



Federated Condition Generalization on Low-dose CT Reconstruction via Cross-domain Learning

Shixuan Chen^{1,3}, Boxuan Cao^{1,3}, Yinda Du^{1,3}, Yaoduo Zhang⁴, Ji He⁴,
Zhaoying Bian^{1,3}, Dong Zeng^{1,2,3}(✉), and Jianhua Ma^{1,3}(✉)

¹ School of Biomedical Engineering, Southern Medical University, Guangdong, China
{zd1989, jhma}@smu.edu.cn

² Department of Radiology, Zhujiang Hospital, Southern Medical University,
Guangdong, China

³ Pazhou Lab (Huangpu), Guangdong, China

⁴ School of Biomedical Engineering, Guangzhou Medical University, Guangdong,
China

Abstract. The harmful radiation dose associated with CT imaging is a major concern because it can cause genetic diseases. Acquiring CT data at low radiation doses has become a pressing goal. Deep learning (DL)-based methods have proven to suppress noise-induced artifacts and promote image quality in low-dose CT imaging. However, it should be noted that most of the DL-based methods are constructed based on the CT data from a specific condition, i.e., specific imaging geometry and specific dose level. Then these methods might generalize poorly to the other conditions, i.e., different imaging geometries and other radiation doses, due to the big data heterogeneity. In this study, to address this issue, we propose a condition generalization method under a federated learning framework (FedCG) to reconstruct CT images on two conditions: three different dose levels and different sampling schemes at three different geometries. Specifically, the proposed FedCG method leverages a cross-domain learning approach: individual-client sinogram learning and cross-client image reconstruction for condition generalization. In each individual client, the sinogram at each condition is processed similarly to that in the iRadon-MAP. Then the CT images at each client are learned via a condition generalization network in the server which considers latent common characteristics in the CT images at all conditions and preserves the client-specific characteristics in each condition. Experiments show that the proposed FedCG outperforms the other competing methods on two imaging conditions in terms of qualitative and quantitative assessments.

This work was supported in part by the NSFC under Grant U21A6005, and Grant 12226004, and Young Talent Support Project of Guangzhou Association for Science and Technology.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-43898-1_5.

Keywords: low-dose CT · image reconstruction · federal learning · generalization · generalization

1 Introduction

Lowering radiation dose is desired in the computed tomography (CT) examination. Various strategies, i.e., lowering incident photons directly (low-mAs), reducing sampling views (sparse-view) and reducing sampling angles (limited-view) can be used for low-dose CT imaging. However, the reconstructed images under these conditions would suffer from severe quality degradation. A number of reconstruction algorithms have been proposed to improve low-dose CT image quality [1–6]. Among them, deep learning (DL)-based methods have shown great promise for low-dose CT imaging, including methods that learn noise distribution features from the image domain to directly reduce noise and artifacts in the reconstructed image [1, 2], as well as methods to improve the reconstruction quality based on the sinogram domain [3]. In addition, cross-domain learning methods are able to learn CT data features from dual domains to construct models that approximate traditional reconstruction process [4–6].

However, most DL-based CT reconstruction methods are condition-specific, i.e., dose-specific, and geometry-specific. In the dose-specific case, these methods are constructed on the dataset at one specific dose level, which might fail to obtain promising result at other dose levels. Centralized learning via collecting data at different dose levels is an alternative way, but it is difficult to collect sufficient data efficiently. In the geometry-specific case, the DL-based methods, especially the cross-domain learning methods, usually reconstruct the final image from the measured sinogram data with a specific imaging geometry that takes the geometry parameters into account during reconstruction. However, the geometry parameters in the scanner are vendor-specific and different from each other. Then the DL-based methods trained on data from one geometry would fail to be transferred to those from the other geometry due to the different characteristics distributions and big data heterogeneity among different geometries. Xia et al. constructed a framework for modulating deep learning models based on CT imaging geometry parameters to improve the reconstruction performance of the DL models under multiple CT imaging geometries [7], but the method did not consider model degradation due to variations in scanning conditions. Multi-task learning methods can be used to address this issue, but they are limited by the tedious design of auxiliary tasks, which leads to lower efficiency [8, 9], and the privacy issues caused by the sharing of data are also limitations of these methods.

