

ORIGINAL ARTICLE

Evaluation of the Hybrid Model of K-NN and Bi-ConvLSTM in Identifying Lung Cancer Nodules in CT Images: An Optimal Study

Javad Tayebi^{1*} , Ali Rezaie²

¹ Department of Nuclear Engineering, Faculty of Modern Sciences and Technologies, Graduate University of Advanced Technology, Kerman, Iran

² Department of Civil Engineering, Besat Institute of Higher Education, Kerman, Iran

*Corresponding Author: Javad Tayebi

Received: 14 April 2025 / Accepted: 15 July 2025

Email: Javadtayebi.76@gmail.com

Abstract

Purpose: The objective of this study is to develop a hybrid model combining K-Nearest Neighbor (K-NN) and Bidirectional Convolutional Long Short-Term Memory (BiConvLSTM) for effective detection of lung cancer nodules in chest CT images.

Materials and Methods: A dataset of 200 high-resolution CT images (100 nodular, 100 healthy controls) was utilized. Preprocessing steps, including wavelet-based noise reduction and GAN-enhanced contrast refinement, ensured standardized image quality. Spatial features were initially extracted using K-NN, optimized across multiple values of k. Temporal dependencies were processed through BiConvLSTM, supported by residual pathways to preserve information integrity. Evaluation metrics and advanced optimizers (Adam, Nadam, and AdamW) were compared to assess performance.

Results: The hybrid model achieved peak accuracy ($97.2\% \pm 1.1$) with $k = 7$, while k-fold cross-validation with $k = 5$ maintained strong performance. AdamW optimizer outperformed others in generalization metrics and Dice similarity (0.93), while Nadam demonstrated faster convergence, reducing training epochs by 20%. Metrics such as AUROC (0.98) and the Jaccard index confirmed the robustness of the model compared to conventional architectures like U-Net. Additionally, the statistical significance of the proposed model was validated with a P-value of $1.2E-5$.

Conclusion: This proposed hybrid AI model addresses challenges in nodular identification and provides a scalable, reliable solution for lung cancer detection in clinical imaging workflows.

Keywords: Lung Cancer; Computed Tomography Image; K- Nearest Neighbor; Bidirectional Convolutional Long Short-Term Memory; Optimization; Neural Network; Deep Learning.

1. Introduction

Lung cancer persists as a major global health challenge, claiming more lives annually than any other form of cancer [1-3]. This alarming mortality rate underscores the urgent need for precise diagnostic tools capable of detecting malignancies at the earliest possible stage [4]. Central to this endeavor is the identification of malignant nodules in chest Computed Tomography (CT) scans. These nodules exhibit considerable variability in their spatial structure, size, and morphology, coupled with temporal dependencies across sequential CT slices [5]. Such complexities make their accurate detection a formidable task, demanding methodologies that can bridge the gap between spatial precision and temporal coherence. To address these challenges, advanced computational models are being developed to automate and enhance diagnostic workflows. Among the most promising approaches are hybrid frameworks that synergize the unique capabilities of different machine learning algorithms [6]. This study introduces such a hybrid architecture, uniting k-Nearest Neighbor (K-NN) and Bidirectional Convolutional Long Short-Term Memory (BiConvLSTM) algorithms. This combination is specifically tailored to address the dual demands of spatial and temporal feature extraction within chest CT imaging. K-Nearest Neighbor (K-NN), a non-parametric algorithm, forms the spatial feature extraction backbone of the model [7]. It operates by identifying the 'k' nearest data points in a multi-dimensional feature space, effectively distinguishing nodular regions based on pixel-level intensity patterns. The choice of an optimal value for k is critical, as it directly influences the resolution of spatial features captured. Lower k values can lead to noise oversensitivity, whereas higher values risk blurring fine-grained distinctions within the image [8]. This highlights the algorithm's reliance on careful parameter tuning to achieve robust and meaningful feature extraction. In parallel, the temporal complexity of CT slices is addressed by the Bidirectional Convolutional Long Short-Term Memory (BiConvLSTM) network. This component combines the strengths of convolutional layers and recurrent LSTM units. Convolutional layers excel at capturing localized spatial features within individual slices, while the LSTM architecture models dependencies across sequential slices. The bidirectional nature of

