International Journal of Computational Science and Engineering Research
ISSN: 3107 - 8605 (Online) , http://www.ijcser.com/
Regular Issue , Vol. 3, Issue. 3 (July – September) , 2026 , Pages: 8 - 18
*Received:* 26 August 2025 ; *Accepted:* 21 June 2026 ; *Published:* 03 July 2026
https://doi.org/10.63328/IJCSER-V3RI3P2
https://www.youtube.com/watch?v=HcsCGEp6u0c
https://archive.org/details/jsp-ijcser-v3ri3p2
Original Paper : Applied Research – Experimental Research Paper

# Lung Cancer Detection from CT Images Using ResNet50

**J.V. G. Prakasa Rao Pyla [1] , Dr. D. Mabuni [2]**

[1] Research Scholar, Department of Computer Science & Technology, Dravidian University, Kuppam, Andhra Pradesh, India ; prakash.pyla@gmail.com
[2] Assistant Professor , Department of Computer Science & Technology, Dravidian University, Kuppam, Andhra Pradesh, India ; mabuni.d@gmail.com

* *Corresponding Author : J.V.G. Prakasa Rao ; prakash.pyla@gmail.com*

**Abstract:** Lung cancer remained one of the leading causes of cancer-related mortality worldwide, making early diagnosis essential for improving treatment opportunities and patient outcomes. Computed tomography (CT) imaging became an important diagnostic modality because of its ability to capture detailed anatomical information from the pulmonary region. At the same time, advances in artificial intelligence and deep learning created opportunities for automated medical image analysis and decision support. Despite this progress, challenges, including limited labelled datasets, moderate generalisation capability, computational constraints, reproducibility concerns, and limited interpretability, continued to influence the development of reliable diagnostic frameworks. This research developed a transfer learning framework based on ResNet50 for lung cancer detection using CT scan images, evaluated classification performance through accuracy, precision, recall, and F1-score, and established a reproducible experimental baseline for future explainable and optimised diagnostic systems. A publicly available CT image dataset containing four diagnostic categories was utilised. Data preparation included image resizing, normalisation, and augmentation before model training. The experimental outcomes demonstrated progressive learning behaviour throughout training, where training accuracy reached 71.32% and validation accuracy reached 62.50%, accompanied by decreasing training and validation loss values. Evaluation on the testing dataset produced an overall classification accuracy of 57%, with weighted precision, recall, and F1-score values of 0.57, 0.57, and 0.50, respectively. Class-level analysis indicated stronger recognition performance for normal CT images and comparatively balanced detection capability for adenocarcinoma, whereas lower performance was observed for large cell carcinoma and squamous cell carcinoma. The findings suggested that transfer learning remained a practical approach for CT-based lung cancer classification under limited data conditions. In addition to experimental evaluation, the study contributed a reproducible implementation framework that may support future comparison, explainability integration, and continued optimisation of diagnostic models.

**Keywords:** Lung Cancer Detection, CT , ResNet50, Transfer Learning, DL, Medical Image Classification.

## 1. Introduction

Lung cancer remained a major public health challenge because of its high incidence and mortality across healthcare systems with different levels of development. Early detection remained clinically important because survival outcomes continued to depend strongly on the stage at which the disease was identified. Although diagnostic and treatment strategies continued to evolve, delayed identification frequently reduced treatment effectiveness and limited patient recovery opportunities [1]. Computed tomography (CT) emerged as one of the most effective imaging modalities for lung cancer detection because of its ability to capture detailed anatomical information and reveal pulmonary abnormalities that may remain unnoticed in conventional imaging approaches (Pehrson et al., 2019). Alongside advances in medical imaging, artificial intelligence and deep learning increasingly became supportive technologies in diagnostic decision-making and disease analysis (Zhi et al., 2023).

An effective automated lung cancer detection framework should operate reliably, consistently, and with sufficient reproducibility to support diagnostic workflows. Ideally, such systems would contribute to earlier diagnosis while

ISSN: 3107 - 8605 (Online) , http://www.ijcser.com/ , Vol. 3, Issue 3 , 2026 , https://doi.org/10.63328/IJCSER-V3RI3P2

reducing variation in interpretation and improving screening efficiency. Despite rapid progress in computational methods, several practical limitations remained unresolved [2]. Deep learning models often depended on limited annotated datasets, exhibited inconsistent performance across different environments, required substantial computational resources, and relied on extensive preprocessing procedures. Concerns surrounding explainability and reproducibility also continued to restrict broader confidence in automated clinical support systems (Khan et al., 2026; S et al., 2024).
Research on CT-based lung cancer classification has undergone considerable transformation over time. Earlier approaches depended heavily on machine learning algorithms supported by manually engineered features extracted from segmented regions of interest. As deep learning matured, convolutional neural networks (CNNs) gradually replaced handcrafted feature design by enabling automatic hierarchical feature extraction directly from medical images [3]. These developments improved classification capability and reduced dependency on manual descriptors, although performance continued to remain sensitive to dataset characteristics and experimental design choices (Al-Yasriy et al., 2020).

