Dynamic Fusion based Federated Learning for COVID-19 Detection

Weishan Zhang¹, Tao Zhou¹, Qinghua Lu², Xiao Wang³, Chunsheng Zhu⁴, Haoyun Sun¹, Zhipeng Wang¹, Fei-Yue Wang³
¹College of Computer Science and Technology, China University of Petroleum (East China), China
²Data61, CSIRO, Australia
³State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China
⁴SUSTech Institute of Future Networks, Southern University of Science and Technology, China

Medical diagnostic image analysis (e.g., CT scan or X-Ray) using machine learning is expected to be an efficient and accurate way to detect COVID-19 infections. However, sharing diagnostic images across medical institutions is usually not allowed due to the concern of patients' privacy, which causes the issue of insufficient datasets for training the image classification model. Federated learning is an emerging privacy-preserving machine learning paradigm which produces an unbiased global model based on the received updates of local models trained by clients without exchanging clients’ local data. However, the default setting of federated learning introduces huge communication cost of transferring model updates and can hardly ensure model performance when data heterogeneity of clients heavily exists. To improve communication efficiency and model performance, in this paper, we propose a novel dynamic fusion based federated learning approach for medical diagnostic image analysis to detect COVID-19 infections. First, we present architecture for dynamic fusion based federated learning systems to analyse medical diagnostic images. Second, we design a decision making mechanism for clients to decides each round’s participation based on the local model performance. Third, we propose an aggregation scheduling method to dynamically select the participating clients based on each participating client’s training time. Fourth, we summarise a category of medical diagnostic image datasets for COVID-19 detection, which can be used by the machine learning community for image analysis. The evaluation results show that the proposed approach is feasible and performs better than the default setting of federated learning in terms of model performance, communication efficiency and fault tolerance.

Index Terms—Federated learning, machine learning, image processing, classification, COVID-19, architecture, AI, CT, X-Ray.

I. INTRODUCTION

The COVID-19 pandemic has introduced an unprecedented global crisis. The rapidly increasing number of COVID-19 cases leads to a severe shortage of test kits and calls for an efficient and accurate way to diagnose COVID-19 infections. To address the issue of the shortage of test kits for COVID-19 diagnosis, researchers have been working on machine learning technologies, especially deep learning, using medical diagnostic images (e.g., CT scan or X-Ray). The model performance is heavily dependent on the training dataset size and diversity. However, like other real-world machine learning driven systems, data hungersis is a critical challenge due to the concern for data privacy. To protect patients’ privacy, sharing medical data, including diagnostic images, across medical institutions is not allowed, which causes the issue of insufficient datasets for model training.

The concept of federated learning was introduced by Google in 2016 as a new machine learning paradigm in a way that produces an unbiased model while preserving data privacy [1], [2]. In each round of training, clients (e.g., organisations, data centers, or mobile/IoT devices) are selected to train a model using local data and send the updates of local models to a central server for aggregation without transferring any local raw data. Instead, only model updates are sent to the central server to preserve data privacy. The central server aggregates the received local models and produces a global model.

Federated learning has the potential to connect isolated medical institutions and train a model for COVID-19 positive case detection in a way that preservers data privacy. However, the default setting of federated learning introduces huge communication cost of transferring model updates (e.g. massive matrices of weights) and does not work well when data heterogeneity of clients heavily exists. To improve communication efficiency and model performance, in this paper, we propose a novel dynamic fusion based federated learning approach for COVID-19 positive case detection. Each client assesses the local model trained and only uploads the model updates when it performs better than the previous version through local assessment. The central server configure the waiting time for each client to send model updates based on all the participating clients’ training time. The evaluation results show that the proposed approach achieves better detection accuracy, fault tolerance and communication efficiency compared to the default setting of federated learning.

The contributions of the paper are as follows:

• We present architecture for dynamic fusion based federated learning systems for medical diagnosis image analysis to detection COVID-19 positive cases. The proposed architecture provides a systematic view of system interactions and serves as a guidance for the design of federated learning systems;
• To improve communication efficiency and model performance, We design a decision making mechanism for clients to decide each round’s participation based on the local model performance.
• We propose an aggregation scheduling method to dy-
We summarise a category of medical diagnosis image datasets for COVID-19 detection, which are collected from a variety of open data sources and can be used by the machine learning community for image analysis.

