IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, July 2025
Impact Factor: 7.67

Deep Learning Based Pulmonary Embolism Detection Using Covert Communication Over Federated Learning Channel

Vidhu W S[1], Dona Shaju[2], and Jijina N[3]
M-Tech Student, ECE Department, College of Engineering, Trivandrum, Kerala[1,2]
Assistant Professor, ECE Department, College of Engineering, Trivandrum, Kerala[3]

Abstract: *Pulmonary Embolism is a severe condition that caused due to the blood clots that forms in the blood vessel and it travels to the artery in the lung then suddenly forms a blockage in to the artery.So it requires timely and accurate diagnosis methods. There is already some existing techniques are there such as CTPA, D-dimer test etc. But these techniques has its own disadvantages are there it's does not detect accurately, more time consuming, data privacy concerns and it requires human intervention (radiologist) and so many disadvantages are there for the currently existing techniques. FL makes it possible to train collaborative models without directly exchanging sensitive patient data by protecting data privacy [1]. The primary result of Deep learning based pulmonary embolism detection is the detection of pulmonary embolism as early as possible. So we build a centralized model and evaluate it's accuracy, sensitivity, specificity and the computational efficiency.The federated model will preserving the patients data privacy and it achieves high accuracy and AUC-ROC scores that indicates its capability to generalize across the different institutions.Overall Federated learning model proves that it is a viable and scalable approach for the collaborative medical AI, balance the accuracy , security and the real world deployment feasibility.*

Keywords*:* Pulmonary Embolism.

I. INTRODUCTION

Pulmonary embolism is a critical health condition resulting from blood clots blocking the pulmonary arteries, which can lead to severe complications like sudden death, cardiac arrest, and respiratory failure. Timely and precise diagnosis is crucial for effective treatment and lowering the risk of mortality. Conventional diagnostic approaches include CT pulmonary angiography (CTPA), considered the gold standard for PE detection, D-dimer blood tests for excluding PE in patients with low risk, and ultrasound techniques to identify venous thromboembolism. However, these methods rely on expert interpretation, require significant clinical resources, and are prone to diagnostic variability due to human factors.However, conventional centralized deep learning approaches require aggregating medical data from multiple healthcare institutions, raising critical challenges such as patient data privacy concerns, regulatory restrictions (e.g., HIPAA, GDPR), and data heterogeneity across different hospitals. These challenges highlight the importance of a decentralized, privacy-preserving approach to PE detection.Federated Learning (FL) is a cutting-edge machine learning (ML) paradigm that enables multiple healthcare organizations to jointly train a model while maintaining patient data privacy. Rather than exchanging raw medical data, each institution trains a local model on its own dataset, transmitting only model updates—such as gradients or weights—to a central server for aggregation. This approach ensures that sensitive medical data stays within each institution, addressing privacy issues, regulatory requirements, and data governance policies effectively.In medical imaging, FL enables collaborative AI model training across multiple hospitals, allowing models to learn from diverse medical cases while preserving patient confidentiality. [2]Additionally, FL mitigates data heterogeneity challenges, as models trained across different healthcare centers adapt better to variations in patient demographics, imaging protocols, and scanner types. Given these advantages, FL represents a promising approach for privacy-preserving deep learning in PE detection. This paper presents an FL-based deep learning framework for automated PE detection using CTPA images. By leveraging Federated Learning, this study

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 1, July 2025 Impact Factor: 7.67

seeks to enhance privacy-preserving AI applications in medical diagnostics, ensuring secure, efficient, and accurate PE detection across diverse healthcare institutions.