Transfer learning further accelerated development in this domain by allowing pretrained image recognition models to be adapted for medical image analysis tasks. Instead of training deep networks entirely from the beginning, transfer learning enabled previously learned feature representations to support classification under limited medical data conditions. Al-Huseiny and Sajit (2021) reported that transfer learning improved lung cancer detection performance while reducing computational burden compared with fully trained models [4]. Similar observations appeared across later studies, where pretrained architectures improved feature extraction efficiency but continued to experience challenges related to class imbalance, overfitting, and dependence on preprocessing strategies (Naseer et al., 2023; Zhi et al., 2023). More recent developments moved toward increasingly sophisticated architectures and multi-stage frameworks. Lung-EffNet demonstrated the application of EfficientNet to CT image classification and reported improved performance under optimised experimental settings (Raza et al., 2023). Other studies incorporated segmentation pipelines and hybrid feature extraction mechanisms to strengthen lesion identification and classification performance (Yang et al., 2026; Lakshmi & Nagaraj, 2025). Explainability approaches such as Grad-CAM were also introduced to improve transparency and interpretation of deep learning predictions (S et al., 2024). Although these methods demonstrated technical progress, many remained computationally intensive and comparatively difficult to reproduce in practical settings. Collectively, the literature suggested a persistent gap between experimentally optimised architectures and practically reproducible implementation frameworks. Existing evidence strongly supported the value of deep learning and transfer learning for lung cancer analysis, yet relatively fewer studies emphasised baseline implementations that balanced methodological simplicity with measurable classification performance [5]. Such reproducible frameworks remain important because they create a foundation for future optimization, explainability integration, and comparative evaluation across architectures. To address this gap, the present work employed a ResNet50-based transfer learning framework for CT image classification. ResNet50 was selected because residual learning supports efficient optimisation and stable feature propagation within deeper architectures. Within medical imaging contexts where labelled datasets are frequently constrained [6], transfer learning provided an effective mechanism for extracting useful image representations while reducing training requirements and maintaining computational feasibility (Al-Huseiny & Sajit, 2021; Zhi et al., 2023).

Based on these observations, the research pursued three objectives. First, it aimed to develop a transfer learning framework based on ResNet50 for lung cancer detection using CT images [7]. Second, it evaluated classification performance using standard measures including accuracy, precision, recall, and F1-score. Third, it sought to establish a reproducible experimental baseline that could support future explainable and optimised diagnostic systems. The contribution of this work extended beyond measuring classification performance alone [8]. From a practical perspective, automated CT image analysis may support AI-assisted screening and clinical decision support. From an academic perspective, the study contributed a reproducible experimental framework that may assist future work involving architecture comparison, explainable AI integration, and continued optimisation of lung cancer detection systems. This paper is organised as follows. Section 2 presents related literature and identifies research gaps. Section 3 describes the methodology and experimental procedures. Section 4 presents the experimental findings, while Section 5 discusses their implications, limitations, and future directions.

## 2. Literature Survey

Lung cancer detection using CT imaging continues to receive considerable research attention because of its direct relevance to early diagnosis, treatment planning, and improved patient outcomes. CT imaging provides high-resolution anatomical information capable of revealing pulmonary abnormalities that may remain undetected using conventional screening approaches. At the same

ISSN: 3107 - 8605 (Online) , http://www.ijcser.com/ , Vol. 3, Issue 3 , 2026 , https://doi.org/10.63328/IJCSER-V3RI3P2

time, advances in artificial intelligence and deep learning shifted medical image analysis from manually engineered workflows toward automated representation learning frameworks capable of extracting clinically meaningful features directly from imaging data [9]. This transition contributed to making automated lung cancer classification both feasible and increasingly valuable within medical imaging research (Pehrson et al., 2019; Zhi et al., 2023; Litjens et al., 2017).

Early approaches to lung cancer recognition relied primarily on conventional machine learning pipelines supported by handcrafted features and explicit segmentation procedures. Although these approaches demonstrated feasibility, their dependence on manual feature engineering reduced scalability and limited adaptability across datasets. With the growing adoption of convolutional neural networks (CNNs), research gradually moved toward architectures capable of learning hierarchical image representations directly from CT scans [10]. Al-Yasriy et al. (2020) demonstrated that CNN-based diagnosis improved classification capability while reducing dependence on handcrafted descriptors. At the same time, these developments highlighted a recurring challenge, namely that classification outcomes remained highly sensitive to training configuration, dataset composition, and preprocessing decisions.

Transfer learning later emerged as a practical response to the limited availability of labelled medical imaging data and became an important transition point in medical image classification research. Rather than developing deep networks entirely from the beginning, pretrained models enabled knowledge transfer from large-scale image datasets into domain-specific healthcare applications [11]. The theoretical foundation for this approach was established by Pan and Yang (2010), who suggested that transferring learned representations could improve performance under limited domain-specific data conditions. This concept proved particularly valuable in medical imaging. Shin et al. (2016) reported that transfer learning improved performance and optimisation in computer-aided diagnostic systems [12]. Similarly, Al-Huseiny and Sajit (2021) implemented transfer learning using GoogLeNet for lung cancer detection and demonstrated its practical applicability in medical image classification. As network architectures became deeper and more computationally demanding, research attention increasingly shifted toward understanding how architectural design influenced classification behaviour and learning efficiency [13]. One influential contribution emerged from residual learning introduced by He et al. (2016), who demonstrated that deeper models could be optimised effectively while preserving stable feature propagation. This development strongly influenced subsequent medical image analysis research and contributed to the widespread adoption of ResNet architectures. Dandil (2018) further proposed an automated computer-aided framework for lung cancer classification using CT scans and demonstrated promising performance. Nevertheless, limitations commonly observed across related work remained evident.

Comparative evidence further indicated that reported improvements were influenced not only by model architecture but also by differences in dataset composition, preprocessing decisions, and evaluation protocols. Pehrson et al. (2019), through a systematic review of pulmonary nodule detection studies using the LIDC-IDRI dataset, identified substantial variation in reported outcomes despite similar research goals [14]. Their analysis indicated that segmentation methods, evaluation procedures, and preprocessing choices often contributed as strongly to performance differences as the underlying network architecture. Similar observations appeared in later work by Khan et al. (2026), who showed that CT image intensity and contrast adjustments significantly affected lung nodule detection outcomes.

Recent developments have placed increasing emphasis on architecture optimisation and integrated learning pipelines. Naseer et al. (2023) proposed a modified framework based on U-Net architecture that incorporated lobe segmentation and nodule detection before classification. Although this improved feature extraction capability, it also increased model complexity and dependence on segmentation quality. Similarly, Lung-EffNet, developed by Raza et al. (2023) demonstrated that EfficientNet-based architectures could achieve improved predictive performance under optimised experimental conditions [15]. Another notable direction in recent literature involved improving interpretability and transparency within diagnostic models. Earlier studies largely emphasised predictive performance, whereas newer approaches increasingly considered explainability. Yogesh Kumaran et al. (2024) introduced an ensemble transfer learning framework combining VGG16, ResNet50, and InceptionV3 with Grad-CAM visualisation techniques to improve interpretability in lung cancer classification. Their findings suggested that explainability mechanisms could strengthen trust and interpretability without necessarily reducing predictive capability. Similar observations appeared across broader medical imaging reviews.

