Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 218 (2023) 1485–1496

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Machine Learning and Data Engineering, 2022

# Simulating Federated Transfer Learning for Lung Segmentation using Modified UNet Model

Sateesh Ambesange[a,*], Annappa B[a], Shashidhar G Koolagudi[a]

[a] *Computer Science Engineering Dept, National Institute of Technology Karnataka, Surathkal,* Mangalore, Karnataka state, India- 575025

**Abstract**

Lung segmentation helps doctors in analyzing and diagnosing lung diseases effectively. Covid -19 pandemic highlighted the need for such artificial intelligence (AI) model to segment Lung X-ray images and diagnose patient covid conditions, in a short time, which was not possible due to huge number of patient influx at hospitals with the limited radiologist to diagnose based on test report in short time. AI models developed to assist doctors to diagnose faster, faces another challenge of data privacy. Such AI Models, for better performance, need huge data collected from multiple hospitals/diagnostic centres across the globe into single place to train the AI models. Federated Learning (FL) framework, using transfer learning approach addresses these concerns as FL framework doesn't need data to be shared to outside hospital ecosystem, as AI model get trained on local system and AI model get trained on distributed data. FL with Transfer learning doesn't need the parallel training of the model at all participants nodes like other FL. Paper simulates Federated Transfer learning for Image segmentation using transfer learning technique with few participating nodes and each nodes having different size dataset. The proposed method also leverages other healthcare data available at local system to train the proposed model to overcome lack of more data. Paper uses pre-trained weights of U-net Segmentation Model trained for MRI image segmentation to lung segmentation model. Paper demonstrates using such similar healthcare data available at local system helps improving the performance of the model. The paper uses Explainable AI approach to explain the result. Using above three techniques, Lung segmentation AI model gets near perfect segmentation accuracy.



*Keywords:* ***Federated Learning, Transfer Learning, Federated Transfer Learning, X-ray Image segmentation, MRI image segmentation, data privacy, Lung image segmentation, U-net Architecture***

## 1. Introduction

Whole world experienced massive Covid 19 pandemic outburst during 2020 - 2021 and faced sever resource crunch across all skill segments including specialist radiologist. Several deaths happened, when patient had covid19 symptom and due to shortage of available radiologist, the diagnosing such patients is delayed by few days. So the need arises for identifying patients who have some symptoms, faster and then send such patient to specialist doctors for further


10.1016/j.procs.2023.01.127

diagnosis, instead of sending all patients to specialist doctors. To accomplish such segregation of patients, several Artificial Intelligence (AI) models are developed during covid pandemic period. These AI models detect lung complication early on by analysing X-ray and other Radiology Images and rate/rank risk level. Hospital/doctors select the patient with higher risk to diagnose first to reduce number of deaths. Such AI models helps the effective utilization of an available small number of specialist doctors and also reduce the number of deaths by sending patient whose cases have to be addressed fast to reduce the overall earth.

Even though these AI models address the resource shortage problem, the AI model faces several challenges like data privacy concerns In recent years, data privacy rules are drafted by several countries to reduce the misuse of health data to avoid health insurance claims, and other kinds of misuse. Data privacy is becoming the new priority of hospitals and patients. The AI models need data and more data to train the model to achieve better performance, hence data privacy is becoming a major hurdle in AI model development.

Another concern of AI models is data aggregation in a single storage space. To accumulate huge data a single place, one need bigger storage space and training AI model on such huge data need huge computing capability. Increases Space and computation power cost more.

Federated Learning(FL) framework addresses these two concerns 1. data privacy issues and 2. Data Aggregation concern, as FL framework works with distributed data. In Federated Learning, model will be learning on the local data at specific node and only learning of model will be transferred and aggregated hence data privacy issues addressed.

There are three kind of FL methods 1. Vertical Federated Learning(VFL), 2. Horizontal Federated Learning(HFL) and 3. Federated Transfer Learning(FTL). The paper opted FTL approach and simulate FTL whole using single system.

Deep Learning models need large data to train model and achieve better performance. Non Availability of large relevant data at each node is another concern of AI model developer. Paper address the issue by leveraging other similar domain data or similar problem type data available at each node to get better performance. This approach is called data driven model development.

The other concern of doctors or patients about AI models using in diagnosis is, the trust as the AI model just gives output, which works like black box model. To address this concern, The paper usage GradCAM[7] to explain the diagnosis result, and help doctors to understand the diagnosis output effectively.

In short, paper attempting to address, data privacy concern, data aggregation concern, concern of having less data and concern of block box output of model by using data driven AI model development approach and simulating FTL.

The paper is structure into 7 sections. The second section describe the Federated learning concepts. Third section describes dataset. Forth section describes, Research Methodology. Fifth section, provides the modified Unet architecture. Sixth section provides result and analyses result. Final section concludes and provides future scope.