Different from the centralized learning, federated learning has potential to the train model on decentralized data without the need to centralized or share data, which provides significant benefits over centralized learning methods [10–13]. Federated learning has made achievements in medical imaging [14] and applications in CT reconstruction, for example, Li et al. presented a semi-centralized federated learning method to promote the generalization performance of the learned global model [15], Yang et al. propose a hypernetwork-based federated

learning method to construct personalized CT imaging models for local clients [16]. However these methods are constructed on image domain and do not consider the perturbations of multi-source CT data on the sinogram domain, thus the local specificity is insufficient and the generalization of the model still needs to be improved.

Inspired by the previous work [4] and federated learning framework, we propose a condition generalization method under a federated learning framework to reconstruct CT images from different conditions, i.e., different dose levels, different geometries, and different sampling schemes. The proposed method is termed as federated condition generalization (FedCG). Specifically, the proposed FedCG method can be treated as the extension of iRadonMAP [4] to the federated learning framework. And it leverages a cross-domain learning approach: individual-client sinogram learning, and cross-client image reconstruction for condition generalization. In each client, the sinogram at each individual condition is processed similarly to that in iRadonMAP, then the latent characteristics of reconstructed CT images at all conditions are processed through the framework via a condition generalization network in the server. The condition generalization network considers latent common characteristics in the CT images at all conditions and preserves the client-specific characteristics in each condition. Different from the existing FL framework, the server in the proposed FedCG holds a large amount of labeled data that is closer to real world. We validate the proposed FedCG method on two simulation studies, including three different dose levels at the same geometry, and three different sampling schemes at the different geometries. The effectiveness of FedCG has been validated with significant performance improvements on both tasks compared with a number of competing methods and FL methods, as well as comprehensive ablation studies.

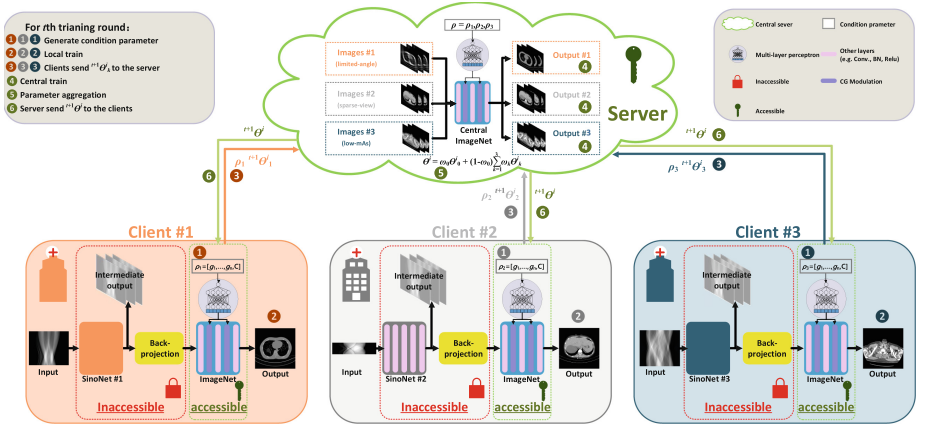


Fig. 1. Pipeline of the proposed FedCG.

2 Method

2.1 iRadonMAP

As a cross-domain learning framework for Radon inversion in CT reconstruction, iRadonMAP consists of a sinogram domain sub-network, an image domain sub-network, and a learnable back-projection layer, which can be written as follows [4]:

$$\tilde{\mu} = F^i \{ F^{R^{-1}} [F^s(p; \theta^s); \theta^{R^{-1}}]; \theta^i \}, \quad (1)$$

where $\tilde{\mu}$ is the final image, p is the sinogram data, $F^s(\cdot; \theta^s)$ and $F^i(\cdot; \theta^i)$ denote the sinogram domain sub-network (i.e., SinoNet) and the image domain sub-network (i.e., ImageNet) of iRadonMAP, $F^{R^{-1}}(\cdot; \theta^{R^{-1}})$ is learnable back projection layer, the details of $F^{R^{-1}}(\cdot; \theta^{R^{-1}})$ are available in [4]. θ denotes the parameters of networks.

Although iRadonMAP can obtain promising reconstruction result, the generalization is still an area that is poorly exploited. When iRadonMAP is trained on CT data with a particular condition (i.e., specific dose level and imaging geometry) and is inferred on CT data with a different condition, it would become unstable as a result of data heterogeneity among different conditions. Federated learning framework is an alternative strategy to address this issue.

