Abstract—In clinical practice, computed tomography (CT) is an important noninvasive imaging technology to provide patients’ anatomical information. However, its potential radiation risk is an unavoidable problem that raises people’s concerns. Recently, deep learning (DL)-based methods have achieved promising results in CT reconstruction, but these methods usually require the centralized collection of large amounts of data for training from specific scanning protocols, which leads to serious domain shift and privacy concerns. To relieve these problems, in this article, we propose a hypernetwork-based physics-driven personalized federated learning method (HyperFed) for CT imaging. The basic assumption of the proposed HyperFed is that the optimization problem for each domain can be divided into two subproblems: local data adaption and global CT imaging problems, which are implemented by an institution-specific physics-driven hypernetwork and a global-sharing imaging network, respectively. Learning stable and effective invariant features from different data distributions is the main purpose of global-sharing imaging network. Inspired by the physical process of CT imaging, we carefully design physics-driven hypernetwork for each domain to obtain hyperparameters from specific physical scanning protocol to condition the global-sharing imaging network, so that we can achieve personalized local CT reconstruction. Experiments show that HyperFed achieves competitive performance in comparison with several other state-of-the-art methods. It is believed as a promising direction to improve CT imaging quality and personalize the needs of different institutions or scanners without data sharing. Related codes have been released at https://github.com/Zi-YuanYang/HyperFed.

Index Terms—Computed tomography (CT), deep learning (DL), federated learning (FL), image reconstruction, physics-driven.

I. INTRODUCTION

COMPUTED tomography (CT) is an important imaging modality in clinical diagnosis since it can noninvasively visualize anatomical information inside the patient’s body. However, with the popularization of CT, related concerns have been expressed about the potential radiation risk that may lead to genetic, cancerous, and other diseases. In practice, switching the voltage/current of X-ray tube and reducing the scanning views are two typically used strategies to lower the radiation dose [1]. However, the above strategies will unavoidably degrade the imaging quality, which produces negative impacts on the subsequent image analysis and diagnosis [2], [3], [4].

Recently, attracted by the encouraging performance of deep learning (DL), researchers are enthusiastic about introducing DL to low-dose CT (LDCT) reconstruction and achieved impressive results [5], [6], [7]. In spite of fruitful results obtained in recent years, current methods are all centralized training (CL)-based and need to collect huge amounts of training data from different institutions without considering privacy issue. Although these data are anonymously transferred, many works have proven that anonymization cannot effectively protect the patients’ privacy [8], [9], [10]. Restricted to privacy, legal, and ethical concerns, since it is difficult to train a stable and effective model just based on a small number of local data, a privacy-preserving multiinstitutional cooperative training method is highly demanded.

As a recently proposed decentralized solution, federated learning (FL) aims to protect data privacy and confidentiality while the model can learn enough information from multiple data sources [11]. The greatest difference between the training processes of CL-based and FL-based frameworks is the transferred contents. In CL-based methods, the patients’ private data is transferred, but in FL-based methods, only gradients are transferred, and the privacy contained in the gradients is negligible compared with the contents transferred in CL-based methods.

One main challenge in FL lies in that the data from different sources are usually nonindependent, identically distributed (non-iid), in which this problem is more severe in CT imaging than the other analysis tasks. Due to the differences in both...
hardware and scanning protocol, the data collected from different scanners or institutions inevitably suffer from this problem. Unfortunately, the global model generated by FL methods can only capture some common statistical characteristics from different institutions and cannot provide specific features for different data sources.

Recently, personalized FL methods have been proposed to address the non-iid problem. These methods aim to train personalized local models rather than a global-shared model based on diverse data distributions. Some approaches utilize hypernetworks to achieve personalization. Generally, hypernetwork aims to generate personalized weights of the targeted network. However, existing hypernetwork-based FL methods usually neglect the physical process and suffer from high additional training costs.

The computer-aided diagnosis framework typically consists of both upstream tasks, such as imaging, and downstream tasks, including segmentation and classification. It is crucial to achieving accurate results in the upstream phase as it directly influences the accuracy of downstream tasks and ultimately aids the subsequent diagnoses. In this article, we focus on proposing a general FL framework for upstream tasks, including postprocessing and reconstruction. Specifically, we propose a hypernetwork-based physics-driven personalized federated learning framework (HyperFed) for CT imaging. The basic assumption of this work is that the global optimization problem can be decomposed into two subproblems, the personalization feature adaption, and the invariant global imaging feature extraction problems. Actually, the scanned targets (human body) are structurally similar, enabling us to effectively leverage vast amounts of data from different clients to collaboratively train a globally shared imaging network for extracting invariant global imaging features, which implicitly contain stable structural information. Since the CT physical imaging process and the imaging quality are highly related to the scanning protocol and geometry parameters [12], the data distributions and reconstructed image qualities are highly associated with these scanning parameters. Based on this observation, we design a physics-driven hypernetwork to fully explore the prior knowledge implicitly contained in the scanning parameters to modulate imaging features for different local data distributions. Hence, the proposed HyperFed comprises two main modules: the institution-specific physics-driven hypernetwork and the global-sharing imaging network. For each institution, corresponding scanning parameters are transformed as personalization features and fed into the institution-specific physics-driven hypernetwork to modulate invariant features extracted by the imaging network, which is learned from different data domains. Since the input dimension of the imaging network is much higher than the hypernetwork, it is hard for the imaging network to converge to a satisfactory solution only based on a small number of local data. One possible solution is to introduce the idea of FL and share the gradients globally to train a robust global model with the data from multiple sources. Institution-specific physics-driven hypernetwork does not participate in the global aggregation and facilitate the personalization for different data domains, and the global-sharing imaging network is responsible for learning the global universal features from different institutions.

The main contributions of this article can be summarized as follows.

1) A HyperFed is proposed for CT imaging in this article. To our best knowledge, this is the first attempt for personalized CT imaging based on FL.
2) The proposed HyperFed is comprised of two main modules, the institution-specific physics-driven hypernetwork and the global-sharing imaging network. The hypernetwork is carefully designed for personalization to alleviate non-iid, and the global-sharing imaging network aims to learn stable and invariant features from different data distributions.
3) The proposed framework is flexible, which can be easily extended to different CT imaging tasks, such as postprocessing and reconstruction.

II. RELATED WORKS

A. CT Image Reconstruction

CT imaging technology has witnessed remarkable innovations in the past decades and improved the diagnostic performance substantially [13]. CT reconstruction methods can be roughly divided into sinogram filtration, iterative reconstruction (IR), and postprocessing.

For sinogram filtration methods, either raw data or log-transformed data need to be filtered before reconstruction, such as structural adaptive filtering and penalized weighted least-squares [12]. For most IR methods, the key idea of solving the reconstruction problem is formulated by combining the prior knowledge to optimize an objective function [14], such as the family of total variation (TV) [15], nonlocal means filter [16] and some other regularization terms [17]. Postprocessing-based methods do not need to have access to the raw data, so they are convenient to be deployed into existing CT systems [18]. These methods achieve satisfactory performance if the inputs strictly follow their prior hypothesis, but they still suffer from high computational complexity and limited flexibility.

