Computer Science and Engineering Research 3 (1): 38-47

Content available at Genesis Publishing Consortium Limited

Computer Science and Engineering Research

Journal homepage: www.genesispcl.com/journals/cser

Original Research

An augmentation-enhanced hybrid deep learning framework for lung cancer classification using the LC15000 histopathological image dataset

Ishmat Jahan[1*], Md. Jubayar Hossain Jamil[1], Khandaker Farhan Labib[1]

[1] *Department of Computer Science and Engineering, Faculty of Science and Engineering, Green University of Bangladesh, Purbachal American City, Rupganj, Narayanganj-1460, Bangladesh*

Article info

Article history
Received: 15 March 2026
Revised: 20 April 2026
Accepted: 29 May 2026
Published: 01 Nov 2026

Keywords
Transfer learning
Image classification
Data augmentation
EfficientNetB5
BiLSTM

Abstract

Lung cancer is one of the leading causes of death from cancer around the world. Early diagnosis is necessary for effective treatment. In Bangladesh, air quality is at an alarming state, especially in the capital city of Dhaka and other major cities, where high levels of air pollution are causing respiratory problems and increasing the vulnerability to lung diseases. This has aggravated the issue of unhealthy lungs even more. Since the quality of air is already poor, there is an immediate need to discover lung cancer through histopathological images, a field in which automation and a deep learning approach can expedite the accurate process of detection. This work deals with a system for lung cancer detection and classification from histopathological images by using deep learning architecture, which employs the LC15000 dataset. The methodology uses CNNs, which is a deep learning technique, to classify histopathological images of lung cancer. Lung cancer images are preprocessed to ensure consistency and suitability for model training, with augmentation techniques applied to enhance the dataset. A carefully selected CNN architecture will be trained on a subset of the dataset and tested on validation and test sets. The model's diagnostic capability will be assessed by important metrics like accuracy, precision, recall, and F1-score. The Hybride Model that we have developed, it's with Augmentation: Accuracy: 0.9967, Precision: 0.9967, Recall: 0.9967, F1 Score: 0.9966 and it's without Augmentation: Accuracy: 0.9913, Precision: 0.9913, Recall: 0.9913, F1 Score: 0.9913. It is expected that the deep learning model will be highly discriminative in classifying lung cancer from benign tissue and contribute to developing an automated diagnostic tool for pathology. This has big implications for Bangladesh, where poorly maintained air quality gives the population a higher rate of lung disease, while also facilitating the integration of AI-driven methods in a clinical setting for more efficient and accessible healthcare solutions.

© 2026 Jahan et al. This is an open access article distributed under the **Creative Commons Attribution 4.0 International License** (www.creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

Lung cancer remains one of the most critical global public health challenges due to its high incidence, aggressive progression, and poor prognosis when diagnosed at advanced stages. It is consistently ranked among the leading causes of cancer-related morbidity and mortality worldwide, accounting for a substantial proportion of all cancer incidences (Wu *et al.*, 2024). According to estimates from the World Health Organization (WHO), more than 1.8 million deaths occur annually due to lung cancer, making it one of the deadliest malignancies globally (Tu *et al.*, 2026; Li *et al.*, 2023; Thandra *et al.*, 2021). The primary reason for this high mortality rate is delayed diagnosis, as the disease is often detected only after it has reached an advanced or metastatic stage. Early-stage detection significantly improves survival rates; however, identifying lung cancer at its earliest stage remains a major clinical challenge (Habbab *et al.*, 2025). This challenge is further compounded by the subtle morphological differences between benign and malignant tissues in histopathological images, which require expert interpretation and are often difficult to distinguish even for experienced pathologists.

The conventional diagnostic process for lung cancer heavily relies on histopathological examination, where tissue samples are observed under a microscope by trained pathologists (Davri *et al.*, 2023). Although this method is considered the diagnostic gold standard, it is inherently subjective, time-consuming, and prone to inter-observer and intra-observer variability. The increasing workload in pathology laboratories further compounds these limitations, often leading to diagnostic delays and potential inconsistencies in interpretation (Faria *et al.*, 2025). As the number of biopsy samples continues to rise globally, there is an increasing need for automated systems that can assist pathologists in improving diagnostic efficiency and accuracy.

The burden of lung cancer is particularly concerning in developing countries such as Bangladesh, where environmental and infrastructural challenges significantly impact public health outcomes. Rapid urbanization and industrialization have led to severe air pollution in major cities, with Dhaka frequently ranking among the most polluted cities in the world according to global air quality indices (Rahman *et al.*, 2025). Prolonged exposure to harmful pollutants such as particulate matter ($PM_{2.5}$), carbon monoxide, and other toxic gases has been strongly associated with respiratory diseases and an increased risk of lung cancer. As a result, lung-related illnesses are becoming increasingly common, and many cases are diagnosed at advanced stages due to limited access to early screening facilities and insufficient awareness (Wang *et al.*, 2025). This highlights a critical gap in healthcare systems and underscores the urgent need for reliable and scalable diagnostic tools.

In recent years, artificial intelligence (AI) and machine learning (ML) have emerged as transformative technologies in medical image

[*]***Corresponding authors***
Email address: ishmatjahan248@gmail.com (Ishmat Jahan)

doi: https://doi.org/10.69517/cser.2026.03.01.0006

ISSN: 3136-8549

analysis. In particular, deep learning techniques, especially convolutional neural networks (CNNs), have demonstrated exceptional capability in automatically learning complex hierarchical features from images without requiring manual feature engineering (Smrity *et al.*, 2026). These models have shown strong performance in various medical imaging tasks, including tumour detection, organ segmentation, and disease classification. Their ability to analyse large volumes of data efficiently makes them particularly suitable for histopathological image analysis, where subtle morphological differences between benign and malignant tissues must be identified with high precision (Ahmad *et al.*, 2025; Rayed *et al.*, 2024).

Despite strong progress in deep learning for lung cancer diagnosis, several key limitations still restrict clinical translation. The most important issue is poor generalizability, as many high-performing models reported in different studies that are trained and tested on relatively clean datasets like LC15000 or LC25000 (Shweikeh *et al.*, 2026; Durgam *et al.*, 2025). In real clinical environments, however, histopathological images vary widely due to differences in staining protocols, imaging quality, and the presence of rare tumour subtypes. Without multi-centre or heterogeneous dataset validation, the reliability of these models in practical deployment remains uncertain (Raza *et al.*, 2026). Another major limitation is the lack of interpretability in many deep learning systems. Although some works, such as Garg and Garg, use techniques like Grad-CAM to highlight important image regions, most high-accuracy models still function as black-box systems (Ennab and Mcheick, 2025; Henna *et al.*, 2025). In medical applications, this lack of transparency is a serious barrier because clinicians need to understand and trust model decisions before integrating AI tools into diagnostic workflows.

In addition, the complexity of many proposed deep learning pipelines presents a barrier to real-world clinical deployment. Several studies rely on multi-stage architectures combining deep feature extraction with ensemble classifiers or hybrid models (Ochoa-Ornelas *et al.*, 2024). While these approaches often yield high performance, they may not be practical in hospital environments where computational resources, processing time, and system simplicity are critical considerations. A clinically viable model must balance accuracy with efficiency and ease of integration into existing workflows (Ali *et al.*, 2025).

There is still limited progress in extending deep learning applications beyond simple binary or categorical classification tasks. More advanced clinical applications such as prognosis prediction, survival analysis, and drug response estimation remain relatively underexplored. Although some recent studies, such as Zhang et al. and Pan et al., have made progress in areas like contrastive clustering and mutation prediction, these approaches are still in early stages of development (Pan *et al.*, 2025; Großerueschkamp and Gerwert, 2020). This represents a significant opportunity for future research to move toward more clinically meaningful outcomes that support personalized cancer treatment and precision oncology.