The remainder of this paper is organized as follows. Section II presents the approach. Section III evaluates the approach. Section IV discusses the related work. Section V concludes the paper and points out the future work.

II. DYNAMIC FUSION BASED FEDERATED LEARNING

In this section, we present a dynamic fusion based federated learning approach for CT scan image analysis to diagnose COVID-19 infections. Section II-A provides an overview of the architecture and discusses how the components and their interactions. Section II-B discusses a dynamic model fusion method to dynamically decide the participating clients and schedule the aggregation based on each participating client’s training time.

A. Architecture

Fig. 1 illustrates the architecture, which consists of two types of nodes: central server and clients. The central server initialises a machine learning job and coordinates the federated learning process, while clients trains a model specified in the learning job using local data and computation resources.

Each client gathers images scanned by the diagnostic imaging equipment through the client data collector and cleans the data (e.g., noise reduction) via the client data pre-processor. All the collected client data are stored in the database of local data hosted by the client. The job creator initialises a model training job (including initial model code and number of aggregation) and configures the initial value of waiting time for clients to return the updates of local model trained. Each participating client downloads the job and trains the model via the model trainer. After a set number of epochs, the model trainer completes this round of training and uploads the training time to the central server. The aggregation scheduler updates the waiting time based on the training time received from participating clients.

The local model assessor on each client compares the current local model trained with the previous version. If the current local model performs better, the client sends a request for model upload to the central server. Otherwise, the client sends a request not to upload the model update for this round. All the clients which do not complete the set number of epochs within the current waiting time are not allowed to participate the aggregation. After the set waiting time, the aggregation scheduler on the central server notifies each client which has sent the upload request to upload the model updates. After aggregation, the global model assessor measures the accuracy of the aggregated global model and sends the global model back to each client for a new round of training.

B. Dynamic Fusion

To improve communication efficiency in federated learning, the proposed dynamic fusion method consists of two decision making points: client participation and client selection. On the client side, each client decides whether to join this round of aggregation based on the performance of new model trained. After completing the set number of epochs, each client assesses the accuracy of the model trained by comparing with the previous version. If the new model performs better, the client sends the participation request to the central server for aggregation. Otherwise, the client requests for skipping this round of aggregation. On the central server side, the model aggregator determines the participating clients based on the waiting time which is the average training time of local models. If a client does not upload the model updates by waiting time, it is excluded by the central server for this round of aggregation. The waiting time of current round is calculated by averaging the previous round’s training time of each client. The initial waiting time is configured by the platform owner.

Fig. 2 illustrates the process of the proposed dynamic fusion method and Algorithm 1 describes the detailed process. The
process starts with creating a learning job by the central server. All the clients download the job from the central server and sets up the local training environment. From the second round, a timer is set for each client based on the average training time of all the participating clients for last round. If a client does not complete the training within the configured time, the central server proceeds the aggregation without any input from this client for this round. On the other hand, if the model trained by the client for this round performs worse than last round, the client sends a request to the central server for skipping this round’s aggregation. Otherwise, the client notifies the central server to upload this round’s model updates.

III. Evaluation
The proposed approach is evaluated via quantitative experiments using three datasets as shown in Table I which include 746 CT images and 2960 X-ray images. The CT dataset has 349 CT images containing 349 images of COVID-19 positive
cases and 397 images of negative cases. The chest X-ray images are from two datasets. The first X-ray dataset has 2905 images which contains 219 images of COVID-19 positive cases, 1341 images of negative cases and 1345 images of viral pneumonia (VP). The second X-Ray dataset consists of 55 COVID-19 positive cases, 200 negative cases, and 200 virus pneumonia. Please note that the CT images are taken from top, while the X-Ray images are taken from the front.

A. Performance of Default Setting of Federated Learning

We evaluate the accuracy of the default setting of federated learning using three different models, ResNet50, ResNet101, and GhostNet. The models were trained using the six groups of datasets shown in Table III. We compared the accuracy of default setting of federated learning (Default_FL) with local learning (LL) which predicts using the model trained on the local dataset. GFL federated learning framework was used for the default setting of federated learning in our experiments.

