© June 2026 | IJIRT | Volume 13 Issue 1 | ISSN: 2349-6002

# Federated Learning Based Lung Cancer Detection Using VGG16 And ResNet50 Models

Dr. S Jeyakumar[1], Ms. S Jenolin Aiswarya[2]

[1]*Professor & Principal, Department of Computer Science and Engineering, Jayaraj Annapackiam CSI College of Engineering, Nazareth, TamilNadu, India*

[2]*PG Scholar, Department of Computer Science and Engineering, Jayaraj Annapackiam CSI College of Engineering, Nazareth, TamilNadu, India*

***Abstract*—Lung cancer remains one of the leading causes of cancer-related mortality worldwide. Computed Tomography (CT) imaging is widely used for diagnosis; however, manual analysis is time-consuming and highly dependent on radiologist expertise. Traditional centralized training of medical data additionally raises significant privacy concerns. This paper proposes a hybrid federated learning framework for lung cancer classification using a VGG16 deep learning model. The system enables decentralized training across multiple clients while preserving patient data privacy through the Federated Averaging (FedAvg) algorithm. The final global model classifies CT scan images into three categories—Benign, Malignant, and Normal—achieving an overall accuracy of 92%, outperforming centralized VGG16 (89%), centralized ResNet50 (84%), and federated ResNet50 (91.34%) baselines. Grad-CAM visualization is incorporated to improve model interpretability and clinical trust.**

***Index Terms—Lung Cancer, CT Imaging, Federated Learning, VGG16, Deep Learning, Medical Image Classification, Privacy Preservation, FedAvg Algorithm, Clinical Decision Support.***

## I. INTRODUCTION

Lung cancer is one of the leading causes of cancer-related deaths worldwide, and early detection is essential for improving patient survival rates. Computed Tomography (CT) scan imaging plays a significant role in identifying lung abnormalities; however, manual analysis is time-consuming and depends heavily on radiologist expertise.

Deep Learning techniques have shown promising results in medical image classification by enabling automatic feature extraction and accurate prediction. In particular, VGG16 is widely used for analyzing CT scan images due to its strong feature extraction capabilities. However, conventional deep learning approaches rely on centralized data training, which raises serious concerns regarding patient data privacy and security. Medical data sharing between institutions is restricted due to confidentiality issues.

To overcome these challenges, this paper proposes a privacy-preserving lung cancer detection system using Federated Learning. In this approach, model training is distributed across multiple clients without sharing raw data. A global server trains a base VGG16 model and distributes the trained weights to clients. Each client performs local training on its own dataset and sends updated weights back to the server. The server aggregates client weights using the Federated Averaging (FedAvg) algorithm to update the global model. The final global model classifies CT scan images into Benign, Malignant, and Normal categories, ensuring data privacy while improving model performance through collaborative learning.

### 1.1. Objectives

The objectives of this work are: (i) to develop a privacy-preserving lung cancer detection system using VGG16 and Federated Learning; (ii) to implement decentralized training ensuring sensitive patient data is not shared between entities; (iii) to improve model accuracy through collaborative learning across multiple clients; and (iv) to serve as an efficient clinical decision-support tool for early lung cancer detection.

## II. RELATED WORK

Federated learning has emerged as a prominent paradigm for privacy-preserving medical image analysis. McMahan et al. [1] introduced the foundational FedAvg algorithm, enabling communication-efficient decentralized training, which forms the backbone of the proposed framework.

Several studies have explored federated learning specifically for lung cancer detection. A decentralized learning framework proposed for distributed hospital networks demonstrated improved generalization in early-stage cancer detection without sharing raw patient data. Another work integrated blockchain with federated learning to enhance security and data integrity during collaborative training.

Ensemble-based federated models have also been proposed to combine predictions from multiple local models, improving classification accuracy in complex multi-stage cancer diagnosis scenarios. Studies on collaborative federated frameworks for multi-disease classification further demonstrated the feasibility of handling heterogeneous datasets across institutions [2–6].

Despite these advances, challenges remain in communication overhead, system scalability, and real-world deployment. The present work addresses these gaps by combining VGG16 transfer learning with a structured federated communication strategy and incorporating Grad-CAM visualization for interpretability.

## III. METHODOLOGY

### 3.1. Dataset

The lung CT scan dataset (IQ-OTH/NCCD) used in this study contains 2,496 images categorized into three classes: Benign, Malignant, and Normal. The dataset was split into training, validation, and testing subsets as shown in Table 1.

