The Comprehensive Investigation of Federated Learning with Its Application in the Medical Image Analysis

Wenxiao Zeng[a]
Computer Science, University of Massachusetts Amherst, Amherst, U.S.A.

Keywords: Federated Learning, Medical Image Analysis, Automatic Diagnosis, Data Privacy.

Abstract: Due to the pandemic in 2019 and the causing medical system crush, it has become necessary for the nations to increase the system capacity of health care services. By combining with machine learning, the newly developed automatic diagnosis can largely save the manpower and time cost required for the current medical systems, thus increase their overall efficiencies. The Federated Learning algorithm, based on the background of rising demand for automatic diagnosis and data privacy, is becoming widely-applied in the medical diagnosis, especially in the sophisticated medical image analysis. This study overviews the current researches of the Federated Learning algorithms in the health care system, including the diagnosis prediction of the Federated Learning-based models towards three specific types of diseases. The study discusses the advantages of Federated Learning in data privacy and heterogeneous massive data processing by the architecture of its workflow. On the other hand, its potential drawbacks are the lack of interpretability and applicability, which can possibly be solved or improved by SHapley Addictive exPlanation (SHAP) algorithm and Dynamic Weighting Translation Transfer Learning (DTTL) algorithm. Its potential safety issue by data transmission, however, though being minimized by the decentralized computation architecture of the Federated Learning, can hardly be fully removed unless the fully distributed algorithm will have developed and replaced its application in the future.

1 INTRODUCTION

The recent sustain and rapid development of Artificial Intelligence (AI) has induced its application in multiple novel environments (Hunt, 2014). On the other hand, under of the tendency of aging society in the world's major industrialized and industrializing states, and also after experiencing the crisis of medical resource run caused by the COVID-19 pandemic in 2019, it has become essential for these states' governments and medic facilities to seek out a scheme of enhancing their current medical systems' efficiency, in response to the pressure leading by the society's rising demand for medic treatment. Naturally, the application of AI algorithms as a type of medic assistant, aiming for providing prediction based on medical record data and thus shortening the time and manpower required for diagnosis, has deservedly become a popular method for this purpose. However, its model-training process also raises a requirement of data privacy to the training algorithms by both the patients and the facilities who build up the dataset's servers, in order to avoid the leakage of sensitive data.

Federated Learning (FL) is a deep-learning algorithm that involves a local model-training process in remote devices or data center, then sends the parameters of locally-trained model to the center server. Thus, it is able to ensure a certain degree of privacy since no raw data is sent to the center server, and also is able to process large-scale and heterogeneous data (Li et al., 2019).

In the practical application, the characteristic of the Federated learning algorithm in processing heterogeneous data makes it adept at the diversity of medical record data, giving it the potential of assisting human doctors in disease diagnosis, or even establishing an automatic analysis system in the near future. As a common and important type of medical record data, the medical images in disease diagnosis, such as ultrasound, X-rays,and Magnetic Resonance Imaging (MRI), are frequently utilized for the Federated Learning model for its importance in

[a] https://orcid.org/0009-0006-2511-4479

Zeng and W.
The Comprehensive Investigation of Federated Learning with Its Application in the Medical Image Analysis.
DOI: 10.5220/0013527800004619
In *Proceedings of the 2nd International Conference on Data Analysis and Machine Learning (DAML 2024)*, pages 532-536
ISBN: 978-989-758-754-2

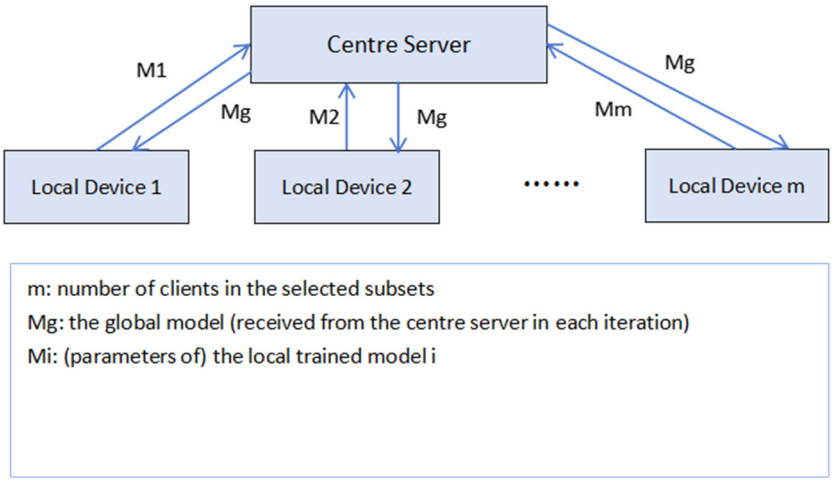

Figure 1: The structure of federated learning (Photo/Picture credit: Original).