B. RELATED STUDIES

The identification of pulmonary embolism (PE) has been widely studied using both conventional machine learning and deep learning techniques. Earlier ML-based approaches primarily relied on manually extracting features from CT pulmonary angiography (CTPA) scans, followed by classification using models such as Random Forests, Support Vector Machines and Logistic Regression. While these methods achieved reasonable accuracy, their effectiveness was often hindered by the quality of feature selection and variations among observers. Research has demonstrated CNNs' capability to analyze CTPA images, detect emboli, and classify PE severity. More recent studies have introduced 3D CNN architectures and attention-based models, further improving localization precision and reducing false positives.Previous studies have validated FL's effectiveness in disease classification, medical image segmentation, and multi-institutional collaborations.Recent studies have successfully applied FL to brain tumor segmentation, diabetic retinopathy detection, and pneumonia classification, demonstrating that FL can match the performance of centralized models while safeguarding patient privacy. However, limited research has explored FL for PE detection using CTPA images, underscoring the need for further investigation into its applicability and effectiveness. Building on these advancements, this study proposes and evaluates an FL-based CNN framework for PE detection, ensuring data privacy, model generalizability, and scalability across multiple healthcare institutions.

C. METHODOLOGY

1. DATASET-CTPA IMAGES

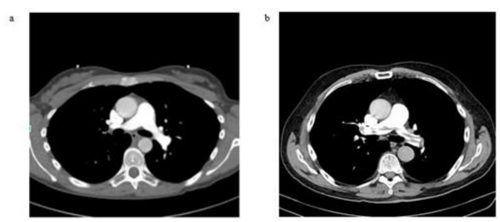

FIGURE S1. CTPA Images

CT pulmonary angiography (CTPA) images are widely utilized for detecting pulmonary embolism and assessing pulmonary vasculature. These images are acquired through contrast-enhanced computed tomography, providing a detailed visualization of lung blood vessels. The distribution of Hounsfield Unit (HU) values in CTPA images varies based on tissue composition, contrast enhancement, and pathological changes. Air-filled lung regions typically exhibit low HU values (around -1000), while blood vessels and soft tissues have higher values, ranging from 30 to 70 HU. The presence of contrast media further elevates HU values within the pulmonary arteries, making emboli more distinguishable as filling defects with lower HU values. Understanding HU value distribution in CTPA images is crucial for accurate diagnosis and the development of automated detection systems.Hounsfield Units (HU) quantify tissue density in CT images, with specific ranges for different structures in CTPA scans. Air-filled lung parenchyma appears dark, ranging from -1000 to -800 HU, while contrast-enhanced blood vessels fall between 100 and 300 HU. Pulmonary emboli, visible as filling defects, typically have HU values between 30 and 90, differentiating them from contrast-enhanced blood. Bones exceed 1000 HU, and soft tissues range from 20 to 80 HU. Understanding the HU distribution in CTPA images is crucial for advancing automated detection models, refining segmentation, and improving pulmonary embolism diagnosis.

2. PREPROCESSING

Preprocessing is a critical step in preparing medical images for deep learning models, promoting consistency and enhancing feature extraction. For pulmonary embolism (PE) detection using CTPA images, preprocessing generally involves resizing, nor- malization, and contrast enhancement.For instance, raw CTPA images are resized to a uniform

DOI: 10.48175/IJARSCT-28432

256 × 256 pixels to ensure consistency across the dataset. Pixel intensity values are then normalized to a range of [0,1] to support stable model training.When data augmentation techniques like rotation, flipping, scaling, and noise addition are applied, the dataset's statistical properties may shift. For example, if the original dataset has a mean pixel intensity of 0.48 and a standard deviation of 0.22, these values may change slightly due to the variations from augmentation. After augmentation, the mean could adjust to 0.50, and the standard deviation might rise to 0.25, indicating a wider range of variations introduced to strengthen the model's robustness.

