http://dx.doi.org/10.14500/aro.12355

A Hybrid Deep Learning Model with Self-Attention for the Classification of Lung Cancer Using Histopathology Image

Lana L. Nahmatwlla[1†] and Abbas M. Ali[2]

[1]Department of Computer Science and Information Technology, College of Science, Salahaddin University, Erbil, Kurdistan Region – F.R. Iraq

[2]Department of Software and Informatics Engineering, College of Engineering, Salahaddin University, Erbil, Kurdistan Region – F.R. Iraq

***Abstract*—Lung cancer remains a prevalent health burden and is one of the leading causes of cancer mortality worldwide. Its high mortality rate is partly attributable to histological heterogeneity and the difficulty of detecting it at early stages. An accurate distinction of lung cancer subtypes in histopathological images is crucial for improving the accuracy of diagnosis and planning an appropriate treatment to improve the quality of life of patients. This study proposes a hybrid deep-learning model for classifying cancer types using histopathology images. The ConvNeXt-Tiny is an extension of the ResNet-50 base architecture. This architecture is inspired by both models and introduces self-attention layers to improve both feature extraction and classification performance, leading to a unique model design. The proposed model and two other deep learning models were trained and tested using the public Lung and Colon Cancer Histopathological Image (LC25000) dataset and a private clinical dataset, and their effectiveness was evaluated. The proposed model outperformed the best classification accuracy among the other architectures (98.73% public and 93.17% private), outperforming baseline models, such as ConvNeXt-Tiny (96.27% public and 89.33% private) and ResNet-50 (94.00% public and 87.67% private). The results confirm the robustness and generalization ability of the proposed architecture.**

***Index Terms*—Deep learning, Histopathology images, Lung cancer classification, Model evaluation, Primary lung cancer.**

I. Introduction

Lung cancer continues to possess one of the highest ranks of cancer-related deaths in the world, and histopathological study is still considered the gold standard of diagnosis (Siegel, Giaquinto and Jemal, 2024). Primary lung cancer is divided based on cell type. Non-small cell lung cancer (NSCLC) is the first kind, which can be further subtyped into squamous cell carcinoma (SCC), large cell carcinoma (LCC), and adenocarcinoma (ADC). The second type is small-cell lung cancer (SCLC) (Howlader, et al., 2020). The purpose of our analysis was to identify distinguishing features for primary lung cancers. Histopathology remains important in the accurate diagnosis of these cancers, making it an essential tool for categorizing tumors by cellular morphology and architecture on examination of lung tissue under a microscope (Nicholson, et al., 2022).

While being the most dependable form of diagnosis, it is also an arduous and subjective manual histological evaluation process. For the same sample, a second review by a pathologist might lead to ambiguous conclusions, particularly in the case of complex or borderline samples. The procedure itself is time-consuming and liable to human mistakes (Komura and Ishikawa, 2018). These limitations have created an urgent need for technological solutions to improve the consistency and efficiency of pathological assessment.

Convolutional neural networks (CNN) have played a transformative role in histopathological image classification (Hua, Li, and Wang, 2024). Architectures, such as ResNet have been widely adopted due to their residual learning design, which facilitates the training of deeper networks and enhances feature extraction from complex tissue structures. ResNet-based models have demonstrated strong performance in classification and patch-level analysis tasks. However, CNNs, such as ResNet are fundamentally limited by their localized receptive fields, which restrict their ability to capture long-range spatial dependencies – an essential capability when interpreting heterogeneous histopathological slides (Kanavos and Mylonas, 2023).

ConvNext-Tiny was one of the models proposed to combine a backbone CNN model while leveraging the architectural design, parameter optimization, and training strategy of vision transformers, such as Swin-T. The ConvNext-Tiny model has outperformed most state-of-the-art CNN and transformer-based models in various computer vision tasks and can be easily scaled up. The ConvNeXt-

ARO-The Scientific Journal of Koya University
Vol. XIII, No. 2(2025), Article ID: ARO.12355. 8 pages
DOI: 10.14500/aro.12355
Received: 13 June 2025; Accepted: 21 August 2025
Regular research paper; Published: 01 September 2025
†Corresponding author's e-mail: lana.nahmatwlla@su.edu.krd

Tiny uses ResNet-50 as the baseline and optimizes its architecture and training strategy using Swin-T transformers (Liu, et al., 2022). Despite adopting the architecture and parameter optimization strategies of vision transformers, ConvNext-Tiny fails to incorporate and leverage one of the core components of the transformer architecture, which is the self-attention mechanism. This study aims to break through these constraints by developing a hybrid deep-learning model for classifying lung cancer in histopathological images.

The proposed model is designed to draw on selected ideas from ResNet50 and ConvNext-Tiny and introduce self-attention. The self-attention mechanism offers an effective and better cross-channel feature correlation that is useful for the overall improvement of feature extraction. The model's generalizability was assessed using two separate datasets. Results are compared with two baseline architectures to substantiate the model's performance in classifying between different lung cancer categories. This paper's remainder is structured as follows. The second section presents the literature review. The third section outlines the methodology adopted in this study. The fourth section reports and briefly discusses the results. The final section presents the study's conclusion.