*Table 1: Dataset split across training, validation, and test subsets.*

| Class | Train | Validation | Test | Total |
|---|---|---|---|---|
| Benign | 665 | 83 | 84 | 832 |
| Malignant | 665 | 83 | 84 | 832 |
| Normal | 665 | 83 | 84 | 832 |
| Total | 1,995 | 249 | 252 | 2,496 |

### 3.2. Image Preprocessing

All CT scan images were resized to 224×224×3 pixels to match the VGG16 input requirements. Pixel values were normalized to the [0,1] range to improve convergence. Data augmentation techniques including random rotation, zoom, and horizontal flipping were applied to the training set to improve generalization and reduce overfitting.

### 3.3. VGG16 Model Architecture

The VGG16 convolutional neural network, pretrained on ImageNet, was adopted as the feature extractor. The original classification head was replaced with a custom head comprising a GlobalAveragePooling2D layer, a Dropout layer (rate=0.5), a Dense layer (256 units, ReLU activation), and a final Softmax output layer for three-class classification. The convolutional base was initially frozen to retain pretrained feature extraction capability; the final layers were subsequently fine-tuned on the CT scan dataset.

### 3.4. Federated Learning Strategy

The federated framework involves a central server and three distributed clients, simulating a multi-hospital environment.

The training proceeds in the following stages:

1) Global Server Training:

The server initializes and trains the VGG16 model for 12 epochs using its local dataset to generate the initial global weights ($W_0$).

2) Client 2 Local Training:

The server distributes $W_0$ to Client 2, which trains locally for 15 epochs followed by 7 fine-tuning epochs. Updated weights $W_2$ are returned to the server.

3) Client 1 and Client 3 Local Training:

The server updates the global model with $W_2$ and distributes the refined weights to Client 1 and Client 3 for parallel local training (5 epochs each). Updated weights $W_1$ and $W_3$ are returned.

4) Weight Aggregation (FedAvg):

The final global model is obtained by aggregating client weights using the Federated Averaging algorithm:

$$W\_global = (n_1 \cdot W_1 + n_3 \cdot W_3) / (n_1 + n_3)$$

where $W_1$, $W_3$ are updated weights from Client 1 and Client 3, and $n_1$, $n_3$ are their respective sample counts. This aggregation ensures proportional contribution based on dataset size.

