Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 5, No.1, May : 2025

A REVIEW ON:ADVANCED SEGMENTATION OF PULMONARY CANCER IN CT IMAGES USING DEEP NEURAL NETWORK

Prof. Jeevan Mangrulkar, Assistant Professor, Department Of Information Technology, JD College Of Engineering And Management, Nagpur, Maharashtra, India
Saurabh Sarve, Student, Department Of Information Technology, JD College Of Engineering And Management, Nagpur, Maharashtra, India
Gaurav Raipurkar, Student, Department Of Information Technology, JD College Of Engineering And Management, Nagpur, Maharashtra, India
Yash Mankar, Student, Department Of Information Technology, JD College Of Engineering And Management, Nagpur, Maharashtra, India
Sahil Kamale, Student, Department Of Information Technology, JD College Of Engineering And Management, Nagpur, Maharashtra, India

ABSTRACT:
Non-small cell lung cancer (NSCLC) is among the globe's top causes of cancer-related death because of late diagnosis and the absence of efficient early detection methods. Deep learning advancements in recent times have revolutionized medical imaging with new avenues for automatic, precise, and reproducible segmentation and diagnosis of lung cancer from computed tomography (CT) and histopathological images. This review article gives a critical overview of the recent developments in deep learning techniques employed for lung cancer segmentation and classification. Besides, it visualizes the new state-of-the-art models, i.e., TransUNet, U-Net with adjustments, probabilistic 3D CNNs, and hybrid classification survival models, i.e., LCSCNet and LCSANet. Moreover, it discusses the incorporation of multi-omics data to improve diagnostic and prognostic accuracy. The paper also introduces top challenges, i.e., model explainability, transfer learning across sets, and computational limits. Future research areas, e.g., explainable AI, federated learning, and multi-modal fusion, are discussed. The paper introduces the potential revolution that deep learning can facilitate in early diagnosis and personalized treatment strategies for lung cancer therapy.
Keywords: Lung Cancer, Deep Learning, CT Imaging, Pulmonary Nodule Segmentation, U-Net, TransUNet, 3D CNN, Histopathology, Survival Prediction, Multi-Omics, Explainable AI, Precision Oncology

INTRODUCTION :
Lung cancer continues to be the leading isolated cause of cancer mortality worldwide and kills almost 1.8 million people annually. Of its various types, non-small cell lung cancer (NSCLC) constitutes nearly 85% of all cases diagnosed. It is predominantly diagnosed at its advanced stages, restricting treatment options and lowering the survival rate among patients. Diagnosis of the disease during early stages with imaging, particularly by low-dose computed tomography (CT), has been one of the prime factors in improving patient prognosis. But manual reporting of CT scans is very time-consuming, susceptible to human error, and with various radiologists [8].
Deep learning (DL), a branch of artificial intelligence, is very promising for the automated detection, segmentation, and classification of lung cancer. Deep neural networks (DNNs) can learn high-level spatial features of medical images and therefore can outperform the traditional image processing techniques both in terms of performance and speed [2][4]. All top-performing architectures among U-Net, TransUNet, and 3D CNNs were shown to be very effective for lung nodule segmentation from CT scans, and hybrid approaches have been proposed to combine classification with survival prediction [8][10].
Apart from imaging, DL in cancer goes further. Future work integrates histopathological examination and multi-omics information— gene expression, microRNA profile, and DNA methylation status —

to deliver predictive and diagnostic capacity [7][10]. Multi modal methods obey the vision of precision oncology, in which tailored treatment is administered based on a full understanding of tumor biology. Despite significant advances, applying these models to real-world clinical environments still has some practical issues. These include the simplicity of the models, usability across a broad range of institutions and computational expense. On this aside, this review strives to detail state-of-the-art deep learning methods for lung cancer segmentation and diagnosis, weigh their advantages and disadvantages, and suggest potential directions of future research and application in the clinic.

LITERATURE REVIEW:

Progress in deep learning for lung cancer diagnosis, diagnosis, i.e., segmentation, classification, and prediction, has been immense. Comparative evaluation of segmentation models—UNet, SegNet, FCN, DeepLabV3+, and TransUNet—established that TransUNet was the best performing on various CT datasets with a DICE coefficient of 0.887. It credits its hybrid architecture with efficiently employing convolutional operations and attention-based operations on transformers to achieve deep local features and long-contextual features [2].

