Lung Cancer Segmentation Using an Enhanced TransUNet++ Architecture

Anoop V.[1*], Karthik G. M.[2], Dr Savitha S.[3], Anil Kumar Muthevi[4], Saravanan Ramamoorthi Agilesh[5], and Ramu B.[6]

[1] Department of Artificial Intelligence and Data Science, Jyothi Engineering College-Thrissur, Jyothi Hills, Panjal Road, Vettikattiri PO, Cheruthuruthy, Thrissur, Kerala 679531, India

[2] School of Computer Science and Engineering (SCOPE), Vellore Institute of Technology, Katpadi Thiruvalam Road, Vellore 632014, Tamil Nadu, India

[3] Department of Computer Science and Engineering, BMS Institute of Technology and Management, Doddaballapur Main Road, Yelahanka, Bengaluru, Karnataka 560064, India

[4] Department of Computer Science and Engineering, Aditya University, Surampalem, Kakinada, Andhra Pradesh 533437, India

[5] Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation, Green Fields, Vaddeswaram, Guntur District, Andhra Pradesh 522502, India

[6] Department of Electronics and Communication Engineering, Geethanjali College of Engineering and Technology, Hyderabad, Telangana 501301, India

*e-mail: vanoopainds@gmail.com

Abstract. Lung cancer is a life-threatening disease in which accurate staging of malignant nodules using computed tomography (CT) scans is critical for reducing mortality. Most existing approaches rely solely on deep learning models. This work proposes an accurate and computationally efficient hybrid deep learning framework for lung cancer analysis, integrating advanced preprocessing, feature extraction, and hybrid network architectures. The pipeline begins with preprocessing steps including resizing, normalization, edge detection, and median filtering to enhance image quality. Texture features are extracted using local binary patterns (LBP), while principal component analysis (PCA) is applied for dimensionality reduction. The optimized features are classified using an EfficientNet-B0 model. For precise segmentation, EfficientNet-B0 is embedded within a Transformer-based UNET++ (TransUNET++), enabling effective modeling of both local details and global contextual dependencies. Evaluated on a benchmark CT dataset, the proposed method achieved 98.58% accuracy, 98.47% sensitivity, 99.23% specificity, and 98.42% precision for classification, along with strong segmentation performance (99.53% Dice similarity coefficient, 98.56% Intersection over Union, 99.73% Hausdorff distance, 98.86% volumetric overlap error), demonstrating high spatial agreement with ground-truth masks.

Keywords: lung cancer; preprocessing; local binary patterns; principal component analysis; EfficientNet-B0; TransUNET++.

Paper #9323 received 14 Aug 2025; revised manuscript received 9 Jan 2026; accepted for publication 18 Jan 2026; published online 23 Mar 2026. doi: 10.18287/JBPE26.12.010305.

1 Introduction

Lung cancer is one of the deadliest malignancies worldwide, with over 2 million cases and 1.76 million deaths reported in 2018. Early and accurate diagnosis is essential for improving survival rates. While chest radiography offers limited diagnostic benefit, CT imaging provides superior sensitivity and spatial resolution and is the preferred modality for lung cancer detection [1]. However, manual interpretation of CT scans is time-consuming and prone to observer variability, motivating the need for automated diagnostic systems. Traditional CAD systems using ML classifiers such as support vector machine (SVM), *k*-Nearest Neighbors (KNN) and Artificial Neural Network (ANN) rely on handcrafted features, which limits robustness and generalization under varying imaging conditions [2, 3]. Recent advances in deep learning, particularly Convolutional Neural Networks (CNNs) such as VGGNet, ResNet, and DenseNet, have enabled end-to-end feature learning and improved performance [4], but they primarily capture local features and struggle with global contextual dependencies. Transformer-based models and hybrid CNN-Transformer architectures address this limitation by integrating attention mechanisms, though challenges such as computational complexity and inconsistent CT quality remain [5].

