International Journal of Mathematical, Engineering and Management Sciences
Vol. 10, No. 2, 464-485, 2025
https://doi.org/10.33889/IJMEMS.2025.10.2.023

Ram Arti Publishers

A Collaborative Privacy Preserved Federated Learning Framework for Pneumonia Detection using Diverse Chest X-ray Data Silos

Shagun Sharma
Chitkara University Institute of Engineering and Technology,
Chitkara University, Punjab, India.
E-mail: shagunsharma7098@gmail.com, shagun.sharma@chitkara.edu.in

Kalpna Guleria
Chitkara University Institute of Engineering and Technology,
Chitkara University, Punjab, India.
Corresponding author: guleria.kalpna@gmail.com, kalpna@chitkara.edu.in

(Received on July 7, 2024; Revised on October 4, 2024 & November 19, 2024; Accepted November 25, 2024)

Abstract
Pneumonia detection from chest X-rays remains one of the most challenging tasks in the traditional centralized framework due to the requirement of data consolidation at the central location raising data privacy and security concerns. The amalgamation of healthcare data at the centralized storage leads to regulatory concerns passed by the governments of various countries. To address these challenges, a decentralized, federated learning framework has been proposed for early pneumonia detection in chest X-ray images with a 5-client architecture. This model enhances data privacy while performing collaborative learning with diverse data silos and resulting in improved predictions. The proposed federated learning framework has been trained with a pre-trained EfficientNetB3 model in the Independent and Identically Distributed (IID) and non-IID data distributions, while the model updation has been performed using federated proximal aggregation. The configuration of the proximal term has been kept as 0.05, achieving an accuracy of 99.32% on IID data and 96.14% on non-IID data. In addition, the proximal term has also been configured to 0.5, resulting the accuracy levels of 92.05% and 96.98% in IID data and non-IID data distributions, respectively. The results of the proposed model demonstrate the effectiveness of the federated learning model in pneumonia detection, highlighting its potential for real-world applications in decentralized healthcare configurations.

Keywords- Federated learning, Decentralized learning, Collaborative learning, Deep learning, Pneumonia detection, EfficientNetB3.

1. Introduction

The COVID-19 pandemic has been recognized as one of the most catastrophic healthcare crises, resulting in devastating loss of life worldwide (Malik et al., 2023; Rushdi and Serag, 2020). The virus spreads through respiratory droplets when an infected individual coughs, sneezes, or talks. Symptoms such as cough, fever, diarrhea, muscle weakness, and loss of smell indicate a viral infection, posing a significant threat to global health and safety (Sharma and Guleria, 2023c). If left unchecked, COVID-19 progresses to pneumonia, a severe and potentially fatal infection that inflames the lungs, filling air sacs with fluid and pus (Chouhan et al., 2020; Sharma and Guleria, 2022; Sharma and Guleria, 2023b). This congestion leads to breathing difficulties, fever, coughing up phlegm, confusion, and a bluish tint to the face and lips due to oxygen deprivation (Mujahid et al., 2022; Sharma and Guleria, 2023c). In pneumonia immune system doesn't respond effectively to the virus, resulting in lung inflammation. This disease affects people with old age and those with existing lung diseases, ranging from mild to fatal results (Srivastav et al., 2023). As per the statistics provided in (Kareem et al., 2023), each year almost 200 million people are affected by pneumonia leading to a 10 times higher fatality rate in comparison to the deadly HIV disease. In the last few years, pneumonia cases have rapidly increased which has developed difficulty in managing hospitalization and resources. However, the advancements in technology with Artificial Intelligence (AI), have transformed

the disease diagnosis process by providing extra assistance with improved identification of the symptoms in the chest X-ray images (Zaffiri et al., 2012). The deep learning (DL) technique has also emerged as additional support to traditional disease detection procedures which automatically examines the image data, extracts the essential features, and provides the potential outcome of the disease existence (Castellano and Vessio, 2022; Lepcha et al., 2023; Mabrouk et al., 2023; Ragnarsdottir et al., 2024; Sharma and Kukreja, 2022; Sharma et al., 2023). However, the DL models pose a challenge of centralized data storage, raising concerns about data privacy and security (Rawat and Suryakant, 2019). The centralized DL-based models require the data to be consolidated at the central location, limiting the model training, which imposes difficulty in accurate disease prediction in healthcare scenarios. Data ownership and control can also be disputed while using DL models for disease prediction. Further, the variation in the data quality may result in diminished performance. The DL models lead to the potential problems of sharing the data across different healthcare providers and can also lead to the challenge of privacy and regulatory concerns passed by the governments of various countries (Goyal et al., 2022; Riedel et al., 2023; Rodrigues et al., 2022; Sharma and Guleria, 2023c). To address these challenges, the federated learning (FL) framework was introduced in 2016, which is a decentralized framework that focuses on privacy preservation, reduces latency, and complies with data regulatory norms (Sharma and Guleria, 2023a). In FL, data remains decentralized, and the server sends models to clients for training, which further reverts model updates to the server after successful training. These client updates are aggregated using the FedAVG method and the updation in the global model on the server side has been done (Makkar and Santosh, 2023; Sharma and Guleria, 2023a). This procedure continues until an improved and accurate disease detection model is developed. In the proposed work, a privacy-preserved FL-based pneumonia detection model has been developed using a pre-trained EfficientNetB3 with FedProx aggregation. This aggregation leads to improved results over traditional FedAVG in non-IID distributions of the data. Further, one of the major research gaps addressed is the configuration of the proximal term to optimal value in the case of non-IID data distribution, which resembles real-life scenarios for achieving better accuracy outcomes in a privacy-preserved FL environment.

The contribution of the proposed work has been provided below:

- This work implements a privacy-preserved federated learning framework using a pre-trained EfficientNetB3 model with FedProx aggregation for handling client heterogeneity and data distribution variability in a decentralized 5-client architecture.
- The FedProx aggregation was configured with the optimal value of the proximal term "μ" for mitigating the convergence issues caused by data variability in each client. This configuration provides adaptive environments in independent and identically distributed (IID) data and performs better regularization in the non-IID distributions.
- The configuration of hyperparameters for training the proposed distributed federated learning model has been configured to 0.001, 32, 5, 15, Adam, and 5 for learning rate, batch size, clients, rounds, optimizer, and epochs, respectively.

The remaining structure of the article contains Section 2, which discusses the existing pneumonia detection models, and Section 3 elaborates on the X-ray dataset and proposed methodology. Section 4 discusses the results obtained by the proposed model and Section 5 provides the conclusion of the work.

2. Related Work

In Kandati and Gadekallu (2023), a particle swarm optimization (PSO) technique was integrated with federated learning (FL) for classifying pneumonia, COVID-19, and normal lungs. The client utilization rates were set at 100%, 50%, 20%, and 10%, corresponding to client values of 10, 5, 2, and 1. The model was configured with key hyperparameters: batch size (10), training epochs (5), and testing epochs (30).

However, a significant research gap exists as the authors failed to disclose the data distribution among clients. In Mabrouk et al. (2023), a FL framework was employed with various models (DenseNet169, MobileNetV2, Xception, InceptionV3, VGG16, DenseNet121, ResNet152v2, and ResNet50). The two top-performing models were aggregated and implemented globally on the server. Meanwhile, Kareem et al. (2023) developed a FL-based approach using VGG19, ResNet50, DenseNet, and AlexNet models, with a configuration of (4) clients and (20) epochs. The dataset exhibited class imbalance, which was addressed through exploratory data analysis (EDA), resulting in a balanced dataset. ResNet50 achieved the highest accuracy of 93% for pneumonia and normal lung classification. In Malik et al. (2023), a DMFL_Net model was proposed, based on a pre-trained DenseNet169 in FL framework, for multiclass classification of tuberculosis, pneumonia, pneumothorax, and lung cancer. Furthermore, Riedel et al. (2023) introduced a ResNetFed model for COVID-19 pneumonia detection in CXR images, utilizing a limited dataset of 1411 images without augmentation. Despite these limitations, the ResNet50 model outperformed existing models in COVID-19 classification. Lastly, Alfiansyah et al. (2024) developed a pneumonia detection model using FL on Mendeley and RSNA datasets (5,856 and 14,863 images, respectively), achieving an accuracy of 91.90% with IID data across three clients. A federated learning (FL) approach for pneumonia detection was employed in (Farkaş et al., 2023), leveraging a deep neural network (DNN) based on the (VGG16) model. The hyperparameters were configured as follows: stochastic gradient descent (SGD) as the optimizer, (25) rounds, and rectified linear unit (ReLU) as the activation function. The client value was varied from (1) to (10), and the performance was evaluated for each client utilization. The results showed the highest accuracy of 96.875% with a single client and the lowest accuracy of 90.47% with (10) clients.

In Malik and Anees (2024), authors implemented VGG19, DenseNet201, InceptionV3, and DenseNet169 models in a federated framework for classifying Covid-19 from other chest diseases, including pneumonia, lung cancer, pneumothorax, consolidation lungs, and tuberculosis. The model training was performed with an augmented dataset using IID and non-IID distributions in (150) rounds, which identified that the FL-DenseNet201 model achieves the highest accuracy of 97.89%, whereas the lowest accuracy was resulted using the FL-VGG19 model as 90%. In Adjei-Mensah et al. (2024), the authors proposed a Cov-Fed model for identifying Covid-19. This FL framework has been implemented on multiple clients using two different datasets of chest X-rays. The development of this framework was performed using a multi-ECA technique to enhance feature maps in medical data. The performance of the Cov-Fed model has been analyzed in terms of accuracy as 87.65% and ROC-AUC as 0.9446. In Naz et al. (2024), the authors implemented DL and FL framework using ResNet50, InceptionV3, and VGG16 models for predicting the presence of Covid-19, lung opacity, pneumonia, and normal in X-ray scans containing (21165) images. This model has been trained with high-class imbalance containing (3616) images for Covid-19, (10192) images for normal lungs, (6012) images for lung opacity, and (1345) images for pneumonia. This model has achieved the highest accuracy with FL-framework as 95% for IID and 91% for non-IID data.

In the existing works, authors have implemented various FL-based pneumonia detection models, however, these models resulted in a lowered convergence rate and higher time complexity, whereas the proposed model implemented using FL-based EfficientNetB3 has improved performance and better convergence rate. The existing models have not been fine-tuned with the varying proximal values, which may result in a higher difference between the client's and server's loss function. In the proposed work, the model has been fine-tuned with the varied values of proximal terms, resulting in higher stability in the performance over the increased number of rounds. The proposed model improves the accuracy of both IID and non-IID distributions, where the proximal term as 0.05 achieves an accuracy of 99.32% on IID data and 96.14% on non-IID data and the proximal term configured to 0.5, resulting in the accuracy levels of 92.05% and 96.98% in IID data and non-IID data distributions, respectively. Whereas the accuracy achieved by Alfiansyah et al. (2024), Adjei-Mensah et al. (2024) and Kareem et al. (2023) have resulted in 93%, 91.9%,

and 87.65%, respectively. The accuracy comparison depicts that the proposed model outperforms other existing models by achieving higher accuracy in both IID and non-IID distributions.

