Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Federated Learning for Privacy-Preserving Medical Image Analysis: A Chest X-Ray Case Study

Dr. Dipali Koshti [1*], Dr. Supriya Kamoji [2], Dr. Harsh Bhor [3], Dr. Archana Said [4], Unik Lokhande [5]
[1] *Department of Electronics and Computer Science, Fr. Conceicao Rodrigues College of Engineering, Mumbai, Maharashtra, India*
[2,5] *Department of Computer Engineering, Fr. Conceicao Rodrigues College of Engineering, Mumbai, Maharashtra, India*
[3] *Department of Information Technology, K.J. Somaiya Institute of Technology, Mumbai, Maharashtra, India*
[4] *AISSMS Institute of Information Technology, Pune, Maharashtra, India*
* *Corresponding Author Email: dipalis@fragnel.edu.in*

ARTICLE INFO

Received: 18 Dec 2024

Revised: 14 Feb 2025

Accepted: 28 Feb 2025

ABSTRACT

A key component of contemporary healthcare, medical image analysis has transformed precision disease diagnosis, treatment planning, and disease monitoring. Advances in medical image processing, especially about the incorporation of deep learning techniques, have greatly improved the precision and efficacy of diagnostic processes. Federated learning protects patient anonymity by enabling cooperative model training across decentralized data sources, keeping sensitive medical data localized and private while still promoting model progress. This paper provides a complete medical image analysis framework using Federated Learning. In a normalized and pre-processed Chest X-ray dataset, several base deep learning models were trained for comparison. These models included a basic CNN model, VGG16, ResNet50, and InceptionV3. After training, hyperparameters were optimized to improve performance. Our experimental results show that the Inception V3 performs better than other two DL models. The best-performing deep learning model was selected as the client's local model. To address privacy concerns, Federated Learning (FL) techniques were employed. FL allows devices to update models locally without sharing raw data. The FedAvg algorithm was used at the server to aggregate the data received from the clients, and the process was performed cyclically. Model weights from individual devices were transmitted to a central server for aggregation. This collaborative approach enables learning while preserving data anonymity.

Keywords:. Federated Learning, Medical Image analysis, Deep Learning.

INTRODUCTION

Medical image analysis is essential to modern healthcare because it enables precise disease diagnosis, treatment planning, and disease monitoring. The accuracy and effectiveness of tasks involving medical image processing have increased because of recent developments in deep learning techniques. Applying several deep learning models in the medical industry now confronts significant obstacles. These challenges include the requirement for reliable and understandable models, handling data privacy issues, making sure that the models apply to a variety of patient demographics and healthcare environments, and smoothly incorporating the models into clinical processes. The interpretability of model predictions and the absence of defined evaluation criteria are further issues that need to be addressed if one is to win over healthcare professionals' confidence and acceptance. Furthermore, larger and more varied datasets are required to effectively train deep learning models, particularly in light of the variations in clinical outcomes and medical imaging data among various patient populations. To fully utilize deep learning to transform healthcare, treatment planning, and medical diagnostics, these issues must be resolved.

1.1 Federated Learning:

Federated learning makes it possible to collaboratively train models using decentralized data sources while protecting data privacy (Soham et al., 2023). The effectiveness of federated and deep learning methods can be seen for a variety of medical imaging tasks, such as image segmentation (Renard, F, 2020), classification, disease identification (Hu, M., 2021) and anomaly detection (Rauniyar, A, 2023). In the fields of drug discovery, clinical decision support systems, healthcare IOT, public health surveillance, and medical research, federated learning makes collaboration easier. It benefits multiple medical disciplines by enabling data analysis while maintaining privacy.

Existing federated learning (FL) algorithms face difficulties in several medical fields. Maintaining model performance while aggregating heterogeneous data is challenging due to source-specific data heterogeneity. Network bandwidth is strained by high communication overhead, and privacy issues with local model updates and

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

communication security still exist. It is difficult to ensure model convergence when data distributions diverge. Effective FL deployment is further hampered by security concerns like model poisoning and resource shortages in healthcare facilities. The implementation of FL in medical contexts is further complicated by regulatory compliance with privacy restrictions.

1.2 Different types of Federated Learning:

Federated learning encompasses various strategies tailored to address distinct challenges in distributed machine learning. As shown in Fig. 1, Each approach offers unique advantages and considerations in optimizing collaboration and efficiency across distributed environments.

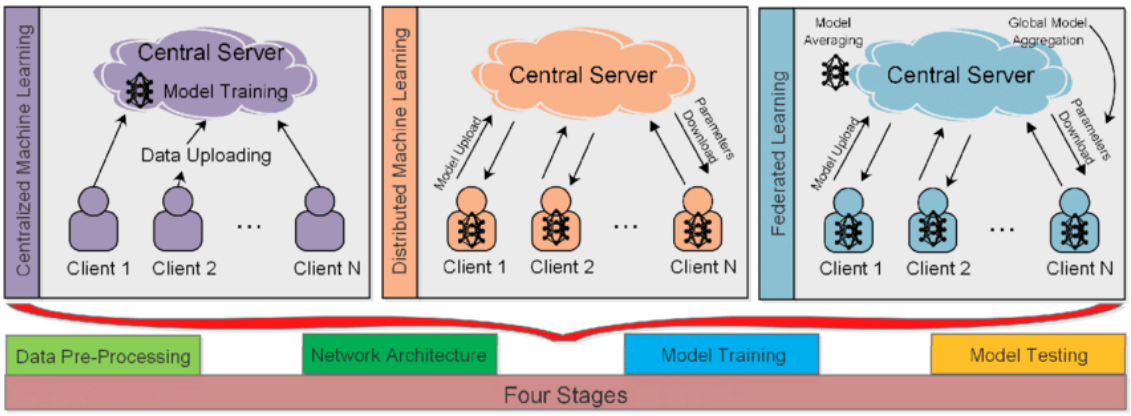


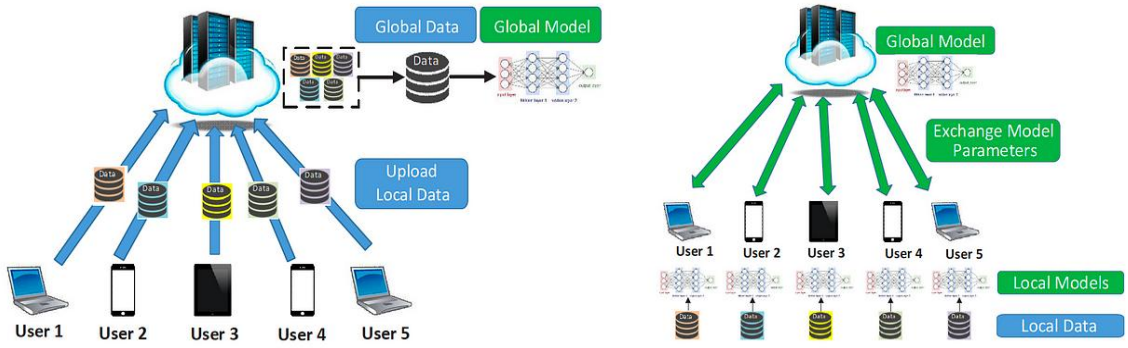
Figure. 1: Federated Learning Framework

Centralized Federated Learning:

Centralized federated learning relies on a central server to orchestrate model training by selecting client devices and collecting updates. This method seems straightforward and yields accurate models. However, it encounters a critical bottleneck issue: network failures can disrupt the entire process. Despite its efficacy, the centralized approach is susceptible to interruptions due to its dependency on a single point of coordination—the central server.

