Diagnosis of Pneumonia from Chest X-Ray Images Using Federated Learning

Rahul Mahto
Dept. of Information Technology
IIIT, Allahabad
prayagraj, India
iit2020022@iiita.ac.in

Saurabh Kumar
Dept. of Information Technology
IIIT, Allahabad
prayagraj, India
iit2020055@iiita.ac.in

Rohit Chowdhury
Dept. of Information Technology
IIIT, Allahabad
prayagraj, India
iit2020043@iiita.ac.in

Asadi Srinivasulu
Dept. of Information Technology
IIIT, Allahabad
prayagraj, India
rwi2023002@iiita.ac.in

Anupam Agrawal, IEEE, Senior Member
Dept. of Information Technology
IIIT, Allahabad
prayagraj, India
anupam@iiita.ac.in

***Abstract*—Pneumonia continues to be a major global health concern, accounting for millions of cases and hundreds of thousands of deaths each year. Ensuring accurate and prompt diagnosis of pneumonia is crucial for effective disease management. However the chest x - ray images data is heterogeneous in nature and has intricate pattern so to train the machine learning model we need huge amount to data but according to General Data Protection Regulation (GDPR) and the Data Protection Act (DPA), data sharing is not allowed by the third party. So We will use a federated learning strategy to overcome these difficulties, in which the model is trained on decentralized data without having access to the original data. Our method entails delivering locally trained models, like ResNet-50, Inception V3, and VGG-16, to client devices. We conducted a separate assessment of the accuracy with three models—namely VGG-16, ResNet-50, and Inception V3 with a federated learning paradigm to diagnosis of pneumonia from chest X-ray images dataset as part of our research work.**

***Keywords*—Federated learning, Vgg-16, ResNet-50, Inception V3**

I. Introduction

Pneumonia, a serious respiratory infection characterized by lung inflammation and fluid buildup in the air sacs, is a significant cause of death worldwide, performing in roughly 2 million losses annually [1]. People who are more likely to get pneumonia are youthful children, the old age people , and people who are immunocompromised. Early and accurate diagnosis is crucial for effective treatment and improved patient outcomes. Traditionally, chest X-ray imaging has been used for diagnosis, but interpreting these images can be challenging due to the need for specialized radiological expertise but Artificial intelligence and machine learning have revolutionized medical image analysis, and now these chest x-ray images can be interpreted by various deep learning algorithm. As the Deep learning models require large volumes of annotated data, which poses challenges related to patient privacy, data security, and logistical difficulties. Federated learning (FL) offers a promising solution by developing robust deep learning models without centralized data collection. Federated learning trains models across multiple decentralized devices or servers, ensuring patient data remains secure within the originating institution.This research aims to explore and validate the efficacy of federated learning in diagnosing pneumonia from chest X-ray images, integrating advanced deep learning techniques with federated learning frameworks.

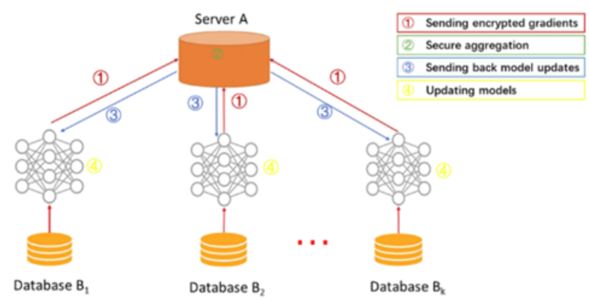

Fig. 1. Federated Learning Diagram[2]

Above is the diagram of federated learning paradigm, Where we have three clients and one server. Each client then trains the model locally using its own data and sends the updated model parameters back to the server. The central server aggregates these updates to refine the global model, which is then redistributed to the devices for further local training. This iterative process continues until the model achieves satisfactory performance.

