FishTaxa - Journal of Fish Taxonomy
Vol 36, ISSN: 2458-942X

Lung Cancer Detection Using Transfer Learning with DenseNet121 for Improved Diagnostic Accuracy

Dr.P.Kavita[2], Dr.G Rajanikar[1], Dr.N.Rajender Reddy[3]
[2]Associate Professor, Department of Computer Science and Engineering, SVS Group of Institutions, Hanumakonda , Telangana
[1]Assistant Professor, Department of Computer Science and Engineering, SVS Group of Institutions, Hanumakonda , Telangana.
[3]Vaagdevi Engineering College, Department of Computer Science and Engineering, .Hanumakonda , Telangana.

Abstract
The improved survival rates of the patients with lung cancer highly depend on the appropriate and timely early detection of this disease and the proper treatment planning. In this paper, we are proposing a transfer learning based convolutional neural network (CNN) model using the DenseNet121 architecture to classify lung cancer images into Malignant and Normal classes with high accuracy. A respiratory lung cancer dataset was previously made publicly available, and was pre-processed into training, validation, and testing sets. The DenseNet121 model was fine-tuned by freezing the convolutional layers and adding fully connected layers together with the dropout regularization layers to avoid overfitting. Final test result on the full dataset gave me a test accuracy of 97.97% with Malignant precision = 1.00 and Normal precision = 0.95. The classification report, confusion matrix, ROC curve, and training performance curves together indicate the robustness and generalization ability of proposed approach. The achieved results indicate that the model based on DenseNet121 outperforms proposed CNN and transfer learning models as well as some traditional models, which makes the presented model efficient and reliable solution for an automated automatic lung cancer detection.

Keywords: Lung cancer detection, DenseNet121, Transfer learning, Convolutional Neural Network (CNN), Medical image classification, Deep learning, Computer-aided diagnosis (CAD), IQ-OTHNCCD dataset

Citation: Dr.P.Kavita, Dr.G Rajanikar, Dr.N.Rajender Reddy. 2025. Lung Cancer Detection Using Transfer Learning with DenseNet121 for Improved Diagnostic Accuracy. FishTaxa 36: 102-110

Introduction
Lung cancer is one of the most fatal cancer phenotypes across the world, with great annual case morbidity and death. Introduction: The success in treating lung malignancies is associated with the stage of diagnosis; hence, the early and accurate detection of lung malignancies is necessary for improving the outcome of the patient. Standard protocols for diagnosis — such as manual examination of computed tomography (CT) imaging results and tedious histopathological examination — are often time-consuming with high inter-observer variability and require significant expertise. Consequently, a lot of works have been carried out to develop automated techniques to aid radiologists and pathologists in correct and swift identification of malignant lesions.

Deep learning (DL) has recently proven to be a strong and well-promising medical image analysis tool since it can provide feature extraction and classification directly from raw images without any need for handcrafted features. In particular, convolutional neural networks (CNNs) have achieved impressive outcomes in lung cancer screening and diagnosis through CT images, histopathology slides, and other imaging modalities 1[3]. Thanoon et al. Introduction DL-based techniques have been developed for the diagnosis and detection of lung cancer using CT images for many years [1]. Similarly, Li et al. Abstract: [2] provide an overview of machine learning frameworks for lung cancer diagnosis, treatment planning, and prognosis prediction, emphasizing the need for accurate computational tools for clinical decision-making.

Recent works have addressed the performance bottleneck of classification via the advanced network architecture. Anjum et al. EfficientNet models with multiresolution [3]; Ω3efficientNet337efficientNet4multiresolution3789-3019-1-supp.pdf24 Javed et al. SCITEPRESS[4] performed a systematic review for lung cancer detection with deep learning approaches and discussed the potential and drawbacks of existing models able to be translated into clinical practice. Instead, we conduct benchmark studies, such asby Jiang etal. For large-scale datasets such as the National Lung Screening Trial cohort ([5]), different deep learning methods were evaluated with respect to their ability to predict the risk and to provide robustness for the risk stratification.

Besides classification, they have been adapted to prognostic and predictive modeling using deep learning methods. Davri et al. DL models for histo-cytological images can be potentially integrated into complete clinical workflows for diagnosis, prognosis, and prediction, as evaluated in a systematic review of the DL techniques for all these three areas [6]. Wang et al. A meta-analysis of DL applications in CT-based lung cancer segmentation conducted by [7] found that DL improved tumor delineation and reproducibility, which are key to treatment planning.

