Preserving Medical Data Privacy with Federated Learning

Adria Binte Habib
School of Data and Sciences
Brac University, Dhaka, Bangladesh
adria.habib@bracu.ac.bd

Abdullah Umar Nasib
School of Data and Sciences
Brac University, Dhaka, Bangladesh
abdullah.umar.nasib@g.bracu.ac.bd

Md. Sabbir Hossain
School of Data and Sciences
Brac University, Dhaka, Bangladesh
md.sabbir.hossain@g.bracu.ac.bd

Annajiat Alim Rasel
School of Data and Sciences
Brac University, Dhaka, Bangladesh
annajiat@gmail.com

***Abstract*—Privacy is an essential human right that is protected by data protection regulations in many regions around the world. These regulations are designed to ensure that people's personal information is handled with care and kept confidential. In the digital age, this becomes more crucial as people engage online and share more of their personal data. One area where data privacy is particularly important is in the medical sector. Medical data is highly sensitive, and patients have a right to expect that their data will be protected and used only for the intended purposes. However, there are many ways in which data privacy can be compromised, such as data breaches or unauthorized access. To address this issue, a federated learning based model can be used to deliver accurate results while maintaining the data privacy of other hospitals' data. Federated learning is a machine learning technique that allows multiple parties to collaborate on a machine learning model without sharing their data directly. In this approach, each hospital trains a local model on its own data and then sends only the model updates to a central server, where they are aggregated to create a global model. By using federated learning, the privacy of each hospital's data is maintained because the data never leaves the local device. Instead, the model updates are sent, which contain only information on the changes to the model's parameters.**

***Index Terms*—Federated Learning, Medical Data, Machine Learning, Data Mining.**

I. Introduction

Modern technologies introduced new technologies in our medical system that involves a significant amount of medical data in daily basis. The threads of data privacy violation is also tremendous along with the advancement of technology. This issue requires an special attention to deal with the privacy thread where Federated Learning comes into play. Like most other machine learning algorithms, Federated Learning is data hungry and that makes it a hot cake to work on protecting medical data.

Healthcare data often becomes fragmented due to the intricate nature of the healthcare system and its operations. For example, clinical records of patients within specific patient groups may only be accessible to certain hospitals. The records contain confidential health details of individuals and are subject to strict privacy laws such as the Health Insurance Portability and Accountability Act (HIPAA), which governs the handling and examination of such information [1]. Modern data mining and machine learning (ML) technologies, such deep learning, which normally needs a vast quantity of training data, are faced with a significant difficulty as a result.

Federated learning has emerged as a prominent paradigm due to its ability to leverage sensitive and distributed data for learning purposes. The approach involves training a global model using a central server, while the data remains stored in local institutions where it originated, rather than being aggregated from multiple sources at once or relying on the traditional approach of discovering and replicating data.

The successful implementation of federated learning (FL) can greatly enhance the practice of precision medicine on a large scale. It can lead to the creation of models that provide unbiased assessments, accurately represent an individual's physiology, account for rare diseases, and uphold privacy and governance protocols. However, FL requires careful technical scrutiny to ensure that the algorithm functions optimally while safeguarding patient privacy and safety. Nevertheless, it offers a solution to the limitations of methods that rely on a single centralized data pool.

The data receives additional protection because federated learning makes it feasible to jointly train a model with data from several users without letting any raw data leave their devices. This is the actual aim of our study. To obtain more accurate medical data while maintaining top-notch data privacy, we will train the medical photos from various clients.

II. Related Works

A noticeable amount of researches has already been done and published on Federated Learning which are also freely available.

In their work, Nicola Rieke and their team aimed to offer comprehensive insights and details to the community about the advantages and impacts of federated learning (FL) in medical applications [2]. They also highlighted the critical considerations and challenges associated with implementing

FL in digital health. Their findings indicated that FL has the potential to overcome the limitations of methods that rely on a single centralized data pool.