## 2. Federated Learning & Federated Transfer Learning

### 2.1. *Federated Learning(FL)*

Federated learning [1,2] firstly coined by google scholars in 2016. In FL, Machine learning model learning from distributed data. In FL no need to collect all data from local system to central system. In FL the model learning on local data is transferred to central system and at central system the learning received from several FL nodes will be aggregated effectively using several model learning aggregating algorithms like FedAvg and others. The federated learning approach address the key concern of data privacy by removing need for sharing data to company or researcher who is building machine learning or deep learning model. Because Once data shared to other parties than hospital or patient will not have more control on such shared data and how people will use such data, even after getting signing on several data usage terms while sharing data.

Federated learning can be tested in three different ways[3], 1. Training over several edge devices, 2. Using distributed computing systems and 3. Simulating Federated learning on single system. The paper attempt to perform simulating on single system.

### *2.2. Federated Transfer Learning(FTL)*

In Federated Transfer learning approach, learning aggregation happens through Transfer learning mechanism and it is a sequential fashion.Transfer Learning(TL) is a method of sharing the learning of model from specific data to another similar model. Using TL most models are pre-trained. Usually most models are pre-trained on generic irrelevant datasets like ImageNet dataset. The model performance depends on 1. what kind of transfer learning applied for the target model and 2. what kind of data the model is pre-trained on. The TL method used will decides model to performance on new target data, even through new data size is small.

The paper select FTL approach because 1. FTL framework is scalable , 2. FTL works well on hybrid participating FL node 3. FTL works with different datasets with ease, ie. data sets with different feature sets, output labels and 4. FTL can be applied to any type of problem and any kind of machine learning or deep learning or speech processing models.

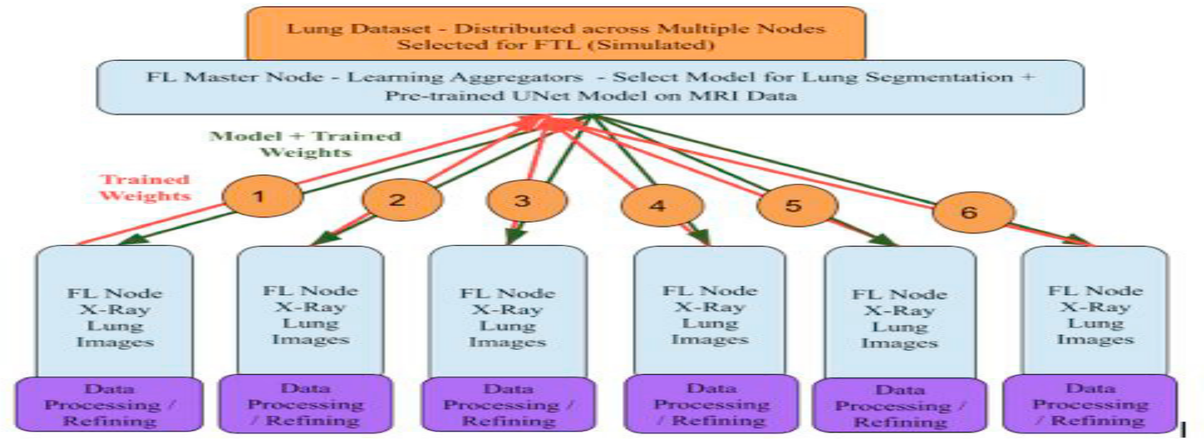



**Fig. 1.Federated Transfer Learning Framework for Lung Segmentation**

## 3. Dataset details

The paper used a part of X-ray Image datasets [5] from two sources first from Montgomery County (MC) chest X-ray set and second Shenzhen dataset - chest X-ray set. The dataset is prepared by National Library of Medicine, Maryland, in association with Shenzhen People’s Hospital. The dataset consists of x-ray images of 326 normal cases, 336 tuberculosis cases. Models are trained on not full datasets, but part of the combined dataset.

Since the overall dataset is small, which is not enough for effective AI model training, the paper pre-trains the proposed model using a transfer learning mechanism before starting to train the AI model on the actual dataset (above dataset).

To pre-train the proposed model, the first proposed model has to train on another dataset. Paper selects LGG MRI segmentation dataset, which consists of the MRI images of Brain Timor and normal images. The selection of such datasets is mainly because the possibility of having such datasets on local systems or participating FL nodes is high. The node or edge device participating in federated learning would be either system at the hospital or diagnostic center, where health data is digitized and stored. Such centers will have enough radiology images, several different type data and hence availability of MRI images is high.

## 4. Research Methodology

There are two common approach for building Deep Learning(DL) segmentation model for X-ray image of Lung organ. First approach is building customer DL models for the datasets. and the second approach is to leverage the existing well known deep learning segmentation architecture which is already proven architecture and then customizes the selected model for target data and then refine the model by tuning hyperparameters of the module to get better performance.

The second approach focuses more on data to get high performance and hence it is called data have driven model development approach. In Data-driven model, the model performance can be improved over available data, by Enhancing the data, using data argumentation, important feature extractions from data, leveraging relevant other data available at that system and using transfer learning and so on.

