UNIVERSITY of MARYLAND
SCHOOL OF MEDICINE

MIDL
Paris 2024

Improving Multi-Center Generalizability of GAN-Based Fat Suppression using Federated Learning

Pranav Kulkarni, Adway Kanhere, Harshita Kukreja, Vivian Zhang, Paul H. Yi, Vishwa S. Parekh

University of Maryland School of Medicine, Baltimore, MD, USA

Federated Learning can improve multi-center generalizability of GANs for medical image synthesis while facilitating privacy-preserving collaborations.

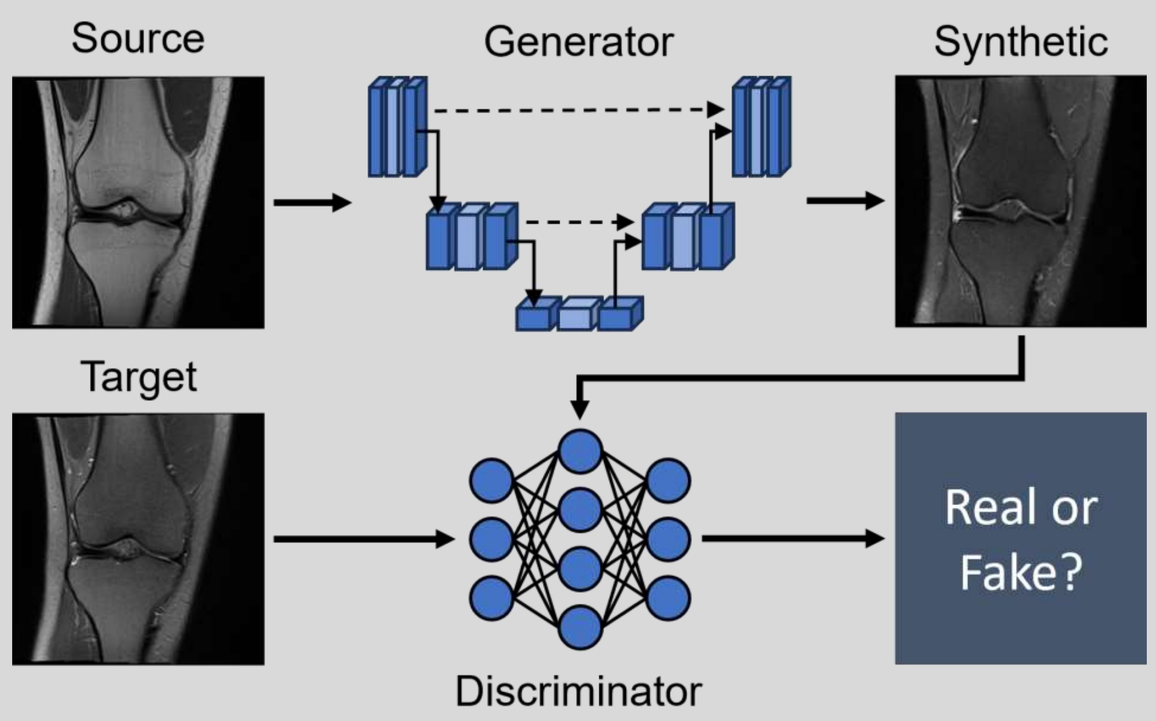

Figure 1. Centralized GAN-based synthesis of FS MRI sequences from non-FS PD sequences. This approach requires all participating centers to aggregate patient data at a single site.

Introduction

- GAN-based MRI synthesis has the potential to accelerate image acquisition and reduce patient discomfort[1].
- One potential use-case is for knee MRIs, where proton density-weighted (PD) and fluid-sensitive fat suppressed (FS) sequences are used to detect abnormalities[2].
- GANs trained on single-site data have poor generalizability to external data and it is impractical to curate multi-center dataset at a single site due to patient privacy[1,3].
- We showed that federated learning (FL) can improve the multi-center generalizability of MRI synthesis while facilitating privacy-preserving multi-institutional collaborations.