2 Literature Survey

The existing literature largely reports individual methodological advances but often lacks a critical synthesis that clearly motivates subsequent research. For instance, Akitoshi et al. proposed a deep learning–based segmentation model for lung cancer detection in chest X-rays, achieving a mean Dice coefficient of 0.52 on malignant lesions; while this demonstrates feasibility, the moderate overlap score and limited dataset size restrict its robustness and clinical generalizability [6]. Mamtha et al. introduced a hybrid framework combining deformable models, Bayesian fuzzy clustering, and the water cycle sea lion optimization algorithm with a Shepard CNN, achieving high diagnostic accuracy; however, the reliance on multiple complex optimization stages significantly increased computational cost, limiting real-time or clinical deployment [7]. Ajni and Anitha developed a multi-object optimized attention network using shuffled shepherd and social ski-driver optimization with deep Renyi entropy fuzzy clustering for lung lobe segmentation, but the use of multiple heuristic optimizers adds algorithmic complexity and reduces interpretability and scalability [8]. Iftikhar et al. presented a three-stage pipeline integrating modified U-Net-based segmentation with a hybrid AlexNet-SVM classifier, yielding strong performance on the LUNA16 dataset; nonetheless, the separation of segmentation and classification required manual post-processing and failed to exploit end-to-end learning advantages [9]. Similarly, Imran et al. combined deep feature extraction with SVMs for early lung cancer detection on the LIDC/IDRI dataset, but the dependence on handcrafted pipeline stages limited adaptability across datasets [10]. Collectively, these studies highlight improvements in segmentation accuracy but reveal persistent challenges related to computational complexity, lack of end-to-end optimization, and limited clinical scalability, thereby motivating the need for a unified, efficient deep learning framework. Overall, existing approaches suffer from poor generalization, insufficient global context modeling, or excessive complexity. To address these gaps, this study proposes a unified hybrid deep learning framework that jointly performs lung cancer classification and segmentation on CT images. The pipeline incorporates preprocessing, LBP and PCA-based feature optimization, EfficientNet-B0 for efficient classification, and a TransUNET++ for attention-driven segmentation.

3 Proposed Methods

The proposed method introduces a comprehensive and robust pipeline for accurate medical image classification and segmentation by integrating advanced preprocessing, feature extraction, and deep learning models. Initially, medical images retrieved from the database (DB) undergo preprocessing, including resizing, normalization, edge detection, and median filtering, to ensure data uniformity, reduce noise, and enhance salient structures. Subsequently, texture features are extracted using LBP, and dimensionality reduction is performed using PCA to preserve discriminative information. The refined features are used to train an EfficientNet-B0 model, chosen for its optimal balance between accuracy and computational efficiency, enabling reliable image classification. For enhanced segmentation and further refinement, the learned features from EfficientNet-B0 are integrated into a TransUNET++, which combines convolutional and self-attention mechanisms to capture both local details and long-range contextual dependencies. This dual-stage framework ensures accurate classification and precise region-level segmentation. Fig. 1 illustrates the architecture of the proposed method.

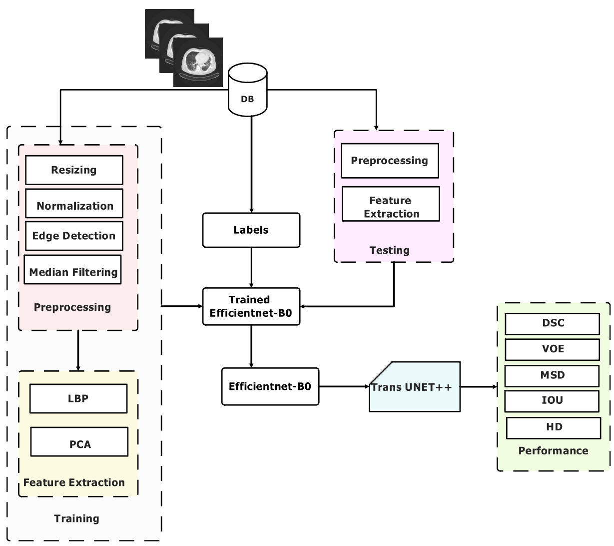

Fig. 1 Block diagram of proposed method.

Table 1 Parameter Specification of EfficientNet-B0.