Decentralized Federated Learning:

In contrast to centralized federated learning, decentralized federated learning operates without a central server. Instead, model updates are shared directly among interconnected edge devices. While this eliminates the risk of single-point failures, the accuracy of the final model hinges on the network topology of these edge devices. Consequently, the effectiveness of decentralized federated learning varies depending on the interconnectedness and reliability of the network. Fig. 2 depicts the centralized and distributed learning techniques.



(a) centralized learning (can be outsourced learning). (b) distributed learning

Figure. 2: Centralized Learning vs. Distributed Learning

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Heterogeneous Federated Learning:

Heterogeneous federated learning addresses the diversity among client devices, which can range from mobile phones to IoT devices with varying hardware, software, and data characteristics. Unlike conventional federated learning methods that assume uniformity among local models, HeteroFL acknowledges and accommodates the heterogeneity of these devices. By training over multiple distinct local models, HeteroFL can converge on a single global model suitable for inference across diverse device types and data modalities. Fig. 3 depicts the heterogeneous learning technique.

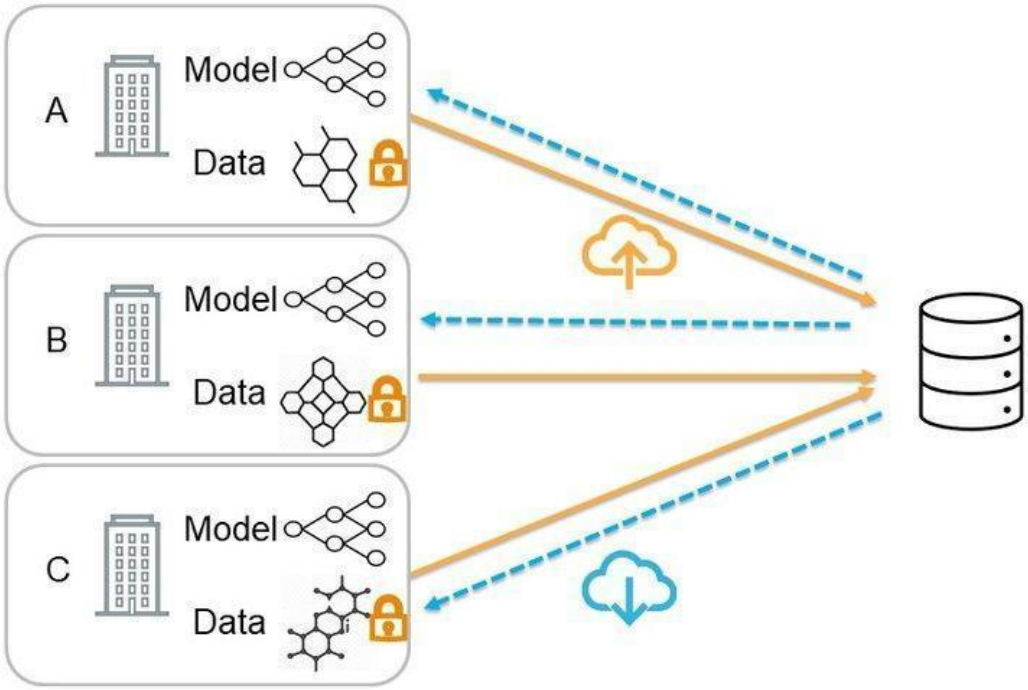


Figure. 3: Heterogeneous Learning

1.3 Federated Learning Algorithms:

Federated learning algorithms comprise a range of methodologies tailored for collaborative model training across distributed devices. These techniques, such as FEDSG, leverage different strategies to facilitate communication and aggregation of model updates among participating devices. While some algorithms prioritize centralized coordination for model convergence, others adopt decentralized approaches to eliminate single points of failure. Additionally, algorithms like FEDAVG aim to accommodate device heterogeneity by adapting to varied hardware, software, and data characteristics. Each algorithm brings its unique strengths and considerations, contributing to the optimization of collaborative learning in distributed environments.

Federated Averaging (FedAvg):

FedAvg (Fig. 4) builds upon the foundation laid by FedSGD, presenting an enhanced algorithm for federated learning. Clients in FedAvg are empowered to perform multiple local gradient descent updates, unlike FedSGD where gradients are directly shared with the central server. Instead of transmitting gradients, clients share their locally tuned model weights with the server. The server then aggregates these weights, effectively averaging the model parameters across all participating clients. This approach ensures model synchronization while accommodating variations in local data distributions. By allowing for local weight tuning before aggregation, FedAvg optimizes model convergence and robustness in heterogeneous federated learning environments.

Federated Stochastic Gradient Descent (FedSGD):

FedSGD as shown in Fig. 5, revolutionizes the conventional stochastic gradient descent (SGD) approach by adapting it to the federated learning setting. Unlike traditional SGD, which computes gradients on mini-batches of data samples, FedSGD treats these mini-batches as different client devices, each housing local data. In FedSGD, the central model is disseminated to these clients, where each client computes gradients using its local data. These gradients are then transmitted back to the central server, which aggregates them in proportion to the number of samples on each

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

client to calculate the gradient descent step. This distributed computation allows for collaborative model training while preserving data privacy and decentralization, hallmark features of federated learning.

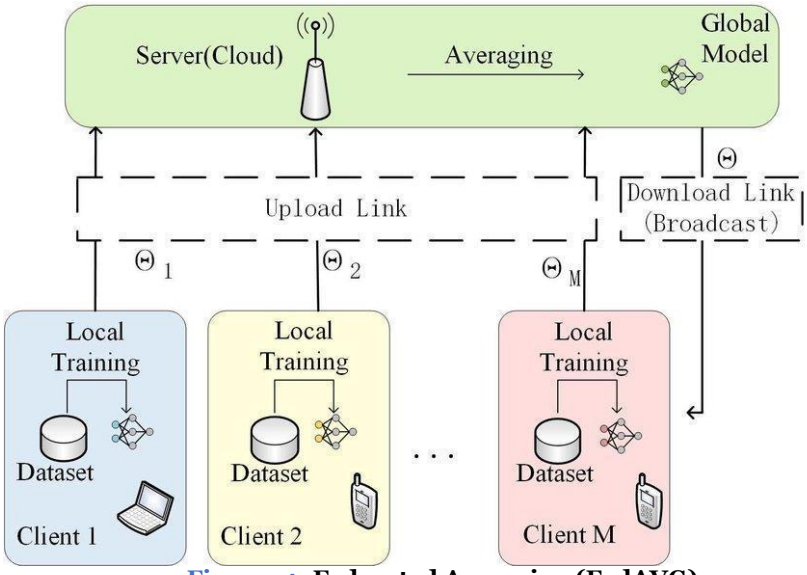


Figure. 4: Federated Averaging (FedAVG)

