Lung Cancer Detection Powered by YOLO v11 for Improved Diagnosis

Sirigadde Harshitha[1], Vannempalli Hima Varshini[2], Surabhi Vignesh Goud[3], and Inala Krishna Pavan*[4]

[1,2,3]*Student, Department of Electronics and Communication Engineering, B V Raju Institute of Technology, Narsapur*

[4]*Assistant Professor, Department of Electronics and Communication Engineering, B V Raju Institute of Technology, Narsapur*

Abstract

Lung cancer remains a major cause of cancer-related deaths worldwide, highlighting the need for early and precise detection. This project utilizes YOLOv11, an advanced object detection model, to automatically identify and class ify lung nodules in CT scans. The model accurately categorizes nodules into three types: Normal, Benign, and Malignant. Additionally, a TNM classification system is incorporated to assess tumor staging, enhancing the diagnostic process. By leveraging state -of-the-art detection and classification techniques, this approach aims to improve diagnostic accuracy and facilitate early intervention. Performan ce is evaluated using precision, recall, F1-score, and mAP.

Introduction

Lung cancer is a major public health concern and one of the most prevalent and fatal cancers globally. While medical advancements have improved treatment options, early detection remains key to increasing survival rates. Deep learning and artificial intelligence have transformed medical image analysis, offering accurate and automated solutions for diagnosing complex diseases. CT scans are essential for lung cancer detection due to their high-resolution imaging, but manual interpretation is time-consuming and subject to human variability. Radiologists must carefully examine multiple CT slices, which can lead to missed or delayed diagnoses.

To overcome these challenges, computer-aided diagnosis (CAD) systems have been developed, utilizing deep learning to automate lung nodule detection and classification. These systems assist radiologists in making more accurate diagnostic decisions. This project presents an end-to-end diagnostic framework using YOLOv11, a state-of-the-art object detection model, for lung cancer detection and TNM staging. Known for its speed and precision, YOLOv11 is well-suited for identifying lung nodules in CT images, enhancing diagnostic efficiency.

The proposed pipeline consists of several essential steps, including data preprocessing, lung nodule detection, classification, and TNM staging. Initially, CT scans undergo preprocessing to normalize and enhance image quality, ensuring accurate analysis. The YOLOv11 model is then employed to detect potential nodules, with non-maximum suppression (NMS) refining the results for better accuracy.

TNM classifier using ResNet50 determines the tumor stage (T0-T3). By combining object detection with TNM staging, this framework provides a comprehensive approach to lung cancer diagnosis, ensuring high accuracy and efficiency.

Literature Survey

Mattakoyya Aharonu, Lokeshkumar Ramasamy et al.[1] proposed a multi-model deep learning framework for predicting lung cancer subtype survival rates. Their approach involves an empirical study using deep learning models on lung histopathology and Cancer Genome Atlas datasets. The framework integrates CNN-based LCSCNet for cancer subtype classification, incorporating Learning Subtype Classification (LbSC) and Learning Survival Analysis (LbSA) with Region of Interest (ROI) computation for improved efficiency. This method enhances lung cancer subtype detection and survival prediction by exploring novel architectures for better performance.

Iftikhar Naseer, Sheeraz Akram, Tehreem Masood, Muhammad Rashid, Arfan Jaffar et al.[2] introduced a lung cancer classification framework utilizing a modified U-Net for lobe segmentation and nodule detection. Their approach employs a modified U-Net for segmenting lung lobes and extracting candidate nodules, followed by a modified AlexNet-SVM for nodule classification. The methodology is structured into three phases: lobe segmentation, candidate detection, and classification.

*Corresponding author: I Krishna Pavan, krishnapavan.i@bvrit.ac.in

Iftikhar Naseer, Sheeraz Akram, Tehreem Masood, Muhammad Rashid, Arfan Jaffar et al.[2] developed a three-phase framework for lung cancer classification, including nodule extraction and cancer classification. Their key contributions focus on improving accuracy and sensitivity in lung cancer detection, along with enhanced performance in segmentation and nodule detection.