This study addresses several research questions: How do pre-trained CNN architectures (VGG16, MobileNetV3-Large, and EfficientNetB5) perform on the LC15000 dataset with and without data augmentation? Can a hybrid CNN-BiLSTM model improve lung cancer classification performance compared to standard CNN models? Does data augmentation enhance the generalization ability of deep learning models in histopathological image classification? Which model provides the most robust and clinically reliable performance for this task? Based on these questions, it is hypothesized that data augmentation improves model generalization, and that the hybrid CNN-BiLSTM model achieves superior performance over traditional CNN architectures in terms of accuracy and overall classification metrics.

This study addresses these challenges by systematically evaluating multiple CNN architectures under both augmented and non-augmented conditions using the LC15000 dataset. The primary aim is to identify robust, generalizable, and efficient models for lung cancer classification. By comparing classical, lightweight, and advanced architectures, as well as introducing a hybrid CNN-BiLSTM model, this research seeks to contribute to the development of more reliable and clinically applicable AI-based diagnostic systems. The findings are expected to provide insights into model performance, generalization behaviour, and the role of data augmentation in improving robustness, ultimately supporting the advancement of automated histopathological analysis in lung cancer diagnosis.

2. Literature review

The rapid growth of AI in medical imaging has significantly transformed the field of histopathological analysis, particularly in cancer diagnosis. Lung cancer detection, in particular, has benefited from deep learning approaches due to the increasing demand for fast, accurate, and reproducible diagnostic systems. Traditional histopathological examination relies on manual microscopic evaluation by expert pathologists, which is inherently subjective, time-consuming, and susceptible to inter-observer and intra-observer variability. These limitations have motivated the development of automated systems capable of assisting in diagnosis and improving consistency and accuracy in clinical decision-making (Arshad *et al.*, 2025; Zhu *et al.*, 2025).

Deep learning, especially CNNs, has emerged as the dominant approach for medical image analysis due to its ability to automatically extract hierarchical feature representations without manual feature engineering. Models such as VGGNet, ResNet, Inception, and EfficientNet have demonstrated strong performance across a variety of medical imaging tasks, including tumor classification and organ segmentation (Gayap and Akhloufi, 2024). Among these, EfficientNet has gained significant attention due to its compound scaling strategy, which efficiently balances network depth, width, and resolution, achieving state-of-the-art performance in many image classification benchmarks (Lin *et al.*, 2023). Transfer learning has also played a crucial role in medical imaging applications, where models pre-trained on large-scale natural image datasets such as ImageNet are fine-tuned on domain-specific medical datasets, enabling improved feature learning even with limited labelled data (Wen *et al.*, 2021).

In addition to model architecture advancements, data augmentation techniques have been widely used to improve generalization and reduce overfitting in deep learning models. Common augmentation strategies such as rotation, flipping, scaling, shifting, and brightness variation artificially increase dataset diversity and simulate real-world variability in medical images (Goceri, 2023). Although augmentation generally improves model robustness, its impact is not always consistent across different architectures and datasets, particularly in histopathological imaging where subtle texture differences are clinically significant (Lerch *et al.*, 2024).

More recently, hybrid deep learning approaches have been explored to enhance classification performance by combining CNNs with sequential learning models such as Long Short-Term Memory (LSTM) networks. These hybrid architectures aim to capture both spatial feature representations and contextual dependencies within feature maps, offering improved discriminatory capability in complex medical imaging tasks (Hayat *et al.*, 2025). Despite these advancements, most existing studies still rely on standalone CNN architectures or handcrafted feature-based pipelines, with limited comparative analysis across multiple models under standardized experimental conditions (Raza *et al.*, 2026). This creates a clear research gap in evaluating how different architectures perform under consistent datasets and training settings, particularly when considering augmentation effects.

Several prior studies have contributed significantly to lung cancer classification using histopathological images. For example, previous studies developed a custom CNN model trained on the LC15000 dataset, achieving very high classification accuracy by leveraging deep convolutional layers to automatically learn discriminative tissue features (Attallah, 2025; Uddin *et al.*, 2024). However, the study

lacked interpretability mechanisms such as Grad-CAM, limiting clinical transparency. Similarly, Talukder *et al.* (2022) combined pretrained CNN feature extractors with ensemble classifiers, demonstrating high performance on the LC25000 dataset, though the complexity of the multi-stage pipeline raised concerns regarding clinical deployability (Ochoa-Ornelas *et al.*, 2025).

Garg and Garg (2020) investigated multiple pre-trained CNN models including ResNet50, DenseNet169, and Xception, and incorporated Grad-CAM visualization to enhance interpretability. Their study highlighted the importance of explainability in clinical AI systems, although performance varied across architectures depending on transfer learning compatibility (Abraham *et al.*, 2025). In a broader context, Parra-Medina *et al.* (2025) conducted a systematic review on deep learning models for predicting oncogenic drivers in non-small cell lung cancer, reporting moderate to high predictive performance but emphasizing the need for multi-center validation to ensure generalizability (Fuster-Matanzo *et al.*, 2025).

Further advancements have extended deep learning applications beyond classification tasks. Großerueschkamp and Gerwert (2020) proposed a contrastive clustering framework for prognosis and drug response prediction, demonstrating the potential of deep learning in precision oncology beyond diagnostic classification. Similarly, Abd El-Aziz *et al.* (2024) introduced a multi-CNN fusion model combining ResNet101V2, NASNetMobile, and EfficientNet-B0, achieving near-perfect classification performance on histopathological datasets, although concerns remain regarding external validation and real-world robustness. Recent developments also include dataset-driven and multi-instance learning approaches. Pan *et al.* (2026) introduced PathGene, a multicentre dataset designed for mutation prediction, and demonstrated that deep learning models can predict genetic alterations such as EGFR, KRAS, and TP53 directly from histopathology images. Another study by Pan *et al.* (2025) proposed SMILE, a scale-aware multiple-instance learning framework for detecting spread-through-air-space (STAS) patterns, achieving strong AUC performance across multicentre datasets.

Lightweight and efficient architectures have also been explored to address computational constraints in clinical deployment. For instance, Pang *et al.* (2022) proposed DCSNet, a lightweight model based on knowledge distillation that achieved high classification accuracy with reduced computational cost and improved interpretability through attention-based visualization. Such approaches are particularly relevant for real-world deployment in resource-limited healthcare settings.

Despite these advancements, several research gaps remain evident. First, most existing models demonstrate high accuracy on curated datasets such as LC15000 and LC25000, but their generalizability to real-world clinical environments remains uncertain due to variations in staining protocols, imaging quality, and biological heterogeneity. Second, interpretability remains limited in many high-performing models, reducing clinical trust and adoption. Third, many studies rely on complex multi-stage or ensemble-based pipelines that are difficult to deploy in real-time clinical workflows. Finally, there is limited comparative analysis of multiple CNN architectures under consistent experimental settings, particularly with systematic evaluation of data augmentation effects.

To address these gaps, this study proposes a comprehensive comparative analysis of multiple deep learning architectures, including VGG16, MobileNetV3-Large, EfficientNetB5, and a hybrid CNN-BiLSTM model, using the LC15000 dataset. The study also systematically evaluates the impact of data augmentation on model performance and generalization (Table 1). The proposed hybrid architecture integrates spatial feature extraction from CNNs with sequential learning from BiLSTM networks to improve classification robustness. This dual experimental design enables a fair comparison of classical, lightweight, and advanced architectures under standardized conditions, contributing to the development of more reliable and clinically applicable AI-based diagnostic systems for lung cancer classification. A summary of related works on deep learning applications in lung cancer histopathol¬ogy is presented in following Table 1.

Table 1. Summary of deep learning-based studies for lung cancer diagnosis using histopathological images.

