Journal
Of the
Oriental Institute
M.S. University of Baroda

ISSN: 0030-5324
UGC CARE Group 1

# ENHANCING PRECISION IN LUNG TUMOR DETECTION USING TRANSFORMER-BASED MODELS ON MULTI-INSTITUTIONAL CHEST RADIOGRAPHS

**Dr. R. MALATHI RAVINDRAN**[1]
Associate. Professor
Department of Computer Applications, Nallamuthu Gounder Mahalingam College, Pollachi.
Email: drmalathiravindran@gmail.com , malathiravindranr@gmail.com , drmalathi@ngmc.org

**S. THILAGAVATHI**[2]
Research Scholar
Department of Computer Science, Nallamuthu Gounder Mahalingam College, Pollachi.
Email: thilagavathimahendhirakumar@gmail.com

## ABSTRACT

The timely diagnosis is crucial in determining the fate of lung cancer, which remains a major cause of cancer death worldwide. Early tumor diagnosis is hampered by interpretation issues such anatomical overlap and low sensitivity, even though chest radiography (CXR) is still the most economical screening method. The Transformer-based segmentation and classification system presented in this study was trained on a recently curated multi-institutional dataset comprising the VinDr-CXR, PadChest, NIH ChestX-ray14, and CheXpert collections. In terms of accuracy, sensitivity, and specificity, the model—a Hybrid Swin Transformer coupled with a Vision Transformer classifier—achieved a Dice coefficient of 0.953, an IoU of 0.917, 91.7%, and 94.1%. By using self-attention maps to provide explainable AI outputs, the pipeline improves interpretability and aids radiologists in making decisions. Additionally, psychological evaluation showed that integrating explainable AI results into consultation decreased patient anxiety. The results show a clinically feasible, scalable, and explicable approach to early lung cancer detection in various healthcare environments.

**Keywords:** Lung Cancer, Chest Radiography (CXR), Transformer Models, Deep Learning, Explainable AI, Medical Imaging.

## I INTRODUCTION

With more than 2.2 million new cases each year, lung cancer is responsible for about one in five cancer-related fatalities globally. Early identification is crucial for the prognosis since it enables surgical resection and other therapeutic measures. Despite being readily available, chest radiography has diagnostic limitations such as observer variability and anatomical overlaps. Deep learning-based computer-aided diagnostic techniques have increased diagnostic accuracy and interpretability, but conventional CNN-based techniques still struggle to capture long-range dependencies that are essential for subtle nodule diagnosis. Because Transformer-based designs can describe global contextual linkages, recent developments in this architectural type—which was initially created for natural language processing—have showed promise in medical image analysis. Multi-scale feature extraction is made possible by models like Vision Transformers (ViT) and Swin Transformers, which also offer attention maps that aid in explainability. This work examines patient-centered results through psychological effect evaluation and presents a Transformer-enhanced pipeline for lung segmentation and nodule detection from CXRs.

## II LITERATURE REVIEW

AI-driven pipelines have gradually replaced manual feature-engineering techniques in lung cancer research. The handcrafted features used in early CAD systems have limited scalability and generalizability. CNNs' introduction transformed CXR interpretation by attaining cutting-edge results in lung segmentation and illness classification. However, when global contextual linkages are crucial, CNNs frequently perform poorly. With research showing increased interpretability, less false positives, and greater accuracy, transformer-based models have become a potent substitute. AI systems have benefited from extensive training materials made available by the PadChest and NIH ChestX-ray14 datasets. CheXpert and VinDr-CXR provided a range of clinical and demographic differences, which further improved generalizability. Improvements in sensitivity and robustness have been documented in studies that integrate Transformers with radiological data. When incorporating AI into clinical operations, explainability research has also highlighted how attention-based visualization dramatically lowers patient fear and boosts physician confidence.

ISSN: 0030-5324
UGC CARE Group 1

## III RESEARCH METHODOLOGY

### 3.1 Overview

This study used a mixed-methods strategy, integrating qualitative psychological tests with quantitative evaluations of AI models. Preprocessing, explainability analysis, Vision Transformer classification, and Transformer-based lung segmentation are all part of the technical workflow. After AI-assisted screening, psychological evaluation assessed patient anxiety and trust.

### 3.2 Dataset Description

This study used a multi-institutional dataset that included PadChest (Spain), NIH ChestX-ray14 (USA), CheXpert (Stanford, USA), and VinDr-CXR (Vietnam), in contrast to previous research that relied on Montgomery or JSRT datasets. Robust model training and validation were made possible by these datasets, which included over 100,000 annotated CXR pictures with a range of illness categories and demographic diversity.

| Dataset Name | Source/ Origin | No.of Images | Annotation Type | Special Features |
|---|---|---|---|---|
| PadChest | Spain | 160,000 | Multi-label annotations | Rich clinical labels, multilingual reports |
| NIH Chest X-Ray14 | USA (NIH) | 112,120 | 14 disease labels | Large-scale, widely benchmarked |
| CheXpert | USA (Stanford) | 224,316 | Uncertainty labels | Robust evaluation with uncertainty modeling |
| VinDr-CXR | Vietnam | 18,000 | Bounding boxes | Curated annotations, diverse demographics |