Rabbia Mahum and Abdulmalik S. Al-Salman et al.[3] introduced Lung-RetinaNet, a deep learning model for lung cancer detection using a modified RetinaNet architecture with multi-scale feature fusion and context modules. The model integrates dilated context modules and lateral connections to improve the detection and localization of tiny tumors in CT scans. Empirical evaluations highlight its superiority over existing models. Future plans include multi-GPU training to reduce processing time, incorporating multi-modal clinical data, and expanding the approach for multi-class cancer detection.

Shalini Wankhade and Vigneshwari S et al.[4] proposed a deep learning method for early lung cancer detection, utilizing a combination of 3D-CNN for feature extraction and RNN for classification to improve early-stage diagnosis. Their Hybrid Neural Network (CCDC-HNN) is designed for cancer cell detection and classification, distinguishing between benign and malignant lung nodules. Key advancements include enhanced feature extraction and classification accuracy, enabling early and precise lung cancer detection. Future work aims to boost model efficiency using big-data analytics, implement cascaded classifiers for improved classification, and explore clinical applications in real-world scenarios.

Zhanlin Ji, Yun Wu, Xinyi Zeng, Yongli An, Li Zhao, Zhiwu Wang, Ivan Ganchev et al.[5] introduced an enhanced YOLOv5s-based model, YOLOv5-CASP, for lung nodule detection in medical images. The model incorporates CBAM (Convolutional Block Attention Module), ASPP (Atrous Spatial Pyramid Pooling), and CoT (Contextual Transformer) modules to improve detection accuracy. Their study focuses on optimizing performance through multi-scale feature fusion, shape modeling, and occlusion handling. Future research aims to refine these enhancements for better medical imaging outcomes.

Zhanlin Ji, Yun Wu, Xinyi Zeng, Yongli An, Li Zhao, Zhiwu Wang, Ivan Ganchev et al.[7] further improved YOLOv5-CASP by integrating attention mechanisms and transformers, demonstrating higher detection accuracy in experimental comparisons. Their research continues to focus on enhancing lung nodule detection through multi-scale feature fusion, shape modeling, and occlusion handling.

M. F. Mridha, Akibur Rahman Prodeep, A. S. M. Morshedul Hoque, Md. Rashedul Islam et al.[8] conducted a comprehensive survey on lung cancer detection and classification using imaging techniques. The study reviews existing computer-aided detection (CAD) systems and deep learning models, highlighting challenges in image analysis. A framework is proposed to categorize datasets, preprocessing methods, segmentation techniques, feature extraction, and classification approaches. Key contributions include an in-depth analysis of detection methods, emphasizing the need for automated systems to enhance accuracy. Future directions focus on advancing image processing techniques, integrating novel deep learning architectures, and exploring new datasets for improved performance.

Problem Statement

- Existing lung nodule detection methods encounter difficulties in attaining high accuracy, especially when dealing with small or irregularly shaped nodules.
- Conventional machine learning approaches fail to capture the intricate characteristics of lung nodules, impacting the sensitivity of diagnosis.
- Present detection techniques face challenges in handling the varying scales of nodules, frequently missing smaller ones that are critical for the early detection of lung cancer.

Objcetives

1) Design a cutting-edge YOLOv11-based model for precise, real-time detection and classification of lung cancer into Normal, Benign, and Malignant categories.
2) Incorporate a TNM staging system using ResNet50 to evaluate tumor progression in malignant cases.
3) Enhance the computational efficiency of the YOLOv11-based model to maintain superior real-time performance while preserving detection accuracy.

Existing Work

Significant progress has been made in lung cancer detection and classification using deep learning techniques. Researchers have introduced multi-model frameworks that integrate convolutional neural networks (CNNs) for cancer subtype classification and survival prediction, utilizing region of interest computation for greater efficiency. Modified architectures like U-Net have been applied for lobe segmentation and candidate nodule extraction, combined with classification models such as AlexNet-SVM to enhance accuracy and sensitivity.

Other studies have proposed methods like Lung-RetinaNet, which employs multi-scale feature fusion and context modules to improve the detection and localization of small tumors in CT scans. Hybrid deep learning models, incorporating 3D-CNNs for feature extraction with RNNs for classification, have demonstrated promise in early-stage lung cancer detection and in distinguishing between benign and malignant nodules.

