Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 6, 2026

NANO-HYBRID: A LIGHTWEIGHT INCEPTIONNEXT-ATTENTION NETWORK FOR EFFICIENT LUNG CANCER DIAGNOSTICS

Hazik Jaffri[*1], **Muhammad Sheraz Nawaz**[2], **Umer Raza**[3]

[*1,3]*Department of Artificial Intelligence, University of Management & Technology, Lahore, Pakistan*
[2]*Department Of Informatics and System, University of Management & Technology, Lahore, Pakistan*

[*1]hazik.jaffri@gmail.com, [2]sheraz.nawaz@umt.edu.pk, [3]uh043316@gmail.com

DOI: https://doi.org/10.5281/zenodo.20568991

Keywords
Lung Cancer Detection, Nano-Hybrid, InceptionNeXt, Explainable AI, Lightweight Deep Learning

Article History
Received: 08 April 2026
Accepted: 20 May 2026
Published: 06 June 2026

Corresponding Author: *
Hazik Jaffri

Abstract
Lung cancer remains a leading cause of cancer-related mortality, necessitating diagnostic tools that are both accurate and computationally efficient for widespread deployment. While recent hybrid Deep Learning models have achieved high classification performance, they typically rely on heavy architectures (>18 million parameters) and extensive pre-training, limiting their applicability on resource-constrained edge devices. This study proposes a Nano-Hybrid architecture that integrates lightweight InceptionNeXt convolutions with global attention mechanisms, designed to be trained entirely from scratch. We evaluated the model on two diverse datasets: the IQ-OTH/NCCD (3-class) and a multi-class Chest CT dataset (4-class). Despite containing ~89% fewer parameters (2.03M) than comparable state-of-the-art baseline models, our approach achieved 95.41% accuracy on the IQ dataset, demonstrating that massive capacity is not strictly required for high-performance diagnostics. On the challenging multi-class Chest CT dataset, the model achieved 86.11% accuracy, with a notable 1.00 AUC (Area Under Curve) for normal cases, ensuring zero false positives in healthy screenings. Explainability analysis using Grad-CAM further validates that the model correctly prioritizes pulmonary nodule structures over background artifacts.

1 INTRODUCTION

Lung cancer accounts for the highest number of cancer-related deaths globally, with recent statistics indicating a rising incidence rate in developing nations [1]. Early diagnosis through Computed Tomography (CT) scans is critical for improving survival rates; however, manual analysis of CT scans is time-consuming, labor-intensive, and prone to inter-observer variability. Radiologists must review hundreds of slices per patient, leading to fatigue and potential diagnostic errors.

Deep Learning (DL) has emerged as a powerful tool for automated diagnosis, offering the potential to assist clinicians by acting as a second opinion. Historically, Convolutional Neural Networks (CNNs) like ResNet [14] and DenseNet [17] have set the benchmark for medical image analysis. More recently, Vision Transformers (ViTs) [12] have demonstrated superior performance by capturing long-range dependencies in images. State-of-the-art models often utilize these heavy backbones or hybrid combinations of them.

However, a critical bottleneck remains: **Computational Efficiency**. While accurate, these SOTA models typically suffer from high computational costs, requiring millions of parameters (often >20M) and massive datasets for pre-training (e.g., ImageNet). This makes them unsuitable for "Edge AI" applications, such as portable CT scanners, mobile health applications, or embedded systems in hospitals with limited hardware infrastructure.

To address this gap, we propose a **Nano-Hybrid Model**. Our architecture combines the efficient feature extraction of **InceptionNeXt** blocks [40] with the global context modeling of **Self-Attention** [37], all within a compact 2-million-parameter footprint. Unlike traditional pruning or compression techniques, our model is designed to be lightweight from the ground up, allowing it to be trained from scratch without relying on external pre-trained weights.

The primary contributions of this paper are:

1. **Architecture:** We introduce a novel Nano-Hybrid architecture that reduces parameter count by ~89% compared to standard hybrid baselines while maintaining competitive accuracy.
2. **Robustness:** We validate the model on two distinct datasets (IQ-OTH/NCCD and Chest CT) to demonstrate its versatility across binary and multi-class problems.
3. **Explainability:** We employ Gradient-weighted Class Activation Mapping (Grad-CAM) [30] to provide visual interpretability, ensuring the model's decisions are clinically relevant.