II. Related Work

In this section, we discuss and summarise existing approaches that contribute to chest X-ray image classification using machine learning methods. The paper titled ”Improved Pneumonia Diagnosis of Radiological Images using Hybrid Loss with Conventional CNN” by Kusum Rajpurohit and Tushar Sandhan focuses on enhancing pneumonia detection from chest X-ray images using a convolutional neural network (CNN) model integrated with a novel hybrid loss function[3]. The study aims to improve the accuracy and efficiency of pneumonia diagnosis which is traditionally done by radiologists through visual inspection of chest X-rays—a process that can be subjective and prone to errors. The CNN model with hybrid loss outperformed other models, achieving an accuracy of 0.91 and an F1-score of 0.91 on the original dataset. The paper ”Pneumonia Detection Using Federated Learning” by Camelia-Raluca Farkas, Radu-Ioan Ciobanu, and Ciprian Dobre addresses the challenge of detecting pneumonia in elderly patients using a federated learning approach[4]. The dataset consisted of

5,863 labeled chest X-ray images, divided into training, testing, and validation sets. These images were distributed among 10 clients to simulate a federated learning environment. Federated learning allows for the distribution of machine learning tasks across multiple devices while preserving data privacy and reducing communication overhead. The federated learning model showed promising results in detecting pneumonia, demonstrating that it is a viable approach for distributed medical data analysis while preserving patient privacy. Federated learning typically involves solving an optimization problem across distributed devices. Algorithms such as Federated Stochastic Gradient Descent (FSGD), Federated Averaging, or Federated Proximal methods may be used to optimize the global model. Luka Racic in his paper ”Pneumonia Detection Using Deep Learning Based on Convolutional Neural Network” proposed a basic CNN model architecture and achieved an accuracy of 0.88 [5]. The dataset [5] used in the paper is provided by Guangzhou Women and Children’s Medical Center, Guangzhou.

III. Methodology

A. Research Design

This study aims to investigate if federated learning can improve the accuracy and efficiency of diagnosing pneumonia while maintaining patient privacy. Pneumonia is a critical lung infection that requires prompt detection to save lives. Federated learning allows computers to work on the same data without sharing private information, making it a valuable tool in healthcare. The study will use chest X-ray images from a public database, deep learning, and federated learning to identify pneumonia. However, challenges such as bad data, high communication costs, and insufficient training resources will be addressed using experimental and analytical methods.

1) Data collection and preprocessing: We have collected chest x-ray images data from kaggle. We will implement techniques like image scaling, normalization, and augmentation on the dataset.

2) Model architecture and selection: In this step we will try various deep learning model like Vgg-16, ResNet-50, and Inception V3 is.

3) Federated learning approach: In order to ensure data privacy we will use federated learning approach. We will apply flower federated learning to establish the clients, server and communication protocols.

4) Evaluation metrics: We will evaluate the performance according to standard metrices. We will calculate the accuracy , precision , recall and f1 score. We will also find confusion matrix for better understanding of our result.

B. Data Collection and Preprocessing

In the preprocessing step of dataset preparation, we perform several operations like augmentation, scaling, and normalization for various reasons:

1) Augmentation: This procedure involves creating supplementary training samples by implementing alterations such as rotation, mirroring, or cropping to pre-existing images. Augmentation helps in increase the diverserity of our dataset by applying these above operations in order to prevent overfitting cause it becomes diverse in nature and can be used for robust detection of image classification task.

2) Scaling: Scaling refers to adjusting the size of images to a consistent dimension. This ensures uniformity across all images, which is essential for inputting them into a neural network. Additionally, it aids in minimizing computational complexity and shortening training duration.

3) Normalization: Normalization involves rescaling the pixel values of the images to a common scale, typically between 0 and 1 or -1 and 1. This process helps in standardizing the input data and makes training more stable and efficient. It also helps in preventing issues like vanishing or exploding gradients during training.

4) Balancing the Dataset: To prevent biased models caused by imbalanced datasets, where there are many more normal cases than pneumonia cases, we’ll balance the dataset. This involves randomly increasing the number of pneumonia cases (oversampling) and decreasing the number of normal cases (undersampling). This ensures that both classes are equally represented in the dataset during training, preventing the model from favoring one class over the other.