Component	Specification
Number of layers	82
Scaling coefficient	ϕ=0 (baseline)
Activation function	Swish
Optimization	Adam
Building blocks	MBConv + squeeze-and-excitation (SE)
Output classes	3 (normal, benign, malignant)

3.1 Preprocessing

In this work, a comprehensive preprocessing pipeline is employed to enhance the quality of lung CT images for reliable classification and segmentation. All images are first resized to 224 × 224 pixels to ensure input uniformity and compatibility with EfficientNet-B0 and TransUNet++. Pixel intensities are then normalized to the range [0, 1] to stabilize training and improve convergence. Edge information is enhanced using the Canny operator, which highlights anatomical boundaries, nodules, and tumor regions. To suppress noise introduced during edge detection while preserving critical structures, median filtering is applied, effectively removing salt-and-pepper noise without blurring essential anatomical details.

3.2 Efficientnet-B0

EfficientNet-B0 is a convolutional neural network designed to achieve state-of-the-art image classification performance with minimal computational cost [11–13]. In this study, EfficientNet-B0 is selected as the classification backbone due to its compound scaling of network depth, width, and input resolution, making it well suited for relatively small medical imaging datasets. The model uses an input image size of 224 × 224 × 3, providing a balance between computational efficiency and feature resolution. The architecture consists of MBConv blocks with integrated squeeze-and-excitation (SE) modules to enhance feature representation. Table 1 summarizes the key parameter settings of the EfficientNet-B0 model used for lung cancer classification.

3.3 Local Binary Pattern

The LBP is a widely used texture descriptor in image processing and computer vision, known for its simplicity and effectiveness in capturing local texture features [14–15]. For a given pixel, its intensity value is compared with those of its P surrounding neighbors positioned on a circle of radius R. The mathematical expression for the LBP code at a pixel location (x_c, y_c) is given by the following Eq.:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) \cdot 2^p \, , \tag{1}$$

where g_c is the gray value of the central pixel, g_p is the gray value of the p^{th} neighboring pixel, and $s(x)$ is a thresholding function defined as:

$$s(x) = \begin{cases} 1 & if\ x \geq 0 \\ 0 & if\ x < 0 \end{cases} . \tag{2}$$

3.4 Principal Component Analysis

The PCA is used in the current research as a dimensionality reduction method in order to project the extracted features into a lower dimensional space with the highest level of significant variance [16−18]. Once the texture features have been extracted with the help of LBP, PCA removes redundancy and enhances the computational efficiency, making the EfficientNet-B0 classifier process the most informative features.

3.5 Enhanced TransUNet++

The suggested segmentation model is based on Enhanced TransUNet++, which is a hybrid architecture that combines CNN-based feature extraction and Transformer-based global context modeling [19−20]. EfficientNet-B0 backbone is used as the encoder to extract rich feature maps in an efficient manner and there are dense skip connections between UNet++ to enhance flow of gradients and reuse of features. The parameters chosen to balance between the segmentation accuracy and the computational feasibility are summarized in Table 2.

Table 2 Parameter specification of TransUNET++.

Component	Specification
Backbone	Vision Transformer (ViT-B/16) pre-trained on ImageNet21k
Patch size	16 × 16
Number of Patches	196
Embedding dimension	768
Number of attention heads	12
Dropout rate	0.1
Skip connections	Multi-scale from UNet++ encoder blocks
Decoder architecture	UNet++ with hierarchical feature fusion
Loss function	Dice loss + cross-entropy loss
Optimizer	AdamW

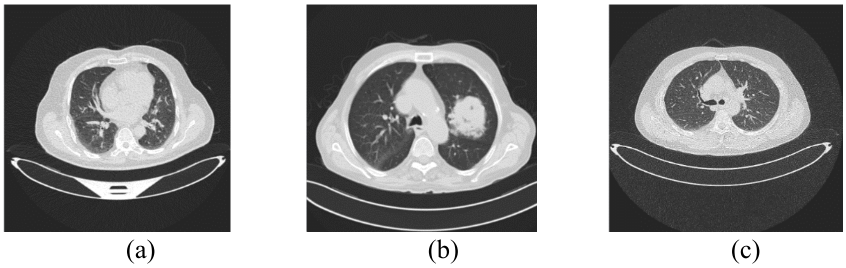

Fig. 2 Sample images from the dataset: (a) benign, (b) malignant, (c) normal.