2.2 Proposed FedCG Method

In this study, inspired by the previous work [4] and federated learning framework, we propose a condition generalization method under a federated learning framework (FedCG) to reconstruct CT images from different conditions as an extension of the cross-domain learning framework iRadonMAP in federal learning. Specifically, the proposed FedCG is characterized by the following aspects:

Individual-client Sinogram Learning. Due to the big data heterogeneity among different conditions and data privacy preservation, as shown in Fig. 1, in the proposed FedCG, the sinogram data at each condition is processed via SinoNet and learnable back-projection layer as in the iRadonMAP, and the corresponding parameters are not exchanged to communication and augment data privacy when the clients have unique distributions. Furthermore, the individual-client sinogram learning strategy allows for processing condition-specific sinogram data, which can vary flexibly across clients and among different conditions. Then ImageNet can be utilized to reconstruct final CT images wherein the corresponding parameters are updated with the help of federated learning strategy. The central server collects the information from all clients without directly sharing the private data of each client.

Central Data Guidance Training. Different from the existing federated learning framework, the central server has labeled data that is closer to real world, as shown in Fig. 1. In each training round, the central server model can obtain well-trained model with the labeled data, and then the corresponding parameters in the ImageNet can be updated as follows:

$${}^{t+1}\theta^i = \omega_0 {}^{t+1}\theta_0 + (1 - \omega_0) \sum_{k=1}^K \omega_k {}^{t+1}\theta_k^i, \quad (2)$$

where ω_0 is the weight of central model, $\omega_k = \frac{n_k}{n}$, where n_k denotes the number of local iterations.

Condition Generalization in FL. To fully consider the unique distribution in each client, inspired by Xia et al. [7], we introduce condition generalization (CG) network to learn the deep features across scanners and protocols. Specifically, the CG network in each client generates a specific normalized parameters vector according to the imaging geometry and scanning protocol:

$$\rho = [g_1, \dots, g_n, C], \quad (3)$$

where ρ represents the condition parameter in each client, g_1, \dots, g_n are normalized imaging geometric parameters. C is a parameter that represents protocol parameter (i.e., dose level, sparse views, and limited angles). In the k th local client, ρ_k is fed into the local model along with the input data. And in the central server, all condition parameters are fed into the central model. As shown in Fig. 1, multilayer perceptron (MLP) which consists of fully connected layers map ρ into high-dimensional condition vectors, and the vectors are used to modulate the feature map of the ImageNet as follows:

$$\hat{f} = h_1(\rho)\tilde{f} + h_2(\rho), \quad (4)$$

h_1, h_2 are MLPs with shared parameters. \tilde{f} is the feature map of network layer, \hat{f} is the modulated feature map.

Then, the total loss function of the proposed FedCG can be written as follows:

$$L_{total} = \omega_0 \|\mu_0^* - \tilde{\mu}_0\|_2^2 + (1 - \omega_0) \left(\sum_{k=1}^K \omega_k \|\mu_k^* - \tilde{\mu}_k\|_2^2 + \sum_{k=1}^K \lambda_k \omega_k \|p_k^* - \tilde{p}_k\|_1 \right), \quad (5)$$

where μ_0^* is the noise-free image in the central server, μ_k^* is the noise-free image in the k th client, $\tilde{\mu}_0$ is the output image of central ImageNet, $\tilde{\mu}_k$ is the output image of iRadonMAP in the k th client, p_k^* is the noise-free sinogram in the k th client, \tilde{p}_k is the intermediate output of SinoNet in the k th client, λ is the parameter that controls the singogram loss function in the local client.

3 Experiments

3.1 Dataset

The experiments are carried out on three publicly available datasets (120,000 CT images) [17–19], six private datasets from different scanners (Data #1: 1100 brain CT images, Data #2: 1500 chest CT images, Data #3: 847 body phantom CT images, Data #4: 1600 body phantom CT images, Data #5: 1561 abdomen CT images, Data #6: 1598 abdomen CT images). In the experiment, two different conditions are presented, i.e., different dose levels with different geometries (Condition #1), and different sampling schemes with different geometries (Condition #2).

In the Condition #1, the three publicly available datasets are collected for the central server, Data #1, Data #2, and Data #3 are selected for three local clients (Client #1, Client #2 and Client #3), respectively. We obtained the corresponding low-dose sinogram data at different dose levels from the normal-dose CT images based on the previous study, respectively [20]. The X-ray intensities of Clients #1, #2, #3 are $5e5$, $2e5$, $1e5$. In Condition #2, the simulated limited-angle CT images, sparse-view CT images, and ultra-low-dose CT images from the publicly available datasets are collected for the central server. Data #4 contains simulated limited-angle cases (120 degrees of parallel beam, with full angles of 180 degrees), Data #5 contains simulated sparse-view cases (144 views, with full views of 1152) and Data #6 contains simulated ultra-low-dose cases (X-ray intensity of $5e4$) are collected for the three local clients, respectively. In the both experiments, ninety percent of data are used for training and the remaining for testing for all clients.