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

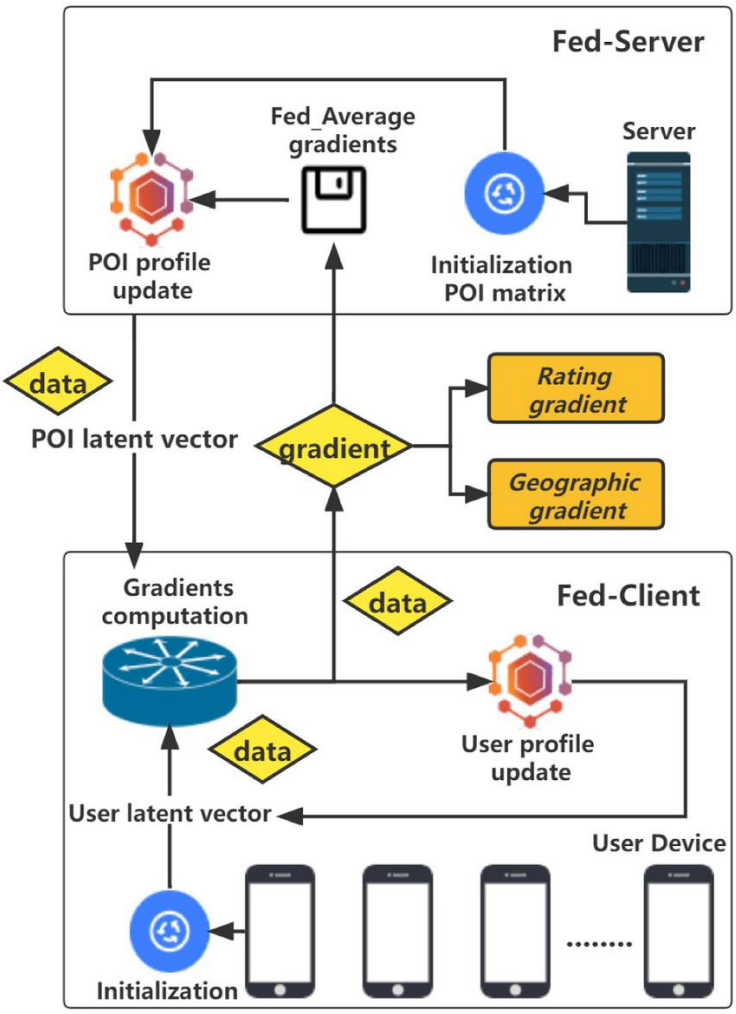


Figure. 5: Federated Stochastic Gradient Descent (FedSGD)

Federated Learning with Dynamic Regularization (FedDyn):
FedDyn introduces dynamic regularization to federated learning, addressing the challenge of heterogeneous data distributions across client devices. While traditional regularization methods aim to enhance generalization by penalizing the loss function, FedDyn adapts regularization to the unique characteristics of federated learning. By dynamically adjusting the regularization term based on factors like data volume and communication cost, FedDyn ensures that local losses converge effectively to the global loss. This personalized regularization approach optimizes model performance across diverse devices and data types, enhancing the efficiency and effectiveness of federated learning in real-world scenarios.

LITERATURE SURVEY

This section presents a detailed review about ML and FL frameworks implemented such as CNN models (VGG16, Inception, ResNet18 etc.) with FedAvg, SecAgg, Basic Fl and Methods of federated data partitioning like Federated Transfer Learning (FTL) and Vertical FL (VFL). (Sohan, M. F.,2023).

2.1 Related work on Federated Learning

Rauniyar, A., (2023) discussed the benefits of Federated Learning (FL) in various healthcare applications, emphasizing its ability to train models on distributed datasets while preserving data privacy. Specifically, they introduce FedHome, a FL framework tailored for personalized in-home health monitoring, which utilizes a

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

lightweight GCAE model to overcome statistical and communication challenges. FedHome demonstrates superior accuracy compared to traditional CNN techniques and reduces communication costs. Moreover, the authors propose a 5G-enabled FL architecture for COVID-19 diagnosis, enabling model sharing among institutions and cooperation with central cloud systems. Their investigation aims to develop AI models with robust generalization capabilities across multinational COVID-19 datasets. Additionally, the authors introduce a decentralized FL architecture with privacy-preserving encryption techniques for healthcare applications, highlighting its advantages in preserving IoT device data privacy within the Internet of Medical Things (IoMT) context. This decentralized approach contrasts with traditional centralized ML methods, offering significant security and privacy benefits. Furthermore, the integration of differential privacy with FL and blockchain technology is proposed for smart home monitoring of IoT data, ensuring both data privacy and security. Lastly, they introduce dynamic fusion-based federated learning for COVID-19 medical image analysis, facilitating collaborative model training across institutions without the need to share patient data, thereby enhancing model performance while preserving privacy.

Zhang et al. (2023) proposed a dynamic fusion method aimed at enhancing communication efficiency within federated learning (FL) systems. Clients autonomously decide to participate based on their model's performance, while the central server selects participants based on waiting time, which is determined by averaging each client's previous round's training time. Clients download the learning job from the central server, initiate local training, and set a timer based on their previous round times. If a client fails to complete training within the specified time or experiences a decline in performance, it may request to skip aggregation; otherwise, it notifies the server to update the model accordingly. The results from 18 experimental groups indicate that dynamic fusion-based federated learning (DF_FL) achieves lower accuracy than the default federated learning setting (D_FL) in only four groups, with differences ranging from 0.57% to 1.711%. Conversely, DF_FL demonstrates higher accuracy in the remaining 14 groups, consistently outperforming the default setting overall. Furthermore, the interference introduced in the fourth group of the dataset does not significantly affect the performance of the fusion-based FL model, highlighting its fault tolerance and robustness.

He, C. et al. (2020) introduced FedGKT, a federated learning framework designed specifically for edge devices. FedGKT integrates both Federated Averaging (FedAvg) and Supervised Learning (SL) approaches, leveraging local Stochastic Gradient Descent (SGD) training to reduce computational demands on edge devices. It adopts a novel approach by transferring knowledge from compact edge Convolutional Neural Networks (CNNs) to a larger cloud server CNN. Through an alternating minimization technique, FedGKT optimizes edge and server models iteratively, thereby improving training efficiency and introducing a novel knowledge distillation paradigm. The compact CNN deployed on edge devices comprises a lightweight feature extractor and classifier, trained locally to ensure uniform output tensor dimensions across all edge nodes. The server model is then trained utilizing features extracted from the edge-side model, aiming to minimize the discrepancy between ground truth labels and soft labels. Bidirectional knowledge transfer b
between edge and server enhances the performance of both models. Ultimately, the final model amalgamates the local feature extractor with the shared server model, resulting in an optimized federated learning framework for edge devices.

Hard, A. et al. (2018) employed federated learning to train an on-device RNN language model for smartphone virtual keyboard next-word prediction. Federated learning utilizes a decentralized computation approach to train neural models, particularly suitable for mobile devices acting as clients. Instead of transferring data to central servers for training, clients process their local data and share model updates with the server. The server then aggregates these updates to create an improved global model. This distributed method remains effective even with unbalanced datasets and non-independent data across clients. The Federated Averaging algorithm plays a crucial role in this process, combining client updates on the server to generate a new global model. Each client computes the average gradient on its local data using stochastic gradient descent, adapting to its dataset size and characteristics. The training of both the server and federated CIFG models demonstrates enhanced top-1 and top-3 recall compared to the baseline n-gram FST model. Despite the larger vocabulary and inclusion of personalized components like user history and contacts LMs in the n-gram model, CIFG achieves significant improvements. Notably, federated CIFG outperforms server-trained CIFG, as evidenced by evaluation on client cache data showing a relative 5% improvement (0.8% absolute) in top-1 recall compared to server-trained CIFG. Although comparisons on server-hosted logs data show similar recall between the two models, it's important to note that the logs may not entirely represent the true performance.

2.2 Related work on Deep Learning

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

In the analysis of dataset variability, it was found that more than half of the methods utilized multiple datasets, with approximately 30% relying solely on private ones Renard, F. et al., (2020). A notable observation was the insufficient description of data augmentation in 13% of articles, indicating a transparency gap in method reporting. Optimization strategies varied, with one article introducing a novel approach of merging results from three deep learning (DL) models, while others lacked discussions on variability management, suggesting potential for further exploration in this area. Hyperparameter explanations were limited, with only one article detailing tuning using a grid search. The training proportion, considered a hyperparameter, exhibited significant variability, ranging from 20% to 95% of the dataset size. However, despite the wide range of training proportions, the predominant use of convolutional neural network (CNN) or recurrent neural network (RNN) architectures suggested diverse approaches across studies (Renard, F. et al., 2020). In terms of middleware and infrastructure, various tool boxes were employed, with only one article utilizing an in-house implementation. While most articles leveraged GPUs for computation, the lack of references to distributed systems indicated limited exploration of distributed computing in DL implementations.