The proposed research method includes the following steps, 1. Data Processing, 2. Simulating distributed data on single system 3. Simulating The Federated Transfer Learning Mechanism. It also includes selecting other relevant data to be used for pre-train the model 4. finally explaining the result using GradCAM[7].

### *4.1. Key Terms of Federated Learning Framework*

Federated Transfer Learning (FTL), an FL framework type that uses Transfer Learning approach. FL Master Node, is the node that selects the FL nodes, who can participate in the model training process. FL Master Node controls the whole FL process. FL Node is system having data and computing power like a hospital system. Trained weights of the model is the value of weights of neurones of the model obtained when the model is trained on data. Pre-Trained model is the machine learning or deep learning model, which learns from a similar model trained on other data using transfer learning. Usually model will be pre-trained and then the model is further trained on new data/actual target data. In the pre-trained model, all weights of neurones of the model have proper values, so that model gives low loss on testing data. if the model is not pre-trained, then the weights of neurones of model have random values. Most of the deep learning Models either pertained on well known datasets, like CIFAR, MINIST or any other standard data from a similar domain for which the model is being developed.

### *4.2. Model development Process Steps*

The process of developing final model involves five key steps. The first step is simulating distributed data. Second step is simulating Federated Transfer Learning approach. Also selecting proposed model for Lung Segmentation. Third step is selecting pre -trained weights for the proposed model, which helps in improving performance of the proposed model. Fourth Step is processing local data for model training. And perform transfer learning on the proposed model first and then further train the model over process local data. Also verify the improvement of the result with local test data. Send back the Model and weights to Master node. Fifth step, is aggregate model learning in sequential fashion at Master node.

### *4.3. Multi Node distributed data Simulation*

The paper uses combined X-ray segmentation dataset provided by institutes as mentioned in above section. To simulate distributed data, paper randomly samples the data, such a way that there won't be any duplication of data in any of the samples. The data can be split into multiple sub-data sets to simulate distributed data using various approaches - balanced sized approach - in terms of the number of samples at each node, and balanced classes approach. The paper splits data into 6 distinct samples from the datasets such that the total data is equal to the sum of 6 samples. Instead of deciding randomly six node essentials and splitting data into 6 sub datasets, the number of Nodes to be created is decided based on the size of the data, and the number of classes in the datasets. For example, if datasets A have more classes and dataset B have less number of classes and both dataset's size is the same, then the number of sub data created for dataset A is less compared to datasets B, because the model needs more data to get trained effectively when datasets have more classes. Another example, the size of dataset A is more than the size of dataset B, then dataset A will be split into more sub-datasets. Also while splitting data into small sub-datasets in such a way that none of sub-datasets consists of data less than 5%. Data can be split equally among nodes or unevenly among nodes. The split well balanced way in terms of distribution class labels at each nodes same, or not balanced way to mimic real world data distribution.

The paper decided to split data unevenly and not balanced way. Data is split into the 6 sub datasets in ratio - 35% , 20% ,05%, 10% , 15%, 15 % Or 35:20:05:10:15:15 . In FL, model is trained parallel on all subdatas on same time, and hence order of FL node / subdata is not that important. In FTL, model trained on FL Node / sub data in sequential fashion, hence order in which model gets trained impact overall performance of the model. Selecting FL Nodes that have more data as initial FL nodes and selecting a few FL nodes having more data size as later FL nodes, will provide better performance. FL node 1: 35% of total data, FL node 2: 20% of total data, FL node 3: 05% of total data, FL node 4: 10%of total data, FL node 5: 15%of total data, FL node 5: 15 % of total data. In FTL model gets trained on each node in the order FL node 1 to FL node 6, in a sequential fashion.

The data split ratio, among multiple FL nodes, fixing the number of FL nodes for the dataset and the order of model training over FL nodes is dependent on the datasets. Its ways datasets to the dataset, but the pattern of sequence of FL node trained will be same, ie select the FL Nodes having max data as first FL node and next FL node with less data than the previous node like that each next FL nodes with reducing data size and final 1 or 2 nodes with increasing data size as mentioned above.

The observation for such data split and sequence of training is initial model will be pre-trained with weights of a model trained on a different dataset. Now the model will be further trained on target data or actual (X-ray) data. The deep learning model needs more data to get better performance. When the model is trained initially with more data, the model's loss(error) will be low and the model would have reached local minima. Later when the model is trained on the next FL node data the model performance either improves slightly or doesn't improve. If the data at FL node have different information or variance, along with a small portion of data with similar distribution or variation as data at previous nodes then the model performance improves. If the model performance is not improved, then the master node will reject the model's learning from the specific node and the master node further progress on learning aggregation.