Model / Method	Accuracy / AUC	Reference
Custom CNN on LC15000, automatic feature learning	99.80%	Mohalder *et al.* (2023)
Pretrained CNNs (VGG16, InceptionV3) + Ensemble (RF, GB) on LC25000	99.05%	Talukder *et al.* (2022)
Transfer learning (ResNet50, DenseNet169, Xception) + Grad-CAM	96–100%	Garg and Garg (2020)
Meta-analysis for EGFR/ALK/TP53 prediction from HandE WSIs	~70–85%	Parra-Medina *et al.* (2025)
EfficientNet-B5 + ResNet50, six-class classification with federated test cohorts	AUC 0.918–0.978	Yang *et al.* (2021)
Systematic review on DL for histological/cyto-logical lung cancer images	N/A	Davri *et al.* (2023)
Multi-CNN fusion (ResNet101V2, NASNetMo-bile, EfficientNet-B0) on LC25000	99.96%	Abd El-Aziz *et al.* (2024)
PathGene: MIL for gene mutation (EGFR, KRAS, TP53) prediction from WSIs	High (mutation-specific)	Pan *et al.* (2026)
SMILE: Scale-aware MIL for detecting STAS in multicenter dataset	AUC 0.87–0.92	Pan *et al.* (2025)
Label-free DL on autofluorescence histology for NSCLC subtype classification	AUC > 0.98	Großerueschkamp and Gerwert (2020)
DCSNet: Lightweight CNN with knowledge distillation and XAI on LC25000	>99%	Pang *et al.* (2022)

3. Methodology

3.1 Workflow

The research follows a structured and reproducible experimental pipeline designed to evaluate multiple deep learning architectures for lung cancer histopathological image classification. The workflow spans dataset acquisition, preprocessing, augmentation, model development, training, evaluation, and comparative analysis, ensuring consistency across all experimental settings (Figure 1).

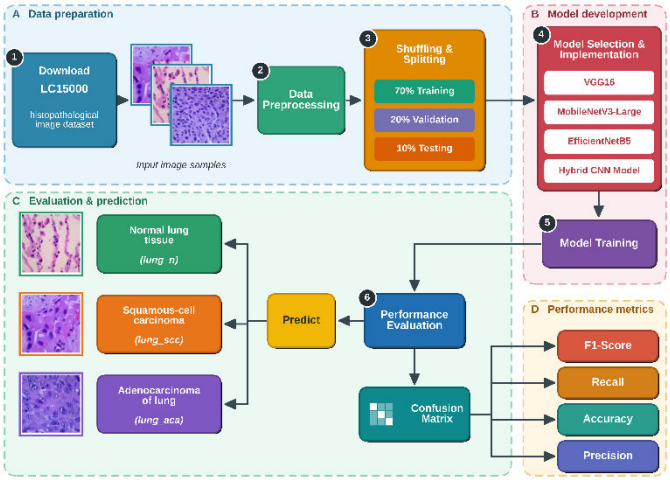

Figure 1. Workflow diagram of methodology.

3.2 Dataset description and acquisition

The LC15000 histopathological image dataset is used in this study. It contains 15,000 high-resolution lung tissue images, equally distributed into three categories: 5000 adenocarcinoma of lung, 5000 squamous-cell carcinoma of lung, and 5000 normal lung tissue images. The dataset is publicly available and widely used in lung

cancer image analysis research due to its balance and quality. For the purpose of this study, the dataset is reformulated into a binary classification task where adenocarcinoma and squamous-cell carcinoma are combined into a single malignant class, while normal tissue represents the benign class. This transformation aligns the dataset with clinically relevant diagnostic requirements for malignancy screening.

3.3 Data preprocessing

A standardized preprocessing pipeline is applied to ensure compatibility with pre-trained CNN architectures and to improve numerical stability during training. All images are resized to 224 × 224 pixels, which matches the input requirements of VGG16, MobileNetV3-Large, and EfficientNetB5. Pixel values are normalized from the original range of 0–255 to a scaled range of 0–1 to stabilize gradient updates and accelerate convergence. Label encoding is performed by converting categorical class labels into binary format, where benign is represented as 0 and malignant as 1. The dataset is shuffled to eliminate ordering bias and then split into training (70%), validation (20%), and testing (10%) subsets. This stratified split ensures balanced class distribution across all subsets and prevents sampling bias during evaluation (Table 2).

Table 2. Dataset distribution across training, validation, and testing sets.

Dataset portion	Percentage	Number of images
Training	70%	3500
Validation	20%	1000
Testing	10%	500

3.4 Data augmentation strategy

To evaluate the effect of data variability on model generalization, two parallel experimental branches are created. In the augmentation branch, real-time transformations are applied only to the training dataset, while validation and test sets remain unchanged to ensure unbiased evaluation. The augmentation techniques include random rotation within ±20 degrees, horizontal flipping, zooming between 0.8 and 1.2, width and height shifts up to ±10%, shear transformation with a range of 0.2, and nearest-neighbor fill mode to preserve image structure.

3.5 Model architectures and implementation

Four deep learning architectures are implemented and evaluated: VGG16, MobileNetV3-Large, EfficientNetB5, and a custom hybrid CNN-BiLSTM model. All models are initialized with ImageNet pretrained weights and fine-tuned for binary classification on the LC15000 dataset. VGG16 consists of 13 convolutional layers and 3 fully connected layers with 3×3 convolutional filters and max pooling operations, making it a deep but computationally expensive baseline model. MobileNetV3-Large incorporates depthwise separable convolutions, squeeze-and-excitation blocks, and the Hard-Swish activation function, making it lightweight and suitable for efficient computation. EfficientNetB5 employs compound scaling of depth, width, and resolution and integrates MBConv blocks with squeeze-and-excitation mechanisms for high representational capacity. The hybrid model integrates MobileNetV3-Small with a Bi-directional Long Short-Term Memory (BiLSTM) network to combine spatial feature extraction with sequential dependency modeling. For all models, the original classification layers are removed and replaced with a global average pooling layer, followed by fully connected dense layers, dropout regularization, and a sigmoid activation layer for binary classification (Figure 2).

3.6 Training procedure

All models are trained using a standardized deep learning pipeline implemented in TensorFlow with a Keras backend. Training is conducted under identical conditions to ensure fair comparison. The Adam optimizer is used with an initial learning rate of 1×10^{-4}, and binary cross-entropy is used as the loss function due to the binary classification setup. A batch size of 32 is used across all experiments. Models are trained for a maximum of 50 epochs; however, early stopping is applied with a patience of 5 epochs based on validation loss to prevent overfitting. Model checkpoints are used to retain the best-performing weights during training. Real-time augmentation is applied only in the augmented branch using Keras ImageDataGenerator. In the non-augmented branch, only rescaling is applied (1/255 normalization). A fixed random seed is used to ensure reproducibility, and identical data splits are maintained across all experiments.

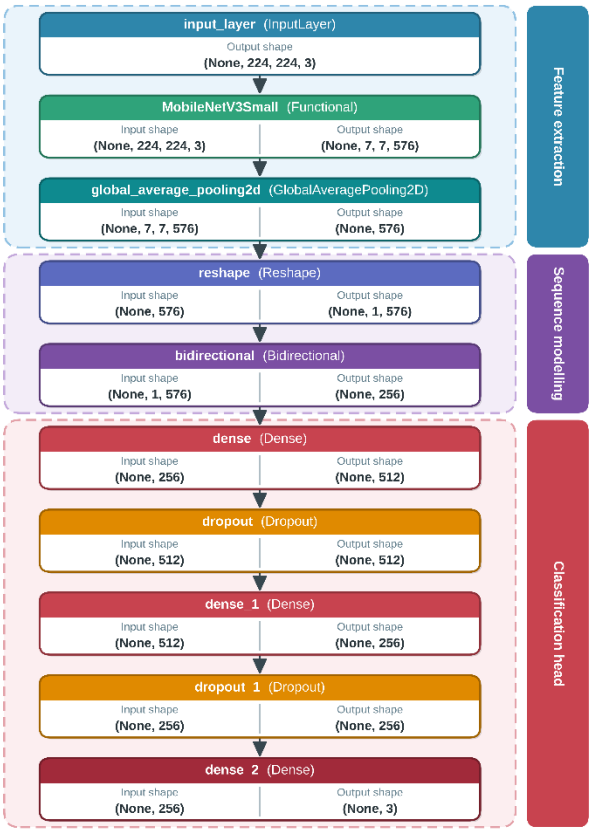

Figure 2. Architecture of the proposed hybrid CNN-BiLSTM model.