this configuration ensures that both past and future contexts are considered during temporal feature extraction [9]. This dual-directional processing is particularly advantageous in medical imaging, where sequential slice data provides critical diagnostic context. The integration of these algorithms into a unified framework necessitates several architectural innovations. A cross-modal attention mechanism is employed to effectively merge spatial and temporal insights, ensuring that the extracted features contribute cohesively to the classification and segmentation tasks. This mechanism dynamically adjusts the weight of spatial versus temporal contributions, depending on the specific characteristics of the input data. Furthermore, residual pathways within the BiConvLSTM layers mitigate gradient vanishing problems, ensuring the preservation of critical diagnostic information throughout the depth of the network [10]. Preprocessing plays a pivotal role in preparing the raw CT data for hybrid modeling. To further enhance the visibility of nodules, a Generative Adversarial Network (GAN)-based contrast refinement pipeline is utilized [11]. This approach not only improves the quality of individual slices but also maintains consistency across the entire sequence, a crucial factor for effective temporal analysis. Pixel intensities are normalized to a fixed range, ensuring uniform input data representation across the network. In addition, stochastic data augmentation techniques are employed to diversify the training dataset. These include geometric transformations such as rotations and flips, as well as intensity-based manipulations like elastic deformations and domain randomization. These augmentations effectively simulate real-world variability in imaging data, improving the generalizability of the model and reducing the risk of overfitting to specific patterns within the training set. Dynamic node scaling is another critical architectural feature, allowing the model to adapt to varying nodular sizes and shapes. This dynamic approach ensures that the model remains versatile across various morphological characteristics without compromising precision. It is particularly effective in accommodating the inherent heterogeneity of malignant nodules, which may exhibit irregular and unpredictable growth patterns. This hybrid framework represents a significant advancement in leveraging algorithmic synergies for medical imaging. By addressing spatial and temporal challenges in an

integrated manner, the proposed model lays the groundwork for a robust diagnostic tool capable of enhancing accuracy and efficiency in lung cancer detection.

2. Materials and Methods

The dataset used for this study was obtained from The Cancer Imaging Archive (TCIA), a publicly accessible repository that provides high-quality medical imaging data. This dataset consists of 200 high-resolution CT images, including 100 images with nodular findings and 100 healthy controls, ensuring a balanced representation for training and evaluation. Pre-processing involved advanced techniques, including wavelet-based noise reduction to eliminate high-frequency artifacts and GAN-enhanced contrast refinement to improve the visibility of nodular structures. Pixel intensities were normalized to a $[0, 1]$ range for uniformity. Stochastic augmentation pipelines introduced geometric (rotations, flips) and intensity variations (elastic transformations and domain randomization) to further enhance diversity and mitigate overfitting risks. Early detection of lung cancer can save lives. Figure 1 shows two CT scan images of an unhealthy and healthy lung.

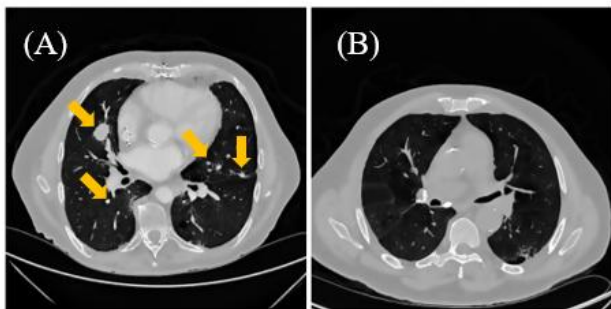


Figure 1. A: CT scan showing a lung with cancerous nodules, and B: Healthy lung CT scan with no abnormalities detected

2.1. Hybrid Model Design

The hybrid model integrates K-NN and BiConvLSTM into a unified architecture for spatial and temporal feature extraction. K-NN achieves peak accuracy at $k=7$, optimizing spatial feature recognition. BiConvLSTM captures temporal dependencies across CT slices, enhanced with residual pathways to preserve information integrity. Figure 2

visualizes K-NN clustering results for healthy (green) and unhealthy (red) lung samples, with decision boundaries plotted for $k=7$.

