IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 7, July 2026

DOI: 10.17148/IARJSET.2026.13716

A Multi-Dataset Stacked Ensemble Framework for Multi-Class Lung Cancer Classification

Dhaval J. Rana[1], Keyur Rana[2]

Computer Engineering, Gujarat Technological University, Gujarat, India[1]

Professor, Department of Computer Engineering, Sarvajanik College of Engineering and Technology-SCET, Gujarat, India[2]

Abstract: Lung cancer is a major contributor to cancer-related deaths worldwide, emphasizing the need for accurate and early diagnostic systems. Although deep learning–based approaches have shown promising performance in classifying lung cancer from computed tomography (CT) images, many existing methods are limited by reliance on single datasets and restricted class categorization. To overcome these challenges, this paper introduces SEMLCC-5X, a stacked ensemble framework for multi-class lung cancer classification. The proposed approach integrates multi-source data by combining two publicly available CT datasets, improving data diversity and generalization. The classification task is extended to a fine-grained five-class problem, including normal, benign, adenocarcinoma, large-cell carcinoma, and squamous-cell carcinoma. The framework employs transfer learning with fine-tuning to train four deep learning base models by integrating custom CNN classification layers with the pre-trained Xception, VGG19, EfficientNetB7, and InceptionV3 architectures. These models are combined using a stacking strategy with a logistic regression meta-learner for optimized prediction aggregation. The SEMLCC-5X model is evaluated using comprehensive metrics, including accuracy, precision, recall, F1-score, and AUC. Experimental results demonstrate superior performance, achieving an accuracy of 98.06% and a macro AUC of 99.58%, outperforming individual models and conventional ensemble methods. In conclusion, SEMLCC-5X provides a robust and accurate framework for multi-class lung cancer classification, with strong potential for integration into computer-aided diagnostic systems and clinical workflows.

Keywords: Medical Image Analysis, Computed Tomography (CT), Stacked Ensemble Learning, Deep Learning, Multi-Class Lung Cancer Classification.

I. INTRODUCTION

Lung cancer ranks among the deadliest diseases in the world and is one of the leading causes of cancer deaths [1]. Statistics of the global burden of cancers reveal that lung cancer is responsible for many new cancer cases and deaths per year [2]. Diagnosis and detection at early stages increase chances of survival and help in better treatment planning [3][4]. CT imaging is used widely for lung cancer diagnosis because CT scans offer good visualization of structures of the lungs and pulmonary nodules. Manual analysis of CT scan images is time consuming and requires expert radiologists especially when mass screenings are performed. Hence, automatic computer-aided diagnostic systems for classification of lung cancer have gained much importance recently [5].

Deep learning methods, particularly Convolutional Neural Networks (CNNs), have proven to be very effective in performing tasks like analysis of medical images and classification of lung cancer [6][7][8][9]. Methods that use transfer learning, for example, VGG19 [10], Xception [11], EfficientNet [12], and InceptionV3 [13], have proven to be very effective in extracting features for classification of CT images. Ensemble learning has also been used to improve the classification task using different deep learning methods. However, there are certain limitations in current methods of classification of lung cancer. Many existing studies rely on a single dataset, which limits the generalization capability of the developed models [14]. Furthermore, most approaches focus on binary or three-class classification problems, while fine-grained multi-class classification of lung cancer subtypes remains relatively underexplored [15]. Existing ensemble models [16] often employ simple averaging or voting mechanisms without optimized meta-learning strategies [17].

In previous research, a stacked ensemble method for lung cancer classification ("SEMLCC") [18] based on transfer learning-based deep learning models was suggested using the "IQ-OTH/NCCD" [19] dataset. While "SEMLCC" architecture was promising in terms of classification performance, it is limited to three-class classification using only one dataset. To overcome the above shortcomings, in this paper, a more advanced stacked ensemble architecture, SEMLCC-5X, for multi-class lung cancer classification from CT images is proposed. This architecture involves the use of two publically available CT image datasets. The classification task is also expanded to five classes, which include normal, benign, adenocarcinoma, large-cell carcinoma, and squamous-cell carcinoma. Transfer learning with fine-tuning is

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 7, July 2026

DOI: 10.17148/IARJSET.2026.13716

employed to develop four heterogeneous deep learning models based on the Xception, VGG19, EfficientNetB7, and InceptionV3 architectures. Stacking is used to combine their predictions, along with logistic regression being utilized as the meta-learning algorithm. The primary contributions of this study are summarized as follows:

1. Construction of a multi-source CT image dataset through combination of two existing datasets for increasing data diversity and generalization capabilities.
2. Development of an advanced stacked ensemble scheme, referred to as SEMLCC-5X, for highly accurate classification into five classes of lung cancer.
3. Application of transfer learning with fine-tuning to develop four heterogeneous deep learning models based on the Xception, VGG19, EfficientNetB7, and InceptionV3 architectures.
4. Implementation of logistic regression-based stacking technique for maximizing accuracy of predictions.
5. The SEMLCC-5X framework demonstrates robust and accurate five-class lung cancer classification.