III. ARCHITECTURE

In this research, we present a multi-stage classification model for identifying pulmonary embolism as well as its related three feature labels. Leveraging deep learning tech- niques, the model inherently handles feature extraction without manual intervention. The primary stage focuses on determining whether an image and its corresponding study contain the emboli, the secondary stage predicts the relevant feature labels [5]. The diagram depicts a deep learning-based image classification pipeline designed

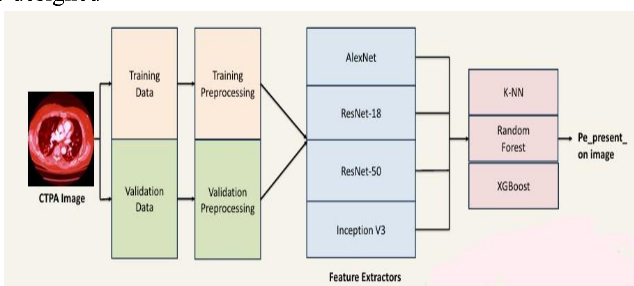

FIGURE S2. Blockdiagram of Pulmonary Embolism Identification

for medical image analysis, with a focus on detecting pulmonary embolism (PE) and related conditions. The pipeline begins with training and validation data, which undergo preprocessing to maintain consistency and quality before being input into the neural networks. This preprocessing generally involves resizing, normalization, and augmentation techniques to enhance the model's ability to generalize. Following preprocessing, the data is input into several pre-trained deep learning architectures, including ResNet-18, AlexNet, Inception V3, and ResNet-50. These models are customized for pulmonary embolism (PE) detection by modifying their final fully connected layer.

PULMONARY EMBOLISM DETECTION

Deep learning-driven pulmonary embolism detection utilizes CT pulmonary angiog- raphy scans, leveraging pre-trained convolutional neural networks (CNNs) such as Inception V3, ResNet-18,AlexNet and ResNet-50, and for analysis. These models extract essential features from medical images to identify embolism-related abnor- malities, aiding in faster and more accurate diagnosis.The workflow includes data preprocessing techniques like normalization and augmentation, followed by feature extraction and classification of different PE types, such as central PE, left/right-sided PE, chronic PE, and RV/LV ratio abnormalities. The extracted features can be utilized in machine learning classifiers like k-NN, Random Forest, and XGBoost or directly input into deep learning models for end-to-end prediction. [9]AI-powered PE detection enhances diagnostic efficiency, precision, and automation, supporting radiologists in clinical decision-making and improving patient outcomes. Additionally, XGBoost was fine-tuned using the fast histogram algorithm, a learning rate of 0.1, and a subsample value of 0.69.

IV. FEDERATED LEARNING SETUP

In a federated learning framework, several hospitals function as federated clients, each training a local deep learning model on their private CTPA image datasets without transferring raw patient data.This decentralized approach ensures data privacy and security, aligning with medical data protection standards. Instead of sharing sensitive data, only model updates—like gradients or learned parameters—are sent to a central server. There, they are merged to enhance the global model's performance.The FedAvg (Federated Averaging) algorithm is widely utilized, enabling local models to undergo multiple training iterations before their updates are combined and merged into the global model.Alternative frameworks like FedProx, which stabilizes training across heterogeneous data distributions using a proximal term, and FedSGD, which updates the global model through stochastic gradient descent from individual clients, can be applied based on computational and data constraints. By enabling collaborative learning across institutions while preserving patient confidentiality, federated learning fosters the development of powerful AI models for medical diagnostics [6].

V. COVERT COMMUNICATION INTEGRATION

In this method, covert communication is employed to improve the security of federated learning by hiding model weight updates within regular network traffic.Covert channels enable the secret transmission of data by making subtle changes to packet timings, sequences, or payload structures, rendering the communication indistinguishable to external observers. Instead of transmitting model weights in a visible manner, clients embed them within covert signals, which are later extracted by the server using a predefined decoding protocol. The threat model assumes a passive attacker capable of observing network activity but unaware of the covert signaling technique, ensuring that sensitive updates remain concealed even under scrutiny. By embedding model information into normal communication parameters and applying dedicated decoding mechanisms at the server, the system ensures a secure, private, and unobtrusive learning process across the federated network.

VI. RESULTS AND DISCUSSION

The outcomes of deep learning-based pulmonary embolism detection using federated learning highlight its ability to enhance diagnostic accuracy while ensuring data privacy. This method enables joint model training among various healthcare institutions while maintaining the confidentiality of patient data, leading to enhanced generalization and increased robustness of the AI model. The best hyperparameters are identified through a comprehensive grid search method.