From the research papers, one of them suggested a hybrid model that involves a U-Net variant for lung lobe segmentation and SVM classifier and AlexNet-based feature extractor. Sequential was an approach that recorded a 97.98% accuracy in classification and 97.70% F1-score and demonstrated the efficacy of segmenting and classifying as a robust system [8]. Deep architectures VGG16, ResNet50, and ConvNeXtSmall were employed together with the current image preprocessing methods including histogram equalization (HE), contrast-limited adaptive histogram equalization (CLAHE), and noise elimination in another paper. The approach was 100% accurate on four-class CT data set and pointed out the very significant role of preprocessing in maximizing the performance of deep learning approaches [4].

A CNN radiomics model was employed to distinguish different subtypes of non-small cell lung cancer (NSCLC), i.e., adenocarcinoma (ADC) and squamous cell carcinoma (SQC), from computed tomography (CT) scans. One of the AUC 0.71 models had concluded that deep learning is most probable to utilize radiologic characteristics significantly correlated with histologic classification in non-invasive diagnosis[3]. Another model used PCA-SMOTE-CNN method with multi omics data—mRNA, miRNA, and DNA methylation—and differentiated lung adenocarcinoma. Accuracy, precision, recall, and F1-score of the model were 0.97, which testifies for the potential of using genomic data and deep learning models to deliver utmost diagnostic accuracy [7].

DenseNet201, one of the first of the new methods, was combined with hand-crafted color histogram features and annotated histopathological images at 99.68% accuracy on the LC25000 dataset. Hand-crafted features and deep learning were also found to be useful while annotating cancer from histology [1]. Furthermore, a high performance state-of-the-art end-to-end probabilistic deep learning architecture was introduced with 3D CNNs for malignancy classification (CADx) and lung nodule detection (CADe) from low-dose CT scans. A Bayesian inference through Monte Carlo dropout integration enabled the system to generate well-calibrated uncertainty estimates and probability predictions. The model attained an FROC of 0.921 over the LUNA16 database and an AUROC of approximately 0.87 over the Kaggle Data Science Bowl dataset, which proclaims its clinically high worthiness [9].

Lastly, a two-pipeline model has been proposed for lung cancer histopathological diagnosis and survival rate estimation. The pipeline used LCSCNet to predict subtyping and LCSANet to predict the survival outcome prognosis of the patient. The pipeline used improved preprocessing techniques such as Macenko stain normalization, semantic segmentation, and ROI extraction using U-Net. LCSCNet used the accuracy of 96.55% on different subtypes of lung cancer, and LCSANet used the accuracy of 95.85% to predict survival outcome prognosis of the patient. The integrated framework enables the ability to facilitate prognostic and diagnostic decision-making in precision oncology scenarios [10].

PROPOSED METHDOLOGY:

Deep learning methods in lung cancer diagnosis are all based on formalized steps across each step of data preprocessing, model construction, training protocols, and testing. Below is a collection of significant methodologies used in research studies.

1. Data Preprocessing

Preprocessing is the process by which model performance and data quality are improved.Roi cropping is one of the methods utilized in CT imaging-based pipelines, which was proved to improve feature segmentation and localization performance.For instance, [2] proved the way in which ROI cropping elimination of unnecessary lung tissue improved segmentation DICE scores significantly.

In histopathology, color changes are normally handled by Macenko stain normalization, in a multi-model survival prediction model [10].Other techniques such as Contrast Limited Adaptive Histogram Equalization (CLAHE) and histogram equalization have been employed effectively with a goal of maximizing nodule detectability in CT scans [4].

2. Model Architectures

Multiple different architectures have been used with deep learning, including:

• TransUNet: The model combines the ability of Convolutional Neural Networks (CNNs) to do spatial coding and transformer global attention ability. [2] achieved a DICE value of 0.887 in using TransUNet in LIDC IDRI and other datasets.

• AlexNet-SVM with modified U Net: It has three phases with U-Net for segmentation, AlexNet for feature extraction, and SVM for classification and accuracy-based classification up to the level of 97.98% [8].