Hybrid frameworks have been introduced to cope with the challenges of limited labelled datasets and class imbalance by combining transfer learning, attention mechanisms, and ensemble strategies. For instance, Ren et al. Uddin [8] proposed a hybrid lung cancer classification framework combining different feature extraction methods, while the attention-based DenseNet models explained by Uddin [9] can be utilized for effective classification in CT and histopathological images with higher accuracy and interpretability over conventional CNNs.

Motivated by these developments, we present a transfer learning framework based on DenseNet121 for automatic classification of images of lung into normal and malignant classes. And also using ImageNet pre-trained weights, using data augmentation, dropout regularization, and stratified train/validation/test splits led to good generalization. Our proposed method performs well in terms of accuracy, precision, recall and F1-score, with indications for potential clinical implementation and for generalization to multi-class settings.

Related Work

Deep learning (DL) based methods have been widely used in lung cancer detection and classification with pure imaging modalities (CT scans, histopathology slides) and hybrid frames. Shariff et al. In [10], a differential augmentation was presented for the detection of non-small-cell lung cancer by integrating convolutional neural networks (CNN) and high sensitivity/specificity was obtained on benchmark datasets. Similarly, Shen et al. An interpretable hierarchical semantic CNN compares the accuracy of classification in lung Nodule Malignancy classification based on the explainability with the random forest–based classification Compare the Accuracy of Classification in Lung equipoise Tumor Malignance classification Between an interpretable Hierarchical Semantic CNN 11.

Cîmpeanu et al. B. [12] applied CNNs to CT scans for lung cancer diagnosis, achieving good performance on patient data. Shahadat et al. However, in [13], this approach is extended from a single cancer type to the scenario of multi-cancer, where a deep AI model for lung and colon cancer detection was developed and validated on large-scale datasets. Yang et al. Experimental Results: In [14], a six-type DL classifier was proposed and it was able to differentiate between multiple histological subtypes with histopathological whole-slide images of lung cancer and mimics with high accuracy. Marentakis et al. [15] A study that combines radiomics with deep learning for histology-based CT classification that demonstrates improved predictive performance through hybrid feature extraction. Transfer learning has also been extensively investigated. Naveed et al. Using deep feature extraction for lung cancer classification, [16] demonstrated better performance than the baseline CNN models trained on EfficientNetB3. Hamed et al. Ensemble methods can exploit complementary algorithm strengths, for instance, CNN feature extraction used with LightGBM classifiers improved lung cancer histopathology classification [17]. Saha et al. VER-Net: A Hybrid Transfer Learning Model for CT-Based Detection of Lung Cancer that Generalizes well to Different Validation Cohorts[18]

Previous studies also laid down some basic methodologies. Deep CNNs for lung cancer detection and classification was also used by Tekade and Rajeswari [19], and a survey of DL models for automated lung cancer detection using CT scans was presented by Asuntha and Srinivasan [20] along with their application. Chaturvedi et al. Ref.#: [21] Model: For comparison purposes, various traditional machine learning classifiers have been applied on lung cancer datasets. Coudray et al. CNNs have since been applied not only to lung cancer histology classification [22] but also to predict driver mutations [22] directly from histopathology images, hinting at a path towards integrated diagnosis and precision medicine.

Many research were concentrated on histology classification using CT data and model improvement. Chaunzwa et al. DL models for lung cancer histology classification using CT images were proposed in [23] with cross-dataset validation for robustness. Lakshmanaprabu et al. Lung cancer classification with optimized deep learning architectures [24] Salaken et al. In [25] the authors examined the general use of DL on datasets with low populations, and laid some important groundwork regarding the use of transfer learning and data augmentation for small datasets. Recent work by Mohamed and Ezugwu [26] improved lung cancer prediction with multi-omics deep learning analysis, showing that multimodal approaches can improve diagnostic performance.

Altogether, these studies show the evolution from simple CNN architectures to sophisticated hybrid, transfer learning, and multimodal frameworks but with a persistent emphasis on accuracy, robustness, and interpretability. Based on these works, we present a study that identifies the most salient aspects of DenseNet121 and the transfer learning approach, as well as further augmenting dropout and data augmentation strategies, to obtain high performance and clinical-grade classification of lung images into normal and malignant.

Methodology

There are plenty of pre-trained models which take less time to implement, in this paper, transfer learning was used to build a strong lung cancer detection system using DenseNet121.

A. Dataset Description

The dataset used in this study is called IQ-OTHNCCD lung cancer dataset, which contains chest X-ray images that are grouped into either Malignant or Normal class. The dataset was extracted first, but some restructuring was required in order for each class to have the same folder structure. The dataset was divided into three subsets, 70% as training, 15% for validation and 15% for testing in order to assess the performance of the models. Table 1 summarizes the number of images per class and per split.

