FedSL: Federated Split Learning for Collaborative Healthcare Analytics on Resource-Constrained Wearable IoMT Devices

Wanli Ni, *Member, IEEE*, Huiqing Ao, Hui Tian, *Senior Member, IEEE*,
Yonina C. Eldar, *Fellow, IEEE*, and Dusit Niyato, *Fellow, IEEE*

***Abstract*—Many wearable Internet of Medical Things (IoMT) devices have limited computing power and small storage space. Additionally, the healthcare data sensed by a single IoMT device is not enough to train a sophisticated deep learning model. To address these challenges, we propose a federated split learning (FedSL) framework that allows for collaborative healthcare analytics on multiple IoMT devices with limited resources. Compared to centralized learning, FedSL can protect user privacy by not sending raw data over wireless networks. Furthermore, FedSL offers more flexibility than other federated learning methods. It enables even low-end IoMT devices to participate in model training and result inference. Experimental results show that our FedSL performs well on medical imaging tasks with different data distributions.**

***Index Terms*—Federated split learning (FedSL), healthcare analytics, Internet of Medical Things (IoMT), user privacy, wearable devices.**

I. Introduction

THE INTERNET of Medical Things (IoMT) enables real-time body data collection in a cost-effective way, which is changing existing healthcare industries [1]. Moreover, with the success of deep learning (DL) in medical diagnosis tasks, online healthcare platforms can now train powerful models using data from massive IoMT devices [2], [3]. This enables patients to get telemedicine services quickly without visiting a clinic, and doctors can make more reliable and faster decisions [4]. But how to use wearable IoMT devices for collaborative healthcare analytics is still a challenge, especially with limited resources.

Due to the highly sensitive nature of medical information, sending IoMT data over wireless networks may pose a significant privacy concern. Recently, federated learning (FL) has emerged as a potential solution to overcome this issue [5], [6]. However, FL requires all devices to train the full DL model, which can be resource-intensive, especially for models with millions of parameters [5], [6]. This is impractical for many small-size and low-cost wearable IoMT devices, such as patch sensors and smart watches [1]. To this end, the split learning paradigm can be adopt to divide the complete DL model into a server-side model and a device-side model [2]. By doing so, IoMT devices only need to train a small portion of the DL model, effectively reducing their computational burden [7]. As a federated variant of split learning, in this letter, we propose a federated split learning (FedSL) framework to alleviate local hardware requirements and support semi-distributed learning while partially preserving user privacy.

Manuscript received 12 February 2023; revised 2 February 2024; accepted 25 February 2024. Date of publication 28 February 2024; date of current version 9 May 2024. The work of Hui Tian was supported by the Beijing Municipal Natural Science Foundation under Grant L232052. *(Corresponding author: Wanli Ni.)*

Wanli Ni is with the Department of Electronic Engineering and the Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing 100084, China (e-mail: niwanli@tsinghua.edu.cn).

Huiqing Ao and Hui Tian are with the State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mail: hqao@bupt.edu.cn; tianhui@bupt.edu.cn).

Yonina C. Eldar is with the Faculty of Mathematics and Computer Science, Weizmann Institute of Science, Rehovot 7610001, Israel (e-mail: yonina.eldar@weizmann.ac.il).

Dusit Niyato is with the School of Computer Science and Engineering, Nanyang Technological University, Singapore (e-mail: dniyato@ntu.edu.sg).

Data is available on-line at https://github.com/niwanli/FedSL-IoMT.

Digital Object Identifier 10.1109/JIOT.2024.3370985

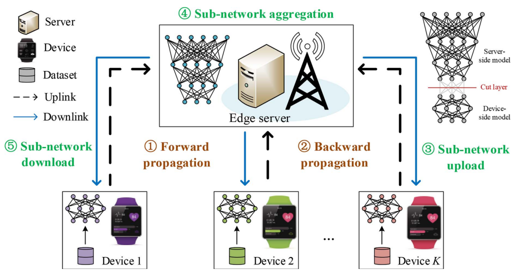

Fig. 1. Illustration of the proposed FedSL framework for IoMT devices.