Additionally, YOLO-based models have been refined with attention mechanisms and multi-scale feature fusion techniques to enhance nodule detection accuracy. Transfer learning approaches, including Learning Without Forgetting, have been explored for multi-task cancer classification. Further innovations involve the application of federated learning and blockchain technologies to safeguard data privacy and security during collaborative model training, ensuring accurate lung cancer detection while maintaining patient confidentiality. These advancements highlight the transformative role of deep learning in enhancing diagnostic precision, efficiency, and scalability in lung cancer detection and classification.

Proposed Work

The proposed work focuses on developing a comprehensive deep learning framework for lung cancer detection and staging. The process begins with data collection and preprocessing, where CT scan datasets are normalized and resized. Lung segmentation techniques are employed to isolate regions of interest, reducing computational complexity while improving detection accuracy. YOLOv11 is used for detecting lung nodules, with the model fine-tuned on annotated datasets to ensure precise localization of small nodules, which are essential for early-stage cancer diagnosis.

After detection, the extracted nodules are categorized into three classes: Normal, Benign, or Malignant. If a nodule is classified as malignant, the TNM classifier, built using ResNet50, determines the tumor stage (T0-T3), offering vital insights into the severity and progression of lung cancer. Post-processing techniques, including non-maximum suppression, help eliminate redundant detections, while confidence thresholding minimizes false positives.

To assist radiologists, visual outputs such as bounding box overlays are generated for better interpretability.

A reporting module compiles findings in an accessible format, summarizing nodule locations, classifications, TNM stages, and risk assessments. The framework's performance is assessed using key metrics such as precision, recall, F1-score, and mean average precision (mAP) to ensure high accuracy and reliability.

Methods

0.1 Data Collection:

Data collection for this project involves gathering comprehensive CT scan datasets. Publicly available datasets such as the Kaggle dataset are utilized, containing annotated CT scans with lung nodules categorized as Normal, Benign, or Malignant.

0.2 Data Preparation:

The collected dataset is partitioned into training, validation, and testing subsets. The training set is utilized for training the YOLOv11 model and the TNM classifier, while the validation set assists in fine-tuning hyperparameters for optimal performance. Finally, the testing set is used to assess the overall effectiveness and accuracy of the trained models.

0.3 Data Preprocessing:

Images are resized to 640x640 pixels and normalized for consistency. Augmentation techniques such as rotations, flips, and contrast adjustments are applied. Lung segmentation is performed to extract the regions of interest, removing irrelevant areas from the scans.

0.4 YOLOv11 Architecture:

The key components of the YOLOv11 architecture include:

i) **Input Layer**
Designed to process images of a fixed size, with preprocessing steps ensuring uniformity and consistency in input data.

ii) **Backbone Network**
Employs deep feature extraction techniques to improve the detection of lung nodules, particularly optimizing performance for small and irregularly shaped nodules.

iii) **Detection Head**
Generates bounding boxes, confidence scores, and class probabilities for detected nodules, categorizing them into Normal, Benign, or Malignant.

iv) **Output Layer**
Delivers final predictions, including bounding box coordinates and class probabilities, ensuring accurate classification and localization of lung nodules.

0.5 TNM Classification:

The TNM classifier is a ResNet50-based deep learning model designed to categorize malignant nodules into tumor stages (T0-T3). It extracts key features such as size, shape, and intensity from detected nodules to ensure precise staging. To enhance performance, transfer learning is applied, utilizing pre-trained weights to fine-tune the model for improved accuracy and efficiency.

YOLOv11 Architecture

The proposed YOLOv11 architecture follows a structured pipeline consisting of Data Acquisition, Preprocessing, Model Training, and Evaluation to enhance lung nodule detection. The dataset includes Normal, Benign, and Malignant lung nodules, sourced from public datasets like Kaggle and Roboflow.