C. Convolution Networks and Machine Learning

CNNs outperform traditional methods in pneumonia detection due to their ability to identify subtle abnormalities and intricate patterns in medical images, and their ability to learn hierarchical representations from raw pixel data. Convolution Neural Network works by using following layers.

1) Convolutional Layer: The convolutional layer is the fundamental element of image processing, involving the application of learnable filters to the input image, resulting in feature maps.

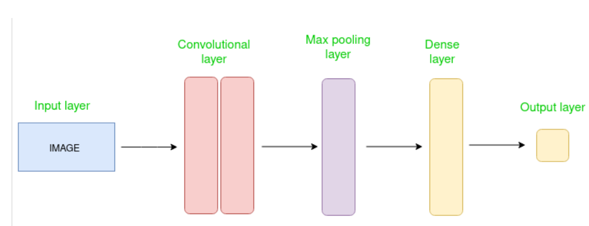

Fig. 2. Simple CNN Architecture[6]

2) Relu Activation Function: CNNs employ the Rectified Linear Unit (ReLU) activation function to enhance network nonlinearity, aiding in learning intricate and asymmetrical data relationships between zero and input values.

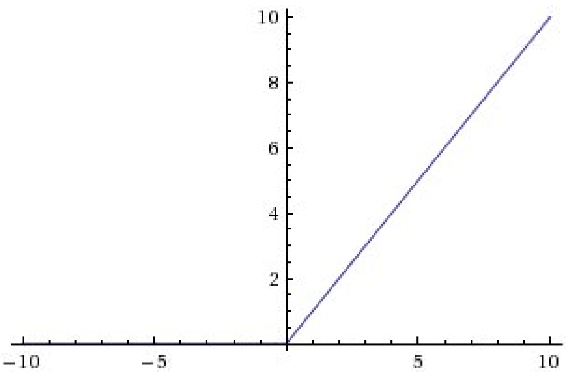

Fig. 3. Rectified Linear Activation Function[7]

The ReLU (Rectified Linear Unit) function is written as follows:

$$f(x) = x^+ = \max(0, x) = \frac{x + |x|}{2} = \begin{cases} x & \text{if } x > 0, \\ 0 & \text{otherwise.} \end{cases}$$

The derivative of the ReLU function is:

$$f'(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x \leq 0. \end{cases}$$

3) Softmax Activation Function: When performing multi-class classification tasks, the output layer of a CNN frequently employs the Softmax activation function. It takes the raw output values from the preceding layer and normalizes them across all classes to create probabilities. The outputs can be interpreted as class probabilities because the Softmax function makes sure that the predicted probabilities add up to one

The standard Softmax function $\sigma : \mathbb{R}^K \to (0, 1)^K$ is defined when $K \geq 1$ by the formula:

$$\sigma(\mathbf{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \quad \text{for} \quad i = 1, \ldots, K \quad \text{and} \quad \mathbf{z} = (z_1, \ldots, z_K)$$

4) Max Pooling layer: The feature maps produced by the convolutional layers are downsampled using the MaxPooling layer.

The feature map dimensions are reduced while the most important data is preserved. The feature map is divided into non-overlapping regions via MaxPooling, and the highest value is chosen within each region. The network's computational complexity is decreased while the most dominating features are extracted using this method. The following is a representation of the MaxPooling:-

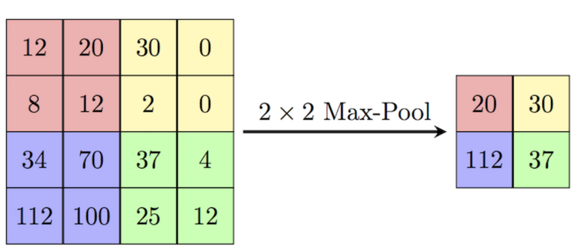

Fig. 4. Max Pooling Layer[8]

5) Flattening: The flattening process allows the feature maps to be converted into a linear format, enabling them to be passed into a fully connected layer for classification task.