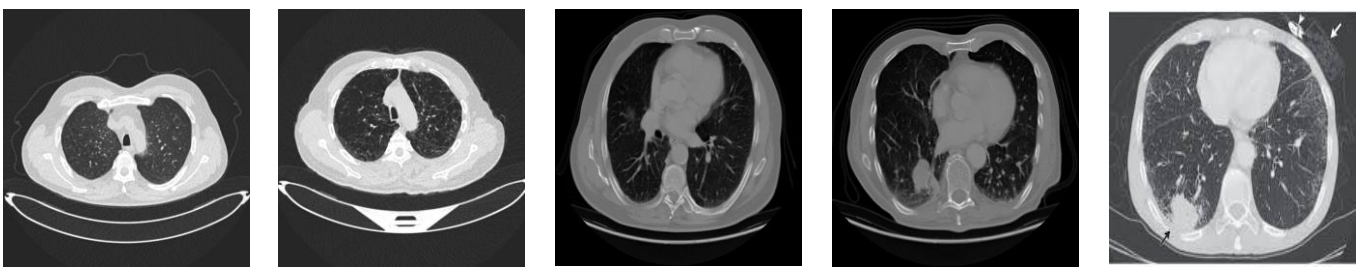
The rest of this paper is structured in the following manner. In Section II, the dataset, data pre-processing, and SEMLCC-5X framework are described. Section III deals with the experimentation and performance analysis. In Section IV, the comparative analysis and discussion are presented. The paper is concluded in Section V.

II. METHODS

A. Dataset Description

Experiments were done on two publicly available thoracic CT images datasets: Kaggle Chest CT-Scan Images Dataset [20] and "IQ-OTH/NCCD" lung cancer dataset, in order to assess the performance of the SEMLCC-5X model. These datasets were integrated to increase diversity in dataset, better class representation and to increase generalization capability of the SEMLCC-5X across different CT image distributions. First dataset named as Dataset A is the publicly available Kaggle Chest CT-Scan Images Dataset. It has 1000 thoracic CT images classified into four clinically relevant classes namely normal, adenocarcinoma, large-cell carcinoma and squamous-cell carcinoma. It has 215 normal images, 338 adenocarcinoma images, 187 large-cell carcinoma images, and 260 squamous-cell carcinoma images. Second dataset named as Dataset B is the "IQ-OTH/NCCD" lung cancer dataset, which was developed through collaborative efforts among oncology institutes in Iraq. This benchmark dataset has 1097 thoracic CT images organized into three different classes namely normal, malignant and benign. It has 416 normal images, 561 malignant images, and 120 benign images. The CT scans have been clinically validated by expert radiologists and oncologists and cover different anatomical sections of chest area.

For the study, the CT scans from the two datasets were merged and organized into five target classes such as normal, benign, adenocarcinoma, large-cell carcinoma, and squamous-cell carcinoma. This inclusion of more than one dataset brings about variability in image characteristics, thereby making the proposed SEMLCC-5X method more robust and generalizable. The merged dataset was then split into training, validation, and test sets in the ratio of 70:15:15. This means that 70% of the CT images would be used for training, 15% for validation and the other 15% for testing. These images have been made anonymous and are only used for research purposes. Sample CT scans of the merged dataset are presented in Fig. 1.



I) Benign **II) Normal** **III) Adenocarcinoma** **IV) Large-cell carcinoma** **V) Squamous-cell carcinoma**

Fig. 1. Representative CT images from the combined lung cancer dataset [19] [20]

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 7, July 2026

DOI: 10.17148/IARJSET.2026.13716

B. Data Pre-processing

For training of the proposed SEMLCC-5X model, a set of pre-processing steps was conducted on the CT image data set in combination. Since the images in the CT scan set were obtained from different public sources and have various properties, pre-processing became essential in order to prepare uniform input data for all deep learning models used in the current work. All images were resized to 224 × 224 pixels. Such image resolution was chosen to ensure a balance between efficiency and retention of visual properties that are important in case of the lungs' segmentation. Additionally, such size of the images is convenient for use in popular transfer learning architectures, including Xception, VGG19, EfficientNetB7, and InceptionV3 with ImageNet pretrained parameters. The original images in the CT scan were grayscale images. Nevertheless, the convolutional architectures that were used in this research required input images in three-channel form. For this reason, each original image was transformed to three-channel image by repeating the same intensity values across all channels. Normalization of intensity values to the range [0,1] was also conducted in order to enhance training stability. The normalization procedure ensures that the neural networks are able to converge more effectively and reduces the effects caused by inconsistent intensity distributions of the images in the combined data sets [21]. Moreover, various methods of data augmentation have been employed during the training phase to increase the robustness and generalization ability of the proposed approach. Data augmentation methods involved random rotation, random zoom, horizontal flip, and shift of images [22]. These methods help in artificially increasing the variety of the training data and prevent overfitting, while retaining clinically meaningful image features. Following the pre-processing step, the dataset is split into three parts: training, validation, and testing in the ratio of 70:15:15. The training part is used for training of the model, the validation part is used to monitor the performance of the model, and the test part is kept out of sight of the training process.