### *4.4. Simulating Federated Transfer Learning & selecting Pre-trained Weights.*

- The federated transfer learning framework selects the modified UNet segmentation model for lung segmentation.
- FTL is simulated over the single system, so the master node does need not to send the selected model to all nodes.
- In FTL, every FL node, other than the first FL node, gets the model pre-trained with model learning from the previous node.
- In some cases, the first FL node has other data on a similar domain, for example, segmentation model, intended to train on Lung Segmentation data, initially trained on another organ segmentation and the model weights saved and through transfer learning model pre-trained with saved model learning.
- The paper proposes, Instead of pre-training proposed model with weights of similar model trained on generic datasets, pre-training proposed model with weights of similar model trained on relevant healthcare care data sets, and radiology datasets, gives better performance.
- The modified UNet model is pre-trained from weights of UNet model trained on MRI brain tumor Image data sets (LGG MRI Segmentation dataset). The selection depends on such data available at the first FL Node, else select any FL node, as first FL node, which has such a dataset. In this case, MRI datasets are assumed to be available at first FL Node. The key point here in selecting the MRI dataset is the data sets are the segmentation dataset, the mean image with masked image data. The first preference would be X-ray segmentation datasets for another body part, else any other radiology image dataset.
- In simulated FTL, the master node mimics sending UNet model to each node, along with learned weights. The proposed model, model learning all is saved in a single place.
- At each federated node, the model has to be pre-trained on model weights send by the master node. In this case, all models and model learning are available in a single place. The code to picks up previous FL node weights and then performs fine transfer learning on the proposed model before training the model on processed local lung x-ray data.
- Master node coordinates with all FL nodes and receives the model learning from all FL nodes. Then the master node aggregates the all-model learning received.

### *4.5. Data processing of local data at FL Node*

- Two sets of datasets used in paper. Montgomery dataset and Shenzhen hospital chest x-ray dataset
- For Montgomery datasets, left and right lung masks are combined to have mask for full lungs. The images are re-sized to 512X512 pixels images. Similarly for the Shenzhen hospital chest x-ray images are resized to 512X512 pixels images.
- The image converted to grayscale and then normalised. The normalised data are converted to binary value 0 or 1 based on threshold value >= 0.5.
- Finally data augmentation performed on data using train_generator from Keras. Basic data augmentation performed are, rotate image upto 20%, expand height and weight each upto 5%. shear_range upto 5%, horizontal flip, zoom the image upto 5% and fill the pixel value by nearest pixel.
- In simulated FTL setup, data processing will be performed on sub data, belongs to the virtual FL node.

### *4.6. Model Training at FL Node on local data*

- After data processing performed on local data or sub data, then processed data is split into training data and validation data.
- First model is trained on the data available at FL node without pre-trained on the weights send by master node. Second time, proposed model is trained on local data, after pre-training the model with model weights supplied by Master node. When model is trained, then check the performance of model by checking on its response on validation data. Performance of both scenario i.e model with pre-trained and model without pre-trained is compared.
- The learning weights of model, which performance better will be sent to master node.

### *4.7. Master Node learning aggregation*

In Federated Transfer Learning, there are two ways of model learning aggregation. First Sequential learning aggregation and another parallel learning aggregation like FedAvg approach. The paper proposes modified sequential learning aggregation method. There are two possible ways of modifying sequential learning in FTL.

First approach: Master Node, works in two phases. In first phase master node sends selected model and pre-trained weights to all FL nodes, who are keen to participate in FL. All FL node, train model on processed local data and send result of model who performance is good between model without using pertained weights and model with per-trained weights. In second phase, select the FL node, whose performance is good at beginning and at last order of sequential learning.

Second approach: Master node sends the selected model and pre-trained weights in random order, or depends on data-size, node with more data size to node with low data size and then node increased data size. The paper approach this method.

## 5. *Modified Unet Model Architecture*

U-Net Architecture is most popular segmentation architecture. Most papers use Unet Architecture for image Segmentation [9,10]. Unet Architecture is influenced from encoder - decoder architecture. Encode and decode approach refines or reduces unwanted clutter in the image and hence good fit for segmentation task. UNet have 3 key components, Encoder Path, Decoder Path and connecting Path. The encode and decoder path need to have equal number of blocs and in Encoder Path the feature map size contacts and apposite happens in decoding path. The connecting path connects decoder and encoder path. Also skip connection from similar feature map is another key feature of Unet Architecture.