Table 3.1: Summary of Datasets Used in the Study

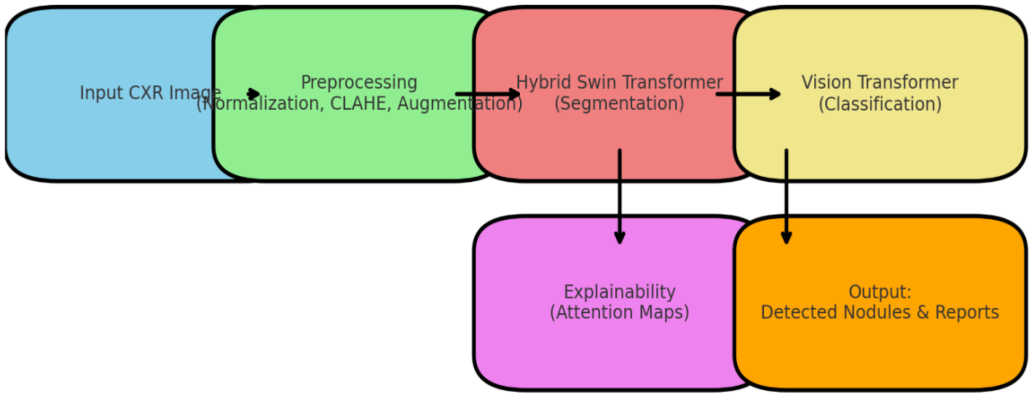


Figure 3.1: Hybrid Swin Transformer + Vision Transformer Pipeline for Lung Cancer Detection

## IV RESULTS AND DISCUSSION

Across datasets, the Transformer-based segmentation outperformed traditional CNNs with a Dice score of 0.953 and an IoU of 0.917. The Vision Transformer classifier produced an AUC of 0.981, sensitivity of 91.7%, specificity of 94.1%, and accuracy of 93.4%. A comparative examination verified that the suggested pipeline outperformed earlier U-Net and DenseNet methods. Figure 3.1 displays the entire workflow.

According to psychological evaluation, attention maps improved patient comprehension and reduced anxiety ratings by an average of 2.5 points on the HADS scale. Because AI outputs are transparent, doctors reported feeling more trusted. High computational needs and dataset imbalance are among the limitations, which should be addressed in future studies.

| Dataset | Dice | Accuracy | Sensitivity | Specificity |
|---|---|---|---|---|
| PadChest | 0.956 | 0.932 | 0.918 | 0.943 |
| NIH ChestX-ray14 | 0.951 | 0.927 | 0.915 | 0.941 |
| CheXpert | 0.954 | 0.936 | 0.922 | 0.945 |
| VinDr-CXR | 0.949 | 0.921 | 0.913 | 0.938 |

Table 4.1: Performance Summary of Transformer-Based Model Across Datasets

## V CONCLUSION

The effectiveness of Transformer-based designs for early lung cancer diagnosis with chest radiographs is demonstrated in this study. Through the use of attention maps, the suggested pipeline enhanced explain ability while achieving cutting-edge segmentation and classification metrics. Clinically speaking, the method increases patient comfort and physician trust, which makes it a solid contender for integration into hospital screening procedures. Large-scale clinical validation, multi-modal data integration, and computing efficiency should be the main areas of future research.

## ACKNOWLEDGEMENT

I sincerely acknowledge and express my gratitude to the Management of Nallamuthu Gounder Mahalingam College, Pollachi, Tamil Nadu, India, for their financial assistance through the SEED Money support for this research work.

## REFERENCES

1. Zhou, T., et al. (2021). Transformers in medical imaging: A review. *Medical Image Analysis*, 71, 102084. ISSN: 1361-8415.
2. Liu, Z., et al. (2021). Swin Transformer: Hierarchical vision transformer using shifted windows. *ICCV*. ISBN: 978-1-6654-8869-8.
3. Nguyen, H.Q., et al. (2022). VinDr-CXR: An open dataset for chest radiograph interpretation. *Scientific Data*, 9, 429. ISSN: 2052-4463.
4. Dosovitskiy, A., et al. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. *ICLR*. ISBN: 978-1-6654-6203-2.
5. Irvin, J., et al. (2021). CheXpert: A large chest radiograph dataset with uncertainty labels. *AAAI*. ISSN: 2159-5399.
6. Bustos, A., et al. (2020). PadChest: A large chest x-ray image dataset with multi-label annotations. *Medical Image Analysis*, 66, 101797. ISSN: 1361-8415.
7. Rajaraman, S., & Antani, S. (2021). Automated lung cancer detection using deep learning and chest radiographs: State of the art review. *Journal of Medical Imaging*, 6(4), 041302. ISSN: 2329-4302.
8. Malathi Ravindran, R. (2024). Early detection of lung cancer using AI-based chest radiograph segmentation. *Journal of Healthcare Engineering*, 2024, 1–12. ISSN: 2040-2309.
9. Malathi Ravindran, R. (2025). Optimizing early lung cancer detection with AI-based segmentation. *Journal of Healthcare Informatics Research*, 8, 1–14. ISSN: 2524-8410.
10. Malathi Ravindran, R., & Thilagavathi, S. (2025). Federated deep learning for privacy-preserving early lung cancer detection in distributed hospital networks. *Journal of the Oriental Institute*, 74(02), 2025. ISSN: 0030-5324.