Federated learning is a machine learning method used to build a global model by aggregating local models trained on different devices without exchanging the raw data. In the biomedical sector, federated learning is utilized to protect sensitive patient data while still allowing multiple medical institutions to collaborate and improve their models. However, there are challenges encountered in using this technology, including statistical challenges related to the variability in data from different sources, system challenges related to the communication and synchronization between devices, and privacy challenges related to the need to keep patient data confidential. To address these challenges, Jei Xu and colleagues provided an overview of potential solutions for each challenge. For instance, to deal with statistical challenges, they suggested using transfer learning, data augmentation, and model aggregation techniques [3]. For system challenges, they proposed using communication-efficient algorithms, error detection and correction methods, and adaptive synchronization strategies. For privacy challenges, they proposed using encryption, differential privacy, and secure aggregation techniques to protect sensitive patient data. In addition, they explored the potential implications and uses of FL technology in healthcare, such as improved prediction accuracy, personalized medicine, and better disease detection. They also discussed the challenges that need to be addressed to ensure the widespread adoption of this technology in the healthcare sector. Overall, their study provides a valuable contribution to the ongoing efforts to advance FL technology in the biomedical field while addressing the challenges faced by practitioners.

Kaiyue and their team conducted an extensive survey that provided a concise overview of the latest federated learning techniques and their advancements [4]. The paper also addressed various unresolved issues in federated learning, along with existing solutions, while highlighting future research directions for the field.

Wei Yang Bryan Lim and their colleagues utilized dynamic contract design in [5] to develop a two-period incentive mechanism that upholds intertemporal incentive compatibility (IIC). This ensures that the contract's self-revealing mechanism holds true across both periods. Their performance evaluation confirmed that their contract design meets the IIC constraints and generates more profits compared to the uniform pricing scheme. These findings demonstrate the effectiveness of their approach in mitigating the adverse effects of information asymmetry.

Adnan Q. and their team conducted a research study that utilized the concept of clustered federated learning (CFL) to develop a collaborative learning framework for automatic multi-modal COVID-19 diagnosis [6]. Their approach was particularly suitable for this task, as it involved collecting visual data such as CT scans, X-rays, and ultrasound images from various medical centers. The collected data could then be used to develop a joint/shared machine learning model using cloud-edge infrastructure. The researchers' approach aimed to diagnose COVID-19 in X-ray and ultrasound images without the need to share data with a cloud or central entity. Instead, the CFL technique allowed the data to remain on the local devices while still enabling collaboration among multiple institutions. The clustered federated learning approach involved grouping the data from different centers into clusters and training local models on each cluster. These local models were then combined into a global model using a weighted averaging method, which allowed the model to learn from the differences in data across the various centers. The results of the study showed that the proposed framework using CFL achieved a high accuracy rate of 98% in diagnosing COVID-19 from X-ray and ultrasound images. The framework also demonstrated good scalability and robustness to noisy data. The researchers' approach provides a promising solution for developing a collaborative learning framework for COVID-19 diagnosis that is secure, privacy-preserving, and scalable. In conclusion, Adnan Q. and their team's research study utilizing clustered federated learning demonstrates the potential of collaborative learning frameworks in the medical field. Their approach of using CFL for COVID-19 diagnosis provides a practical solution for developing accurate and secure machine learning models while preserving the privacy of sensitive medical data.

III. Proposed Workflow

The proposed model consists of two main sub sections. The first one is to classification of data and the second one is to detection which are shown in Figure 1 and Figure 2 respectively.

The architecture diagram i.e Fig 1. explains how exactly the solution works at the system level. The steps to detect the most effective DL model for RBC abnormality classification are displayed here.

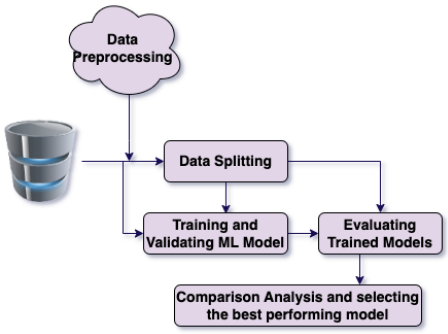

Fig. 1. Proposed Workflow of Classification of Data

In data preprocessing stage, firstly we have normalized the RBC images with Principal Component Analysis which is also known as PCA and modified them to a resolution of 128x128 pixels where it becomes more workable.

Secondly, we have split our available data in a ratio of 1:7:2 where 10% was utilized for the model's validation process where as 70% of the image was picked randomly for the purpose of training the models. The remaining 20% data was

used for validation task. Finally, in our research, we have trained and made comparisons between the three DL models (VGG19, ResNet50, and Inception v3).