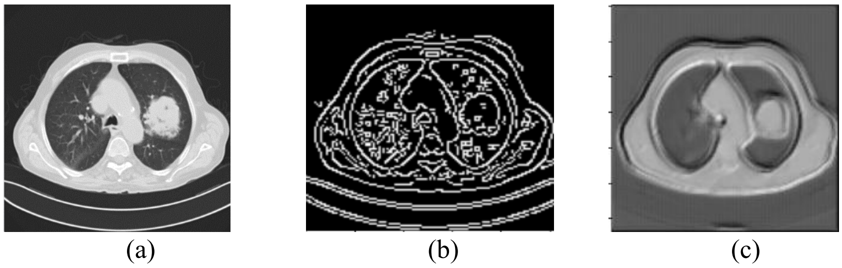

Fig. 3 Sample segmented image: (a) original image, (b) edgemap, (c) predicted mask.

4 Results and Discussions

The proposed framework was implemented in Python on the Google Colab platform using an NVIDIA Tesla T4 GPU (16 GB VRAM), an Intel Xeon CPU, and 13 GB RAM. The software environment included Python 3.10.11, TensorFlow 2.7.0, Keras 2.6.0, and supporting libraries such as OpenCV, Scikit-learn, and Matplotlib.

4.1 Dataset

The IQ-OTH/NCCD lung cancer dataset was collected at the Iraq Oncology Teaching Hospital/National Center for Cancer Diseases during the fall of 2019. It consists of CT scan slices from 110 subjects, including cancer patients at different stages and healthy individuals, totaling 1,190 images [21]. The dataset is categorized into three classes: normal (55 cases), benign (15 cases), and malignant (40 cases). All scans were acquired in DICOM format using a Siemens SOMATOM CT scanner with a protocol of 120 kV, 1 mm slice thickness, window width of 350–1200 HU, window center of 50–600 HU, and breath-hold at full inspiration. The dataset was split into training, validation, and testing sets in a 70:15:15 ratio, ensuring patient-level separation to avoid data leakage. Fig. 2 presents sample images from the dataset.

4.2 Performance Metrics

The proposed hybrid EfficientNet-B0-TransUNet++ model was tested on the basis of common metrics of classification and segmentation. To be classified, the metrics used to determine the capability of the model to identify the benign, malignant, and normal cases correctly were accuracy [22], sensitivity, specificity, precision, and F_1-score. To perform segmentation, the measures of Dice similarity coefficient (DSC), Intersection over Union (IoU) [23], Hausdorff distance (HD) [24], volume overlap error (VOE) [25], and mean surface distance (MSD) [26] were used to measure the overlap and spatial consistency between predicated masks and ground truth.

Fig. 3 illustrates a sample segmentation result obtained using UNET++. Table 3 summarizes the class-wise performance of the proposed classification and segmentation framework across the benign, malignant, and normal classes. The proposed method demonstrates consistently high classification accuracy, achieving 98.12% for benign cases and 98.00% for both malignant and normal cases. It reports the class-wise performance of the stand-alone EfficientNet-B0-based classification model, with parameters including input size 224 × 224, batch size 16, learning rate 1×10^{-4}, and training for 50 epochs using the Adam optimizer.

Table 3 Class-wise performance of the proposed EfficientNet-B0-based classification model.

Performance	Benign	Malignant	Normal
Accuracy, %	98.12	98.00	98.00
Sensitivity, %	97.80	98.20	98.10
Specificity, %	98.40	97.90	98.20
Precision, %	98.00	97.85	97.90
F_1-score, %	97.90	98.02	98.00

Table 4 Performance of classification and segmentation.

Classification		Segmentation	
Accuracy, %	98.58	DSC	99.53
Sensitivity, %	98.47	IoU	98.56
Specificity, %	99.23	HD	99.73
Precision, %	98.42	VOE	98.86
F_1-score, %	99.01	MSD	99.60

Table 5 Comparative performance of proposed method.