II. Literature Review

To explore the classification of lung cancer using histopathology images and deep learning, we reviewed relevant studies from reputable journals. The literature is synthesized in two complementary forms: Key studies are discussed in paragraph format to outline the deep learning models, datasets, and classification outcomes, while Table I summarizes additional works by architecture, data source, and performance.

Understanding the present situation, identifying research gaps, and guiding the development of our suggested model are the goals of this review. Wei, et al. (2019) proposed a ResNet-based model for classifying five lung ADC subtypes from whole-slide histopathology images, achieving a kappa score of 0.525 and 66.6% agreement with pathologists, surpassing the inter-pathologist rate of 62.7%.

Ahmed, et al. (2021) applied transfer learning using Inception-V3 and VGG-16 on the Kimia Path24 dataset, which covers 24 tissue types, reporting accuracies of 80% and 77%, respectively. Yang, et al. (2021) utilized EfficientNet-B5 for multiclass classification across the SYSUFH, SZPH, and TCGA datasets, achieving a precision of up to 99%, recall of up to 100%, and accuracy ranging from 86% to 89%.

Mehmood, et al. (2022) utilized a modified AlexNet on LC25000 with 227 × 227 inputs, achieving an accuracy of 98.8%. Ijaz, et al. (2023) combined features from ResNet-50 and EfficientNet-B0 with the Grey Wolf Optimization algorithm, achieving 98.73% accuracy through soft voting. Tummala, et al. (2023) applied EfficientNetV2 with an input resolution of 224 × 224, achieving 99.97% accuracy.

El-Ghany, et al. (2023) fine-tuned ResNet-101 on LC25000, achieving 99.94% accuracy with precision, recall, and F1-score around 99.84%Tortora, et al. (2023) developed a multimodal framework integrating CT, histopathology, and clinical data to predict NSCLC survival, achieving an AUC of 0.909 using handcrafted features and Random Forest. Gowthamy and Ramesh (2024) introduced a two-stage

TABLE I
Summary of Existing Methods for Lung Cancer Classification

Study	Input image size	Dataset	Pre-trained (Yes/No)	Methodology	Accuracy (%)
Hatuwal and Thapa (2020)	180×180	LC25000	No	Custom CNN	97.20
Masud, et al. (2021)	64×64	LC25000	Not specified	Custom CNN	96.33
Kumar, et al. (2022)	224×224	LC25000	Yes (ImageNet)	DenseNet121 and RF	98.60
Hattori, et al. (2023)	800×800	WSI.	No	Switching Discriminator	90.9
Rajasekar, et al. (2023)	224×224	histopathology image	Yes (ImageNet)	CNN	93.64
				CNN GD	97.86
				VGG-16	96.52
				Inception V3	93.54
				VGG-19	92.17
				ResNet-50	93.47
Anjum, et al. (2023)	224×224	LC25000	Yes (ImageNet)	EfficientNetB0	95.87
	240×240			EfficientNetB1	96.26
	260×260			EfficientNetB2	97.24
	300×300			EfficientNetB3	95.63
	380×380			EfficientNetB4	96.83
	456×456			EfficientNetB5	94.31
	512×512			EfficientNetB6	93.76
	600×600			EfficientNetB7	95.59
Sumon, et al. (2024)	224×224	LC25000	Yes (ImageNet)	Deep learning+SVM approach for classification	96.6
Gong, et al. (2025)	224×224	739 Diff-Quik–stained cytology samples (6 classes)	Yes (ImageNet)	ResNet-18 (enhanced)	84.05 (AI only) 90.52–93.83 (AI+experts)
Kriegsmann, et al. (2020)	299×299	Quality-controlled histopathology patches	Yes (ImageNet)	InceptionV3	95
Devi, et al. (2024)	300×300	LC25000	No	Simple CNN	85

ensemble that combines multiple CNNs with Electric Eel Optimization; achieving 98.96% accuracy on LC25000 using weighted averaging. Table I presents additional studies relevant to the classification of lung cancer.

While existing studies using public datasets, such as LC25000, have shown promising results, they often lack clinical diversity, do not encompass all primary lung cancer subtypes, and rely on artificial augmentation. Moreover, no prior research has focused on histopathological lung cancer classification using clinical data from the Kurdistan region. To address these limitations, this study introduces a hybrid deep-learning model developed to improve classification performance on both public and private clinical datasets.

III. Materials and Methods

The development of the lung cancer classification model is outlined in this section. Both public and private clinical image datasets are utilized in the study. Two baseline architectures were also introduced for performance comparison in addition to the proposed model. Models were assessed using the standard metrics, i.e., accuracy, precision, recall, and F1-score.

A. Dataset

The present work is based on two sets of data: Public and private clinical data. The public dataset was directly used for training and testing models, while the private dataset was clinically collected and pre-processed for model building.