An improved fuzzy clustering algorithm has been developed by Hu, M. et al., (2020) to address the challenges posed by the complexity of human brain tissues and noise in MRI images. Named SIPMFCM (Fuzzy Clustering based on Spatial Information Fusion PM), this algorithm integrates various techniques such as kernel distance metric, membership constraints, and a regularization parameter ρ to improve segmentation accuracy and detail retention. By incorporating local spatial information, SIPMFCM aims to enhance segmentation, particularly in scenarios with high-intensity noise, resulting in smoother image edges. The algorithm modifies the traditional FCM expression using kernel distance measurement, thereby improving fuzzy division distinction and incorporating neighborhood space restriction to enhance robustness against image noise. By integrating spatial function into the membership function, SIPMFCM aims to achieve better performance in brain image processing and disease diagnosis prediction while ensuring safety.

Autoencoder (AE), as described by Puttagunta, M et al., (2021) , serves as a model for unsupervised representation learning, where it encodes input x into z and decodes it back to x' through a hidden layer h. The process involves encoding, decoding, and minimizing reconstruction error. Restricted Boltzmann machines (RBMs) function as Markov Random Fields (MRFs) with visible and hidden units, operating independently of each other. They establish energy for states {v,h}, define joint probability distributions, and compute conditional distributions. Deep Belief Networks (DBNs) are formed by stacking RBMs, enabling feature extraction and hierarchical representation learning. They consist of visible and hidden layers, constructing a directed generative model.Convolutional Neural Networks (CNNs) leverage shared weights, local receptive fields, and spatial sub-sampling for pattern recognition [9]. They handle unstructured data through convolution operations, which reduce parameters and allow deeper networks. Pooling methods such as max-pooling and average pooling are employed to select superior features or averages in regions. Fully connected layers at the end process flattened feature maps from previous layers.

Various CNN-based techniques have been proposed for the analysis of medical images, demonstrating promising results. One method focused on lung disease categorization using CT scans, leveraging a dataset comprising 14696 image patches. Another approach utilized a convolutional classification restricted Boltzmann machine for lung CT image analysis (Anwar, S. M. etr al., 2018). Additionally, a method based on multiple instances of deep learning was applied for body organ recognition. For colon cancer detection, a method combining a locality-sensitive deep learning algorithm with CNN was employed. Moreover, a CNN-based system aimed at content-based medical image retrieval using radiographic images showed effectiveness. Another method involved a hybrid thyroid module diagnosis system that utilized pre-trained CNNs, achieving satisfactory results. Furthermore, CNN-based techniques for diagnosing breast cancer and diabetic retinopathy demonstrated high accuracy. Additionally, deep neural networks like GoogLeNet and ResNet were employed for multi-class classification of Alzheimer's disease patients, showing promising results.

In a study by Yan et al., various methods were utilized for classification, employing different features and classifiers. Initially, Bag of Words (BoW) combined with Scale-Invariant Feature Transform (SIFT) descriptors and linear regression (LR) achieved 62.21% precision, 63.37% recall, and 62.78% F1 score. BoW and SIFT with Support Vector Machine (SVM) improved performance to 63.72%, 64.63%, and 64.17%, respectively. Histogram of Oriented Gradients (HOG) with LR/SVM achieved higher precision, recall, and F1 score: 67.74%, 68.71%, and 68.22%, and 76.39%, 76.75%, and 76.57%, respectively. CNN surpassed all, achieving 92.25%, 92.21%, and 92.23% precision, recall, and F1 score (Anwar, S. M.et al., 2018).

The method involves two main steps: bag preparation and Multiple-Instance Learning (MIL). Initially, patches are extracted from Whole Slide Images (WSIs) to generate mosaics. Then, a MEM model is utilized for MIL, ensuring

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

privacy through a combination of local training and central aggregation Adnan, M. et al.(2022). Representative patches, known as mosaics, are extracted from WSIs using a selection method. This process entails removing non-tissue regions, clustering patches, and randomly selecting a subset to compose each mosaic. These mosaics are subsequently converted into bags for MIL, with feature vectors derived from a pre-trained DenseNet model. The MEM model comprises memory units within a block, generating a permutation invariant representation from input sequences. Each unit transforms sequences into attention vectors, capturing relationships among elements. The resulting output sequence is invariant to permutations, a crucial characteristic for MIL tasks.

Research Gap

Federated Learning (FL) enables collaborative model training across decentralized data sources while preserving privacy. Various FL frameworks, including CNN models like VGG16, Inception, and ResNet18, alongside techniques such as FedAvg, SecAgg, and Basic FL, have been developed. Methods like Federated Transfer Learning (FTL) and Vertical FL (VFL) have also been explored for federated data partitioning. However, FL faces challenges such as data privacy risks, communication overhead, and optimizing model performance across decentralized datasets, particularly in domains like medical image analysis. Resource constraints on edge devices further complicate FL implementation. Similarly, in Deep Learning (DL), interpretability, data scarcity, computational complexity, overfitting, biases, and lack of explainability pose significant challenges. Addressing these limitations is essential for realizing the full potential of FL and DL across various domains.

Contribution:

Contribution of the paper is as follows:

1. The paper proposes a comprehensive framework that integrates Federated Learning (FL) with deep learning models for medical image analysis. This approach ensures that sensitive patient data remains private while still enabling the development of effective diagnostic models across multiple decentralized sources.
2. The paper provides a detailed evaluation of multiple deep learning models, including a basic CNN, VGG16, ResNet50, and InceptionV3, for the task of chest X-ray analysis. The experimental results demonstrate that InceptionV3 outperforms the other models, offering valuable insights into model selection for medical image tasks.
3. One of the key contributions is the practical implementation of Federated Learning to address privacy concerns. By using FL, the models are trained collaboratively across different devices without transferring sensitive data, preserving patient privacy while improving model performance.
4. The paper employs the FedAvg algorithm for aggregating model updates from multiple clients. This decentralized approach not only improves the model over time but also demonstrates the feasibility of FL in healthcare applications, where data privacy is crucial.

METHODS

Federated Learning is at the forefront of ensuring robust data privacy and security in tandem with training deep learning models. In our methodology, we distribute a global model to local devices, each of which houses sensitive medical images. These devices independently train the model using their data, obviating the need to share actual images. Instead, only model updates—such as weights and biases—are transmitted to a central server, where they undergo aggregation to refine the global model. This iterative process fosters collaborative learning across diverse datasets while safeguarding sensitive information from centralization.

Our approach not only facilitates collaborative model training across multiple healthcare institutions and distributed data sources but also prioritizes data privacy. It empowers healthcare professionals to access a wealth of medical imaging datasets from various institutions without the necessity of physically relocating or centralizing the data. By integrating the potent capabilities of deep learning with Federated Learning, our methodology promotes privacy-conscious collaboration, serving as a valuable resource for radiologists and healthcare providers alike. Timely and accurate diagnoses are pivotal for effective treatment planning, mitigating medical errors, and ultimately enhancing patient outcomes.

Our architecture that simulates the federated system involves dividing the COVIDx CXR-4 dataset (Wu, Y. et al., 2023) into 10 virtual clients. Random 8 clients are chosen to train a deep learning model for 10 epochs, and this process repeats for 3 rounds. The resulting weights, biases, accuracy losses, and other parameters are transmitted to a central server. The server then employs the Federated Averaging (FedAvg) algorithm to enhance the overall performance of the model.