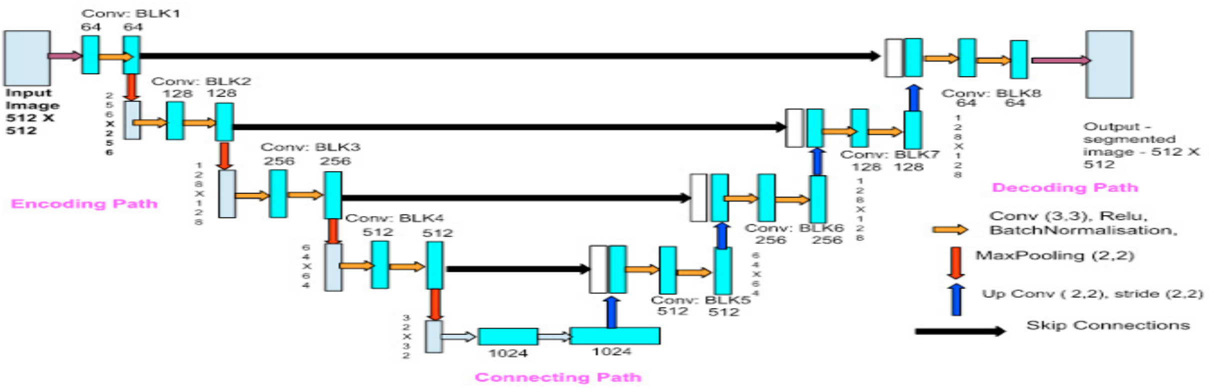



Fig. 2. Proposed Modified Unet Architecture

Unet Model - Detailed Model Architecture
Input: (512. 512, 1) OutPut: Conv2D ( 1, 1X1 Kernel, Sigmoid Activation)
Hyper parameters of the UNet Architecture, Optimiser : Adam with Learning Rate 1e-5 and Loss Function DICE COEF LOSS. Batch_Size = 2, EPOCH = 56.

To save the best performing model weights, ModelCheckpoint method used to save the weights of best performing model in terms of validation loss ( least value).

### 5.1. *Transfer Learning*

Transfer Learning performed on the proposed model with pre-trained weights of model on MRI Image Segmentation dataset. The output layer weights are is not transferable, input layer also excludes. Rest all layers are pre-trained with weights and set layers as trainable layer for further training of the model on new data. Paper attempted transfer weights layer by layer and setting layer by layer as trainable. Started Fine tuning with a few layers and upto all layers without last output layer.

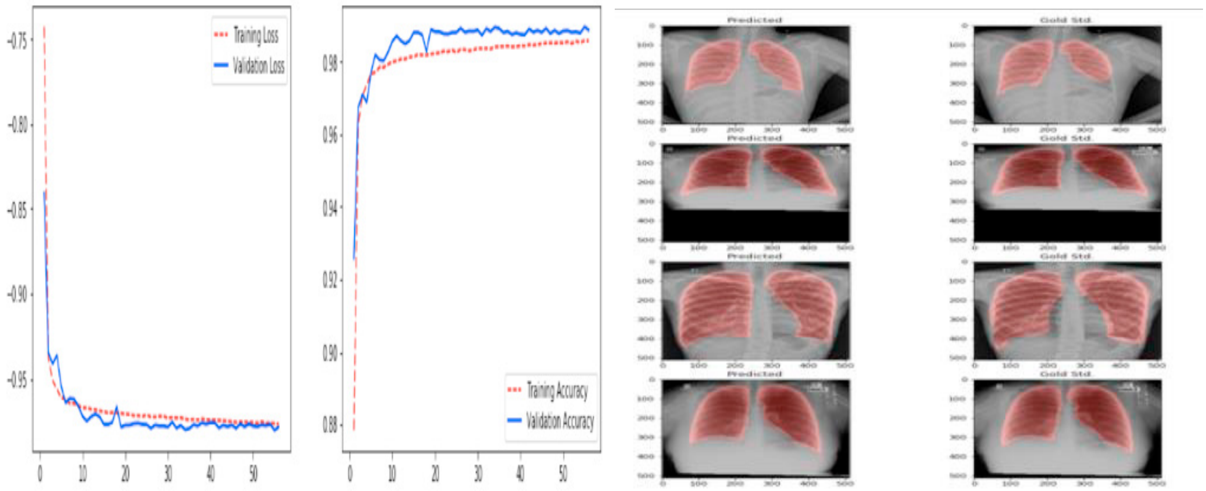
## 6. *Results and result analysis*

### 6.1. *Execution setup*

The Kaggle Notebook with GPU is used to execute the Segmentation Model. The Two datasets are added to the Kaggle notebook, one MRI Brain Tumour Segmentation dataset and X-Ray segmentation dataset.

### 6.2. *Performance of FL Nodes*

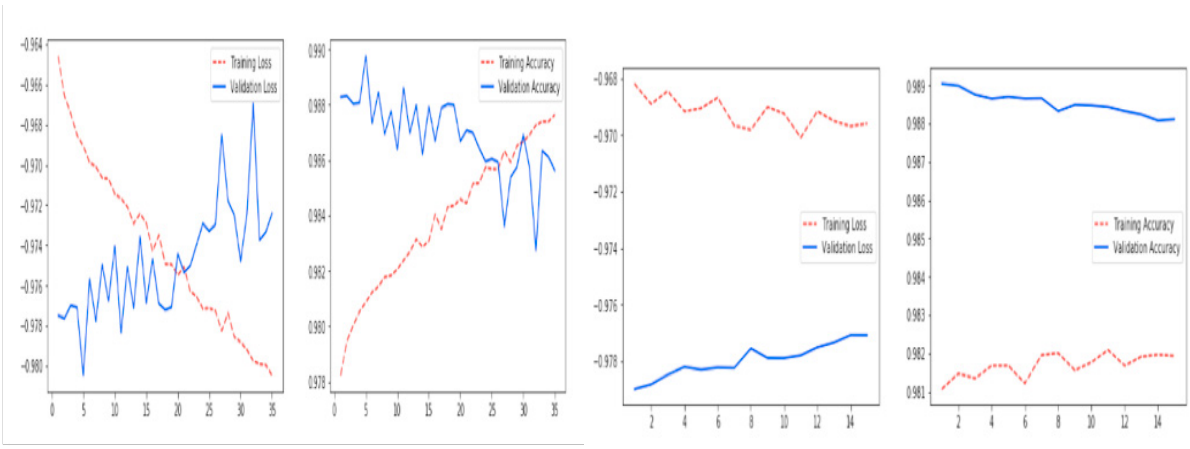
Instead of using 6 physical systems, the paper simulated 6 virtual nodes and full datasets are divided into 6 small chunks of the dataset. Model executed on small chunks of data and model weights are saved. In line with Federated learning instead of sending the model to each node, and collecting model learning, here, running the t model on small chunks and saving the model learning on the same system, which is used before model training on other data. Fine-tuned Transfer learning was used before model training on data at that specific node. Data processing and splitting the data into training and testing sets will be performed at each node.