Beyond individual architectures, broader review studies highlighted persistent methodological challenges affecting reproducibility and comparison across studies [16]. Zhi et al. (2023), through literature review and experimentation on pulmonary nodule segmentation methods, reported that variations in datasets, preprocessing decisions, and evaluation metrics continued to

ISSN: 3107 - 8605 (Online) , http://www.ijcser.com/ , Vol. 3, Issue 3 , 2026 , https://doi.org/10.63328/IJCSER-V3RI3P2

complicate reproducibility and comparison. These factors reduced confidence in direct performance comparisons across studies [17]. In addition, many investigations prioritised overall accuracy while giving comparatively less attention to balanced measures such as precision, recall, and F1-score. Across the reviewed studies, several recurring patterns became evident. Deep learning and transfer learning consistently improved feature extraction capability relative to traditional approaches, yet performance gains could not be attributed solely to architecture selection. Dataset characteristics, preprocessing pipelines, training procedures, explainability mechanisms, and evaluation criteria all played meaningful roles in determining outcomes (Pehrson et al., 2019; Zhi et al., 2023; Raza et al., 2023). At the same time, conflicting findings persisted regarding whether increasingly complex pipelines offered meaningful advantages over simpler transfer learning implementations.

Despite continued methodological progress, several unresolved issues remained and directly motivated the present research. Many existing studies concentrated on maximising algorithmic performance while introducing architectural and computational complexity that reduced reproducibility. Cross-study comparison remained difficult because of variations in datasets and experimental conditions. In addition, despite increasing attention toward explainable artificial intelligence, comparatively fewer efforts have focused on establishing reproducible baseline frameworks suitable for future extension. Accordingly, the present work positioned itself as a practical continuation of existing efforts rather than a claim of methodological novelty [18]. The research developed a ResNet50-based transfer learning framework for lung cancer detection using CT images, evaluated performance using accuracy, precision, recall, and F1-score metrics, and established a reproducible baseline intended to support future explainable diagnostic research. In summary, the literature consistently demonstrated the growing importance of deep learning and transfer learning for CT-based lung cancer detection while simultaneously revealing challenges related to reproducibility, balanced evaluation, and interpretability [19]. The present study responded to these observations by adopting a ResNet50-based transfer learning framework intended to provide a reproducible experimental baseline for future explainable and optimised diagnostic research.

## 3. Research Methodology

### *3.1. Research Design and Experimental Setting*

This research adopted an experimental deep learning framework to investigate automated lung cancer detection from computed tomography (CT) images through transfer learning. An experimental design was selected because the objective was to develop, train, and evaluate an image classification model under controlled computational conditions [20]. Deep learning approaches have become increasingly relevant in medical imaging because of their capability to learn hierarchical representations directly from image data without relying on manually engineered features (Litjens et al., 2017). Existing literature emphasised not only improvements in detection performance but also the importance of reproducible, balanced, and interpretable frameworks. In response to these considerations, the methodology implemented a practical ResNet50-based transfer learning framework designed to support future extension toward explainable diagnostic systems. Transfer learning and CNN architectures have shown practical suitability for medical image classification under limited data conditions (Shin et al., 2016). The experiment was conducted using an openly available chest CT image dataset obtained through Kaggle and containing four diagnostic categories: adenocarcinoma, large cell carcinoma, normal, and squamous cell carcinoma. Model development [21], training, and testing were performed using cloud-based computational resources with GPU acceleration to support efficient training and reproducible execution.

### *3.2. Dataset Description and Image Preprocessing*

A publicly available CT image dataset containing four diagnostic categories, namely adenocarcinoma, large cell carcinoma, normal, and squamous cell carcinoma, was employed in the experiment. The selection of a public dataset supported transparency, reproducibility, and comparability with related medical image classification studies. Before training, all CT images underwent a structured preprocessing pipeline. Images were resized to satisfy the input dimensional requirements of the pretrained ResNet50 architecture. Standard normalisation procedures were applied to scale pixel intensity values and stabilise optimisation during model learning. Similar normalisation strategies have been widely adopted in transfer learning-based medical imaging studies because they improve convergence behaviour and reduce sensitivity to input variation (Shorten & Khoshgoftaar, 2019). Data preparation additionally incorporated augmentation procedures to improve image diversity and reduce overfitting risk [22]. Augmentation operations were applied during training through controlled transformation of images while preserving medically meaningful characteristics. Previous studies demonstrated that augmentation improves model robustness and supports generalisation under limited training conditions (Perez & Wang, 2017). To independently monitor learning progression and evaluate predictive performance, the dataset was divided into training, validation, and testing subsets. The training partition was used for parameter optimization [23], whereas the validation subset supported monitoring and hyperparameter observation during model development.

### *3.3. Deep Learning Architecture and Transfer Learning Technique*

The proposed classification framework employed transfer learning using the pretrained ResNet50 architecture for CT-based lung cancer detection. ResNet50 was originally developed to address optimisation degradation in deep neural networks through residual learning mechanisms that support efficient feature propagation (He et al., 2016). The architecture was selected because of its demonstrated performance in image classification and its increasing applicability to medical imaging environments where labelled data availability is often constrained [24]. Transfer learning enabled reuse of feature representations learned from large-scale datasets and adapted them to the target CT image classification task [Figure. 1]. This approach was particularly suitable because manually annotated medical image collections are frequently limited (Pan & Yang, 2010; Shin et al., 2016).

The architecture of the proposed model is illustrated in Figure 1. Before entering the ResNet50 backbone, CT images passed through preprocessing stages including resizing, normalisation, and augmentation. The network initialization used pretrained ImageNet weights [25], while the original classification head was removed to support multi-class classification requirements. Additional classification layers were incorporated into the pretrained feature extraction network.