II. Framework Design

As illustrated in Fig. 1, we propose a FedSL framework for resource-limited wearable IoMT devices to train a shared DL model collaboratively using the medical healthcare data collected. The set of devices is denoted as $\mathcal{K} = \{1, 2, \ldots, K\}$. The layers of the DL model are denoted as $\mathcal{L} = \{L_1, L_2, \ldots, L_N\}$. According to the depth of neural networks, the full DL model is divided into two subnetworks [2]. Specifically, these high-level layers $\mathcal{L}_S = \{L_{c+1}, L_{c+2}, \ldots, L_N\}$ are kept at the server, while the low-level layers $\mathcal{L}_D = \{L_1, L_2, \ldots, L_c\}$ are stored on IoMT devices, where L_c is the cut layer [3]. FedSL's model splitting paradigm offers several advantages. Similar to FL, FedSL can fully utilize the distributed data collected by wearable IoMT devices while preserving user privacy. This is achieved by transferring only the intermediate features extracted by the device-side model to the server, enabling collaborative model training without disclosing sensitive raw data. Unlike FL, which requires clients with sufficient computing and storage resources, FedSL enables computing-limited and storage-constrained IoMT devices to participate in the model training process. This is because only shallow neural networks are computed on these devices, reducing computational and storage requirements. Compared to the scheme in [7], FedSL only needs one edge server to coordinate model training, and thus the communication flow is different. Specifically, as shown in Fig. 2, each training round of FedSL has two stages.

1) *Stage 1 (Multiuser Split Learning):* At the beginning, all devices feed subnetworks with mini-batch data samples in sequence or parallel. Next, the smashed data obtained at the cut layer and the label are sent to

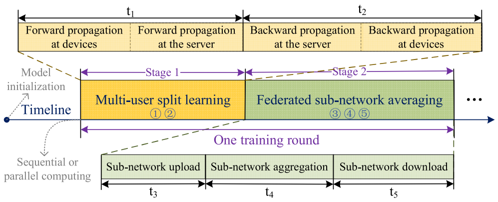

Fig. 2. One training round of FedSL in the IoMT-based healthcare system.

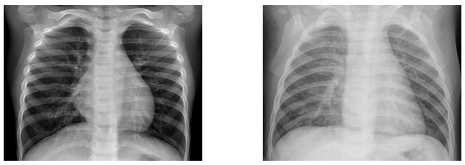

Fig. 3. Normal chest X-ray image (left) versus viral pneumonia image (right).

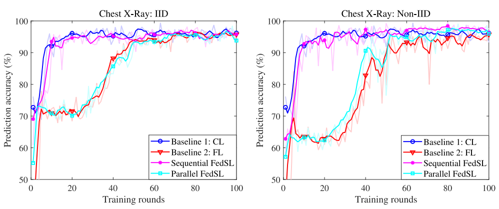

Fig. 4. Accuracy on the chest X-Ray data set with IID and non-IID settings.

the edge server for completing the rest of the forward propagation [7]. Then, with the loss values obtained at the output layer, the gradients are computed for different devices and back propagated from the server to devices.

2) *Stage 2 (Federated Subnetwork Averaging):* In the sequel, to share knowledge and accelerate convergence, all devices upload their updated subnetwork parameters to the server. After aggregation, the averaged subnetwork parameters are sent back to all devices for the next training round [5], [6]. The learning process terminates when the loss value stabilizes or the maximum number of training rounds is reached.

III. Experimental Results

We conduct simulations in an IoMT-based healthcare monitoring system with $K = 5$ devices. Specifically, an 18-layer ResNet is employed to classify Chest X-Ray (two classes as shown in Fig. 3) and optical coherence tomography (OCT, four classes) images [4] with independent identically distributed (IID) and non-IID settings. The cut layer is set as $L_c = 3$, the learning rate is 10^{-4}, and the maximum number of training rounds is 400. For comparison, we consider four schemes: 1) centralized learning (CL); 2) FL; 3) sequential FedSL; and 4) parallel FedSL.