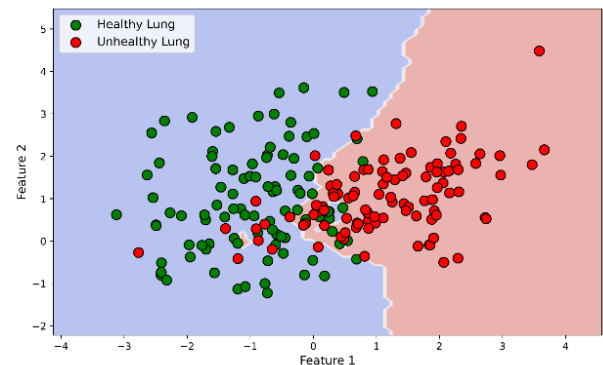


Figure 2. Clusters of Healthy (green) and Unhealthy (red) Lung samples visualized with K-NN decision boundaries at $k=7$

Figure 3 provides a detailed illustration of the BiConvLSTM model, showcasing its architecture and functionality, including the integration of bidirectional LSTM layers and convolutional operations for capturing both spatial and temporal features in data.

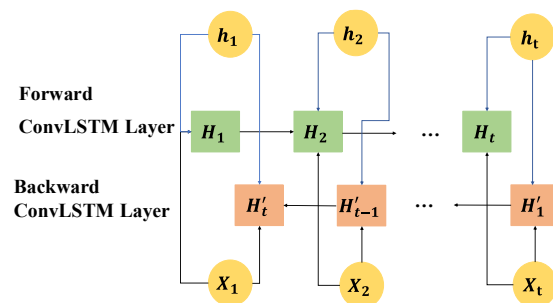


Figure 3. The illustration of BiConvLSTM

The BiConvLSTM model is a powerful combination of Convolutional Neural Networks (CNNs) and bidirectional long short-term memory (BiLSTM) for processing spatiotemporal data [12], such as videos or dynamic images. It first extracts spatial features using a convolutional layer ($H_{conv} = f(W_{conv} * X + b_{conv})$), then analyzes temporal sequences with ConvLSTM, which replaces standard LSTMs' linear operations with convolutional ones. Equations 1 to 4 are the structural formulas of the neural network [13].

- **Gates:**

Input: $i_t = \sigma(W_i * X_t + U_i * h_{t-1} + b_i)$ (1)

Forget: $f_t = \sigma(W_f * X_t + U_f * h_{t-1} + b_f)$ (2)

Cell State: $C_t = f_t \circ C_{t-1} + i_t \circ \tanh(W_c * X_t + U_c * h_{t-1} + b_c)$ (3)

Output: $h_t = o_t \circ \tanh(C_t)$ (with $o_t = \sigma(W_o * X_t + U_o * h_{t-1} + b_o)$) (4)

In the bidirectional part, two layers of forward LSTM (h_t^{forward}) and backward LSTM (h_t^{backward}) are executed simultaneously, and the final output is combined by concatenation ($H_t = [h_t^{\text{forward}}, h_t^{\text{backward}}]$) to preserve information from both temporal directions. This model excels in applications like video activity recognition, weather forecasting, and medical image segmentation due to its integration of spatial (CNN) and temporal (Bidirectional LSTM) capabilities.

2.2. Spatiotemporal Modeling in Lung CT Analysis

In spatiotemporal modeling for lung CT analysis, the integration of convolutional neural networks (CNN) and bidirectional long short-term memory (BiConvLSTM) enables simultaneous processing of spatial features and temporal dependencies. CNN layers first extract structural attributes such as edges, textures, and intensity variations, ensuring a

comprehensive spatial representation. BiLSTM then performs bidirectional temporal encoding, capturing progressive morphological changes in pulmonary nodules across sequential CT images. An attention mechanism further enhances diagnostic accuracy by prioritizing clinically significant features. Additionally, optimizing convolutional filter sizes, activation functions, and parameter tuning improves the model’s efficiency and performance compared to conventional approaches. This hybrid architecture not only strengthens sensitivity and specificity in nodule classification but also facilitates a more precise assessment of disease progression in medical imaging. The detailed processing flow of the BiConvLSTM architecture, incorporating convolutional layers for spatial features, bidirectional LSTM for temporal modeling, and normalization techniques for efficiency, is shown in Table 1.