Global Average Pooling was applied to reduce feature dimensionality before representation learning. Feature outputs were processed through a Dense layer containing 256 neurons with ReLU activation. A Dropout layer with a rate of 0.5 was included to reduce overfitting. Final classification probabilities across adenocarcinoma, large cell carcinoma, normal, and squamous cell carcinoma classes were generated using a Softmax output layer [26]. Training included selective fine-tuning in which earlier network layers retained general visual representations while later layers adapted to CT-specific imaging characteristics.

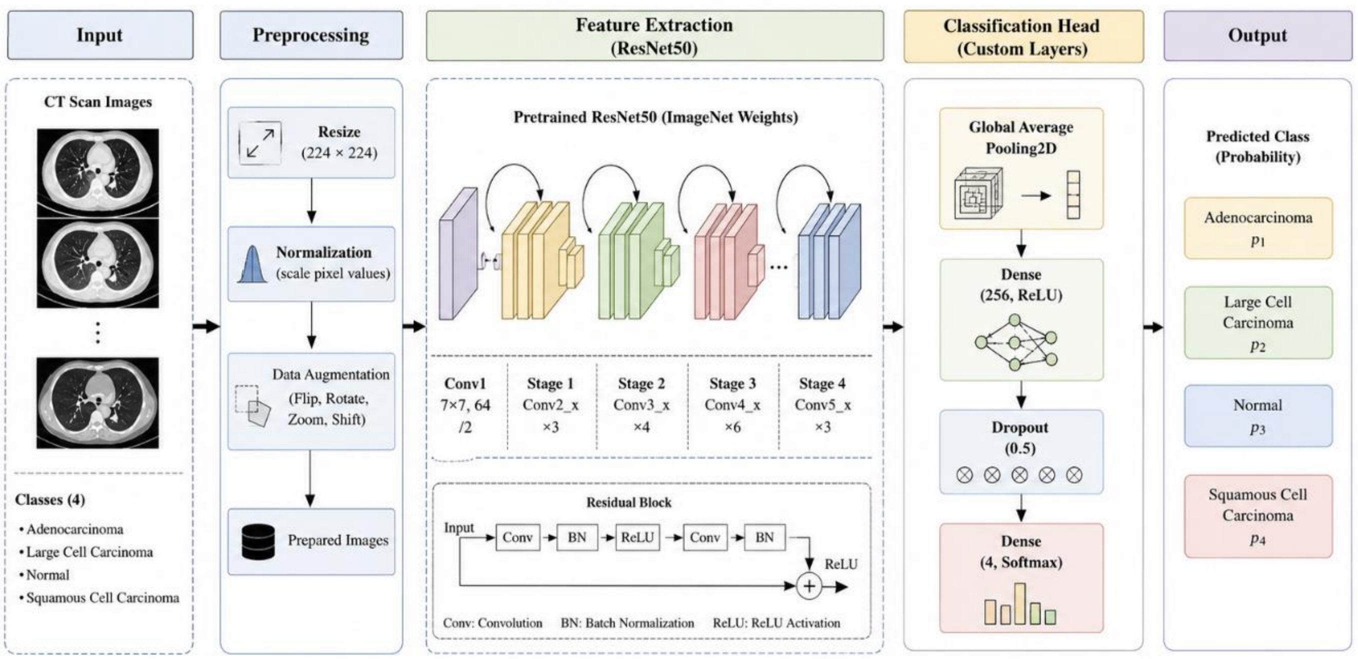



**Figure. 1** Proposed ResNet50-Based Transfer Learning Architecture for Detecting Lung Cancer via CT Scans.

### *3.4. Steps for Training Model*

Model training was performed after completion of the preprocessing and architecture configuration stages. The Adam optimizer was selected because of its adaptive learning capability and stable convergence characteristics (Kingma & Ba, 2015). Learning rate control mechanisms were implemented to maintain stable optimisation behaviour throughout training. Categorical cross-entropy loss was employed as the optimisation objective for measuring divergence between predicted outputs and target categories. Mini-batch training was conducted with the selected batch size for 25 epochs [27]. Callback mechanisms for learning rate adjustment and training monitoring were incorporated to improve optimisation stability and reduce unnecessary parameter updates. GPU-supported execution was used throughout experimentation to improve computational efficiency and reduce training duration. These training decisions aligned with commonly adopted practices in deep learning-based medical image classification studies (Goodfellow et al., 2016).

ISSN: 3107 - 8605 (Online) , http://www.ijcser.com/ , Vol. 3, Issue 3 , 2026 , https://doi.org/10.63328/IJCSER-V3RI3P2

### *3.5. Evaluation Approach and Criteria*

Model performance was evaluated using multiple complementary metrics to obtain a balanced assessment of classification behaviour. Accuracy was employed to measure overall classification correctness across all diagnostic categories. Precision evaluated the reliability of positive predictions, while recall assessed the ability of the framework to identify relevant diagnostic classes [28]. F1-score was selected because it provides a balanced assessment of precision and recall and reduces misleading interpretation under uneven class distributions (Sokolova & Lapalme, 2009).

Evaluation further included analysis of training and validation accuracy curves, together with loss behaviour, to assess learning progression and identify possible overfitting trends. As illustrated in Figure-4, Confusion matrix analysis was performed to obtain class-level insight into prediction behaviour and misclassification tendencies. This evaluation strategy enabled a broader assessment of classification performance and aligned with accepted medical image analysis practices [29]. The overall methodological framework was designed to prioritise reproducibility, computational practicality, and transparent experimental implementation.