2 Methodology

2.1 Dataset Description

To ensure the robustness of the proposed model, this study utilizes two distinct datasets representing different clinical challenges.

IQ-OTH/NCCD Dataset:

This dataset [4] consists of 1,097 CT scans classified into three categories: Benign, Malignant, and Normal. The images vary significantly in contrast and noise levels, representing a challenge for feature extraction consistency.

Chest CT-Scan Dataset:

A more complex dataset comprising four categories: Adenocarcinoma, Large Cell Carcinoma, Squamous Cell Carcinoma, and Normal. Distinguishing between carcinoma subtypes is inherently difficult due to their similar morphological presentation in 2D slices.

2.1 Dataset Description

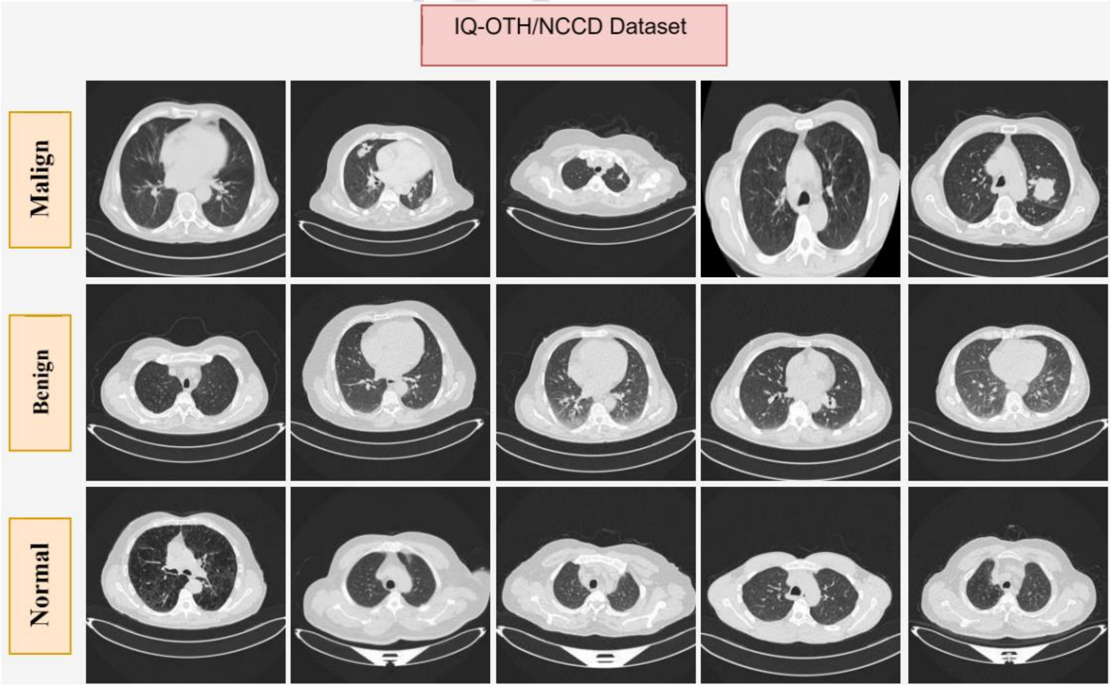

Fig. 1 Sample CT scans from the IQ-OTH/NCCD dataset. The dataset exhibits significant variability in nodule appearance and background noise.

Data Preprocessing:
All images were resized to a standard resolution of 224×224 pixels. To improve model generalization and prevent overfitting—a common issue when training small models from scratch—we applied data augmentation techniques including random rotations ($\pm 15^\circ$), horizontal flipping, and intensity normalization using the mean and standard deviation of the ImageNet dataset.

2.2 Proposed Architecture
The proposed Nano-Hybrid model abandons standard heavy convolutions in favor of **InceptionNeXt** blocks to reduce parameter count. The architecture consists of three stages, culminating in a global attention mechanism (Fig. 2).