C. Proposed model SEMLCC-5X Architecture

1) Proposed Framework Overview

The SEMLCC-5X framework is a stacked ensemble model developed for multi-class lung cancer classification using thoracic CT Scan images. This model incorporates both the concepts of transfer learning [23] and ensemble learning in a single framework. The SEMLCC-5X model has been divided into three key parts as shown in Fig. 2.

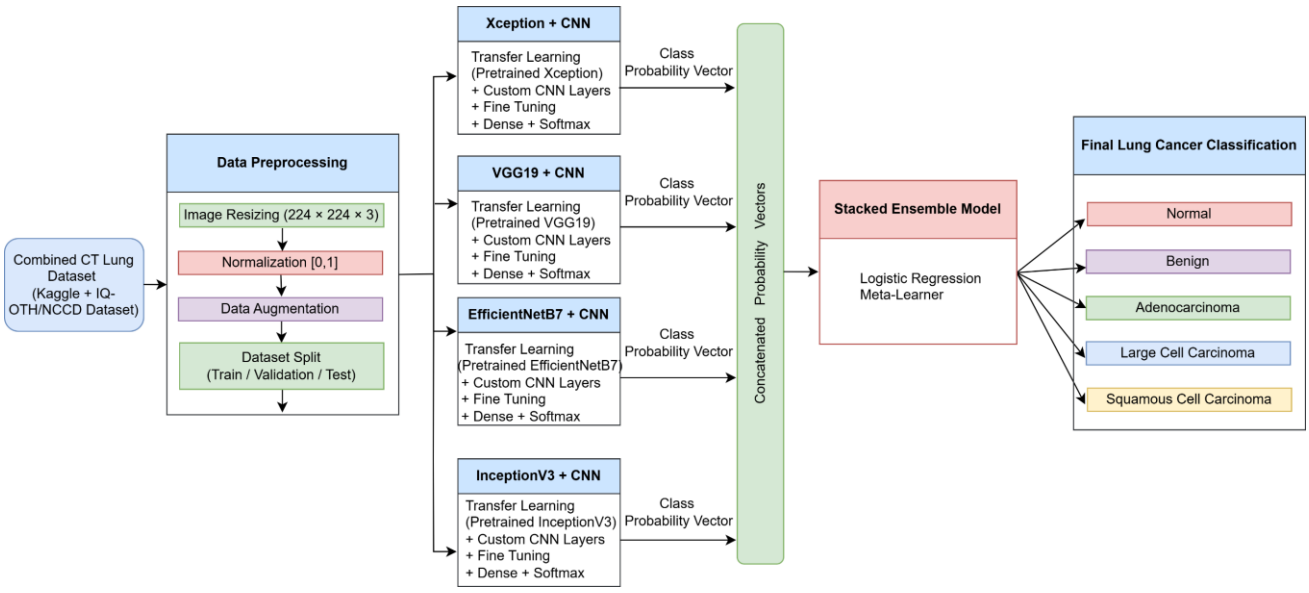


Fig. 2. SEMLCC-5X stacked ensemble framework for multi-class lung cancer classification architecture.

Algorithm 1. Proposed SEMLCC-5X Framework

(1) **Input:** CT lung images from Kaggle and IQ-OTH/NCCD datasets
(2) **Output:** Predicted lung cancer class label ∈ {Normal, Benign, Adenocarcinoma, Large-Cell Carcinoma, Squamous-Cell Carcinoma}
(3) Acquire CT lung images from both datasets
(4) Merge datasets and organize images into five classes.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 7, July 2026

DOI: 10.17148/IARJSET.2026.13716

(5) Resize all CT images to 224 × 224 × 3.
(6) Normalize pixel intensities to the range [0,1]
(7) Apply data augmentation for improved dataset balancing and generalization
(8) Split dataset into training, validation, and testing subsets
(9) Initialize four transfer learning base models:
    a) Xception
    b) VGG19
    c) EfficientNetB7
    d) InceptionV3
(10) **For** each base model do
(11) Load pretrained ImageNet weights
(12) Add custom CNN classification layers
(13) Apply transfer learning with fine-tuning
(14) Fine-tune the model using the Adam optimizer and a categorical cross-entropy
(15) Generate class probability predictions on validation and test datasets
(16) **End for**
(17) Concatenate probability vectors obtained from all trained base models
(18) Train Logistic Regression meta-learner using stacked prediction vectors
(19) Generate final SEMLCC-5X stacked ensemble prediction
(20) Evaluate Stacked Ensemble models using:
Accuracy, Precision, Recall, F1-score, AUC
(22) Generate confusion matrices and ROC-AUC curves for class-wise analysis
(23) Output final predicted lung cancer class label

2) Base Model Architectures

The SEMLCC-5X framework employs four heterogeneous transfer learning models based on the Xception, VGG19, EfficientNetB7, and InceptionV3 architectures as base learners. Each model is initialized using pretrained ImageNet weights and fine-tuned for five-class lung cancer classification. Additional CNN classification layers with Softmax activation are incorporated to generate probability distributions corresponding to normal, benign, adenocarcinoma, large-cell carcinoma, and squamous-cell carcinoma classes.