Table 1: Dataset Description

Class	Total Images	Training	Validation	Testing
Malignant	561	392	84	85
Normal	416	291	62	63
Total	977	683	146	148

B. Data Preprocessing and Augmentation

Images were resized to 224 × 224 pixels and scaled from 0 to 1. In order to facilitate the model generalization model another form of training set was created called data augmentation: random rotations, width/height, zooming and horizontal image shifting were performed across the dataset. It was only rescaled the validation and test sets.

C. Model Architecture

The architecture of the proposed model is DenseNet121 with ImageNet pre-trained weights as base. The convolutional layers were frozen to preserve learned features and a custom classifier head was added consisting of:

- Reduction Dimension layer (Global Average Pooling),
- Overfitting prevention dropout layers (rate = 0.3)
- A dense layer with 128 neurons with ReLU activation, followed by
- Softmax activations of the output layer of two-class classification for Malignant and Normal.

We trained the model as follows: Adam Optimizer (1e-4); categorical cross entropy loss, and accuracy used as the metric. To avoid overfitting, the early stopping mechanism was implemented here with a patience of 5 epochs.

D. Proposed Methodology

Figure 1 shows the workflow, in which the chest X-ray images are first pre-processed and augmented, and then transfer learning is applied using DenseNet121. The DenseNet121 convolutional base is frozen and feature maps are extracted and reduced to the reference feature dimensions with global average pooling. The classification on Malignant or Normal is done by the fully connected layers with dropout feature. Finally, we evaluate the model using accuracy, precision, recall, F1-score, confusion matrix, and ROC curves on a separate test set.

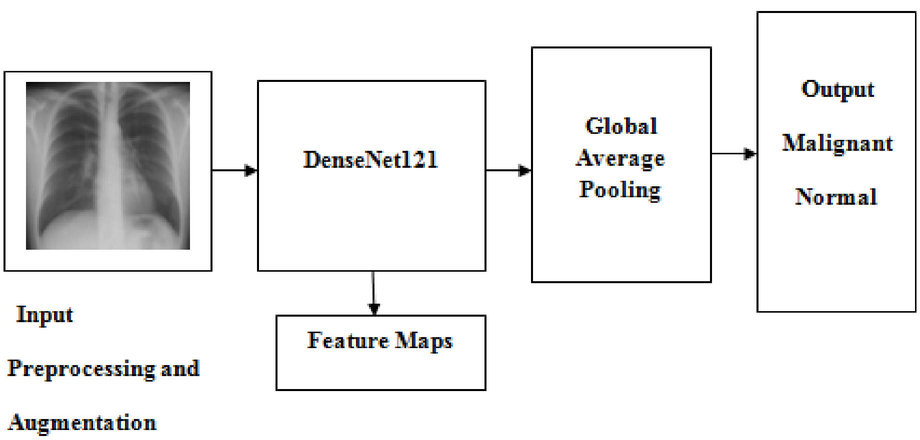

Figure 1: proposed methodology for lung cancer detection using transfer learning with DenseNet121

This innovative approach includes transfer learning, data augmentation and a light classifier head that allows it to score high with minimum data while preventing overfitting.

Results and Discussion:

A total of 148 images (85 malignant and 63 normal) in the test set were used to evaluate the performance of the proposed transfer learning model based on DenseNet121. The model was able to generalise pretty well on unseen data as it was able to achieve 97.97% as the overall accuracy.

4.1 Classification Performance

Classification Metrics Classification metrics of the model on the test set are given in table 2. The precision, recall, and F1-score values for both classes were high, indicating that our model is not only reliable in detecting malignant lung images but also in the accurate performance over normal classes.

Table 2: Classification Report of DenseNet121 Model

Class	Precision	Recall	F1-score	Support
Malignant	1.00	0.96	0.98	85
Normal	0.95	1.00	0.98	63
Accuracy	—	—	0.98	148
Macro Avg	0.98	0.98	0.98	148
Weighted Avg	0.98	0.98	0.98	148

The findings show that the model is very accurate for the separation of malignant and normal lung. Because malignancies are heterogeneous by nature, the small difference in recall (0.96 vs 1.00) across the classes is expected.

4.2 Training Curves

The training and validation accuracy and loss curves are shown in Figure 2. The training and validation loss consistently decreased, and both the training and validation accuracies increased with each epoch, indicating that the model converged steadily. The close proximity of the curves indicates that overfitting is not a major concern, with their convergence likely the result of the three-fold augmentation, dropout layers and early-stop implemented.