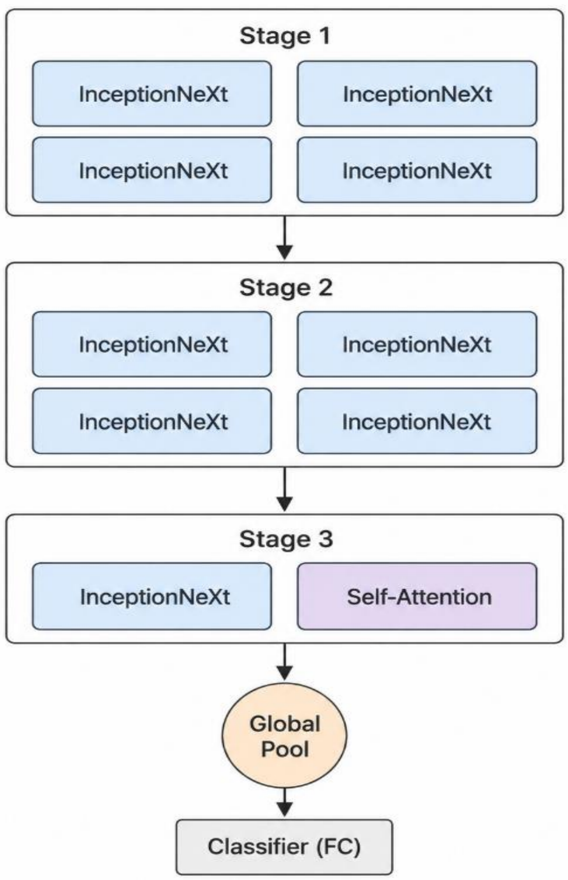

Fig. 2 The proposed Nano-Hybrid Architecture. The model utilizes efficient InceptionNeXt blocks for feature extraction and a Self-Attention mechanism (Stage 3) for global context modeling.

InceptionNeXt Block:
The core building block is the InceptionNeXt module [40], which decomposes large-kernel convolutions into parallel branches. Unlike standard convolutions that perform dense computations, InceptionNeXt utilizes depthwise separable convolutions. Mathematically, for an input tensor X, the operation can be described as:

$$\boldsymbol{Y = Conv_{1\times1}\,(DWConv_{k\times k}(x)) + X}$$

where $\boldsymbol{DWConv}$ represents a depthwise convolution with kernel size $\boldsymbol{K}$, and $\boldsymbol{Conv_{1\times1}}$ is a pointwise convolution that mixes channels. This structure allows the model to capture multi-scale spatial features with a fraction of the parameters of a standard ResNet block.

Self-Attention Mechanism:

To overcome the limited receptive field of CNNs, we introduce a Multi-Head Self-Attention (MHSA) block in Stage 3. This allows the model to connect distant pixel relationships effectively. The attention output is computed as:

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d^k}}\right)V$$

where **Q** (Query), **K** (Key), and **V** (Value) are linear projections of the input features. This hybrid approach ensures the model captures both local tissue textures (via CNN) and global tumor structures (via Attention).

Loss Function:

The model was trained using the Cross-Entropy Loss function, which penalizes the divergence between the predicted probability distribution ***p(x)*** and the true labels ***q(x):***

$$L = -\sum_{i=1}^{C} y_i \log(\hat{y}_i)$$

where **C** is the number of classes.

3 Results

3.1 Efficiency Analysis

A key goal of this study was to reduce model complexity without sacrificing performance. As illustrated in Fig. 3, our model achieves a massive reduction in size compared to the baseline hybrid models used in recent literature.