Medical Image Dataset:

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

COVIDx CXR-4 (Wu, Y. et al., 2023), dataset contains 84,818 images from 45,342 subjects, and includes separate validation and test sets. The COVIDx CXR-4 dataset, available on Kaggle, is a collection of chest X-ray images specifically curated for the detection of COVID-19. With the global outbreak of the COVID-19 pandemic, there has been a growing interest in leveraging medical imaging, such as chest X-rays, for the diagnosis of the disease. This dataset aims to provide researchers and practitioners with a standardized and curated collection of chest X-ray images to facilitate the development and evaluation of machine learning models for COVID-19 detection. The dataset comprises chest X-ray images obtained from various sources, including public repositories, research articles, and healthcare institutions. These images are labeled according to their respective classes, including COVID-19 positive, viral pneumonia (non-COVID), bacterial pneumonia, and normal (healthy) cases. Each image is accompanied by metadata providing information such as patient demographics, clinical history, and imaging parameters.

Base DL Model for local devices:

We implemented three DL models - Inception V3, Res-Net-50 and VGG-19 as a base model for local machines. We selected InceptionV3 as the model for our federated learning (FL) along with a simple CNN network for comparison as the base deep learning models due to its exceptional accuracy and robust performance with the specific dataset we are working with. InceptionV3's proven ability to capture intricate features and hierarchical representations in image data makes it an ideal choice for our task. Its pre-trained weights, trained on large-scale datasets, enable us to harness the power of transfer learning, saving us valuable time and resources in training a model from scratch.

Federated Average (FedAvg) algorithm

Federated Averaging (FedAvg) orchestrates model training via a central server that hosts the shared global model. However, the critical optimization steps occur locally on individual client devices, where the chosen deep learning model is applied. These local devices train their respective models using their unique data and computational resources. Subsequently, the locally trained weights and other relevant parameters are transmitted back to the central server.

At the central server, sophisticated averaging techniques are employed to aggregate the parameters received from the clients. By combining these parameters, the global model is further refined. This collaborative approach allows the model to learn from the diverse data distributed across the client devices, ultimately enhancing its effectiveness and efficiency.

FedAvg's decentralized architecture offers several advantages. Firstly, it addresses privacy concerns by keeping raw data localized on client devices, mitigating the risks associated with centralized data storage. Secondly, it promotes scalability by distributing the computational load across multiple devices, enabling large-scale training without overburdening any single server. Additionally, FedAvg adapts to varying data distributions and device capabilities, making it suitable for diverse deployment scenarios.

In Fig. 6 the flowcharts begin with a base model that starts with pre-processing techniques used to normalize the data as much as possible. To achieve a higher quality image result, techniques like applying noise reduction filters. Resampling uses methods to standardize pixel sizes and spatial orientations so that it's important for some analyses to ensure uniform resolution across all images. Pixel values, when scaled to a standard range, can aid in the consistent training of deep learning models. The dataset is split as 70% training data, 20% testing data and 10%

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

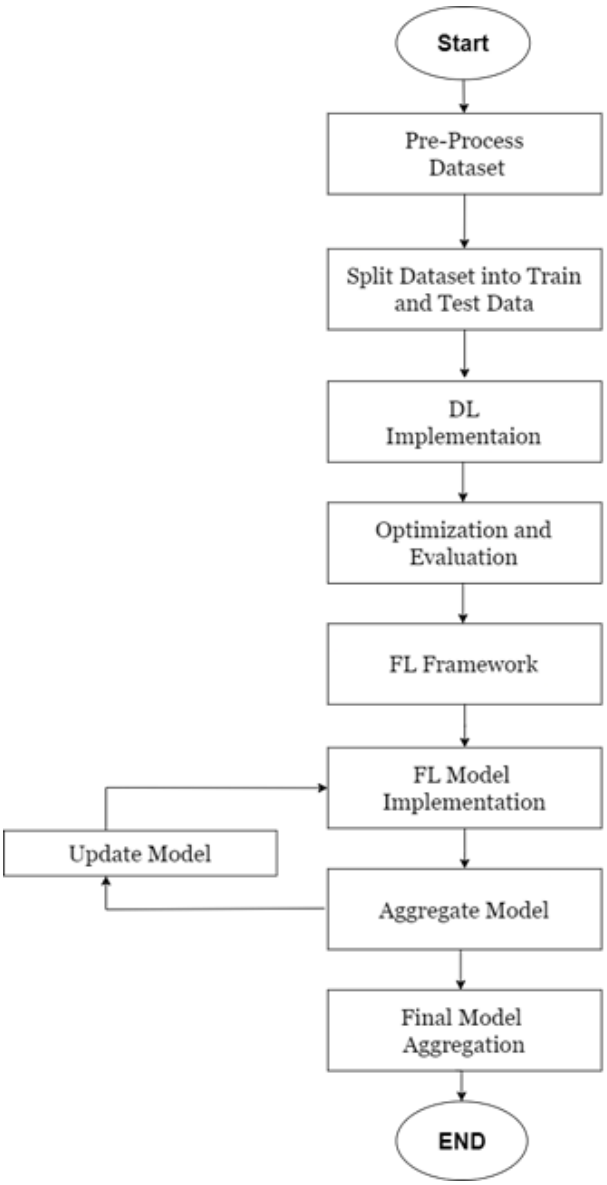


Figure. 6: Proposed System Flowchart

validation data. DL implementation stands for the base model which in our case in ResNet50 which is then hypertuned to achieve utmost performance that optimizes the results. FL framework to implement FedAvg algorithm. Each device conducts its own model updates using its unique dataset, without sharing the actual data. The device will train its model on its unique data and send the model's weights to the central server. For a designated number of rounds, the model updates are rounded up by the central server using the FedAvg algorithm that finds the average of the model weights. The localized devices then revive the aggregated model for extra tweaking. In this way, the model learns without jeopardizing the confidentiality of the data.

RESULTS

To implement the federated learning framework on top of basic models, libraries like PySyft are crucial, especially for federated learning. PySyft enhances PyTorch by adding features for protecting privacy and ensuring secure federated learning, allowing smooth incorporation with current deep learning processes. We use PySyft to deploy federated learning methods like Federated Averaging, guaranteeing data privacy and security among decentralized clients.

Federated learning algorithm with CNN and InceptionV3:

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Utilizing Federated Averaging (FedAvg) with a basic CNN network as well as InceptionV3 shows great potential for enhancing medical image analysis in a federated learning system. In the scenario of the basic CNN network, FedAvg manages the joint training process on various client devices while also maintaining the privacy of data. Every client independently trains its CNN model with its own medical images to keep patient data secure on the device. The model parameters trained locally are combined on a central server through averaging methods to continually improve the overall model. This method allows the CNN model to gain insights from varied data from various medical facilities, enhancing its ability to generalize and be resilient. While integrating FedAvg with InceptionV3 simultaneously, it bolsters the analysis capabilities because InceptionV3 excels in capturing complex features in medical images with its multi-scale convolutional architecture. The federated learning framework enhances accurate diagnosis and treatment planning in medical image analysis tasks by utilizing the scalability and privacy-preserving elements of FedAvg with the advanced feature extraction capabilities of InceptionV3, all while ensuring patient privacy.

Performance Evaluation of deep learning algorithms:

Table1 presents the performance metrics of three different deep learning (DL) algorithms—InceptionV3, VGG-19, and ResNet-50—on both training and validation datasets. For the training data, InceptionV3 achieved a low loss of 0.0183 and a high accuracy of 99.31%. On the validation data, the loss slightly increased to 0.0958, while the accuracy remained relatively high at 93.75%. VGG-19 yielded a higher loss of 0.2392 and a slightly lower accuracy of 88.75% on the training data. However, on the validation set, it showed a comparable performance with a loss of 0.2617 and a slightly higher accuracy of 90.62%. ResNet-50 performed well on the training data with a moderate loss of 0.1630 and a high accuracy of 93.77%. However, its performance decreased notably on the validation set, with a substantially higher loss of 0.6918 and a reduced accuracy of 68.75%.