6) Fully connected Layer: This layer takes the input from the preceding layer, typically a flattened one dimensional vector from the previous layers, and perform the final classification task.

7) Output Layer: The above all transformation converts the raw input of each class into a probability score, indicating the likelihood of the input belonging to each class and helps in giving the desired output.

D. Model Architecture and Selection

We have evaluated the performance of three deep learning architectures: ResNet50, InceptionV3, and VGG16, in the context of pneumonia diagnosis. Accuracy is the primary criterion for model selection, as it helps healthcare practitioners trust diagnostic information. Interpretability is crucial for transparency and confidence among clinicians, and computational efficiency is essential for practical deployment in healthcare settings with limited resources. The evaluation process involves rigorous experimental procedures, including training, validation, and test sets. Hyperparameters will be fine-tuned using the validation set, and model performance will be assessed using the testing set. Interpretability will be examined by visualizing activation maps of chest X-ray images to identify significant areas contributing to the model's predictions. Computational efficiency will be assessed by measuring training and inference times. The chosen architecture will serve as the foundation for further research and analysis, ultimately aiding in the development of a precise, understandable, and computationally efficient pneumonia detection system.

1) VGG 16: Simonyan and Zisserman presented VGG16 in 2014, which is a widely recognized design of a convolutional neural network. The name of the group is derived from the Visual Geometry Group at the University of Oxford. VGG16 is composed of a total of 16 layers, specifically 13 convolutional layers and 3 fully connected layers. The distinguishing features of VGG16 are its straightforwardness, consistency, and the utilization of compact 3x3 filter dimensions. Although VGG16 has a simple design, it has proven to be extremely efficient in a wide range of computer vision tasks.

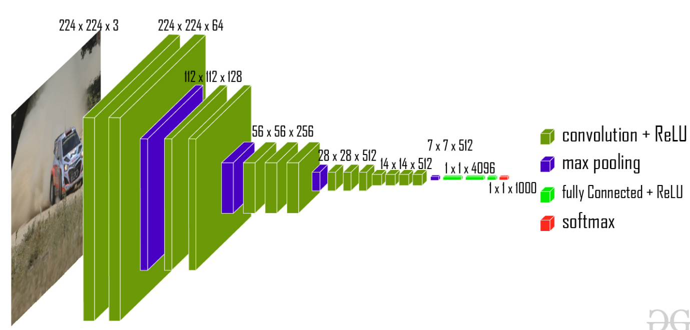

Fig. 5. VGG-16 architecture[9]

In the VGG16 architecture, max-pooling layers are inserted after groups of convolutional layers to reduce the spatial dimensions. The classifier, located at the end of the network, consists of fully connected layers. While VGG16 achieves high accuracy, its major drawback is its large number of parameters. This results in high computational costs for both training and usage.

2) ResNet 50: ResNet50, part of the ResNet family introduced by He et al. in 2016, addresses the vanishing gradient problem. It's a deep neural network comprising 50 layers, hence the name ResNet50. By incorporating skip connections, also known as shortcuts, ResNet50 enables the direct flow of gradients through the network. This innovation allows for the training of extremely deep networks without suffering from the vanishing gradient issue.

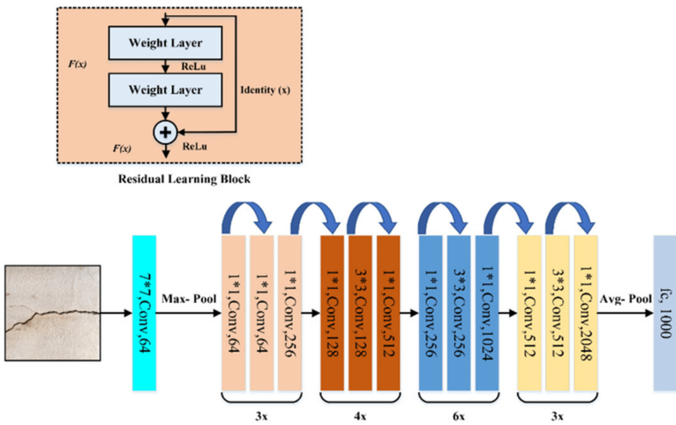

Fig. 6. ResNet50 architecture[10]

In the ResNet50 architecture, each residual block consists of multiple convolutional layers. Identity skip connections play a crucial role in propagating information across the network, facilitating the training of deeper models without performance degradation. ResNet50 has demonstrated exceptional performance in addressing image classification challenges, consistently achieving state-of-the-art results.