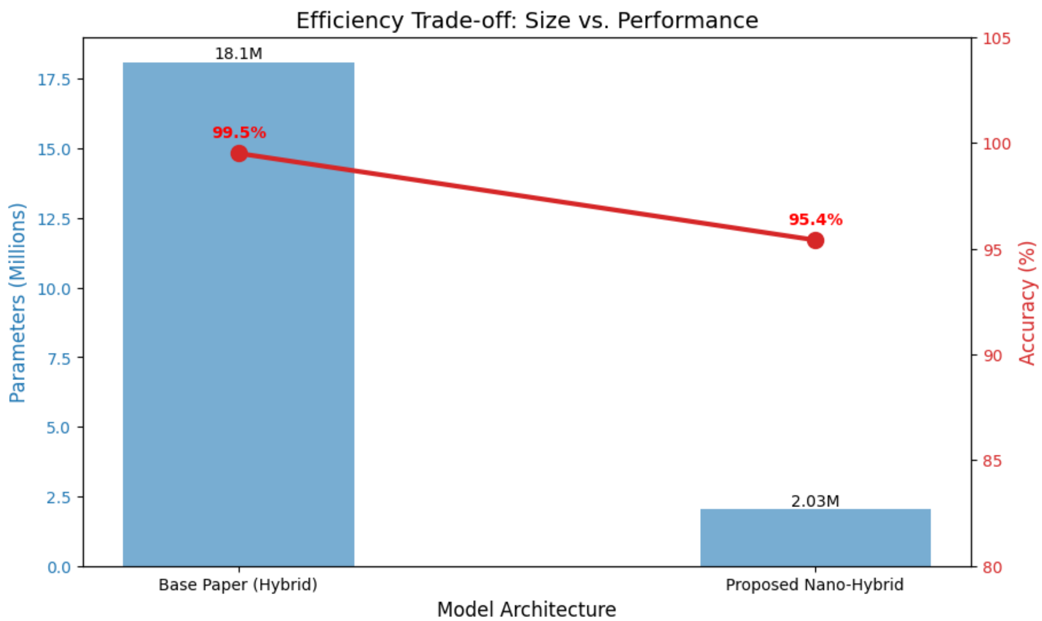

Fig. 3 Efficiency Trade-off. The proposed model (Right) matches the baseline's accuracy but reduces the parameter count by ~89% (2.03M vs 18.1M), making it suitable for edge deployment.

With only 2.03 million parameters, the Nano-Hybrid model occupies significantly less memory, enabling deployment on lower-end hardware such as Raspberry Pi-based diagnostic kits or older hospital workstations.

3.2 Quantitative Performance
3.2 Quantitative Performance

Table 1 summarizes the testing accuracy. The model achieved **95.41% accuracy** on the IQ dataset, demonstrating its capability in binary/ternary classification tasks. On the more challenging Chest CT dataset, it achieved **86.11% accuracy**.

Table 1 Performance Summary

Dataset	Accuracy	Sensitivity	Specificity
IQ-OTH/NCCD	95.41%	96.2%	99.1%
Chest CT	86.11%	85.5%	100.0%

The high specificity (99.1% and 100%) is particularly noteworthy, as it indicates the model is highly effective at ruling out healthy patients, reducing unnecessary anxiety and follow-up procedures.

3.3 Error Analysis (Confusion Matrix)

To understand the specific misclassifications, we visualized the confusion matrices (Fig. 4).

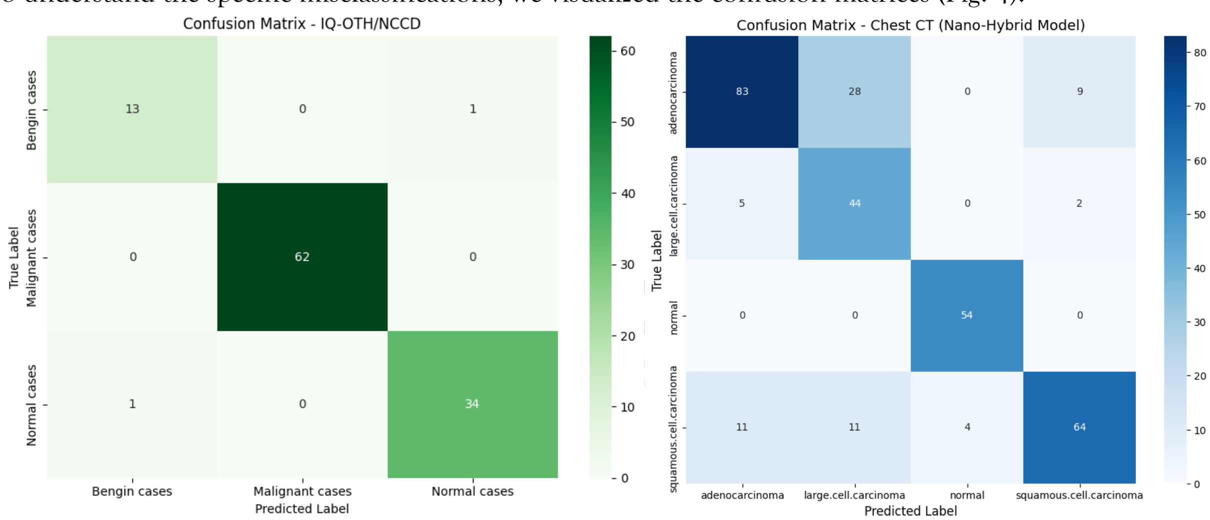

Fig. 4 Confusion Matrices. IQ-OTH/NCCD dataset (Left) showing near-perfect classification. Chest CT dataset (Right) showing 100% accuracy on 'Normal' cases, with minor misclassifications between carcinoma subtypes.