## 4. Results and Discussions

The experimental evaluation examined the performance of the proposed ResNet50-based transfer learning framework for lung cancer detection using CT scan images. Performance assessment was conducted through analysis of the learning process and multiclass classification outcomes across training, validation, and testing stages. The model was trained for 25 epochs using pre-processed CT images [Figure. 2]., and performance was measured as illustrated in Figure-2, Figure-3, Figure-4 and Table-1 using accuracy, loss, confusion matrix analysis, precision, recall, and F1-score, respectively [30]. Training behaviour demonstrated progressive improvement throughout the learning process. Training accuracy increased from approximately 44% to 71.32%, indicating continuous adaptation of the model to image classification patterns within the training data.

At the same time, training loss decreased from approximately 1.17 to 0.68, reflecting reduced prediction error during optimisation and stable convergence of the training process using transfer learning with pretrained ResNet50. Validation performance followed a different but gradually improving trend. During the initial training stages, validation accuracy exhibited fluctuations before becoming comparatively more stable in later epochs. As illustrated in Figure 2, **validation** accuracy ultimately reached 62.50%, and as illustrated in Figure 3, while validation loss decreased from approximately 1.53 to 0.85. Minor variations remained during later epochs, although the learning curves indicated reduced separation between training and validation behaviour across the experimental process.

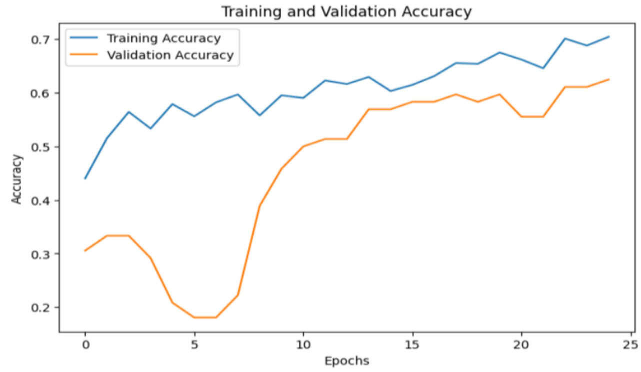



**Figure. 2** Training and Validation Accuracy Curve


**Figure. 3** Training and Validation Loss Curve

Testing performance was evaluated using an independent test dataset consisting of 315 CT scan images. Under the current experimental configuration, the framework achieved a final testing accuracy of 57%, indicating measurable multiclass classification capability across the four diagnostic categories. To provide a broader evaluation beyond overall accuracy, additional classification measures, including precision, recall, and F1-score, were calculated. Weighted performance values for precision, recall, and F1-score reached 0.57, 0.57, and 0.50, respectively. Macro-averaged values were recorded as 0.59 for precision, 0.56 for recall, and 0.51 for F1-score. Class-level evaluation revealed differences in predictive behaviour across diagnostic categories. The normal class demonstrated the strongest recognition performance, achieving precision, recall, and F1-score values of 0.64, 1.00, and 0.78, respectively, indicating complete identification of normal CT images within the testing subset. The adenocarcinoma category produced comparatively balanced results with precision, recall, and F1-score values of 0.52, 0.82, and 0.64. Performance differed across the remaining cancer categories [Table.1]. Large cell carcinoma achieved a

ISSN: 3107 - 8605 (Online) , http://www.ijcser.com/ , Vol. 3, Issue 3 , 2026 , https://doi.org/10.63328/IJCSER-V3RI3P2

precision of 0.63, a recall of 0.33, and an F1-score of 0.44. Squamous cell carcinoma recorded a precision of 0.56, recall of 0.10, and an F1-score of 0.17, indicating lower recognition consistency relative to other categories within the experimental dataset. As illustrated in Figure-4, the confusion matrix provided additional insight into category-level prediction behaviour and classification distribution patterns. The framework correctly identified 98 adenocarcinoma images, 17 large cell carcinoma images, 54 normal images, and 9 squamous cell carcinoma images. Misclassification occurred more frequently for squamous cell carcinoma, where several instances were assigned to alternative diagnostic categories. In comparison, normal CT images demonstrated stronger classification consistency, while adenocarcinoma predictions remained comparatively stable.

| Actual Class ↓ / Predicted Class → | Adenocarcinoma | Large Cell Carcinoma | Normal | Squamous Cell Carcinoma |
|---|---|---|---|---|
| Adenocarcinoma | 98 | 8 | 10 | 4 |
| Large Cell Carcinoma | 24 | 17 | 7 | 3 |
| Normal | 0 | 0 | 54 | 0 |
| Squamous Cell Carcinoma | 66 | 2 | 13 | 9 |

**Figure. 4** Confusion Matrix of CT Scan Classification Results

A detailed summary of precision, recall, F1-score, and support values for each diagnostic category is presented in Table 1.

| Class | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|
| Adenocarcinoma | 0.52 | 0.82 | 0.64 | 120 |
| Large Cell Carcinoma | 0.63 | 0.33 | 0.44 | 51 |
| Normal | 0.64 | 1.00 | 0.78 | 54 |
| Squamous Cell Carcinoma | 0.56 | 0.10 | 0.17 | 90 |
| Accuracy | — | — | 0.57 | 315 |
| Macro Average | 0.59 | 0.56 | 0.51 | 315 |
| Weighted Average | 0.57 | 0.57 | 0.50 | 315 |

**Table. 1** Classification Report

Overall, the experimental findings demonstrated that the ResNet50-based transfer learning framework learned discriminative visual representations from CT scan images and produced measurable multiclass classification performance under the selected experimental conditions. The reported learning curves, as illustrated in Figure 4 confusion matrix outcomes, and as illustrated in Table-1 evaluation metrics collectively reflected consistent execution of the experimental pipeline and reproducible reporting of model behaviour.