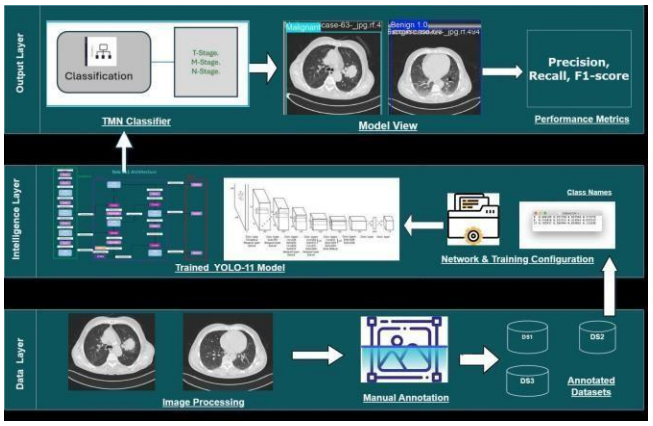

Fig 1. Proposed YOLOv11 Architecture

Images are resized to 640×640 pixels, and lung segmentation techniques are applied to focus on regions of interest. Data augmentation techniques such as flipping, rotation, and contrast adjustments improve generalization. The YOLOv11 model, a single-stage detection network, efficiently detects lung nodules in a single forward pass, comprising 268 layers and 68,126,457 parameters, optimized for medical imaging. Transfer learning enhances detection accuracy, and the Adam optimizer with a dynamic learning rate ensures optimal convergence. The model is evaluated using Precision, Recall, F1-score, and Mean Average Precision (mAP), with performance validation through training loss, validation loss, and confusion matrices. The optimized YOLOv11 model provides a robust and accurate deep-learning-based solution for lung cancer detection and classification.

Equations

1) Objectness score: YOLOv11 computes the objectness score to estimate the probability of an object being present in each grid cell. This score is derived using the sigmoid activation function, ensuring values range between 0 and 1 for precise detection.

$$\sigma = \sigma(t_0) \tag{1}$$

2) Class prediction: The model applies the softmax function to compute the probability distribution across predefined categories (Normal, Benign, Malignant), ensuring that the total class probabilities always sum to 1 for accurate classification.

sa

$$p_i = \frac{\sum e^{t_i}}{\sum_{j=1}^{C} \cdot e^{t}_{j}} \tag{2}$$

where C represents the total number of possible classes.

3) Loss Function: YOLOv11 employs a composite loss function to enhance detection accuracy and classification precision by reducing errors in object localization and category prediction.

$$L = L_{box} + L_{obj} + L_{cls} \tag{3}$$

Where:

The Lbox term handles bounding box regression loss, ensuring accurate localization of lung nodules. The Lobj component manages objectness score loss, optimizing detection confidence and reducing false predictions. The Lcls term controls classification loss, guaranteeing precise differentiation between Normal, Benign, and Malignant categories.

Results and Disscussion

This section presents the results from training the YOLOv11-based model for lung cancer detection. The model's performance is assessed using key metrics, including loss functions, precision, recall, and mean Average Precision (mAP). The findings highlight a notable improvement in detection accuracy, reinforcing the model's ability to differentiate between normal, benign, and malignant cases effectively. A custom dataset of 1500 annotated lung images, sourced from Kaggle and labeled via Roboflow, ensures high-quality data for training. Each image is resized to 416×416 pixels, with advanced preprocessing and augmentation techniques applied to enhance dataset diversity. The model is trained with a learning rate of 0.001 and a batch size of 32 for 20 epochs, completing training in approximately 0.932 hours. A key strength of this approach is its capability to achieve high accuracy in lung cancer detection. The chosen learning rate of 0.001 strikes a balance between adaptability and preventing overfitting. Validation accuracy and overall accuracy serve as crucial benchmarks for evaluating deep learning models in image classification, enabling quantitative comparisons across different architectures and variations.