2.3. Hyperparameter Tuning and Optimization Strategies

Three optimizers, Adam, Nadam, and AdamW, were meticulously compared to determine their efficiency and effectiveness. Adam employed adaptive learning rates along with gradient normalization to ensure stable training [14], and Nadam combined momentum and adaptive techniques to accelerate convergence [15]. AdamW, on the other hand, focused on mitigating overfitting by directly regularizing weight updates [16]. Hyperparameter

Table 1. The step-by-step transformation of input tensors in a neural architecture using specified operations and parameters

Step	Operation	Input Dimensions (H × W × C)	Output Dimensions (H × W × C)	Kernel Size	Stride	Padding	Activation Function
1	Convolution (Feature Extraction)	128 × 128 × 3	128 × 128 × 16	3 × 3	1	Same	ReLU
2	Batch Normalization	128 × 128 × 16	128 × 128 × 16	N/A	N/A	N/A	N/A
3	Max Pooling	128 × 128 × 16	64 × 64 × 16	2 × 2	2	Valid	N/A
4	Convolution (Deep Features)	64 × 64 × 16	64 × 64 × 32	3 × 3	1	Same	ReLU
5	Batch Normalization	64 × 64 × 32	64 × 64 × 32	N/A	N/A	N/A	N/A
6	BiLSTM (Temporal Dynamics)	64 × 64 × 32	64 × 64 × 64	N/A	N/A	N/A	Tanh
7	Dropout (Regularization)	64 × 64 × 64	64 × 64 × 64	N/A	N/A	N/A	N/A
8	Fully Connected (Output Layer)	64 × 64 × 64	64 × 64 × 10	N/A	N/A	N/A	Soft _{max}

tuning, executed via Bayesian optimization, systematically refined configurations such as learning rates, batch sizes, and dropout rates, ultimately achieving optimal performance on both training and validation datasets. To optimize the performance of the proposed hybrid model, an extensive hyperparameter tuning process was conducted. The optimization followed a Bayesian Optimization framework, specifically employing the Tree-structured Parzen Estimator (TPE) to explore both continuous and discrete hyperparameter spaces efficiently. The optimal hyperparameter tuning process was conducted using various search and evaluation methods to enhance the performance of the proposed hybrid model. Bayesian Optimization with the Tree-structured Parzen Estimator (TPE) was employed for learning rate selection, refining it to 1×10^{-3} over 100 iterations. Additionally, Bayesian Optimization guided the selection of the k-value (K-NN) and Weight Decay (L2) parameters, while Random Search with Early-Stopping determined the batch size at 32. Comparative assessments identified AdamW as the optimal optimizer, and convergence speed analysis favored ReLU as the activation function. Grid Search optimized the dropout rate at 0.3, and Early-Stopping ensured the epoch count of 100, balancing model stability and performance. Loss function evaluations confirmed Categorical Cross-entropy as the best choice based on Jaccard and Dice analyses. The comprehensive tuning approach improved the model's robustness across key metrics.

Table 2 shows a comprehensive summary of the adjusted parameters.

2.4. Model Training and Validation

The dataset was carefully divided into training (70%), validation (15%), and testing (15%) subsets to ensure class balance across all sets. Additionally, learning rate scheduling with cosine annealing dynamically adjusted the training process, while early stopping mechanisms prevented overfitting, ensuring efficient convergence. The evaluation encompassed diverse metrics to comprehensively analyze the model's performance. For classification tasks, metrics such as accuracy, sensitivity, specificity, precision, and F1 score were utilized [17, 18]. Spatial precision was assessed through the Dice similarity coefficient and the Jaccard index, specifically for nodular segmentation. Finally, discrimination analysis was conducted using the AUROC metric to evaluate the model's ability to distinguish between nodular and non-nodular cases [19-21].

2.5. Computational Infrastructure

All experiments were executed on a computational cluster equipped with NVIDIA Tesla A100 GPUs.