3.7 Validation and testing strategy

The dataset is partitioned into training (70%), validation (20%), and testing (10%) subsets, ensuring no overlap between sets. The validation set is used during training to monitor model performance, guide early stopping, and prevent overfitting. The test set is completely isolated and used only for final performance evaluation. The validation set consists of 20% of the dataset (3,000 images), while the test set consists of 10% (1,500 images).

3.8 Regularization techniques

Multiple regularization strategies are incorporated to improve generalization and reduce overfitting. Dropout is applied in fully connected layers across all models, with dropout rates ranging from 0.3 in lightweight architectures such as MobileNetV3-Large to 0.5 in deeper models such as VGG16, EfficientNetB5, and the hybrid model. Data augmentation acts as an additional regularization mechanism by introducing controlled variability in the training data. Early stopping is applied based on validation loss with a patience of 5 epochs, ensuring training halts at optimal convergence. Light L2 weight regularization is selectively applied in the hybrid model to penalize overly complex weight distributions and improve stability.

3.9 Evaluation methodology

Model evaluation is conducted exclusively on the test set, which contains 10% of the dataset (1,500 images) and remains unseen during training and validation. Each model is evaluated under both experimental conditions: with augmentation and without augmentation, resulting in eight total experimental configurations. Performance is measured using accuracy, precision, recall, F1-score,

binary cross-entropy loss, confusion matrix, and ROC-AUC curves. Accuracy measures overall correctness, while precision and recall assess class-specific predictive performance. The F1-score provides a balanced measure of precision and recall, particularly important in medical classification tasks where false negatives are critical. The confusion matrix is used to analyze misclassification patterns, while ROC-AUC evaluates the trade-off between sensitivity and specificity. All metrics are computed using TensorFlow/Keras evaluation functions and systematically recorded for comparative analysis across models and experimental conditions.

4. Results

4.1 Performance evaluation and model performance comparison

The evaluation was conducted using a test dataset comprising 2,250 histopathological images, representing 15% of the total LC15000 dataset. All samples were strictly excluded from the training and validation phases to ensure unbiased performance assessment. Each model was evaluated under two experimental conditions, with and without data augmentation, using identical preprocessing pipelines and training configurations to maintain comparability.

The performance was assessed using test accuracy, precision, recall, F1-score, and test loss computed via binary cross-entropy or categorical cross-entropy depending on the model structure. In addition, validation accuracy and validation loss were monitored during training to assess convergence behavior and detect overfitting. Confusion matrices and ROC curves were used to evaluate class-wise performance and discriminative capability. Without augmentation, EfficientNetB5 achieved the highest performance with an accuracy of 0.996, followed by VGG16 and the Hybrid model (0.9913), while MobileNetV3-Large recorded 0.9907 (Table 3).

Table 3. Performance of CNN models without data augmentation on LC15000 dataset (accuracy, precision, recall, F1-score).

Model	Accuracy	Precision	Recall	F1-score
EfficientNetB5	0.996	0.996	0.996	0.996
VGG16	0.9913	0.9914	0.9913	0.9913
MobileNetV3-Large	0.9907	0.9909	0.9907	0.9907
Hybrid	0.9913	0.9913	0.9913	0.9913

Under augmentation, the Hybrid CNN-BiLSTM model achieved the best performance with an accuracy of 0.9967, followed by VGG16 (0.988), EfficientNetB5 (0.9773), and MobileNetV3-Large (0.9727) (Table 4).

Table 4. Performance of CNN models with data augmentation on LC15000 dataset (accuracy, precision, recall, F1-score).

Model	Accuracy	Precision	Recall	F1-score
EfficientNetB5	0.9773	0.9775	0.9773	0.9774
VGG16	0.988	0.988	0.988	0.988
MobileNetV3-Large	0.9727	0.9739	0.9727	0.9727
Hybrid	**0.9967**	0.9967	0.9967	0.9967

4.2 VGG16 performance analysis

The VGG16 model was evaluated using ImageNet-pretrained weights with a modified binary classification head. Training was conducted for up to 20 epochs with early stopping based on validation loss. VGG16 demonstrated fast convergence in the non-augmented setting, with rapid reduction in training and validation loss during early epochs followed by stabilization. In the augmented setting, convergence was slower but smoother, indicating improved generalization and reduced overfitting due to increased data variability. In the non-augmented setting, VGG16 achieved an accuracy of 0.9913 (Figure 3), while in the augmented setting it achieved 0.988 (Figure 4).

Although there was a slight reduction in accuracy with augmentation, the model exhibited improved robustness to unseen variations in histopathological images. In the non-augmented case, the model correctly classified all normal lung tissue samples with zero error, while minor misclassifications occurred between adenocarcinoma and squamous cell carcinoma. In the augmented case, the normal class remained perfectly classified; however, slightly increased confusion was observed between cancer subtypes, indicating a minor reduction in class boundary sharpness due to augmentation-induced variability.

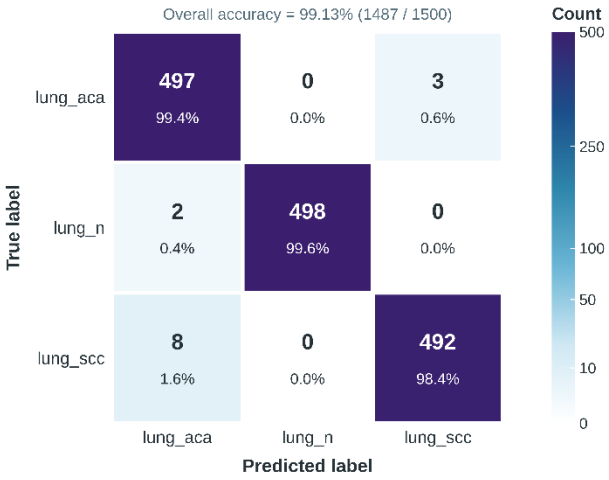

Figure 3. Confusion matrix of VGG16 without data augmentation.

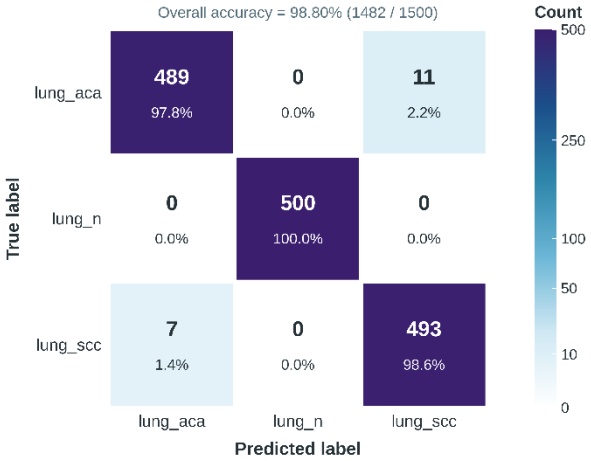

Figure 4. Confusion matrix of VGG16 with data augmentation.

4.3 MobileNetV3-large performance analysis

MobileNetV3-Large, a lightweight CNN architecture optimized for efficiency, was evaluated under both experimental conditions. The model showed stable convergence with steadily increasing training and validation accuracy. Minimal divergence between training and validation curves indicated low overfitting and stable learning behavior across both settings. Without augmentation, MobileNetV3-Large achieved an accuracy of 0.9907, while with augmentation the accuracy dropped to 0.9727 (Figures 5 and 6).

Precision, recall, and F1-score followed a similar trend, indicating that augmentation slightly reduced discriminative performance in this architecture. The augmented model showed increased misclassification between adenocarcinoma and squamous cell carcinoma, while the non-augmented model demonstrated better class separation with fewer errors.