Recently, attracted by the impressive performance of DL in different fields, researchers are enthusiastic about introducing DL to LDCT reconstruction and achieved encouraging results. For example, Chen et al. [19] introduced the residual structure into the convolutional autoencoder and proposed a residual encoder-decoder convolutional neural network (RED-CNN) for low-dose denoising. You et al. [20] incorporated convolutional neural network (CNN), residual learning as well as generative adversarial network (GAN) techniques for CT super resolution. Yang et al. [21] trained a GAN with perceptual loss, which is defined as the distance between the reconstructed and the reference images in feature space and achieved satisfactory performance. Zhang et al. [22] extended GAN to reconstruct images from two domains by designing a comprehensive generator. Chen et al. [23] unrolled the IR model optimized by gradient descent method into a network...
and learned the regularization terms and balancing parameters using training data. According to the work proposed by [24], the modified proximal forward-backward splitting method was unrolled into a residual reconstruction network corresponding to the updates of data fidelity and regularization terms. Xia et al. [25] proposed a parameter-dependent framework (PDF) to introduce the scanning parameters into CT image denoising/reconstruction and achieved remarkable improvement. In spite of fruitful results obtained in recent years, current models are all CL-based, which is at a risk of privacy leakage during the process of data collection and severely restricts the promotion of DL-based methods in real clinical scenarios.

B. Federated Learning

FL is a promising decentralized solution that aims to preserve data privacy while the model can learn enough information from multiple data sources [26], [27], [28]. Typically, McMahan et al. [29] proposed a classical FL method FedAvg, which learns the global model by averaging local models of different parties. FedProx [30] is considered as improved FedAvg, which constrains the local models close to the global model. Similarly, Li et al. [31] proposed the model-contrastive (MOON) method to constrain the local models close to the global model by minimizing the contrastive distances between them. Benefiting from the privacy-preserving characteristic, FL was introduced in different medical tasks in recent years [32], [33], [34], [35]. We also notice that there are some contemporary works about medical imaging. For example, Guo et al. [36] proposed an intermediate latent feature alignment method for magnetic resonance image (MRI) reconstruction. However, this method needs to know the target domain data, which is difficult in practice. Besides, Sattler et al. [37] and Dinh et al. [38] introduced novel approaches aimed at mitigating the non-iid problem in multitask learning within the context of FL. The mentioned methods above all aimed to train a global-shared model, so they may suffer from serious non-iid problems.

C. Personalized FL

To relieve the non-iid problem, several works were witnessed in personalized FL. For example, Zhang et al. [39] proposed a semi-asynchronous FL framework for short-term solar power forecasting and achieved satisfactory performance. Huang et al. [40] proposed to use unlabeled public data for communication and construct cross-correlation matrix to learn a generalizable representation. Liang et al. [41] proposed LG-FedAvg to combine local representation learning and global federated training. However, the local representation methods were designed for high-level tasks, which cannot be embedded into low-level imaging tasks. In [42], Ditto was proposed to simultaneously learn local and global models via a global-regularized multitasks learning framework and achieved impressive performance. Ma et al. [43] proposed to identify the importance of each layer from different clients to optimize the personalized model aggregation. Li et al. [44] proposed FedBN, which alleviates the feature shift using local batch normalization (BN) to achieve personalization in local institutions. Shamsian et al. [45] proposed pFedHN, which utilized a hypernetwork to generate the parameters of local models, which works well in simple local models. However, this method is difficult to keep its performance in CT imaging, due to the increased enormous parameters. Inspired by pFedHN and FedBN, Li et al. [46] proposed using a hypernetwork to generate personalized projection matrices for self-attention layers, allowing the personalization of client-specific queries, keys, and values. However, this method’s applicability is restricted to its strict requirement that the client-side models must be transformer-based, which heavily limits its generality. Zhao et al. [47] proposed to learn a personalized feature space for each client by identifying which models should be selected for collaboration. Chen et al. [48] proposed a cyclic knowledge distillation method to distill the high-level semantic features from other local models to achieve personalized without a central server. However, the above methods are more suitable for classification tasks, where the high-level semantic features of each category in different clients are typically similar and not easily applicable to imaging tasks. Hanzely and Richtárik [49] added the regularization term to calculate the distance between local models and the global model to control the optimization degree. Some researchers also introduced personalized FL into medical tasks [50]. For example, researchers proposed several FL-based COVID-19 diagnosis methods [51], [52], [53], but these tasks were designed for the analysis task not imaging task. Feng et al. [54] proposed a personalized model for MRI denoising, which shares the encoder rather than the whole model. However, CT reconstruction is heavily impacted by the scanning protocol and geometry parameters, which is quite different from MRI reconstruction and leads to a serious non-iid problem. Besides, this method cannot be embedded into IR-based methods, which greatly limits its promotion. To relieve the above problems, we propose a physics-driven personalized CT reconstruction framework in this article. To our best knowledge, it is the first attempt in this field. Furthermore, the proposed model is flexible which can be easily extended to different CT imaging tasks, such as postprocessing and reconstruction.

III. METHODOLOGY

A. Problem Statement

In general, CT reconstruction can be formulated as [55]

$$\min_{x} \frac{1}{2} \|Ax - y\|_2^2 + R(x)$$

where $\|\cdot\|_2^2$ is the $L_2$ norm and $A$ represents the system matrix. $x$ and $y$ denote the image be reconstructed and the measured data, respectively. $R(\cdot)$ denotes the regularization term, which is usually formulated with prior knowledge.

Besides, LDCT image denoising can be formulated as [56]

$$\min_{w} \|F(x_l, w) - x_n\|_2^2$$

where $x_l$ is the low-dose image and $x_n$ is the corresponding normal-dose image. $F$ is the objective model to push $x_l$ to
x_n as close as possible. w denotes the parameters of imaging network.

For personalized FL, institutions attempt to get personalized models based on local data and improve the personalized performance in a certain FL way, such as sharing some layers. Assuming that there are K institutions, the learning process can be formulated as [57]

\[
\min_{\theta_1, \ldots, \theta_K} \left( \sum_{k=1}^{K} \frac{1}{p_k} \mathbb{E}_{(x_i, y_i) \sim D_k} \left[ \| F_k(x_i, y, \theta_k) - x_n \|^2 \right] \right) \tag{3}
\]

where \( F_k \) represents the optimization model in the kth institution, \( p_k \) stands for the weight of the kth institution in global optimization, and \( D_k \) denotes the dataset in the kth institution. The purpose of personalized FL is to search the optimal local models for different institutions without data sharing.