### 4.1. Discussion

The experimental findings demonstrated that the proposed ResNet50-based transfer learning framework achieved measurable performance in multiclass lung cancer detection using CT scan images. As illustrated in Figure 2 Training accuracy increased to 71.32% after 25 epochs, while validation accuracy reached 62.50%, accompanied by progressive reduction in both training and as illustrated in Figure 3 validation loss values throughout the learning process. These observations suggested that pretrained visual representations were successfully adapted to the CT classification task. At the same time, the remaining difference between training and validation outcomes indicated that improvements in learning performance did not translate uniformly across unseen data. Similar behaviour has been reported in prior deep learning studies in medical imaging, where model adaptation remained sensitive to dataset composition and experimental settings (Litjens et al., 2017; Al-Huseiny et al., 2021). Evaluation on the independent testing dataset provided additional perspective regarding classification capability under practical conditions. The framework achieved testing accuracy of 57%, together with precision and recall values of 0.57 and an F1-score of 0.50. Macro-averaged precision, recall, and F1-score values reached 0.59, 0.56, and 0.51, respectively. Rather than indicating uniform predictive behaviour across all categories, these outcomes reflected differences in how effectively the model learned discriminative characteristics across diagnostic classes.

Class-level analysis further clarified these variations. The Class-wise classification performance is summarised in Table 1. The normal category achieved the strongest recognition performance, reaching precision, recall, and F1-score values of 0.64, 1.00, and 0.78, respectively. Adenocarcinoma also demonstrated comparatively balanced behaviour with a precision of 0.52, a recall of 0.82, and an F1-score of 0.64. In contrast, large cell carcinoma produced acceptable precision but lower recall performance, while squamous cell carcinoma recorded substantially lower recall despite moderate precision values. As illustrated in Figure-4, the Confusion matrix analysis revealed that squamous cell carcinoma images were frequently assigned to alternative categories, particularly adenocarcinoma. Similar challenges associated with morphological overlap and uneven recognition behaviour have been reported in CT-based lung cancer classification studies (Dandil, 2018; Al-

ISSN: 3107 - 8605 (Online) , http://www.ijcser.com/ , Vol. 3, Issue 3 , 2026 , https://doi.org/10.63328/IJCSER-V3RI3P2

Yasriy et al., 2020). When viewed alongside earlier research, the obtained performance remained lower than studies reporting higher classification accuracy through larger datasets, segmentation-assisted pipelines, extensive augmentation procedures, or architecture optimisation approaches (Pehrson et al., 2019; Naseer et al., 2023; Raza et al., 2023). Direct comparison should nevertheless be interpreted carefully because experimental conditions varied substantially across studies. Several earlier investigations relied on curated datasets such as LIDC-IDRI, incorporated explicit nodule segmentation stages, or employed more extensive preprocessing and optimisation strategies. The present experiment instead prioritised implementation simplicity and reproducibility by adopting a ResNet50 transfer learning framework using publicly available CT images. Under these conditions, the observed behaviour aligned more closely with studies showing that transfer learning performance remained strongly dependent on dataset characteristics and preprocessing choices (Shin et al., 2016; Zhi et al., 2023).

From a practical perspective, the experiment suggested that transfer learning continued to provide meaningful value for medical image classification even under constrained data conditions. Reduction in loss values together with improvements in validation performance indicated that pretrained representations contributed useful image features without requiring complete model training from the beginning. These observations remained consistent with earlier findings showing that pretrained CNN architectures reduced computational requirements while supporting feature transfer in medical imaging environments (Pan & Yang, 2010; Litjens et al., 2017).

Several limitations should be considered when interpreting the present findings. Classification accuracy remained moderate, which may have been influenced by limited dataset size and uneven class distribution. Although learning improved during training, as illustrated in Figure 2, differences between training and validation outcomes suggested some degree of overfitting. In addition, advanced augmentation procedures and explainability approaches such as Grad-CAM were not incorporated into the current experimental framework. Earlier work suggested that interpretable diagnostic models may strengthen confidence and understanding of prediction behaviour in medical applications (S et al., 2024). Future work should therefore focus on strengthening augmentation strategies, extending fine-tuning of pretrained layers, and conducting comparative evaluation using architectures such as EfficientNet and DenseNet under controlled experimental conditions. Integration of explainable artificial intelligence methods, including Grad-CAM, together with evaluation using larger datasets such as LIDC-IDRI, may further improve interpretability and support more consistent classification outcomes (Pehrson et al., 2019; Raza et al., 2023).

## 5. Conclusion and Future Work

The present investigation examined the applicability of a ResNet50-based transfer learning framework for automated lung cancer detection using CT scan images. The study pursued three objectives: developing a transfer learning framework for multiclass lung cancer classification, evaluating performance using accuracy, precision, recall, and F1-score, and establishing a reproducible baseline that could support future development of diagnostic systems.

Experimental outcomes demonstrated that the framework learned meaningful visual representations from CT images and achieved measurable multiclass classification performance under controlled experimental conditions. The training performance improved progressively throughout the learning process, reaching a training accuracy of 71.32% with training loss reduced to approximately 0.68, indicating successful adaptation of pretrained feature representations to the target classification task. The Validation performance also improved during experimentation and reached 62.50% accuracy, together with reduced validation loss, as illustrated in Figure 3, suggesting that the framework maintained learning capability across unseen samples despite intermediate fluctuations.

Class-level evaluation provided additional insight into the behaviour of the proposed framework. The normal CT category achieved the strongest recognition performance and demonstrated complete recall within the testing subset. Adenocarcinoma classification also remained comparatively balanced, indicating that the extracted feature representations supported recognition of this category under the selected conditions. At the same time, lower recall and F1-score values observed for large cell carcinoma and squamous cell carcinoma reflected remaining challenges in distinguishing among certain diagnostic categories. Confusion matrix analysis further highlighted differences in recognition consistency and class overlap. From a broader perspective, the findings contributed to ongoing research at the intersection of deep learning and AI-assisted medical imaging by indicating that transfer learning remained applicable even when domain-specific medical data were limited. Consistent with earlier work in medical image analysis, pretrained architectures continued to reduce training requirements while preserving useful feature extraction capability (Shin et al., 2016; Litjens et al., 2017). Rather than proposing a new architecture, the contribution of this work centred on establishing a comparatively reproducible implementation

ISSN: 3107 - 8605 (Online) , http://www.ijcser.com/ , Vol. 3, Issue 3 , 2026 , https://doi.org/10.63328/IJCSER-V3RI3P2

framework that may serve as a reference point for future experimentation.