4.4 EfficientNetB5 performance analysis

EfficientNetB5, based on compound scaling, demonstrated strong feature extraction capability and stable learning behavior. The model exhibited smooth convergence with consistently decreasing loss and stable validation performance. It maintained strong generalization in both augmented and non-augmented settings. EfficientNetB5

achieved 0.996 accuracy without augmentation and 0.9773 with augmentation (Figures 7 and 8).

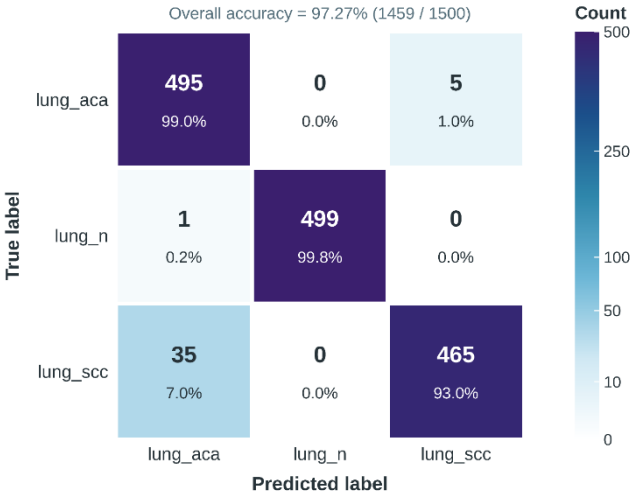

Figure 5. Confusion matrix of MobileNetV3-Large with data augmentation.

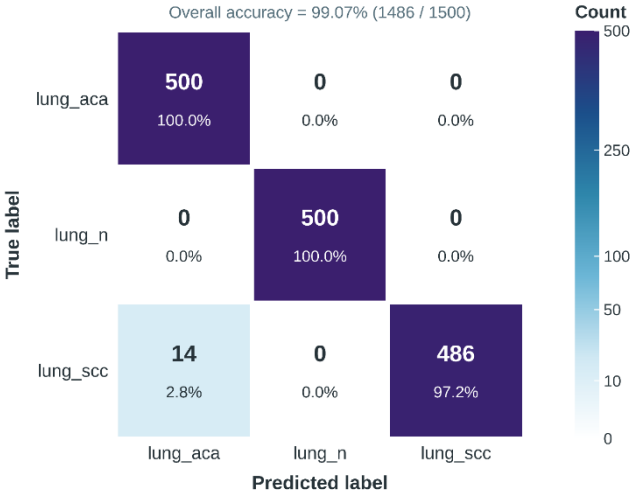

Figure 6. Confusion matrix of MobileNetV3-Large without data augmentation.

Despite the reduction under augmentation, it remained one of the strongest standalone models in this study. In the non-augmented setting, only a very small number of misclassifications occurred. In the augmented setting, limited confusion was observed mainly between adenocarcinoma and squamous cell carcinoma classes, reflecting high but slightly reduced class separability.

4.5 Hybrid CNN–BiLSTM model performance analysis

The hybrid model combines MobileNetV3Small as a feature extractor with a Bi-directional LSTM layer for sequential feature modeling. The model demonstrated rapid convergence, with validation accuracy stabilizing early in training. The inclusion of BiLSTM layers improved feature interaction learning, resulting in strong class separability. ROC analysis showed an AUC value of 1.00 across all classes, indicating near-perfect discrimination. The hybrid model achieved an accuracy of 0.9913 without augmentation and 0.9967 with augmentation, making it the best-performing model under augmented conditions (Figures 9 and 10).

Precision, recall, and F1-score followed similar trends, indicating balanced and stable classification performance. The augmented hybrid model showed minimal misclassification between adenocarcinoma and squamous cell carcinoma, while maintaining strong performance across all classes.

ROC analysis confirmed perfect separability with AUC = 1.00, indicating highly reliable classification performance (Figure 11).

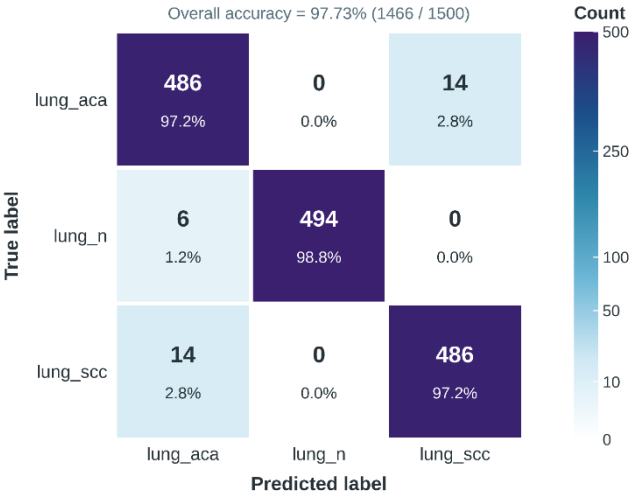

Figure 7. Confusion matrix of EfficientNetB5 with data augmentation.

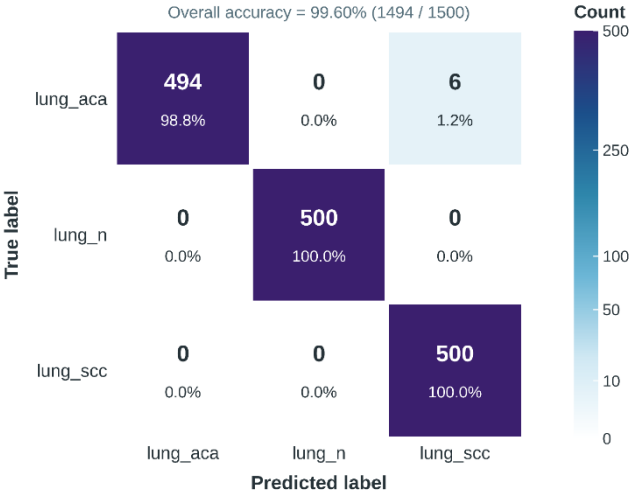

Figure 8. Confusion matrix of EfficientNetB5 without data augmentation.

4.6 Impact of data augmentation

Data augmentation had a measurable influence on both model generalization and convergence behavior. Although augmentation slightly reduced peak accuracy in some models, it consistently improved robustness and reduced overfitting. In non-augmented settings, models such as VGG16 and EfficientNetB5 achieved faster convergence and lower training loss but showed higher sensitivity to overfitting. In contrast, augmented models demonstrated smoother training curves and closer alignment between training and validation performance, particularly in EfficientNetB5 and the hybrid model.

5. Discussion

The results demonstrate that deep learning-based architectures can achieve very high performance in histopathological lung cancer classification using the LC15000 dataset. Across all experiments, both classical and modern CNN architectures consistently achieved accuracy above 97%, confirming the strong suitability of convolutional feature extraction for histopathological image analysis. The consistently high performance aligns with previous studies reporting strong results of CNN-based models in lung cancer classification tasks (Abd El-Aziz *et al.*, 2024; Mohalder *et al.*, 2023).

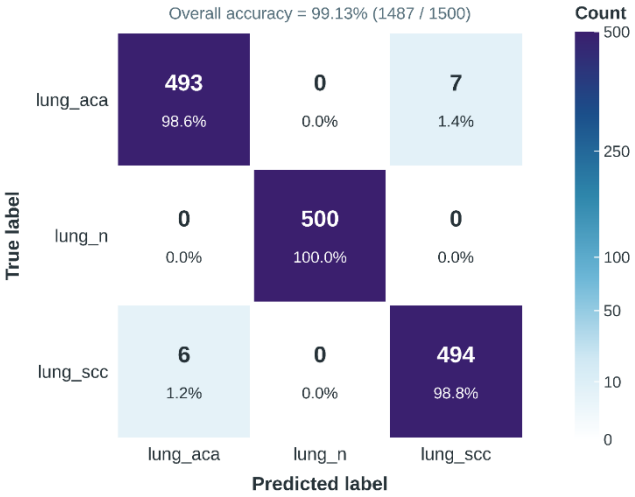

Figure 9. Confusion matrix of the hybrid model without data augmentation.