Methods	Accuracy	Sensitivity	Speciifcty	Precision
U-Net [1]	96.82	96.33	96.91	97.35
U-Net [2]	97.85	96.69	97.24	97.96
CNN [3]	96.00	96.00	97.00	97.00
VGG16 [21]	96.00	96.00	98.00	93.00
This work	98.58	98.47	99.23	98.42

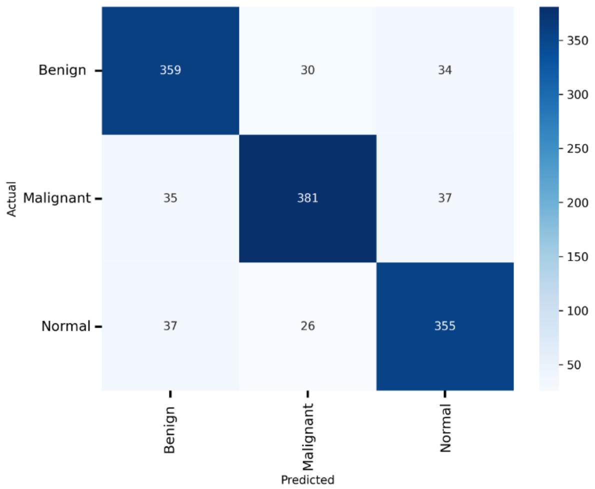

Fig. 4 Confusion matrix of proposed method.

Table 4 summarizes the performance of the proposed method for classification and segmentation. The model achieved high classification accuracy (98.58%), sensitivity (98.47%), specificity (99.23%), precision (98.42%), and F_1-score (99.01%). For segmentation, it obtained a Dice score of 99.53% and an IoU of 98.56%, demonstrating strong agreement with the ground truth.

Table 5 demonstrates comparative performance of proposed method. Fig. 4 illustrates the confusion matrix for the proposed classification method.

5 Conclusion

This paper proposes a robust hybrid deep learning framework for lung cancer classification and segmentation. The pipeline begins with image preprocessing, including resizing, intensity normalization, and noise removal using edge detection and median filtering. Texture features are extracted using LBP, while PCA reduces dimensionality and retains discriminative information. The optimized features are classified using the lightweight EfficientNet-B0 model, achieving a favorable accuracy–efficiency trade-off. For segmentation, a TransUNET++ integrates convolutional feature extraction with self-attention to capture both local details and global contextual dependencies, enabling accurate delineation of complex anatomical structures and pathological regions.

Disclosures

The authors declare that they have no conflicts of interest or competing financial interests related to this work.

References

1. A. Gupta, A. Kumar, and K. Rautela, "UDCT: lung cancer detection and classification using U-Net and DARTS for medical CT images," Multimedia Tools and Applications 84, 19065–19085 (2025).
2. R. Deepika, P. Shanmugam, K. Moorthi, P. M. Kumar, and S. L. Swarna, "Optimized transfer learning model for lung cancer stage classification using computed tomography images," Proceedings of International Conference on IoT Based Control Networks and Intelligent Systems (ICICNIS), IEEE, 912–917 (2024).
3. M. R. Islam, M. N. H. Sarker, M. K. E. S. Miah, and M. H. Hossain, "Enhancing lung cancer screening with an automated deep learning system: a resource-efficient approach," International Journal of Imaging Systems and Technology 35(6), e70225 (2025).
4. S. Nageswaran, G. Arunkumar, A. K. Bisht, S. Mewada, J. S. Kumar, M. Jawarneh, and E. Asenso, "Lung cancer classification and prediction using machine learning and image processing," BioMed Research International 2022, 1755460 (2022).
5. B. Dunn, M. Pierobon, and Q. Wei, "Automated classification of lung cancer subtypes using deep learning and CT-scan based radiomic analysis," Bioengineering 10(6), 690 (2023).
6. A. Shimazaki, D. Ueda, A. Choppin, A. Yamamoto, T. Honjo, Y. Shimahara, and Y. Miki, "Deep learning-based algorithm for lung cancer detection on chest radiographs using the segmentation method," Scientific Reports 12(1), 727 (2022).
7. M. V. Shetty, S. Tunga, "Optimized deformable model-based segmentation and deep learning for lung cancer classification," The Journal of Medical Investigation 69(3-4), 244–255 (2022).
8. A. K. Ajai, A. Anitha, "Clustering-based lung lobe segmentation and optimization-based lung cancer classification using CT images," Biomedical Signal Processing and Control 78, 103986 (2022).

