MICCAI

This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

FedMRL: Data Heterogeneity Aware Federated Multi-agent Deep Reinforcement Learning for Medical Imaging

Pranab Sahoo[1][0000−0002−0784−1908], Ashutosh Tripathi[2], Sriparna Saha[1], and Samrat Mondal[1]

[1] Department of Computer Science and Engineering
Indian Institute of Technology Patna, India
[2] Rajiv Gandhi Institute of Petroleum Technology, India
{pranab_2021cs25,sriparna,samrat}@iitp.ac.in

Abstract. Despite recent advancements in Federated Learning (FL) for medical image diagnosis, addressing data heterogeneity among clients remains a significant challenge for practical implementation. A primary hurdle in FL arises from the non-independent and identically distributed (non-IID) nature of data samples across clients, which typically results in a decline in the performance of the aggregated global model. This study introduces FedMRL, a novel federated multi-agent deep reinforcement learning framework designed to address data heterogeneity. FedMRL incorporates a novel loss function to facilitate fairness among clients, preventing bias in the final global model. Additionally, it employs a multi-agent reinforcement learning (MARL) approach to calculate the proximal term (μ) for the personalized local objective function, ensuring convergence to the global optimum. Furthermore, FedMRL integrates an adaptive weight adjustment method using a Self-organizing map (SOM) on the server side to counteract distribution shifts among clients' local data distributions. We assess the proposed approach using two publicly available real-world medical datasets, and the results demonstrate that FedMRL significantly outperforms state-of-the-art techniques, showing its efficacy in addressing data heterogeneity in federated learning. The code can be found here `https://github.com/Pranabiitp/FedMRL`.

Keywords: Federated learning · Heterogeneity · Reinforcement learning.

1 Introduction

Deep learning (DL) algorithms have demonstrated significant achievements in medical image analysis tasks [15], [19], [18], [17]. However, creating effective DL-based models typically requires gathering training data from various medical centers, such as hospitals and clinics, into a centralized server. Obtaining patient data from multiple centers presents challenges due to privacy concerns, legal restrictions on data sharing, and the logistical difficulty of transferring large data volumes [20]. Researchers have increasingly employed Federated Learning (FL) as a solution, enabling medical image classification with decentralized data from multiple sources while preserving privacy [25], [3]. Unlike models trained independently at individual sites, FL can leverage a more diverse and extensive dataset, resulting in improved performance and increased generalizability [13], [11]. The efficacy of federated training encounters challenges due to data heterogeneity within local hospital datasets, resulting in performance degradation in real-world healthcare applications. This heterogeneity manifests in various forms: some hospitals may possess more data from the early stages of the diseases, while others collect data with severe conditions, leading to label distribution skew. Additionally, variations in data quantity among hospitals, with larger institutions having more patient data compared to community clinics, contribute to quantity skew. Moreover, differences in imaging acquisition protocols and patient populations further exacerbate feature distribution skew [24]. FedAvg, a foundational FL algorithm, while successful in many scenarios, exhibits diminished efficacy in heterogeneous data settings [13]. To address this challenge, FedProx [8] introduced a proximal term (μ) into the conventional optimization objective to penalize large updates in the model parameters. However, selecting an optimal value for the μ in FedProx presents a challenge, as traditional methods like trial and error or heuristics may not effectively adapt to heterogeneous data distributions.

Distribution shifts within each hospital's private dataset often result in scenarios where the global model performs better for certain hospitals but neglects others. The study [9] introduced q-fairness optimization problems in FL, where the parameter q guides the loss function to desirable outcomes. Huang et al. [6] focused on fairness and robustness by dynamically selecting local centers for training, but this approach may not be suitable for the medical domain due to limited participant hospitals. Lyu et al. [12] proposed a collaborative, fair FL framework to enforce convergence to different models, addressing fairness differently. Another challenge is that existing methods commonly train the global model by minimizing the average training losses of all local clients [13], [10], [22]. However, these approaches lack performance guarantees for individual hospitals as they prioritize average training results, leading to divergent performance across participants [8]. This challenge is further exacerbated in real-world scenarios where data from medical centers exhibit discrepancies in both size and distribution [4]. Motivated by the aforementioned challenges, we introduce FedMRL, a novel framework that addresses these issues through three distinct components. Our main contributions are:

- The FedMRL framework introduces a novel method for calculating adaptive μ values by leveraging the QMIX algorithm from Multi-agent Reinforcement Learning (MARL). This approach accounts for client-specific factors such as data distribution, volume, and performance feedback, facilitating dynamic regularization adjustments during FL training.
- We propose incorporating a novel loss function into individual client objectives to promote fairness, aiming to mitigate discrepancies between each

client's loss and the global loss. This will reduce potential biases towards specific clients or groups of clients within the system.