3) Inception V3: InceptionV3, devised by Szegedy et al. in 2015, is another notable convolutional neural network design. Its primary aim is to strike a balance between computational efficiency and model complexity. InceptionV3 achieves this by employing multiple parallel convolutional operations with different kernel sizes. These operations are organized into inception modules, enabling the network to capture features at various scales efficiently.

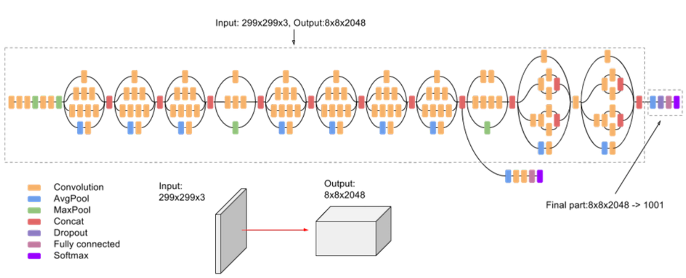

Fig. 7. InceptionV3 architecture[11]

In the InceptionV3 architecture, a series of inception modules precede fully connected layers for categorization. These modules facilitate the extraction of intricate patterns by capturing features at different sizes and resolutions. InceptionV3 strikes a harmonious balance between model complexity and computational efficiency, making it suitable for environments with resource constraints.

E. Federated Learning Approach

1) Overview Of Federated Learning: A decentralized machine learning paradigm called federated learning enables model training on numerous distributed data sources while maintaining data privacy and confidentiality.

2) Communication Between Clients and Server: Effective communication between clients and servers is crucial for coordination, model updates, and aggregation in a federated learning system. The gRPC framework, known for its high performance, facilitates this process by allowing bidirectional streaming and using Protocol Buffers data serialization.

gRPC offers several advantages for client-server communication in federated learning:

Communication in Federated Learning

- Efficient and Scalable Communication : -GRPC utilizes the HTTP/2 protocol, which supports multiplexing and asynchronous processing. This enhances communication efficiency by enabling the simultaneous handling of multiple requests and responses. Consequently, gRPC is well-suited for efficiently managing large federated learning systems comprising numerous clients.

- Bidirectional Streaming : - Federated learning involves iterative updates exchanged between clients and the server. With gRPC's bidirectional streaming capability, the server can transmit model commands and parameters to clients while simultaneously receiving model updates from them. This bidirectional data exchange accelerates federated learning processes and facilitates real-time collaboration.

- Strong Message Serialization : - GRPC defines message structures exchanged between clients and the server using Protocol Buffers, a binary serialization standard that is language-independent. With Protocol Buffers, we can define message formats using a clear and extensible syntax, ensuring data consistency and interoperability across various programming languages utilized by clients and the server. Additionally, Protocol Buffers' strong typing further guarantees data integrity and compatibility.

- Security and Authentication
  GRPC offers native support for Transport Layer Security (TLS) encryption, which guarantees secure communication between clients and servers. This feature plays a crucial role in maintaining the confidentiality and integrity of data transmitted during federated learning processes. Additionally, gRPC provides authentication mechanisms, such as token-based authentication or mutual TLS authentication, to verify the identities of clients and servers, enhancing overall security measures.

3) Model Aggregation: In federated learning, model aggregation is a vital process where the server aggregates the model updates from multiple clients and create an updated global model.

- **Weighted Averaging:** It combine the updates from different nodes. Each node is assigned a weight that determines its influence on the global model update. The weight can be determined based on various factors such as the amount or quality of data available on the node, the computational resources of the device, or the importance of the node's data.