Table 1 provides a comprehensive review of the existing pneumonia detection FL models.

Table 1. A comprehensive review of the existing pneumonia detection federated learning models.

References	Models	Dataset details	Parameters	Research gaps
Kandati and Gadekallu (2023)	PSO with FL	Dataset 1: 5,856 images Dataset 2: 317 images	Epochs: 5 (Train), 30 (Test) Batch size: 32	The limited dataset has been used by the authors, which lead to overfitting challenges.
Mabrouk et al. (2023)	FL-DenseNet169, FL-MobileNetV2, FL-Xception, FL-InceptionV3, FL-VGG16, FL-DenseNet121, FL-ResNet152v2, FL-ResNet50	5,856 images	Rounds: 7, Batch size: 16, Epochs: 20, Learning rate: 0.00001, Optimizer: Adam	The authors configured a small value of learning rate and batch size, which may lead to high computational costs.
Kareem et al. (2023)	ResNet50, VGG19, AlexNet, DenseNet	5,856 images	Epochs: 20, Clients: 4	Limited information on the hyperparameter configuration has been provided by the authors.
Malik et al. (2023)	VGG16, DenseNet169, VGG19	101017 images of different lung diseases	Rounds: 100, Epochs: 100, Clients: 3	The model may result in high complexity due to a large value of rounds and epochs.
Riedel et al. (2023)	FL-ResNet50	1411 images	Rounds: 25, Optimizer: SGD, Aggregation: FedAVG, Batch size: 6	The authors have utilized a limited number of images for training the FL model.
Alfiansyah et al. (2024)	DenseNet	Mendeley Dataset: 5,856 images, RSNA Dataset: 14,863 images	Aggregation: FedAVG, Clients: 3	The limited details on the hyperparameters have been provided.
Farkaş et al. (2023)	VGG16	5863 images	Rounds: 25, Optimizer: SGD	The image class balance is present in the dataset, leading to overfitting challenges.
Malik and Anees (2024)	VGG19, DenseNet201, DenseNet 169, and InceptionV3	101017 images of chest diseases	Rounds: 150, Clients utilization: 100%, 50%, and 25%	The dataset has not been balanced.
Adjei-Mensah et al. (2024)	Cov-Fed	Chest X-ray dataset containing SARS, MERS, covid-19, and pneumonia lungs	Clients utilization: 100%	The authors provided limited information on the hyperparameter tuning to the proposed Cov-Fed model
Naz et al. (2024)	ResNet50, InceptionV3, and VGG16	21165 chest X-ray images	Clients: 10 and 20, Rounds: 10, Epochs: 20, 50, and 70, Batch size: 8, Optimizer: AdaDelta, Learning rate: 0.005	This model has a slow convergence rate due to a smaller configured value of batch size. It may lead to increasing computational overhead.

3. Material and Methods

This section elaborates on the dataset curated for training and the proposed methodology used for pneumonia detection.

3.1 Dataset

The chest X-ray dataset has been curated from Kaggle containing 5,856 images of pneumonia and normal lungs (Mooney, 2018). This dataset contained some distorted images which have been removed from the dataset remaining with 4800 CXRs. The images were of different sizes which have been resized to 300X300 pixels for training the EfficientNetB3 model. The sample images present in the dataset are provided in **Figure 1**.

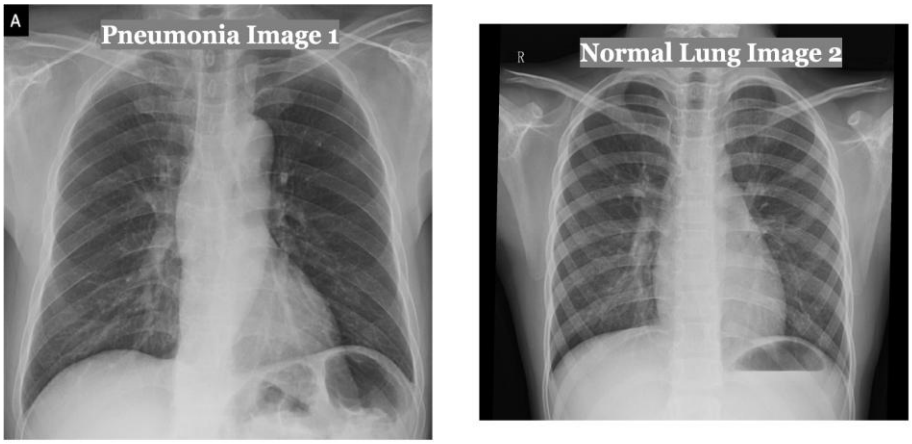

Figure 1. Sample CXRs in the dataset (Mooney, 2018).

3.2 Proposed Methodology

FL is a revolutionary approach to DL that enables the training of shared models on decentralized data from different sources, without compromising data privacy. In traditional DL models, data is typically centralized and processed on a single server, which raises data ownership, privacy, and security concerns. FL solves these challenges by allowing data to remain distributed among various locations, such as hospitals, mobile devices, and other organizations taking participation in the framework, and only sharing model updated parameters with a central server as shown in **Figure 2**. This model not only protects private information but also leverages the collective knowledge from diverse data sources, resulting in more improved, accurate and robust model development. FL has the potential to perform in an improved and accurate way in various applications, including healthcare, edge computing, and finance which transforms the learning and training procedure of the traditional DL models with decentralized and privacy-preserved frameworks. The architecture of the proposed FL model has been provided in **Figure 3**.

***Federated Averaging* (*FedAVG*)**

FedAVG is an aggregation technique used in FL for model update aggregation. It enables the model to improve performance based on the updated values. This technique periodically aggregates the updates computed by the local clients and further improves the global model on the server using aggregated averaged updates. This process preserves data by neither sharing it with the server nor with the peer clients. It was introduced to improve model performance while preserving data privacy (Konečný et al., 2016). It allows various participating devices to collaboratively train themselves in a shared environment without

sharing the data either with clients or servers. This procedure iteratively averages the updates and performs model upgradation leading to a single, unified model generation with optimal outcomes from distributed data sources. The FedAVG technique is a basic aggregation technique that was initially designed by Google to introduce security in the distributed environment. This aggregation is efficient, scalable, and flexible; however, it works optimally with IID distribution in comparison to non-IID distribution. The FedAVG aggregation is shown in Algorithm 1.

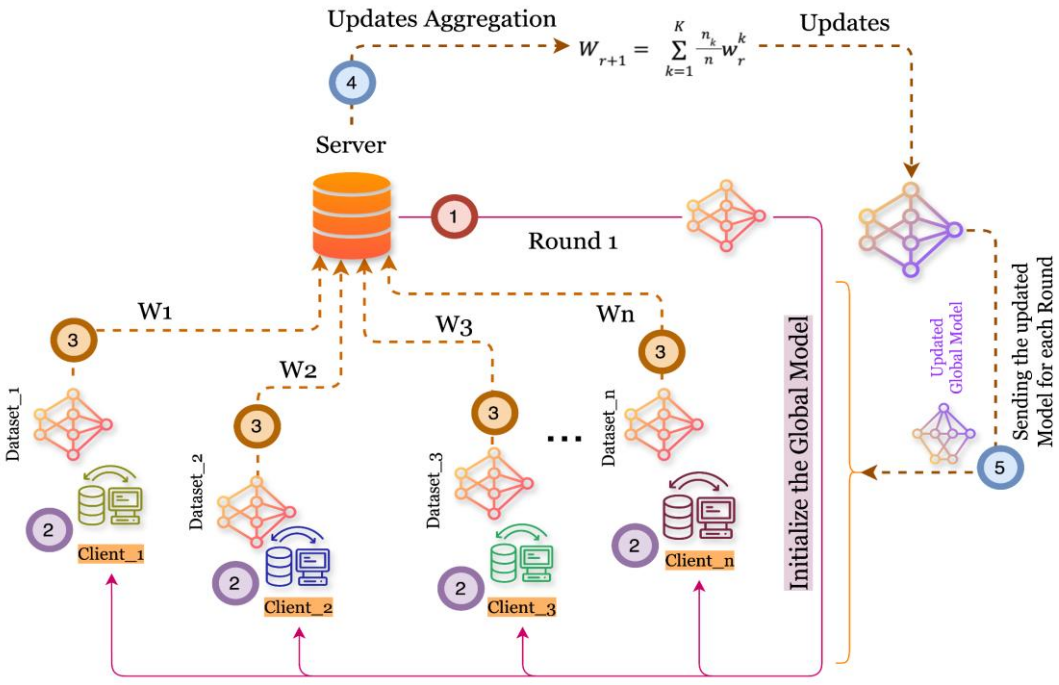

Figure 2. A schematic architecture of the federated learning model.

Algorithm 1:
Step 1: for each epoch "e" = 1, 2, 3, … , E do
Step 2: for each client k = 1, 2, 3, …, K do
Step 3: Sampling a batch of data B_k from client k dataset
Step 4: Computing model updates Δw_k using B_k and local model w_k
Step 5: Update local model $w_k = w_k - \eta \Delta w_k$
Step 6: end for
Step 7: Computing global model update using $\Delta w = \frac{1}{K} \sum_{k=1}^{K} \Delta w_k$
Step 8: Updating global model using $w = w - \eta \Delta w$
Step 9. end for

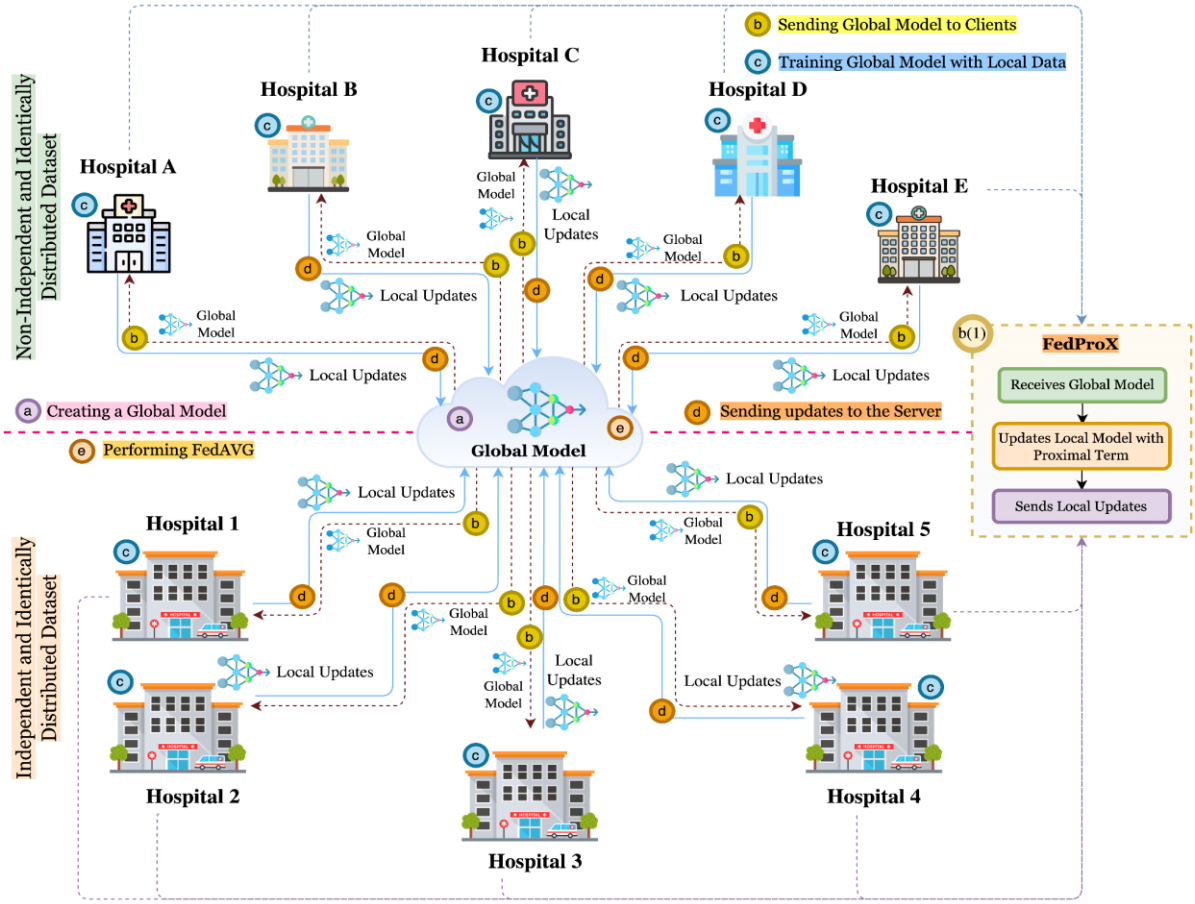

Figure 3. Architecture of the proposed federated learning model.