• Monte Carlo Dropout 3D CNN: This uses Bayesian on 3D CNNs and dropout Bayesian for representing uncertainty, achieving to the value of 0.921 on LUNA16 in FROC and with an AUROC of approximately 0.87 on Kaggle data [9].

•LCSCNet and LCSANet: These are specifically trained models for subtype detection and survival rate prediction from histopathological images with a classification accuracy of 96.55% and survival prediction accuracy of 95.85% [10].

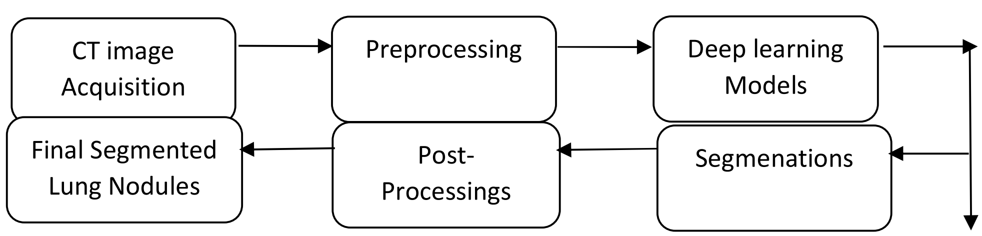
• PCA-SMOTE-CNN: PCA was applied for dimensionality reduction and SMOTE for class balancing in this model and it classified lung adenocarcinoma from omics data with 97% accuracy [7].

3. Evaluation Metrics

For segmentation, DICE coefficient and Intersection over Union (IoU) are general quality measures of segmentation [2].

For classification, Recall, Precision, Accuracy, F1-score are general evaluation metrics used (Shatnawi et al., 2025)(Naseer et al., 2023).

For prognosis models, AUROC metric was used for the assessment [9], whereas accuracy in survival prediction was used as a measure in [10]



COMPARATIVE ANALYSIS:

4.1. Performance and Objectives of the Models

• Segmentation Models :

TransUNet was the best one at segmentation of CT data and achieved the DICE score of 0.887 on LIDC-IDRI with ROI cropping preprocessing [2].Its transformer module enhances the long-range context awareness ability, which is vital in the step of distinguishing between lung nodules and tissues. The Modified U-Net with AlexNet-SVM achieved combined lobe segmentation, nodule detection, and classification on an extremely high accuracy rate of 97.98% across the LUAN16 dataset. This is an indicator that an extremely high accuracy rate can be achieved using the segmentation techniques in combination with traditional classifiers [8].

• Classification Models:

Enhanced CNNs attained 100% accuracy for 4-class CT dataset, where preprocessing techniques such as histogram equalization and noise removal were the benefit [4]. PCA-SMOTE-CNN model on omics data attained 97% accuracy, which ensured to include genomic features along with deep learning models [7].

• Probabilistic Frameworks :

The 3D Probabilistic Deep Learning System defined FROC as 0.921 and AUROC of approximately 0.87, demonstrating its effectiveness in malignancy classification and risk decision-making from low-dose CT scans [9].

• Histopathology-Based Survival Models: LCSCNet and LCSANet model reached 96.55% in lung cancer subtypes classification and in predicting survival rate with 95.85% accuracy, which shows the clinical importance of using diagnostic and prognostic features [10].

4.2. Insights

•Best Segmentation Model: TransUNet [2] is the optimum model to employ for task-specific segmentation tasks because it involves a hybrid model which uses transformers and convolutional networks.

•Best for Integrated Diagnosis and Prognosis: LCSCNet + LCSANet [10] does the classification and survival prediction best, most important for decision-making based on individuality.

•Simplest to use in clinics: Ozdemir et al.'s [9] Bayesian 3D CNN engineered model is the pioneer in emphasizing calibrated uncertainties, simplest to make decisions off when used in real-time settings.

CHALLENGES:

While there are optimistic developments, on the implementation side of deep learning for lung cancer diagnosis, there are also plenty of challenges.

5.1.Data Scarsity and Imbalance

Labeled lung cancer CT images and histopathological slide data are limited (and occasionally imbalanced). Mohamed and Ezugwu [7] addressed this using SMOTE, but most models still remain vulnerable to overfitting and poor generalization.