### B. Architecture of HyperFed

Since scanning protocol and geometry parameters heavily impact the physical CT reconstruction process, it is reasonable to consider that these parameters contain information to guide the imaging network to predict the normal-dose CT. This idea motivates us to take full advantage of these parameters to alleviate non-iid problem and improve imaging performance. Specifically, the architecture of HyperFed is composed of an institution-specific physics-driven hypernetworks and a global-sharing imaging network. The learning process of current imaging networks can be considered as \( F_{\text{imag}}(x, y, w) \). The inputs are fed into the model \( F_{\text{imag}} \) parameterized with \( w \), and no constraint is applied. It is a common trick to reduce the complexity when all the inputs strictly follow a uniform distribution. However, it is hard to meet this assumption in real situations and non-iid problem usually appears. To deal with this problem, the institution-specific hypernetwork is implemented as a regulator to modulate the learning process. Then, the learning process of HyperFed is reformulated as \( F_{\text{imag}}(x, y, w, \theta) \), where \( \theta \) is output from the hypernetwork \( H(\cdot) \) parameterized with \( \xi \). Specifically, the physical parameter vector \( g \), which contains important scanning and geometry parameters in physical imaging, is fed into \( H(\cdot) \) to generate the scaling factor set \( \gamma \) and the bias set \( \beta \), which are used to modulate the features of \( F_{\text{imag}} \). This process is formulated as

\[
\gamma, \beta = H(g, \xi) \tag{4}
\]

For simplicity, in this article, \( g \) includes the number of detector bins, pixel length, detector bin length, the distance between the source and rotation center, the distance between the detector and rotation center as well as the photon number of incident X-rays. Normalization and logarithmization are performed for some elements in \( g \) with large magnitudes, such as the numbers of view and detector bins and the photon number of incident X-rays. The normalization is defined as

\[
g_j = \frac{g_j - \min(g_i)}{\max(g_i) - \min(g_i)} \tag{5}
\]

where \( j \) is the element index in \( g \), \( \max(\cdot) \) and \( \min(\cdot) \) represent the maximum and minimum functions, respectively.

Fig. 1 shows the architecture of our proposed HyperFed. Since the proposed hypernetwork is a plus-and-play network, its structure is highly related to the imaging network. In this article, our hypernetwork is a three-layer fully-connected network, which is used to generate \( \gamma \) and \( \beta \). The input of the hypernetwork is a 7-D physical parameter vector \( g \), and the output sizes of the following two linear layers are 256 and 512, respectively. The final output dimensionality depends on the imaging network’s feature dimension, which is double the imaging feature dimension. To reduce the number of learnable parameters, in RED-CNN, encoder and decoder layers share two groups of modulation factors, and all IR blocks share the same modulation factors in LEARN.\(^1\) Then, the modulation function can be formulated as follows:

\[
F_{\text{imag}}^*(y) = \gamma F_{\text{imag}}(y) + \beta. \tag{6}
\]

The modulation operation is applied on the feature maps from different modules in the imaging network and (6) is reformulated as

\[
F_{\text{imag}}^*(y) = \gamma_i F_{\text{imag}}(y) + \beta_i \tag{7}
\]

where \( F_{\text{imag}}(y) \) and \( F_{\text{imag}}^*(y) \) represent the feature map from the \( i \)th module and its corresponding modulated feature map, respectively. \( \gamma_i \) and \( \beta_i \) represent the regularization factor and bias for the \( i \)th module, respectively.

As we mentioned above, the imaging network in the proposed HyperFed is flexible for different tasks. Imaging units in the proposed architecture vary for different imaging methods. For example, the imaging unit can be convolution layers for the postprocessing method, such as RED-CNN [19]. On the other side, for the unrolled iteration methods, such as LEARN [23], it denotes an unrolled iteration module.

### C. Implementation of HyperFed

In this article, we propose a hypernetwork-based personalized FL framework to relieve non-iid problem in CT image reconstruction. Similar to other FL methods, such as FedAvg and FedProx, HyperFed updates the hypernetwork and imaging network locally and only averages the imaging network in the server. The concept of the training process is depicted in Fig. 2. Similar to FedBN, the data normalization method is

\(^1\)To ensure reproducibility and help readers to understand each detail, we will release codes about our all model combinations.
Algorithm 1 Main Steps of HyperFed

\textbf{Input}: $D \triangleq \cup_{k \in K} D^k$, data from $K$ institutions; $T$, the number of communication rounds; $E$, the number of local epochs.

\textbf{Output}: The optimized parameters $w$ of the imaging network, and $\xi_k$ of the hypernetwork in the $k$th institution.

1: Server executes:
2: \hspace{1em} Randomly initialize $w^{ini}$ and $\xi^{ini}$ on the server and deliver them to each institution.
3: \hspace{1em} for $t = 1, 2, \ldots, T$ do
4: \hspace{2em} for $k = 1, 2, \ldots, K$ in parallel do
5: \hspace{3em} send $w^t$ to $k$th institution
6: \hspace{3em} $w_k^t \leftarrow \text{Institution Local Training}(k, w^t)$
7: \hspace{2em} end for
8: \hspace{1em} $w^{t+1} \leftarrow \sum_{k=1}^{K} |D^k| \text{avg} w_k^t$
9: end for
10: Institution Local Training:
11: $w_k^{(1,1)} \leftarrow w^t$
12: for $e = 1, 2, \ldots, E$ do
13: \hspace{1em} for $(y, x, y_n)$ in $D^k$ do
14: \hspace{2em} $\gamma, \beta \leftarrow H_L(g)$
15: \hspace{2em} $l \leftarrow \text{MSE Loss}(F_{\text{imag}}(y) + \beta, x_n)$
16: \hspace{2em} $w_k^{(e+1)}, \xi_k^{(e+1)} \leftarrow \text{Adam}(w_k^{(e)}, \xi_k^{(e)}, \eta)$
17: \hspace{2em} end for
18: \hspace{1em} end for
19: $w_k^t \leftarrow w_k^{(E)}$
20: return $w_k^t$ to the server

D. Convergence Analysis

In this section, we provide a convergence analysis for Algorithm 1. At first, we define that $L^k(\delta^k) \triangleq \mathbb{E}_{d^k \sim D^k} [L^k(\delta^k, d^k)]$, and $\delta^k$ denotes the set of parameters of imaging network and hypernetwork at $k$th institution at $t$th communication round, and $G = \mathbb{E}_{d^k \sim D^k} \nabla L^k(\delta^k, d^k)$. The other definitions can be found in Algorithm 1. Then, there are two basic assumptions, which are usually employed in current FL literature [44], [59], [60], [61].

Assumption 1 (Smoothness): Each local loss function $L^k$ is smooth with modulus $M$.

Assumption 2 (Bounded Variances and Second Moments): There exists constants $\sigma > 0$ and $G > 0$ such that $\mathbb{E}_{d^k \sim D^k} \nabla L^k(\delta^k, d^k) \leq G^2, \forall \delta^k, \forall k$.

Lemma 1: Under Assumption 1, Algorithm 1 can ensure that $\mathbb{E}[\|\delta^k - \delta^e\|_2^2] \leq 10\lambda^2E^2G^2, \forall k, \forall t$.