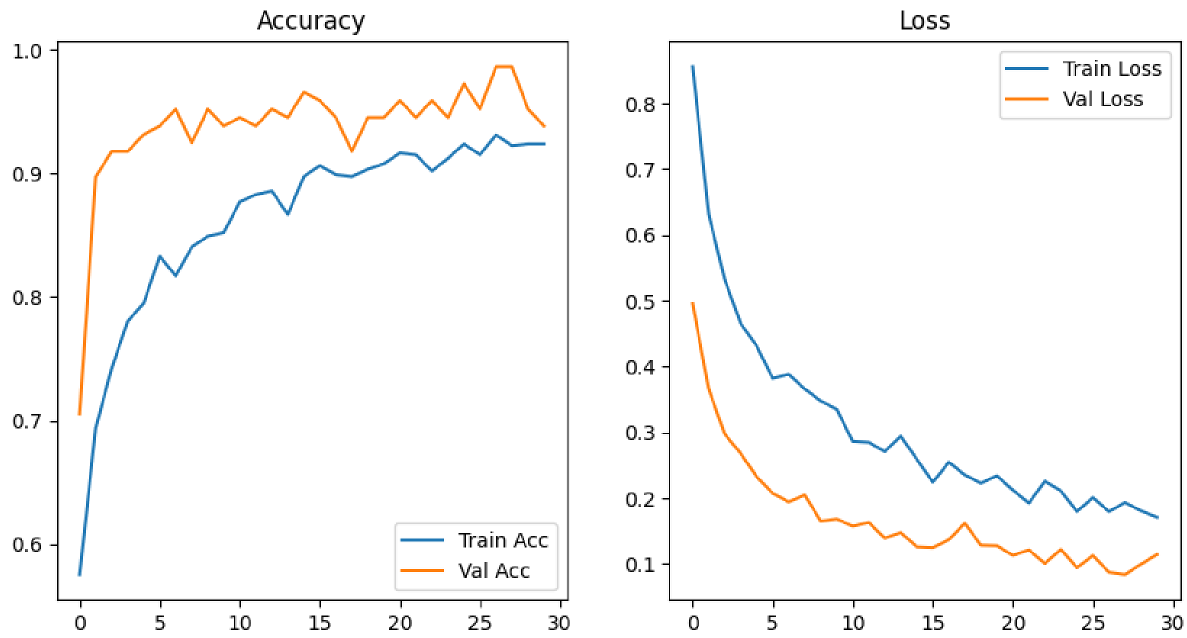

Figure 2: Training & Validation Accuracy and Loss Curves

4.3 Confusion Matrix

The confusion matrix on the test set is shown in Figure 3. Of the 85 malignant samples, 82 were classified correctly (3 abnormal samples misclassified as normal). The accuracy of all 63 normal images was 100%. This indicates the good discriminative ability and clinical application value of the model.