We trained the three models using six groups of configured datasets. There are eighteen groups of experiments in total. The results are shown in Table IV. There are 17 experiments in which the accuracy of default setting of federated learning is higher than the minimum accuracy of each individual local learning. The above results indicates that federated learning can enhance the model generalization and improve model performance. Please note that interference is introduced in the fourth group of dataset for each model, where images of negative cases are marked as COVID-19. However, the model trained by federated learning can still achieve relatively high accuracy, which shows that the default setting of federated learning can ensure fault tolerance.

B. Performance of Dynamic Fusion based Federated Learning

1) Accuracy

To evaluate the accuracy of dynamic fusion based federated learning (DF_FL), we conducted experiments using the dataset configuration illustrated in Table III and compared the results with default setting of federated learning (Default_FL). The results are presented in Fig. 3, Fig. 4, Fig. 5 respectively for each type of model. The results show that in the 18 groups of experiments, there are 14 groups in which dynamic fusion based federated learning (DF_FL) has higher accuracy than the default setting of federated learning (Default_FL). In the rest four groups, the accuracy of DF_FL is lower than Default_FL for 1.711%, 0.57%, 0.57%, and 1.141% respectively. Overall, the proposed dynamic fusion based federated learning...
approach performs better compared to the default setting of federated learning in terms of accuracy.

2) Training Time

To evaluate the communication efficiency of the proposed dynamic fusion based federated learning, we recorded the training time during the above experiments. The training epochs of all the three clients are set to 90, and the maximum network speed is configured as 10 MB/s for model upload/download (10MB/s). The recorded training time is illustrated in Fig. [6] The results show that in GhostNet, the proposed dynamic fusion based federated learning does not lower the training time, while there is an apparent effect on ResNet50 and ResNet101. The training time of ResNet50 is reduced by 8-10 minutes, while the training time of ResNet101 is decreased by 25-30 minutes.

Since we found that the proposed dynamic fusion based federated learning cannot reduce the training time of GhostNet network in the above experiments, thus we further study the influence factor. After measuring the single model transmission time (we do not include the experiment results due to paper length limit), we observe that the GhostNet is smaller and with less parameters compared to the other two networks. Thus, GhostNet costs less time for model transmission (which is 2.2s on average), which results in no change on GhostNet training time. By contrast, ResNet50 and ResNet101 have more parameters which takes longer time to transmit the model updates. Thus, there is an apparent improvement on these two networks in terms of communication efficiency.

We can conclude that applying the proposed dynamic fusion based approach can significantly reduce the training time when the network is poor and the model has large amounts of parameters.

IV. RELATED WORK

The concept of federated learning is first proposed by Google in 2016 [1], which initially focuses on cross-device learning. For example, Google adopts federated learning to predict search suggestions, next words and emojis, and the learning of out-of-vocabulary words [3], [5]. The scope of federated learning is then extended to cross-silo learning, e.g.,

| Algorithm | Group No. | LL on Client1 | LL on Client2 | LL on Client3 | Default_FL |
|-----------|-----------|---------------|---------------|---------------|------------|
| GhostNet  | 1         | 0.48288973    | 0.87452491    | 0.89923954    | 0.88022814 |
|           | 2         | 0.8878327     | 0.90874525    | 0.89923954    | 0.88403042 |
|           | 3         | 0.88973384    | 0.8460076     | 0.89163498    | 0.8973384  |
|           | 4         | 0.87832699    | 0.49420658    | 0.8973384     | 0.7756654  |
|           | 5         | 0.8669201     | 0.93155894    | 0.8460076     | 0.878327   |
|           | 6         | 0.91634981    | 0.83460076    | 0.8878327     | 0.90114068 |
| ResNet50  | 1         | 0.20532319    | 0.88022814    | 0.85931559    | 0.94106464 |
|           | 2         | 0.90494297    | 0.88403042    | 0.89163498    | 0.93346008 |
|           | 3         | 0.8878327     | 0.85171103    | 0.89543726    | 0.9486692  |
|           | 4         | 0.8923954     | 0.48098859    | 0.88973384    | 0.92585551 |
|           | 5         | 0.90874525    | 0.90874525    | 0.87642586    | 0.9678061  |
|           | 6         | 0.90114068    | 0.91064639    | 0.90494296    | 0.97338403 |
| ResNet101 | 1         | 0.4904943     | 0.86692015    | 0.87262357    | 0.91634981 |
|           | 2         | 0.83840304    | 0.878327      | 0.85931559    | 0.93346008 |
|           | 3         | 0.90304183    | 0.90304183    | 0.90304183    | 0.94296578 |
|           | 4         | 0.89923954    | 0.50760456    | 0.89543726    | 0.92395437 |
|           | 5         | 0.91254753    | 0.89923954    | 0.87623574    | 0.96577947 |
|           | 6         | 0.92775665    | 0.93155894    | 0.86882129    | 0.97528517 |
for different organisations or data centers [2], [6]. For example, Sheller et al. [7] build a segmentation model using the brain tumor data from different medical institutions.