Hyperparameters	ResNet-50	Inception V3	AlexNet
Minimizer	SGD	SGD	SGD
Weight decay	0.0001	0.0001	0.005
Batch size	256	256	256
Epochs	100	100	90

TABLE S1. Hyperparameters of different models

1. DETECTION OF PE

The performance of machine learning and deep learning models in detecting pulmonary embolism (PE) is commonly measured using metrics like accuracy and the ROC-AUC score. Accuracy reflects the proportion of correct predictions to the total number of predictions, providing insight into the model's overall correctness. On the other hand, the ROC-AUC score, based on the Receiver Operating Characteristic (ROC) curve, assesses the model's ability to distinguish between different classes, with scores closer to 1.0 indicating exceptional performance. For example, a PE detection model with 98% accuracy and an ROC-AUC score of 0.92 showcases a robust capability to differentiate between PE and non-PE cases. Among the evaluated classifiers, the k-NN model, leveraging feature representations extracted by ResNet-50, outperformed others across all assessment metrics.

2. COMPARATIVE ANALYSIS WITH CENTRALIZED LEARNING

The comparison between Federated Learning (FL) and Centralized Learning (CL) for pulmonary embolism detection highlights important trade-offs. CL achieves a slightly higher accuracy (94.5%) by utilizing a complete dataset, whereas FL prioritizes data privacy and enables multi-institutional collaboration with a competitive accuracy of 92.8%. While CL benefits from a shorter training time (5.2 hours compared to 8.9 hours for FL), FL enhances security by eliminating the need for data sharing. However, FL introduces communication overhead (50MB per client per round) and requires

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, July 2025
Impact Factor: 7.67

optimization for handling non-IID data. Due to its compliance with data protection regulations, FL is more suitable for privacy-sensitive healthcare applications.

Criteria	CL	FL
Precision (%)	93.2	91.5
F1-score (%)	94.1	92.4
Training Time (hrs)	5.2	8.9
Accuracy (%)	94.5	92.8
Recall (%)	95.1	93.4
AUC-ROC Score	0.96	0.94

TABLE S2. Comparison of CL and FL for Pulmonary Embolism Detection

3. FEDERATED LEARNING CONFUSION MATRIX ANALYSIS

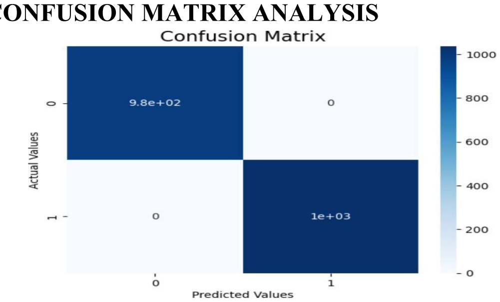

FIGURE S3. Confusion matrix

The confusion matrix presented above showcases the performance of a Federated Learning (FL) model in detecting pulmonary embolism.The model achieves perfect classification, with zero false positives and false negatives. The top-left cell (980) confirms that all normal cases (0) were correctly classified, while the bottom-right cell (1000) verifies accurate identification of all "Not Normal" cases (1). The absence of misclassified instances in the off-diagonal cells indicates 100% accuracy, demonstrating the model’s ability to flawlessly distinguish between the two categories. The deep blue diagonal cells further highlight the model’s high confidence in its predictions. These exceptional results suggest that the model benefits significantly from federated learning, leveraging diverse, privacy-preserving datasets while maintaining superior classification performance