Public dataset

The proposed lung cancer classification model and the baseline models were developed and tested based on the LC25000. A total of 25,000 histopathological images of five lung and colon tissue classes are included. Three classes of lungs were selected for our study: ADC, SCC, and normal tissue (NT). There are 5000 color images per class, generated by augmenting a set of 250 original images with rotation, zoom, and contrast modifications. The dataset is publicly accessible on Kaggle and is widely used in computer-aided diagnosis research (Borkowski, et al., 2019). Fig. 1 shows example images in the dataset.

Private dataset

This research collected a custom histopathology image dataset from biopsies obtained at PAR Hospital in Erbil, Kurdistan Region, Iraq. The dataset includes four lung cancer subtypes: ADC, SCC, LCC, and SCLC. The private dataset focuses on primary lung cancer types, excluding normal lung tissue, as the study aimed to address the more challenging task of distinguishing between different primary lung cancer subtypes. The histological slides were hand-stained by the pathology laboratory staff at the hospital using standard hematoxylin and eosin staining protocols. The images were captured at ×40 magnification using a standard light microscope. Data were collected with hospital permission and ethical approval to use anonymized patient samples. The approval letter has been submitted as a supplementary file for editorial review.

The original images (2992 × 2992 pixels) were resized to a standard resolution of 712 × 712 pixels to keep fine morphological details. These patches were then resized to the network input size for model training. No stain normalization or noise removal was applied to the histopathology images to maintain the original tissue appearance and ensure that inherent morphological features were preserved for accurate analysis, as detailed in Table II and stored in JPG format for compatibility with deep learning frameworks. Fig. 2 presents representative examples from each of the lung cancer subtypes. To ensure balanced class representation and reduce bias during model training, 1500 image patches were selected per class.

The data were split into three parts: A training set (70%), a validation set (20%), and a testing set (10%). Data augmentation, including flipping, rotation, zooming, brightness adjustment, and color jittering, was applied only to the training set to enhance generalization and simulate histological variability.

B. Proposed Model

The proposed architecture draws inspiration from the residual structure of ResNet-50 and the ConvNext-Tiny block design (Liu, et al., 2022), using their strengths in a standalone model. We revisited the popular ResNet-50 and made vital modifications to the original architecture. One of the most visible additions in our proposal is the inclusion of three self-attention layers between the network's second and fourth stages as depicted in Fig. 3. The proposed model integrates a self-attention mechanism within the CNN pipeline, where the input features are processed. At this point, features are forwarded through two parallel branches. One branch performs bottleneck-like convolutional blocks to capture local spatial information. At the same time, the other takes a self-attention mechanism to capture global dependencies. The two branches are concatenated together and supplied to the rest of the CNN layers. Thereafter, a global average pooling layer is applied to aggregate spatial information into a compact representation, which is then forwarded to a fully connected layer within the classification head.

The purpose of these layers is to utilize the cross-stage and cross-channel to learn correlated features for a better

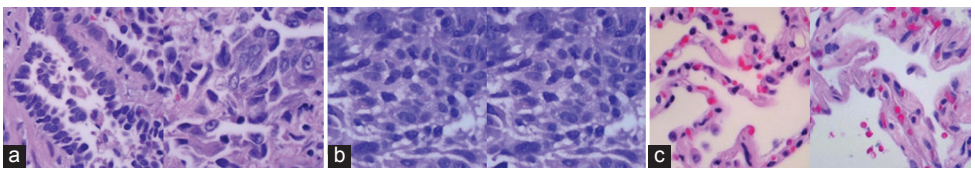

Fig. 1. Sample histopathology images from the LC25000 dataset: (a) Adenocarcinoma, (b) Squamous cell carcinoma, and (c) Normal lung tissue.

TABLE II
SUMMARY OF PRIVATE DATASET

Cancer type	No. of patients	No. of slides	Average slide per patient	Average images per slide	Average patches per image	Total patches	Total patches used for each class
Adenocarcinoma	38	50	2	5	7	1750	1500
Squamous cell carcinoma	35	48	2	5	7	1680	1500
Large cell carcinoma	25	43	2	5	7	1505	1500
Small cell carcinoma	31	47	2	5	7	1645	1500

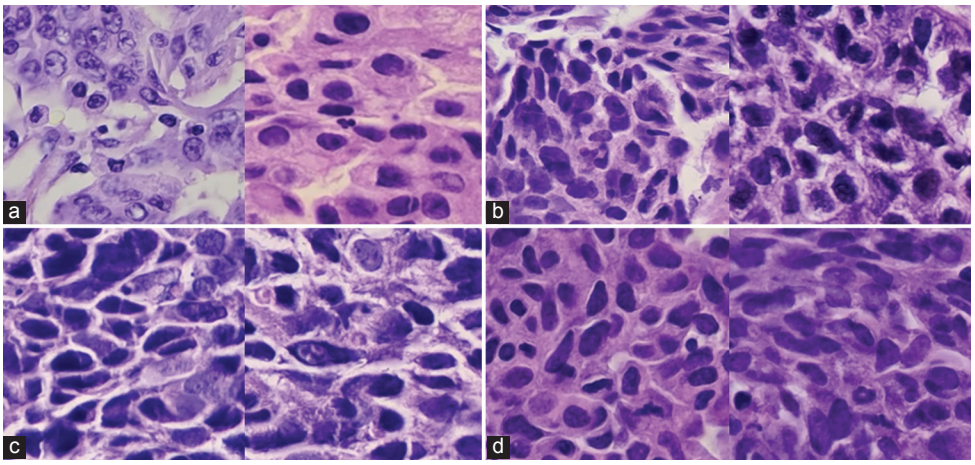

Fig. 2. Sample histopathology images from the private clinical dataset: (a) Adenocarcinoma, (b) Squamous cell carcinoma, (c) Large cell carcinoma, and (d) Small cell lung cancer.