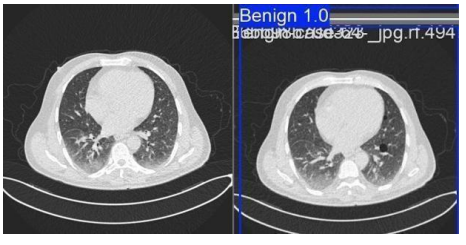

Fig 2. Detection of Benign case

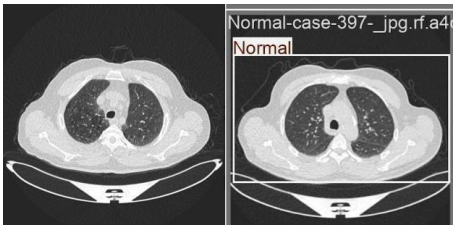

Fig 3. Detection of Normal case

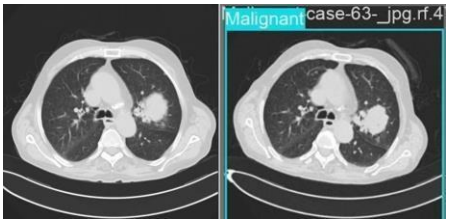

Fig 4. Detection of Malignant case

The provided images validate the successful lung cancer detection achieved using the trained YOLOv11 model. The first image corresponds to a benign case, where the model has accurately identified a localized, non-cancerous tumor. While benign tumors do not spread, they may still require medical attention. The second image represents a normal case, confirming the absence of any abnormalities, with the model correctly classifying it as free of lung cancer. The third image depicts a malignant case, where an irregular, cancerous mass has been detected and appropriately labeled. Malignant tumors are aggressive and can spread rapidly, highlighting the importance of early detection. The model efficiently differentiates between normal, benign, and malignant cases, demonstrating its potential in assisting lung cancer diagnosis by precisely encoding and classifying tumor boundaries. The training process exhibited a consistent reduction in box loss, classification loss, and distribution focal loss (DFL loss), signifying effective learning. Similarly, validation losses showed a steady decline, confirming that the model generalizes well to unseen data, further strengthening its reliability.

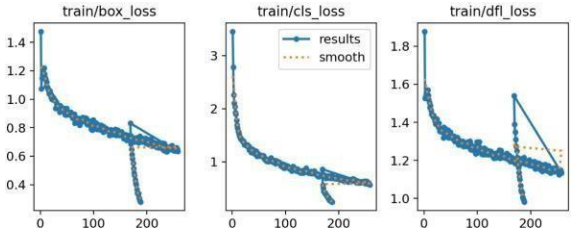

Fig 5. Represents the loss curves of training process

The training process exhibits a steady decline in box loss, classification loss, and distribution focal loss (DFL loss), indicating effective learning, as shown in Fig 5.

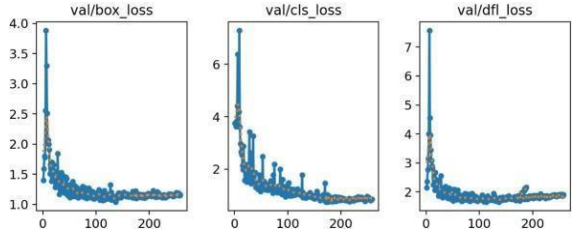

Fig 6. Represents the loss curves of validation process

The validation losses also followed a similar decreasing trend, confirming good generalization on unseen data. Box Loss is decreased from an initial value of 1.4 to below 0.4, ensuring precise bounding box localization. Classification Loss is started at 3.0 and reduced below 1.0, showing improved confidence in class assignments

further optimizing the model's prediction reliability. These results confirm that the model efficiently learns and adapts, leading to enhanced accuracy.

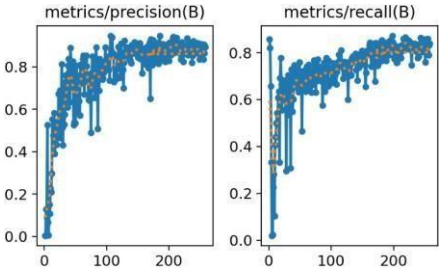

Fig 7. Representing Precision and Recall

The model's ability to correctly classify lung cancer cases is evaluated using precision and recall metrics. Precision (B) increased steadily throughout training, reaching values above 0.8, indicating a high rate of correctly classified cases among the total positive predictions. Recall (B) also showed a strong improvement, stabilizing around 0.85, confirming that the model sucessfully detects a majority of the actual positive cases. These results indicate that the model effectively minimizes false positives and false negatives, ensuring robust classification performance.