**a** **b**

Fig. 3. a. Accuracy & Loss Function of Training and testing Data of FL Node 1. Right side plot, Accuracy Plot: X-axis : Epoch and Y-axis: Model Binary Accuracy range from 0 to 1. Left Side Plot, Loss Plot and X:axis Epoch and Y-Axis: Loss function value. b. Chest Lung Segmentation Results at FL Node 1. Fist Column pictures segments Output of Model, Second column pictures golden standard, actual expected result

The Initial Node improves the performance of Model and hence select the Node, which have all classes, and large datasets.

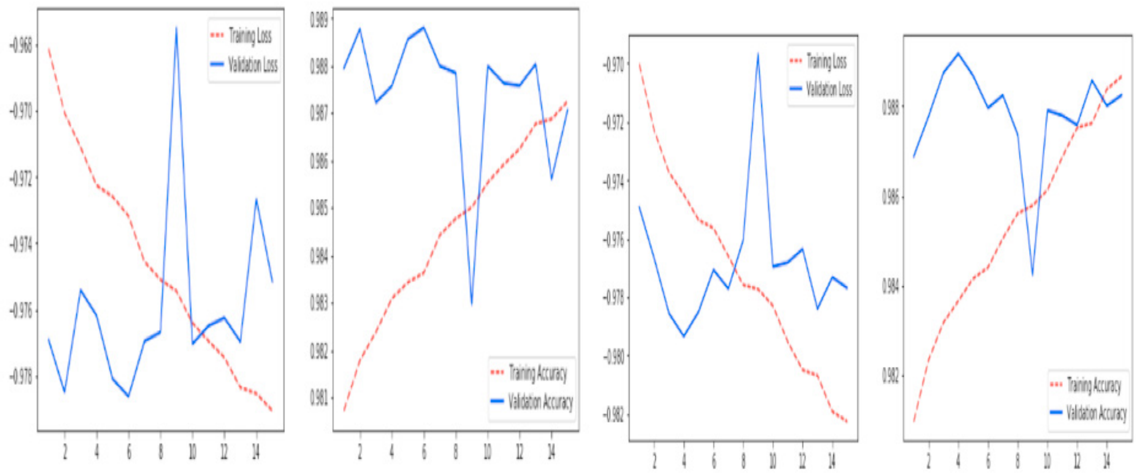



**a** **b**

Fig. 4. a. Accuracy & Loss Function of Training and testing Data of FL Node 2. Right side plot, Accuracy Plot: X-axis : Epoch and Y-axis: Model Binary Accuracy range from 0 to 1. Left Side Plot, Loss Plot and X:axis Epoch and Y-Axis: Loss function value. b. Accuracy & Loss Function of Training and testing Data of FL Node 3. Right side plot, Accuracy Plot: X-axis : Epoch and Y-axis: Model Binary Accuracy range from 0 to 1. Left Side Plot, Loss Plot and X:axis Epoch and Y-Axis: Loss function value

The learning improved for training data and learning decreased for validation data. It indicates an overfitting condition. Master FL Node, will discard this learning if there is a big gap between validation accuracy and training accuracy. In FL Node 2, the difference in accuracy is not big, and hence Master Node transfer FL Node 2 weights to FL Node 3.

Performance of FL Node 3 improved over FL Node 1 and observed no overfitting of model. The Master FL Node forward the learning of the Model to next FL Node.




**a** **b**

Fig. 5. a. Accuracy & Loss Function of Training and testing Data of FL Node 4 .Right side plot, Accuracy Plot: X-axis : Epoch and Y-axis: Model Binary Accuracy range from 0 to 1. Left Side Plot, Loss Plot and X:axis Epoch and Y-Axis: Loss function value. b. Accuracy & Loss Function of Training and testing Data of FL Node 5. Right side plot, Accuracy Plot: X-axis : Epoch and Y-axis: Model Binary Accuracy range from 0 to 1. Left Side Plot, Loss Plot and X:axis Epoch and Y-Axis: Loss function value

Performance of FL Node 4 improved over FL Node 1 and no overfitting of the model. The Master FL Node forwards the learning of the Model to the next FL Node.