Federated Proximal (*FedProx*)

It is an aggregation technique of FL architecture that is built on the basis of the FedAVG aggregation. This technique has been specially designed for real-life applications, where the data at different locations are spread in non-IID distributions and also generated with varying computational resources. The FedProx has introduced a term called "proximal term" for the local training that improves the convergence along with the stability of the FL model. It is also proposed by Google in the year (2020) to address the challenges of FedAVG aggregation including sensitivity to device heterogeneity and non-IID data. In practical FL scenarios, the devices having different distributions of the data and computational powers may result in inconsistent and inappropriate updates leading to diverged and slow performance at the global model. The FedProx adds a proximal term for improving the local training function, resulting in the local model updates to get closer to the global model. The updation in the FedProx aggregation leads to an improvement in stability, and convergence, enhancing the robustness in handling the outliers, and data heterogeneity challenges. The procedure of FedProx has been discussed in Algorithm 2.

The EfficientNetB3 model has been utilized as the initial model for training clients with their local dataset. This pre-trained model has a series of existing models with versions (B0), (B1), and (B2). It contains (3) depth wise multipliers which results in an efficient and powerful tool for image classification. The series of

EfficientNet was introduced by Google in 2020 with the aim of developing the models best suited for image classification tasks with improved efficiency (Sharma and Guleria, 2023a). This model works on the AutoML framework that uses neural network architecture to automatically identify the best-performing model. In the proposed work, the FL framework has been developed using a pre-trained EfficientNetB3 model. To enhance the convergence rate and performance of the proposed FL-EfficientNetB3 model, the values of the hyperparameters have been thoroughly configured. The participating clients in the FL-framework have been configured to (5) with IID and non-ID distributions. The local epochs were set to (5), whereas, learning rate and batch size of (0.001) and (32) have been used, respectively. Further, the FedProx aggregation has been utilized for updates aggregation, using the values of the proximal term (μ) as (0.5) and (0.05). A detailed information on the hyperparameters configuration in the proposed model has been provided in **Table 2**.

Algorithm 2:
Step 1: for each epoch "e" = 1, 2, 3, … , E do
Step 2: for each client k = 1, 2, 3, …, K do
Step 3: Sampling a batch of data B_k from client k dataset
Step 4: Computing model updates Δw_k using B_k and local model w_k
Step 5: Computing proximal term as $P_k = \mu (w_k - w)$
Step 6: Update local model $w_k = w_k - \eta(\Delta w_k + P_k)$
Step 7: end for
Step 8: computing global model update using $\Delta w = \frac{1}{K}\sum_{k=1}^{K} \Delta w_k$
Step 9: updating global model using $w = w - \eta \Delta w$
Step 10. end for

Table 2. Hyperparameter configuration of the proposed FL framework with EfficientNetB3 model.

Parameters	Value	Parameters	Value
Client	5	Rounds	15
Learning Rate	0.001	Aggregation Technique	FedProx
Batch Size	32	Local Epochs	5
Proximal Term (μ)	0.5 and 0.05	Client Utilization	100%

4. Results and Discussion

This section discusses the performance outcomes of the proposed FL model with IID and non-IID distributions with (μ) values configured to (0.5) and (0.05).

Table 3 depicts the performance of the proposed FL model for pneumonia detection with the proximal term "(μ)" configured to (0.05) and (0.5) in terms of accuracy, loss, recall, precision, and F1-score values.

Table 3. Performance outcomes of the proposed FL model for pneumonia detection with Proximal term "μ" configured to 0.05 and 0.5.

Accuracy (%)	Precision	Recall	F1-score	Lowest Loss
Proximal Term "μ = 0.05" with IID Distribution				
99.32%	0.9613	0.9614	0.9613	0.1494
Proximal Term "μ = 0.05" with non-IID Distribution				
96.14%	0.9613	0.9614	0.9613	0.1494
Proximal Term "μ = 0.5" with IID Distribution				
92.05%	0.9237	0.9205	0.9222	0.2240
Proximal Term "μ = 0.5" with non-IID Distribution				
96.98%	0.957	0.9564	0.961	0.0545

Figure 4 depicts the accuracy of the proposed FL model with 5 clients in IID data when the value of the proximal term "μ" has been set to 0.5 and 0.05. The highest accuracy of μ= 0.05 has been achieved at round 9 as 99.32%, whereas, for μ= 0.5, it has been resulted at its peak at round 10 as 92.05%.

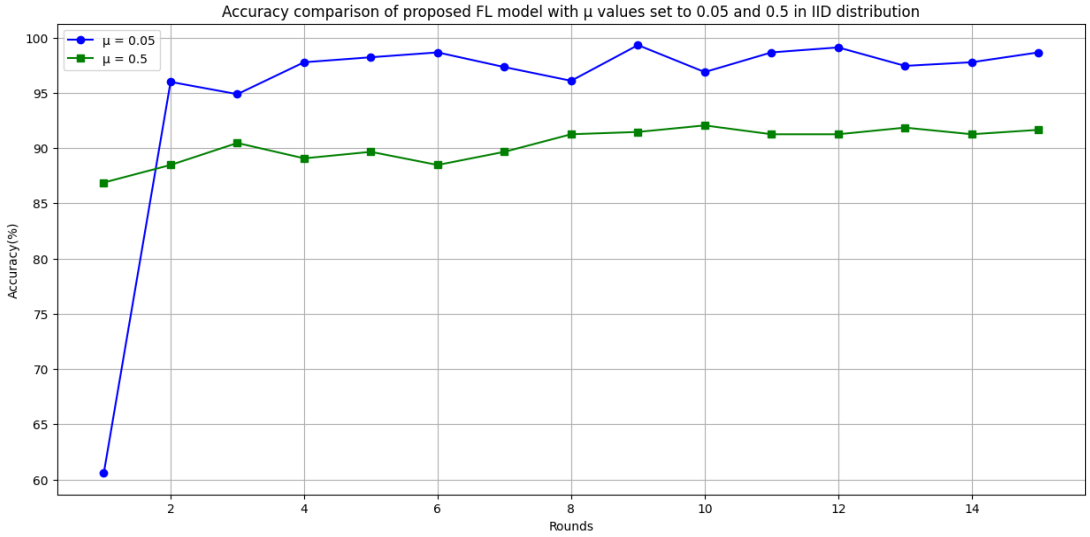

Figure 4. Accuracy of the proposed FL model with μ values set to 0.05 and 0.5 in IID distribution.

Figure 5 shows the recall of the proposed FL model with 5 clients in IID data when the value of the proximal term "μ" has been set to 0.5 and 0.05. The highest recall of μ= 0.05 has been achieved at round 9 as 0.9614, whereas, for μ= 0.5, it has resulted at its peak at round 10 as 0.9205.

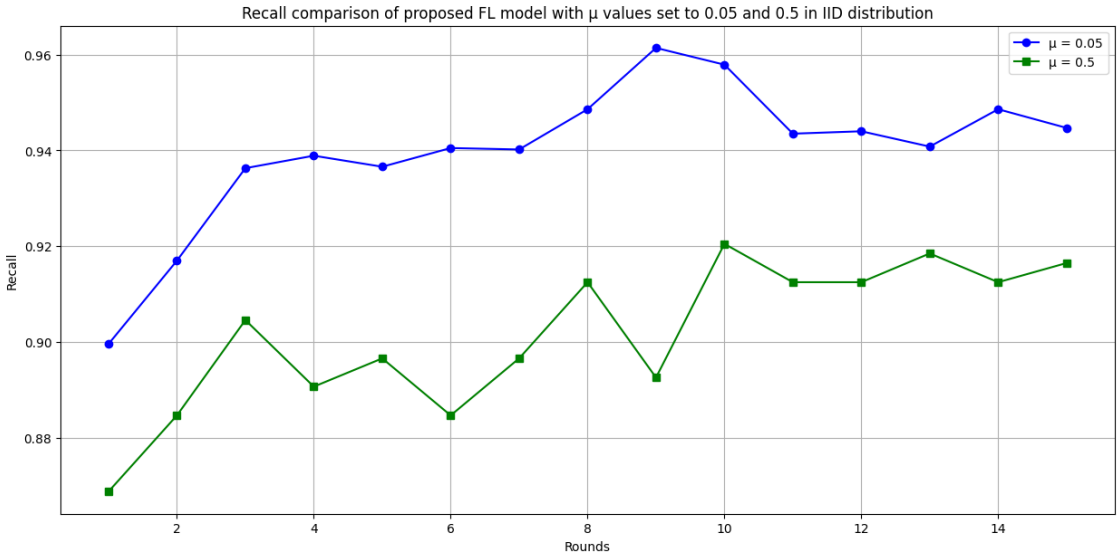

Figure 5. Recall of the proposed FL model with μ values set to 0.05 and 0.5 in the IID distribution.

Figure 6 represents the precision of the proposed pneumonia detection model with 5 clients in IID data where the proximal term "μ" has been set to 0.5 and 0.05. The highest precision of μ= 0.05 has been achieved at round 9 as 0.9613, whereas, for μ= 0.5, it has been resulted at its peak at round 10 as 0.9237.