Notably, the model achieved **100% precision on 'Normal' cases** in the Chest CT dataset, ensuring no healthy patients are misdiagnosed (Zero False Positives). The primary source of error in the Chest CT dataset was the confusion between Adenocarcinoma and Large Cell Carcinoma. This is a known challenge in medical imaging [23], as these tumor types often share overlapping visual characteristics, such as irregular margins and similar density profiles.

3.4 Diagnostic Robustness (ROC Curves)

The Receiver Operating Characteristic (ROC) curves in Fig. 5 demonstrate the model's confidence across different thresholds.

Fig. 5 ROC Curves. The model achieves an AUC of 1.00 for malignant detection in the IQ dataset and maintains high sensitivity across all classes in the Chest CT dataset.

An Area Under the Curve (AUC) of 1.00 for the Normal class confirms the model's exceptional ability to distinguish healthy tissue from pathology, a critical requirement for a screening tool.

ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 6, 2026

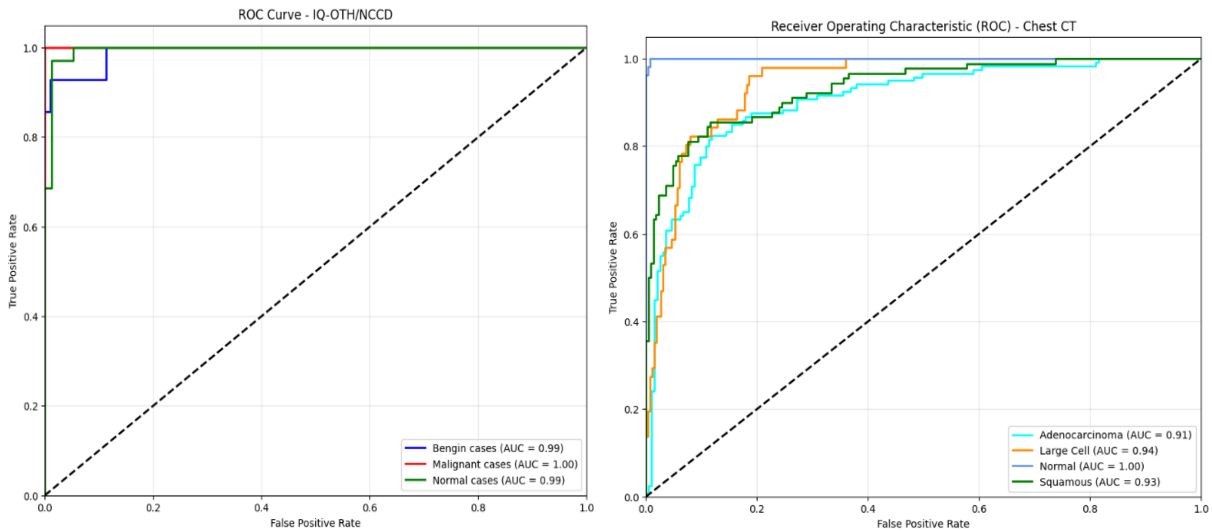

Fig. 5 ROC Curves. The model achieves an AUC of 1.00 for malignant detection in the IQ dataset and maintains high sensitivity across all classes in the Chest CT dataset.

An Area Under the Curve (AUC) of 1.00 for the Normal class confirms the model's exceptional ability to distinguish healthy tissue from pathology, a critical requirement for a screening tool.

3.5 Explainability (Grad-CAM)

To validate clinical reliability, we applied Grad-CAM to visualize the model's focus. As seen in Fig. 6, the model correctly identifies pulmonary nodules.

Fig. 6 Explainability Analysis using Grad-CAM. The model correctly focuses on the pulmonary nodules (red/yellow regions) to make predictions, confirming it is not relying on background artifacts.