3) Probability-Level Stacking Mechanism

The trained base models generate class probability vectors that are utilized for ensemble learning. To improve classification robustness, a probability-level stacking mechanism is introduced using Logistic Regression as the meta-learner. The concatenated probability outputs from all base models are provided as input to the meta-learner, enabling adaptive ensemble weighting and improved decision boundary learning compared to conventional averaging-based ensembles. The detailed stacking process is illustrated in Fig. 3.

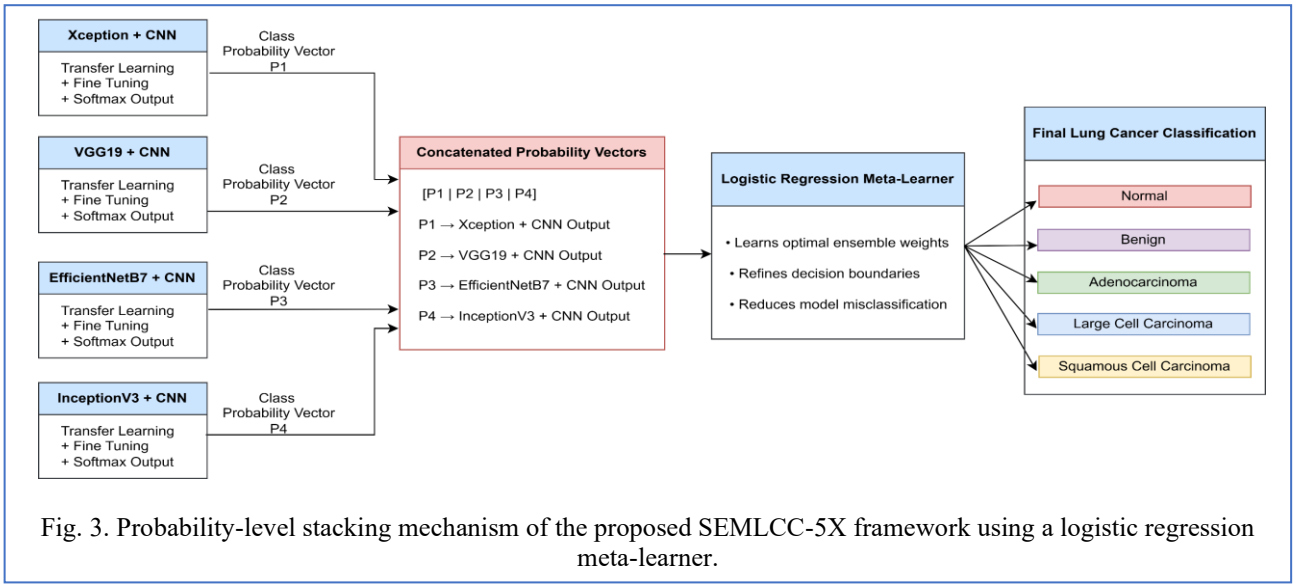


Fig. 3. Probability-level stacking mechanism of the proposed SEMLCC-5X framework using a logistic regression meta-learner.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 7, July 2026

DOI: 10.17148/IARJSET.2026.13716

D. Training Configuration

The proposed SEMLCC-5X framework was developed and trained using the TensorFlow and Keras deep learning frameworks. Training was performed on the combined thoracic CT image dataset obtained from the Kaggle Chest CT-Scan Images Dataset and the "IQ-OTH/NCCD" lung cancer dataset. Prior to training, all CT images were resized to 224 × 224 × 3 to maintain a consistent input format for all transfer learning architectures. Pixel intensities were normalized through rescaling operations to improve numerical stability during optimization.