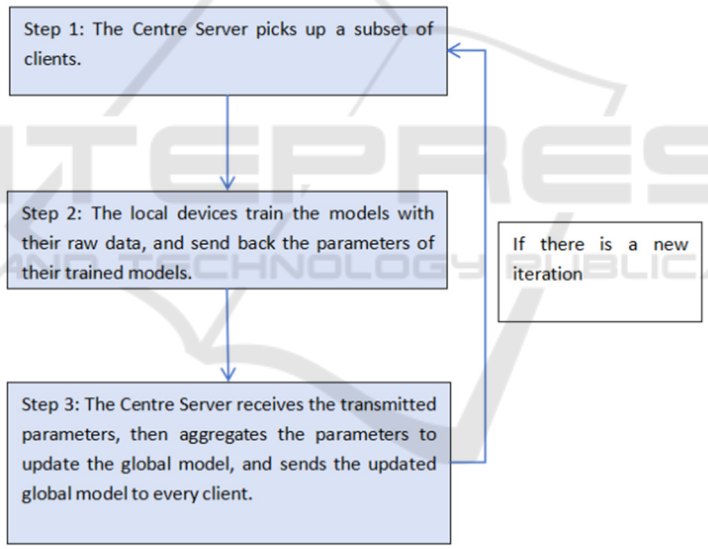

Figure 2: The workflow of federated learning (Photo/Picture credit: Original).

diagnosing many ailments and diseases such as cancers (Nazir et al., 2023).

The confidentiality endowed by the Federated learning algorithm, on the other hand, makes it appropriate for privacy protection in the hospitals, since it avoids the large and common restriction of direct access to these image collections by training the model locally with raw data and sending only the parameters of the local model to the center server for the final training (Silva et al., 2023).

Considering the significance of medical images to disease diagnosis, and the massive potential of Automatic Medical Image Analysis in improving the therapeutic outcome and medical system efficiency by the Federated Learning trained models (Silva et al., 2023), it is necessary to offer an overview that summarize the current research and application of Federated Learning in medical image analysis, including its practical application cases, contribution, limitation, challenges, and further potentials in the

future. In the following sections, the article shall present them with elaboration.

2 METHOD

2.1 Preliminaries of Federated Learning

The Federated Learning shown in Figure 1 and Figure 2 is able to avoid the two problems, massive processing and data privacy, by its unique workflow. In each of its iterations, it selects a subset among its thousands or millions of clients, computes a local model in the local device, then sends the parameters of this locally trained model to the central server for further aggregation, and finally sends back the aggregated global model to each client and starts a new iteration (McMahan et al., 2023). The local training process keeps raw data stationary in the local device or service and is not shared by other clients or transmitted the central server, therefore the Federated Learning can guarantee the privacy of clients. Furthermore, the local training process avoids computation with heterogeneous type of data, since the data type within the same client device is more likely to be homogeneous. Meanwhile, the partition and distribution of model-training procedure into clients' local device allows Federated Learning to dispose of enormous amount of data without being limited by the number of available worker nodes.

2.2 Federated Learning-Based Predication in Diagnosis

2.2.1 COVID-19

In the study of Darzi et al, the Federated Learning methods are utilized for COVID-19 detection (Darzi et al., 2024). The researchers adopt the datasets from Public Hospital of Sao Paulo (HSPM), Brezil, and Tongji Hospital of Wuhan, China, and adopt Convolutional Neural Networks (CNNs) as the mode type for their study. As a type of respiratory infection, the images used for training are chest CT-scans from both COVID-19 positive and healthy subjects, resized to 224×224 pixels with interpolation (Darzi et al., 2024). As a comparative study, the researchers select five different Federated Learning algorithms—Centralized Data Sharing (CDS), Federated Averaging (FedAvg), Federated Stochastic Gradient Descent (FedSGD), Single Weight Transfer (SWT), and Stochastic Weight Transfer (STWT)—and compare their performances under different numbers of Federated iterations and clients (Darzi et al., 2024). In addition, the study also provides an insight into the challenge of catastrophic forgetting, which states the forgetting of previous information when a learning is training upon new datasets, calling for a future method for solving this issue (Darzi et al., 2024).