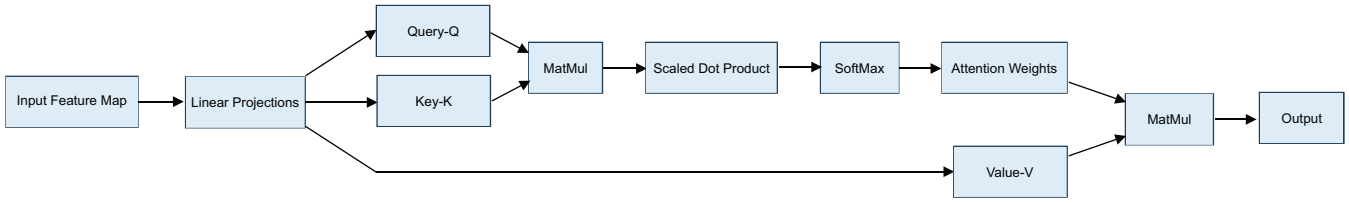

Fig. 3. Schematic diagram of the self-attention mechanism.

representation of the input feature map. As a departure from the ResNet-50, self-attention enables neural networks to dynamically highlight semantically important regions in an image while reducing the influence of irrelevant information (Khan, et al., 2022). It projects the input feature map into Query, Key, and Value vectors to compute attention scores using scaled dot-product attention, followed by softmax normalization as shown in Equations 1-4 (Raschka, Mirjalili and Raschka, 2022).

$$Q = XW^{Q},\ K = XW^{K},\ V = XW^{KV} \quad (1)$$

$$Score = \frac{QK^{T}}{\sqrt{d_k}} \quad (2)$$

$$A = softmax\ (Score) \quad (3)$$

$$Output = A.V \quad (4)$$

Here, X, denotes the input feature map; W^{Q}, W^{K} and W^{V} are learnable weight matrices used to project the input into the query (Q), key (K), and value (V), representations; d_k, is the dimensionality of the key vectors, used as a scaling factor in the attention score computation; and A represents the attention weight matrix that determines the contribution of each value vector V to the final output. As shown in Fig. 4, the diagram provides a detailed illustration of the self-attention mechanism.

The stage ratio of the proposed model has been arranged in a 1:1:3:1 compute ratio for efficient computation and performance improvement, as opposed to the stage compute ratio of ResNet-50, which is (3, 4, 6, 3). For each stage, the convolution process is repeated in a number of 3, 3, 9, and 3 before taking the global averaging of the overall feature. Moreover, the larger filter sizes of 5, 7, 9, and 11 are used in the stages instead of the fixed 3 × 3 size used in the original ResNet-50. This choice was made to capture both fine-grained cellular details and larger tissue structures, enabling the network to learn multiscale features important for histopathology image classification. This choice is also inspired by the ConvNet-Tiny and Swin-T transformer architecture.

Another micro-design change is the activation function and its frequency of use in ResNet-50 instead of ReLU, the GELU activation function is used, providing a smoother output that helps improve the flow of gradients during training and results in improved performance in terms of final accuracy on the validation set, as seen from

http://dx.doi.org/10.14500/aro.12355

the higher validation accuracy achieved in comparative experiments. We preserve the batch normalization of the original architecture rather than the layer normalization proposed in ConNext-Tiny. A detailed summary of the training parameters and techniques is provided in Table III. The hyperparameters were rigorously optimized through systematic grid search and evaluation on the validation set to ensure stable training and maximize model performance. These settings were selected based on their consistent improvement in validation metrics across multiple experiments. The proposed model and both the baseline models were implemented using Python with the TensorFlow library. Training was carried out for 200 epochs on an Intel i9 (3.4GHz) CPU system along with an NVIDIA RTX 4090 GPU. To make sure we are making a fair comparison, all models were evaluated under the same experimental conditions.