Proof: For the convenience of proof, we introduce intermediate variable $\delta^t_{\text{imp}} = (w_{t-1}, \xi_{t-1})$, where $w_{t-1} = (1/N) \sum_{k=1}^{N} w_k^{t-1}$. According to the lemma 1 of [62], we know that $\mathbb{E}[\|\delta^t_{\text{imp}} - \delta^e_{\text{imp}}\|_2^2] = \mathbb{E}[\|w_{t-1} - w_{t-1}^e\|_2^2] \leq 4\lambda^2E^2G^2, \forall k, \forall t$. Furthermore, we note that $\delta^e_{\text{imp}}$ can be obtain by

$$\delta^k = \delta^t_{\text{imp}} - \lambda \sum_{e=1}^{E} G^k,$$  

Thus, we have

$$\mathbb{E}[\|\delta^k - \delta^e\|_2^2] \leq 10\lambda^2E^2\mathbb{E}[\|G^k\|_2^2]$$

performed at local in HyperFed. We note that the proposed HyperFed is expected to predict the reconstruction results for different scanners based on different critical physical parameters. However, these measured CT data follow different distributions and is against the assumption of FedBN that BN is helpful if the output obeys a specific distribution [58]. This mismatching will compromise the model performance. To circumvent this obstacle, the hypernetwork is proposed to modulate the feature maps of the imaging network, which can be approximately considered as physics-driven self-normalization.

As discussed above, the cooperation among different institutions is hard to implement due to the privacy and security concerns in the real situation. HyperFed is proposed to allow different institutions to cooperate in training an imaging network in a privacy-preserving way. In HyperFed, local data can only be accessed by its own institution, and only the local gradients are aggregated in the server, which can effectively protect patients’ privacy. In the proposed HyperFed, the model is initialized in the server and delivered to different institutions at first. For each institution, the local model is trained with its own data by minimizing the following loss:

$$L^k = \mathbb{E}_{(x,y) \sim D^k} \frac{1}{2} \|F^k(\delta^k, y, g) - x\|^2_F$$  

where $F^k(\cdot)$ denotes the local network (including both hypernetwork and imaging network) at the $k$th institution parameterized by $\delta^k$.

Assuming $w_k$ and $\xi_k$ represent the parameters of $F_{\text{imag}}$ and $H(\cdot)$ at the $k$th institution, respectively, the parameter optimization can be formulated as

$$w_{k}^{p+1}, \xi_{k}^{p+1} = (w_{k}^{p}, \xi_{k}^{p}) - \lambda \nabla (w_{k}^{p}, \xi_{k}^{p}) L^k$$  

where $w_{k}^{p}$ and $\xi_{k}^{p}$ represents the $p$th training epoch of $w_k$ and $\xi_k$, and $\lambda$ stands for the learning rate.

After every several training epochs, the gradient of the imaging network will be delivered to the server for aggregation, and the hypernetwork is preserved at local for modulation. The main steps of our proposed HyperFed are listed in Algorithm 1.

Fig. 2. Concept of the training process in the proposed HyperFed.
Based on (11), we have
\[ \mathbb{E}\left[ \left\| \delta_t^i - \delta_{t-1}^i \right\|^2 \right] \leq 10M N^2 \lambda^2 E^2 G^2. \]

Proof: Based on Lemma 1, we can have
\[ \mathbb{E}\left[ \left\| \sum_{k=1}^{N} \nabla L^k (\delta_t^i) - \sum_{k=1}^{N} \nabla L^k (\delta_{t-1}^i) \right\|^2 \right] \leq 10M N^2 \lambda^2 E^2 G^2. \]

So far, we have proved that the gradient between adjacent models is bounded. Furthermore, we need to prove that the entire training process is bounded. Then, we can get
\[ \mathbb{E}\left[ \left\| \sum_{i=1}^{T} \sum_{k=1}^{N} \nabla L^k (\delta_t^i) - \sum_{k=1}^{N} \nabla L^k (\delta_{t-1}^i) \right\|^2 \right] \leq 10M N^2 T^2 \lambda^2 E^2 G^2. \]
TABLE I

| Geometry Parameters and Dose Levels in Different Institutions for Different Tasks (Postprocessing) |
|--------------------------------------------------|
| Institution #1 | Institution #2 | Institution #3 | Institution #4 | Institution #5 |
|----------------|----------------|----------------|----------------|----------------|
| Number of views | 512 \ 1024 | 512 \ 88 | 364 \ 1024 | 400 \ 128 | 384 \ 108 |
| Number of detector bins | 368 \ 315 | 315 \ 768 | 330 \ 768 | 350 \ 512 | 350 \ 512 |
| Detector bin length (mm) | 1.33 \ 0.66 | 1.40 \ 0.78 | 1.39 \ 1.00 | 1.20 \ 1.20 | 1.40 \ 0.50 |
| Distance between the source and rotation center (mm) | 2.57 \ 0.72 | 3.00 \ 0.58 | 2.60 \ 0.62 | 2.20 \ 1.40 | 2.50 \ 0.40 |
| Distance between the detector and rotation center (mm) | 2.59 \ 0.50 | 450 \ 350 | 400 \ 500 | 500 \ 400 | 500 \ 400 |
| Intensity of X-rays | 0.5e5 \ 1e5 | 0.6875e5 \ 1e6 | 0.875e5 \ 3e4 | 1.0625e5 \ 2.5e5 | 1.25e5 \ 5e5 |

Fig. 5. Results of w/o FL, FedAvg, FedProx, FedBN, Ditto, pFedHN, and HyperFed under different geometries and dose levels for the postprocessing task. The display window is $[-160, 240]$ HU.

is equal, and only their noise distributions are different, which is caused by different settings of scanning protocols or scanners. In Fig. 4, it can be seen that the heterogeneity is noticeable among different CT scans in both postprocessing and reconstruction cases.

For simplicity, we assume that each institution only has one type of data, and the data transmission is strictly prohibited. RED-CNN [19] and LEARN [23], which are one of the most representative methods in postprocessing and unrolled reconstruction networks are adopted, respectively. The learning rate is set to $1 \times 10^{-4}$, the number of local training epochs is set to 3, and the number of communication rounds are set to 1000 and 200 for RED-CNN and LEARN, respectively. We compare the proposed HyperFed with FedAvg [29], FedProx [30], FedBN [44], Ditto [42], pFedHN [45], and the original imaging models without FL, which is dubbed as w/o FL. For FedProx, the penalty constant hyperparameter is set as $1 \times 10^{-4}$. FedBN achieves personalization with BN layers, which are added after each convolution layer of the imaging networks in this article. Adam [64] is used to optimize the network, and mean squared error (MSE) is adopted as the loss function. All codes are implemented in PyTorch and the experiments are performed on a NVIDIA GTX 3090 GPU.