To enhance the diversity of the training samples and improve the generalization capability of the proposed framework, several augmentation operations were incorporated into the pre-processing pipeline. Some of the transformations done on images involved rotation, zooming, horizontal flipping, and spatial shift. The augmented training helped in improving robust feature extraction from the images under various imaging conditions, without causing overfitting of the network. The SEMLCC-5X framework employed four heterogeneous transfer learning architectures based on Xception, VGG19, EfficientNetB7, and InceptionV3 as base learners. Each model utilized pretrained ImageNet weights [24] for initialization, followed by fine-tuning on lung CT images to adapt the learned representations to the target classification task. The classification layers in each of the models provided five-class probability outputs through the use of Softmax activation function. Training was performed using the Adam optimizer together with the categorical cross-entropy loss function. For the purposes of training, a learning rate of 0.0001 and a batch size of 16 were employed. Each of the models was trained for up to 100 epochs using training and validation sets that were different. Probability outputs from the trained base models were further used for logistic regression based stacked ensemble learning. Best model configuration was ensured by monitoring the validation performance and storing the best performing weights through checkpointing. Evaluation of the final SEMLCC-5X model was performed on an independent testing set using various performance metrics, including accuracy, precision, recall, F1 score, ROC-AUC.

Table 1. Configuration of Training Parameters for the Proposed SEMLCC-5X Framework

Training Parameter	Configuration
Image size	224 × 224 × 3
Mini-batch size	16
Optimizer	Adam
Initial learning rate	0.0001
Output activation function	Softmax
Number of training epochs	100
Normalization	Rescaling using 1./255
Meta-learner	Logistic Regression
Loss function	Categorical Cross-Entropy

III. RESULT AND PERFORMANCE EVALUATION

A. Evaluation Metrics

Evaluating the performance of AI models for lung cancer diagnosis requires appropriate use of statistical metrics, particularly due to the clinical consequences of misclassification. This section outlines the key metrics used to assess classification models, especially in the context of imbalanced medical datasets.

Accuracy (ACC): Shows the capability of the machine learning model to correctly classify instances into their classes. High accuracy indicates better classification performance [25].

Sensitivity (Recall): Refers to the capability of the model to correctly classify the samples belonging to each lung cancer class. It is very important to have high sensitivity in order to reduce the number of false negatives [25].

Specificity (True Negative Rate): Shows the capability of the model to correctly classify the samples which do not belong to a particular class. Having high specificity reduces additional medical examination and related patient anxiety [26].

Precision (Positive Predictive Value): The rate of instances predicted to belong to a particular class that are correctly classified. It matters especially when there are additional consequences from false positives [27].

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 7, July 2026

DOI: 10.17148/IARJSET.2026.13716

F1 Score: It is the harmonic mean between the precision and recall measures since both of these have to be considered while working with unbalanced datasets [28].

Area Under the ROC Curve (AUC): t is a measure of discrimination power of a model across the different lung cancer classes at different decision thresholds. Higher values of AUC indicate better performance in terms of classification [26]. In order to check the efficiency of the proposed SEMLCC-5X method, the performance of each of the transfer learning base models is individually analyzed before performing the ensemble learning process. The architectures being used include Xception, VGG19, EfficientNetB7, and InceptionV3. All of the models are trained using the same pre-processing steps and parameterization techniques.

