ChestFed: Model-Contrastive Federated Learning for Cardiopulmonary Disease Classification in Chest X-rays using Multiple Datasets

Tianhao Li
tianhao@utexas.edu

Cyprian Zander
zandercyprian27@gmail.com

Ajay Jaiswal
ajayjaiswalhi@gmail.com

Justin Rousseau
justin.rousseau@austin.utexas.edu

Ying Ding
ying.ding@ischool.utexas.edu

Abstract

Artificial intelligence in medical imaging has emerged to be a topic with high demand in medical practice in the recent years. However, limited data availability due to strict patient privacy policy becomes a main barrier in this area. Federated learning enables multiple parties to collaboratively train a machine learning/deep learning model without sharing their local data. Model-contrastive federated learning, as a novel federated learning framework, is designed to handle the heterogeneity of local data distribution by using contrastive learning across parties. In this work, we applied the model-contrastive federated learning in multiple chest x-ray datasets to derive a global model for disease diagnosis. Our experiment shows that using federated learning only on two datasets, our model outperforms the model trained in one single dataset by 4%, which indicates the potential to apply federated learning on several chest x-ray datasets to achieve higher accuracy without the need to share local data.

1 Introduction

Artificial intelligence (AI) is revolutionizing healthcare and, in particular, medical imaging. Health innovations applying machine learning (ML) and deep learning (DL) in radiology account for more than half of the total AI innovations in health. Advancing AI in medical imaging brings extraordinary benefits with better accuracy, lower cost, and higher efficiency. However, the broad application of AI techniques in healthcare is currently hindered by limited dataset availability due to the strict legal and ethical requirements to protect patient privacy [6]. For example, in chest x-ray medical imaging area, there are only several publicly available datasets: NIH ChestX-ray8 dataset [14] containing 108,948 frontal view chest x-ray images of 32,717 unique patients with 8 common diseases; VinDr-CXR dataset [13] containing 18,000 frontal view chest x-ray images with a total of 28 findings and diagnoses; and MIMIC-CXR dataset [5], the largest chest x-ray dataset now, containing 377,110 images in both frontal view and lateral view from 65,379 patients, etc. The total number of images in these datasets is even less than one tenth of that in a single ImageNet [3] with 14,197,122 images. In this situation, how to fully use the data in different datasets with strict patient privacy policy becomes an important problem to be solved.

Federated learning enables multiple parties to jointly learn a machine learning/deep learning model without exchanging their local data. A classic algorithm of federated learning is FedAvg [11]. In FedAvg, there are several local models from different parties and one global model as the target. In each round, the updated local models of the parties are transferred to the global server, which aggregates them to update the global model. The raw data is not exchanged during the learning

35th Conference on Neural Information Processing Systems (NeurIPS 2021), Sydney, Australia.

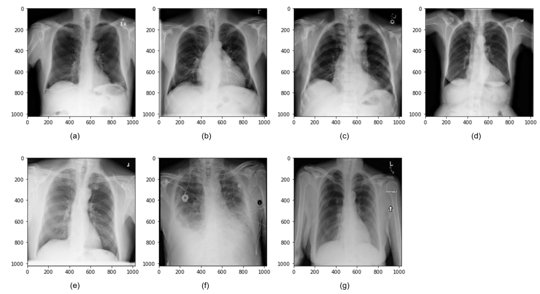

Figure 1: Example images of diseases in both datasets: (a) Atelectasis, (b) Cardiomegaly, (c) Effusion, (d) Infiltration, (e) Nodule/Mass, (f) Pneumonia, (g) Pneumothorax.

process. In medical imaging, this means each hospital can train its data locally without make it publicly, which satisfies the strict patient privacy policy to a great extent. Model-Contrastive Federated Learning (MOON) [8] is an advanced algorithm of federated learning. It applies contrastive learning in a model-wise way. MOON decreases the distance between the representations learned by the local models and the representations of the aggregated global model, and increase the distance between the representations learned by the local models and the representations of previous round learned by the local models. By doing that, it can effectively deal with the heterogeneity of local data distribution across parties, which is extremely suitable for the medical imaging situation with heterogeneous datasets due to different equipment and measurement of different hospitals.