B. Experiments for the Postprocessing Task

Fig. 5 shows several typical slices denoised using different methods. It can be seen that the reconstructed images using FedAvg and FedProx still contain some noise or artifacts. The possible reason lies in that these methods are apt to extract global features rather than personalized features for different local models. Since the dataset for RED-CNN without FL for each institution only has 80 samples, which is relatively small, the details of the results are not well preserved. FedBN performs better than FedAvg and FedProx, since the personalized BN layers can modulate the global features. Even though its personalized performance is better than FedAvg and FedProx, it still cannot effectively remove all the noise and artifacts. Although Ditto has better results than FedAvg and FedProx, its performance is still limited. The reason lies in that its local models are heavily affected by its global model. Besides, its
Table II

|                  | w/o FL | FedAvg | FedProx | pFedHN | HyperFed |
|------------------|--------|--------|---------|--------|----------|
|                  | PSNR   | SSIM   | PSNR    | SSIM   | PSNR     | SSIM   |
| Institution #1   | 36.57  | 0.9306 | 37.81   | 0.904  | 36.78    | 0.9317 |
| Institution #2   | 39.21  | 0.9544 | 37.74   | 0.9417 | 36.65    | 0.9342 |
| Institution #3   | 38.68  | 0.9512 | 38.11   | 0.9425 | 37.06    | 0.9317 |
| Institution #4   | 40.23  | 0.9614 | 38.77   | 0.948  | 38.16    | 0.9458 |
| Institution #5   | 39.31  | 0.9551 | 38.11   | 0.9438 | 37.20    | 0.9393 |
| Average          | 38.88  | 0.9545 | 38.10   | 0.9433 | 37.17    | 0.9365 |
| STD              | 2.10   | 0.0679 | 1.69    | 0.0650 | 1.62     | 0.0674 |
| CC               | 0.9989 | 0.9987 | 0.9984  | 0.9989 | 0.9989   | 0.9986 |
|                  |        |        |         |        |          |        |

Fig. 6. Horizontal profiles of the results of w/o FL, FedAvg, FedProx, FedBN, Ditto, pFedHN, and HyperFed.

Fig. 7. Boxplots of PSNR based on w/o FL, FedAvg, FedProx, FedBN, Ditto, pFedHN, and HyperFed for the postprocessing task. (a)–(e) Results of institution #1–#5, respectively, and (f) average result of all institutions.

The quantitative results of the whole testing set are shown in Table II. It can be seen that HyperFed achieved the best performance in most institutions compared with other methods. Furthermore, we offer boxplots in Fig. 7 to demonstrate the stability of different methods. It can be noticed that HyperFed shows better stability and performance in comparison with other methods. The main reason is that the imaging network of HyperFed learns from different sources and has more data than the original RED-CNN, and the results support that our imaging network can extract stable features. Although both FedAvg and FedProx are trained with mixed data, the performance is not satisfactory due to the diversity of data sources. Although the quantitative results of Ditto are relatively good, its results lose lots of details, as shown in Fig. 5. All parameters generated from the client index, which contains limited scanning information, and enormous model parameters make training difficult. Both factors limit the performance of pFedHN. Personalized BN layers can learn the modulation parameters to help local models improve personalized performances. However, BN layers are not perfectly suitable for CT imaging, since the outputs of different institutions do not obey the specific distribution, resulting in an unstable imaging performance.

C. Experiments for the Reconstruction Task

In this section, LEARN, which belongs to unrolled reconstruction network, is included to validate the performance.
Fig. 8. Results of w/o FL, FedAvg, FedProx, FedBN, and HyperFed under different geometries and dose levels for the reconstruction task. The display window is [−160, 240] HU.

TABLE III

|                        | w/o FL | FedAvg | FedProx | FedBN | HyperFed |
|------------------------|--------|--------|---------|-------|----------|
| **PSNR**               |        |        |         |       |          |
| Institution #1         | 46.26  | 44.47  | 43.87   | 45.25 | 45.74    |
| Institution #2         | 39.63  | 39.35  | 38.59   | 38.76 | 39.74    |
| Institution #3         | 38.79  | 38.96  | 38.73   | 38.78 | 39.47    |
| Institution #4         | 38.41  | 40.51  | 40.51   | 40.45 | 40.71    |
| Institution #5         | 40.46  | 40.99  | 40.42   | 38.78 | 41.39    |
| **Average**            | 40.71  | 40.54  | 39.91   | 40.30 | 41.44    |
| **SSIM**               | 0.9889 | 0.9544 | 0.9771  | 0.9889| 0.9853   |
| **STD**                | 0.9991 | 0.9992 | 0.9990  | 0.9991| 0.9994   |

and generalization of the proposed HyperFed. However, the comparison with pFedHN is not conducted in this subsection. This is primarily due to the substantial training cost of the combination of the IR-based methods and pFedHN that our hardware (one NVIDIA GTX 3090 GPU) cannot afford it. Discussions regarding external learning parameters can be found in Section IV-F, which explore the impact of these external parameters on the generality. Some typical slices reconstructed using different methods are shown in Figs. 8 and 9 demonstrates the profiles along the green dotted lines in Fig. 8. Similar to the results of postprocessing task, it can be observed that there are still some noise and artifacts left in the results of FedAvg and FedProx, especially for the sparse-view cases. Due to the non-iid problem, simply combining with FL and adopting more training data from different sources cannot improve the performance and even degrade the results a little. For FedBN, limited improvement is achieved compared with FedAvg and FedProx. As mentioned above, the reason lies in that the output distributions of different institutions are not specific, which is against the basic assumption of FedBN. Our method still performs best for the unrolled reconstruction network in both visual inspection and quantitative matrices.

The quantitative results of the whole testing set are shown in Table III. It can be seen that HyperFed achieved the best
performance in most cases compared with other methods. As shown in the boxplots in Fig. 10, HyperFed has the most stable performance in the comparison with other methods. The original LEARN without FL achieves good average quantitative performance and satisfactory imaging results in some institutions, but it cannot maintain its performance in each institution. Compared with the original LEARN, the performance of FL-based methods is more stable for different institutions. HyperFed integrates the advantages of CL and FL, achieving promising imaging performance in each institution.

From the above experiments for both postprocessing and reconstruction tasks, we can see that it is difficult to achieve satisfactory imaging performance without considering non-iid problem and directly applying existing FL methods cannot efficiently explore the power of big data. It is reasonable to divide the global optimization problem into the global imaging problem and the local feature domain adaption problem. Specifically, $F_{\text{imag}}$ is designed for learning some common imaging features from different institutions, and $H(\cdot)$ attempts to modulate the extracted features from $F_{\text{imag}}$ to alleviate the non-iid issue. Meanwhile, results support our motivation that physical parameters may contain much information that can be used to modulate the invariant global imaging features.

### D. Experiments on the Large-Scale Training Set

As shown in the previous experiments, some FL-based methods and the original imaging models without FL cannot work well with a small local training set. In this section, we compare the performance of these methods trained with a large-scale local dataset. RED-CNN is used as the backbone network. The other settings about hyperparameters are exactly the same as the previous experiments and the only difference is the size of the training set. Similar to the previous training settings, patients are randomly divided into five groups and each institution contains 200 images to simulate the training data. Compared with other methods, HyperFed can still keep its remarkable performance, which supports that the proposed architecture can improve the imaging quality regardless of the size of the training set.