3.2 Implementation Details

FedCG is constructed by Pytorch toolbox [21], training with an NVIDIA RTX A6000 graphics processing unit with 48 GB memory, and the CT simulation and reconstruction are carried out by the Astra toolbox [22]. The iRadonMAPs in all the local clients and the ImageNet in the central server are optimized by the RMSProp optimizer, and the learning rate of all the models is $2e-5$. The number of training rounds is set to 1000, and the central ImageNet has 100 iterations per round while each local client has 10. ω_0 in the Eq. 2 is empirically set to 0.6. More details on the imaging geometries and architecture of the iRadonMAP can be found in the supplementary materials.

4 Result

In this work, five algorithms are selected for comparison. The classical FBP algorithm and the iRadonMAPs trained on condition-specific dataset, FedAvg [10], Fedprox [11] and FedBN [12]. Peak signal-to-noise ratio (PSNR), structural similarity index (SSIM) and root mean square error (RMSE) are used to quantify reconstruction performance.

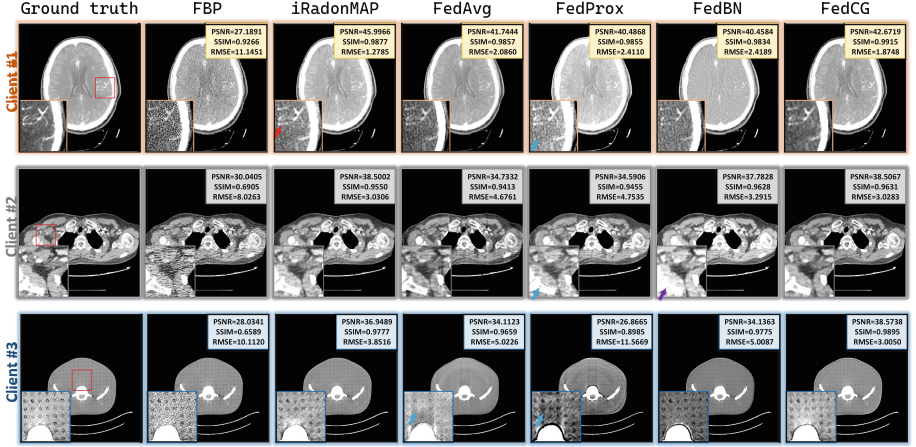


Fig. 2. Reconstruction results of Condition #1. The display windows for CT images at Client #1, Client #2, Client #3 are $[-150, 150]$, $[-150, 150]$, $[-200, 200]$ HU, respectively. The display windows for zoomed-in ROIs are $[-10, 150]$, $[-30, 50]$, $[-150, 150]$ HU, respectively.

4.1 Result on Condition #1

Figure 2 shows results reconstructed by all the competing methods on Condition #1 wherein the normal-dose FBP images are chosen as ground truth for comparison. The results show that the proposed FedCG produces the sharpest images with fine details at all clients, as apparent from the zoomed-in regions of interest (ROIs). Although iRadonMAP can suppress noise-induced artifacts to some extent, it might introduce undesired artifacts as indicated by the red arrows. FedAvg and FedProx can also produce sharp images, but the reconstruction of the fine structure details is less detailed than that in FedCG results. Moreover, the shifted values occur in both FedAvg and FedProx results, as indicated by the blue arrows. FedBN can reconstruct the textures with less noise, but with unsatisfactory performance in the fine texture recovery as indicated by purple arrows. Furthermore, the quantitative measurements also indicate that the proposed FedCG can obtain the best performance among all the competing methods. The possible reason might be that the labeled data in the central server provide sufficient prior information to promote FedCG reconstruction performance. More experimental results are listed in the supplementary materials.