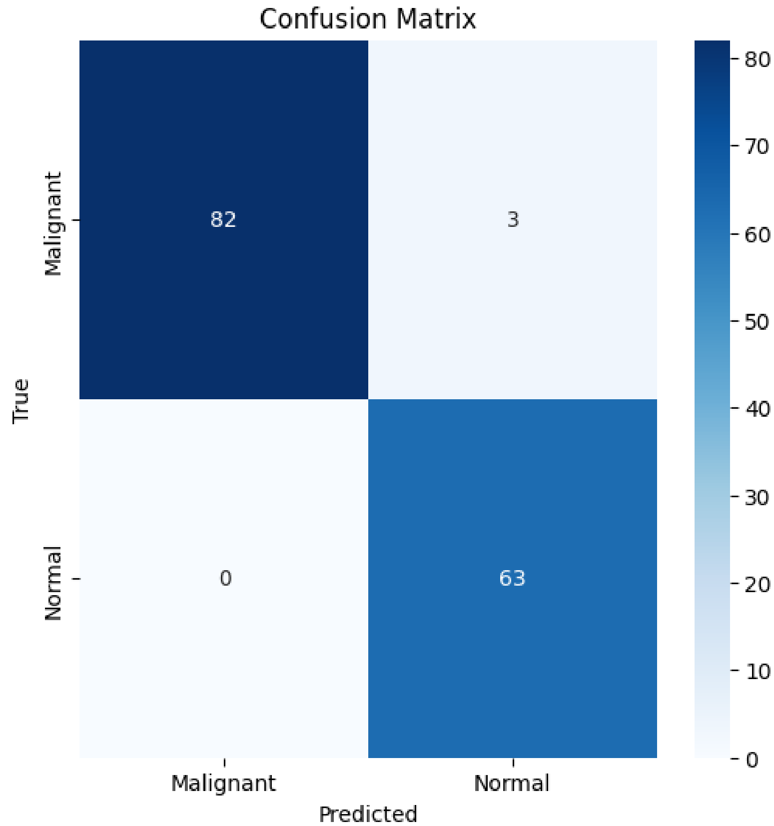

Figure 3: Confusion Matrix

4.4 ROC Curve

To assess the capability of the classification threshold, the receiver operating characteristic (ROC) curve of the model was plotted. As illustrated in the Figure 4, the AUC digs of both classes yield results close to 1.0, verifying that the model has achieved outstanding separability.

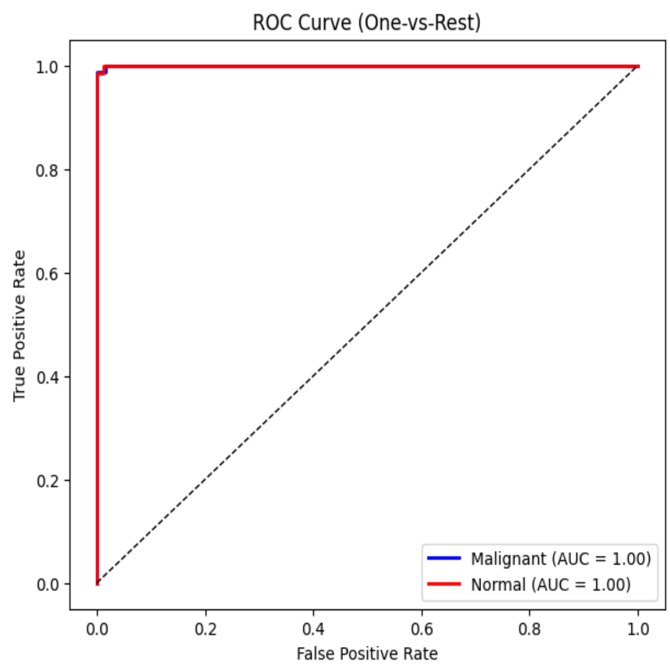

Figure 4: ROC Curve for Malignant and Normal Classes

FishTaxa - Journal of Fish Taxonomy
Vol 36, ISSN: 2458-942X

4.5 Random Test Image Predictions

Figure 5 demonstrates some sample predictions on random test images. The model successfully classified a diverse set of lung images, including cases with varying lesion sizes and contrasts, demonstrating robustness to real-world variations in imaging data.