Table 1: Comparison of Deep Learning Model Performance on Training and Validation Data

	For Training Data		For Validation Data	
DL Algorithms	Loss	Accuracy	Loss	Accuracy
InceptionV3	0.0183	0.9931	0.0958	0.9375
VGG-19	0.2392	0.8875	0.2617	0.9062
ResNet-50	0.1630	0.9377	0.6918	0.6875

Thus, InceptionV3 showed the best overall performance, maintaining high accuracy on both training and validation data with relatively low loss. VGG-19 performed decently but slightly lagged behind InceptionV3 in terms of accuracy. ResNet-50 exhibited good performance on the training data but struggled to generalize well to the validation set, indicating potential overfitting.

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

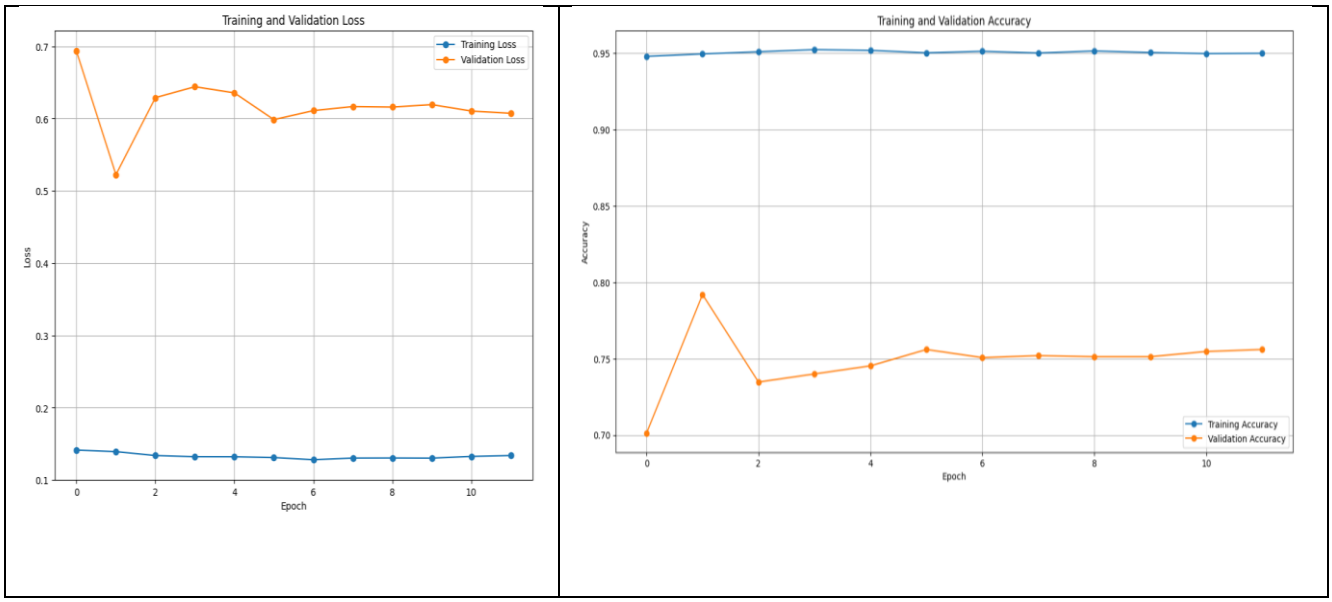


Figure 7: Resnet model Training and Validation COVID-19 losses and accuracies

Fig 7 shows the training and validation accuracy of a ResNet model on COVID-19 X-ray data. Both lines rise with training epochs, indicating the model is learning effectively and avoiding overfitting. This suggests the ResNet model has potential for accurate chest X-ray classification. This graph shows the training loss of a ResNet model for COVID-19 X-ray data. The decreasing loss suggests the model is learning the patterns in the data, which is a positive step towards accurate chest X-ray classification.

Accuracy and Loss of VGG -19 Model for COVID Dataset

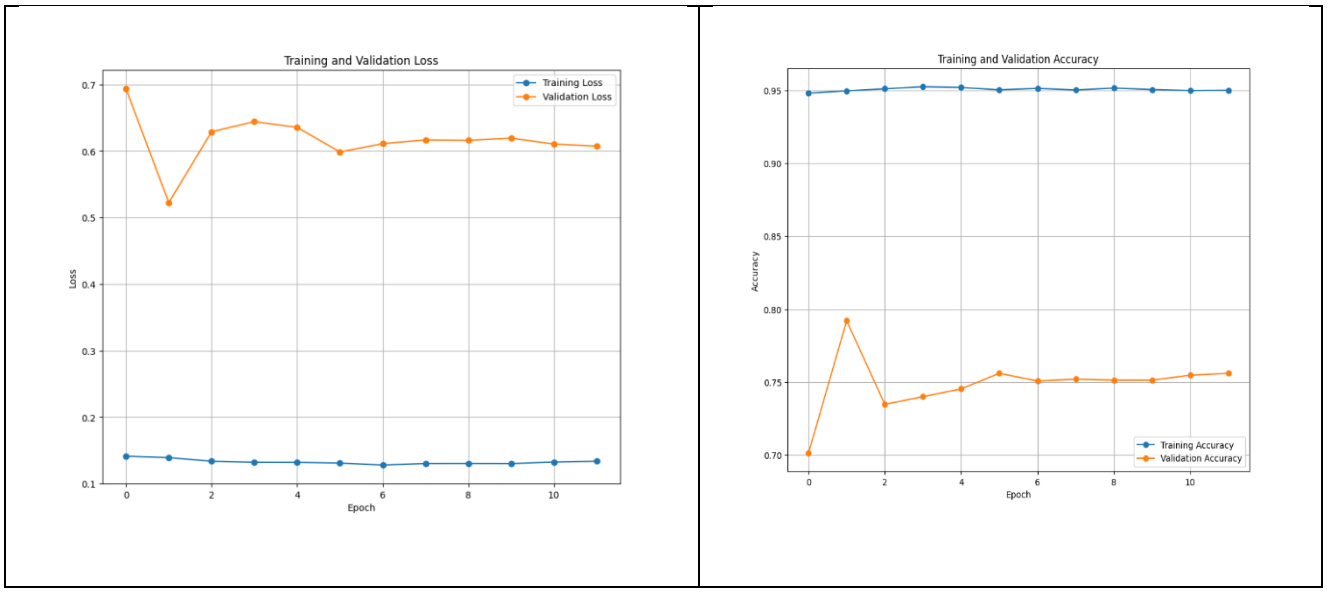


Figure.8: VGG -19 model Training and Validation Loss and accuracies

Fig. 8 shows the training and validation accuracy of a ResNet model on COVID-19 X-ray data. Both lines rise with training epochs, indicating the model is learning effectively and avoiding overfitting. This suggests the ResNet model has potential for accurate chest X-ray classification. This graph shows the training loss of a ResNet model for COVID-19 X-ray data. The decreasing loss suggests the model is learning the patterns in the data, which is a positive step towards accurate chest X-ray classification.