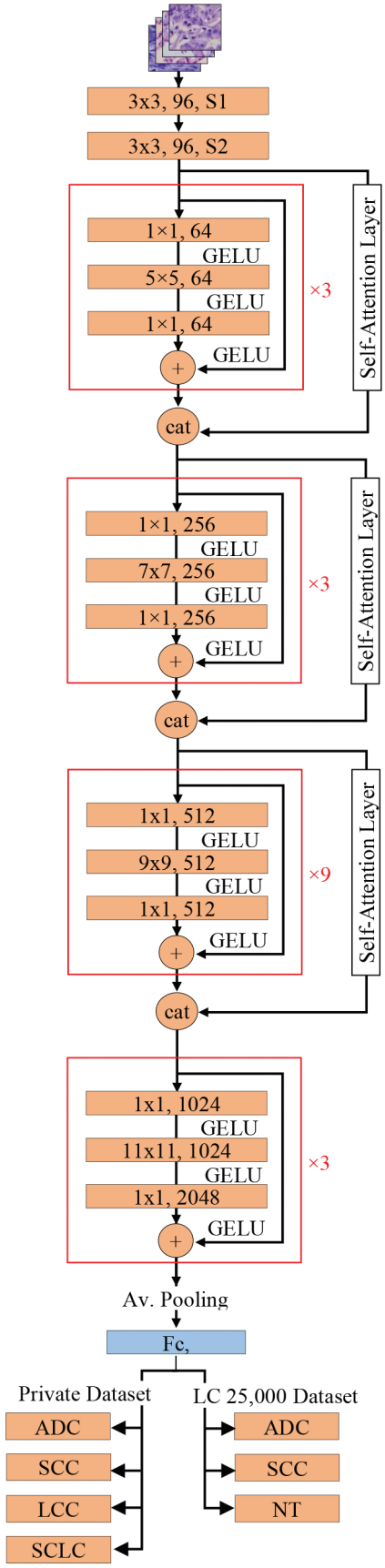

Fig. 4. Architecture of the proposed deep learning model.

IV. Results and Discussion

This section evaluates the proposed model against two baselines, ConvNext-Tiny and ResNet-50, on the LC25000 dataset and private datasets. It includes training curve analysis for the proposed model, confusion matrices for all models, separate performance tables, and a comparative discussion of results

A. Training and Validation Performance

Training and validation performance plots of the proposed model on the public and private datasets are illustrated in Figs. 5 and 6, respectively. One can observe stable convergence for 200 epochs in both figures, accompanied by a consistent reduction in loss and an improvement in validation accuracy. Training is smooth in the public dataset, and although the private one has slight fluctuations, it tends to stabilize. The downward trend of the loss demonstrates effective learning and limited overfitting. Superior accuracy was obtained on the LC25000 dataset, which was augmented before splitting. On the other hand, the training dataset of the private was first split, and augmentation was performed on the training data only. Validation and testing were conducted in the same manner as those without augmentation, which also contributed to its greater variability and made it more clinically realistic, albeit more challenging.

B. Confusion Matrix

To evaluate classification performance, Figs. 7 and 8 present the confusion matrices of the proposed model, ConvNext-

TABLE III
Training Hyperparameters

Image size	224×224
Optimizer	AdamW
Learning Rate	5e-5
Weight decay	0.05
Optimizer momentum	β_1=0.9, β_2=0.999
Batch size	64
Training epoch	200
Learning Rate schedule	Cosine decay
Random erasing	0.2

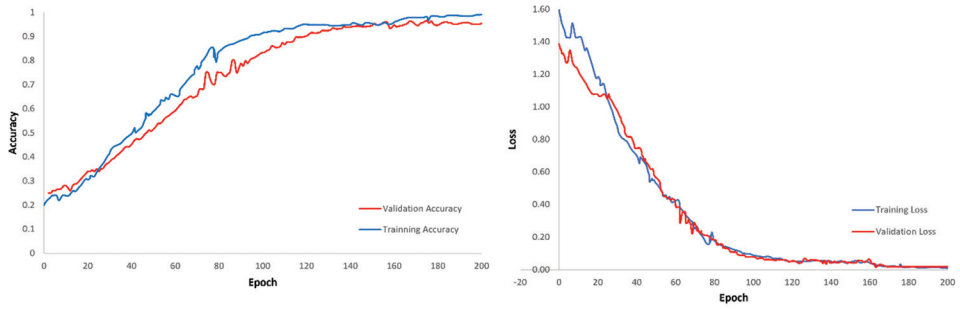

Fig. 5. The accuracy and loss curve of the proposed model on the LC25000 dataset.

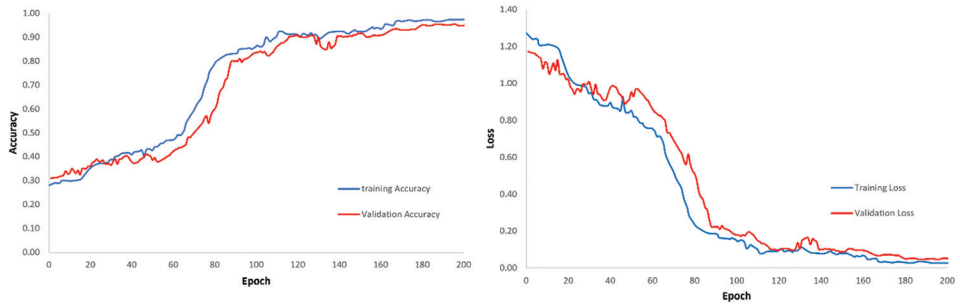

Fig. 6. The accuracy and loss curve of the proposed model on the private dataset.