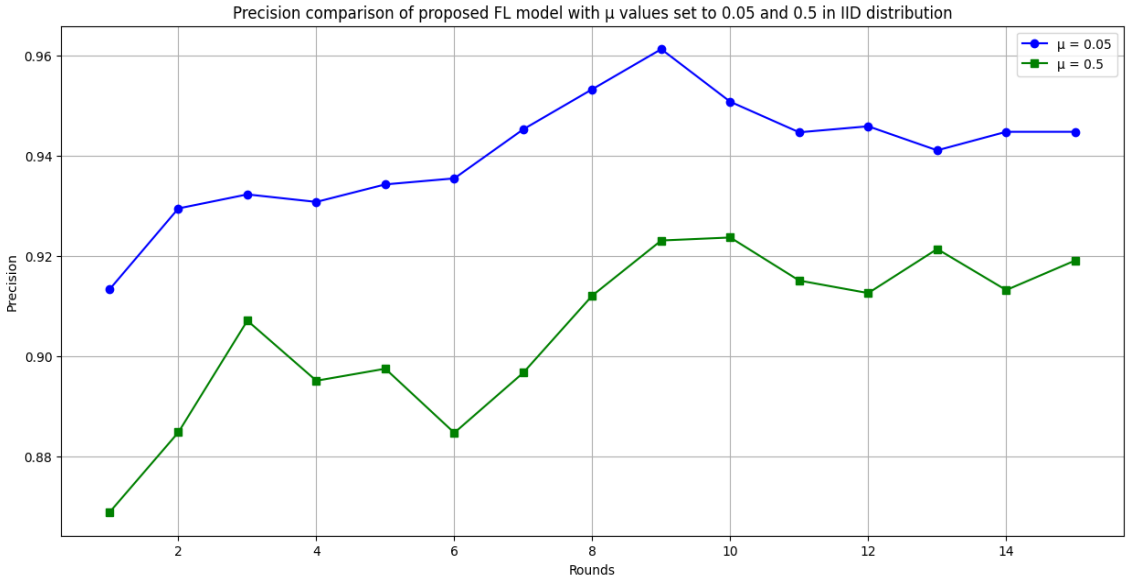

Figure 6. Precision of the proposed FL model with μ values set to 0.05 and 0.5 in the IID distribution.

Figure 7 shows the F1-score of the proposed model with 5 clients in IID data where the proximal term "μ" has been set to 0.5 and 0.05. The highest F1-score of μ= 0.05 has been achieved at round 9 as 0.9613, whereas, for μ= 0.5, it has been resulted at its peak at round 10 as 0.9222.

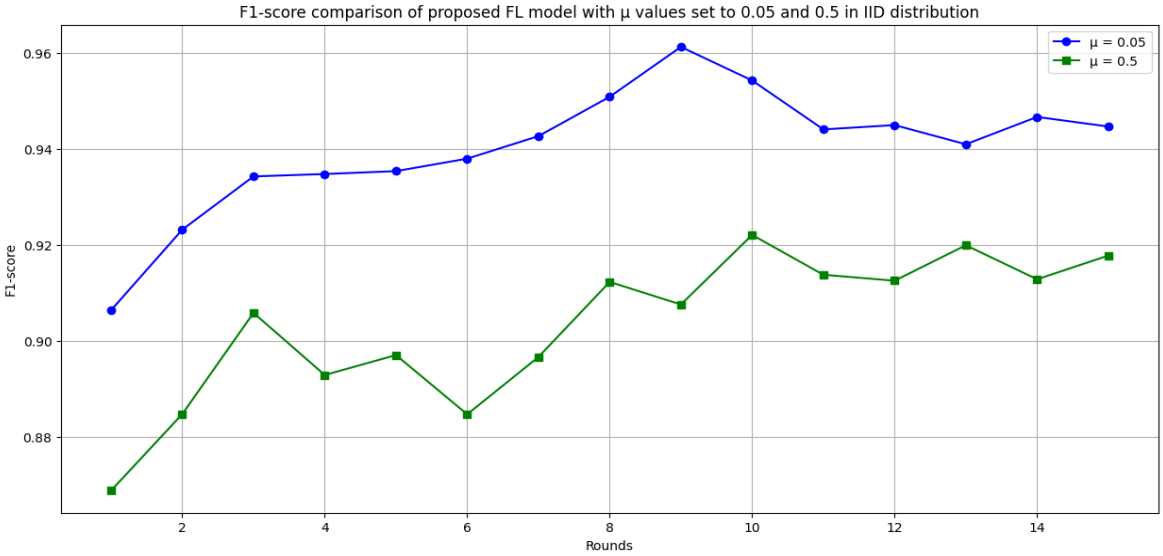

Figure 7. F1-score of the proposed FL model with μ values set to 0.05 and 0.5 in the IID distribution.

Figure 8 shows the loss of the proposed model with 5 clients in IID data where the proximal term "μ" has been set to 0.5 and 0.05. The minimum loss of μ= 0.05 has been resulted at round 9 as 0.1494, whereas, for μ= 0.5, it has been resulted at its minimum value at round 10 as 0.2240.

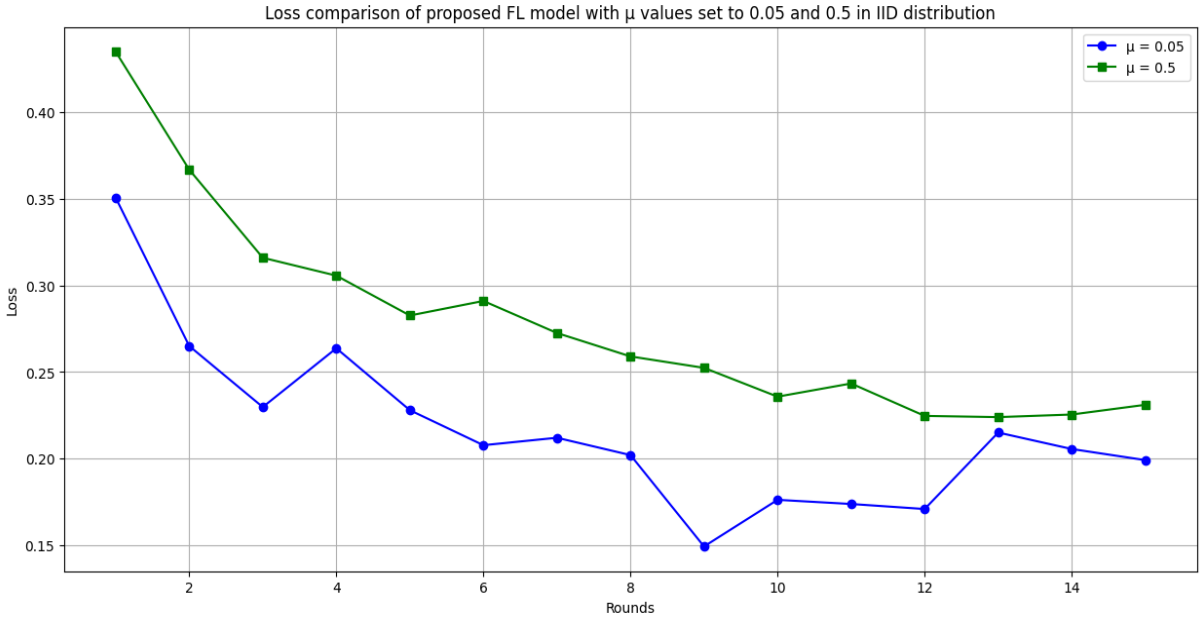

Figure 8. Loss of the proposed FL model with μ values set to 0.05 and 0.5 in IID distribution.

Figure 9 represents the accuracy of the proposed model with 5 clients in non-IID data where the proximal term "μ" has been set to 0.5 and 0.05. The highest accuracy of μ= 0.05 has been identified at round 12 as 96.14%, whereas, for μ= 0.5, it has been resulted at its peak at round 10 as 96.98%.

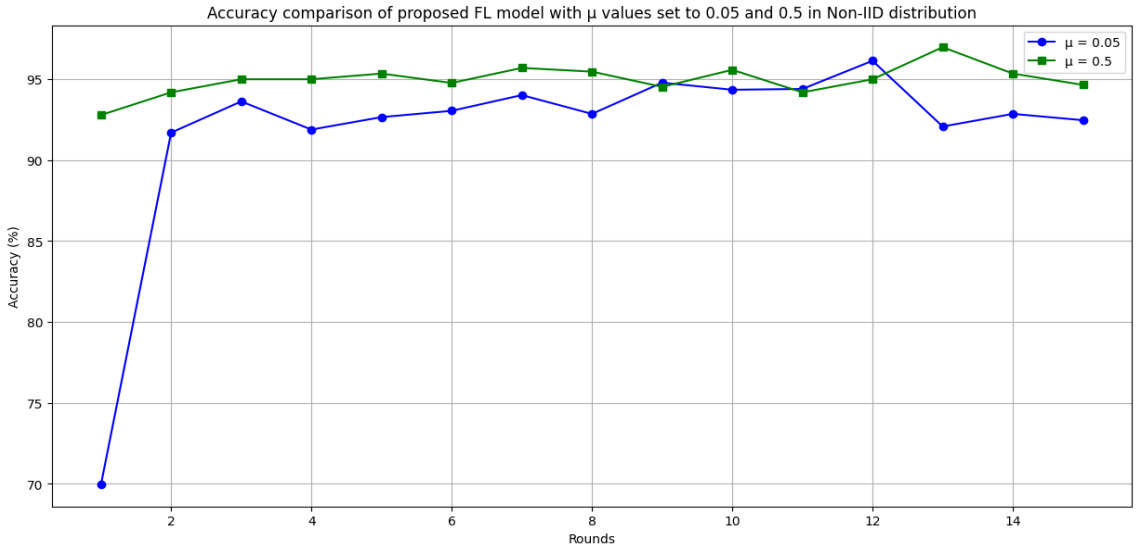

Figure 9. Accuracy of the proposed FL model with μ values set to 0.05 and 0.5 in non-IID distribution.

Figure 10 depicts the recall of the proposed model with 5 clients in non-IID data where the proximal term "μ" is configured to 0.5 and 0.05. The highest precision with μ= 0.05 and 0.5 have been identified at rounds 12 and 15 as 0.9614 and 0.9564, respectively.

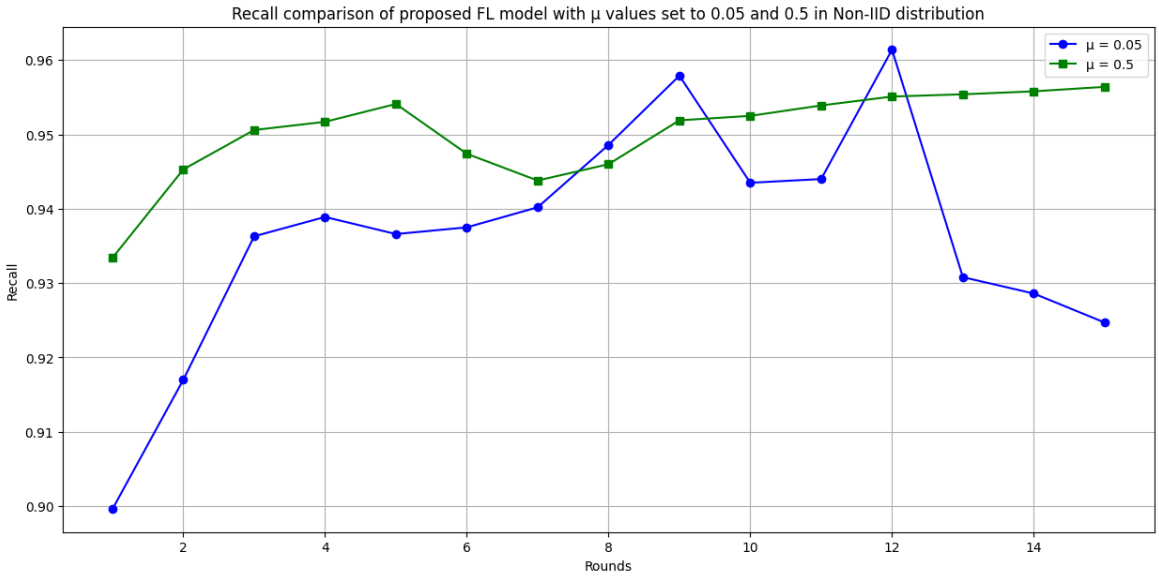

Figure 10. Recall of the proposed FL model with μ values set to 0.05 and 0.5 in non-IID distribution.