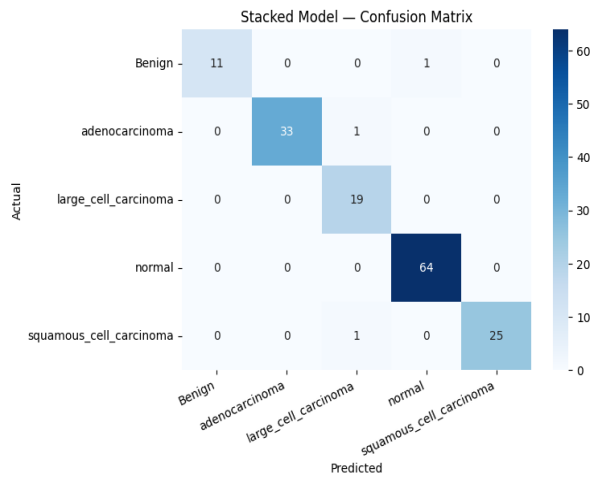


Fig. 4. Confusion Matrix

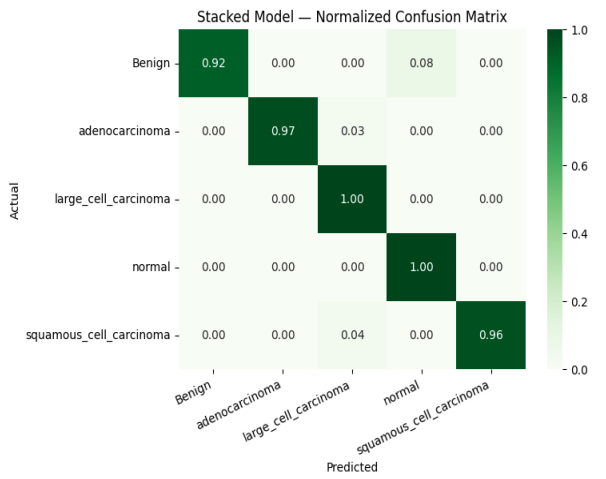


Fig. 5. Normalized Confusion Matrix

As depicted in Fig. 4, the SEMLCC-5X classifier was able to correctly classify most of the CT images, where predictions were clustered along the main diagonal line. Very few cases of misclassification occurred, but mostly for the benign category owing to its similarity with some malignant cases, indicating the success of the proposed stacked ensemble. Figure 5 shows the confusion matrix, whereby high values along the diagonal line depict high recall within each individual class and low values outside the diagonal line signify low classification errors. These findings support the robustness and balanced nature of the proposed framework for all the five classes. As illustrated in Fig. 6, the ROC curves are very close to the upper-left corner, and the SEMLCC-5X classifier had a macro AUC value of 99.58%. This shows high discriminative ability for all lung cancer classes. The classification report in Fig. 7 depicts high precision, recall, and F1-score values for all the classes. High weighted-average measures also show that the proposed framework can handle class imbalances well.

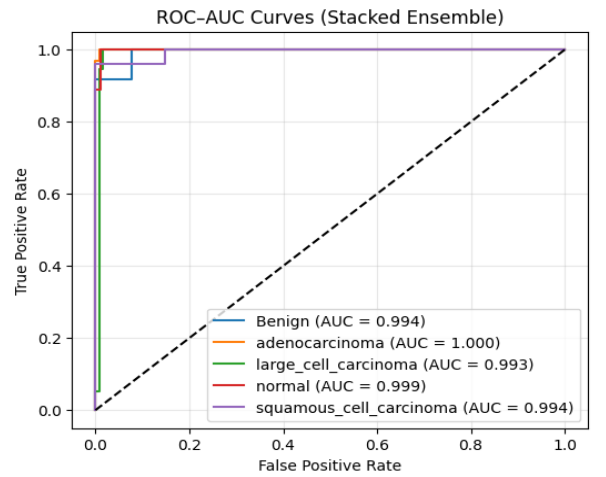


Fig. 6. ROC curves and AUC analysis

	Precision	Recall	F1-score	Support
Benign	1.0000	0.9167	0.9565	12
Adenocarcinoma	1.0000	0.9706	0.9851	34
Large Cell Carcinoma	0.9048	1.0000	0.9500	19
Normal	0.9846	1.0000	0.9922	64
Squamous Cell Carcinoma	1.0000	0.9615	0.9804	26
accuracy			0.9806	155
macro avg	0.9779	0.9698	0.9728	155
weighted avg	0.9820	0.9806	0.9807	155

Fig. 7. Classification Report

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 7, July 2026

DOI: 10.17148/IARJSET.2026.13716

IV. COMPARATIVE ANALYSIS WITH INDIVIDUAL CNN MODELS

Comparative analysis of the proposed SEMLCC-5X approach relative to other state-of-the-art methods is carried out in Table 2 below. The direct comparisons between different methods should be done carefully because most of the methods were suggested for solving problems of two- or three-class lung cancer classification based on one publicly available CT data set. On the contrary, the suggested SEMLCC-5X approach deals with a more complicated problem of five classes classification by using two CT data sets. Despite these differences, the proposed framework achieved competitive performance with an accuracy of 98.06%, precision of 97.79%, recall of 96.98%, F1-score of 97.28%, and a macro AUC of 99.58%. These results demonstrate the effectiveness of the proposed probability-level stacking strategy in improving feature representation and generalization for fine-grained lung cancer classification.

TABLE 2. Performance Comparison of the Proposed SEMLCC-5X Ensemble Model with Existing State-of-the-Art Methods

Methods	Model Performance Metrics(%)					
	Classes	Accuracy	Recall	Precision	F1-Score	AUC
SEMLCC-5X (Proposed)	**5**	98.06	96.98	**97.79**	**97.28**	99.58
LRSE-LCC [29]	3	**98.19**	**100.00**	97.00	96.42	99.90
SEMLCC [18]	3	96.89	96.89	97.04	96.89	**100.00**
UDCT [30]	3	96.82	97.50	98.70	98.24	-
MobileNetV2-SGRU [31]	3	96.83	96.83	96.78	96.78	95.82
iSOA Algorithm [32]	3	96.58	95.38	84.16	91.53	-

Furthermore, Table 3 presents a comparative analysis of the proposed SEMLCC-5X framework and its constituent transfer learning models. The experimental results demonstrate that the stacked ensemble consistently achieves superior performance compared with the Xception, VGG19, EfficientNetB7, and InceptionV3-based CNN models across all evaluation metrics. This improvement is attributed to the complementary feature representations learned by the heterogeneous base models and their effective integration through the Logistic Regression meta-learner. Consequently, the proposed SEMLCC-5X framework provides a robust and reliable solution for automated multi-class lung cancer classification.