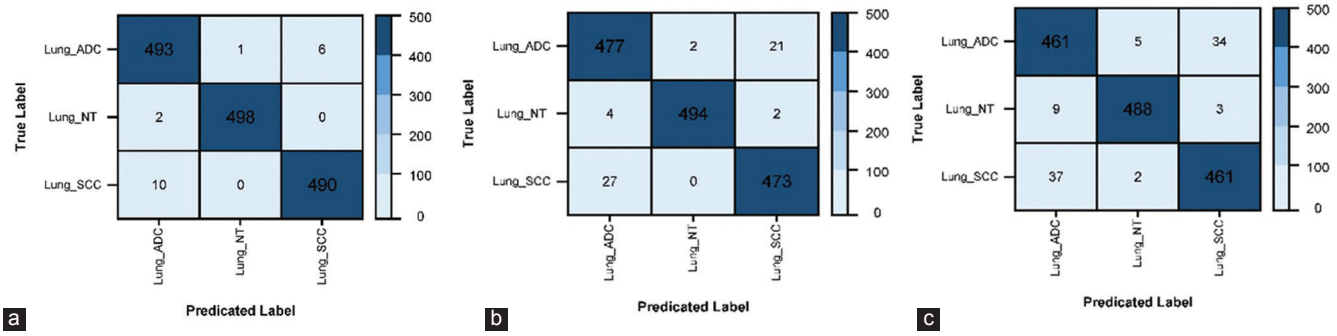

Fig. 7. Confusion matrices on the LC25000 test set: (a) Proposed model, (b) ConvNext-Tiny, and (c) ResNet-50.

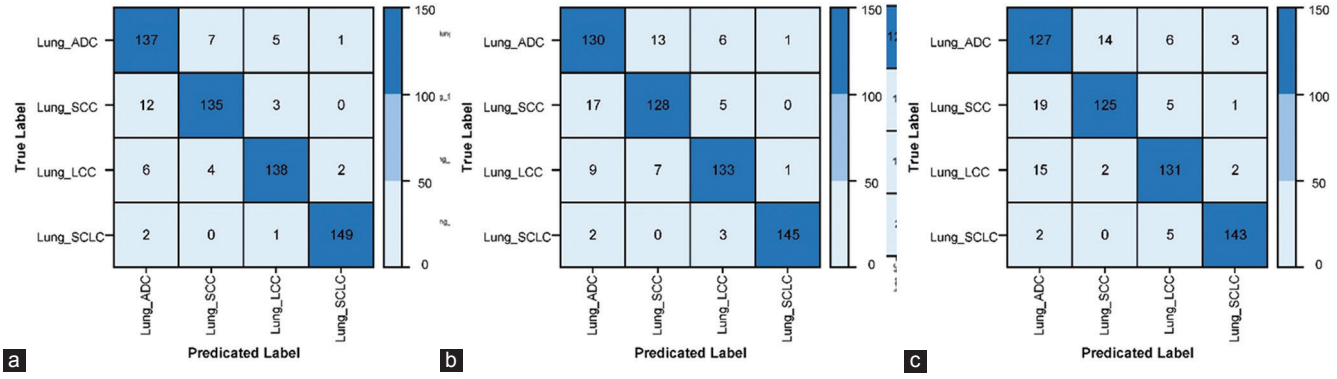

Fig. 8. Confusion matrices on the private test set. (a) Proposed Mode,(b) ConvNext-Tiny, (c) ResNet-50.

Tiny, and ResNet-50 on the public and private datasets. The proposed model demonstrated superior performance, distinguishing ADC from SCC. Overall, accuracy was higher on the LC25000 dataset, which includes benign tissue that is easier to classify and less likely to be misclassified. In contrast, the private dataset, containing only cancer pathological cases, posed greater challenges and offered a more clinically realistic evaluation of generalizability.

TABLE IV
RESULTS SUMMARY OF PROPOSED MODEL

Dataset type	Training accuracy (%)	Validation accuracy (%)	Test accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
LC 25,000	99.13	98.51	98.73	98.74	98.73	98.73
Private	97.45	95.53	93.17	93.19	93.17	93.16

TABLE V
RESULTS SUMMARY OF CONVNEXT-TINY

Dataset type	Training accuracy (%)	Validation accuracy (%)	Test accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
LC 25,000	98.92	96.98	96.27	96.29	96.27	96.27
Private	94.96	91.48	89.33	89.47	89.33	89.38

TABLE VI
RESULTS SUMMARY OF RESNET-50

Dataset type	Training accuracy (%)	Validation accuracy (%)	Test accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
LC 25,000	98.34	95.21	94.00	94.03	94.00	94.01
Private	93.91	89.88	87.67	87.91	87.67	87.73

C. Quantitative Evaluation and Discussion

Among the three evaluated models, the proposed model demonstrated the highest classification accuracy, achieving 98.73% on the LC25000 dataset and 93.17% on the private clinical dataset. While all models performed better on the public dataset, the lower performance on the private dataset reflects the complexity of clinically representative, non-augmented data before partitioning. An extended overview of performance metrics through training, validation, and testing steps is discussed in Tables IV-VI. ConvNext-Tiny produced the second-highest results, followed by ResNet-50. The suggested model performed better than both of them consistently in precision, recall, and F1-score, and was more robust and generalizable, validating its robustness and generalizability ability in diverse data settings. These findings suggest its potential to assist histopathologists in future clinical applications and contribute meaningfully to the early and accurate classification of lung cancer.

