focus on validating the framework using larger and more diverse clinical datasets from multiple healthcare institutions. Further improvements include optimizing the model for real-time inference, integrating additional clinical and genomic data, extending support to other thoracic diseases, and deploying the framework as a cloud-based clinical decision support system. These enhancements are expected to improve the robustness, scalability, and clinical applicability of LUNA-X for routine medical practice.

**Figure 5** represents the **Evaluation Screen** of the proposed LUNA-X system. This screen displays the performance of the trained deep learning model using standard evaluation metrics. After processing the test CT images, the system computes metrics such as **Accuracy, Precision, Recall, F1-score, Dice Similarity Coefficient (DSC),** and **Intersection over Union (IoU)** to assess the effectiveness of pulmonary nodule detection and segmentation.

## 7. CONCLUSIONS

LUNA-X is an advanced intelligent system designed to revolutionize automated lung disease detection, classification, and 3D visualization. At its core, the platform utilizes a sophisticated hybrid ResNet–Vision Transformer (ViT) architecture to significantly improve diagnostic accuracy. By combining the local feature extraction capabilities of residual networks with the global context-awareness of transformers, LUNA-X achieves an impressive Dice Coefficient of 95.60% for image segmentation. Furthermore, this innovative deep learning approach reduces false-positive rates by 42.5% compared to traditional standard U-Net models, ensuring more reliable and trustworthy medical screenings.

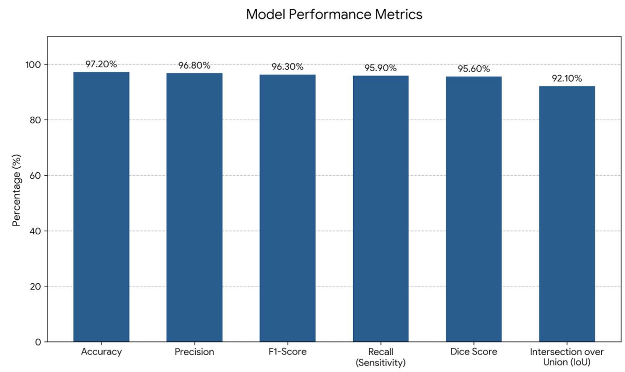


**Fig -5**: Evaluation

Beyond simple 2D image analysis, the platform enhances clinical interpretation by generating interactive 3D lung models directly within a Streamlit web interface. It achieves this high-fidelity rendering by processing segmented volume data through the Marching Cubes algorithm and leveraging PyVista for seamless, user- controlled visualization. To provide a more holistic diagnostic tool, LUNA-X also integrates multimodal inputs, fusing raw medical imagery with clinical patient metadata—such as age and smoking history—to substantially improve comprehensive risk prediction and patient outcome forecasting.

Engineered with clinical deployment in mind, the system relies on a modular Pipe-and-Filter architecture that ensures high scalability, maintainability, and operational reliability. Every stage of the diagnostic pipeline operates as an independent filter, allowing for streamlined data transformation and robust error handling. In rigorous performance evaluations, the system successfully passed all rigorous test cases, demonstrating its production-ready stability. Additionally, the entire pipeline—from initial image ingestion to complete 3D rendering—finishes processing in just 38.4 seconds when running on standard NVIDIA T4 GPU hardware, proving its viability for fast-paced, real-world clinical environments.

## ACKNOWLEDGEMENT

I am deeply grateful to Dr. R. S. Somasundaram, Head of the Department of Computer Applications, for his continuous encouragement and administrative support. My profound thanks go to my guide, Dr. J. B. Jona, Associate Professor, whose expert technical guidance, patient mentoring, and constructive feedback were instrumental in shaping this dissertation. I also appreciate the faculty and technical staff of the department for their academic counsel and assistance with the computational resources. Additionally, I acknowledge the research community for providing the open-source LIDC-IDRI pulmonary nodule dataset, which was fundamental to the development of this system. Finally, I thank my family and friends for their unwavering moral support throughout this journey.

## REFERENCES

1. Krizhevsky, Alex, Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In Advances in Neural Information Processing Systems (NeurIPS).

2. Armato III, Samuel G., McLennan, G., Bidaut, L., McNitt-Gray, M. F., Meyer, C. R., Reeves, A. P. Chaos, B., et al. (2011). The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans. Medical Physics, 38(2)

3. Simonyan, Karen, & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv preprint arXiv:1409.1556.

4. He, Kaiming, Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision & Pattern Recognition (CVPR).

5. Setio, Arnaud A. A., Traverso, A., de Bel, T., Berens, M. S. N., van den Bogaard, C., Cerello, P., Chen, H., et al. (2017). Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge. Medical Image Analysis, 42, 1–13

6. Ronneberger, Olaf, Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI).

7. Milletari, Fausto, Navab, N., & Ahmadi, S. A. (2016). V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. In 2016 Fourth International Conference on 3D Vision (3DV).

8. Zhou, Zongwei, Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. (2018). UNet++: A Nested U-Net Architecture for Medical Image Segmentation. In Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support.

9. Oktay, Ozan, Schlemper, J., Le Folgoc, L., Lee, M., Heinrich, M., et al. (2018). Attention U-Net: Learning Where to Look for the Pancreas. arXiv preprint arXiv:1804.03999.

10. Dou, Qi, Chen, H., Yu, L., Qin, J., & Heng, P. A. (2017). Multilevel contextual 3-D CNNs for false positive reduction in pulmonary nodule detection. IEEE Transactions on Biomedical Engineering, 64(7), 1558–1567.

11. Zhu, Weikang, Liu, C., Fan, W., Xie, X., & Yuille, A. L. (2018). DeepLung: Deep 3D Dual Path Nets for Automated Pulmonary Nodule Detection and Classification. In Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV).

12. Dosovitskiy, Alexey, Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., et al. (2020). An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. arXiv preprint arXiv:2010.11929.

13. Chen, Jieneng, Lu, Y., Yu, Q., Luo, X., Adeli, E., Wang, Y., et al. (2021). TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation. arXiv preprint arXiv:2102.04306.

14. Liu, Ze, Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., et al. (2021). Swin Transformer: Hierarchical -Vision Transformer using Shifted Windows. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV).

15. Cao, Hu, Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., & Wang, M. (2021). Swin-Unet: Unet-like Pure Transformer for Medical Image Segmentation. arXiv preprint arXiv:2105.05537.

16. Lorensen, William E., & Cline, H. E. (1987). Marching Cubes: A High Resolution 3D Surface Construction Algorithm. ACM SIGGRAPH Computer Graphics, 21(4), 163–169.

17. Ardila, Diego, Kiraly, A. P., Bharadwaj, S., Choi, B., Reicher, J. J., Peng, L., et al. (2019). End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nature Medicine, 25(6), 954–961.

18. Srivastava, Saurabh, Jain, S., & Gupta, A. (2021). Multimodal fusion of clinical and imaging features for lung cancer prediction using deep learning. Journal of Medical Systems, 45(9), 1–12.