Table IV shows the quantitative results, and all the methods can achieve improved performance due to the increase of training samples. In this case, HyperFed still has competitive performance compared with other methods. However, pFedHN cannot maintain its performance in each client, because its hypernetwork needs to generate too many parameters, which greatly increases the task difficulty. Although the original RED-CNN achieves better results in some cases, its stability is not improved as shown in the profiles and boxplots in Figs. 13 and 12, respectively. FedAvg and FedProx maintain their stability in all the experiments, but the problem for these methods is that they cannot adjust extracted features based on the characteristics of different domains. In other words, the cost of stability is the personalization. The two methods cannot work well for some specific cases and lack of personalization leads to unsatisfactory reconstruction. The reason lies in that these methods attempt to minimize the average imaging loss of different institutions rather than the best local imaging reconstruction. FedBN cannot achieve significant benefits from the expansion of the training set like FedAvg and FedProx. FedBN attempts to use personalized BN layers to improve the personalized performance, but with the expansion of the training set, the gaps between output distributions of different institutions are also enlarged, compromising the model performance. As shown in the results, our hypernetwork can take full use of the CT geometry parameters to modulate the feature maps of the imaging network. On the other words, HyperFed can balance well between the stability and the imaging features.
performance in the learning process. HyperFed integrates the advantages of both training modes and achieves stable and competitive performance.

E. Ablation Study

In this section, we have conducted various ablation studies to demonstrate the effectiveness of different components in HyperFed. Related experiments follow the same experimental settings as described in Section III-B, and the results are presented in Table V. The symbols “+” and “−” represent the imaging networks without and with the hypernetwork, respectively. It can be observed that the hypernetwork effectively improves imaging performance by alleviating serious domain gap problems. Furthermore, we also conduct experiments to explore the effectiveness of our learning strategy. We only aggregate the hypernetwork rather than the imaging network in each round, denoted as “HyperFed ◦.” Due to the substantial heterogeneity of different scanning parameters, learning a stable imaging feature becomes challenging. Meanwhile, we conduct additional experiments to investigate the modulation scope of the hypernetwork. “HyperFed *” and “HyperFed ◦” represent the hypernetwork modulating only the encoder or decoder, respectively. The results indicate that modulating all layers, only encoder, or only decoder yields similar performances. To ensure the consistency and generality of the method, we recommend the hypernetwork to modulate all layers of imaging networks, as there are no encoders or decoders in the IR-based methods.

We have demonstrated the effectiveness of our method in improving the performance of both CNN-based and IR-based imaging networks. Furthermore, we also validate our method on a transformer-based imaging network called U-former [65]. Since U-former is a postprocessing method, we utilize the same data as described in Section IV-B in this experiment. Specifically, we set the number of local training epochs to 1 and the number of communication rounds to 500. The results...
are presented in Table VI, the symbols “†” and “‡” represent the imaging networks without and with the hypernetwork, respectively. The results indicate that HyperFed can achieve promising performance even when the imaging network is transformer-based. The above experiments indicate that the designed physics-driven hypernetwork can effectively enhance the imaging network performance, and the proposed learning paradigm HyperFed can mitigate the heterogeneity issue.

F. Discussion

In the experiments, our method outperforms pFedHN and requires much fewer external parameters than pFedHN. The numbers of parameters for both pFedHN and our HyperFed with RED-CNN as the baseline are listed in Table VII. pFedHN utilizes linear layers to generate all parameters from a 100-D feature vector. This linear relationship between the external parameters and the model’s parameters limits the applicability of pFedHN to IR-based or transformer-based methods, which typically involve a large number of learning parameters. Moreover, pFedHN only generates parameters from the client index, which contains minimal information about the imaging process. In contrast, HyperFed incorporates only a small number of external parameters, making it more compatible with transformer-based or IR-based methods. In addition, the physical scanning parameters give a physics-driven modulation process to alleviate the domain gap. By considering these factors, HyperFed surpasses pFedHN in terms of flexibility, effectiveness, and applicability to various imaging methods.

Another issue we must mention is that our method does not introduce extra communication costs compared to other FL methods. Specifically, our method consists of a hypernetwork and an imaging network, but only the imaging network is transferred to the server for aggregation and hypernetwork is only used for personalization in each client. As a result, although the number of network parameters may slightly increase (the hypernetwork is quite small), our communication cost remains unchanged. Moreover, considering that the hypernetwork is locally hosted, the imaging network is globally shared, and all clients only transfer the parameters of the same model (imaging network) as other FL methods do, the communication latency of all these FL methods is equal.

According to the results presented in Sections IV-B–IV-D, the baseline model w/o FL occasionally achieves better performance than other FL-based methods. This observation is
consistent with our expectation that the non-iid problem will have a negative impact on FL and enforcing different institutions to train a global-shared model cooperatively may not be effective. To alleviate this problem, the proposed HyperFed aims to address the serious data heterogeneity problem in CT imaging by leveraging prior knowledge about the physical parameters of the imaging process. By incorporating this knowledge, we strive to achieve a better balance between the benefits of FL and the latent information of CT imaging data from different clients.

In this article, we consider the scenario where different institutions utilize various CT scanners with different low-dose technologies. In practical settings, doctors typically adjust the "distance between the source and rotation" to accommodate the "distance between the detector and rotation" to accommodate patients of different sizes. Other parameters are generally fixed in the scanners. To validate the robustness and generality of our proposed HyperFed, we have conducted additional experiments. Specifically, we simulate scenarios where we fix the other parameters and sample random values within the range of $[-100, 100]$. These values are then added to the distance parameters mentioned above. The results are shown in Table VIII, which demonstrates that our method consistently achieves the best performance. Besides, from both II and VIII, we can find that each method has similar performances across the two experiments, which suggests that the methods are not sensitive to these two parameters. Considering the great performance gaps between different institutions in the above experiments, we can find that the performance is sensitive to "the number of views" and "intensity of X-rays," which directly affect the radiation doses.

V. CONCLUSION

Current CT imaging networks are CL-based without considering any privacy issue, and there are inevitable non-iid problem caused by different scanning protocols and scanners in CT imaging. Current FL methods cannot achieve satisfactory performance for they do not consider the physical process. To relieve the above problems, we propose a hypernetwork-based physics-driven personalized FL framework for CT imaging, dubbed HyperFed. The proposed HyperFed is meticulously designed and effective, which can be considered a plug-and-play method. Our method can be flexibly embedded into different imaging networks, including CNN-based, IR-based, and transformer-based networks. The basic architecture of HyperFed is composed of an institution-specific physics-driven hypernetwork and a global-sharing imaging network. The global sharing imaging network ensures to extract the stable and invariant imaging features, and the hypernetwork modulates the feature domains to adapt local data. Experimental results show that the method w/o FL can personalize CT reconstruction, but some details are lost when the number of training samples is small. FL-based methods can keep stable performance and recover more details for all institutions, but they cannot reconstruct high-quality CT images for each domain. The proposed HyperFed effectively balances the tradeoff between both factors and it can reconstruct high-quality CT images for each institution in the comparisons with other methods in both qualitative and quantitative aspects.