D. Limitations and Future Work

This study utilized a patch-based dataset. While this approach was practical for computational scaling and reduced training times, the results were still based on a patch-based dataset, not suitable for real-world diagnostic purposes. By using semi-uniform cropped image patches from a relatively homogeneous dataset, the classification was straightforward, completely ignoring the complexity and variability of whole-slide histopathology images. Thus, the model might become trained to recognize particular aspects of training data, and its accuracy under these controlled conditions may not entirely represent its performance in wider and less-structured clinical scenarios.

This high level of classification accuracy is in agreement with previous studies using the same standard patch-based experimental design, which supports the fact that the proposed model is reliable under this methodology. However, this also suggests that the strong performance is mainly attributable to the controlled and simplified nature of the task, and may not directly translate to more challenging, clinically realistic settings. Underscoring the need for further validation on whole-slide or heterogeneous datasets.

Furthermore, the LC25000 dataset includes augmented images, which, while helpful in increasing dataset size, may inadvertently introduce data leakage between training and testing splits. Future work should therefore prioritize the development and rigorous evaluation of models on whole-slide images to ensure greater generalizability and clinical relevance.

V. Conclusion

This study developed deep-learning model for lung cancer classification inspired by the architectural strengths of ConvNext-Tiny and Restnet-50 while also incorporating self-attention layers to enhance its learning capabilities. Rather than integrating these models, they were used as baseline references to inform the design of the proposed architecture, which enables the model to capture both local and global contextual dependencies among pixels, thereby improving feature representation and classification. Public and private datasets are used for training and testing models.

The suggested architect reached higher accuracy than baselines, ConvNext-Tiny and ResNet-50, on both public and private datasets. The results would enable the model to classify lung cancer in an early and accurate manner, and such an interesting research direction could be thoroughly investigated in future studies of computer-aided pathology.

REFERENCES

Ahmed, S., Shaikh, A., Alshahrani, H., Alghamdi, A., Alrizq, M., Baber, J., and Bakhtyar, M., 2021. Transfer learning approach for classification of histopathology whole slide images. *Sensors*, 21(16), p.5361.

Anjum, S., Ahmed, I., Asif, M., Aljuaid, H., Alturise, F., Ghadi, Y.Y., and Elhabob, R., 2023. Lung cancer classification in histopathology images using multiresolution efficient nets. *Computational Intelligence and Neuroscience*, 2023, pp.7282944.

Borkowski, A.A., Bui, M.M., Thomas, L.B., Wilson, C.P., DeLand, L.A., and Mastorides, S.M., 2019. *Lung and Colon Cancer Histopathological Image Dataset (LC25000)*. [Preprint].

Devi, B.R., Sharma, A., Ramesh, S., and Kumar, R., 2024. Efficient NetB3 for enhanced lung cancer detection: Histopathological image study with augmentation. *Journal of Information Technology Management*, 16(1), pp.98-117.

El-Ghany, S.A., Azad, M., and Elmogy, M., 2023. Robustness fine-tuning deep learning model for cancers diagnosis based on histopathology image analysis. *Diagnostics*, 13(4), p.699.

Gong, W., Liu, X., Wang, Y., Zhang, H., and Zhao, J., 2025. Evaluation of an enhanced ResNet-18 classification model for rapid on-site diagnosis in respiratory cytology. *BMC Cancer*, 25(1), p.10.

Gowthamy, J., and Ramesh, S., 2024. A novel hybrid model for lung and colon cancer detection using pre-trained deep learning and KELM. *Expert Systems with Applications*, 252, p.124114.

Hattori, H., Morikawa, T., Matsumoto, T., Shibata, H., Yamamoto, K., Takeda, Y., Nakamura, S., Fujita, T., and Nakano, T., 2023. Tumor-identification method for predicting recurrence of early-stage lung adenocarcinoma using digital pathology images by machine learning. *Journal of Pathology Informatics*, 14, p.100175.

Hatuwal, B.K., and Thapa, H.C., 2020. Lung cancer detection using convolutional neural network on histopathological images. *International Journal of Computer Trends and Technology*, 68(10), pp.21-24.

Howlader, N., Noone, A.M., Krapcho, M., Miller, D., Brest, A., Yu, M., Ruhl, J., Tatalovich, Z., Mariotto, A., Lewis, D.R., Chen, H.S., Feuer, E.J., and Cronin, K.A., 2020. The effect of advances in lung-cancer treatment on population mortality. *New England Journal of Medicine*, 383(7), pp.640-649.

Hua, W., Li, C., and Wang, X., 2024. Review of convolutional neural network models and image classification. *Academic Journal of Science and Technology*, 10(3), pp.178-184.