- Our solution involves a server-side adaptive weight adjustment method using a self-organizing map (SOM). This prioritizes contributions from clients with similar data distributions, determined by cosine similarity between local and global models.
- Experiments conducted on two real-world medical datasets on severe degree of data heterogeneity ($\eta = 1$) show that FedMRL outperforms several state-of-the-art FL techniques.

2 Problem Statement

Assuming there are H hospitals, each represented by $h \in [1, 2, \cdots, H]$, and possessing privately labeled data denoted by D^h, the aim is to train a generalized global model over the combined dataset $D = \bigcup_{h=1}^{H} D^h$. The global objective function is represented in Eq. 1.

$$\arg\min_{w} L(w) = \sum_{h=1}^{H} \frac{|D^h|}{|D|} L_h(w) \tag{1}$$

The local objective function $L_h(w)$ in client h, which quantifies the local empirical loss over the data distribution D^h, is represented in Eq. 2.

$$L_h(w) = \mathbb{E}_{x \sim D^h}[l_h(w; x)] \tag{2}$$

In this context, l_h denotes the loss function utilized by client h, and w represents the global model parameters. While the above fixed weighted averaging method offers an unbiased global model estimation in the presence of independent and identically distributed (IID) training samples across clients, non-IID distributions, stemming from device and user heterogeneity, lead to slower convergence and reduced accuracy [26]. To address this challenge, we propose FedMRL, integrating a novel fairness term into the local objective function, dynamically determining the proximal term for each hospital through a MARL-based approach, and employing a self-organizing map-based aggregation method at the server. The final local objective function is represented as shown in Eq. 3.

$$\arg\min_{w} L(w) = \sum_{h=1}^{H} \frac{|D^h|}{|D|} L_h(w) + \frac{\mu}{2}\|w - w^t\|^2 + L_{fair}(w) \tag{3}$$

3 Proposed Framework

This section provides a comprehensive overview of the proposed FedMRL, which consists of three contributions such as calculating adaptive personalized μ Value, novel loss function, and server-side adaptive weight Aggregation using SOM. The overall algorithm is represented in Algo. 1. The architecture details are shown in Fig. 1.

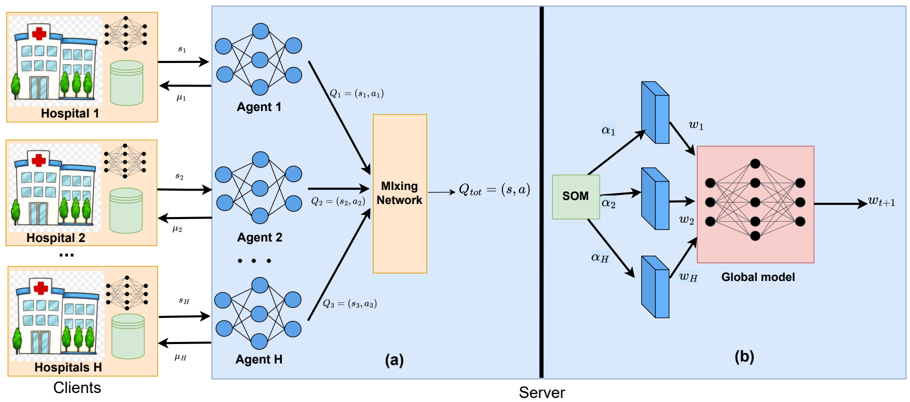

Fig. 1. The proposed architecture comprises clients and a server: (a) Clients transmit their model weights to the global server, while agents corresponding to clients retrieve the corresponding state s_i from the global state s and compute the respective μ_i, subsequently sharing it with the corresponding hospitals. (b) Global weight aggregation is facilitated using SOM, where α_i denotes the weight adjustment factor, and w_i represents the local model weights utilized to derive the final global model w_{t+1} for subsequent communication rounds.