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Accuracy and Loss of InceptionV3 Model for COVID Dataset

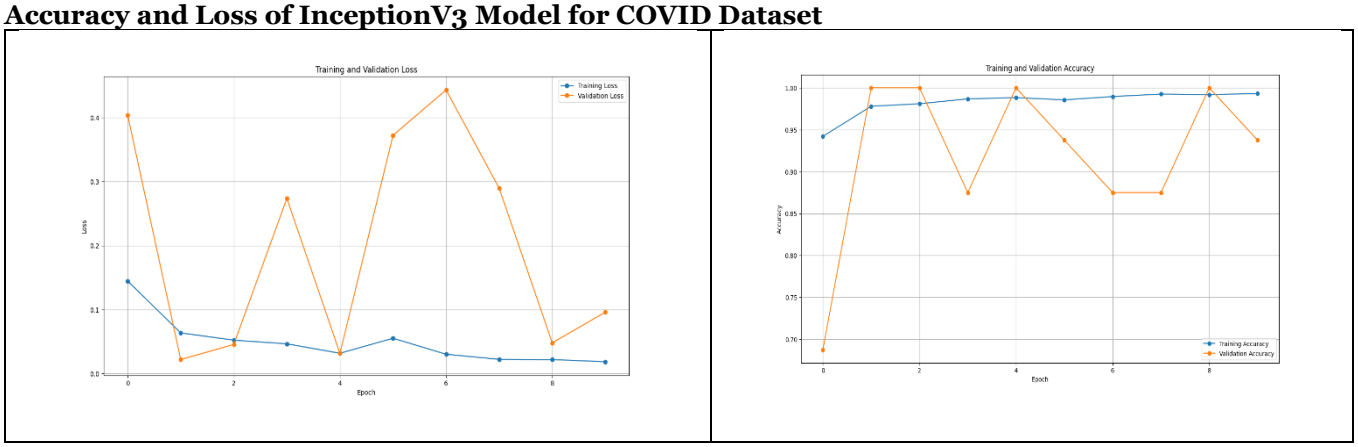


Figure. 9: Inception V3 model Training and Validation Loss and accuracies

The graph in Fig. 9 shows the Inception V3 model's accuracy on COVID-19 X-ray data . Both training and validation accuracy rise with epochs, indicating the model is learning effectively and avoiding overfitting. This suggests good potential for accurate chest X-ray classification. This graph shows the training loss of an Inception V3 model for COVID-19 X-ray data. The decreasing loss signifies the model is learning the patterns in the data, which is a positive step towards accurate chest X-ray classification.

Accuracy of Client data across the rounds on the deep learning model

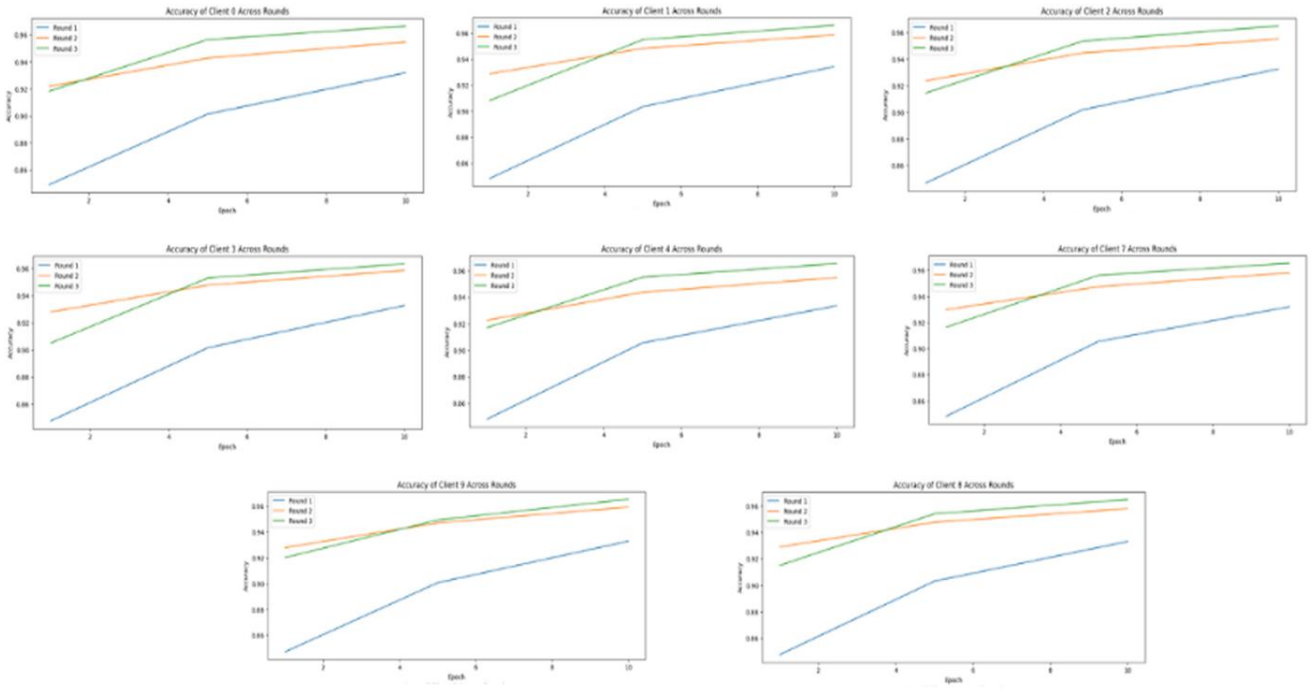


Figure. 10: Trained Parameters of 8 clients

The graphs in Fig. 10 likely shows client-side performance within a federated learning framework, focusing on a single client across three training rounds (1, 2, and 3). The X-axis likely represents training epochs within each round, where the client trains its local model on its own data. The Y-axis likely shows two metrics: client accuracy (how well the client's model performs on its own data) and client loss (how well the model fits the client's data). Ideally, client accuracy should increase over epochs within each round, indicating the model is learning. Similarly, client loss should decrease within each round, signifying a better fit between the model and the client's data. It's important to remember that this graph only reflects the client's performance, and the server-side evaluation provides a broader picture of the

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

federated learning process. As the number of clients grows and multiple training iterations occur within each client, the accuracy progressively improves.

Server Accuracy:

In federated learning, devices train a model collaboratively without sharing data. The server evaluates the global model on unseen data (hold-out set). The provided graph likely shows training epochs (x-axis) and server-side performance metrics (y-axis). Server loss (how well the model fits training data) should decrease, while server accuracy (correct predictions on unseen data) should increase over epochs (ideally across many rounds). This initial analysis of the first 3 rounds offers insights into the model's early learning progress.

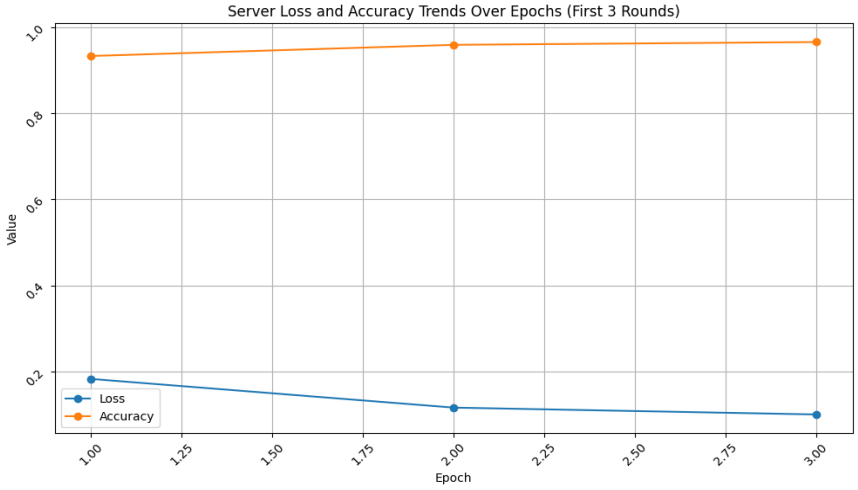


Figure. 11: Accuracy and Loss on the Server using FedAvg

Client Accuracies Over the rounds:

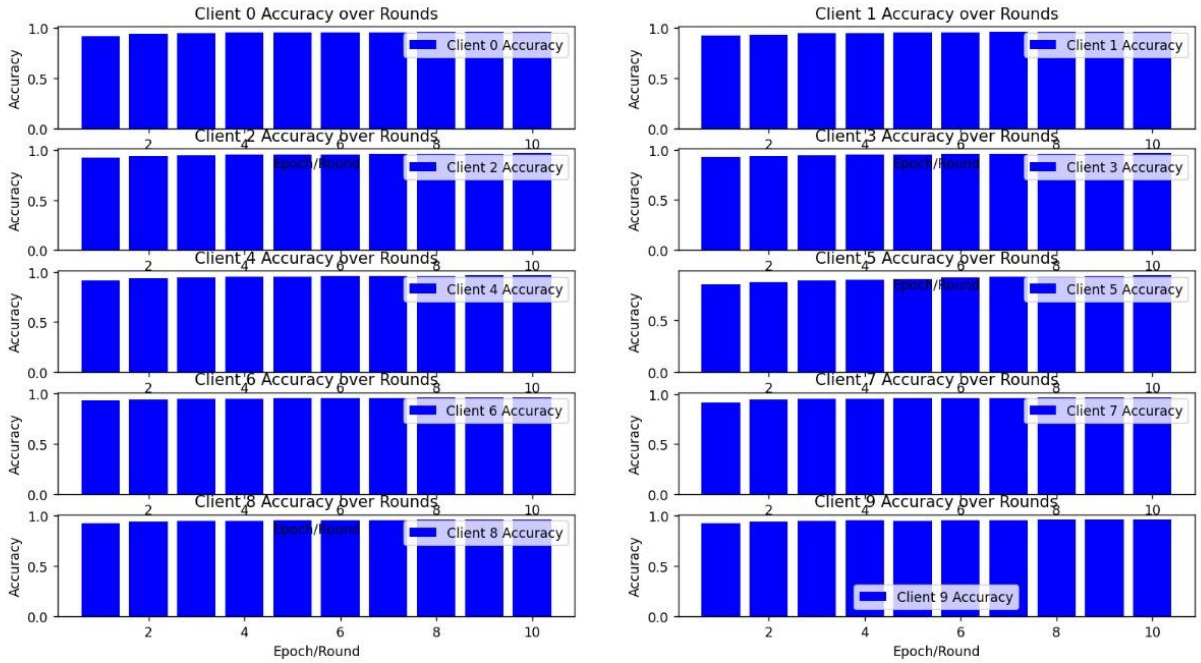


Figure 12: Client 0 to Client 9 accuracies over the 10 epochs

From Fig. 12 we can observe that each client's accuracy increases as the no. of epochs increases. with each epoch the client performance improves. We trained each client for 10 epochs.

CONCLUSION

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Federated learning enables collaborative model training across multiple healthcare institutions data sources while preserving data privacy. Deep and federated learning allow access to a range of medical imaging datasets from different institutions without centralizing the data. This study explores the ability to leverage deep learning's powerful capabilities in medical image analysis while respecting privacy and collaboration constraints through federated learning. This is a helping hand to the small community of radiologists along with health professionals for effective treatment planning, reducing medical errors, and improving patient outcomes. The implementation of Federated Learning (FL) using PyTorch, as demonstrated through Python scripts, showcases the practical application of collaborative model training while prioritizing data privacy. Through techniques such as noise reduction filters and pixel value standardization, image data normalization is achieved, facilitating the training of deep learning models like ResNet50 and Inception. The dataset division into validation, testing, and training subsets ensures robust model evaluation. FL, enacted with algorithms such as FedAvg and FedSGD, empowers devices to update models locally, preserving data anonymity. By orchestrating communication between clients and a central server, FL enables decentralized model training across distributed clients, enhancing scalability and privacy. The incorporation of neural network architectures like Inception augments feature extraction capabilities, pivotal for tasks such as image classification. While the provided Python scripts offer a comprehensive framework for FL implementation, considerations for dataset specifics and metrics aggregation across clients.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFRENCES

[1] Anwar, S. M., Majid, M., Qayyum, A., Awais, M., Alnowami, M., & Khan, M. K. (2018). Medical image analysis using convolutional neural networks: a review. *Journal of medical systems*, *42*, 1-13.

[2] Adnan, M., Kalra, S., Cresswell, J. C., Taylor, G. W., & Tizhoosh, H. R. (2022). Federated learning and differential privacy for medical image analysis. *Scientific reports*, *12*(1), 1953.

[3] Antunes, R. S., André da Costa, C., Küderle, A., Yari, I. A., & Eskofier, B. (2022). Federated learning for healthcare: Systematic review and architecture proposal. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *13*(4), 1-23.

[4] Bharati, S., Mondal, M. R. H., Podder, P., & Prasath, V. S. (2022). Federated learning: Applications, challenges and future directions. International Journal of Hybrid Intelligent Systems, 18(1-2), 19-35.

[5] Hard, A., Rao, K., Mathews, R., Ramaswamy, S., Beaufays, F., Augenstein, S., ... & Ramage, D. (2018). Federated learning for mobile keyboard prediction. *arXiv preprint arXiv:1811.03604*.

[6] He, C., Annavaram, M., & Avestimehr, S. (2020). Group knowledge transfer: Federated learning of large cnns at the edge. *Advances in neural information processing systems*, *33*, 14068-14080.

[7] Hu, M., Zhong, Y., Xie, S., Lv, H., & Lv, Z. (2021). Fuzzy system based medical image processing for brain disease prediction. *Frontiers in Neuroscience*, *15*, 714318.

[8] Puttagunta, M., & Ravi, S. (2021). Medical image analysis based on deep learning approach. *Multimedia tools and applications*, *80*(16), 24365-24398.

[9] Rauniyar, A., Hagos, D. H., Jha, D., Håkegård, J. E., Bagci, U., Rawat, D. B., & Vlassov, V. (2023). Federated learning for medical applications: A taxonomy, current trends, challenges, and future research directions. *IEEE Internet of Things Journal*, *11*(5), 7374-7398.

[10] Renard, F., Guedria, S., Palma, N. D., & Vuillerme, N. (2020). Variability and reproducibility in deep learning for medical image segmentation. *Scientific Reports*, *10*(1), 13724.

[11] Sohan, M. F., & Basalamah, A. (2023). A systematic review on federated learning in medical image analysis. *IEEE Access*, *11*, 28628-28644.

[12] Singh, A., Sengupta, S., & Lakshminarayanan, V. (2020). Explainable deep learning models in medical image analysis. Journal of imaging, 6(6), 52.

[13] Suganyadevi, S., Seethalakshmi, V., & Balasamy, K. (2022). A review on deep learning in medical image analysis. International Journal of Multimedia Information Retrieval, 11(1), 19-38.

[14] Wu, Y., Gunraj, H., Tai, C. E. A., & Wong, A. (2023). COVIDx CXR-4: An Expanded Multi-Institutional Open-Source Benchmark Dataset for Chest X-ray Image-Based Computer-Aided COVID-19 Diagnostics. arXiv preprint arXiv:2311.17677.

[15] Zhang, W., Zhou, T., Lu, Q., Wang, X., Zhu, C., Sun, H., ... & Wang, F. Y. (2021). Dynamic-fusion-based federated learning for COVID-19 detection. *IEEE Internet of Things Journal*, *8*(21), 15884-15891.