Ijaz, M., Qureshi, S.A., Khan, M.A., Ashraf, I., Zahid, U., Yasin, A., Zhang, Y.D., and Wang, S., 2023. DS²LC³Net: A decision support system for lung colon cancer classification using fusion of deep neural networks and normal distribution-based gray wolf optimization. *ACM Transactions on Asian and Low-Resource Language Information Processing*. New York: ACM Publications, p.3625096.

Kanavos, A., and Mylonas, P., 2023. Deep learning analysis of Histopathology images for breast cancer Detection: A Comparative Study of ResNet and VGG Architectures. In: *2023 18th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP)*. IEEE, pp.1-6.

Khan, S., Naseer, M., Hayat, M., Zamir, S.W., Khan, F.S., and Shah, M., 2022. Transformers in vision: A survey. *ACM Computing Surveys*, 54(10), p.200.

Komura, D., and Ishikawa, S., 2018. Machine learning methods for histopathological image analysis. *Computational and Structural Biotechnology Journal*, 16, pp.34-42.

Kriegsmann, M., Haag, C., Weis, C.A., Steinbuss, G., Warth, A., Zgorzelski, C., Muley, T., Winter, H., Eichhorn, M.E., Eichhorn, F., Kriegsmann, J., Christopolous, P., Thomas, M., Witzens-Harig, M., Sinn, P., Von Winterfeld, M., Heussel, C.P., Herth, F.J.F., Klauschen, F., Stenzinger, A., and Kriegsmann, K., 2020. Deep learning for the classification of small-cell and non-small-cell lung cancer. *Cancers*, 12(6), p.1604.

Kumar, N., Sharma, M., Singh, V.P., Madan, C., and Mehandia, S., 2022. An empirical study of handcrafted and dense feature extraction techniques for lung and colon cancer classification from histopathological images. *Biomedical Signal Processing and Control*, 75, p.103596.

Liu, Z., Mao, H., Wu, C.Y., Feichtenhofer, C., Darrell, T., and Xie, S., 2022. A ConvNet for the 2020s. In: *Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, pp.11966-11976.

Masud, M., Sikder, N., Abdullah-Al Nahid, A., Bairagi, A.K., and AlZain, M.A., 2021. A machine learning approach to diagnosing lung and colon cancer using a deep learning-based classification framework. *Sensors (Basel, Switzerland)*, 21(3), p.748.

Mehmood, S., Ghazal, T.M., Khan, M.A., Zubair, M., Naseem, M.T., Faiz, T., and Ahmad, M., 2022. Malignancy detection in lung and colon histopathology images using transfer learning with class selective image processing. *IEEE Access*, 10, pp.25657-25668.

Nicholson, A.G., Tsao, M.S., Beasley, M.B., Borczuk, A.C., Brambilla, E., Cooper, W.A., Dacic, S., Jain, D., Kerr, K.M., Lantuejoul, S., Noguchi, M., Papotti, M., Rekhtman, N., Scagliotti, G., Van Schil, P., Sholl, L., Yatabe, Y., Yoshida, A., and Travis, W.D., 2022. The 2021 WHO classification of lung tumors: Impact of advances since 2015. *Journal of Thoracic Oncology*, 17(3), pp.362-387.

Rajasekar, V., Vaishnnave, M.P., Premkumar, S., Sarveshwaran, V., and Rangaraaj, V., 2023. Lung cancer disease prediction with CT scan and histopathological images feature analysis using deep learning techniques. *Results in Engineering*, 18, p.101111.

Raschka, S., Mirjalili, V., and Raschka, J., 2022. *Machine Learning with PyTorch and Scikit-Learn: Develop Machine Learning and Deep Learning Models with Scikit-Learn and PyTorch*. Packt Publishing, Birmingham.

Siegel, R.L., Giaquinto, A.N., and Jemal, A., 2024. Cancer statistics, 2024. *CA: A Cancer Journal for Clinicians*, 74(1), pp.12-49.

Sumon, R.I., Mazumdar, M.A.I., Uddin, S.M.I., and Kim, H.C., 2024. Exploring Deep Learning and Machine Learning Techniques for Histopathological Image Classification in Lung Cancer Diagnosis. In: *2024 International Conference on Electrical, Computer, and Energy Technologies (ICECET)*. IEEE, pp.1-6.

Tortora, M., Cordelli, E., Sicilia, R., Nibid, L., Ippolito, E., Perrone, G., Ramella, S., and Soda, P., 2023. RadioPathomics: Multimodal learning in non-small cell lung cancer for adaptive radiotherapy. *IEEE Access*, 11, pp.47563-47578.

Tummala, S., Yedla, R.N., Ahmed, K., Katukojwala, A., Vankayalapati, S.L.V., and Garikapati, V., 2023. An explainable classification method based on complex scaling in histopathology images for lung and colon cancer. *Diagnostics*, 13(9), p.1594.

Wei, J.W., Tafe, L.J., Linnik, Y.A., Vaickus, L.J., Tomita, N., and Hassanpour, S., 2019. Pathologist-level classification of histologic patterns on resected lung adenocarcinoma slides with deep neural networks. *Scientific Reports*, 9, p.3358.

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, p.80.