9. I. Naseer, S. Akram, T. Masood, M. Rashid, and A. Jaffar, "Lung cancer classification using modified U-Net-based lobe segmentation and nodule detection," IEEE Access 11, 60279–60291 (2023).
10. I. Shafi, S. Din, A. Khan, I. D. L. T. Díez, R. J. P. Casanova, K. T. Pifarre, and I. Ashraf, "An effective method for lung cancer diagnosis from CT scan using deep learning-based support vector network," Cancers 14(21), 5457 (2022).
11. B. Kalpana, S. K. Kannaiah, "Enhancing lung cancer prediction with EfficientNet and stochastic optimization," Proceedings of 4th International Conference on Computer, Communication, Control & Information Technology (C3IT), IEEE (2024).
12. Y. Said, A. A. Alsheikhy, T. Shawly, and H. Lahza, "Medical image segmentation for lung cancer diagnosis based on deep learning architectures," Diagnostics 13(3), 546 (2023).
13. A. A. Nafea, M. S. Ibrahim, M. M. Shwaysh, K. Abdul-Kadhim, H. R. Almamoori, and M. M. AL-Ani, "A deep learning algorithm for lung cancer detection using EfficientNet-B3," Wasit Journal of Computer and Mathematics Science 2(4), 68–76 (2023).
14. R. Sun, Y. Pang, and W. Li, "Efficient lung cancer image classification and segmentation algorithm based on an improved Swin Transformer," Electronics 12(4), 1024 (2023).
15. P. M. Shakeel, M. A. Burhanuddin, and M. I. Desa, "Automatic lung cancer detection from CT image using improved deep neural network and ensemble classifier," Neural Computing and Applications 34, 9579–9592 (2022).
16. S. Ram, W. Tang, A. J. Bell, R. Pal, C. Spencer, A. Buschhaus, C. R. Hatt, M. P. diMagliano, A. Rehemtulla, J. J. Rodríguez, S. Galban, and C. J. Galban, "Lung cancer lesion detection in histopathology images using graph-based sparse PCA network," Neoplasia 42, 100911 (2023).
17. A. Alsadoon, G. Al-Naymat, A. H. Osman, B. Alsinglawi, M. Maabreh, M. R. Islam, "DFCV: a framework for evaluating deep learning in early detection and classification of lung cancer," Multimedia Tools and Applications 82(28), 44387–44430 (2023).
18. S. Al-Ofary, H. O. Ilhan, "Classification of PCA-based reduced deep features by SVM for diagnosing lung and colon cancer," Proceedings of 5th International Congress Human-Computer Interaction, Optimization and Robotic Applications (HORA), IEEE (2023).
19. Y. He, J. Yun, J. Li, Y. Wang, and J. Xu, "Multicenter deep learning study for lung tumor subtype segmentation in CT images using Swin-UNETR, nnU-Net, and TransUNet," Journal of Radiation Research and Applied Sciences 18(3), 101721 (2025).
20. J. K. Gupta, J. Kaur, "Enhancing lung cancer diagnosis: Edge-TransUNet for segmentation and Lightweight-MLCAN for classification," IETE Journal of Research, 1–10 (2025).
21. S. Hamza-Cherif, T. Tariq, Z. A. Elaouaber, and M. Mohammed, "Explainable lung cancer classification using VGG16 and Grad-CAM," Proceedings of International Conference on Artificial Intelligence, Smart Technologies and Communications (AISTC 2025), Atlantis Press, 5–11 (2025).
22. B. He, W. Hu, K. Zhang, S. Yuan, X. Han, C. Su, J. Zhao, G. Wang, G. Wang, and L. Zhang, "Image segmentation algorithm of lung cancer based on neural network model," Expert Systems 39(3), e12822 (2022).
23. S. B. Shuvo, An automated end-to-end deep learning-based framework for lung cancer diagnosis by detecting and classifying lung nodules, arXiv preprint arXiv:2305.00046v1 (2023).
24. A. R. Bushara, "A deep learning-based lung cancer classification of CT images using augmented convolutional neural networks," ELCVIA 21(1), 130–142 (2022).
25. S. Agarwal, S. Thakur, and A. Chaudhary, "Prediction of lung cancer using machine learning techniques and their comparative analysis," Proceedings of 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), IEEE, 1–5 (2022).
26. B. Deepesh, T. Latha, "FPGA-based MRI brain tumor segmentation using modified FCM method," Circuits, Systems, and Signal Processing 44, 1–20 (2025).