- **Federated Averaging(FedAvg):** Federated averaging is different from the weighted averaging because in this equal weights are assigned to nodes irrespective of importance of data present at node and computational efficiency of device.

Below is the diagram of whole sequence of federated learning approach.

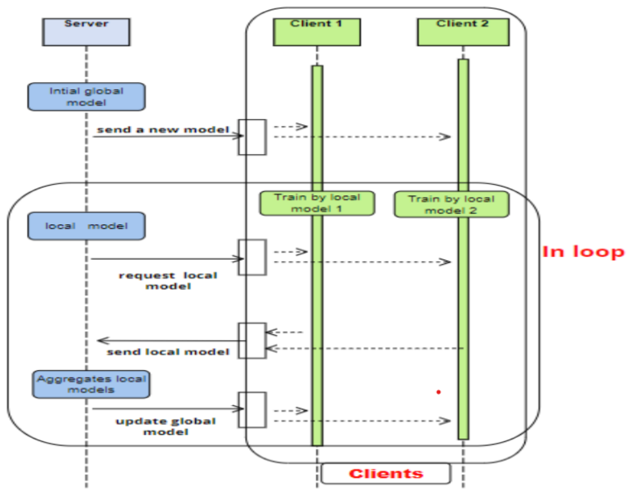

Fig. 8. Federated Learning Sequence Diagram

IV. Dataset Description

The dataset consists of 5856 chest X-ray images, with 1583 labeled "Normal" and 4273 labeled "Pneumonia." The dataset is divided into three subsets: training, validation, and test sets to ensure robustness and generalizability. The gradient descent method is used to train the model's parameters on the training set, which contains 4192 images. The validation set evaluates the model's performance, guiding hyperparameter tuning and model selection. The test set assesses the model's performance on unobserved data, using 626 images. The dataset's separation into training, validation, and test sets allows for a thorough assessment of the model's performance and guarantees its ability to generalize to new data. To maintain a balanced and diverse dataset, a representative distribution of pneumonia and normal patients is preserved in each subgroup. The final distribution of the dataset can be observed in the table.

TABLE I
Distribution of data in the dataset

	Total	Normal	Pneumonia	Percentage (%)
Train	4192	1082	3110	72
Validation	1038	266	772	18
Test	626	235	391	10

Fig. 9. Examples of Chest X-Rays in Patients with Pneumonia[12]

Fig. 9, displaying chest X-ray Images samples from individuals with pneumonia (Source: Kaggle), offers visual aids that facilitate comprehension and recognition of pneumonia in the research setting.

V. Results and Analysis

We have analyzed the accuracy of our model, namely VGG-16, ResNet-50, and InceptionV3, following the federated learning paradigm. We have calculated the accuracy, precision, support, and recall for a better understanding of the accuracy parameters. The following performance metrics were achieved:

TABLE II
VGG16 performance metrics

Class	Precision	Recall	F1-score	Support
Normal	0.87	0.82	0.85	235
Pneumonia	0.90	0.94	0.92	391
Accuracy		0.90		
Fed Avg	0.90	0.89	0.89	626
Weighted Avg	0.90	0.90	0.90	626

Table II shows the accuracy metrics of the VGG-16 architecture. In this case, we have achieved an accuracy of 0.90.

TABLE III
ResNet50 performance metrics

Class	Precision	Recall	F1-score	Support
Normal	0.80	0.88	0.84	235
Pneumonia	0.92	0.87	0.89	391
Accuracy		0.87		
Fed Avg	0.86	0.87	0.86	626
Weighted Avg	0.88	0.87	0.87	626

Similarly, Table III shows the accuracy metrics of the ResNet50 architecture. In this case, we have achieved an accuracy of 0.87.

TABLE IV
InceptionV3 Performance metrics

Class	Precision	Recall	F1-score	Support
Normal	0.89	0.80	0.84	235
Pneumonia	0.89	0.94	0.91	391
Accuracy		0.89		
Fed Avg	0.89	0.87	0.88	626
Weighted Avg	0.89	0.89	0.89	626

Similarly, Table IV shows the accuracy metrics of the InceptionV3 architecture. In this case, we have achieved an accuracy of 0.89.