3.1 Adaptive Personalized μ Value

For the dynamic adaptation of μ, we frame it as a multi-agent reinforcement learning problem, with each hospital h having an agent on the server side. Each agent h observes its state s_i from the overall environment state s_t, selects an action a_i based on the current policy π_t, and the agents' actions collectively form the joint action A_i. The environment transitions to the next state s_{i+1} according to the state transition function $P(s_{i+1}|s_i, a_i)$, iterating until completion or predefined criteria are met. In the proposed work, we integrate QMIX [16], a prominent Q-learning algorithm for cooperative MARL in the decentralized paradigm that represents an advancement over Value-Decomposition Networks (VDN) [21]. Essentially, VDN assesses the influence of each agent on the collective reward, considering that the joint action-value $Q_{\text{tot}}(s, a)$ can be decomposed into N individual Q-functions for N agents, with each Q-function relying solely on local state-action history, represented in Eq. 4.

$$Q_{\text{tot}}(s, a) = \sum_{j=1}^{N} Q_j(s_j, a_j, \theta_j) \tag{4}$$

State: The state of the environment at round t is $s_t = [s_{t,1}, s_{t,2}, \cdots, s_{t,h}]$ representing the data distribution among clients and performance feedback $s_{t,i}$ is defined in Eq. 5.

$$s_{t,i} = (E_c, P_c, acc_c, \text{loss}_c) \tag{5}$$

We have represented the entropy of the dataset in the c^{th} client (E_c) as defined in Eq. 6 and the ratio of the data samples in the c^{th} client to the data samples in all clients (P_c), are represented in Eq. 7 [25].

$$E_c = \sum_{m=0}^{M-1} \frac{N_m}{N_c} \log\left(\frac{N_m}{N_c}\right) \tag{6}$$

$$P_c = \frac{N_c}{N} \tag{7}$$

N_m, N_c, and N denote the number of data samples belonging to the m^{th} class in the c^{th} client, the total number of data samples in the c^{th} client, and the total data samples in FL sourced from local clients, respectively. Additionally, the c^{th} local model's accuracy and training loss, denoted as acc_c and $loss_c$, respectively, are dynamic metrics that evolve across communication rounds, serving as performance feedback for the local models. (E_c) and (P_c) depict the features of the local datasets and remain constant throughout the training.

Action: We consider the proximal term μ as our action-space, a continuous value ranging from 0 to 1, and initialize it at a minimal value (≈ 0.00001) at the beginning of FL training. Subsequently, following the first communication round, agents determine μ based on the current policy, resulting in a joint action denoted as $A_t = [a_{t,1}, a_{t,2}, \cdots, a_{t,h}]$, where $a_{t,i}$ represents the proximal term value assigned to agent h_i.

Reward: The observed reward value in round t is denoted as $r_t = e^{(acc_t - \zeta)} - 1$, where e represents the natural constant, ζ denotes the target accuracy, and acc_t signifies the global model's test accuracy at round t. The QMIX agent maximizes the anticipation of the cumulative discounted reward (R) through training, as described in Eq. 8, where $\gamma \in (0, 1]$ is the discount factor.

$$R = \sum_{t=1}^{T} \gamma_{t-1} r_t = \sum_{t=1}^{T} \gamma_{t-1} \left(e^{(acc_t - \zeta)} - 1\right) \tag{8}$$

3.2 Loss Function

Our proposed novel loss function in FedMRL effectively mitigates the impact of distribution shifts in client datasets, ensuring fairness and consistent performance across all participating institutions. By incorporating a fairness term into the local objective function of individual hospitals, FedMRL adjusts model parameters to achieve uniform training loss across all H hospitals, inspired by the Mean Square Error (MSE) loss commonly used in regression models. The fairness term is formulated as an optimization problem aiming to minimize the sum of squares of differences in loss between each of the H hospitals and the global loss, as presented in Eq. 9.

$$L_{fair} = \sum_{h=1}^{H} \left(F_h(w) - F(w)\right)^2 \tag{9}$$

where, w represents the parameter of the global model, H denotes the number of hospitals, $F_h(w)$ signifies the local loss function of individual hospital h based on its local data, and $F(w) = \sum_{h=1}^{H} F_h(w)$ represents the global loss function. The proof of the above-proposed loss function can be found in Section 1 of the supplementary material.