TABLE 3. Performance Comparison Between Individual CNN Models and the Proposed SEMLCC-5X Ensemble Framework

Methods	Evaluation Parameters (%)				
	Accuracy	Recall	Precision	F1-Score	AUC
SEMLCC-5X (Proposed)	**98.06**	**96.98**	**97.79**	**97.28**	**99.58**
Xception-based CNN	96.77	95.00	97.00	95.00	99.00
VGG19-based CNN	92.90	87.00	94.00	89.00	99.00
EfficientNetB7-based CNN	96.77	96.00	95.00	96.00	99.00
InceptionV3 + CNN	96.77	94.00	97.00	95.00	99.00

V. CONCLUSION

This paper presented SEMLCC-5X, a multi-dataset stacked ensemble framework for automated five-class lung cancer classification using thoracic CT images. The proposed framework integrates two publicly available CT image datasets to improve dataset diversity and enhance the generalization capability of the classification model. Four heterogeneous transfer learning architectures based on Xception, VGG19, EfficientNetB7, and InceptionV3 were employed as base

DOI: 10.17148/IARJSET.2026.13716

learners, while a Logistic Regression meta-learner integrated their class probability outputs using a probability-level stacking strategy. Experimental evaluation demonstrated that the proposed SEMLCC-5X framework achieved superior classification performance compared with the individual transfer learning models. The proposed model attained an overall accuracy of 98.06%, precision of 97.79%, recall of 96.98%, F1-score of 97.28%, and a macro AUC of 99.58%, indicating excellent discrimination capability and reliable multi-class classification performance. The confusion matrix, ROC analysis, and classification report further confirmed the robustness of the proposed framework across all five lung cancer categories.

The results demonstrate that probability-level stacking effectively exploits the complementary strengths of multiple deep learning architectures, leading to improved feature representation, reduced classification errors, and enhanced generalization. The findings of this study demonstrate that the SEMLCC-5X framework is capable of providing robust and interpretable multi-class lung cancer classification from CT images, indicating its potential for integration into clinical decision support systems. Although the obtained results are encouraging, further evaluation using independent datasets collected from different healthcare institutions is essential to assess the model's robustness and generalizability. Future research will concentrate on improving computational efficiency for practical deployment, enhancing the stacking strategy and explainability mechanisms, and evaluating the framework in real-world clinical workflows to increase its reliability and adoption in medical practice.

REFERENCES

[1] World Health Organization (WHO), “WHO Lung Cancer, Geneva: World Health Organization, 2025.” [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/lung-cancer

[2] American Cancer Society, “American Cancer Society, Key Statistics for Lung Cancer, Atlanta: American Cancer Society, 2025.,” 2025, [Online]. Available: https://www.cancer.org/cancer/types/lung-cancer.html.

[3] H. Kadara et al., “Early diagnosis and screening for lung cancer,” Cold Spring Harb. Perspect. Med., vol. 11, no. 9, pp. 1–20, 2021, doi: 10.1101/cshperspect.a037994.

[4] International Agency for Research on Cancer (IARC), “International Agency for Research on Cancer (IARC), Cancer Incidence and Mortality, Both sexes, in 2022,” 2022, [Online]. Available: https://gco.iarc.who.int/today/en/dataviz/bars

[5] X. Chen et al., “Recent advances and clinical applications of deep learning in medical image analysis,” Med. Image Anal., vol. 79, p. 102444, 2022, doi: 10.1016/j.media.2022.102444.

[6] C. Szegedy, S. Reed, P. Sermanet, V. Vanhoucke, and A. Rabinovich, “Going deeper with convolutions,” pp. 1–12.

[7] G. C. Forte et al., “Deep Learning Algorithms for Diagnosis of Lung Cancer: A Systematic Review and Meta-Analysis,” Cancers (Basel)., vol. 14, no. 16, pp. 1–11, 2022, doi: 10.3390/cancers14163856.

[8] M. Li, Y. Jiang, Y. Zhang, and H. Zhu, “Medical image analysis using deep learning algorithms,” Front. Public Heal., vol. 11, no. November, pp. 1–28, 2023, doi: 10.3389/fpubh.2023.1273253.

[9] A. A. Abe, M. Nyathi, A. A. Okunade, W. Pilloy, B. Kgole, and N. Nyakale, “A robust deep learning algorithm for lung cancer detection from computed tomography images,” Intell. Med., vol. 11, no. January, p. 100203, 2025, doi: 10.1016/j.ibmed.2025.100203.

[10] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” Sep. 2014, Accessed: Dec. 19, 2023. [Online]. Available: http://arxiv.org/abs/1409.1556

[11] F. Chollet, “Xception: Deep Learning with Depthwise Separable Convolutions,” Oct. 2016, Accessed: Dec. 18, 2023. [Online]. Available: http://arxiv.org/abs/1610.02357

[12] M. Tan and Q. V. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” May 2019, Accessed: Dec. 21, 2023. [Online]. Available: http://arxiv.org/abs/1905.11946

[13] C. Szegedy, V. Vanhoucke, S. Ioffe, and J. Shlens, “Rethinking the Inception Architecture for Computer Vision.”