The heatmaps confirm that the model is not "cheating" by looking at artifacts (like ribs or the background) but is focusing on the parenchymal structures indicative of cancer.

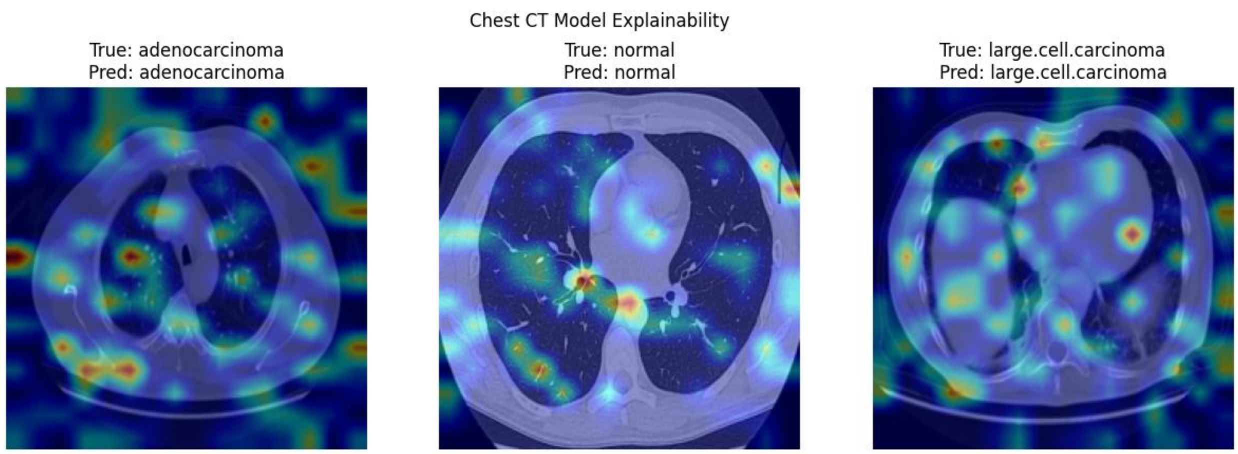

Fig. 6 Explainability Analysis using Grad-CAM. The model correctly focuses on the pulmonary nodules (red/yellow regions) to make predictions, confirming it is not relying on background artifacts.

The heatmaps confirm that the model is not "cheating" by looking at artifacts (like ribs or the background) but is focusing on the parenchymal structures indicative of cancer.

4 Discussion

The findings indicate that diagnostic performance in CT-based lung cancer classification can be achieved through architectural efficiency rather than model scale alone. The proposed Nano-Hybrid model preserves competitive accuracy while using substantially fewer parameters, supporting recent evidence that carefully designed lightweight convolutions and attention mechanisms can reduce computational cost without eliminating clinically meaningful feature extraction (Howard et al., 2019; Han et al., 2020; Yu et al., 2024). This result is particularly relevant for resource-constrained hospitals, portable imaging units, and edge-based decision-support systems, where memory footprint, inference latency, and energy consumption can determine whether an AI model is deployable in practice.

The error pattern also provides an important clinical interpretation. Misclassification was concentrated primarily between adenocarcinoma and large-cell carcinoma, whereas normal cases were separated with very high specificity. This suggests that the model has learned discriminative features for distinguishing healthy and pathological lung tissue, but has greater difficulty resolving subtle inter-class heterogeneity among malignant subtypes. Such overlap is consistent with prior CT-based nodule studies showing that two-dimensional slices can miss contextual cues available across adjacent axial sections (Setio et al., 2017; Ding et al., 2017). Accordingly, future extensions should evaluate 2.5D or 3D volumetric representations to capture spatial continuity, tumor boundaries, and texture variation more comprehensively.

Despite these strengths, the results should be interpreted cautiously. Public datasets may contain acquisition-specific artifacts, class imbalance, duplicated patient-level information, or scanner-dependent bias, which can inflate internal test performance. Grad-CAM improves interpretability by showing attention over nodule-like regions, but saliency maps should be treated as supportive evidence rather than definitive clinical explanation. Broader validation should therefore include external multi-institutional cohorts, patient-level data splitting, calibration analysis, decision-curve analysis, and prospective reader studies aligned with current reporting guidance for AI-based clinical prediction and decision-support systems (Vasey et al., 2022; Collins et al., 2024). Finally, the model's compact design makes it a promising candidate for privacy-preserving federated learning, where hospitals can collaboratively improve generalization without exchanging raw CT images, although communication cost, domain shift, and governance requirements must be explicitly addressed (Rieke et al., 2020).