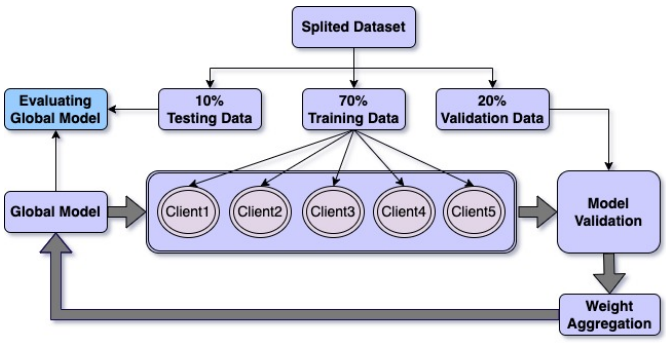

Fig. 2. Detailed Proposed Model

The experiment starts with the use of the global model named the VGG19, ResNet50, and Inception v3. Next, five clones of this global model is made to be sent to 5 separate clients. These clients, instead of working centrally, are decentralized and train the cloned models separately on their own separate data sets. After they are done running their models, the weights found from the training sessions are passed up to the central global model. After this, the 5 separate readings are merged and averaged to find a single value. After this, the global model is tested likewise.

IV. Input Dataset and Reprocessing

A. Data Collection and Data Analysis

We used a dataset that is a skin cancer dataset for our research. This dataset [7] was created by The International Skin Imaging Collaboration and has 2357 photos of both malignant and benign oncological illnesses. (ISIC). With the exception of melanomas and moles, which have a slightly higher representation in the photos, all other images were grouped based on their categorization using ISIC, and each subgroup contains an equal number of images. The dataset comprises various skin conditions, such as basal cell carcinoma, dermatofibroma, melanoma, nevus, pigmented benign keratoses, seborrheic keratoses, squamous cell carcinoma, and vascular lesions.

B. Data Preprocessing

As mentioned in the previous section, the dataset contains 9 distinct types of skin cancer information. To distribute the data among the 5 client-server systems used in this study, each type of skin cancer information was assigned to a specific client. The distribution was done in such a way that each client received an equal number of information for all types of skin cancer. However, in the earlier phase, the dataset was not evenly distributed, and thus there were some remaining pieces of information. To ensure that the distribution was even, the researchers chose to exclude the remaining information and considered it as outliers. Therefore, the data was only distributed among the clients in an equal and fair manner, without any bias or inconsistencies.

V. Machine Learning Module

A. Algorithms

1) VGG19: VGG19 is a convolutional neural network (CNN) architecture developed by the Visual Geometry Group (VGG) at the University of Oxford. It is a variant of the VGG family of models, which includes VGG16, VGG11, and others. The VGG19 architecture consists of 19 layers, including 16 convolutional layers and 3 fully connected layers. The convolutional layers are grouped into 5 blocks, each containing multiple 3x3 convolutional layers followed by a max pooling layer. The first two blocks have 2 convolutional layers each, while the remaining three have 4 convolutional layers each. The fully connected layers consist of a 4096-node layer followed by a 1000-node layer, which is the output layer for the ImageNet dataset.

We chose VGG19 because weights are readily available with other frameworks, like as Keras, and may be adjusted and applied in user satisfactory way. The layers of our model are given below in figure 3.

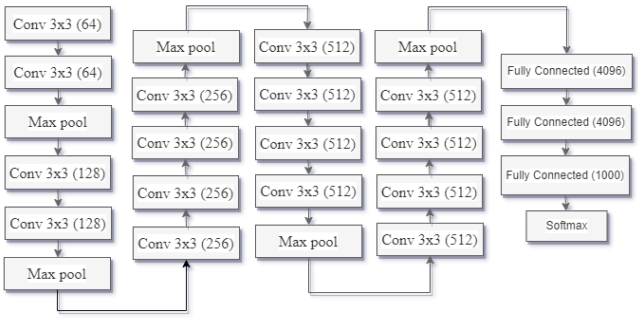

Fig. 3. Structure of our model VGG19

2) Resnet50: We selected this approach because this approach can also be used for jobs that are not computer vision-related to add depth and lower processing costs. The reason behind the lower processing cost is the deep residual learning framework. The original mapping becomes H(x):=F(x)+x as

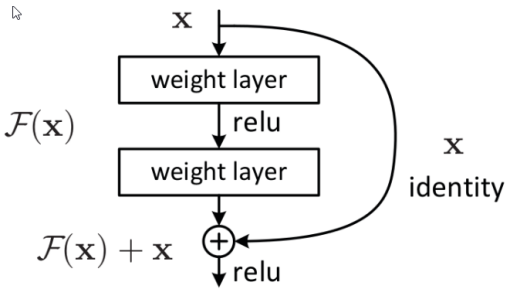

Fig. 4. Residual Learning Framework

can be seen in Figure 4 since they expressly allowed the layers to fit a residual mapping and denoted that as H(x).