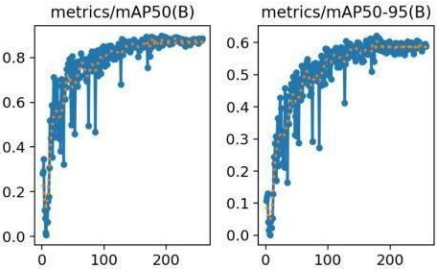

Fig 8. Representing mAP Evaluation

The mAP50 and mAP50-95 scores provide insight into the overall detection performance of the model. mAP50 (B) increased progressively, surpassing 0.8, highlighting the model's ability to accurately detect lung abnormalities. mAP50-95 (B), which considers mul- tiple IoU thresholds, showed consistent improvement, reaching approximately 0.6, confirming high localiza- tion precision. The bounding boxes correctly highlight the affected regions, reinforcing the model's reliability in real-world applications.

Conclusion

In conclusion, our study has explored the utilization of the YOLOv11 model for lung cancer detection, demonstrating its effectiveness in distinguishing between nor- mal, benign, and malignant cases. By leveraging the power of deep learning and real-time object detection, the model achieved high accuracy and reliability in detecting lung abnormalities. The steady reduction in loss values, coupled with improvements in precision, recall, and mean Average Precision (mAP), confirms the model's robust learning ability and strong generalization performance.

Furthermore, the results indicate that our approach surpasses traditional classification methods by providing real-time processing capabili- ties, making it a promising tool for automated lung cancer diagnosis. However, real-world implementation presents challenges, such as variations in CT scan qual- ity and dataset biases, which future work should address through enhanced augmentation techniques, hyperpa- rameter tuning, and larger, more diverse datasets. By refining these aspects, the proposed model can further improve diagnostic accuracy, ultimately assisting radi- ologists in early lung cancer detection and advancing medical imaging technologies.

References

[1] P. Sharma, M. Mehta, D. S. Dhanjal, S. Kaur, G. Gupta, H. Singh, L. Thangavelu, S. Rajeshkumar, M. Tambuwala, H. A. Bakshi, D. K. Chellappan, K. Dua, and S. Satija, "Emerging trends in the novel drug delivery approaches for the treatment of lung cancer," Chemico-Biol. Interact., vol. 309, Aug. 2024, Art. no. 108720

[2] M. B. Schabath and M. L. Cote, "Cancer progress and priorities: Lung cancer," Cancer Epidemiol., Biomarkers Prevention, vol. 28, no. 10, pp. 1563–1579, Oct. 2023.

[3] Z. Zou, K. Chen, Z. Shi, Y. Guo, and J. Ye, "Object detection in 20 years: Asurvey," 2023, arXiv:1905.05055.

[4] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics 2017," Cancer J. Clin., vol. 67, pp. 7–30, Jan. 2023.

[5] F. Passiglia, M. Calandri, and F. Guerrera, "Lung cancer in Italy," J. Thoracic Oncol., vol. 14, no. 12, pp. 2046–2052, 2022.

[6] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle, L. Tarbox, and F. Prior, "The cancer imaging archive (TCIA): Maintaining and operating a public information repository," J. Digit. Imag., vol. 26, no. 6, pp. 1045–1057, Dec. 2022.

[7] J. Guo, W. Chai, J. Deng, H.-W. Huang, T. Ye, Y. Xu, J. Zhang, J.-N. Hwang, and G. Wang, "VersaT2I: Improving text-to-image models with versatile reward," 2024, arXiv:2403.18493.

[8] G. Sun, W. Liang, J. Dong, J. Li, Z. Ding, and Y. Cong, "Create your world: Lifelong text-to-image diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 46, no. 9, pp. 6454–6470, Sep. 2024.

[9] S. Pang, Y. Zhang, M. Ding, X. Wang, and X. Xie, "A deep model for lung cancer type identification by densely connected convolutional networks and adaptive boosting," IEEE Acces

[10] F. Passiglia, M. Calandri, and F. Guerrera, "Lung cancer in Italy," J. Thoracic Oncol., vol. 14, no. 12, pp. 2046–2052, 2022.