Federated Generalized Learning on Non-IID Medical Imaging with Virtual Homogeneous Generation and Adversarial Domain Adaptation

YAN WENHAO

Department of Science and Informatics
Muroran Institute of Technology

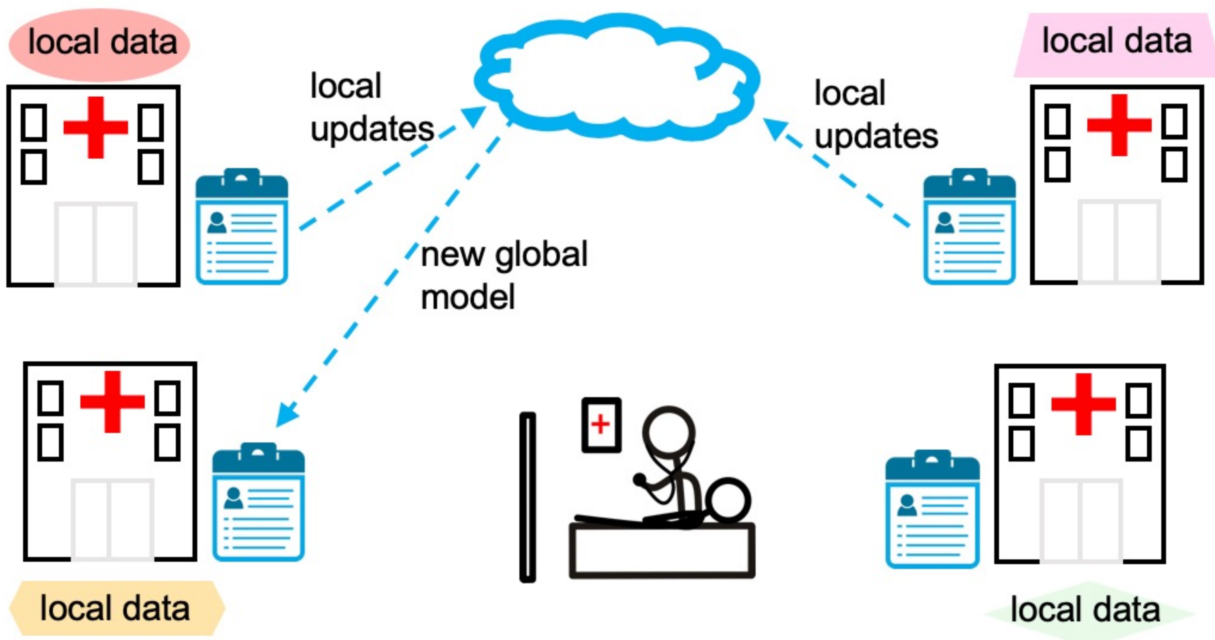
Background

- **Raising Privacy Concern in Intelligent Systems**
  - EU：《General Data Protection Regulation》(GDPR) (2018)
  - WHO: 《Ethics and Governance of Artificial Intelligence for Health: WHO Guidance》(2021)
  - 中 国: 《数据安全法》，《个人信息保护法》(2021)
  - 日 本:《経済安全保障推進法》(令和４年法律第43号)

https://www.nikkei.com/article/DGXZQOGN30F340Q1A730C2000000/
https://www.ft.com/content/c1c57ff8-85c5-4d42-921c-d371f13f6e94

Federated Learning

- **Unlike traditional centralized learning, FL always keeps data storaged in local client for privacy preserving.**