They then allowed the non-linear layers to fit another mapping F(x):=H(x)x. The model did not require the addition of any new parameters, and the computational time was kept under control thanks to these shortcut identity mappings.

3) Inception v3: The Inception V3 model employs a range of techniques to optimize the network for better model adaptation. This model is an upgrade over the Inception V1 and V2 models in terms of efficiency and depth. Despite the additional layers, it maintains the same speed as its predecessors (Figure 5). One of the key features that sets Inception V3 apart is the use of auxiliary classifiers as regularizers (Figure 5). This approach helps to improve the accuracy of the model by adding additional training signals. Moreover, this method is computationally less expensive, which makes it ideal for applications that require real-time performance. Overall, the Inception V3 model offers significant improvements over its predecessors in terms of both accuracy and efficiency.

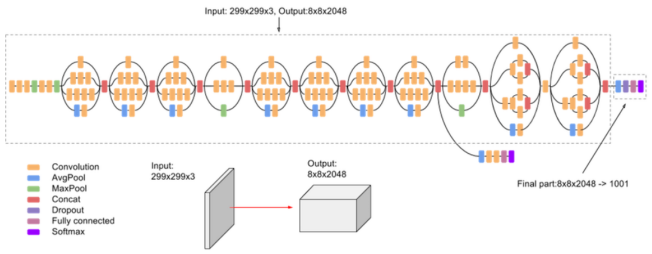

Fig. 5. Inception v3 Framework

VI. Result Analysis

For image identification applications, a deep convolutional neural network architecture called VGG19 is frequently employed. Although VGG19 is a strong architecture for image recognition, its suitability for federated learning depends on a number of variables, including the particular federated learning setup, the size and complexity of the dataset, and the available computational resources.

VGG19 is a computationally costly architecture in general, and training it on sizable datasets might demand a lot of computing power. In federated learning environments where data is dispersed across various devices with constrained computational resources, this can make it difficult to apply VGG19. It can be a successful architecture for federated learning, though, if the dataset is modestly sized and the devices taking part in the process have enough processing capacity. In these circumstances, the VGG19 architecture can be fine-tuned on the local data of each device, and the models that are created can be combined to create a global model that performs well on the entire dataset.

In general, the specific use case and the available resources determine whether VGG19 is appropriate for federated learning. It may not always be the greatest option for federated learning environments, despite the fact that it can be a powerful architecture for image recognition tasks. A deep convolutional neural network architecture called ResNet-50 has attained cutting-edge performance on various image identification applications. Because ResNet-50 balances model complexity and accuracy while remaining computationally efficient, it is a good architecture for federated learning in general. The capability of ResNet-50 to learn rich feature representations from images is one of its advantages for federated learning. This makes it ideal for applications like object detection and segmentation that call for fine-grained picture processing. Additionally, compared to other deep neural network architectures like VGG19, ResNet-50 has a smaller number of parameters, which can make it simpler to train on federated learning setups. This is due to the fact that federated learning requires building models on data that is spread over a number of devices, each of which has a finite amount of computational power. ResNet-50 is simpler to deploy in federated learning contexts since it takes less CPU and memory to train thanks to its fewer parameters. Because ResNet-50 balances model complexity and accuracy while being computationally effective, it is a good architecture for federated learning overall. It is ideal for federated learning setups since it can learn rich feature representations and has a manageable number of parameters.

Another deep convolutional neural network design frequently employed for image identification tasks is Inception v3. It is renowned for its effective use of computational resources and has attained state-of-the-art performance on numerous image recognition benchmarks. Generally speaking, depending on the particular use case and the resources available, Inception v3 can be a good architecture for federated learning. It is particularly suited for applications that need fine-grained image analysis, such as object detection and segmentation because it can learn multi-scale feature representations. Additionally, Inception v3 is more computationally efficient and simpler to train on federated learning setups than other deep neural network architectures because it has a smaller number of parameters than other deep neural network architectures. In federated learning contexts, when data is dispersed across numerous devices with constrained processing capabilities, this can be very crucial.