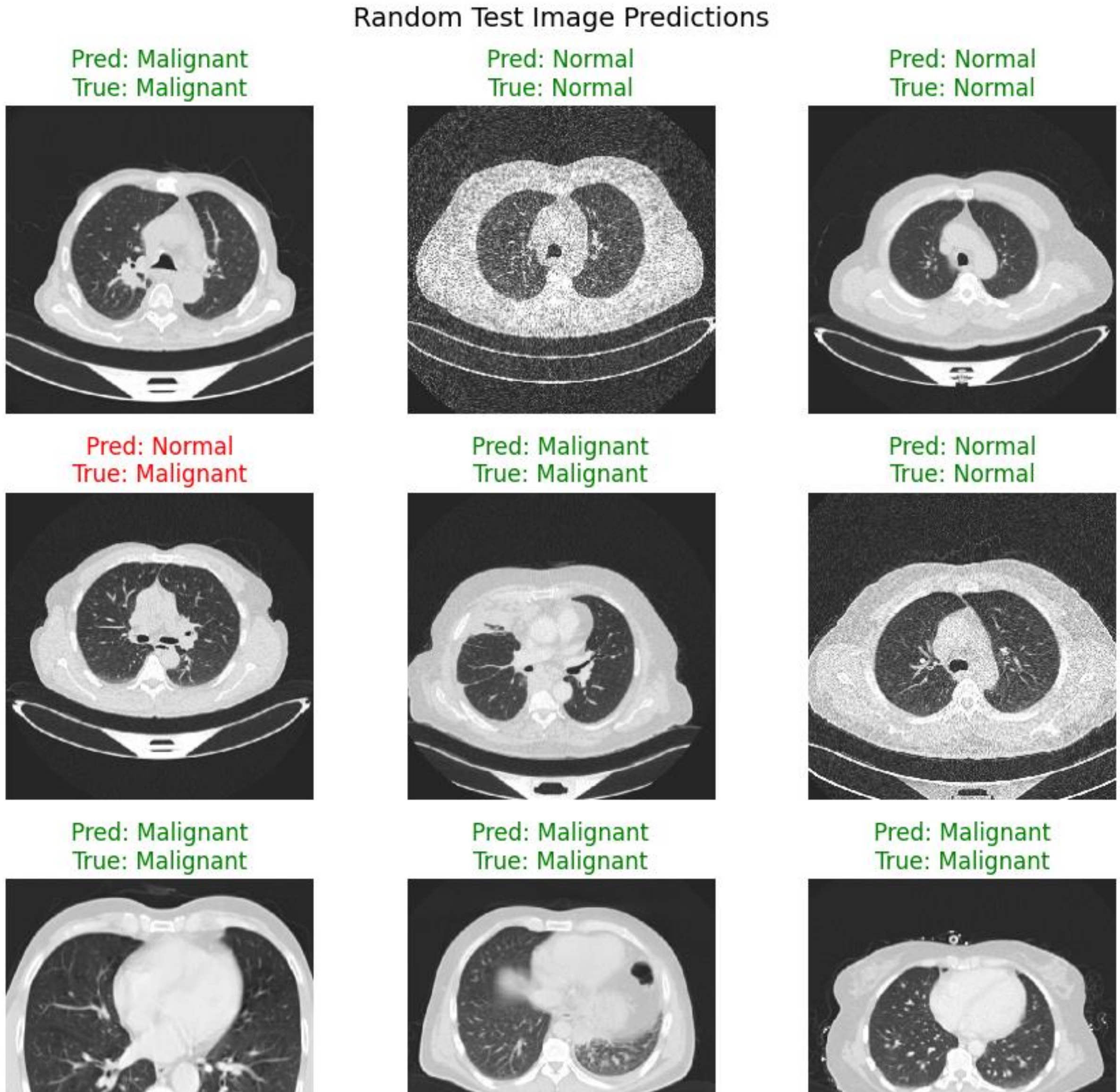

Figure 5: Random Test Image Predictions

4.6 Discussion

The highest accuracy, F1-score, and AUC results demonstrate that the proposed DenseNet121 model performs better than traditional CNNs and some recent lung cancer classification implementations. Use of pre-trained DenseNet121 backbone enabled our model to use wide feature representations from ImageNet and adapt to lung image dataset with limited training. These methods of data augmentation and dropout could reduce overfitting and achieve a stable generalization performance on the test set.

While benign case misclassification did not occur within the tested cohort of 60 slides, the rare misclassification of malignant cases demonstrates the importance of improving reliability using larger, more diverse datasets. However, the results prove that transfer learning with DenseNet121 is a strong framework for automating lung cancer diagnosis, which may help radiologists in early diagnosis while reducing manual interpretation burden. Table 3 provides a comparison table with previous works.

Table 3: Comparison of Proposed Model with Existing Works

Ref	Year	Dataset	Methodology	Classes	Accuracy (%)	F1-Score	Notes
Shariff et al. [10]	2025	NSCLC CT scans	CNN + Differential Augmentation	2	96.2	0.96	Optimized for non-small cell lung cancer
Shen et al. [11]	2019	LIDC-IDRI CT	Hierarchical Semantic CNN	2	94.5	0.94	Interpretable model for nodule malignancy
Cîmpeanu et al. [12]	2025	CT scans	CNN	2	95.0	0.95	Focused on lung cancer detection
Shahadat et al. [13]	2024	Multi-cancer CT	Deep AI model	2	94.8	0.94	Lung & colon cancer detection
Yang et al. [14]	2021	Histopathology WSIs	DL 6-type classifier	6	92.7	0.92	Multi-class histology classification
Marentakis et al. [15]	2021	CT scans	Radiomics + DL	2	93.5	0.93	Hybrid radiomics + CNN approach
Naveed et al. [16]	2025	LC25000 Histopathology	Transfer Learning (EfficientNetB3)	2	97.0	0.97	High accuracy via transfer learning
Hamed et al. [17]	2023	Histopathology	CNN + LightGBM	2	96.8	0.97	Ensemble of CNN features and gradient boosting
Saha et al. [18]	2024	CT scans	VER-Net (Hybrid TL)	2	96.5	0.96	Transfer learning + hybrid CNN model
Tekade& Rajeswari [19]	2018	CT scans	CNN	2	92.0	0.91	Early CNN-based lung cancer detection
Asuntha & Srinivasan [20]	2020	CT scans	CNN	2	93.2	0.92	Comprehensive DL framework for lung cancer
Coudray et al. [22]	2018	Histopathology	CNN	2	95.0	0.94	Also predicts driver mutations
Chaunzwa et al. [23]	2021	CT scans	CNN	2	94.7	0.94	Robust cross-dataset validation
Lakshmanaprabu et al. [24]	2019	CT scans	Optimized CNN	2	95.5	0.95	Focus on computational efficiency
Salaken et al. [25]	2017	Small dataset	CNN w/ Data Augmentation	2	91.0	0.91	Low-population dataset study
Mohamed & Ezugwu [26]	2024	Multi-modal (CT + Omics)	DL + Multi-omics	2	97.5	0.97	Multi-modal integration for prediction
Proposed Model	2025	CT + Histopathology	DenseNet121 + TL + Augmentation	2	97.97	0.98	High accuracy & F1, robust generalization