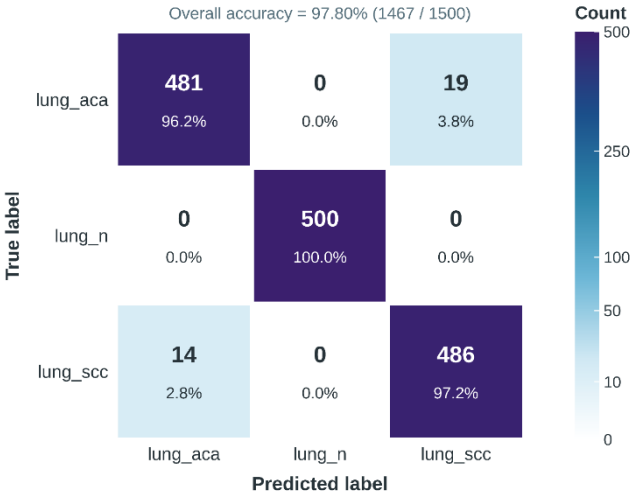

Figure 10. Confusion matrix of the hybrid model with data augmentation.

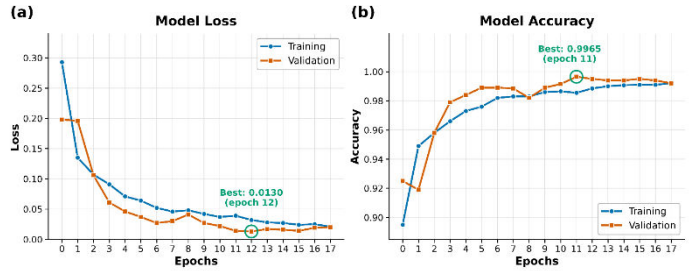

Figure 11. Training and validation accuracy/loss curves of the hybrid model with augmentation.

However, a key observation from this study is that model ranking is not fixed but varies depending on whether data augmentation is applied. In non-augmented settings, EfficientNetB5 performed best, indicating that its compound scaling strategy is highly effective in learning discriminative histopathological features from relatively clean and well-structured datasets. In contrast, under augmented conditions, the hybrid CNN–BiLSTM model achieved the best performance, suggesting that feature fusion and sequential modeling improve robustness when the data distribution becomes more diverse (Shafay *et al.*, 2025). This shift in performance ranking highlights an important methodological insight: model superiority in medical imaging is context-dependent and strongly influenced by preprocessing strategy and data variability rather than architecture alone.

VGG16 showed strong baseline performance in both experimental conditions, confirming its reliability as a classical deep learning architecture for medical image classification. The slight reduction in accuracy after augmentation suggests that while VGG16 benefits from increased data diversity, it may also be sensitive to excessive transformation of fine-grained histopathological features (Hussain *et al.*, 2024). The confusion matrix patterns indicate that VGG16 consistently classified normal lung tissue with perfect or near-perfect accuracy, which reflects strong feature separability between normal and diseased tissues. However, misclassifications between adenocarcinoma and squamous cell carcinoma remained present in both settings. This suggests that these two classes share overlapping morphological characteristics that are difficult to distinguish even with deep convolutional filters (Nayeem *et al.*, 2025; Shahadat *et al.*, 2024). The observed trade-off between accuracy and robustness is consistent with earlier findings that deeper CNNs can overfit to subtle texture-level differences when trained on limited or moderately diverse datasets (Martins *et al.*, 2026).

MobileNetV3-Large demonstrated stable learning behavior and efficient convergence across both experimental settings, reflecting its design goal of computational efficiency. Despite its lightweight structure, it achieved competitive accuracy close to deeper models in the non-augmented scenario, confirming that efficient architectures can still capture meaningful histopathological features (Rahaman *et al.*, 2026). However, the noticeable drop in performance under augmentation indicates that MobileNetV3-Large is more sensitive to aggressive data transformations. The increased misclassification between adenocarcinoma and squamous cell carcinoma suggests that lightweight architectures may have limited capacity to retain fine-grained morphological distinctions when feature distributions are heavily altered (Ochoa-Ornelas *et al.*, 2024). This observation supports prior studies indicating that lightweight CNNs may trade representational capacity for computational efficiency, which can affect their robustness under highly variable training conditions (Raza *et al.*, 2026).

EfficientNetB5 consistently demonstrated strong performance in both settings, particularly in the non-augmented scenario where it achieved the highest overall accuracy. This confirms the effectiveness of compound scaling in balancing network depth, width, and resolution for histopathological feature learning (Abd El-Aziz *et al.*, 2025). The relatively small performance drop under augmentation indicates that EfficientNetB5 maintains strong generalization capability even when exposed to artificially increased data variability. However, the slight increase in misclassification between cancer subtypes suggests that augmentation may introduce subtle feature ambiguity that affects highly sensitive models (Ishaq *et al.*, 2025). These findings are consistent with existing literature where EfficientNet variants have been shown to outperform traditional CNNs in medical imaging due to their optimized scaling strategy (Rahman *et al.*, 2026).

The hybrid CNN–BiLSTM model demonstrated the most interesting behavior among all evaluated architectures. Its superior performance under augmented conditions suggests that combining spatial feature extraction (CNN) with sequential dependency modeling (BiLSTM) enhances the model's ability to generalize in complex feature spaces. The near-perfect AUC value (1.00) indicates that the hybrid model achieves excellent class separability, which is particularly important in medical diagnostics where minimizing false negatives is critical. The early stabilization of validation accuracy further indicates strong learning efficiency and effective feature fusion. This improvement aligns with recent studies suggesting that hybrid architectures can enhance representation learning by capturing complementary feature dimensions beyond traditional CNNs (Hayat *et al.*, 2025). The slight improvement in performance under augmentation also suggests that the hybrid architecture

benefits from increased data diversity more than individual CNN models, likely due to its enhanced capacity to integrate heterogeneous feature representations.

Data augmentation played a crucial role in shaping model generalization behavior across all architectures. While it slightly reduced peak accuracy in some cases, particularly for EfficientNetB5 and MobileNetV3-Large, it consistently improved robustness and reduced overfitting. In non-augmented settings, models exhibited faster convergence and lower training loss, but this was often accompanied by higher sensitivity to training data distribution. In contrast, augmented models showed smoother learning curves and closer alignment between training and validation performance, indicating improved generalization (Wanyonyi *et al.*, 2025). This trade-off between accuracy and generalization is well-documented in deep learning literature, particularly in medical imaging contexts where dataset size and variability are often limited (Goceri, 2023).

From a clinical perspective, the results suggest that deep learning models can provide highly accurate decision-support tools for lung cancer histopathology classification. However, the variability in model performance across augmentation settings highlights the importance of selecting architectures not only based on accuracy but also based on robustness and deployment context. The strong performance of the hybrid model under augmented conditions suggests its potential suitability for real-world clinical environments where data variability is high. Conversely, EfficientNetB5 may be more suitable for controlled or high-quality imaging environments due to its superior performance on clean datasets (Hassan *et al.*, 2026).

6. Conclusions

This study evaluated VGG16, MobileNetV3-Large, EfficientNetB5, and a MobileNetV3Small–BiLSTM hybrid model for lung cancer classification using the LC15000 histopathological dataset under with and without data augmentation settings. All models achieved high diagnostic performance, with EfficientNetB5 performing best without augmentation (accuracy 0.996), while the hybrid model achieved the highest performance under augmentation (accuracy 0.9967, AUC 1.00), indicating strong discriminative capability when spatial and sequential feature learning are combined. Overall, augmentation improved model generalization and reduced overfitting but slightly lowered peak accuracy in some architectures due to increased intra-class variability. EfficientNetB5 proved most effective for high-precision standalone classification, whereas the hybrid model showed superior robustness and generalization in augmented conditions; therefore, a hybrid CNN-based framework with optimized augmentation is recommended for practical histopathological lung cancer diagnosis.