Although communication efficiency can be improved by only sending model updates instead of raw data, federated learning systems still need multiple rounds of communications during training to achieve model convergence. Many researchers work on methods to reduce communication rounds
Another way to reduce communication cost and latency is to extend aggregation, including selective aggregation [10], aggregation scheduling [11], asynchronous aggregation [12], temporally weighted aggregation [13], controlled averaging algorithms [14], iterative round reduction [8], and shuffled model aggregation [15]. Further, model compression methods are utilised to reduce the communication cost that occurs during the model parameters and gradients exchange between clients and the central server [16]. Additionally, communication techniques are introduced to improve communication efficiency, e.g., over-the-air computation technique [17], multi-channel random access communication mechanism [18].

Federated learning is considered as an effective way to improve model performance and address statistical and system heterogeneity issues since models are trained locally [19]. However, challenges still exist in dealing with non-IID data to keep model performance. To deal with the above issue, many researchers work on training data clustering [20], multi-stage local training [21], and multi-task learning [19]. Also, some works [22], [23] focus on incentive mechanism design to motivate clients to participate the machine learning jobs.

Federated learning has been recently adopted in CT or X-Ray image processing for COVID-19 positive case detection [24], [25]. However, the above studies do not consider the communication efficiency and model accuracy issues of federated learning. In this paper, we propose an dynamic fusion-based approach to improve communication efficiency and adopt reinforcement learning to improve model performance.

V. CONCLUSION

This paper proposes a novel dynamic fusion based federated learning approach to improve accuracy and communication efficiency while preserving data privacy for COVID-19 detection. First, we present architecture for dynamic fusion based federated learning using images (CT scan, X-Ray) across multiple medical institutions. Second, we design a decision making mechanism for clients to decide each round’s participation based on the local model performance. Third, we propose an aggregation scheduling method for client selection based on each participating client’s training time. Fourth, we provide a category of image datasets for COVID-19 detection which can be used by machine learning community. The evaluation results show that the proposed approach is feasible and performs better than the default setting of federated learning in terms of model accuracy and communication efficiency.

REFERENCES

[1] H. B. McMahan, E. Moore, D. Ramage, and B. A. y Arcas, “Federated learning of deep networks using model averaging,” CoRR, vol. abs/1602.05629, 2016. [Online]. Available: http://arxiv.org/abs/1602.05629
[2] S. K. Lo, Q. Lu, C. Wang, H. Paik, and L. Zhu, “A systematic literature review on federated machine learning: From a software engineering perspective,” ArXiv, vol. abs/2007.11534, 2020.
[3] T. Yang, G. Andrew, H. Eichner, H. Sun, W. Li, N. Kong, D. Ramage, and F. Beaufays, “Applied federated learning: Improving google keyboard query suggestions,” CoRR, vol. abs/1812.02903, 2018. [Online]. Available: http://arxiv.org/abs/1812.02903
[4] M. Chen, R. Mathews, F. Ouyang, and F. Beaufays, “Federated learning of out-of-vocabulary words,” 2019.
[5] S. Ramaswamy, R. Mathews, K. Rao, and F. Beaufays, “Federated learning for emoji prediction in a mobile keyboard,” 2019.
Fig. 6: Training time.