Algorithm 1: Steps of the proposed FedMRL.

Input: Set of hospitals H, initial global model weights w^0, target accuracy ζ, number of communication rounds T
Output: Global model parameter w_t

Initialize $i = 0$;
while $i < T$ **do**
    Each hospital downloads the weights of the initial global model, performs local SGD training for 1 epoch by minimizing the local objective function represented in Eq. 3, and uploads the local model weights to the server;
    Execute QMIX agent as described in Subsection 3.1;
    Execute SOM as described in Subsection 3.3;
    Share w^{t+1} back to all the clients for the next communication round;
    $i \leftarrow i + 1$;
end

3.3 Server side Weight Aggregation

FedAvg uniformly averages client model updates, neglecting individual data distributions, which can impede performance in non-IID scenarios. In contrast, self-organizing map (SOM)-based weight adjustment considers client model similarity to the global model, significantly impacting clients with more representative data [7]. This adaptivity effectively addresses non-IID distribution challenges, allowing personalized adjustments based on client and global model similarity, thus enhancing performance for clients with unique data distributions. We initialize the SOM grid shape as (5,5), and its weights are randomly initialized. Distances between the SOM weights and each hospital's local weights determine the Best Matching Unit (BMU) on the SOM grid. The influence of local weights on each SOM neuron is calculated based on its distance to the BMU and current sigma value, updating the neuron weights accordingly. Weights for each local model are computed from the SOM weights and similarity metrics, ensuring accurate representation through normalization. Cosine similarity metrics between local and global models, combined with distances, determine weights (α_i) for each local model, favoring higher similarity for increased weight. During SOM weight updates, the influence of each local model scales by its similarity metric, with higher similarity models exerting greater impact. Normalized similarity metrics ensure proportional weighting, are responsive to changes in local models over time and maintain fairness in highly non-IID data settings. The aggregation is performed according to Eq. 10, where w_{t+1} denotes the global model parameters for the $(t+1)^{th}$ communication round, w_t^h signifies the local model weights of the h^{th} hospital, and α_h represents the weight factor for the h^{th} hospital.

$$w_{t+1} = \frac{1}{H} \sum_{h=0}^{H-1} \omega_t^h \cdot \alpha_h \tag{10}$$

4 Dataset and Experimental Results

4.1 Datasets

We have chosen two distinct benchmark datasets pertinent to real-world medical contexts to evaluate the effectiveness of the proposed FedMRL framework for highly heterogeneous scenarios. The ISIC 2018 dataset, notable for its contributions to skin cancer detection, provides diverse dermoscopy images captured from various anatomical regions. This dataset encompasses 7,200 images categorized into 7 distinct classes [1]. Additionally, we have utilized the Messidor dataset [2], consisting of 1,560 authentic fundus images tailored for grading diabetic macular edema across five severity levels. Each client's dataset is split into 20% for testing and 80% for training. For the test dataset on the server side, we have utilized the validation split of ISIC-2018 and randomly curated 240 images from the Messidor dataset.

4.2 Implementation Details

We have constructed non-IID data partitions following the methodology outlined in [14]. Using 80% of each dataset for training purposes, we organize the data based on their labels and segment each class into 200 shards. Subsequently, clients create local datasets by sampling from these shards according to the probabilities represented in Eq. 11.

$$\text{pr}(x) = \begin{cases} \eta \in [0,1], & \text{if } x \in \text{class } j, \\ N(0.5, 1), & \text{otherwise.} \end{cases} \tag{11}$$

Samples from other classes have a Gaussian distribution, but the client samples from a particular class j with a constant probability η. More samples in a given class are indicated by higher η values, which produce more diverse datasets. For all the experiments, we have utilized $\eta = 1.0$. Following [23], we employ DenseNet121 [5] as the backbone. For the Fedprox baseline experiment, we selected the value of μ from the set $\{0.001, 0.1, 0.4\}$, and for the Fednova baseline experiment, we chose the proximal SGD value from the set $\{0.001, 0.1, 0.2\}$ to yield the best results.

4.3 Comparison with State-of-the-arts

To assess the efficacy of the proposed FedMRL, we compare them against four baseline approaches: FedAvg, FedProx, FedNova, and FedBN. Table 1 presents

Table 1. Results comparison with state-of-the-art methods on two datasets.