2.2.2 Brain Tumor

Brain Tumor is an aggregation of abnormal cells within the brain. As a type of disease that exists in the most delicate and complex structure of the human body, it is challenging to diagnose, confirm its location in which section of the brain, and judge its characteristic of being benign or malignant. However, it is crucial for this classification since it determines the specific therapy required to be taken by the patient, while this classification in the diagnosis requisites high accuracy for any error occurring in the diagnosis and surgeries might cause catastrophic and permanent brain damage to the patient. The new research by Albalawi et al aims to develop a new model for the classification of Brain Tumor by utilizing Federated Learning in the training procedure (Albalawi et al., 2024). The researchers adopted the modified VGG16 architecture, a type of CNN model optimized for brain MRI images, for training the resized 128×128 pixels images in their datasets (Albalawi et al., 2024). The main objective of this research is to enhance the accuracy of brain tumor classification, therefore the dataset is composed of 4 types of MRI images: Glioma, Meningioma, No Tumor, and Pituitary.

2.2.3 Skin Cancer

As a frequently occurred disease in the environments of highly intensive direct sunlight or radiation, the early-stage skin cancers are very likely to be confused with normal skin inflammation by their similar outer symptoms, and cause the patients to miss the optimal treatment period. Therefore, the effective diagnosis of early skin cancer is critical for the patient's later therapy and potential recovery. It is research made by AlRakhami et al that aims for achieving an automatic diagnosis of early skin cancer, constructed by the Federated Learning and through the architecture of Deep Convolutional Neural Networks (DCNNs) (AlRakhami et al., 2024). The researchers utilize three heterogeneous datasets: ISIC 2018 Dataset, PH2 Dataset, and Combined Dataset, for training and testing the outcomes of model prediction, and thus compare their effectiveness (AlRakhami et al., 2024). During the experiments, the researchers would perform the model training in different centralized or federated architectures in order to observe the results of effectiveness in different index of evaluation metrics (AlRakhami et al., 2024).

3 DISCUSSIONS

Admittedly, in past research and experiments the Federated Learning have presented its superiority in privacy and heterogeneous data processing and have maintained a relatively high level of accuracy in most of its application circumstances. However, its potential drawbacks and challenges have also been discovered in these experiments, by either certain features of the outcome data or the mathematical structure of its foundational workflow. For some of the illustrated issues, there has been modification or expansion for the Federated Learning algorithms towards their corresponding resolution; for some others, however, it can only be expected to be solved through the further development of other deep learning algorithms.

One of the most common and directly meet problems is that the models trained by the Federated Learning algorithms lack interpretability. It is admittedly that the Federated Learning trained models can often attain great performance in the prediction accuracy, but the researchers can hardly comprehend how and why this is managed to happen since the complexity of the neural network, the commonly used architecture for the models trained by the Federated Learning algorithms, makes it challenge to be understood by human researchers. Such drawbacks can be non-problematic in the result-orientated applications; for instance, it is not very necessary for human researchers and programmers to fully understand why and how an image-recognition model can tell all the cat images from the mixed image dataset, as long as it has achieved its task by attaining acceptable accuracy. Nevertheless, the researches in medic and biology often acquires the understanding of the mechanism behind the diagnosis, no matter if it is aimed for adding the credibility for the diagnosis or provoking further and more subtle researches towards these newly discovered mechanism. In this instance, unfortunately, the Federated Learning algorithm can offer little aid for the potential further researches.

One possible method in adding more interpretability to the Federated Learning algorithm is fusing it with SHapley Addictive exPlanation (SHAP) algorithm. In the study of Lundberg et al, it is firstly mentioned that Shapley values can be introduced in the field of machine learning (Lundberg et al., 2017). The SHAP algorithm can give out a contribution value (SHAP value) of each attribution for the eventual outcome of model prediction accuracy, which can be both positive for the contribution of reaching higher model accuracy and negative for the opposite situation. If the researchers are able to combine Federated Learning with SHAP, it is possible for the researchers to achieve model explainability by analyzing the SHAP values of all the attributes, considering the fact that there have been cases of SHAP algorithm being applied separately in the domain of medical diagnosis.