[14] I. Castiglioni et al., “AI applications to medical images: From machine learning to deep learning,” Phys. Medica, vol. 83, no. February, pp. 9–24, 2021, doi: 10.1016/j.ejmp.2021.02.006.

[15] S. Tang, R. Ma, Q. Li, Y. Bai, and S. Chen, “Classification of Benign and Malignant Pulmonary Nodules Based on the Multiresolution 3D DPSECN Model and Semisupervised Clustering,” IEEE Access, vol. 9, pp. 43397–43410, 2021, doi: 10.1109/ACCESS.2021.3060178.

[16] T. G. Dietterich, “Ensemble Methods in Machine Learning,” pp. 1–2, 2000.

[17] W. Li et al., “Machine Learning Model of ResNet50-Ensemble Voting for Malignant–Benign Small Pulmonary Nodule Classification on Computed Tomography Images,” Cancers (Basel)., vol. 15, no. 22, 2023, doi: 10.3390/cancers15225417.

[18] D. J. Rana and K. Rana, “ScienceDirect ScienceDirect SEMLCC : Stacked Ensemble Model with Transfer Learning SEMLCC : Stacked Ensemble Model with Transfer Learning for Lung Cancer Classification from CT

DOI: 10.17148/IARJSET.2026.13716

Images for High-Accuracy Lung Cancer Classification from CT Images," Procedia Comput. Sci., vol. 258, pp. 2584–2596, 2025, doi: 10.1016/j.procs.2025.04.520.

[19] Hamdalla F. Al-Yasriy, "The IQ-OTH/NCCD lung cancer dataset -- Kaggle." Accessed: Jan. 27, 2024. [Online]. Available: https://www.kaggle.com/datasets/hamdallak/the-iqothnccd-lung-cancer-dataset

[20] "Chest CT-Scan images Dataset." [Online]. Available: https://www.kaggle.com/datasets/mohamedhanyyy/chest-ctscan-images/data

[21] D. Singh and B. Singh, "Investigating the impact of data normalization on classification performance," Appl. Soft Comput., vol. 97, p. 105524, 2020, doi: https://doi.org/10.1016/j.asoc.2019.105524.

[22] A. Mikołajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," in 2018 International Interdisciplinary PhD Workshop (IIPhDW), 2018, pp. 117–122. doi: 10.1109/IIPHDW.2018.8388338.

[23] I. W. Harsono, S. Liawatimena, and T. W. Cenggoro, "Lung nodule detection and classification from Thorax CT-scan using RetinaNet with transfer learning," J. King Saud Univ. - Comput. Inf. Sci., vol. 34, no. 3, pp. 567–577, 2022, doi: 10.1016/j.jksuci.2020.03.013.

[24] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in Advances in Neural Information Processing Systems, F. Pereira, C. J. Burges, L. Bottou, and K. Q. Weinberger, Eds., Curran Associates, Inc., 2012. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

[25] T. Saito and M. Rehmsmeier, "The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets," pp. 1–21, 2015, doi: 10.1371/journal.pone.0118432.

[26] L. To and T. H. E. Editor, "Receiver-operating characteristic curve analysis in diagnostic , prognostic and predictive between sensitivity and specificity with change of test cut-offs Sensitivity and specificity in diagnostic accuracy : ROC solid or lost in translation ? Receiver-operating characteristic curve analysis in diagnostic , prognostic and predictive biomarker research : trade-off between sensitivity and specificity with change of test cut-offs References Email alerting service," vol. 62, no. 11, pp. 11–12, 2009, doi: 10.1136/jcp.2009.069591.

[27] J. Davis and M. Goadrich, "The Relationship Between Precision-Recall and ROC Curves," pp. 233–240, 2006.

[28] D. Zhang, J. Wang, and X. Zhao, "Estimating the Uncertainty of Average F1 Scores," pp. 317–320.

[29] D. J. Rana, "LRSE-LCC : A Lightweight Residual CNN with Squeeze-and-Excitation Attention for Lung Cancer Classification from CT Image," vol. 8, no. 2, pp. 657–671, 2026.

[30] A. Gupta, A. Kumar, and K. Rautela, "UDCT: lung Cancer detection and classification using U-net and DARTS for medical CT images," Multimed. Tools Appl., vol. 84, no. 18, pp. 19065–19085, 2025, doi: 10.1007/s11042-024-19801-9.

[31] A. Bagheri Tofighi, A. Ahmadi, and H. Mosadegh, "Improving lung cancer detection via MobileNetV2 and stacked-GRU with explainable AI," Int. J. Inf. Technol., vol. 17, no. 2, pp. 1189–1196, 2025, doi: 10.1007/s41870-024-02045-z.

[32] C. Yan and N. Razmjooy, "Optimal lung cancer detection based on CNN optimized and improved Snake optimization algorithm," Biomed. Signal Process. Control, vol. 86, Sep. 2023, doi: 10.1016/j.bspc.2023.105319.