TABLE V

COMPARISON OF THE PROPOSED MODEL WITH RECENT STUDIES

Models	Accuracy	Classification
Festa(federated Split Task-Agnostic Learning[13]	0.89	Covid-19
DenseNet121[14]	0.89	HealthCare
DenseNet[15]	0.81	Chest Disease
VGG16	90	Pneumonia
InceptionV3	89	Pneumonia
ResNet50	87	Pneumonia

Tables II, III, and IV offer comprehensive performance metrics such as precision, recall, F1-score, and support for the VGG16, ResNet50, and InceptionV3 models in pneumonia detection. These metrics are essential for assessing each model's efficacy in identifying pneumonia cases, providing crucial insights within the research framework, and for comparative analysis with other studies depicted in Table V.

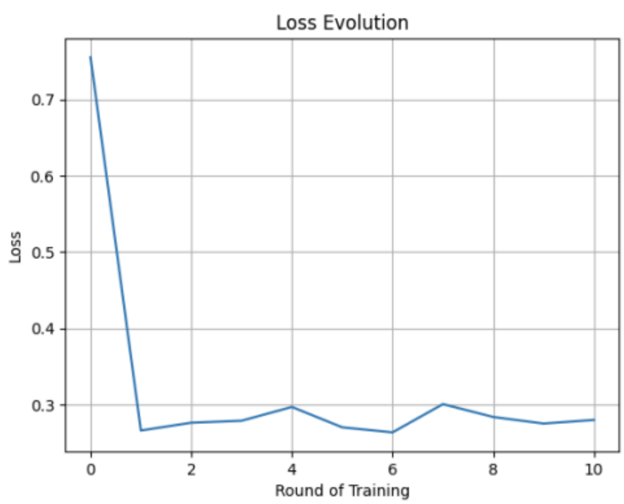

Fig. 10. The graph shows Loss function over round of training[VGG16 Model]

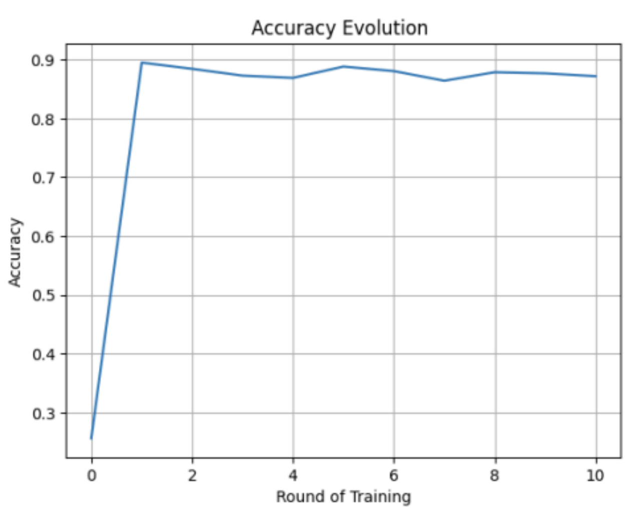

Fig. 11. The graph shows Accuracy function over round of training[VGG16 Model]

VI. CONCLUSION AND FUTURE WORK

This research work presents a novel approach for diagnosis of pneumonia from chest x-ray images using federated learning. We have used Visual Geometry Group(VGG-16), Residual Network(ResNet-50) and Inception V3 model for client side training. We have achieved an accuracy of 0.90 during training local clients with vgg - 16 model while an accuracy of 0.87 with ResNet-50 model and an accuracy of 0.89 when we used Inception V3.The future work could be Exploring the integration of more advanced neural network architectures like DenseNet, Efficient Net, or Transformer-based models could potentially unlock even greater efficiencies and accuracies in diagnosing pneumonia within the federated learning paradigm.

REFERENCES

[1] R. S. Gereige and P. M. Laufer, “Pneumonia,” *Pediatr. Rev.*, vol. 34, no. 10, pp. 438–456, Oct. 2013. doi: 10.1542/pir.34-10-438.