In this work, we develop ChestFed by applying MOON for disease classification on two chest x-ray datasets: NIH ChesX-ray8 and VinDr-CXR. There are 7 disease labels in common between the two datasets. Figure 1 shows the example chest x-ray images of these diseases. We evaluate the ChestFed on the NIH ChesX-ray8 dataset. Without federated learning, the mean AUC is 75% using the model trained only by NIH ChesX-ray8 dataset. However, when using ChestFed on both datasets, the mean AUC is improved to 79% tested by NIH ChesX-ray8 dataset. This significant improvement indicates that ChestFed helps develop the global model by aggregating models trained from different chest x-ray dataset, which implies the possibility that federated learning can learn a better global model from different medical imaging datasets by fully exploring the data without breaking the patient privacy policy.

2 Related Work

Federate learning become a center focus in medical domains in recent years. FedAvg [11] is a classic algorithm of federated learning, which aggregates the local models by simply averaging the model weights. FedProx [9] introduces a proximal term into the loss during local training, which is computed based on the l2-norm distance between the global model and the local models. SCAFFOLD [7] corrects the local updates by introducing control variates, which is used to correct the gradients in local training. MOON [8] conducts contrastive learning in model-level by comparing the representations learned by different models. All these algorithms are based on the assumption ahat the global model is able to learn a better representation than the local models trained on local datasets.

There is not much work in chest x-ray using federated learning. Most of them are related to COVID-19 detection and localization [10, 15]. [1] uses multiple pretrain models and optimizers on federate learning to classify vital pneumonia, bacterial pneumonia and normal patient, in which ResNet-18 and Momentum + SGD give the best accuracy. [2] divides the CheXpert [4] dataset into 5 sites and use federated learning to classify 14 diseases, in which the AUC using federated learning is slightly higher than the one not using federated learning. Up to now, there is no work in medical imaging using contrastive learning combined with federated learning.

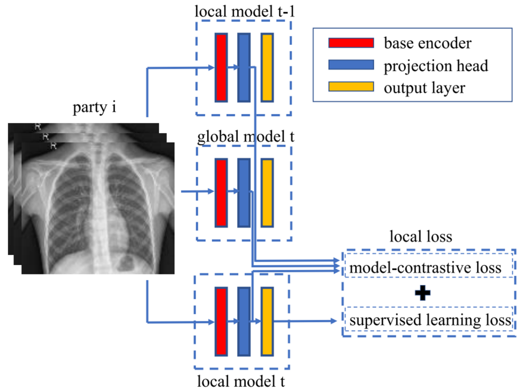

Figure 2: Our ChestFed model. Base encoder: ResNet-50; Projection head: 2-layer MLP.

3 ChestFed

ChestFed structure is based on MOON. Figure 2 is an illustration of our ChestFed structure. For the specific network used in the local model or global model, we use pretrained ResNet-50 as our base encoder. Our projection head is a 2-layer multi-layer perceptron (MLP). The output dimension of the projection head is set to 256. We use cross-entropy loss as our supervised loss. Our local loss contains two parts, cross-entropy loss and model-contrastive loss. The model-contrastive loss is

$$l_{model-contrastive} = -\log \frac{\exp(sim(z, z_{glob})/\tau)}{\exp(sim(z, z_{glob})/\tau) + \exp(sim(z, z_{prev})/\tau)} \tag{1}$$

where τ denotes a temperature parameter, z_{glob} denotes the representation of the image from the current global model, z denotes the representation of the image from the current local model, and z_{prev} denotes the representation of the image from the previous global model. The total loss is computed by

$$l_{total} = l_{cross-entropy} + \mu l_{model-contrastive} \tag{2}$$

where μ is a hyper-parameter to control the weight of model-contrastive loss.