Methods

- We used two datasets:
  1. An internal University of Maryland (UMB) dataset with $n = 151$ non-FS PD and FS sequence pairs in axial and coronal planes.
  2. The FastMRI dataset with $n = 7{,}171$ non-FS PD and FS sequence pairs in sagittal and coronal planes[4].
- We randomly sampled training ($n = 80$) and testing ($n = 20$) sequence pairs, registered them (non-FS PD fixed; FS, moving), and extracted slices in the imaging plane.
- We used pix2pix[5], a conditional GAN comprised of a U-Net generator and PatchGAN discriminator, to synthesize FS sequences (target) from non-FS PD sequences (source).
- We trained four models:
  1. A single-site model with UMB dataset ("Baseline-UMB") (**Fig. 1**).
  2. A single-site model with the FastMRI dataset ("Baseline-FastMRI").
  3. A centrally aggregated model with both datasets ("Central").
  4. A two-client FL model with both datasets distributed at each client (**Fig. 2**). At the end of each epoch, local weights are aggregated using FedGAN[6] at the central server.
- We compared the mean SSIM $\pm$ SD between the ground-truth and synthetic FS sequences across all four models for both test sets using Wilcoxon signed-rank tests. Statistical significance was defined as $p < 0.05$.

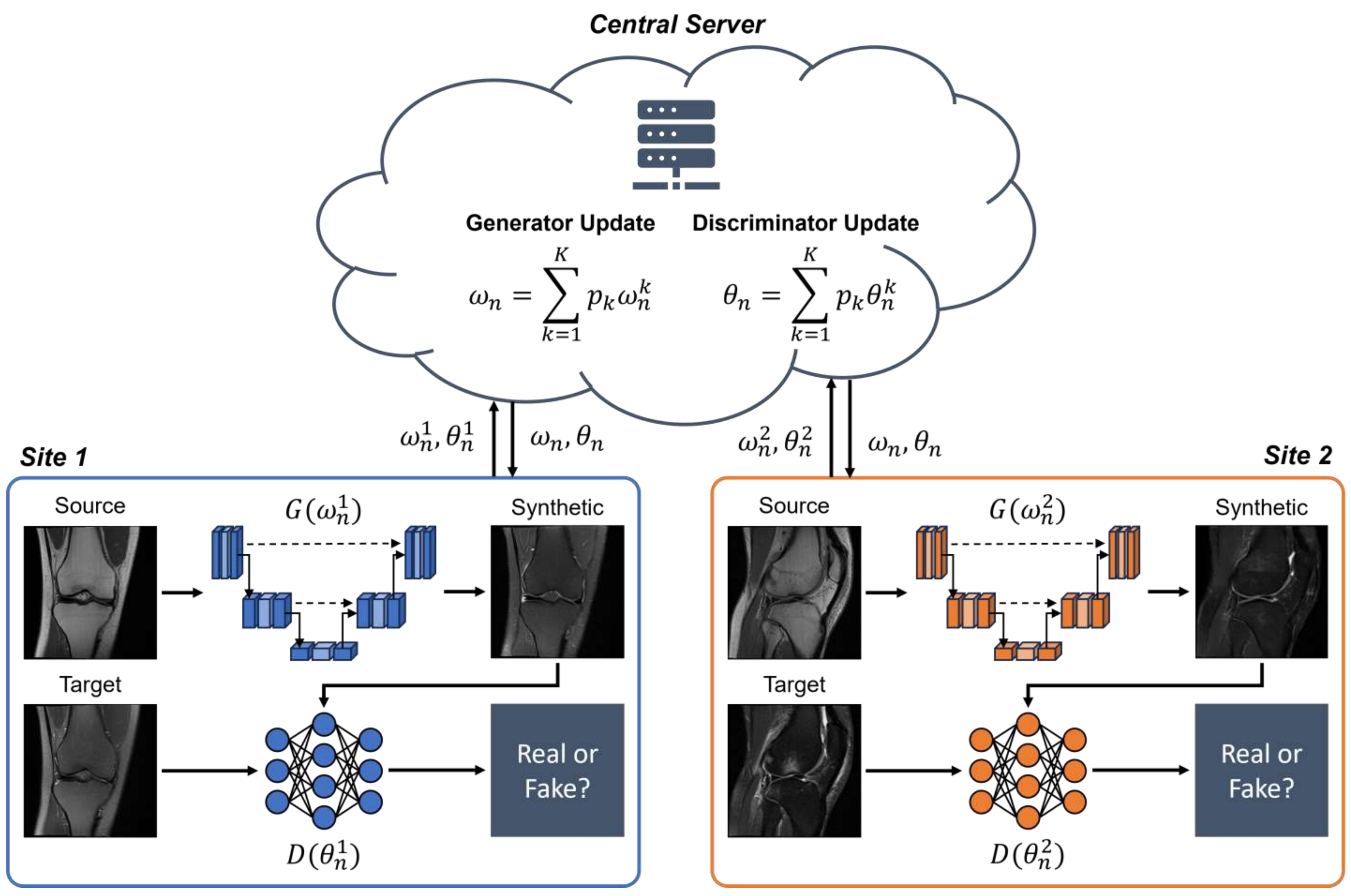

Figure 2. Privacy-preserving multi-center GAN-based synthesis of FS MRI sequences from non-FS PD sequences using FL. Each client updates their local GAN using private data and then communicates the weights of the generator and discriminator models to the central server. Using FedGAN[6], the central server aggregates the weights and communicates them back to each client.

Results

Model	Test Set			
	UMB		FastMRI	
Baseline-UMB	0.64 ± 0.13	$p = 0.63$	0.46 ± 0.11	$p < 0.001$
Baseline-FastMRI	0.46 ± 0.11	$p < 0.001$	0.58 ± 0.12	$p = 0.99$
Central	0.64 ± 0.13	$p = 0.74$	0.58 ± 0.12	$p = 0.93$
FL	0.63 ± 0.13	–	0.58 ± 0.12	–