[2] Y. Ko, “Horizontal Federated Learning,” Medium. [Online]. Available: https://medium.com/disassembly/architecture-of-federated-learning-a36905c1d225.

[3] K. Rajpurohit and T. Sandhan, “Improved Pneumonia Diagnosis of Radiological Images using Hybrid Loss with Conventional CNN,” in *Proc. ICMOCE*, May 2023, pp. 512–517. doi: 10.1109/ICMOCE57812.2023.10166471.

[4] C. R. Farkas, R. I. Ciobanu, and C. Dobre, “Pneumonia Detection Using Federated Learning,” arXiv preprint arXiv:2305.16370, 2023.

[5] L. Racić, T. Popović, S. Ćakić, and S. Šandi, “Pneumonia detection using deep learning based on CNN,” in *Int. Conf. Inf. Technol.*, 2021.

[6] ”Simple CNN Architecture,” GeeksforGeeks [Online]. Available: https://www.geeksforgeeks.org/introduction-convolution-neural-network.

[7] DANB, ”Rectified Linear Activation Function,” Kaggle. [Online]. Available: https://www.kaggle.com/code/dansbecker/rectified-linear-units-relu-in-deep-learning.

[8] ”Max Pooling Layer,” PaperswithCode. [Online]. Available: https://paperswithcode.com/method/max-pooling.

[9] ”VGG-16 Architecture,” GeeksforGeeks. [Online]. Available: https://media.geeksforgeeks.org/wp-content/uploads/20200219152207/new41.jpg.

[10] ”ResNet50 Architecture,” WisdomML. [Online]. Available: https://wisdomml.in/wp-content/uploads/2023/03/resnet.png.

[11] ”InceptionV3 Architecture,” GoogleCloud. [Online]. Available: https://cloud.google.com/tpu/docs/inception-v3-advanced.

[12] ”Examples of Chest X-Rays in Patients with Pneumonia,” Kaggle. [Online]. Available: https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia.

[13] S. Park, G. Kim, J. Kim, B. Kim, and J. C. Ye, “Federated Split Vision Transformer for COVID-19 CXR Diagnosis using Task-Agnostic Training,” arXiv preprint arXiv:2111.01338, 2021. [Online]. Available: https://arxiv.org/abs/2111.01338.

[14] T. V. Nguyen, M. A. Dakka, S. M. Diakiw, M. D. VerMilyea, M. Perugini, J. M. M. Hall, and D. Perugini, “A novel decentralized federated learning approach to train on globally distributed, poor quality, and protected private medical data,” *Sci. Rep.*, vol. 12, Art. no. 8888, 2022. [Online]. Available: https://www.nature.com/articles/s41598-022-12833-x.

[15] K. V. Priya and J. D. Peter, “A federated approach for detecting the chest diseases using DenseNet for multi-label classification,” *Complex Intell. Syst.*, vol. 8, pp. 3121–3129, Jul. 2021. doi: 10.1007/s40747-021-00474-y.

[16] C. Matsoukas, J. F. Haslum, M. Sorkhei, M. Soderberg, and K. Smith, “What makes transfer learning work for medical images: Feature reuse other factors,” in *Proc. IEEE/CVF CVPR*, 2022.

[17] K. L. O’Brien, L. J. Wolfson, J. P. Watt, E. Henkle, M. Deloria-Knoll, N. McCall, E. Lee, K. Mulholland, O. S. Levine, T. Cherian, et al., “Burden of disease caused by streptococcus pneumoniae in children younger than 5 years: global estimates,” *The Lancet*, 2009.

[18] M. Tan and Q. V. Le, “EfficientNet: Rethinking model scaling for convolutional neural networks,” 2020.

[19] Y. T. Alaoui, A. Berrahhou, K. Douge, I. Belabed, and E. M. Jaara, “Classification of chest pneumonia from x-ray images using new architecture based on ResNet,” in *Proc. IEEE ICECOCS*, 2022.