5.2. Model Interpretability

State-of-the-art current deep learning models, other than TransUNet or 3D CNNs, are black boxes that restrict clinical interpretability and use [9]. Nanosimulation models are required to facilitate risk-based decision making in oncology where such models need to simulate not only the efficacy of treatment but also other secondary treatment expenses such as toxicity.

5.3.Computational Requirements

Training and prediction using high-resolution 3D images and large genomic data sets are both time-consuming processes computationally. More time is required for processing using probabilistic methods like Monte Carlo dropout [9].

5.4 .Cross-Domain Generalization

Most of the models are learned and validated on particular datasets. Variances in imaging acquisition protocols (e.g., scanners or histologic stains) will dictate their performance when deployed at different institutions [10].

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 5, No.1, May : 2025

FUTURE DIRECTIONS :

Greater clinical relevance as well as usability can be facilitated by some directions in which some areas have to be explored as future work.

6.1. Explainable AI (XAI)

Adding attention maps, class activation maps (CAM), or SHAP values would make the models more transparent with regard to decision-making processes. This is necessary to win over clinicians and to achieve practical implementation in real-world applications [9].

6.2. Multi-Modal Data Fusion

Merging CT imaging with histopathological slides, genomic data, and clinical histories may give a better understanding of tumor behavior. Techniques such as PCA-SMOTE-CNN [7] are already on this path.

6.3. Federated Learning

Federated learning solves the concern of privacy and enhances model generalizability by enabling several institutions to collaboratively train a model without gaining access to patient information. The mechanism helps in preventing the problem incurred due to heterogeneity between domains [10].

6.4. Lightweight and Real-Time Inference

Model enhancement towards use in real-time or low-resource settings, e.g., on mobile device-based CT viewers or edge devices, can enhance access among rural or underserved healthcare providers [8].

CONCLUSION:

The application of deep learning for diagnosing and segmenting lung cancer is a major advancement in imaging and medical oncology. This paper has discussed some of the deep neural network architectures ranging from simple U-Net variants to advanced models like TransUNet [2], 3D probabilistic CNNs [9], and multi-modal approaches for survival prediction and classification [10].

CT-image sensitive segmentation models like TransUNet and Modified U-Net have recorded remarkable detection accuracy for lung nodules with DICE scores reaching clinical standards [2][8]. Classification networks, especially ensemble CNNs and hybrid models, have recorded near error-free performance in detecting different subtypes of lung cancer [4]. Outside imaging, the application of multi-omics data and histopathological images using deep learning has enabled highly individualized and data-driven diagnostic and prognostic approaches [7][10].

All these achievements notwithstanding, there are several issues that bar simple clinical release of such AI systems. They are issues of interpretability, generalization between domains, and computational constraints, especially in real-time or in low resource environments [9][10]. Moreover, lack of standardized data sets as well as institution-to-institution model validation processes is a threat to trustworthiness as well as replicability.

In the times ahead, there is a need for research to pay greater attention to explainable AI, federated learning, and multi-modal use of data to achieve the highest levels of reliability and personalization. Light, real-time models that can be used in clinical environments also must be created so that the distance between laboratory innovation and clinical utility can be covered.

In summary, deep learning has emerged as a game-changing technology in pulmonary oncology that provides machine learning-based, accurate, and scalable solutions to lung cancer early detection, subtype diagnosis, and survival prediction. Through ongoing multidisciplinary research by AI researchers, clinicians, and biomedical researchers, these technologies have the potential to significantly enhance diagnosis and patient outcomes.

REFERENCES:

[1] Noaman N. F., et al. (2024), "Advancing Oncology Diagnostics: AI-Enabled Early Detection of Lung Cancer Through Hybrid Histological Image Analysis", *IEEE Access*, 2024. https://doi.org/10.1109/ACCESS.2024.3397040

[2] Chen W., Wang Y., Tian D. & Yao Y. (2023), "CT Lung Nodule Segmentation: A Comparative Study of Data Preprocessing and Deep Learning Models", *IEEE Access*, Vol. 11, pp. 34925–34937, 2023. https://doi.org/10.1109/ACCESS.2023.3265170