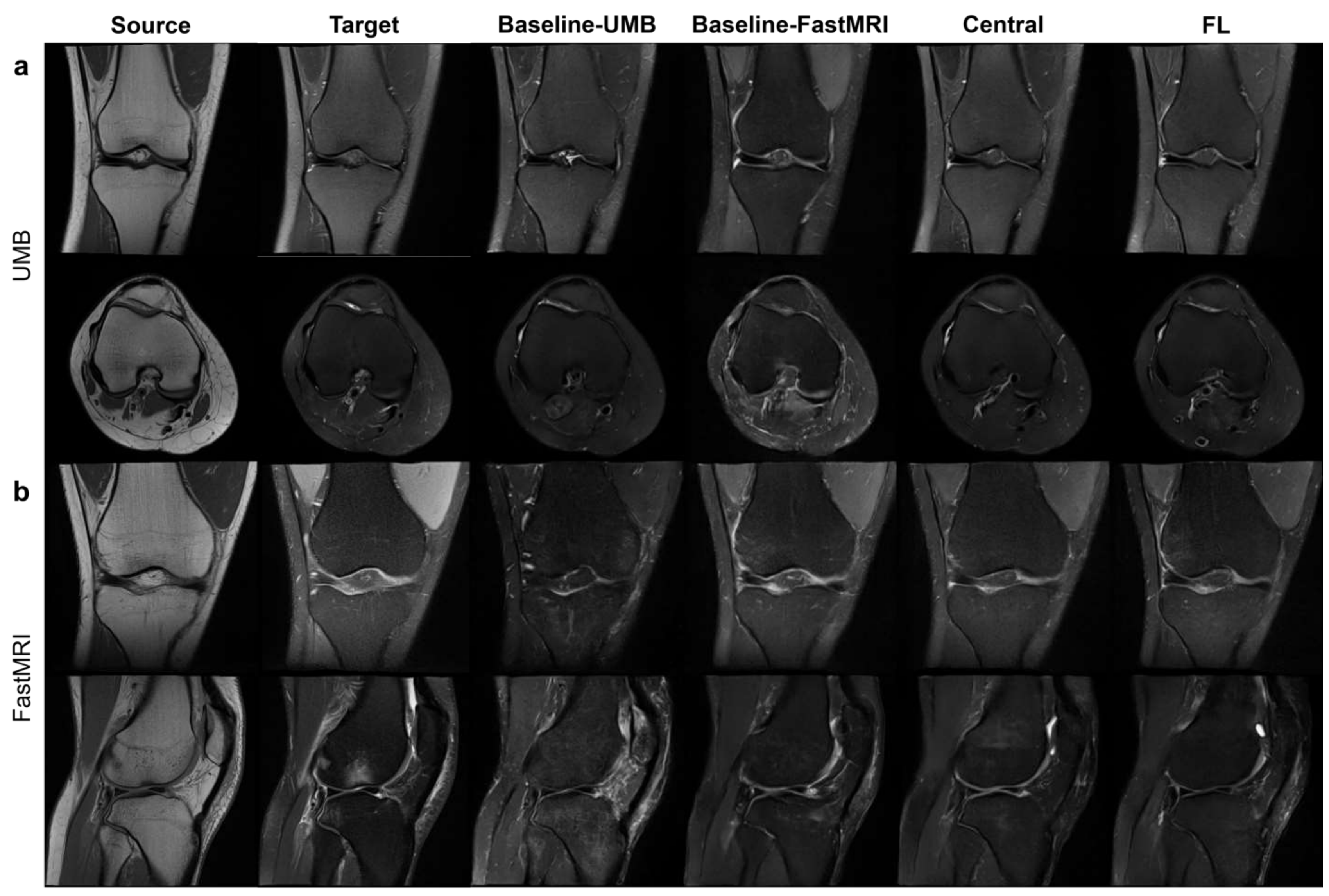

Figure 3. Examples of MRI sequences from the **(a)** UMB and **(b)** FastMRI test sets. The non-FS PD sequences (column 1) and ground-truth FS sequences (column 2) are shown with their corresponding synthetic FS sequences from the Baseline-UMB (column 3), Baseline-FastMRI (column 4), Central (column 5), and FL (column 6) models.

Discussion

- Single-site models had poor generalizability to external data despite high performance on local data.
- The FL model exhibited significantly higher performance on external data compared to single-site models.
- FL improved multi-center generalizability of FS MRI synthesis in a privacy-preserving way.
- Since our work is preliminary, our GANs were trained on a small subset and result in sub-optimal performance which can be alleviated by training on larger datasets.
- Our preliminary results represent an exciting step towards synthetic MRIs becoming a clinical reality.

References

1. Dar, S., ... & Cukur, T. (2019). Image synthesis in multi-contrast MRI with conditional GAN. *IEEE Transactions on Medical Imaging*, 38(10):2375-2388.
2. Lee, S., ... & Kim, J. (2011). Proton density-weighted MR imaging of the knee: fat suppression versus without fat suppression. *Skeletal Radiology*, 40:189-195.
3. Dalmaz, O., ... & Çukur, T. (2024). One model to unite them all: Personalized federated learning of multi-contrast MRI synthesis. *Medical Image Analysis*, 103121.
4. Knoll, F., ... & Lui, Y. W. (2020). fastMRI: A publicly available raw k-space and DICOM dataset of knee images for accelerated MR image reconstruction using machine learning. *Radiology: Artificial Intelligence*, 2(1):e190007.
5. Isola, P., ... & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1125-1134.
6. Rasouli, M., ... & Rajagopal, R. (2020). FedGAN: Federated generative adversarial networks for distributed data. *arXiv preprint arXiv:2006.07228*.

Pranav Kulkarni
University of Maryland School of Medicine
pkulkarni@som.umaryland.edu
@itspranavk
itspranavk

Read the full paper!