Performance of FL Node 5 improved over FL Node 1 and no overfitting of the model. The Master FL Node forwards the learning of the Model to the next FL Node. Alternatively, the Master Node compares all FL Node performance as the master node keeps all the weights. Hence the Master Node selects the Best performing weights. Master Node also keeps track of the FL nodes, where performance has not improved, so that it can investigate further the dataset variation between the datasets of FL nodes, where performance improves and where performance drops.


Fig. 6. Accuracy & Loss Function of Training and testing Data of FL Node 6. Right side plot, Accuracy Plot: X-axis : Epoch and Y-axis: Model Binary Accuracy range from 0 to 1. Left Side Plot, Loss Plot and X:axis Epoch and Y-Axis: Loss function value

Table 1. Performance of All Nodes

| FL Node | Best Validation Loss | Best Val Accuracy | Perf Improved or Not |
|---|---|---|---|
| FL Node 1 | 0.9799 | 0.9894 | Yes (2nd) |
| FL Node 2 | 0.9805 | 0.9898 | Yes (1st) |
| FL Node 3 | 0.979 | 0.989 | No |
| FL Node 4 | 0.9786 | 0.9888 | No |
| FL Node 5 | 0.9794 | 0.9892 | Yes (3rd) |
| FL Node 6 | 0.9795 | 0.9892 | No |

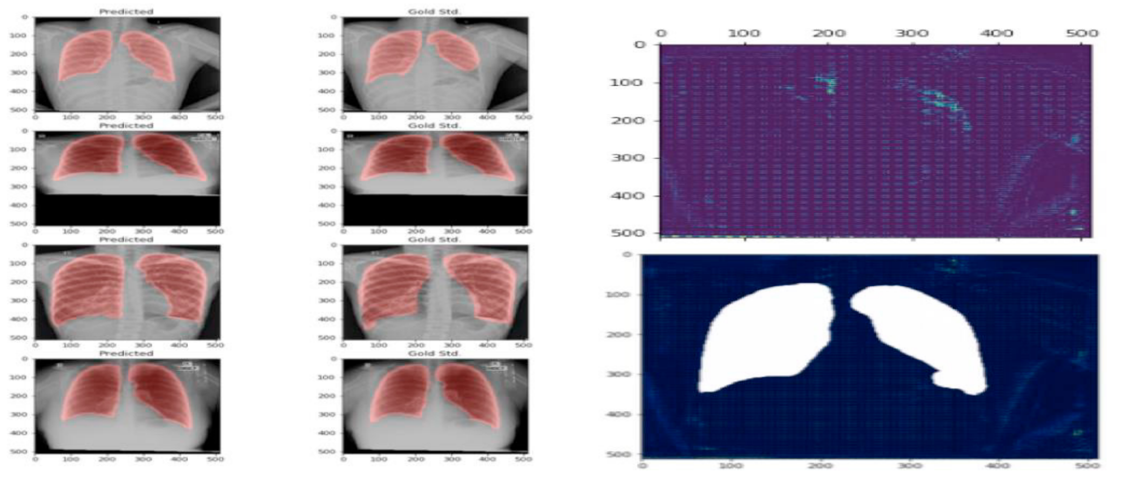
The paper[11] uses three datasets, two are same used here. The paper[11] computed result on individual datasets and performance comparison as mentioned in table 2

Table 2. **Overall Result comparison with A-LugSeg(2022)**

| Model | Dataset - Used | Accuracy |
|---|---|---|
| A-LugSeg | Montgomery County chest x-ray dataset (MC) | 0.97 |
| A-LugSeg | Japanese Society of Radiological Technology dataset (JSRT) + MC | 0.97 |
| proposed Our Model | MC + ShenZhen hospital Chest X-ray dataset (SZCX) | 0.9892 |

key findings are 1. In The Federated Transfer Learning method communication cost will be reduced, by avoiding transferring the weights of Node 3, 4 and 6 to master node. 2. Number of EPOCH essential for training UNet Model reduced, in federated Transfer learning, from **earlier 100+ epoch** to **single digit epoch. Overall learning time is reduced.** 3. The data-driven approach provides better performance even using fewer data compare to the model-driven approach as resulted in table 2

*The segmentation result of the best model is displayed for few sample validation data as follows*




a b

Fig. 7. a. Segmentation Results of Final Model - selected weights of best performing Model. Fist Column pictures segments Output of Model, Second column pictures golden standard, actual expected result. b. GradCAM based heat map of Segmented Image. X-axis and Y axis image size in pixels. Top Figure GradCAM heat map and bottom picture based on heat map selected relevant part of image.

GradCam[7] used for explaining the result of the segmentation, key portion of the Image, which got more important using heat map. Actually, segmentation doesn't need explain ability, as segmentation difference Image with Colour code in results shows the deviation of result with colour coded. The GradCam gives which area high important.