Figure 11 shows the precision of the proposed model with 5 clients in non-IID data where the proximal term "μ" is configured to 0.5 and 0.05. The highest precision with μ= 0.05 and 0.5 have been identified at rounds 12 and 15 as 0.9613 and 0.957, respectively.

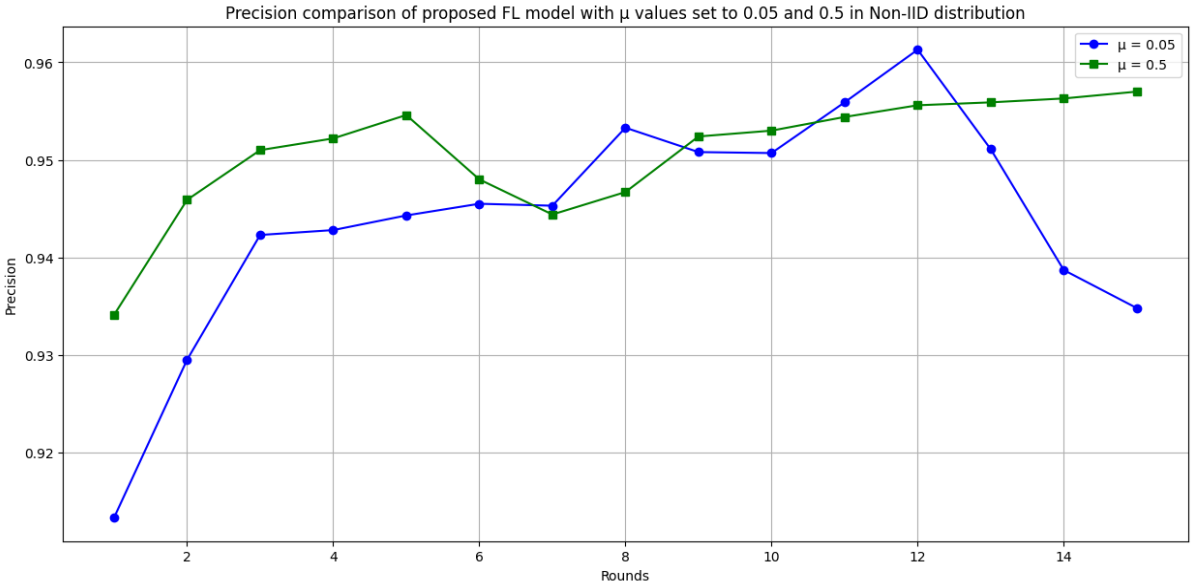

Figure 11. Precision of the proposed FL model with μ values set to 0.05 and 0.5 in non-IID distribution.

Figure 12 shows the F1-score of the proposed model with 5 clients in non-IID data where the proximal term "μ" is configured to 0.5 and 0.05. The highest value of the F1-score with μ = 0.05 and 0.5 resulted in 0.9613 and 0.961 at rounds 12 and 15, respectively.

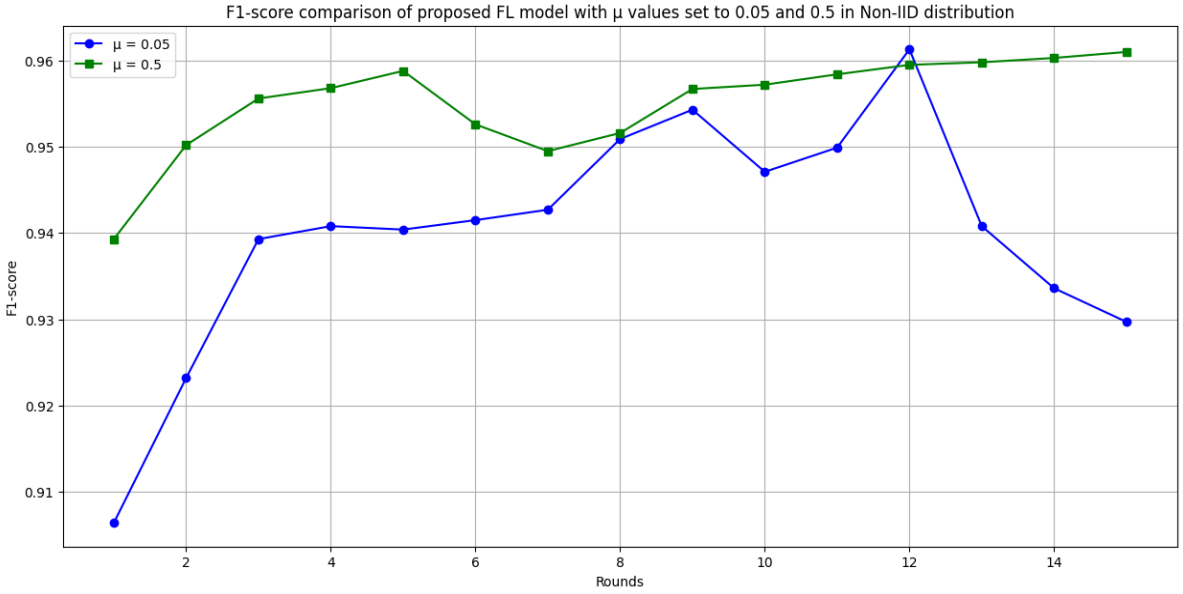

Figure 12. F1-score of the proposed FL model with μ values set to 0.05 and 0.5 in non-IID distribution.

Figure 13 shows the loss identified by the proposed model with 5 clients in non-IID data where the proximal term "μ" configured to 0.5 and 0.05. The minimum value of loss with μ = 0.05 and 0.5 have been resulted as 0.1494 and 0.0545 at rounds 12 and 15, respectively.

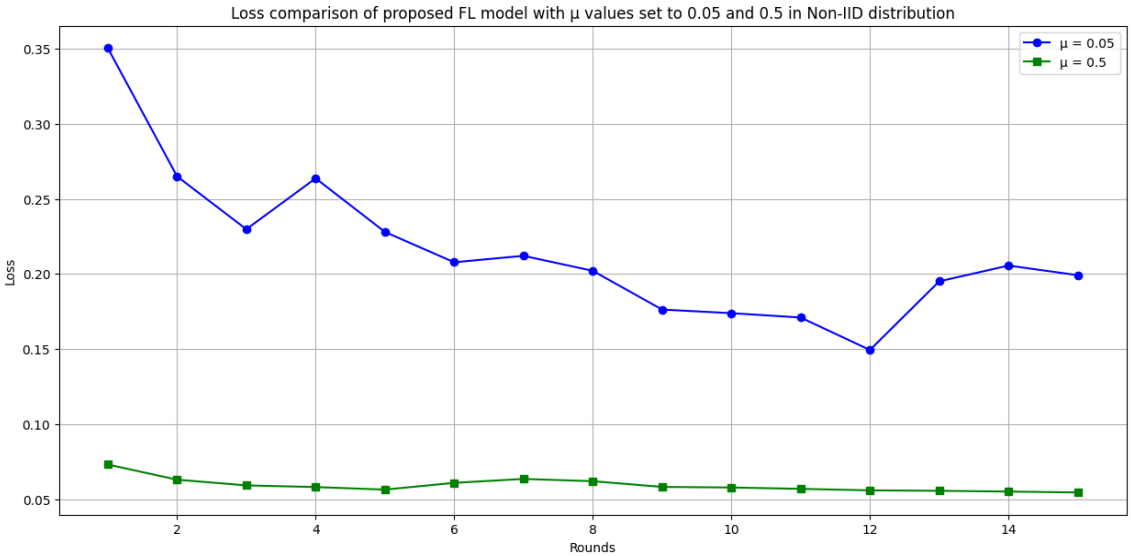

Figure 13. Loss of the proposed FL model with μ values set to 0.05 and 0.5 in non-IID distribution.

The results of the proposed model depict that the pneumonia detection performance has been resulted as the best with the configuration of the proximal term μ = 0.05 in IID distribution as 99.32%, however, the loss value has not been identified as the best. The minimum loss value of 0.0545 has been resulted by non-IID distribution with proximal term μ as 0.5 and the accuracy was resulted at its peak as 96.98%. Further, it has been also analyzed that the choice of μ value and data distribution (IID vs. non-IID) has a significant impact on the performance outcomes of the model. The minimum value of proximal term μ yields higher accuracy, while higher μ values under non-IID distributions optimize the loss better. In addition, the performance of the model is relatively stable across different distributions (IID vs. non-IID) for the same μ values; however, a higher μ value under non-IID distribution provides the best trade-off between loss optimization and performance metrics.

4.1 Performance Comparison of Various Clients in IID and Non-IID Distributions with μ Values Set to 0.05 and 0.5

This section provides the results obtained by the proposed FL model with respect to each client in IID and non-IID distribution with μ values set to 0.05 and 0.5.

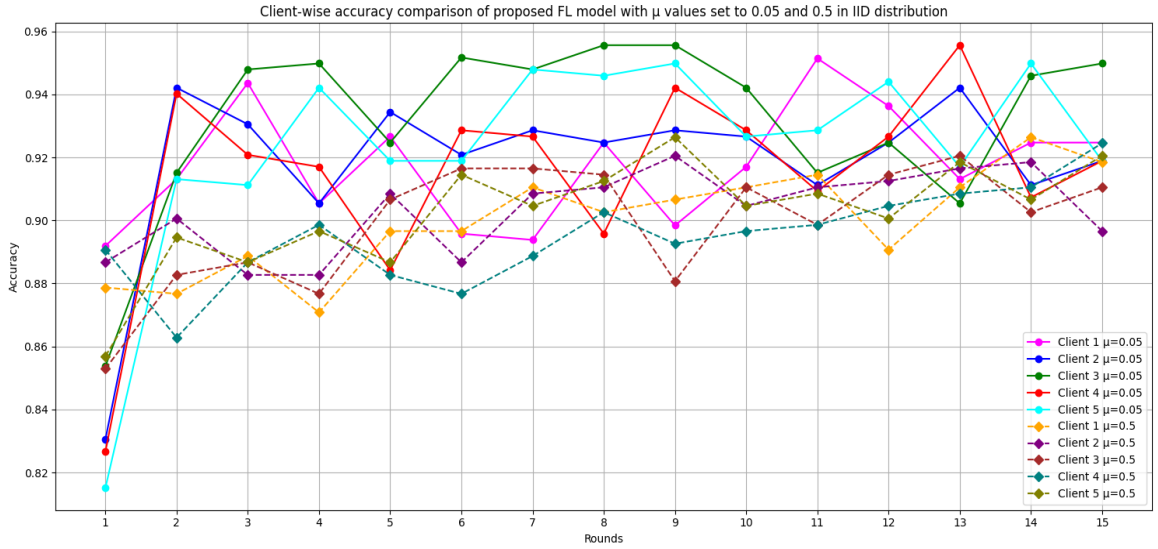

Figure 14. Accuracy comparison of various clients in IID distribution with μ values set to 0.05 and 0.5.