Fig. 4 shows the learning performance of training ResNet-18 on the Chest X-Ray data set with IID and non-IID settings. From this figure, we have the following observations. In the first half of the training process, the sequential FedSL scheme is able to achieve similar performance to CL, while the parallel FedSL scheme is close to the FL baseline. This is because in the sequential FedSL scheme, the server only needs to train one server-side model for all devices, while in the parallel FedSL scheme, the server needs to compute different subnetworks for different devices in parallel. When the number of training rounds is greater than 50, the performance gap between any two schemes is very small. This is due to the fact that the DL model is well trained in the second half of all schemes, and thereby all schemes achieve almost similar prediction accuracy. Compared to the IID setting, these results in the non-IID case fluctuate more significantly and the sequential and parallel FedSL schemes slightly outperform the traditional CL and FL benchmarks, respectively. This shows that our FedSL has good adaptability to the non-IID data sets. Similar observations can be seen in Fig. 5, due to the page limit, we omit the analysis here for brevity. Please go to our site at https://github.com/niwanli/FedSL-IoMT for more details. Overall, the findings above reveal the effectiveness and superiority of the proposed FedSL in healthcare systems.

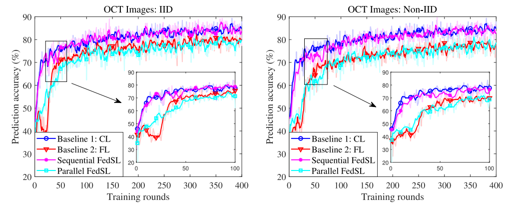

Fig. 5. Prediction accuracy on the OCT data set with IID and non-IID settings.

IV. Conclusion

In this letter, we proposed the FedSL framework to implement medical data analysis on multiple wearable IoMT devices. This framework is friendly to resource-constrained, low-cost terminals in comparison to traditional CL and FL schemes, especially when the size of the trained DL model is large. Furthermore, since only intermediate features are shared between servers and devices, our FedSL can protect user privacy partially without compromising data accessibility, thus making it suitable for privacy-sensitive cases. Experiments showed that our FedSL achieves the similar performance as FL, and the gap between FedSL and CL is small.

References

[1] A. Gatouillat et al., "Internet of Medical Things: A review of recent contributions dealing with cyber-physical systems in medicine," *IEEE Internet Things J.*, vol. 5, no. 5, pp. 3810–3822, Oct. 2018.

[2] P. Vepakomma et al., "Split learning for health: Distributed deep learning without sharing raw patient data," 2018, *arXiv:1812.00564*.

[3] M. Zhang, L. Qu, P. Singh, J. Kalpathy-Cramer, and D. L. Rubin, "SplitAVG: A heterogeneity-aware federated deep learning method for medical imaging," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 9, pp. 4635–4644, Sep. 2022.

[4] D. S. Kermany et al., "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, Feb. 2018.

[5] T. Gafni, N. Shlezinger, K. Cohen, Y. C. Eldar, and H. V. Poor, "Federated learning: A signal processing perspective," *IEEE Signal Process. Mag.*, vol. 39, no. 3, pp. 14–41, May 2022.

[6] L. You, S. Liu, Y. Chang, and C. Yuen, "A triple-step asynchronous federated learning mechanism for client activation, interaction optimization, and aggregation enhancement," *IEEE Internet Things J.*, vol. 9, no. 23, pp. 24199–24211, Dec. 2022.

[7] C. Thapa, P. C. M. Arachchige, S. Camtepe, and L. Sun, "SplitFed: When federated learning meets split learning," in *Proc. AAAI*, 2022, pp. 8485–8493.