4.2 Result on Condition #2

Figure 3 shows the results reconstructed by all the competing methods on Condition #2 wherein the normal-dose FBP images are chosen as ground truth for comparison. It can be observed that the iRadonMAP can produce promising reconstruction results that are closest to the ground truth as it is trained with the

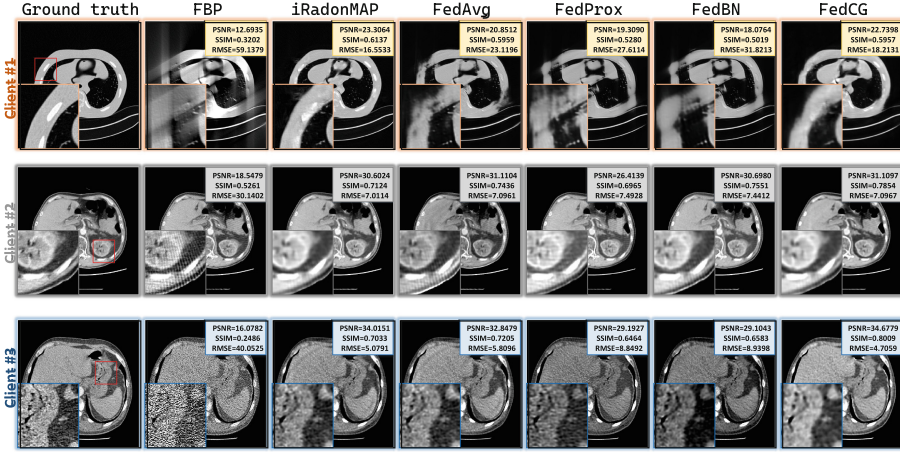


Fig. 3. Reconstruction results of Condition #2. The display windows of Client #1, Client #2, Client #3 are $[-1024, 400]$, $[-100, 100]$, $[-200, 200]$ HU, respectively, with the same in zoomed-in ROIs

geometry-specific data at each client. The proposed FedCG method outperforms other three FL-based methods in terms of artifact suppression and structure details recovery at all cases. The experimental results demonstrate that the proposed FedCG method can simultaneously reconstruct CT images from different geometries and obtain similar performance with the iRadonMAP, indicating strong generalization ability of FedCG.

4.3 Ablation Experiments

ω_0 plays a key role in the proposed FedCG reconstruction performance, then we conduct ablation experiments with different ω_0 on settings Condition #1. Figure 4 shows the mean value of PSNR, SSIM, and RMSE with different ω_0 settings. From the results, when $\omega_0 = 0$, FedCG approaches to FedAvg with poor performance. When $\omega_0 \in [0, 0.2]$, both Client #1 and Client #2 obtain degraded reconstruction performance and Client #3 obtains improved reconstruction performance. And when $\omega_0 > 0.3$, we can see that the reconstruction accuracy increases with increasing ω_0 for Client #1 and Client #2, but reconstruction accuracy for Client #3 degrades with increasing ω_0 . Setting $\omega_0 = 0.6$ yields the best results with the weighted central server parameters leading substantial improvement in all quantitative measurements for Client #1 and Client #2, while setting $\omega_0 = 0.3$ for client #3. The possible reason might be that the similar information are shared among the central server and Client #1, Client #2, but there is large heterogeneity between central server and Client #3. When $\omega_0 > 0.6$, the performance for all clients degrades obviously due to an overly large weight which affects the unique information of each client.

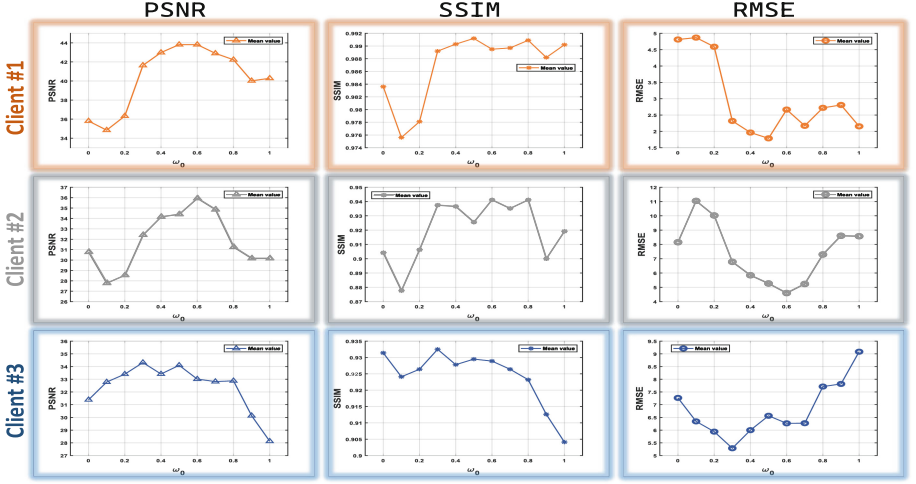


Fig. 4. Mean values of FedCG’s quantitative metrics under different ω_0 .