Figure 14 compares the client-wise accuracy obtained by μ values set to 0.05 and 0.5. This comparison shows that the μ values of 0.5 has obtained less accuracy in comparison to μ values configured to 0.05.

Figure 15 shows the client-wise precision obtained by μ values set to 0.05 and 0.5. This graph depicts that the highest values of precision have been obtained when the value of μ is configured to 0.05.

Figure 16 represents the client-wise recall obtained by μ values set to 0.05 and 0.5. This comparison shows that the highest values of recall are achieved with the μ value as 0.05, whereas, the μ value set to 0.5 results in less recall.

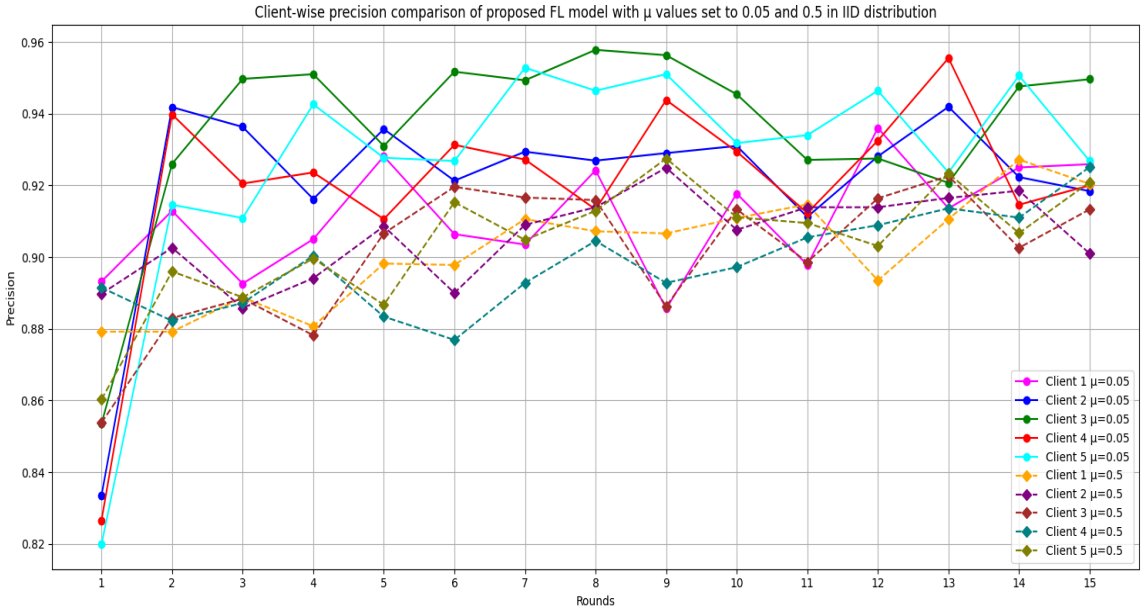

Figure 15. Precision comparison of various clients in IID distribution with μ values set to 0.05 and 0.5.

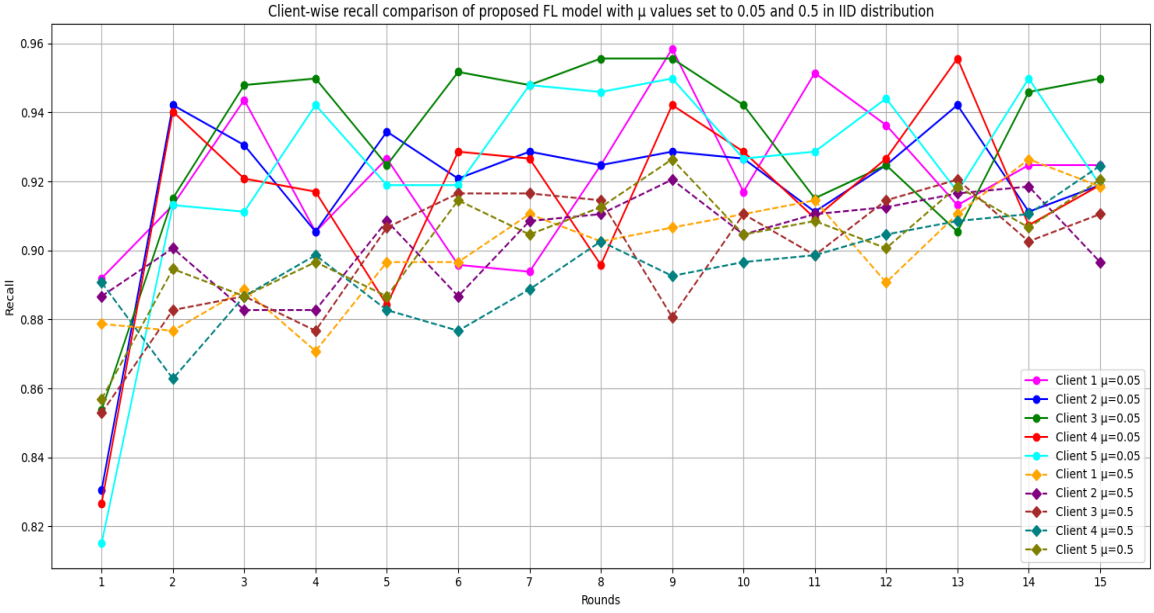

Figure 16. Recall comparison of various clients in IID distribution with μ values set to 0.05 and 0.5.

Figure 17 depicts the client-wise F1-score of the proposed model obtained with μ values set to 0.05 and 0.5. The highest F1-score values have been achieved with the μ= 0.05, whereas, μ=0.5 underperforms in terms of achieving the best values of F1-score.

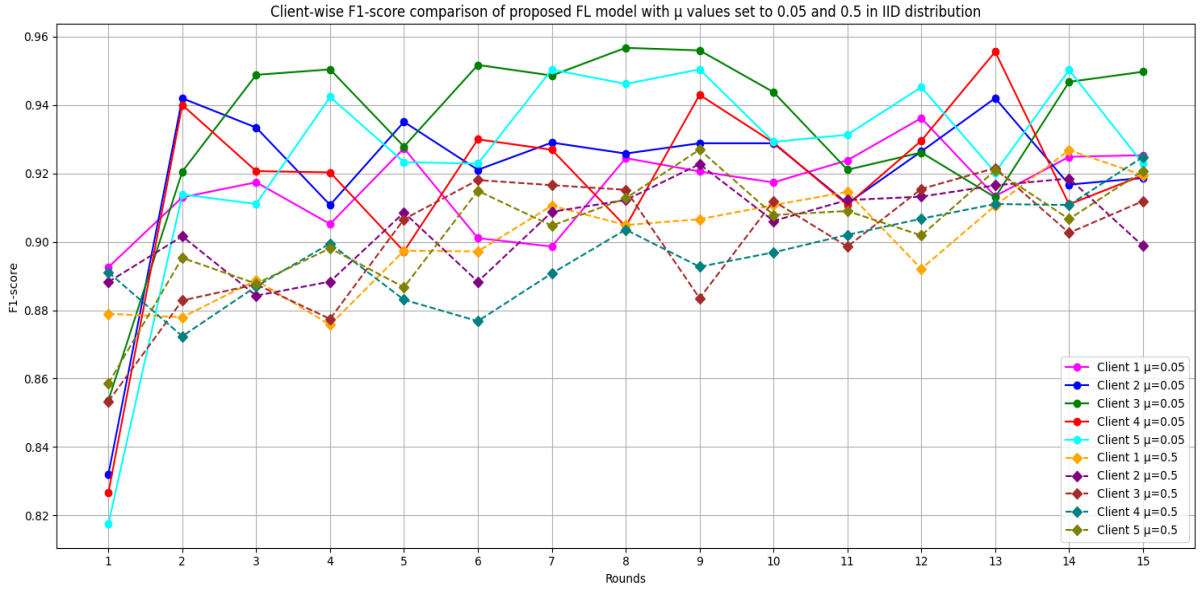

Figure 17. F1-score comparison of various clients in IID distribution with μ values set to 0.05 and 0.5.

Figure 18 illustrates the client-wise accuracy comparison of the proposed model obtained with μ values set to 0.05 and 0.5 in non-IID distributions. The highest accuracy values are obtained by μ= 0.05, whereas, μ=0.5 underperforms in terms of achieving the best values of the accuracy.

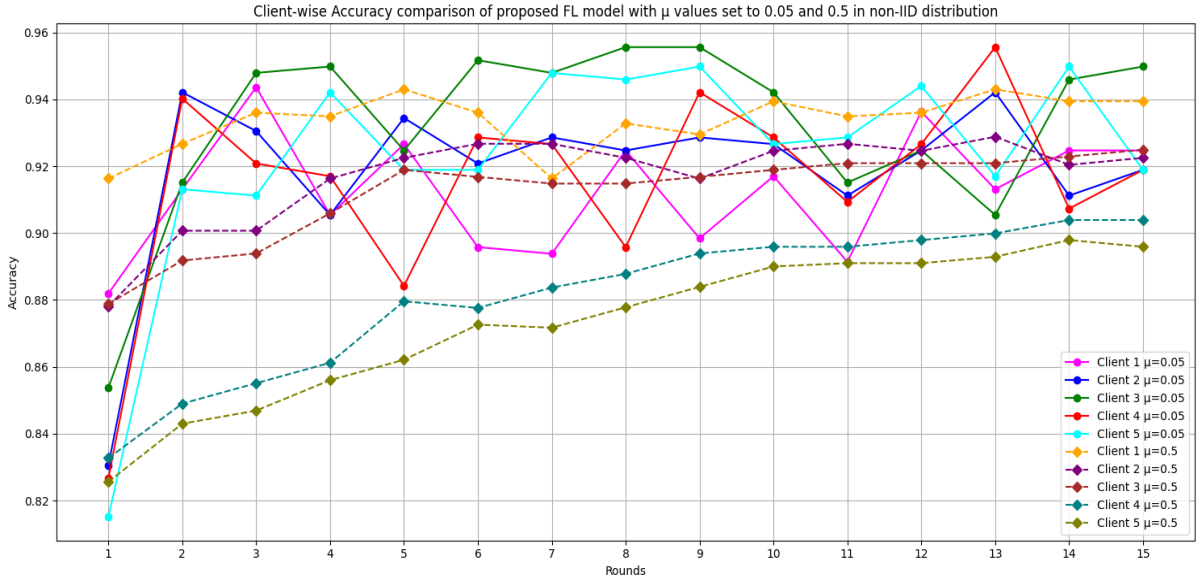

Figure 18. Accuracy comparison of various clients in non-IID distribution with μ values set to 0.05 and 0.5.