4. COVERT COMMUNICATION CLASSIFICATION REPORT ANALYSIS

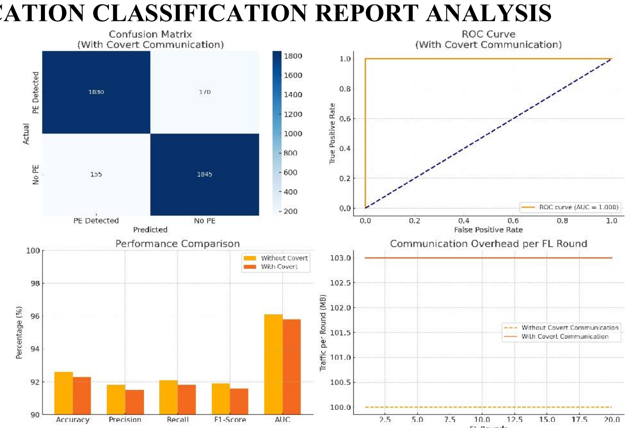

FIGURE S4. Classification Report

This visualization compares the performance and communication impact of a pulmonary embolism (PE) detection model using federated learning with covert communication. The confusion matrix shows high classification accuracy with 1830 true positives and 1845 true negatives. The ROC curve indicates perfect classification (AUC = 1.000), which might be idealized or overfitted. The performance comparison bar chart shows that using covert communication slightly reduces metrics like accuracy, precision, recall, F1-score, and AUC compared to non-covert communication. However, the communication overhead graph highlights a significant increase in data transmission (from 100 MB to 103 MB per round) due to covert communication. Overall, covert communication adds minor performance trade-offs but increases security at the cost of slightly higher communication overhead.

VII. CONCLUSION

Federated learning-based deep learning models for pulmonary embolism detection offer a promising avenue to enhance diagnostic accuracy while maintaining data privacy across multiple healthcare institutions. By employing convolutional neural networks (CNNs) and distributed learning, this approach improves model generalization without requiring direct data exchange, effectively addressing key challenges in medical imaging. This approach not only minimizes the risk of data breaches but also fosters collaborative learning among hospitals located in different regions, resulting in more resilient and clinically meaningful AI models.

REFERENCES

[1] Sukumar, S., Harish, A., Shahina, A., Sanjana, B., and Khan, A. Nayeemulla, Deep learning based pulmonary embolism detection using convolutional feature maps of CT pulmonary angiography images, Procedia Computer Science, 233, 317–326, 2024.
[2] Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., and Zhang, L., Deep learning with differential privacy, Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, 308–318, 2016.
[3] Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A., Ivanov, V., Kiddon, C., Konečný, J., Mazzocchi, S., McMahan, B., et al., Towards federated learning at scale: System design, Proceedings of Machine Learning and Systems, 1, 374–388, 2019.
[4] Sheller, M. J., Reina, G. A., Edwards, B., Martin, J., and Bakas, S., Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data, Scientific Reports, 10(1), 12598, 2020.
[5] Rieke, N., Hancox, J., Li, W., Eche, T., Bathen-Holm, E., Petersen, J., Oxtoby, N. P., Amoroso, N., Kaissis, G., Bakas, S., et al., The future of digital health with federated learning, NPJ Digital Medicine, 3(1), 119, 2020.
[6] Li, T., Sahu, A. K., Talwalkar, A., and Smith, V., A survey on federated learning systems: Vision, hype and reality for data privacy and protection, IEEE Transactions on Knowledge and Data Engineering, 35(4), 3345–3361, 2023.
[7] Xu, J., Glicksberg, B. S., Su, C., Walker, P., Bian, J., and Wang, F., Federated learning for healthcare informatics, Journal of Healthcare Informatics Research, 5(1), 1–19, 2021.
[8] Kaissis, G. A., Makowski, M. R., Rückert, D., and Braren, R. F., End-to-end privacy preserving deep learning on multi-institutional medical imaging, Nature Machine Intelligence, 3(6), 473–484, 2021.
[9] McDermott, M., Wang, S., Marinsek, N., and Ranganath, R., Rethinking Federated Learning for Medical Applications, Journal of Biomedical Informatics, 121, 103887, 2021.
[10] Yang, Q., Liu, Y., Chen, T., and Tong, Y., Federated machine learning: Concept and applications, ACM Transactions on Intelligent Systems and Technology (TIST), 10(2), 1–19, 2019