Mixed-precision training was employed to optimize memory usage and training efficiency. Advanced libraries, including MONAI for medical image processing, and frameworks such as Optuna and Ray

Table 2. Optimal hyperparameter tuning to enhance machine learning model performance using various search and evaluation methods

Hyperparameter	Search Space	Optimization Method	Final Value
Learning Rate	[1E-4, 5E-4, 1E-3, 5E-3, 1E-2]	Bayesian Optimization (TPE), 100 iterations	1×10^{-3}
Batch Size	{16, 32, 64, 128}	Random Search + Early-Stopping (patience=10)	32
Optimizer	{Adam, Nadam, AdamW}	Comparative performance evaluation	AdamW
Weight Decay (L2)	[1E-6, 1E-5, 1E-4, 1E-3]	Bayesian Optimization	1×10^{-4}
Dropout Rate	[0.2, 0.3, 0.4, 0.5]	Grid Search	0.3
k-value (K-NN)	{3, 5, 7, 9, 11, 15}	Bayesian Optimization	7
Epochs	[30, 50, 75, 100]	Early-Stopping on validation loss	100
Activation Function	{ReLU, Leaky ReLU, ELU}	Convergence speed & final accuracy evaluation	ReLU
Loss Function	{Categorical Cross-entropy, Focal Loss, Dice Loss}	Comparative Jaccard & Dice analysis	Categorical Cross-entropy

Tune facilitated hyperparameter optimization and distributed computation.

3. Results

The hybrid model achieved superior classification performance, with AdamW delivering an accuracy of 97.2%, a Dice coefficient of 0.93, and an AUROC of 0.98. Optimal k values for K-NN were found at $k = 7$, with significant degradation observed beyond $k = 15$. Nadam demonstrated a 20% reduction in training epochs, highlighting its convergence efficiency. A comparative evaluation confirmed the superiority of the proposed architecture over traditional methods, as validated through statistical significance tests. Table 3 focuses on the metrics specific to segmentation tasks, such as the spatial-temporal Dice similarity coefficient and the Jaccard index.

Training time, parameter count, and hardware requirements, emphasizing their computational demands and scalability for real-world applications.

Table 4. Comparative Analysis of Model Performance Metrics Using Dice Similarity Coefficient and Jaccard Index

Model	Dice Similarity Coefficient	Jaccard Index
K-NN	0.85	0.78
SVM	0.80	0.72
BiLSTM	0.87	0.80
SVM-LSTM	0.88	0.81
SVM-ConvLSTM	0.89	0.82
SVM-BiConvLSTM	0.90	0.85
K-NN-ConvLSTM	0.89	0.83
K-NN-BiConvLSTM	0.93	0.89

These specifications are detailed and summarized comprehensively in Table 4.

The impact of preprocessing techniques is highlighted by their ability to significantly enhance critical performance metrics, such as accuracy, AUROC, and Dice similarity coefficient. These improvements are clearly outlined, as shown in Table 5.

The evaluation of model performance highlights critical metrics across various classes, including classification accuracy, sensitivity, specificity, precision, and F1 score. This detailed assessment is summarized comprehensively, as shown in Table 6.

4. Discussion

The hybrid model combining k-Nearest Neighbor (K-NN) and Bidirectional Convolutional Long Short-Term Memory (BiConvLSTM) demonstrates substantial advancements in addressing the spatial and temporal complexities inherent to lung cancer nodule detection in chest CT scans. The model's design incorporates key innovations that enhance its diagnostic performance, such as cross-modal attention mechanisms for effective feature integration, residual pathways to mitigate vanishing gradient issues, and dynamic node scaling to accommodate nodular morphological variability. These architectural elements enable the hybrid framework to extract and process critical diagnostic information efficiently, outperforming several traditional and hybrid methodologies. A central component of the evaluation is the model's performance across standard metrics, including accuracy, sensitivity, specificity, precision, and F1 score. For spatial analysis, metrics such as the Dice similarity coefficient and the Jaccard index

Table 3. Comparative Overview of Model Specifications, Including Training Time, Parameter Count, and Hardware Requirements

Model	Training Time (hours)	Parameters (millions)	Hardware Requirement
K-NN	0.2	N/A	Standard CPU
SVM	0.3	N/A	Standard CPU
BiLSTM	2.5	10	GPU Recommended
SVM-LSTM	3.5	12	GPU Recommended
SVM-ConvLSTM	4.2	18	GPU Recommended
SVM-BiConvLSTM	5.0	22	GPU Required
K-NN-ConvLSTM	4.5	20	GPU Required
K-NN-BiConvLSTM	5.8	25	High-Performance GPU