Figure 19 shows the client-wise precision comparison of the proposed FL model obtained with μ values set to 0.05 and 0.5 in non-IID distributions. The graph depicts that the precision value has been achieved highest with μ = 0.5 in comparison to μ = 0.05.

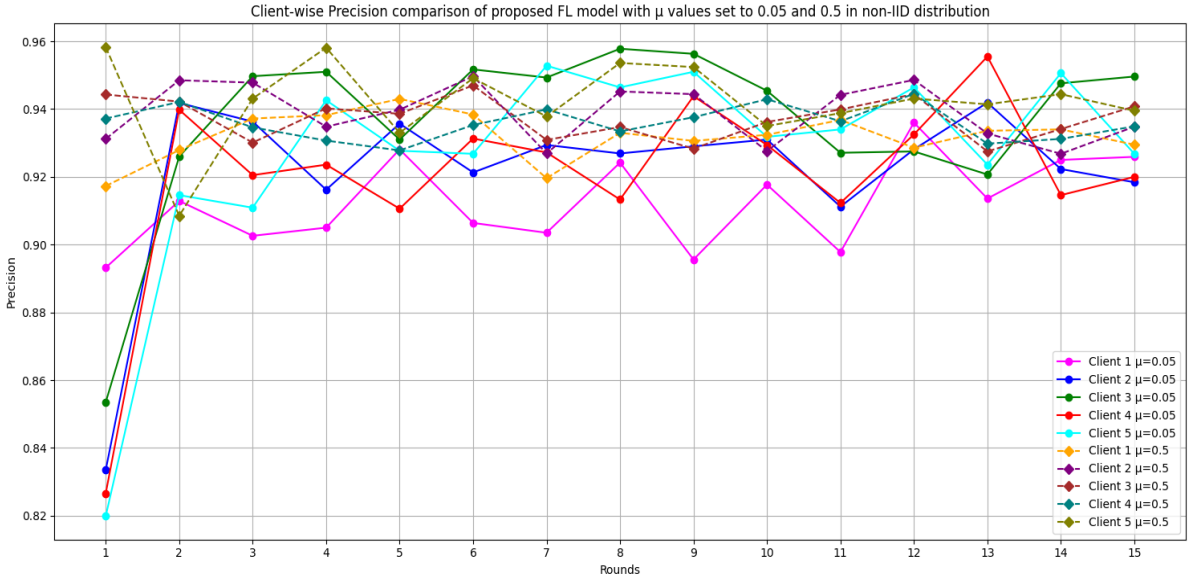

Figure 19. Precision comparison of various clients in non-IID distribution with μ values set to 0.05 and 0.5.

Figure 20 depicts the client-wise recall comparison of the proposed FL model obtained with μ values set to 0.05 and 0.5 in non-IID distributions. The highest value of the recall has been obtained by client 3 with μ = 0.5 in comparison to μ = 0.05.

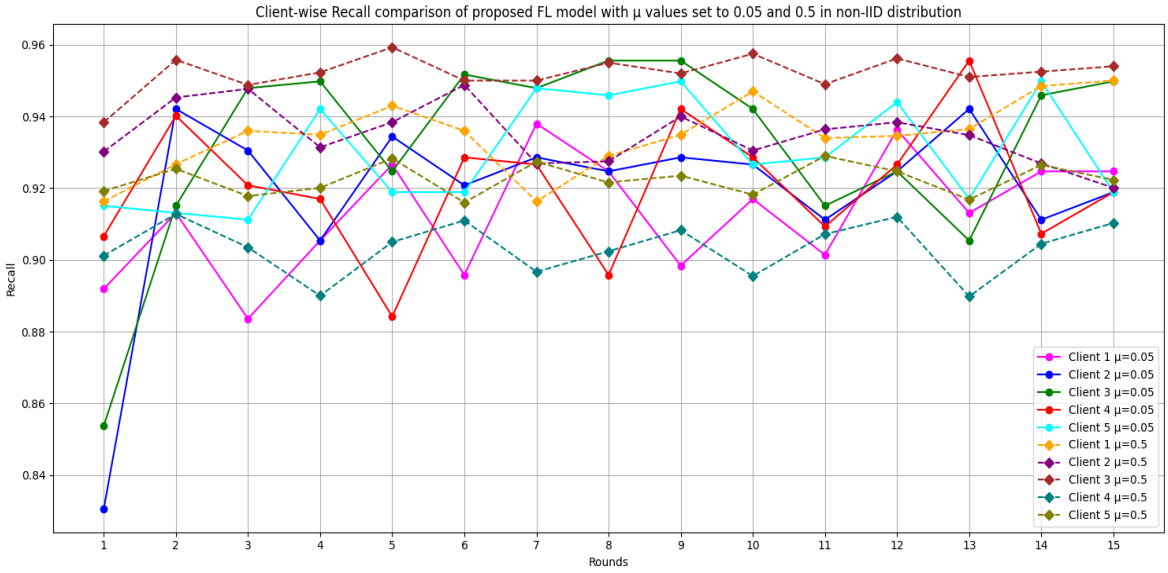

Figure 20. Recall comparison of various clients in non-IID distribution with μ values set to 0.05 and 0.5.

Figure 21 depicts the client-wise F1-score comparison of the proposed FL model obtained with μ values set to 0.05 and 0.5 in non-IID distributions. The highest F1-score value has been obtained by client 3 with μ = 0.05 in comparison to μ = 0.5.

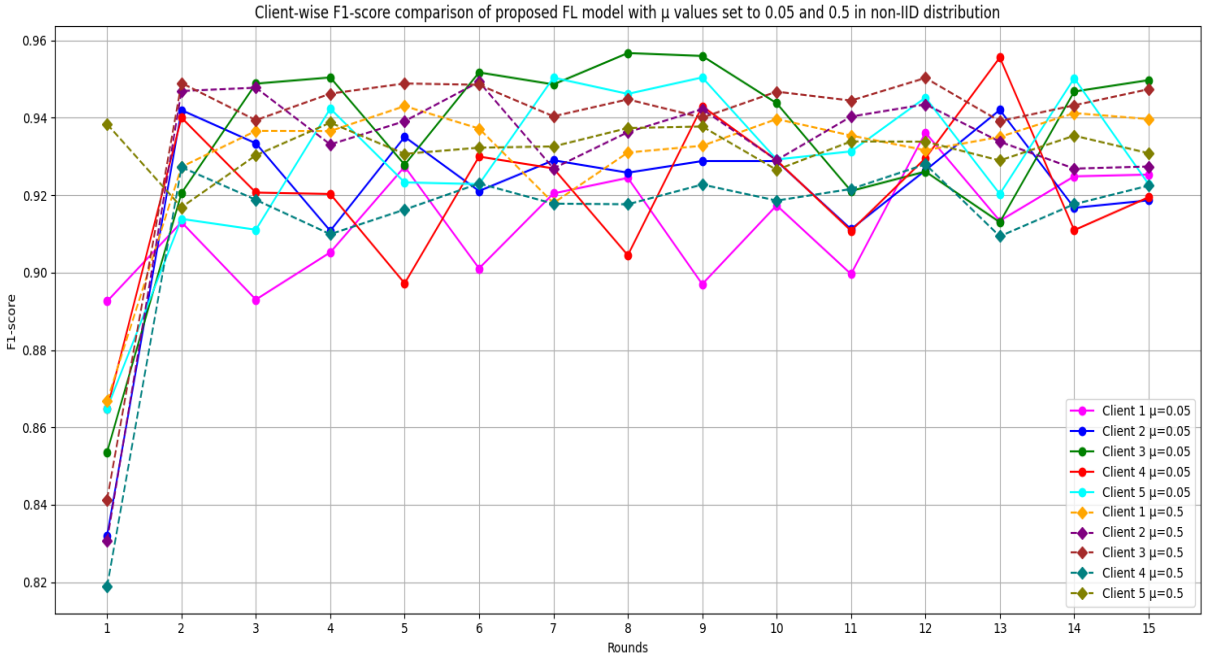

Figure 21. F1-score comparison of various clients in non-IID distribution with μ values set to 0.05 and 0.5.

The client-wise comparative results obtained by the proposed model show that the highest performance has been achieved with the IID distributions in μ value set to 0.05, whereas, in comparison to μ as 0.5, the best results have been obtained in non-IID distributions. This shows that the model is highly capable of recognizing the presence of pneumonia in diverse datasets with varying distributions. The model has also been identified as effective and improved over other centralized models due to privacy preservation. The results depict that the model is robust to the inherent heterogeneity of the data, which leads to the model being optimal and easily adapting non-IID distributions in real-time healthcare sectors. It has also been observed that the proposed FL-EfficientNetB3 model is highly robust due to efficient design, enhanced adaptability, and balanced scaling.

4.2 Accuracy Comparison of the Proposed FL Model with Existing Models

This section compares the proposed model's accuracy with the existing model, the model FL-ResNet50 presented by Kareem et al. (2023) has resulted in an accuracy of 93%, whereas the model proposed by Alfiansyah et al. (2024) has achieved an accuracy of 91.9% and in Farkaş et al. (2023), it has been resulted as 90.47%. The Cov-Fed model proposed by Adjei-Mensah et al. (2024) has also achieved an accuracy of 87.65% for classifying covid-19 from other lung diseases. On the other side, the proposed FL pneumonia detection model has achieved improved accuracy for both IID and non-IID distributions as 99.32% and 96.14% in proximal term configuration as 0.05 and 92.05% and 96.98% for proximal term configuration as 0.5, respectively. These outcomes indicate that the proposed model is highly accurate in terms of pneumonia classification, with both IID and non-IID distributions, as provided in **Figure 22**.