5 Conclusion

In this work, we propose a condition generalization method under a federated learning framework to reconstruct CT images on different conditions. Experiments on different dose levels with different geometries, and different sampling schemes with different geometries show that the proposed FedCG achieves improved reconstruction performance compared with the other competing methods at all the cases qualitatively and quantitatively.

References

1. Chen, H., Zhang, Y., Kalra, M.K., Lin, F., Chen, Y., Liao, P., Zhou, J., Wang, G.: Low-dose CT with a residual encoder-decoder convolutional neural network. *IEEE Trans. Med. Imaging* **36**(12), 2524–2535 (2017)
2. Peng, S., et al.: Noise-conscious explicit weighting network for robust low-dose CT imaging. In: *Medical Imaging 2023: Physics of Medical Imaging*, vol. 12463, pp. 711–718. SPIE (2023)
3. Meng, M., et al.: Semi-supervised learned sinogram restoration network for low-dose CT image reconstruction. In: *Medical Imaging 2020: Physics of Medical Imaging*, vol. 11312, pp. 67–73. SPIE (2020)
4. He, J., Wang, Y., Ma, J.: Radon inversion via deep learning. *IEEE Trans. Med. Imaging* **39**(6), 2076–2087 (2020)
5. Li, D., Bian, Z., Li, S., He, J., Zeng, D., Ma, J.: Noise characteristics modeled unsupervised network for robust CT image reconstruction. *IEEE Trans. Med. Imaging* **41**(12), 3849–3861 (2022)
6. Chen, S., He, J., Li, D., Zeng, D., Bian, Z., Ma, J.: Dual-domain modulation for high-performance multi-geometry low-dose CT image reconstruction. In: *Medical Imaging 2023: Physics of Medical Imaging*, vol. 12463, pp. 687–693. SPIE (2023)

7. Xia, W., et al.: CT reconstruction with PDF: parameter-dependent framework for data from multiple geometries and dose levels. *IEEE Trans. Med. Imaging* **40**(11), 3065–3076 (2021)
8. Vandenhende, S., Georgoulis, S., Proesmans, M., Dai, D., Van Gool, L.: Revisiting multi-task learning in the deep learning era. *arXiv preprint [arXiv:2004.13379](https://arxiv.org/abs/2004.13379)* vol. 2, no. 3 (2020)
9. Thung, K.H., Wee, C.Y.: A brief review on multi-task learning. *Multimedia Tools Appl.* **77**, 29705–29725 (2018)
10. 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)
11. Li, T., Sahu, A.K., Zaheer, M., Sanjabi, M., Talwalkar, A., Smith, V.: Federated optimization in heterogeneous networks. *Proc. Mach. Learn. Syst.* **2**, 429–450 (2020)
12. 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](https://arxiv.org/abs/2102.07623)* (2021)
13. Peng, X., Huang, Z., Zhu, Y., Saenko, K.: Federated adversarial domain adaptation. *arXiv preprint [arXiv:1911.02054](https://arxiv.org/abs/1911.02054)* (2019)
14. Rieke, N., et al.: The future of digital health with federated learning. *NPJ Digit. Med.* **3**(1), 119 (2020)
15. Li, D., et al.: Semi-centralized federated learning network for low-dose CT imaging. In: *Medical Imaging 2023: Physics of Medical Imaging*, vol. 12463, pp. 1024–1028. SPIE (2023)
16. Yang, Z., Xia, W., Lu, Z., Chen, Y., Li, X., Zhang, Y.: Hypernetwork-based personalized federated learning for multi-institutional CT imaging. *arXiv preprint [arXiv:2206.03709](https://arxiv.org/abs/2206.03709)* (2022)
17. McCollough, C.: TU-FG-207A-04: overview of the low dose CT grand challenge. *Med. Phys.* **43**(6Part35), 3759–3760 (2016)
18. Gava, U., et al.: Unitobrain (2021)
19. Heller, N., et al.: The kits19 challenge data: 300 kidney tumor cases with clinical context, CT semantic segmentations, and surgical outcomes. *arXiv preprint [arXiv:1904.00445](https://arxiv.org/abs/1904.00445)* (2019)
20. Zeng, D., et al.: A simple low-dose x-ray CT simulation from high-dose scan. *IEEE Trans. Nucl. Sci.* **62**(5), 2226–2233 (2015)
21. Paszke, A., et al.: Automatic differentiation in pytorch (2017)
22. Van Aarle, W., et al.: Fast and flexible X-ray tomography using the ASTRA toolbox. *Opt. Exp.* **24**(22), 25129–25147 (2016)