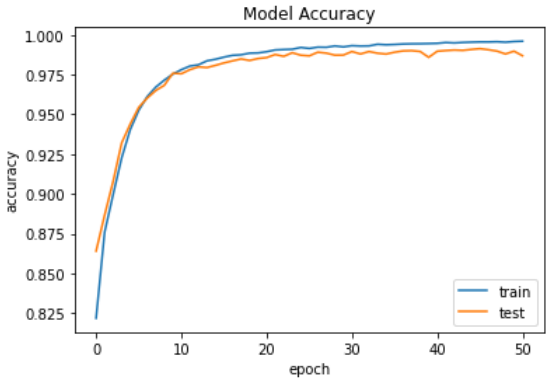

Fig. 6. Learning curve of Resnet50

It's important to remember that Inception v3 can still be a sizable and computationally demanding architecture, making

it unsuitable for all federated learning settings. Other, more computationally effective topologies, like MobileNet or ShuffleNet, may be more suitable in these circumstances. Inception v3 can, in general, be a useful architecture for federated learning, especially when fine-grained picture analysis is needed. However, other architectures might be more suitable in some circumstances. It depends on the specific use case and the resources available.

Regarding federated learning, each of the three architectures has its own advantages and disadvantages. It can be difficult to apply VGG19 in federated learning setups with restricted resources since it can be computationally demanding. VGG19 is a powerful architecture for image recognition problems. ResNet-50 is a powerful image recognition architecture that is also computationally efficient, which makes it ideal for federated learning environments. Given its reputation for learning multi-scale feature representations and computational efficiency, Inception v3 is a good option for some federated learning scenarios.

In our case, Resnet50 performed best out of the other models. It performed with 99% accuracy. In the below figure, the learning curve of the model can be seen. The model accuracy can be detected at around 99% here in figure 6.

Since the dataset consists of over 2000 images with colored images of the cancerous cells on the skin. In other words, the area of the dark skin infected. I have 32 gigabytes of ram along with a 2060 super Nvidia GPU. With a dataset consisting of over 2000 images of skin cancer cells and a 2060 super Nvidia GPU with 32 gigabytes of RAM. So with this setup and not well-grained image details that are critical to the classification of the skin cancer cells, resnet50 performs well. The reason behind it's performance is the residual learning framework.

VII. Conclusion

Protecting data privacy is crucial, especially in sensitive fields like healthcare, as it is crucial to maintain trust in online interactions. Federated learning is a promising solution that can help achieve accurate results while preserving data privacy. With the help of federated learning, data can be kept local and not shared, which can address the issue of compromised privacy in the medical sector. The future of federated learning is bright, with ongoing research exploring new ways to improve model accuracy, reduce communication costs, and address security concerns. As the need for privacy-preserving machine learning continues to grow, federated learning is poised to play an increasingly important role in the development of advanced AI models that can be deployed in various industries.

References

[1] L. O. Gostin, "National health information privacy: regulations under the health insurance portability and accountability act." [Online]. Available: 10.1001/jama.285.23.3015

[2] K. Clarkson and M. Reza, *Was That Sarcasm? -A Survey of Machine Learning Models for Classifying Sarcastic Comments on Reddit Using Word Embeddings*, 12 2018.

[3] J. X. . B. S. . G. . C. S. . P. W. . J. B. . F. Wang1, "Federated learning for healthcare informatics," *Journal of Healthcare Informatics Research (2021)*, pp. 1–19, 11 2020. [Online]. Available: https://doi.org/10.1007/s41666-020-00082-4

[4] K. Z. . X. S. . C. Z. . S. YU, "Challenges and future directions of secure federated learning: a survey," *PeerJ. Computer science*, vol. 12, 2021. [Online]. Available: https://doi.org/10.1007/s11704-021-0598-z

[5] W. Y. B. L. . S. G. . Z. X. . D. N. . C. L. . C. M. . M. Guizani, "Dynamic contract design for federated learning in smart healthcare applications," *IEEE Internet of Things Journal*, vol. 8, pp. 16 853–16 862, 10 2020. [Online]. Available: 10.1109/JIOT.2020.3033806

[6] A. Q. . K. A. . M. A. A. . A. A.-F. . J. Qadir, "Collaborative federated learning for healthcare: Multi-modal covid-19 diagnosis at the edge," *IEEE Open Journal of Computer Society*, vol. 3, pp. 172–184, 09 2022. [Online]. Available: 10.1109/OJCS.2022.3206407

[7] A. Katanskiy, "Skin cancer isic," Aug 2019. [Online]. Available: https://www.kaggle.com/datasets/nodoubttome/skin-cancer9-classesisic