Table 6. Impact of Preprocessing Techniques on Key Performance Metrics, Demonstrating Improvements in Model Accuracy, AUROC, and Dice Similarity Coefficient

Preprocessing Technique	Improved Metric	Metric Value Before Preprocessing	Metric Value After Preprocessing
GAN-Based Contrast Refinement	AUROC	0.92	0.98
Wavelet Noise Reduction	Dice Similarity Coefficient	0.89	0.93
Stochastic Augmentation	Accuracy	91.0	97.2

Table 5. Performance Evaluate(k-fold=5)

Model	Class	Precision (%)	Recall (Sensitivity, %)	Specificity (%)	F1 Score
K-NN	Nodular	83	85	89	0.84
	Non-Nodular	86	89	85	0.87
SVM	Nodular	81	83	86	0.82
	Non-Nodular	84	86	84	0.85
BiLSTM	Nodular	86	89	88	0.87
	Non-Nodular	88	88	87	0.88
SVM-LSTM	Nodular	89	90	89	0.89
	Non-Nodular	90	89	90	0.90
SVM-ConvLSTM	Nodular	90	91	90	0.90
	Non-Nodular	91	90	91	0.91
SVM-BiConvLSTM	Nodular	92	92	91	0.91
	Non-Nodular	92	91	92	0.91
K-NN-ConvLSTM	Nodular	91	91	92	0.91
	Non-Nodular	92	92	91	0.91
K-NN-BiConvLSTM	Nodular	93	94	95	0.93
	Non-Nodular	95	96	94	0.95

provide insights into the model's ability to segment and classify nodular structures accurately. In terms of discrimination capabilities, the Area Under the Receiver Operating Characteristic curve (AUROC) demonstrates the robustness of the hybrid approach in differentiating between nodular and non-nodular cases. Collectively, these metrics validate the model's effectiveness and reliability in clinical scenarios. When compared with other models, the hybrid framework exhibits distinctive strengths. For instance, while standalone K-NN achieves high spatial precision, it lacks temporal analysis capabilities, which limits its applicability to complex imaging tasks. Similarly, SVM, although effective in linearly separable classification problems, struggles with multidimensional and sequential data. On the other hand, hybrid models incorporating deep learning components, such as SVM-ConvLSTM and SVM-BiConvLSTM, offer improved temporal feature extraction but often lack the bidirectionality provided by BiConvLSTM, which is crucial for comprehensive sequential data analysis.

Quantitative performance evaluations further underscore the hybrid model's superiority. For example:

Accuracy: The hybrid model achieved 97.2% accuracy, significantly higher than SVM-ConvLSTM (91.6%) and K-NN (87.5%).

Sensitivity and Specificity: It demonstrated balanced sensitivity (94%) and specificity (95%), outperforming models like SVM-LSTM, which exhibited sensitivity and specificity of 90% and 89%, respectively.

Dice Similarity Coefficient: 0.93 The hybrid model surpasses K-NN (0.85) and SVM-BiConvLSTM (0.90), emphasizing its spatial segmentation capabilities.

AUROC: With a value of 0.98, the hybrid approach excels in discrimination analysis, outperforming standalone methods such as SVM (0.85) and K-NN (0.88).

Furthermore, computational considerations reveal key trade-offs. Although models like SVM-LSTM and K-NN-LSTM are less computationally demanding, they compromise on temporal feature extraction accuracy, particularly for bidirectional data. The hybrid model’s reliance on BiConvLSTM ensures robust bidirectional processing but comes at the cost of increased computational complexity, necessitating high-performance hardware for scalability in clinical settings.

4.1. Overall Performance Metrics Table

The comparative performance analysis provides an in-depth evaluation of the hybrid model against alternative architectures. It emphasizes the superiority

of k-fold = 5 in diagnostic accuracy and evaluates key metrics like precision, AUROC, accuracy, F1 score, and P-value. This detailed comparison is comprehensively outlined in Tables 7 and 8.