Acknowledgments

The author would like to express his sincere gratitude to Department of Computer Science and Engineering, Green University of Bangladesh for the logistic support to conduct the study.

Funding information

No external or internal fund was received to conduct the study.

Ethical approval statement

Not applicable.

Data availability

All available data are presented in the article.

Informed consent statement

Not applicable.

Conflict of interest

The authors declare no conflict of interest.

Authors' contribution

Ishmat Jahan: Conceptualization, formal analysis, writing-original draft preparation, review and editing; **Md. Jubayar Hossain Jamil:** formal analysis, writing-original draft preparation, review and editing; **Khandaker Farhan Labib:** formal analysis, writing-original draft preparation, review and editing. The authors has read and approved the final version of the published editorial.

References

Abd El-Aziz AA, Mahmood MA and Abd El-Ghany S, 2024. Advanced deep learning fusion model for early multi-classification of lung and colon cancer using histopathological images. Diagnostics, 14(20): 2274. https://doi.org/10.3390/diagnostics14202274

Abd El-Aziz AA, Mahmood MA and El-Ghany SA, 2025. EfficientNet-B3-based automated deep learning framework for multiclass endoscopic bladder tissue classification. Diagnostics, 15(19): 2515. https://doi.org/10.3390/diagnostics15192515

Abraham LA, Palanisamy G and Veerapu G, 2025. Transparent brain tumor detection using DenseNet169 and LIME. Scientific Reports, 15: 28185. https://doi.org/10.1038/s41598-025-13233-7

Ahmad IS, Dai J, Xie Y and Liang X, 2025. Deep learning models for CT image classification: a comprehensive literature review. Quantitative Imaging in Medicine and Surgery, 15: 962–1011. https://doi.org/10.21037/qims-24-1400

Ali TE, Ali FI, Dakić P and Zoltan AD, 2025. Trends, prospects, challenges, and security in the healthcare internet of things. Computing, 107: 28. https://doi.org/10.1007/s00607-024-01352-4

Arshad MF, Chowdhury AT, Sharif Z, Islam MS Bin, Sumon MSI, Mohammedkasim A, Chowdhury MEH and Pedersen S, 2025. Artificial intelligence and machine learning in lung cancer: Advances in imaging, detection, and prognosis. Cancers, 17(24): 3985. https://doi.org/10.3390/cancers17243985

Attallah O, 2025. Lung and colon cancer classification using multiscale deep features integration of compact convolutional neural networks and feature selection. Technologies, 13(2): 54. https://doi.org/10.3390/technologies13020054

Davri A, Birbas E, Kanavos T, Ntritsos G, Giannakeas N, Tzallas AT and Batistatou A, 2023. Deep learning for lung cancer diagnosis, prognosis and prediction using histological and cytological images: A systematic review. Cancers, 15(15): 3981. https://doi.org/10.3390/cancers15153981

Durgam R, Panduri B, Balaji V, Khadidos AOAO, Khadidos AOAO and Selvarajan S, 2025. Enhancing lung cancer detection through integrated deep learning and transformer models. Scientific Reports, 15: 15614. https://doi.org/10.1038/s41598-025-00516-2

Ennab M and Mcheick H, 2025. Advancing AI interpretability in medical imaging: A comparative analysis of pixel-level interpretability and Grad-CAM models. Machine Learning and Knowledge Extraction, 7: 12. https://doi.org/10.3390/make7010012

Faria N, Campelos S and Carvalho V, 2025. Enhancing histopathological lung cancer diagnosis through explainable deep learning models: A methodological framework proposal. Information, 16(9): 740. https://doi.org/10.3390/info16090740

Fuster-Matanzo A, Picó-Peris A, Bellvís-Bataller F, Jimenez-Pastor A, Weiss GJ, Martí-Bonmatí L, Lázaro Sánchez A, Bazaga D, Banna GL, Addeo A, Camps C, Seijo LM and Alberich-Bayarri Á, 2025. Prediction of oncogene mutation status in non-small cell lung cancer: a systematic review and meta-analysis with a special

focus on artificial intelligence-based methods. European Radiology, 36(3): 2157–2185. https://doi.org/10.1007/s00330-025-11962-x

Garg S and Garg S, 2020. Prediction of lung and colon cancer through analysis of histopathological images by utilizing Pre-trained CNN models with visualization of class activation and saliency maps. 2020 3rd Artificial Intelligence and Cloud Computing Conference, 38–45. https://doi.org/10.1145/3442536.3442543

Gayap HT and Akhloufi MA, 2024. Deep machine learning for medical diagnosis, application to lung cancer detection: A review. BioMedInformatics, 4: 236–284. https://doi.org/10.3390/biomedinformatics4010015

Goceri E, 2023. Medical image data augmentation: techniques, comparisons and interpretations. Artificial Intelligence Review, 56(11): 12561–12605. https://doi.org/10.1007/s10462-023-10453-z

Großerueschkamp F and Gerwert K, 2020. Label-free digital pathology by infrared imaging. Biomedical Spectroscopy and Imaging, 9(1–2): 5–12. https://doi.org/10.3233/BSI-200196

Habbab FM, Bédard ELR, Joy AA, Alam Z, Abraham AG and Roa WHY, 2025. Early detection of lung cancer: A review of innovative milestones and techniques. Journal of Clinical Medicine, 14(21): 7812. https://doi.org/10.3390/jcm14217812

Hassan NMH, Bayoumy AS and Mahmoud MHM, 2026. Hybrid EfficientNet B4 and SVM framework for rapid and accurate bone cancer diagnosis from X-rays. Scientific Reports, 16: 8156. https://doi.org/10.1038/s41598-026-38801-3

Hayat MT, Allawi YM, Alamro W, Sultan SM, Abadleh A, Kang H and Zreikat AI, 2025. A Hybrid Convolutional Neural Network–Long Short-Term Memory (CNN–LSTM)–attention model architecture for precise medical image analysis and disease diagnosis. Diagnostics, 15(21): 2673. https://doi.org/10.3390/diagnostics15212673

Henna S, Alcaraz JML, Rathnayake U and Amjath M, 2025. An interpretable deep learning framework for medical diagnosis using spectrogram analysis. Healthcare Analytics, 8: 100408. https://doi.org/10.1016/j.health.2025.100408

Hussain M, Thaher T, Almourad MB and Mafarja M, 2024. Optimizing VGG16 deep learning model with enhanced hunger games search for logo classification. Scientific Reports, 14: 31759. https://doi.org/10.1038/s41598-024-82022-5

Ishaq A, Ullah FUM, Hamandawana P, Cho DJ and Chung TS, 2025. Improved efficientNet architecture for multi-grade brain tumor detection. Electronics, 14(4): 710. https://doi.org/10.3390/electronics14040710

Lerch L, Huber LS, Kamath A, Pöllinger A, Pahud de Mortanges A, Obmann VC, Dammann F, Senn W and Reyes M, 2024. DreamOn: a data augmentation strategy to narrow the robustness gap between expert radiologists and deep learning classifiers. Frontiers in Radiology, 4: 1420545. https://doi.org/10.3389/fradi.2024.1420545

Li C, Lei S, Ding L, Xu Y, Wu X, Wang H, Zhang Z, Gao T, Zhang Y and Li L, 2023. Global burden and trends of lung cancer incidence and mortality. Chinese Medical Journal, 136(13): 1583–1590. https://doi.org/10.1097/CM9.0000000000002529

Lin C, Yang P, Wang Q, Qiu Z, Lv W and Wang Z, 2023. Efficient and accurate compound scaling for convolutional neural networks. Neural Networks, 167: 787–797. https://doi.org/10.1016/j.neunet.2023.08.053

Martins OO, Oosthuizen CC and Desai DA, 2026. A robust fusion–pruning framework for lightweight CNNs: Application to apple leaf disease diagnosis under real-world image corruptions. Results in Engineering, 30: 110612. https://doi.org/10.1016/j.rineng.2026.110612