[6] P. Kairouz, H. B. McMahan, B. Avent, A. Bellet, M. Bennis, A. N. Bhagoji, K. Bonawitz, Z. Charles, G. Cormode, R. Cummings et al., “Advances and open problems in federated learning,” arXiv preprint arXiv:1912.04977, 2019.

[7] M. J. Sheller, G. A. Reina, B. Edwards, J. Martin, and S. Bakas, “Multi-institutional deep learning modeling without sharing patient data: A feasibility study on brain tumor segmentation,” in Brainlesion: Glioma,

[8] J. Mills, J. Hu, and G. Min, “Communication-efficient federated learning for wireless edge intelligence in iot,” IEEE Internet of Things Journal, pp. 1–1, 2019.

[9] S. Silva, B. A. Gutman, E. Romero, P. M. Thompson, A. Altmann, and M. Lorenzi, “Federated learning in distributed medical databases: Meta-analysis of large-scale subcortical brain data,” in 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), April 2019, pp. 270–274.

[10] J. Kang, Z. Xiong, D. Niwato, Y. Zou, Y. Zhang, and M. Guizani, “Reliable federated learning for mobile networks,” IEEE Wireless Communications, vol. 27, no. 2, pp. 72–80, 2020.

[11] H. H. Yang, Z. Liu, T. Q. S. Quek, and H. V. Poor, “Scheduling policies for federated learning in wireless networks,” IEEE Transactions on Communications, vol. 68, no. 1, pp. 317–333, Jan 2020.

[12] C. Xie, S. Koyejo, and I. Gupta, “Asynchronous federated optimization,” 2019.

[13] T. Nishio and R. Yonetani, “Client selection for federated learning with heterogeneous resources in mobile edge,” in ICC 2019 - 2019 IEEE International Conference on Communications (ICC), May 2019, pp. 1–7.

[14] S. P. Karimireddy, S. Kale, M. Mohri, S. J. Reddi, S. U. Stich, and A. T. Suresh, “Scaffold: Stochastic controlled averaging for on-device federated learning,” arXiv preprint arXiv:1910.06378, 2019.

[15] B. Ghazi, R. Pagh, and A. Velingker, “Scalable and differentially private distributed aggregation in the shuffled model,” 2019.

[16] L. Wang, W. Wang, and B. Li, “Cmv: Mitigating communication overhead for federated learning,” in 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS), July 2019, pp. 954–964.

[17] K. Yang, T. Jiang, Y. Shi, and Z. Ding, “Federated learning via over-the-air computation,” IEEE Transactions on Wireless Communications, pp. 1–1, 2020.

[18] J. Choi and S. R. Pokhrel, “Federated learning with multichannel aloha,” IEEE Communications Magazine, pp. 1–1, 2019.

[19] L. Corinzia and J. M. Buhmann, “Variational federated multi-task learning,” arXiv preprint arXiv:1906.06268, 2019.

[20] F. Sattler, K.-R. Müller, and W. Samek, “Clustered federated learning: Model-agnostic distributed multi-task optimization under privacy constraints,” arXiv preprint arXiv:1910.01991, 2019.

[21] Y. Jiang, J. Konečný, K. Rush, and S. Kannan, “Improving federated learning personalization via model agnostic meta learning,” arXiv preprint arXiv:1909.12488, 2019.

[22] Y. J. Kim and C. S. Hong, “Blockchain-based node-aware dynamic weighting methods for improving federated learning performance,” in 2019 20th Asia-Pacific Network Operations and Management Symposium (APNOMS), Sep. 2019, pp. 1–4.

[23] H. Kim, J. Park, M. Bennis, and S. Kim, “Blockchain-enhanced on-device federated learning,” IEEE Communications Letters, pp. 1–1, 2019.

[24] B. Liu, B. Yan, Y. Zhou, Y. Yang, and Y. Zhang, “Experiments of federated learning for covid-19 chest x-ray images,” 2020.

[25] R. Kumar, A. A. Khan, S. Zhang, W. Wang, Y. Abudris, W. Amin, and J. Kumar, “Blockchain-federated-learning and deep learning models for covid-19 detection using ct imaging,” 2020.