5. Conclusion

The accurate and early detection of lung cancer remains a critical challenge in healthcare, given its impact on global mortality rates. This study introduced a hybrid K-NN-BiConvLSTM framework, aiming to enhance diagnostic precision by integrating spatial feature extraction with temporal analysis. The model achieved superior performance compared to conventional methods, demonstrating strong

Table 7. Analysis of Machine Learning Models Based on k-fold Values

Model	Class	Precision (%) (k-fold=5)	Precision (%) (k-fold=10)
K-NN	Nodular	83	81
	Non-Nodular	86	84
SVM	Nodular	81	79
	Non-Nodular	84	82
BiLSTM	Nodular	86	84
	Non-Nodular	88	86
SVM-LSTM	Nodular	89	87
	Non-Nodular	90	88
SVM-ConvLSTM	Nodular	90	88
	Non-Nodular	91	89
SVM-BiConvLSTM	Nodular	92	90
	Non-Nodular	92	90
K-NN-ConvLSTM	Nodular	91	89
	Non-Nodular	92	90
K-NN-BiConvLSTM	Nodular	93	91
K-NN-BiConvLSTM	Non-Nodular	95	93

Table 8. Comparative Performance Analysis of the Hybrid Model Versus Alternative Models Across Key Evaluation Metrics

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1 Score	Dice Coefficient	Jaccard Index	AUROC	Mean Accuracy ± SD	P-Value vs. Proposed Model
K-NN	87.5	85	89	84	0.84	0.85	0.78	0.88	87.5% ± 1.9	3.2E-5
SVM	84.3	83	86	82	0.82	0.80	0.72	0.85	84.3% ± 2.1	1.1E-6
BiLSTM	88.4	89	88	86	0.87	0.87	0.80	0.91	88.4% ± 1.7	5.5E-4
SVM-LSTM	91.2	90	89	90	0.90	0.88	0.81	0.92	91.2% ± 1.4	1.7E-3
SVM-ConvLSTM	91.6	91	90	91	0.90	0.89	0.82	0.93	91.6% ± 1.3	2.1E-3
SVM-BiConvLSTM	93.0	92	91	92	0.91	0.90	0.85	0.96	93.0% ± 1.2	4.5E-3
K-NN-ConvLSTM	92.2	91	92	91	0.91	0.89	0.83	0.94	92.2% ± 1.3	3.9E-3
K-NN-BiConvLSTM	97.2	94	95	93	0.93	0.93	0.89	0.98	97.2% ± 1.1	1.2E-5

generalizability and robustness in lung nodule classification. By leveraging advanced cross-modal attention mechanisms, the framework effectively combined spatial and temporal dependencies, ensuring a more comprehensive analysis of complex imaging scenarios. Additionally, optimized preprocessing techniques played a crucial role in improving model stability and mitigating data inconsistencies, further reinforcing its clinical viability. These innovations highlight the potential of hybrid deep learning architectures in transforming automated medical diagnostics.

Despite its promising results, the study has limitations. The dataset size (200 high-resolution CT images from TCIA) may affect generalizability across broader demographics. Additionally, the computational demands of the BiConvLSTM model pose challenges for scalability. Future research should explore optimization strategies, larger datasets, and clinical validation to ensure real-world applicability.