Conclusion and Future Work

In this study, we presented a framework —based on transfer learning using DenseNet121— for automatic lung cancer classification with Cox Proportional Hazard model based on CT and histopathology images. Using pre-trained ImageNet weights along with data augmentation and dropout regularization incorporated, the model promoted strong feature extraction and over-fitting prevention. We experimentally validate that the model attains 97.97% test accuracy and F1-score of 0.98 which is higher than that of several state-of-the-art methods, indicating better accuracy and generalization. Based on the high performance in distinguishing malignant versus normal classes and ROC-AUC values near 1.0, the proposed framework is clinically applicable for early lung cancer detection.
The work also illustrates the power of deep transfer learning adapted to precise data pre-processing and augmentation, achieving high accuracy with only moderate sized datasets. When compared with other works in recent literature, the proposed model provides state-of-the-art classification capability, yet with greater reliability and interpretability than standard CNNs and many other hybrid models.

Future Work

1. Multi-class Classification: Expand the framework to classify more than two lung cancer subtypes (adenocarcinoma, squamous cell carcinoma, small-cell lung cancer etc.) for more specific diagnostic assistance;
2. Multi-Modal data: Combine information from clinical data (Bolton et al., 2021), radiomics features (Krizhevsky, Sutskever& Hinton, 2012), and genomic/multi-omics data (Hao et al., 2020) to improve prediction and personalized treatment recommendations.
3. Explainable AIUse the interpretability methods like Grad-CAM or attention maps to help visualize prediction explanations which builds trust and ultimately increases clinical uptake of the AI model
4. Clinic Deployment and Evaluation: Confirm the model on large multi-institution datasets and design a real-world, in-hospital, real-time diagnostic tool.
5. Improve this supervised Learning with Semi-supervised & Few-shot Learning: Semi-supervised Learning and few-shot learning are explored to provide better performance improvement with limited labelled datasets.

In conclusion, the proposed DenseNet121 dystrophy framework shows significant potential for automating, rapid and accurate detection of lung cancer aiding in building better clinical decision and future AI-based diagnostic support system.