[3] Chaunzwa T. L., Hosny A., Xu Y., Shafer A., Diao N., Lanuti M. & Aerts H. J. (2021), "Deep learning classification of lung cancer histology using CT images", *Scientific Reports*, Vol. 11(1), pp. 5471, 2021. https://doi.org/10.1038/s41598-021-84630-x

[4] Shatnawi M. Q., Abuein Q. & Al-Quraan R. (2025), "Deep learning-based approach to diagnose lung cancer using CT-scan images", *Intelligence-Based Medicine*, Vol. 11, Article ID: 100188, 2025. https://doi.org/10.1016/j.ibmed.2024.100188

[5] Yu H., Zhou Z. & Wang Q. (2020), "Deep Learning Assisted Predict of Lung Cancer on Computed Tomography Images Using the Adaptive Hierarchical Heuristic Mathematical Model", *IEEE Access*, Vol. 8, pp. 86396–86409, 2020. https://doi.org/10.1109/ACCESS.2020.2992645

[6] Zhang L., Zhang J., Tan T., Teng X., Sun X., Zhao H. & Litjens G. (2021), "Deep Learning Methods for Lung Cancer Segmentation in Whole-Slide Histopathology Images—The CDC@LungHP Challenge 2019", *IEEE Journal of Biomedical and Health Informatics*, Vol. 25(2), pp. 429–440, 2021. https://doi.org/10.1109/JBHI.2020.3039741

[7] Mohamed T. I. A. & Ezugwu A. E. (2024), "Enhancing Lung Cancer Classification and Prediction With Deep Learning and Multi-Omics Data", *IEEE Access*, Vol. 12, pp. 59880–59892, 2024. https://doi.org/10.1109/ACCESS.2024.3394030

[8] Naseer I., Akram S., Masood T., Rashid M. & Jaffar A. (2023), "Lung Cancer Classification Using Modified U-Net Based Lobe Segmentation and Nodule Detection", *IEEE Access*, Vol. 11, pp. 60279–60290, 2023. https://doi.org/10.1109/ACCESS.2023.3285821

[9] Ozdemir O., Russell R. L. & Berlin A. A. (2020), "A 3D Probabilistic Deep Learning System for Detection and Diagnosis of Lung Cancer Using Low-Dose CT Scans", *IEEE Transactions on Medical Imaging*, Vol. 39(5), pp. 1417–1426, 2020. https://doi.org/10.1109/TMI.2019.2947595

[10] Aharonu M. & Ramasamy L. (2024), "A Multi-Model Deep Learning Framework and Algorithms for Survival Rate Prediction of Lung Cancer Subtypes With Region of Interest Using Histopathology Imagery", *IEEE Access*, Vol. 12, 2024. https://doi.org/10.1109/ACCESS.2024.3484495

[11] Khan M. A., Tariq U., Hussain N. & Rehman A. (2025), "Explainable AI for Lung Cancer Detection via a Custom CNN on CT Scans", *Scientific Reports*, Vol. 15(1), pp. 1459, 2025. https://www.nature.com/articles/s41598-025-97645-5

[12] Alam M. M., Sharif M. I. & Altameem A. (2024), "Multi-Objective Deep Learning for Lung Cancer Detection in CT Images", *Diagnostics*, Vol. 14(1), pp. 63, 2024. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC12000487/

[13] Zhu X., Lin Y., Li W. & Wang J. (2025), "LungPath: AI-Driven Histologic Pattern Recognition in Early-Stage Lung Adenocarcinoma", *Journal of Thoracic Oncology*, Vol. 20(4), pp. 750–760, 2025. https://pubmed.ncbi.nlm.nih.gov/39263012/

[14] Sharma P., Gupta D. & Singh S. (2025), "Deep Learning Innovations in the Detection of Lung Cancer: A Comprehensive Review", *Cognitive Computation*, Vol. 17, pp. 100–115, 2025. https://link.springer.com/article/10.1007/s12559-025-10408-2

[15] Wei J., Wang L., Chen Y. & Huang R. (2024), "A Benchmark of Deep Learning Approaches to Predict Lung Cancer Risk Using CT Imaging", *Scientific Reports*, Vol. 14(1), pp. 5390, 2024. https://www.nature.com/articles/s41598-024-84193-7