Another existing drawback of the Federated Learning that is due to its workflow is the lack of applicability in all circumstances, a concession made for its universality. What has caused this issue is the difference in data distribution between heterogeneous datasets: the decentralized characteristic of its training process allows the Federated Learning algorithm to process datasets with heterogeneous data types, but the performance of the trained model can also be impaired by their difference, making the prediction accuracy of the aggregated model in an epoch lower than the accuracy of the locally trained model in the application of that particular dataset. Some of researchers who have discovered this phenomenon in their experiments also call it “forgetting” (Darzi et al., 2024), stating that when receiving the heterogeneous data from a new dataset for further training, the model “forgets” some of the information from the previous old dataset and thus experience a decrease in its predication accuracy for the old dataset.

In response to this problem, the solution proposed by Yu et al in their research is the Dynamic Weighting Translation Transfer Learning (DTTL), which builds up a dynamic translation between imbalanced classes in two domains and minimizes the cross-entropy between the domains to reduce domain difference (Yu et al., 2024).

Although frequently being described as a type of decentralized deep learning algorithm, as a matter of fact the Federated Learning algorithm is only a type of partially distributed and coordinated algorithm, which still requires a centre server as a coordinator for model updates. Though still safer than the full centralized algorithm, the remain centralized architecture of the Federated Learning can lead to potential safety loopholes, especially in the transmission from local devices to the centre server. If being intercepted and decrypted, the parameters of the locally trained model from a specific dataset are exposed to the hacker; though it is not the raw data that is disclosed, the hackers might still be able to learn the features of data distribution in this dataset and manage to guess out a portion of the raw data based on the parameters disclosed.

Unless being replaced by the development of fully distributed algorithms, this potential safety issue can always be problematic and hardly be removed, for it relays on the architecture of Federated Learning’s workflow. A possible improvement for this issue is adopting more complex and secure encryption for the transmission; however, it would surely increase the

time cost of encryption and decryption during the data transmission, and thus reduce the time efficiency of the Federated Learning.

4 CONCLUSIONS

In summary, based on the architecture of its workflow, the Federated Learning has successfully achieved the requisites of an automatic diagnosis system used in the domain of health care services: massive and heterogeneous data procession, and data privacy. This study illustrates this architecture, the fundamental mechanism of Federated Learning, and its applications in the medical image analysis that serves for the researches and development of AI diagnosis towards certain specific diseases. While reviewing these researches, the study has summarized the defects of Federated Learning algorithm that has appeared in the past experiments, the lack of interpretability and applicability, and also suggested the theoretical feasible solutions towards each of the two drawbacks. However, it is also suggested that the safety issue laying in the data transmission can be ameliorated but hardly removed by the Federated Learning's partially distributed architecture, while the development of fully distribution algorithms and their potential replacement of the Federated Learning algorithms in the future might be the eventual resolution of this issue.

REFERENCES

Albalawi, E., T.R., M., Thakur, A., et al. 2024. Integrated approach of federated learning with transfer learning for classification and diagnosis of brain tumor. BMC Medical Imaging, 24(1), 110.

AlRakhami, M. S., AlQahtani, S. A., & Alawwad, A. 2024. Effective skin cancer diagnosis through federated learning and deep convolutional neural networks. Applied Artificial Intelligence, 38(1).

Darzi, E., Sijtsema, N. M., & van Ooijen, P. M. A. 2024. A comparative study of federated learning methods for COVID-19 detection. Scientific Reports, 14(1).

Hunt, E. B. 2014. Artificial intelligence. Academic Press.

Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. 2019. Federated learning: Challenges, methods, and future directions. IEEE Signal Processing Magazine, 37(3), 50-60.

Lundberg, S., & Lee, S. I. 2017. A unified approach to interpreting model predictions. arXiv preprint arXiv:1705.07874 [cs.AI].

McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & Aguera y Arcas, B. 2023. Communication-efficient learning of deep networks from decentralized data. arXiv preprint arXiv:1602.05629v4 [cs.LG].

Nazir, S., & Kaleem, M. 2023. Federated learning for medical image analysis with deep neural networks. Diagnostics, 13(9).

Silva, F. R. da, Camacho, R., & Tavares, J. M. R. S. 2023. Federated learning in medical image analysis: A systematic survey. Electronics, 13(1).

Yu, C., Pei, H., & Sparavigna, A. C. 2024. Dynamic weighting translation transfer learning for imbalanced medical image classification. Entropy, 26(5).