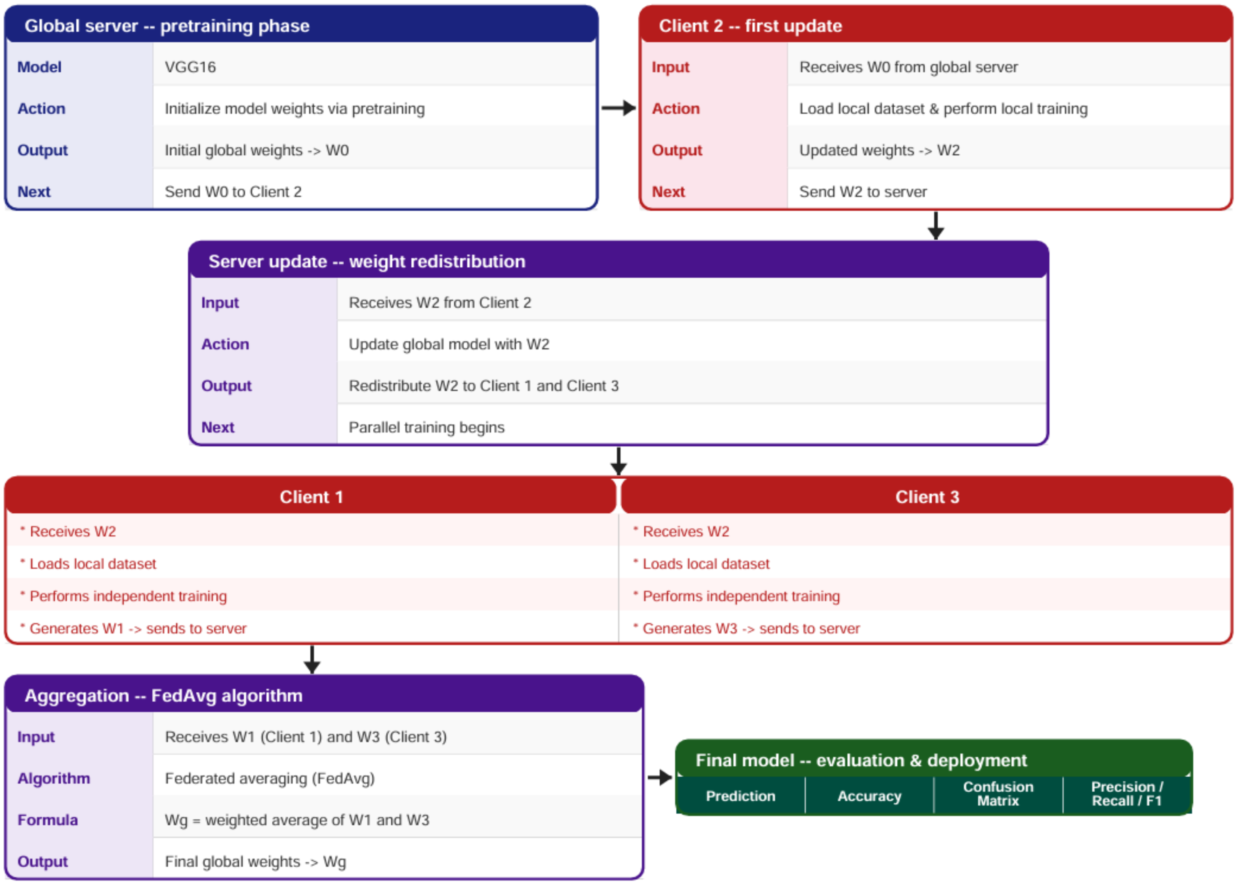


*Figure 1: System Architecture of the Federated Learning-Based Lung Cancer Detection Framework*

3.5. Training Parameters

The model was trained using the Adam optimizer (learning rate = 0.0001) with categorical crossentropy loss. Batch size was set to 32. Activation functions used were ReLU (hidden layers) and Softmax (output layer).

## IV. EXPERIMENTAL RESULTS AND DISCUSSION

### 4.1. Experimental Setup

All experiments were implemented in Python using TensorFlow/Keras in a GPU-enabled Google Colab environment. Supporting libraries included NumPy, Matplotlib, OpenCV, and Scikit-learn. The test dataset was held out entirely from training and validation to ensure unbiased evaluation.

### 4.2. Performance of Centralized VGG16 (Baseline)

The centralized VGG16 model achieved an overall classification accuracy of 89%. It performed well for Malignant (F1 = 0.98) and Normal (F1 = 0.88) classes, while the Benign class showed lower recall (0.18) and F1-score (0.30), likely due to class imbalance and visual similarity with other classes. Table 2 presents the full classification report.

*Table 2: Classification report of the centralized VGG16 model.*

| Class | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|
| Benign | 0.86 | 0.18 | 0.30 | 33 |
| Malignant | 0.99 | 0.97 | 0.98 | 153 |
| Normal | 0.80 | 0.98 | 0.88 | 123 |
| Overall Accuracy | | | 0.89 | 309 |
| Macro Avg | 0.88 | 0.71 | 0.72 | 309 |
| Weighted Avg | 0.90 | 0.89 | 0.87 | 309 |

### 4.3. Performance of Centralized ResNet50 (Baseline)

The centralized ResNet50 model achieved an overall accuracy of 84%. The Malignant class attained strong performance (F1 = 0.92); however, the Benign class showed low precision (0.46) with high recall (0.97), indicating a tendency to over-predict benign cases. The Normal class achieved a precision of 0.90 with moderate recall (0.77). See Table 3.

*Table 3: Classification report of the centralized ResNet50 model.*

| Class | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|
| Benign | 0.46 | 0.97 | 0.62 | 29 |
| Malignant | 1.00 | 0.86 | 0.92 | 148 |
| Normal | 0.90 | 0.77 | 0.83 | 78 |
| Overall Accuracy | | | 0.84 | 255 |
| Macro Avg | 0.78 | 0.86 | 0.79 | 255 |
| Weighted Avg | 0.91 | 0.84 | 0.86 | 255 |

### 4.4. Performance of Proposed Federated VGG16

The proposed federated VGG16 model achieved an overall accuracy of 92%, the highest among all evaluated models. The Malignant class attained perfect classification (Precision = Recall = F1 = 1.00), demonstrating the model's strong ability to detect cancerous cases. The Normal class also showed strong performance (F1 = 0.90). The Benign class achieved lower recall (0.44), consistent with its limited sample size and feature similarity to other classes. Table 4 summarizes the results.

*Table 4: Classification report of the proposed federated VGG16 model.*

| Class | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|
| Benign | 0.67 | 0.44 | 0.53 | 9 |
| Malignant | 1.00 | 1.00 | 1.00 | 42 |
| Normal | 0.86 | 0.94 | 0.90 | 32 |
| Overall Accuracy | | | 0.92 | 83 |
| Macro Avg | 0.84 | 0.79 | 0.81 | 83 |
| Weighted Avg | 0.91 | 0.92 | 0.91 | 83 |

### 4.5. Performance of Federated ResNet50 (Comparative)

The federated ResNet50 model achieved an overall accuracy of 91.34%. Although it showed perfect precision for the Malignant class (1.00), overall classification was weaker than the proposed federated VGG16 model. The deeper ResNet50 architecture required more training data and exhibited reduced convergence in the federated setting, suggesting VGG16 is better suited for this dataset and configuration. Table 5 presents detailed results.