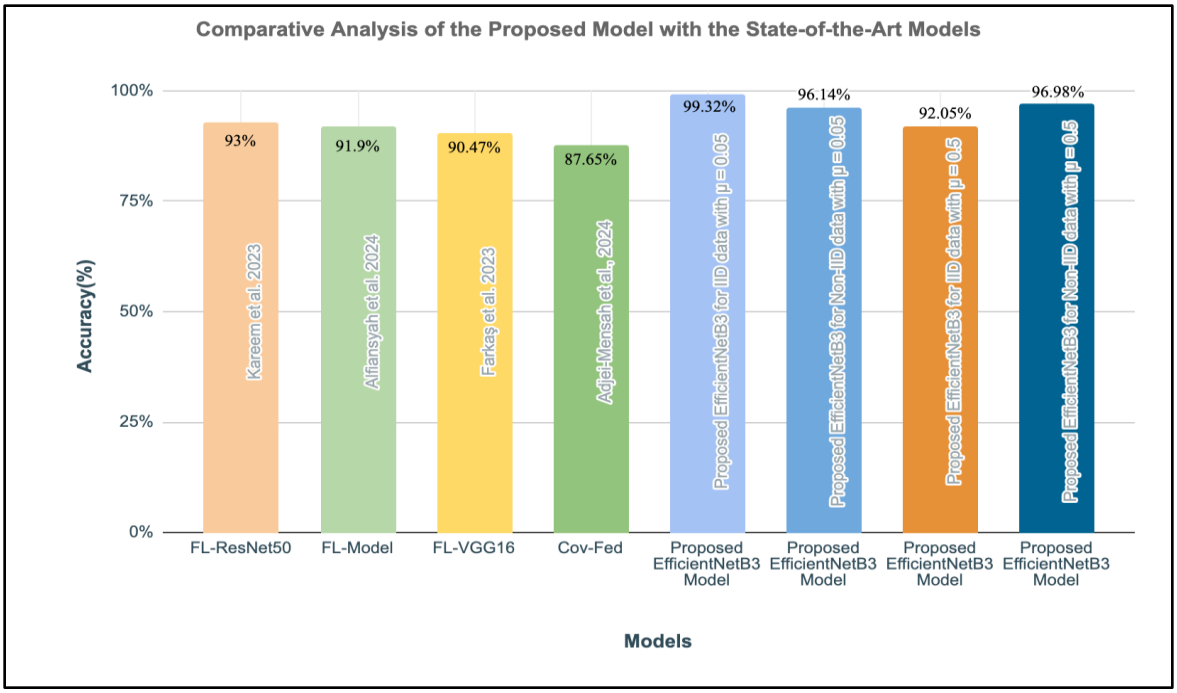

Figure 22. Accuracy comparison of the proposed FL model with existing techniques.

In this work, the FL-EfficientNetB3 model outperforms in comparison to other state-of-the-art models namely, FL-VGG16, Cov-Fed, and FL-ResNet50. The reason behind the better performance of the proposed FL-EfficeintNetB3 model is the efficient design, low complexity, enhanced adaptability, and balanced scaling. Further, the proposed model also leads to fast convergence, resulting in better performance in IID as well as non-IID distributions of the data. In the existing models, it has been observed that the FedAVG technique was applied for both IID and non-IID distributions to aggregate the local model updates. The FedAVG has been identified as the better-performing technique with the IID distribution, and it leads to a lower convergence rate if applied to non-IID distributions. Therefore, the local model update aggregation in the proposed work has been performed using the FedProx technique to deal with the diverse non-IID distribution spread over numerous clients, resulting in improved accuracy.

5. Conclusion and Future Scope

The traditional pneumonia detection centralized deep learning frameworks pose significant challenges, including data consolidation at the central location, data privacy concerns, and regulatory issues. The DL models require the collection of a vast volume of data to provide precise predictions. However, healthcare institutes have various data ownership concerns, which pose significant challenges in data collection. Therefore, in the proposed work, a 5-client decentralized federated learning architecture has been proposed for pneumonia detection with IID and non-IID distributions of the data. This model ensures data privacy while enabling decentralized learning in diverse datasets and achieving the most optimal performance outcomes. The proposed model has been evaluated on two different proximal term values configured to 0.05 and 0.5 resulting in accuracies as 99.32% and 92.05% with IID distribution and 96.14% and 96.98% with non-IID distributions, respectively. This decentralized framework develops a robust, accurate, and privacy-preserving model for early pneumonia detection, potentially improving prediction outcomes and transforming healthcare services.

Conflicts of Interest
No conflicts of interest have been associated with this publication.

AI Disclosure
The author(s) declare that no assistance is taken from generative AI to write this article.

Acknowledgments
The authors would like to thank the editors and reviewers for their valuable comments and suggestions. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Adjei-Mensah, I., Zhang, X., Agyemang, I.O., Yussif, S.B., Baffour, A.A., Cobbinah, B.M., Sey, C., Fiasam, L.D., Chikwendu, I.A., & Arhin, J.R. (2024). Cov-Fed: federated learning-based framework for COVID-19 diagnosis using chest X-ray scans. *Engineering Applications of Artificial Intelligence*, *128*(107448), 107448. https://doi.org/10.1016/j.engappai.2023.107448.

Alfiansyah, A., Widiarti, H., Reynaldo, V., Widjaja, W., & Bobelin, L. (2024). A framework for multi institutional collaboration for pneumonia screening by means of federated learning. In *Electronics, Communications and Networks* (pp. 135-142). IOS Press. Macao, China.

Castellano, G., & Vessio, G. (2022). A deep learning approach to cLustering vIsUal artS. *International Journal of Computer Vision*, *130*(11), 2590-2605.

Chouhan, V., Singh, S.K., Khamparia, A., Gupta, D., Tiwari, P., Moreira, C., Damaševičius, R., & de Albuquerque, V.H.C. (2020). A novel transfer learning based approach for pneumonia detection in chest X-ray images. *Applied Sciences*, *10*(2), 559. https://doi.org/10.3390/app10020559.

Farkaș, C.R., Ciobanu, R.I., & Dobre, C. (2023). Pneumonia detection using federated learning. In *2023 24th International Conference on Control Systems and Computer Science* (pp. 160-165). IEEE. Bucharest, Romania.

Goyal, B., Dogra, A., & Sangaiah, A.K. (2022). An effective nonlocal means image denoising framework based on non-subsampled shearlet transform. *Soft Computing*, *26*(16), 7893-7915.

Kandati, D.R., & Gadekallu, T.R. (2023). Federated learning approach for early detection of chest lesion caused by covid-19 infection using particle swarm optimization. *Electronics*, *12*(3), 710. https://doi.org/10.3390/electronics12030710.

Kareem, A., Liu, H., & Velisavljevic, V. (2023). A federated learning framework for pneumonia image detection using distributed data. *Healthcare Analytics*, *4*(100204), 100204. https://doi.org/10.1016/j.health.2023.100204.

Konečný, J., McMahan, H.B., Yu, F.X., Richtárik, P., Suresh, A.T., & Bacon, D. (2016). Federated learning: strategies for improving communication efficiency. In *Proceedings of the 29th Conference on Neural Information Processing Systems*, (pp. 5-10), Curran Associates, Inc., Barcelona, Spain.

Lepcha, D.C., Dogra, A., Goyal, B., Goyal, V., Kukreja, V., & Bavirisetti, D.P. (2023). A constructive non-local means algorithm for low-dose computed tomography denoising with morphological residual processing. *PloS One*, *18*(9), e0291911. https://doi.org/10.1371/journal.pone.0291911.

Mabrouk, A., Redondo, R.P.D., Elaziz, M.A., & Kayed, M. (2023). Ensemble federated learning: an approach for collaborative pneumonia diagnosis. *Applied Soft Computing*, *144*, 110500. https://doi.org/10.1016/j.asoc.2023.110500.

Makkar, A., & Santosh, K.C. (2023). SecureFed: federated learning empowered medical imaging technique to analyze lung abnormalities in chest X-rays. *International Journal of Machine Learning and Cybernetics*, *14*(8), 2659-2670.

Malik, H., & Anees, T. (2024). Federated learning with deep convolutional neural networks for the detection of multiple chest diseases using chest x-rays. *Multimedia Tools and Applications*, *83*(23), 63017-63045.

Malik, H., Naeem, A., Naqvi, R.A., & Loh, W.K. (2023). DMFL_Net: a federated learning-based framework for the classification of COVID-19 from multiple chest diseases using X-rays. *Sensors*, *23*(2), 743.

Mooney, P. (2018). *Chest X-Ray images (Pneumonia)* [Dataset]. https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia.

Mujahid, M., Rustam, F., Álvarez, R., Mazón, J.L.V., Díez, I.D.L.T., & Ashraf, I. (2022). Pneumonia classification from X-ray images with inception-V3 and convolutional neural network. *Diagnostics*, *12*(5), 1280. https://doi.org/10.3390/diagnostics12051280.

Naz, S., Phan, K., & Chen, Y.P.P. (2024). Centralized and federated learning for COVID-19 detection with chest X-ray images: implementations and analysis. *IEEE Transactions on Emerging Topics in Computational Intelligence*, *8*(4), 2987-3000.

Ragnarsdottir, H., Ozkan, E., Michel, H., Chin-Cheong, K., Manduchi, L., Wellmann, S., & Vogt, J.E. (2024). Deep learning based prediction of pulmonary hypertension in newborns using echocardiograms. *International Journal of Computer Vision*. *132*, 2567-2584. https://doi.org/10.1007/s11263-024-01996-x.

Rawat, V., & Suryakant. (2019). A classification system for diabetic patients with machine learning techniques. *International Journal of Mathematical, Engineering and Management Sciences*, *4*(3), 729-744.

Riedel, P., von Schwerin, R., Schaudt, D., Hafner, A., & Späte, C. (2023). ResNetFed: federated deep learning architecture for privacy-preserving pneumonia detection from COVID-19 chest radiographs. *Journal of Healthcare Informatics Research*, *7*(2), 203-224.

Rodrigues, M., Mayo, M., & Patros, P. (2022). Surgical tool datasets for machine learning research: a survey. *International Journal of Computer Vision*, *130*(9), 2222-2248.

Rushdi, A.M.A., & Serag, H.A.M. (2020). Solutions of ternary problems of conditional probability with applications to mathematical epidemiology and the COVID-19 pandemic. *International Journal of Mathematical, Engineering and Management Sciences*, *5*(5), 787-811.

Sharma, R., & Kukreja, V. (2022). Amalgamated convolutional long term network (CLTN) model for lemon citrus canker disease multi-classification. In *2022 International Conference on Decision Aid Sciences and Applications* (pp. 326-329). IEEE. Chiangrai, Thailand.

Sharma, S., & Guleria, K. (2022). Pneumonia detection from chest X-ray images using transfer learning. In *2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)*(pp. 1-6). IEEE. Noida, India.

Sharma, S., & Guleria, K. (2023a). A comprehensive review on federated learning based models for healthcare applications. *Artificial Intelligence in Medicine*, *146*(102691), 102691.

Sharma, S., & Guleria, K. (2023b). A deep learning based model for the detection of pneumonia from chest X-ray images using VGG-16 and neural networks. *Procedia Computer Science*, *218*, 357-366.

Sharma, S., & Guleria, K. (2023c). A systematic literature review on deep learning approaches for pneumonia detection using chest X-ray images. *Multimedia Tools and Applications*, *83*(8), 24101-24151. https://doi.org/10.1007/s11042-023-16419-1.

Sharma, S., Guleria, K., Kumar, S., & Tiwari, S. (2023). Deep learning based model for detection of vitiligo skin disease using pre-trained Inception V3. *International Journal of Mathematical, Engineering and Management Sciences*, *8*(5), 1024-1039.

Srivastav, S., Guleria, K., & Sharma, S. (2023). Deep learning-based convolutional neural network for Covid-19 and pneumonia prediction. In *2023 2nd International Conference on Futuristic Technologies* (pp. 1-5). IEEE. Belagavi, Karnataka, India.

Zaffiri, L., Gardner, J., & Toledo-Pereyra, L.H. (2012). History of antibiotics. from salvarsan to cephalosporins. *Journal of Investigative Surgery: The Official Journal of the Academy of Surgical Research*, *25*(2), 67-77.

Publisher's Note- Ram Arti Publishers remains neutral regarding jurisdictional claims in published maps and institutional affiliations.