REFERENCES

[1] T. L. Slovis, “The ALARA concept in pediatric CT: Myth or reality?” Radiology, vol. 223, no. 1, pp. 5–6, Apr. 2002.
[2] P. Bao, J. Zhou, and Y. Zhang, “Few-view CT reconstruction with group-sparsity regularization,” Int. J. Numer. Methods Biomed. Eng., vol. 34, no. 9, Sep. 2018, Art. no. e13101.
[3] Y. Chen et al., “Low-dose CT via convolutional neural network,” Biomed. Opt. Exp., vol. 8, no. 2, pp. 679–694, 2017.
[4] W. Wu et al., “Deep learning based spectral CT imaging,” Neural Netw., vol. 144, pp. 342–358, Dec. 2021.
[5] D. Wu, K. Kim, and Q. Li, “Computationally efficient deep neural network for computed tomography image reconstruction,” Med. Phys., vol. 46, no. 11, pp. 4763–4776, Nov. 2019.
[6] D. Wu, K. Kim, and Q. Li, “Low-dose CT reconstruction with Noise2Noise network and testing-time fine-tuning,” Med. Phys., vol. 48, no. 12, pp. 7657–7672, Dec. 2021.
[7] M. Ikuta and J. Zhang, “A deep convolutional gated recurrent unit for CT image reconstruction,” IEEE Trans. Neural Netw. Learn. Syst., vol. 34, no. 12, pp. 10612–10625, Dec. 2023.
[8] A. Narayanan and V. Shmatikov, “Robust de-anonymization of large sparse datasets,” in Proc. IEEE Symp. Secur. Privacy (SP), May 2008, pp. 111–125.
[9] C. G. Schwarz et al., “Identification of anonymous MRI research participants with face-recognition software,” New England J. Med., vol. 381, no. 17, pp. 1684–1686, Oct. 2019.
[10] G. Gürsoy, T. Li, S. Liu, E. Ni, C. M. Brannon, and M. B. Gerstein, “Functional genomics data: Privacy risk assessment and technological mitigation,” Nature Rev. Genet., vol. 23, no. 4, pp. 245–258, Apr. 2022.
[11] G. Kaissis, M. R. Makham, D. R. Ruckert, and R. F. Brame, “Secure, privacy-preserving and federated machine learning in medical imaging,” Nature Mach. Intell., vol. 2, no. 6, pp. 305–311, Jun. 2020.
[12] W. Xia et al., “Spectral CT reconstruction—ASSIST: Aided by self-similarity in image-spectral tensors,” IEEE Trans. Comput. Imag., vol. 5, no. 3, pp. 420–436, Sep. 2019.
[13] Z. Lu et al., “M2NAS: Multi-scale and multi-level memory-efficient neural architecture search for low-dose CT denoising,” IEEE Trans. Med. Imag., vol. 42, no. 3, pp. 850–863, Mar. 2023.
[14] D. L. Donoho, “Compressed sensing,” IEEE Trans. Inf. Theory, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
[15] Y. Zhang, W. Zhang, Y. Lei, and J. Zhou, “Few-view image reconstruction with fractional-order total variation,” J. Opt. Soc. Amer. A, Opt. Image Sci., vol. 31, no. 5, pp. 981–995, 2014.
[16] Y. Chen et al., “Bayesian statistical reconstruction for low-dose X-ray computed tomography using an adaptive-weighting nonlocal prior,” Computerized Med. Imag. Graph., vol. 33, no. 7, pp. 495–500, Oct. 2009.
[17] Y. Chen et al., “Improving abdomen tumor low-dose CT images using a fast dictionary learning based processing,” Phys. Med. Biol., vol. 58, no. 16, pp. 5803–5820, Aug. 2013.
[18] M. Aharon, M. Elad, and A. Bruckstein, “rM-SVD: An algorithm for designing overcomplete dictionaries for sparse representation,” IEEE Trans. Signal Process., vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
[19] H. Chen et al., “Low-dose CT with a residual encoder-decoder convolutional neural network,” IEEE Trans. Med. Imag., vol. 36, no. 12, pp. 2524–2535, Dec. 2017.
[20] C. You et al., “CT super-resolution GAN constrained by the identical, residual, and cycle learning ensemble (GAN-CIRCLE),” IEEE Trans. Med. Imag., vol. 39, no. 1, pp. 188–203, Jan. 2020.

| TABLE VIII |
| --- |
| RESULTS OF ROBUSTNESS EXPERIMENTS |
| PSNR | SSIM |
| w/o FL | 38.88 | 0.9545 |
| FedAvg | 36.47 | 0.9271 |
| FedFoxy | 36.94 | 0.9336 |
| FedBN | 38.99 | 0.9437 |
| Ditto | 38.67 | 0.9484 |
| pFedBN | 37.78 | 0.9424 |
| HyperFed | 39.61 | 0.9551 |

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
B. Wei, J. Li, Y. Liu, and W. Wang, “Non-IID federated learning with J. Xu, W. Du, Y. Jin, W. He, and R. Cheng, “Ternary compression for ANG et al.: HYPERNETWORK-BASED PHYSICS-DRIVEN PERSONALIZED FEDERATED LEARNING FOR CT IMAGING 3149

W. Xia et al., “CT reconstruction with PDF: Parameter-dependent W. Huang, M. Ye, and B. Du, “Learn from others and be yourself in

T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, “Federated optimization in heterogeneous networks,” in Proc. Mach. #include <bits/stdc++.h>

T. Li, S. Hu, A. Beirami, and V. Smith, “Ditto: Fair and robust federated learning through personalization,” in Proc. Int. Conf. Mach. 2021, pp. 6357–6368.

X. Ma, J. Zhang, S. Guo, and W. Xu, “Layer-wised model aggregation for personalized federated learning,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2022, pp. 10082–10091.

X. Li, M. Jiang, X. Zhang, M. Kamp, and Q. Dou, “FedBN: Federated learning on non-HD features via local batch normalization,” 2021, arXiv:2102.07623.

A. Shamsian, A. Navon, E. Fetaya, and G. Chechik, “Personalized federated learning using hypernetworks,” in Proc. Int. Conf. Mach. Learn., 2021, pp. 9489–9502.

H. Li et al., “FedTP: Federated learning by transformer personalization,” IEEE Trans. Neural Netw. Learn. Syst., pp. 1–15, May 2023.

Y. Zhao et al., “Personalized federated few-shot learning,” IEEE Trans. Neural Netw. Learn. Syst., pp. 1–11, Jul. 2022.

Y. Chen, W. Lu, X. Qin, J. Wang, and X. Xie, “MetaFed: Federated learning among federations with cyclic knowledge distillation for personalized healthcare,” IEEE Trans. Neural Netw. Learn. Syst., pp. 1–122, Jul. 2023.