Mohalder RD, Hossain KA, Sarkar JP, Paul L, Raihan M and Talukder KH, 2023. Lung cancer detection from histopathological images using deep learning. In: Satu MS, Moni MA, Kaiser MS and Arefin MS (Eds.), Machine Intelligence and Emerging Technologies. MIET 2022. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 490 (pp. 201–212). Springer. https://doi.org/10.1007/978-3-031-34619-4_17

Nayeem MD, Shishir MEH, Rahman MT, Juwel ZC, Sutradhar S, Chaki S, Rahman MS and Ali ABMS, 2025. A hybrid architecture with separable convolutions and attention for lung and colon cancer detection. Array, 28: 100591. https://doi.org/10.1016/j.array.2025.100591

Ochoa-Ornelas R, Gudiño-Ochoa A and García-Rodríguez JA, 2024. A hybrid deep learning and machine learning approach with mobile-efficientNet and grey wolf optimizer for lung and colon cancer histopathology classification. Cancers, 16(22): 3791. https://doi.org/10.3390/cancers16223791

Ochoa-Ornelas R, Gudiño-Ochoa A, García-Rodríguez JA and Uribe-Toscano S, 2025. A robust transfer learning approach with histopathological images for lung and colon cancer detection using EfficientNetB3. Healthcare Analytics, 7: 100391. https://doi.org/10.1016/j.health.2025.100391

Pan L, Li X, Dou Y, Song Q, Luo J, Liang Q and Peng S, 2025. SMILE: A scale-aware multiple instance learning method for multicenter STAS lung cancer histopathology diagnosis. Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence, 9836–9844. https://doi.org/10.24963/ijcai.2025/1093

Pan L, Luo J, Nie C, Fan S, Wang X, Peng S and Liang Q, 2026. A deep learning framework integrating tumor microenvironmental features accurately predicts multiple driver gene mutations in lung cancer pathology images. Cancer Research, 86(5): 1319–1337. https://doi.org/10.1158/0008-5472.CAN-25-0582

Pang J, Zhang H, Zhao H and Li L, 2022. DcsNet: a real-time deep network for crack segmentation. Signal, Image and Video Processing, 16(4): 911–919. https://doi.org/10.1007/s11760-021-02034-w

Parra-Medina R, Guerron-Gomez G, Mendivelso-González D, Gil-Gómez JH, Alzate JP, Gomez-Suarez M, Polo JF, Sprockel JJ and Mosquera-Zamudio A, 2025. Deep learning in histopathology images for prediction of oncogenic driver molecular alterations in lung cancer: a systematic review and meta-analysis. Translational Lung Cancer Research, 14(5): 1756–1769. https://doi.org/10.21037/tlcr-2024-1196

Rahaman J, Paul P, Chowdhury A and Quamruzzaman M, 2026. A customized MobileNetV3Large-based deep learning framework for plant disease detection. Discover Artificial Intelligence, 6: 110. https://doi.org/10.1007/s44163-025-00733-8

Rahman MM, Hasanuzzaman ABM, Chisty MA, Alam E, Islam MK and Islam ARMT, 2025. Perceived-air pollution and self-reported health status: a study on air pollution-prone urban area of Bangladesh. Frontiers in Public Health, 13: 1382471. https://doi.org/10.3389/fpubh.2025.1382471

Rahman T, Deb N, Larguech S, Moniruzzaman M, Kumer Ghosh A, Ahmed Sami A, Abd Aziz AB and Salem Al-Bawri S, 2026. Optimized CNN framework with VGG19, EfficientNet, and Bayesian optimization for early colon cancer detection. Scientific Reports, 16: 4128. https://doi.org/10.1038/s41598-025-34262-2

Rayed ME, Islam SMS, Niha SI, Jim JR, Kabir MM and Mridha MF, 2024. Deep learning for medical image segmentation: State-of-the-art advancements and challenges. Informatics in Medicine Unlocked, 47: 101504. https://doi.org/10.1016/j.imu.2024.101504

Raza A, Hanif F and Mohammed HA, 2026. Clinical validation of lightweight CNN architectures for reliable multi-class classification of lung cancer using histopathological imaging techniques. Scientific Reports, 16: 6512. https://doi.org/10.1038/s41598-026-36652-6

Shafay M, Hassan T, Owais M, Hussain I, Khawaja SG, Seneviratne L and Werghi N, 2025. Recent advances in plant disease detection: challenges and opportunities. Plant Methods, 21: 140. https://doi.org/10.1186/s13007-025-01450-0

Shahadat N, Lama R and Nguyen A, 2024. Lung and colon cancer detection using a deep AI model. Cancers, 16(22): 3879. https://doi.org/10.3390/cancers16223879

Shweikeh E, Al-Rajab M, Lu J, Xu Q, Joy M, Ahmed A and Chang H, 2026. A deep learning model to enhance lung cancer detection using 'Dual-Branch' model classification approach. PLOS One, 21: e0339404. https://doi.org/10.1371/journal.pone.0339404

Smrity TA, Muntaqim MZ and Kafi HM, 2026. Recent advancements in machine learning and deep learning for early detection of breast cancer: A comprehensive review. Innovative Practice in Breast Health, 8: 100050. https://doi.org/10.1016/j.ibreh.2026.100050

Talukder MA, Islam MM, Uddin MA, Akhter A, Hasan KF and Moni MA, 2022. Machine learning-based lung and colon cancer detection using deep feature extraction and ensemble learning. Expert Systems with Applications, 205: 117695. https://doi.org/10.1016/j.eswa.2022.117695

Thandra KC, Barsouk A, Saginala K, Aluru JS and Barsouk A, 2021. Epidemiology of lung cancer. Współczesna Onkologia, 25: 45–52. https://doi.org/10.5114/wo.2021.103829

Tu J, Tao S, Zhao LK, Yin FQ, Qi ZY, Gao TY, Meng L, Mao M and Zhang C, 2026. Comprehensive evaluation of worldwide lung cancer burden with epidemiological forecasts through mid-century. Critical Public Health, 36: 2659417. https://doi.org/10.1080/09581596.2026.2659417

Uddin AH, Chen YL, Akter MR, Ku CS, Yang J and Por LY, 2024. Colon and lung cancer classification from multi-modal images using resilient and efficient neural network architectures. Heliyon, 10(9): e30625. https://doi.org/10.1016/j.heliyon.2024.e30625

Wang M, Kim RY, Kohonen-Corish MRJ, Chen H, Donovan C and Oliver BG, 2025. Particulate matter air pollution as a cause of lung cancer: epidemiological and experimental evidence. British Journal of Cancer, 132(11): 986–996. https://doi.org/10.1038/s41416-025-02999-2

Wanyonyi M, Morris ZN, Musyoka FM and Kitavi DM, 2025. Enhanced machine learning and hybrid ensemble approaches for Coronary Heart Disease prediction. PLOS One, 20(12): e0328338. https://doi.org/10.1371/journal.pone.0328338

Wen Y, Chen L, Deng Y and Zhou C, 2021. Rethinking pre-training on medical imaging. Journal of Visual Communication and Image Representation, 78: 103145. https://doi.org/10.1016/j.jvcir.2021.103145

Wu Z, Xia F and Lin R, 2024. Global burden of cancer and associated risk factors in 204 countries and territories, 1980–2021: a systematic analysis for the GBD 2021. Journal of Hematology & Oncology, 17: 119. https://doi.org/10.1186/s13045-024-01640-8

Yang H, Chen L, Cheng Z, Yang M, Wang J, Lin C, Wang Y, Huang L, Chen Y, Peng S, Ke Z and 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: 80. https://doi.org/10.1186/s12916-021-01953-2

Zhu E, Muneer A, Zhang J, Xia Y, Li X, Zhou C, Heymach JV, Wu J and Le X, 2025. Progress and challenges of artificial intelligence in lung cancer clinical translation. Npj Precision Oncology, 9: 210. https://doi.org/10.1038/s41698-025-00986-7

Publisher's note
Genesis Publishing Consortium Limited pledges to maintain a neutral stance on jurisdictional claims shown in published maps and institutional affiliations.