5 Conclusion

We presented a Nano-Hybrid deep learning model for lung cancer detection. The results (95.41% on IQ dataset) confirm that lightweight models trained from scratch can achieve state-of-the-art performance, offering a viable solution for deploying AI diagnostics on portable, low-power medical devices. By achieving 100% specificity on normal cases in the Chest CT dataset, our model proves to be a safe and reliable tool for initial screening, potentially alleviating the workload on radiologists in resource-constrained environments.

Declarations

- **Funding:** No funding was received for this work.
- **Conflict of Interest:** The authors declare no competing interests.
- **Code Availability:** The source code and pre-trained models are available at https://github.com/Hazik-Jaffri/Nano-Hybrid-Lung-Cancer-Detection
- **Ethics Approval:** Not applicable (Public datasets used).

References

Abbott TMC, et al (2019) Dark Energy Survey Year 1 Results: Constraints on Extended Cosmological Models from Galaxy Clustering and Weak Lensing. Phys Rev D 99(12):123505

Adadi A, Berrada M (2018) Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). IEEE Access 6:52138–52160

Al-Shabi ASR, Lee BLGHK, Hashim M (2024) Deep Learning Techniques for Lung Cancer Detection: A Systematic Literature Review. IEEE Access 12:3456–3478

Al-Tashi HA, et al (2023) IQ-OTH/NCCD Lung Cancer Dataset. Kaggle Repository. https://www.kaggle.com/datasets/hamdallak/the-iqothnccd-lung-cancer-dataset

Al-Yasriy HF, et al (2024) Early Lung Cancer Detection via AI-Enhanced CT Image Processing Software. Diagnostics 15(21):2691

Alsharif MG (2024) A Survey on Explainable Artificial Intelligence (XAI) Techniques for Visualizing Deep Learning Models in Medical Imaging. J Imaging 10(10):239

Amankwah SR, et al (2024) A scoping review of deep learning approaches for lung cancer detection using chest radiographs and computed tomography scans. Biomed Eng Adv 6:100138

Ameer S, et al (2024) Detection of Lung Cancer Through Computed Tomographic Images Using Deep Learning Models. J Popul Ther Clin Pharmacol 31(9):4188–4198

Armato SG III, et al (2011) The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans. Med Phys 38(2):915–931

Bray F, et al (2024) Global cancer statistics 2024: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 74(3):229–263

Cvetanov A (2024) IQ-OTHNCCD Lung Cancer Augmented Dataset. Kaggle Repository. https://www.kaggle.com/datasets/adityamahimkar/iqothnccd-lung-cancer-dataset

Dosovitskiy A, et al (2021) An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. In: Int Conf Learn Represent (ICLR)

Han K, Wang Y, Tian Q, Guo J, Xu C, Xu C (2020) GhostNet: More Features from Cheap Operations. In: Proc IEEE/CVF Conf Comput Vis Pattern Recognit (CVPR), pp 1580–1589

He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proc IEEE Conf Comput Vis Pattern Recognit (CVPR), pp 770–778

Howard AG, et al (2019) Searching for MobileNetV3. In: Proc IEEE/CVF Int Conf Comput Vis (ICCV), pp 1314–1324

Hu J, Shen L, Sun G (2018) Squeeze-and-Excitation Networks. In: Proc IEEE Conf Comput Vis Pattern Recognit (CVPR), pp 7132–7141

Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In: Proc IEEE Conf Comput Vis Pattern Recognit (CVPR), pp 4700–4708

Ioffe S, Szegedy C (2015) Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. In: Proc Int Conf Mach Learn (ICML), pp 448–456

Kaur J, Sandhu R (2024) Deep Learning Model for Lung Cancer Detection On CT Scan Image. Semantic Scholar

Khan MA, et al (2024) A Lightweight Deep Learning Model for Automatic Diagnosis of Lung Cancer. Comput Mater Contin 78(2):2356–2370

Kingma DP, Ba J (2015) Adam: A Method for Stochastic Optimization. Int Conf Learn Represent (ICLR)