Several limitations remained important when interpreting the reported outcomes. Classification performance remained moderate and indicated opportunities for further improvement. Limited dataset size, potential class imbalance, fluctuations during validation, restricted augmentation procedures, and the absence of explainable artificial intelligence components represented factors that may have influenced experimental behaviour. These challenges have also been discussed in related lung cancer classification research. Future investigations may strengthen the current framework through more robust augmentation strategies, deeper fine-tuning of pretrained layers, and comparative evaluation using architectures such as EfficientNet and DenseNet under consistent experimental settings. Integration of explainable artificial intelligence approaches, including Grad-CAM, may further improve interpretability. Evaluation on larger public datasets such as LIDC-IDRI may additionally support broader assessment and more consistent classification performance. Taken together, the present work contributed a reproducible transfer learning implementation for CT-based lung cancer detection and provided a practical foundation for future explainable and optimised diagnostic research.

## References


[1]. Al-Huseiny, M. S., & Sajit, A. S. (2021). Detection of Lung Cancer through GoogLeNet Transfer Learning Model.

[2]. Al-Yasriy, H. F., Al-Husieny, M. S., Mohsen, F. Y., Khalil, E. A., & Hassan, Z. S. (2020). Diagnosis of lung cancer using CNN through CT scans.

[3]. Dandil, E. (2018). Computer-Aided Pipeline for Automatic Lung Cancer Classification Based on Computed Tomography Images. Journal of Healthcare Engineering. https://doi.org/10.1155/2018/9409267

[4]. Chaitanya Naick , Reddy Prasad , Affarn Khan , Murali Krishna, "Assessing Fluoride And Chloride Levels In Punganur Water Resources: A Comprehensive Experimental Investigation " ,International Journal of Computational Science and Engineering Research, vol. 1, no. 3, p. 1, July. 2024, doi: https://doi.org/10.63328/IJCSER-V1RI3P1

[5]. M. Dharani , Nasreen Sultana Quadri , T. Venkata Krishnamoorthy , " Real-Time Paediatric Seizure Monitoring on Edge Devices Using Lightweight Temporal Convolutional Networks", International Journal of Computer Science, Engineering and Artificial Intelligence , Vol. 3, Issue. 1 ( January – March), Pages: 8 – 12, 2026 . DOI: https://doi.org/10.63328/IJCSEAI-V3RI1P2

[6]. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

[7]. D. B. Bhuvannagari and S. M, "Automated and Explainable Kidney Abnormality Detection from CT Images Using CNN-LSTM Architecture," International Journal of Research and Development in Engineering Sciences, vol. 7, no. 4, p. 39, Aug. 2025, doi: 10.63328/ijrdes-v7ri4p7.

[8]. Sivaram Murugan , Graph Neural Networks with Riemannian Brain Connectivity Analysis for EEG-Based Neurological Classification , International Journal of Computer Science , Engineering and Artificial Intelligence , Vol. 2, Issue. 4 ( October – December ) , 2025 , Pages: 1 – 6. DOI : https://doi.org/10.63328/IJCSEAI-V2RI4P1

[9]. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

**[10].** Khan, S., Noor, M. N., Ashraf, I., Masud, M. I., & Aman, M. (2026). Impact of CT intensity and contrast variability on deep-learning-based lung-nodule detection: A systematic review of preprocessing and harmonization Strategies (2020–2025) *Diagnostics*, *16*(2), 201, https://doi.org/10.3390/diagnostics16020201

[11]. Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimisation. International Conference on Learning Representations (ICLR).

[12]. R. K. S, T. Y, and S. C. U, "Implementation of Pulmonary Disease Detection Model using Deep Learning," International Journal of Research and Development in Engineering Sciences, vol. 7, no. 2, p. 19, Mar. 2025, doi: 10.63328/ijrdes-v7ri2p4.

[13]. Ravi Samikannu, Retrieval-Augmented Multi-Agent Architecture for Hallucination Reduction in Large Language Models, International Journal of Computer Science , Engineering and Artificial Intelligence , Vol. 2, Issue. 4 ( October – December) , 2025 , Pages: 7 – 13. DOI : https://doi.org/10.63328/IJCSEAI-V2RI4P2

[14]. Lakshmi, G., & Nagaraj, P. (2025). Lung cancer detection and classification using optimized CNN features and Squeeze-Inception-ResNeXt model.

[15]. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., Van der Laak, J. A. W. M., Van Ginneken, B., & Sánchez, C. I. (2017). *A survey on deep learning in medical image analysis*. *Medical Image Analysis*, *42*, 60–88. https://doi.org/10.1016/j.media.2017.07.005

[16]. Ujval Mutyala ," Embodied Agentic AI for Autonomous Task Planning and Execution in Dynamic Environments ", International Journal of Computer Science , Engineering and Artificial Intelligence , Vol. 2, Issue. 4 ( October – December ) , 2025 , Pages: 22 – 30. DOI : https://doi.org/10.63328/IJCSEAI-V2RI4P4

[17]. Naseer, I., Akram, S., Masood, T., Rashid, M., & Jaffar, A. (2023). Lung cancer classification using modified U-Net based lobe segmentation and nodule detection.

[18]. Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. IEEE Transactions on Knowledge and Data Engineering, 22(10), 1345–1359.

[19]. Pehrson, L. M., Nielsen, M. B., & Lauridsen, C. A. (2019). Automatic pulmonary nodule detection applying deep learning or machine learning algorithms to the LIDC-IDRI database: A systematic review.

[20]. Perez, L., & Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning, *arXiv*. https://doi.org/10.48550/arXiv.1712.04621.

[21]. Raza, R., Zulfiqar, F., Khan, M. O., Arif, M., Alvi, A., Iftikhar, M. A., & Alam, T. (2023). Lung-EffNet: Lung cancer classification using EfficientNet from CT-scan images.

[22]. Shin, H. C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., & Summers, R. M. (2016). Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Transactions on Medical Imaging, 35(5), 1285–1298.

[23]. Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. Journal of Big Data, 6(1).