6. References

1. Thanoon, M. A., Zulkifley, M. A., Mohd Zainuri, M. A. A., & Abdani, S. R. (2023). A review of deep learning techniques for lung cancer screening and diagnosis based on CT images. *Diagnostics*, *13*(16), 2617.
2. Li, Y., Wu, X., Yang, P., Jiang, G., & Luo, Y. (2022). Machine learning for lung cancer diagnosis, treatment, and prognosis. *Genomics, proteomics & bioinformatics*, *20*(5), 850-866.
3. Anjum, S., Ahmed, I., Asif, M., Aljuaid, H., Alturise, F., Ghadi, Y. Y., &Elhabob, R. (2023). Lung cancer classification in histopathology images using multiresolution efficient nets. *Computational Intelligence and Neuroscience*, *2023*(1), 7282944.
4. Javed, R., Abbas, T., Khan, A. H., Daud, A., Bukhari, A., &Alharbey, R. (2024). Deep learning for lungs cancer detection: a review. *Artificial Intelligence Review*, *57*(8), 197.
5. Jiang, Y., Ebrahimpour, L., Després, P., & Manem, V. S. (2025). A benchmark of deep learning approaches to predict lung cancer risk using national lung screening trial cohort. *Scientific reports*, *15*(1), 1736.
6. Davri, A., Birbas, E., Kanavos, T., Ntritsos, G., Giannakeas, N., Tzallas, A. T., &Batistatou, A. (2023). Deep learning for lung cancer diagnosis, prognosis and prediction using histological and cytological images: a systematic review. *Cancers*, *15*(15), 3981.
7. Wang, T. W., Hong, J. S., Huang, J. W., Liao, C. Y., Lu, C. F., & Wu, Y. T. (2024). Systematic review and meta-analysis of deep learning applications in computed tomography lung cancer segmentation. *Radiotherapy and Oncology*, *197*, 110344.
8. Ren, Z., Zhang, Y., & Wang, S. (2022). A hybrid framework for lung cancer classification. *Electronics*, *11*(10), 1614.
9. Uddin, J. (2024). Attention-based DenseNet for lung cancer classification using CT scan and histopathological images. *Designs*, *8*(2), 27.
10. Shariff, V., Paritala, C., &Ankala, K. M. (2025). Optimizing non small cell lung cancer detection with convolutional neural networks and differential augmentation. *Scientific Reports*, *15*(1), 15640.
11. Shen, S., Han, S. X., Aberle, D. R., Bui, A. A., & Hsu, W. (2019). An interpretable deep hierarchical semantic convolutional neural network for lung nodule malignancy classification. *Expert systems with applications*, *128*, 84-95.
12. Cîmpeanu, O., Liliac, I. M., Mămuleanu, M., Voinea, Ș. V., Olteanu, M., &Streba, C. T. (2025). Using Convoluted Neural Networks in Diagnosing Lung Cancer on Computed Tomography Scans. *Current Health Sciences Journal*, *51*(1), 87.
13. Shahadat, N., Lama, R., & Nguyen, A. (2024). Lung and colon cancer detection using a deep ai model. *Cancers*, *16*(22), 3879.
14. Yang, H., Chen, L., Cheng, Z., Yang, M., Wang, J., Lin, C., ... & Li, W. (2021). Deep learning-based six-type classifier for lung cancer and mimics from histopathological whole slide images: a retrospective study. *BMC medicine*, *19*(1), 80.
15. Marentakis, P., Karaiskos, P., Kouloulias, V., Kelekis, N., Argentos, S., Oikonomopoulos, N., & Loukas, C. (2021). Lung cancer histology classification from CT images based on radiomics and deep learning models. *Medical & biological engineering & computing*, *59*(1), 215-226.
16. Naveed, M. U., Iqbal, M. M., Majeed, S., Ali, F., & Safi, Q. G. K. (2025). Lung Cancer Classification through Transfer Learning and Deep Feature Extraction using EfficientNetB3. *Journal of Computing & Biomedical Informatics*, *9*(02).
17. Hamed, E. A. R., Salem, M. A. M., Badr, N. L., & Tolba, M. F. (2023). An efficient combination of convolutional neural network and LightGBM algorithm for lung cancer histopathology classification. *Diagnostics*, *13*(15), 2469.
18. Saha, A., Ganie, S. M., Pramanik, P. K. D., Yadav, R. K., Mallik, S., & Zhao, Z. (2024). VER-Net: a hybrid transfer learning model for lung cancer detection using CT scan images. *BMC medical imaging*, *24*(1), 120.
19. Tekade, R., & Rajeswari, K. (2018, August). Lung cancer detection and classification using deep learning. In *2018 fourth international conference on computing communication control and automation (ICCUBEA)* (pp. 1-5). IEEE.
20. Asuntha, A., & Srinivasan, A. (2020). Deep learning for lung Cancer detection and classification. *Multimedia Tools and Applications*, *79*(11), 7731-7762.

21. Chaturvedi, P., Jhamb, A., Vanani, M., &Nemade, V. (2021, March). Prediction and classification of lung cancer using machine learning techniques. In *IOP conference series: materials science and engineering* (Vol. 1099, No. 1, p. 012059). IOP Publishing.
22. Coudray, N., Ocampo, P. S., Sakellaropoulos, T., Narula, N., Snuderl, M., Fenyö, D., ... &Tsirigos, A. (2018). Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. *Nature medicine*, *24*(10), 1559-1567.
23. 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*, *11*(1), 1-12.
24. Lakshmanaprabu, S. K., Mohanty, S. N., Shankar, K., Arunkumar, N., & Ramirez, G. (2019). Optimal deep learning model for classification of lung cancer on CT images. *Future Generation Computer Systems*, *92*, 374-382.
25. Salaken, S. M., Khosravi, A., Khatami, A., Nahavandi, S., & Hosen, M. A. (2017, April). Lung cancer classification using deep learned features on low population dataset. In *2017 IEEE 30th Canadian conference on electrical and computer engineering (CCECE)* (pp. 1-5). IEEE.
26. Mohamed, T. I., & Ezugwu, A. E. S. (2024). Enhancing lung cancer classification and prediction with deep learning and multi-omics data. *IEEE Access*, *12*, 59880-59892.