4 Experiment

NIH ChestX-ray8 dataset [14] contains 108,948 chest x-ray images. VinDr-CXR dataset [13] contains 18,000 chest x-ray images. In each of the two datasets, we only use the images with the 7 disease labels mentioned in Figure 1. For each dataset, We use 70% for training, 10% for validation, and 20% for testing. All the original images are resized to $256 * 256$ for processing and pixel intensities are normalized to $[-1, 1]$. We use pretrained weights of ResNet-50 provided by [12], and SGD optimizer with learning rate 0.01 and weight decay 10^{-5}. The batch size is set to 64. The number of local epochs is set to 20. Table 1 shows the AUC results of NIH ChesX-ray8 dataset trained in a single datasets and trained together with VinDr-CXR dataset using ChestFed. The significant 4% increase demonstrates the effectiveness of ChestFed on chest x-ray disease classification.

	Single	ChestFed
Atelectasis	0.75	0.78
Cardiomegaly	0.85	0.89
Effusion	0.83	0.82
Infiltration	0.67	0.76
Nodule/Mass	0.67	0.79
Pneumonia	0.70	0.73
Pneumothorax	0.79	0.79
Mean	0.75	0.79

Table 1: AUC comparison of the model trained in a single dataset and the model trained in two datasets by ChestFed. The evaluations are all made in the NIH ChestX-ray8 dataset. A significant increase in performance illustrates the benefits of ChestFed.

References

[1] S. Banerjee, R. Misra, M. Prasad, E. Elmroth, and M. H. Bhuyan. Multi-diseases classification from chest-x-ray: A federated deep learning approach. In *Australasian Joint Conference on Artificial Intelligence*, pages 3–15. Springer, 2020.

[2] A. Chakravarty, A. Kar, R. Sethuraman, and D. Sheet. Federated learning for site aware chest radiograph screening. In *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, pages 1077–1081. IEEE, 2021.

[3] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pages 248–255. Ieee, 2009.

[4] J. Irvin, P. Rajpurkar, M. Ko, Y. Yu, S. Ciurea-Ilcus, C. Chute, H. Marklund, B. Haghgoo, R. Ball, K. Shpanskaya, et al. Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pages 590–597, 2019.

[5] A. E. Johnson, T. J. Pollard, N. R. Greenbaum, M. P. Lungren, C.-y. Deng, Y. Peng, Z. Lu, R. G. Mark, S. J. Berkowitz, and S. Horng. Mimic-cxr-jpg, a large publicly available database of labeled chest radiographs. *arXiv preprint arXiv:1901.07042*, 2019.

[6] G. A. Kaissis, M. R. Makowski, D. Rückert, and R. F. Braren. Secure, privacy-preserving and federated machine learning in medical imaging. *Nature Machine Intelligence*, 2(6):305–311, 2020.

[7] S. P. Karimireddy, S. Kale, M. Mohri, S. J. Reddi, S. U. Stich, and A. T. Suresh. Scaffold: Stochastic controlled averaging for on-device federated learning. 2019.

[8] Q. Li, B. He, and D. Song. Model-contrastive federated learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10713–10722, 2021.

[9] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith. Federated optimization in heterogeneous networks. *arXiv preprint arXiv:1812.06127*, 2018.

[10] B. Liu, B. Yan, Y. Zhou, Y. Yang, and Y. Zhang. Experiments of federated learning for covid-19 chest x-ray images. *arXiv preprint arXiv:2007.05592*, 2020.

[11] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*, pages 1273–1282. PMLR, 2017.

[12] D. Misra, T. Nalamada, A. U. Arasanipalai, and Q. Hou. Rotate to attend: Convolutional triplet attention module. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 3139–3148, 2021.

[13] H. Q. Nguyen, K. Lam, L. T. Le, H. H. Pham, D. Q. Tran, D. B. Nguyen, D. D. Le, C. M. Pham, H. T. Tong, D. H. Dinh, et al. Vindr-cxr: An open dataset of chest x-rays with radiologist's annotations. *arXiv preprint arXiv:2012.15029*, 2020.

[14] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2097–2106, 2017.

[15] D. Yang, Z. Xu, W. Li, A. Myronenko, H. R. Roth, S. Harmon, S. Xu, B. Turkbey, E. Turkbey, X. Wang, et al. Federated semi-supervised learning for covid region segmentation in chest ct using multi-national data from china, italy, japan. *Medical image analysis*, 70:101992, 2021.