Dataset	Method	ACC	AUC	Pre	Recall	F1
ISIC-2018	Fedavg [13]	72.82	89.85	74.69	72.82	73.74
	Fedprox [8]	72.90	90.26	74.30	73.21	73.75
	Fednova [22]	67.88	86.60	70.51	67.88	69.17
	FedBN [10]	71.71	90.18	73.46	71.71	72.57
	FedMRL	**73.50**	**80.01**	**77.04**	**71.90**	**74.38**
Messidor	Fedavg [13]	55.83	86.07	51.96	55.83	53.82
	Fedprox [8]	57.08	85.99	56.71	57.08	56.89
	Fednova [22]	50.00	82.21	50.70	50.00	50.34
	FedBN [10]	55.83	85.07	57.10	55.83	56.45
	FedMRL	**57.91**	**85.82**	**58.23**	**57.91**	**58.06**

the results of all the models in terms of Accuracy (ACC), Area under the ROC Curve (AUC), Precision (Pre), Recall, and F1-score (F1). The results highlight FedMRL's superior performance over baseline algorithms, demonstrating significant accuracy improvements across both datasets. Specifically, FedMRL outperforms state-of-the-art methods by 0.92%, 0.81%, 7.64%, and 2.43% for the ISIC-2018 dataset, and by 3.59%, 1.43%, 13.65%, and 3.59% for the Messidor dataset, compared to FedAvg, FedProx, FedNova, and FedBN, respectively. These improvements come from the advantage of using MARL to learn the data heterogeneity in the network and adaptively optimize the local objective of clients. Additionally, our novel loss function promotes fairness among clients, while the SOM-based adaptive weight adjustment method for aggregation enhances convergence to a better global optimum. We conducted an ablation study to assess the impact of the proposed components. Without the novel loss function, accuracy reached 70.06% on ISIC-2018 and 51.25% on the Messidor dataset. Similarly, omitting the SOM method resulted in accuracies of 68.0% on the ISIC dataset and 52.45% on the Messidor dataset.

5 Conclusion

This study addresses the challenge of data heterogeneity in federated learning while ensuring fair contributions from the decentralized participants. Our framework demonstrates robustness to non-IID data distribution across clients and outperforms existing benchmarks in two medical datasets. While effective in mitigating the challenges posed by non-IID data distributions, FedMRL may encounter scalability issues when dealing with a large number of clients. Furthermore, the computational overhead associated with calculating personalized μ values and performing server-side adaptive weight aggregation using SOM may impose additional computational burdens, particularly in resource-constrained environments. In future work, we envision exploring distributed computing and resource-sharing models to alleviate the computational burden and aim to enhance FedMRL's scalability and efficiency while exploring its applicability in diverse domains.

Disclosure of Interests The authors declare that there are no competing interests to report in the paper.

References

1. Codella, N., Rotemberg, V., Tschandl, P., Celebi, M.E., Dusza, S., Gutman, D., Helba, B., Kalloo, A., Liopyris, K., Marchetti, M., et al.: Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic). arXiv preprint arXiv:1902.03368 (2019)
2. Decencière, E., Zhang, X., Cazuguel, G., Lay, B., Cochener, B., Trone, C., Gain, P., Ordonez, R., Massin, P., Erginay, A., et al.: Feedback on a publicly distributed image database: the messidor database. Image Analysis & Stereology **33**(3), 231–234 (2014)
3. Feki, I., Ammar, S., Kessentini, Y., Muhammad, K.: Federated learning for covid-19 screening from chest x-ray images. Applied Soft Computing **106**, 107330 (2021)
4. Hosseini, S.M., Sikaroudi, M., Babaie, M., Tizhoosh, H.: Proportionally fair hospital collaborations in federated learning of histopathology images. IEEE transactions on medical imaging (2023)
5. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4700–4708 (2017)
6. Huang, T., Lin, W., Wu, W., He, L., Li, K., Zomaya, A.: An efficiency-boosting client selection scheme for federated learning with fairness guarantee. IEEE Transactions on Parallel and Distributed Systems p. 1–1 (2020). https://doi.org/10.1109/tpds.2020.3040887, `http://dx.doi.org/10.1109/TPDS.2020.3040887`
7. Kohonen, T.: The self-organizing map. Proceedings of the IEEE **78**(9), 1464–1480 (1990)
8. Li, T., Sahu, A.K., Zaheer, M., Sanjabi, M., Talwalkar, A., Smith, V.: Federated optimization in heterogeneous networks. Proceedings of Machine learning and systems **2**, 429–450 (2020)
9. Li, T., Sanjabi, M., Beirami, A., Smith, V.: Fair resource allocation in federated learning. arXiv preprint arXiv:1905.10497 (2019)
10. Li, X., Jiang, M., Zhang, X., Kamp, M., Dou, Q.: Fedbn: Federated learning on non-iid features via local batch normalization. arXiv preprint arXiv:2102.07623 (2021)
11. Lu, M.Y., Chen, R.J., Kong, D., Lipkova, J., Singh, R., Williamson, D.F., Chen, T.Y., Mahmood, F.: Federated learning for computational pathology on gigapixel whole slide images. Medical image analysis **76**, 102298 (2022)
12. Lyu, L., Xu, X., Wang, Q.: Collaborative fairness in federated learning (2020)
13. McMahan, B., Moore, E., Ramage, D., Hampson, S., y Arcas, B.A.: Communication-efficient learning of deep networks from decentralized data. In: Artificial intelligence and statistics. pp. 1273–1282. PMLR (2017)
14. McMahan, H.B., Moore, E., Ramage, D., Hampson, S., y Arcas, B.A.: Communication-efficient learning of deep networks from decentralized data (2023)

15. Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., et al.: Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv preprint arXiv:1711.05225 (2017)
16. Rashid, T., Samvelyan, M., de Witt, C.S., Farquhar, G., Foerster, J., Whiteson, S.: Monotonic value function factorisation for deep multi-agent reinforcement learning (2020)
17. Sahoo, P., Saha, S., Mondal, S., Chowdhury, S., Gowda, S.: Computer-aided covid-19 screening from chest ct-scan using a fuzzy ensemble-based technique. In: 2022 International Joint Conference on Neural Networks (IJCNN). pp. 1–8. IEEE (2022)
18. Sahoo, P., Saha, S., Mondal, S., Chowdhury, S., Gowda, S.: Vision transformer-based federated learning for covid-19 detection using chest x-ray. In: International Conference on Neural Information Processing. pp. 77–88. Springer (2022)
19. Sahoo, P., Saha, S., Sharma, S.K., Mondal, S., Gowda, S.: A multi-stage framework for covid-19 detection and severity assessment from chest radiography images using advanced fuzzy ensemble technique. Expert Systems with Applications **238**, 121724 (2024)
20. Sahoo, P., Sharma, S.K., Saha, S., Mondal, S.: A federated multi-stage light-weight vision transformer for respiratory disease detection. In: International Conference on Neural Information Processing. pp. 300–311. Springer (2023)
21. Sunehag, P., Lever, G., Gruslys, A., Czarnecki, W.M., Zambaldi, V., Jaderberg, M., Lanctot, M., Sonnerat, N., Leibo, J.Z., Tuyls, K., Graepel, T.: Value-decomposition networks for cooperative multi-agent learning (2017)
22. Wang, J., Liu, Q., Liang, H., Joshi, G., Poor, H.V.: Tackling the objective inconsistency problem in heterogeneous federated optimization (2020)
23. Wang, X., Shen, T., Yang, S., Lan, J., Xu, Y., Wang, M., Zhang, J., Han, X.: A deep learning algorithm for automatic detection and classification of acute intracranial hemorrhages in head ct scans. NeuroImage: Clinical **32**, 102785 (2021)
24. Yan, R., Qu, L., Wei, Q., Huang, S.C., Shen, L., Rubin, D., Xing, L., Zhou, Y.: Label-efficient self-supervised federated learning for tackling data heterogeneity in medical imaging. IEEE Transactions on Medical Imaging (2023)
25. Yue, G., Wei, P., Zhou, T., Song, Y., Zhao, C., Wang, T., Lei, B.: Specificity-aware federated learning with dynamic feature fusion network for imbalanced medical image classification. IEEE Journal of Biomedical and Health Informatics (2023)
26. Zhao, Y., Li, M., Lai, L., Suda, N., Civin, D., Chandra, V.: Federated learning with non-iid data. arXiv preprint arXiv:1806.00582 (2018)