*Table 5: Classification report of the federated ResNet50 model.*

| Class | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|
| Benign | 0.56 | 1.00 | 0.72 | 14 |
| Malignant | 1.00 | 0.96 | 0.98 | 74 |
| Normal | 1.00 | 0.79 | 0.89 | 39 |
| Overall Accuracy | | | 0.91 | 127 |
| Macro Avg | 0.85 | 0.92 | 0.86 | 127 |
| Weighted Avg | 0.95 | 0.91 | 0.92 | 127 |

### 4.6. Comparative Analysis

Table 6 provides a comprehensive comparison of all evaluated models across key dimensions including privacy, accuracy, and interpretability. The proposed federated VGG16 model achieves the best accuracy while offering high privacy preservation and Grad-CAM-based interpretability, making it the most suitable framework for real-world clinical deployment.

*Table 6: Comparative analysis of all models.*

| Parameter | Centralized VGG16 | Centralized ResNet50 | Federated ResNet50 | Proposed Federated VGG16 |
|---|---|---|---|---|
| Training Approach | Centralized | Centralized | Federated | Federated |
| Data Sharing | All data centrally stored | All data centrally stored | Model weights only | Model weights only |
| Privacy | Low | Low | High | High |

| | | | | |
|---|---|---|---|---|
| FedAvg Aggregation | No | No | Yes | Yes |
| Interpretability | Standard | Standard | Standard | Grad-CAM |
| Accuracy (%) | 89.00 | 84.00 | 91.34 | 92.00 |
| Performance | Good | Moderate | Satisfactory | Best |

## V. CONCLUSION

This paper presented a privacy-preserving lung cancer detection system using Federated Learning with a VGG16 architecture for classification of lung CT scan images into Benign, Malignant, and Normal categories. The proposed federated VGG16 model achieved an overall accuracy of 92%, outperforming centralized VGG16 (89%), centralized ResNet50 (84%), and federated ResNet50 (91.34%) baselines.

The experimental results demonstrate that federated learning can achieve classification performance comparable to or superior to centralized approaches while ensuring that raw patient data never leaves local client environments. The integration of Grad-CAM visualization further enhances clinical interpretability and trust. The proposed system provides a practical, scalable, and privacy-compliant solution for deployment in multi-hospital and distributed healthcare environments.

Future work will explore differential privacy mechanisms, advanced aggregation strategies, and extension to larger and more heterogeneous clinical datasets to further improve robustness and generalization.

## REFERENCES


[1] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. Aguera y Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," in *Proc. 20th Int. Conf. Artificial Intelligence and Statistics (AISTATS)*, 2017. [Online]. Available: arXiv:1602.05629

[2] F. R. da Silva, R. Camacho, and J. M. R. S. Tavares, "Federated Learning in Medical Image Analysis: A Systematic Survey," Electronics, vol. 13, no. 1, Art. no. 47, 2024.

[3] M. Adnan, S. Kalra, J. C. Cresswell, G. W. Taylor, and H. R. Tizhoosh, "Federated Learning and Differential Privacy for Medical Image Analysis," Scientific Reports, vol. 12, Art. no. 1953, 2022.

[4] H. E. Kim, A. Cosa-Linan, N. Santhanam, M. Jannesari, M. E. Maros, and T. Ganslandt, "Transfer Learning for Medical Image Classification: A Literature Review," BMC Medical Imaging, vol. 22, Art. no. 69, 2022.

[5] R. Javed, T. Abbas, A. H. Khan, A. Daud, and A. Bukhari, "Deep Learning for Lung Cancer Detection: A Review," Artificial Intelligence Review, vol. 57, Art. no. 197, 2024.

[6] H. Guan, P.-T. Yap, A. Bozoki, and M. Liu, "Federated Learning for Medical Image Analysis: A Survey," arXiv preprint arXiv:2306.05980, 2023.

[7] M. Haj Fares and A. M. S. E. Saad, "Towards Privacy-Preserving Medical Imaging: Federated Learning with Differential Privacy and Secure Aggregation Using a Modified ResNet Architecture," arXiv preprint arXiv:2412.00687, 2024.

[8] M. Joshi, A. Pal, and M. Sankarasubbu, "Federated Learning for Healthcare Domain: Pipeline, Applications and Challenges," ACM Transactions on Computing for Healthcare, 2022.

[9] R. Hassan and M. Ali, "Review of Lung Cancer Detection and Classification Using Deep Learning," AIIT Journal of Artificial Intelligence and Technology, vol. 3, no. 2, pp. 85–94, 2023.

[10] M. Varchagall, N. P. Nethravathi, R. Chandramma, *et al.*, "Using Deep Learning Techniques to Evaluate Lung Cancer Using CT Images," *SN Computer Science*, vol. 4, Art. no. 173, 2023.