[24]. Sokolova, M., & Lapalme, G. (2009). Beyond accuracy, precision, recall and F-measure: A review of evaluation metrics.

[25]. Yang, X., Duan, A., Jiang, Z., Li, X., Wang, C., Wang, J., & Zhou, J. (2026). Segmentation and classification of lung cancer images using deep learning. *Applied Sciences*, *16*(2), 628. https://doi.org/10.3390/app16020628.

[26]. Yogesh Kumaran, S., Jospin, J. J., Mahesh, T. R., Bhatia, S., Alzahrani, S., & Alojail, M. (2024). Explainable lung cancer classification with ensemble transfer learning of VGG16, ResNet50 and InceptionV3 using Grad-CAM.

[27]. Zhi, L., Jiang, W., Zhang, S., & Zhou, T. (2023). Deep neural network pulmonary nodule segmentation methods for CT images: Literature review and experimental

ISSN: 3107 - 8605 (Online) , http://www.ijcser.com/ , Vol. 3, Issue 3 , 2026 , https://doi.org/10.63328/IJCSER-V3RI3P2

[28]. Lava Kumar Vandrangi ," Neuromorphic Event-Camera Systems for Ultra-Low Latency Aero-Elastic Monitoring in High-Speed Vehicles " , International Journal of Computer Science , Engineering and Artificial Intelligence , Vol. 2, Issue. 2 ( April – June ) , 2025 , Pages: 1 – 7. DOI : https://doi.org/10.63328/IJCSEAI-V2RI2P1

[29]. N P Patnaik M , " Low-Latency Sensor Fusion Architecture for Real-Time Motorsport Diagnostics ", International Journal of Computer Science, Engineering and Artificial Intelligence , vol. 1, no. 1, p. 1-7, December 2024, DOI: https://doi.org/10.63328/IJCSEAI-V1RI1P1

[30]. H. K. Yenugonda, S. R. V. Reddy, G. D, and G. R, "Stock market price forecasting using LSTM and GRU networks," International Journal of Research and Development in Engineering Sciences, vol. 7, no. 2, Mar. 2025, doi: 10.63328/ijrdes-v7ri2p3.

## How To Cite:

J.V. G. Prakasa Rao Pyla 1 , Dr. D. Mabuni , " Lung Cancer Detection from CT Images Using ResNet50 ", International Journal of Computational Sciences and Engineering Research , Volume 3, issue 3 , pp 8 - 18 , July 2026.
CrossRef DOI: https://doi.org/10.63328/ijcser-v3ri3p2

## Declaration

**Conflicts of Interest:** The authors declare no conflict of interest.

**Author Contribution:** All authors wrote the main manuscript text and also consent to the submission.

**Ethical approval:** Not applicable.

**Consent to Participate:** All authors consent to participate.

**Funding:** Not applicable, and No funding was received

**Institutional Review Board Statement:** Not applicable.

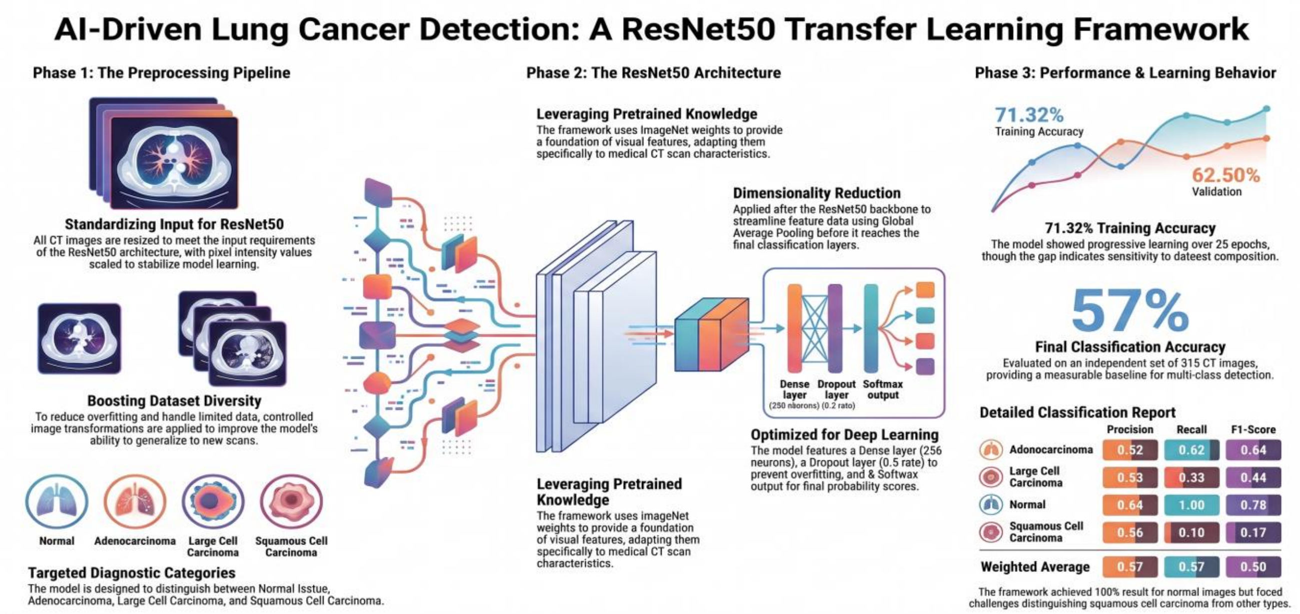
**Informed Consent Statement:** Not applicable.

**Personal Statement:** We declare with our best of knowledge that this research work is purely Original Work and No third party material used in this article drafting. If any such kind material found in further online publication, we are responsible only for any judicial and copyright issues.

**Acknowledgements**

We thank everyone who inspired our work.

: : Overview of the Paper for Researchers : :



ISSN: 3107 - 8605 (Online) , http://www.ijcser.com/ , Vol. 3, Issue 3 , 2026 , https://doi.org/10.63328/IJCSER-V3RI3P2