F. Hanzely and P. Richtárik, “Federated learning of a mixture of global and local models,” 2020, arXiv:2002.05516.

R. S. Antunes, C. André da Costa, A. Küderle, I. A. Yari, and B. Eskofier, “Federated learning for healthcare: Systematic review and architecture proposal,” ACM Trans. Intell. Syst. Technol., vol. 13, no. 4, pp. 1–23, Aug. 2022.

R. Kumar et al., “Blockchain-federated-learning and deep learning models for COVID-19 detection using CT imaging,” IEEE Sensors J., vol. 21, no. 14, pp. 16301–16314, Jul. 2021.

L. M. Florescu et al., “Federated learning approach with pre-trained deep learning models for COVID-19 detection from unsegmented CT images,” Life, vol. 12, no. 7, p. 958, Jun. 2022.

L. C. Fourati and S. Ayed, “Federated learning toward data preprocessing: COVID-19 context,” in pp. 1–6, 2021.

C.-M. Feng, Y. Yan, S. Wang, Y. Xu, L. Shao, and H. Fu, “Specificity-preserving federated learning for MR image reconstruction,” 2021, arXiv:2112.05752.

W. Wu, D. Hu, C. Niu, H. Yu, V. Vardhanabhuti, and G. Wang, “DRONE: Dual-domain residual-based optimization network for sparse-view CT reconstruction,” IEEE Trans. Med. Imag., vol. 40, no. 11, pp. 3002–3014, Nov. 2021.

T. Bai, B. Wang, D. Nguyen, and S. Jiang, “Probabilistic self-learning framework for low-dose CT denoising,” Med. Phys., vol. 48, no. 5, pp. 2258–2270, May 2021.

Y. Mansour, M. Mohri, J. Ro, and A. Theertha Suresh, “Three approaches for personalization with applications to federated learning,” 2020, arXiv:2002.10619.

B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, “Enhanced deep residual networks for single image super-resolution,” in IEEE Proc. Conf. Comput. Vis. Pattern Recognit. Workshops (ICCV Workshops), Jul. 2017, pp. 1132–1140.

Y. Tan et al., “Fedproto: Federated prototype learning across heterogeneous clients,” in Proc. AAAI Conf. Artif. Intell., vol. 36, 2022, pp. 8432–8440.

S. S. Azam, S. Hosseinalipour, Q. Qiu, and C. Brinton, “Recycling model updates in federated learning: Are gradient subspaces low-rank?” in Proc. Int. Conf. Learn. Represent., pp. 1–12.

J. Wang, Q. Liu, H. Liang, G. Joshi, and H. V. Poor, “Tackling the objective inconsistency problem in heterogeneous federated optimization,” in Proc. Adv. Neural Inf. Process. Syst. (NIPS), vol. 33, Dec. 2020, pp. 7611–7623.

H. Yu, S. Yang, and S. Zhu, “Parallel restarted SGD with faster convergence and less communication: Demystifying why model averaging works for deep learning,” in Proc. AAAI Conf. Artif. Intell., Jul. 2019, vol. 33, no. 1, pp. 5693–5700.

C. McCollough, “TU-PG-207A-04: Overview of the low dose CT grand challenge,” Med. Phys., vol. 43, pp. 3759–3760, Jun. 2016.

D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” 2014, arXiv:1412.6980.

Z. Wang, X. Cun, J. Bao, W. Zhou, J. Liu, and H. Li, “Uformer: A general U-shaped transformer for image restoration,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., Jun. 2022, pp. 17683–17693.
Wenjun Xia (Member, IEEE) received the B.A. and Ph.D. degrees from Sichuan University, Chengdu, China, in 2012 and 2021, respectively.

Now, he is an Assistant Researcher at the School of Cyber Science and Engineering, Sichuan University, and the Key Laboratory of Data Protection and Intelligent Management, Ministry of Education, Sichuan University. His research interests include medical imaging, compressive sensing, and deep learning.

Zexin Lu received the B.S. degree in computer science from the Beijing University of Chemical Technology, Beijing, China, in 2018. He is currently pursuing the Ph.D. degree in software engineering with Sichuan University, Chengdu, China.

His research interests include X-ray computed tomography, machine learning, and deep learning.

Yingyu Chen received the B.S. degree in computer science from the School of Software, Sichuan University, Chengdu, China, in 2021, where he is currently pursuing the Ph.D. degree with the College of Computer Science.

His research interests include uncertainty quantification, machine learning, and medical image analysis.

Xiaoxiao Li (Member, IEEE) received the Ph.D. degree from Yale University, New Haven, CT, USA, in 2020.

She has been an Assistant Professor at the Department of Electrical and Computer Engineering (ECE), The University of British Columbia (UBC), Vancouver, BC, Canada, since August 2021. Before joining UBC, she was a Post-Doctoral Research Fellow at the Computer Science Department, Princeton University, Princeton, NJ, USA. In the last few years, she has had over 30 papers published in leading machine learning conferences and journals, including Neural Information Processing Systems (NeurIPS), International Conference on Machine Learning (ICML), International Conference on Learning Representations (ICLR), Medical Image Computing and Computer Assisted Intervention Society (MICCAI), Information Processing in Medical Imaging (IPML), British Machine Vision Association (BMVC), IEEE TRANSACTIONS ON MEDICAL IMAGING, and Medical Image Analysis. Her research interests include range across the interdisciplinary fields of deep learning and biomedical data analysis, aiming to improve the trustworthiness of AI systems for healthcare.

Dr. Li’s work has been recognized with several best paper awards at international conferences.

Yi Zhang (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees in computer science and technology from the College of Computer Science, Sichuan University, Chengdu, China, in 2005, 2008, and 2012, respectively.

From 2014 to 2015, he was with the Department of Biomedical Engineering, Rensselaer Polytechnic Institute, Troy, NY, USA, as a Post-Doctoral Researcher. He is currently a Full Professor with the School of Cyber Science and Engineering, Sichuan University, and the Director of the Deep Imaging Group (DIG). He has authored more than 80 articles in the field of image processing. These articles were published in several leading journals, including IEEE TRANSACTIONS ON MEDICAL IMAGING, IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING, Medical Image Analysis, European Radiology, and Optics Express, and reported by the Institute of Physics (IOP) and during the Lindau Nobel Laureate Meeting. His research interests include medical imaging, compressive sensing, and deep learning.

Dr. Zhang received major funding from the National Key Research and Development Program of China, the National Natural Science Foundation of China, and the Science and Technology Support Project of Sichuan Province, China. He is a Guest Editor of the International Journal of Biomedical Imaging, Sensing and Imaging, and an Associate Editor of IEEE TRANSACTIONS ON MEDICAL IMAGING and IEEE TRANSACTIONS ON RADIATION AND PLASMA MEDICAL SCIENCES.