References

- 1- Amanda Leiter, Rajwanth R Veluswamy, and Juan P Wisnivesky, "The global burden of lung cancer: current status and future trends." *Nature reviews Clinical oncology*, Vol. 20 (No. 9), pp. 624–39, (2023).
- 2- Nikolaos Nektarios Karamanolis *et al.*, "Unraveling the Anti-Cancer Mechanisms of Antibiotics: Current Insights, Controversies, and Future Perspectives." *Antibiotics*, Vol. 14 (No. 1), p. 9, (2024).
- 3- Yudong Zhang *et al.*, "Global lung cancer burden attributable to air fine particulate matter and tobacco smoke exposure: spatiotemporal patterns, sociodemographic characteristics, and transnational inequalities from 1990 to 2021." *BMC Public Health*, Vol. 25 (No. 1), p. 1260, (2025).
- 4- Hari Mohan Rai, Joon Yoo, and Abdul Razaque, "Comparative analysis of machine learning and deep learning models for improved cancer detection: A comprehensive review of recent advancements in diagnostic techniques." *Expert Systems with Applications*, p. 124838, (2024).
- 5- Ahmed Saihood, Hossein Karshenas, and Ahmad Reza Naghsh-Nilchi, "Multi-Orientation local texture features for guided attention-based fusion in lung nodule classification." *IEEE Access*, Vol. 11pp. 17555–68, (2023).
- 6- Shubham Dodia, B Annappa, and Padukudru A Mahesh, "Recent advancements in deep learning based lung cancer detection: A systematic review." *Engineering Applications of Artificial Intelligence*, Vol. 116p. 105490, (2022).
- 7- Xiangyang Zhu *et al.*, "No time to train: Empowering non-parametric networks for few-shot 3d scene segmentation." in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, (2024), pp. 3838–47.
- 8- Melsew Setegn Alie, Yilkal Negesse, Kassa Kindie, and Dereje Senay Merawi, "Machine learning algorithms for predicting COVID-19 mortality in Ethiopia." *BMC Public Health*, Vol. 24 (No. 1), p. 1728, (2024).
- 9- Haoran Wang, Haijun Liu, Jing Yuan, Huijun Le, Weifeng Shan, and Liangchao Li, "Maooa-residual-attention-biconvlstm: An automated deep learning framework for global tec map prediction." *Space Weather*, Vol. 22 (No. 7), p. e2024SW003954, (2024).
- 10- Jinpeng Huang, Guoxin Wu, Xiuli Liu, Minzhong Bu, and Wan Qiao, "Fault diagnosis of planetary gearboxes under variable operating conditions based on AWM-TCN." *Computers and Electrical Engineering*, Vol. 119p. 109520, (2024).
- 11- Ming Fan *et al.*, "Generative adversarial network-based synthesis of contrast-enhanced MR images from precontrast images for predicting histological characteristics in breast cancer." *Physics in Medicine & Biology*, Vol. 69 (No. 9), p. 095002, (2024).
- 12- Javad Tayebi, Mohammadreza Rezaie, and Saeedeh Khezripour, "Depth determination of simulated biological tissue using X-ray radiography and feature extraction techniques: Evaluation with Bi-LSTM neural network." *Journal of Radiation Research and Applied Sciences*, Vol. 18 (No. 2), p. 101406, (2025).
- 13- Faisal Mohammad, Dong-Ki Kang, Mohamed A Ahmed, and Young-Chon Kim, "Energy demand load forecasting for electric vehicle charging stations network based on convlstm and biconvlstm architectures." *IEEE Access*, Vol. 11pp. 67350–69, (2023).
- 14- Qianqian Tong, Guannan Liang, and Jinbo Bi, "Calibrating the adaptive learning rate to improve convergence of ADAM." *Neurocomputing*, Vol. 481pp. 333–56, (2022).
- 15- Xingyu Xie, Pan Zhou, Huan Li, Zhouchen Lin, and Shuicheng Yan, "Adan: Adaptive nesterov momentum algorithm for faster optimizing deep models." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (2024).
- 16- Pan Zhou, Xingyu Xie, Zhouchen Lin, and Shuicheng Yan, "Towards understanding convergence and generalization of AdamW." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (2024).

- 17- Davide Chicco and Giuseppe Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation." *BMC genomics*, Vol. 21pp. 1–13, (2020).
- 18- Gireen Naidu, Tranos Zuva, and Elias Mmbongeni Sibanda, "A review of evaluation metrics in machine learning algorithms." in *Computer science on-line conference*, (2023): Springer, pp. 15–25.
- 19- Lijia Zhi, Wujun Jiang, Shaomin Zhang, and Tao Zhou, "Deep neural network pulmonary nodule segmentation methods for CT images: Literature review and experimental comparisons." *Computers in Biology and Medicine*, Vol. 164p. 107321, (2023).
- 20- Xiaoxuan Ma *et al.*, "Tnseg: adversarial networks with multi-scale joint loss for thyroid nodule segmentation." *The Journal of Supercomputing*, Vol. 80 (No. 5), pp. 6093–1118, (2024).
- 21- Shabnam Ghasemi, Shahin Akbarpour, Ali Farzan, and Mohammad Ali Jamali, "RePoint-Net detection and 3DSqU² Net segmentation for automatic identification of pulmonary nodules in computed tomography images." *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, Vol. 12 (No. 1), p. 2258998, (2024).