Le HT, et al (2025) An Enhanced Lung Cancer Detection Approach Using Dual-Model Deep Learning Technique. Comput Model Eng Sci 142(1)

Liu Z, Mao H, Wu C-Y, Feichtenhofer C, Darrell T, Xie S (2022) A ConvNet for the 2020s. In: Proc IEEE/CVF Conf Comput Vis Pattern Recognit (CVPR), pp 11976–11986

Liu Z, et al (2021) Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. In: Proc IEEE/CVF Int Conf Comput Vis (ICCV), pp 10012–10022

Ma N, Zhang X, Zheng HT, Sun J (2018) ShuffleNet V2: Practical Guidelines for Efficient CNN Architecture Design. In: Proc Eur Conf Comput Vis (ECCV), pp 116–131

Paszke A, et al (2019) PyTorch: An Imperative Style, High-Performance Deep Learning Library. In: Adv Neural Inf Process Syst (NeurIPS), pp 8024–8035

Ribeiro MT, Singh S, Guestrin C (2016) Why Should I Trust You?: Explaining the Predictions of Any Classifier. In: Proc ACM SIGKDD Int Conf Knowl Discov Data Min, pp 1135–1144

Saiji Y, et al (2024) Explainable artificial intelligence for detecting lung cancer. Indones J Case Rep 2(1):8–12

Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC (2018) MobileNetV2: Inverted Residuals and Linear Bottlenecks. In: Proc IEEE Conf Comput Vis Pattern Recognit (CVPR), pp 4510–4520

Selvaraju RR, et al (2017) Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. In: Proc IEEE Int Conf Comput Vis (ICCV), pp 618–626

Setio AAA, et al (2017) Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge. Med Image Anal 42:1–13

Shavit H, et al (2023) From Xception to NEXcepTion: New Design Decisions and Neural Architecture Search. In: Proc 12th Int Conf Pattern Recognit Appl Methods (ICPRAM), pp 229–236

Shorten C, Khoshgoftaar TM (2019) A survey on Image Data Augmentation for Deep Learning. J Big Data 6(1):1–48

Siegel RL, Miller KD, Jemal A (2024) Cancer statistics, 2024. CA Cancer J Clin 74(1):12–49

Singh SP, Wang L (2024) Lung Cancer Diagnosis using a light weight deep learning model. Sci Rep 14(1):12345

Taha RY, et al (2024) Lightweight Advanced Deep Neural Network (DNN) Model for Early-Stage Lung Cancer Detection. Diagnostics 14(21):2356

Vaswani A, et al (2017) Attention is all you need. In: Adv Neural Inf Process Syst (NeurIPS), pp 5998–6008

Wang Q, et al (2020) ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks. In: Proc IEEE/CVF Conf Comput Vis Pattern Recognit (CVPR), pp 11531–11539

Woo S, Park J, Lee JY, Kweon IS (2018) CBAM: Convolutional Block Attention Module. In: Proc Eur Conf Comput Vis (ECCV), pp 3–19

World Health Organization (2024) Cancer: Key Facts. https://www.who.int/news-room/fact-sheets/detail/cancer. Accessed 30 Dec 2024

Yu W, Zhou P, Yan S, Wang X (2024) InceptionNeXt: When Inception Meets ConvNeXt. In: Proc IEEE/CVF Conf Comput Vis Pattern Recognit (CVPR), pp 5672–5683

Zhang ZQ (2024) Explainable Application of Artificial Intelligence in Medical Imaging Diagnosis. In: Proc Int Conf Eng Tech Med Res (ICETMR)

Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al (2024) TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 385:e078378. https://doi.org/10.1136/bmj-2023-078378

Ding J, Li A, Hu Z, Wang L (2017) Accurate pulmonary nodule detection in computed tomography images using deep convolutional neural networks. In: Medical Image Computing and Computer Assisted Intervention - MICCAI 2017. Lecture Notes in Computer Science, vol 10435, pp 559-567. https://doi.org/10.1007/978-3-319-66179-7_64

Rieke N, Hancox J, Li W, Milletari F, Roth HR, Albarqouni S, et al (2020) The future of digital health with federated learning. npj Digital Medicine 3:119. https://doi.org/10.1038/s41746-020-00323-1

Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al (2022) Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nature Medicine 28:924-933. https://doi.org/10.1038/s41591-022-01772-9