## 7. *Conclusion and Feature Scope*

Federated Transfer learning is best suited for any number of FL nodes and FL nodes with varying computational and connectivity capacities. The reason for selecting the FTL approach is data privacy and avoiding the complexity of data aggregation in a single place. The Lung segmentation model (modified Unet Model) is trained in a simulated Federated Transfer learning approach and achieves the best possible accuracy 98.92%. The key reasons for such improved performance 1. Leveraging MRI dataset and model is pre-trained with weights obtained from training modified Unet model on Brain segmentation data. 2. Splitting the data in specific ways as mentioned in the research methodology 3. Modified FTL sequential learning method. The algorithm keeps track of the sub-datasets or FL node, which provides the best performance, and which degrades performance. Then the master node can select the FL node learning and pass on to the next node. The key difference about FTL and other FL is, here not just the model is sent, model learning weights are also sent. The model performance compared to another recent model A-LunSeg model proves that data driven approach provides better results compared to model driven approach. The proposed method is data driven approach. Here the model used is Unet and added more layers to Unet and Batch normalization and hyper parameter tuning. But the most effort put in data organization, data processing, and effective utilization of transfer learning compared to the model-driven approach, most time put in customer building model for the dataset. The data-driven approach also includes providing quality, unbiased data, and a lot more, and will always speed up model development performance and provides better performance. The Federated Transfer learning approach is also useful for selecting which type of AI model architecture works well with datasets at a node and later analyzing data's impact on model performance and understanding more about datasets and refining the model.

Another observation is, not all FL node training improves the performance of the model and hence communication between Master FL and FL node can be improved by reducing communication by avoiding transferring the learning from FL Nodes. To perform that Master node has to send the model, model weights, and current max performance to FL node.

The future scope of this work is to improve the performance of the model. Experiments for performance improvements can be tried with select replacement of MRI dataset, by using various other model evaluation parameters like DICE coefficient and exploring option of class balance at each node using class balancing techniques.

## References

[1] H. Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, Blaise Agüera y Arcas,(2017) "Communication-Efficient Learning of Deep Networks from Decentralised Data", https://arxiv.org/abs/1602.05629.

[2] Jakub Konecn´y, H. Brendan McMahan , Daniel Ramage, Peter Richt´arik, (2016) "Federated Optimization: Distributed Machine Learning for On-Device Intelligence", https://arxiv.org/pdf/1610.02527.pdf

[3] Chaoyang He , Songze Li, Jinhyun So, Xiao Zeng, Mi Zhang, Hongyi Wang, Xiaoyang Wang, Praneeth Vepakomma, Abhishek Singh, Hang Qiu, Xinghua Zhu, Jianzong Wang, Li Shen, Peilin Zhao, Yan Kang, Yang Liu , Ramesh Raskar, Qiang Yang, Murali Annavaram∗ ,(2020) Salman Avestimehr, "FedML: A Research Library and Benchmark for Federated Machine Learning", https://arxiv.org/pdf/2007.13518v4.pdf

[4] Dimitrios Dimitriadis, Kenichi Kumatani, Robert Gmyr, Yashesh Gaur, Sefik Emre Eskimez, (2020) "Federated Transfer Learning with Dynamic Gradient Aggregation", https://arxiv.org/abs/2008.02452

[5] Jaeger S, Candemir S, Antani S, Wáng YX, Lu PX, Thoma G. Two public chest X-ray datasets for computer-aided screening of pulmonary diseases. *Quant Imaging Med Surg*. 2014;4(6):475-477. doi:10.3978/j.issn.2223-4292.2014.11.20

[4] R. Uppin, S. Ambesange, Sangameshwar, S. Aralikatti and M. G. V, "Respiratory Sound Abnormality Classification using Multipath Deep Learning Method," *2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA)*, 2021, pp. 903-915, doi: 10.1109/ICIRCA51532.2021.9544881.

[5] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 618-626, doi: 10.1109/ICCV.2017.74.

[5] Priyanka Rastogi, Kavita Khanna, Vijendra Singh, LeuFeatx: Deep learning–based feature extractor for the diagnosis of acute leukemia from microscopic images of peripheral blood smear,Computers in Biology and Medicine, Volume 142, 2022

[5] Rastogi, P., Khanna, K. & Singh, V. Gland segmentation in colorectal cancer histopathological images using U-net inspired convolutional network, Neural Computation and Application, pp. 1- 12, (2021). https://doi.org/10.1007/s00521-021-06687-z

[5] Ronneberger, O., Fischer, P., Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science(), vol 9351. Springer, Cham. https://doi.org/10.1007/978-3-319-24574-4_28

[5] Peng, T., Gu, Y., Ye, Z., Cheng, X., & Wang, J. (2022). A-LugSeg: Automatic and explainability-guided multi-site lung detection in chest X-ray images. *Expert Systems with Applications*, *198*, [116873]. https://doi.org/10.1016/j.eswa.2022.116873