Centralized Steps

1.Send data directly to server

2.Central training

3.Send back model

FL Steps

1.Local training

2.Send encrypted gradients to server

3.Aggreagte the gradients

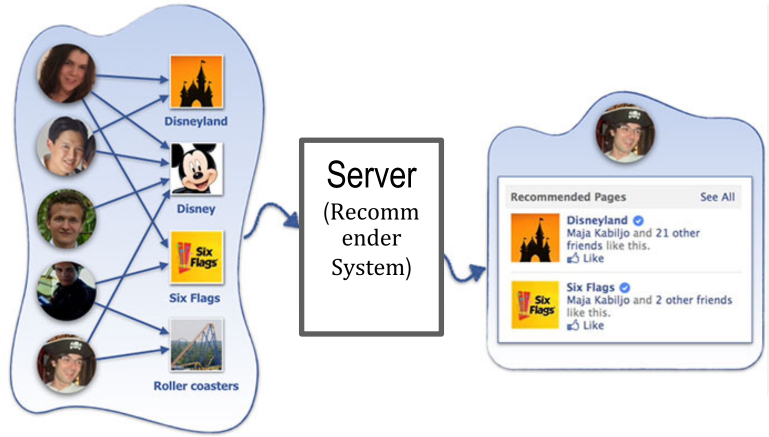
4.Send back model

McMahan, Brendan, et al. "Communication-efficient learning of deep networks from decentralized data." *Artificial intelligence and statistics*. PMLR, 2017.
https://blog.ml.cmu.edu/2019/11/12/federated-learning-challenges-methods-and-future-directions/

Non-iid Problems in FL

- **Label Skew**

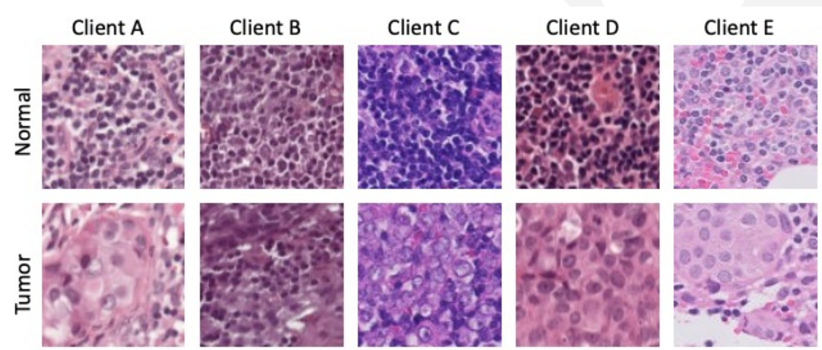
$$P_i(x,y) \to P_i(y|x)P_i(x)$$
$$P_i(x,y) \to P_i(x|y)P_i(y)$$



Example: Distributed Recommender System

- **Feature Skew**

$$P_i(x,y) \to P_i(y|x)P_i(x)$$
$$P_i(x,y) \to P_i(x|y)P_i(y)$$



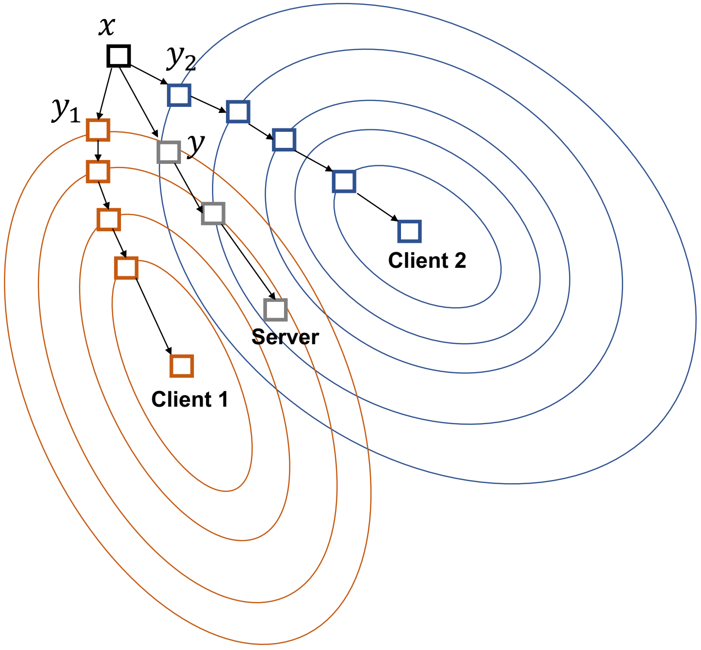
Example: Medical Imaging from Different Institutions

Different

Same

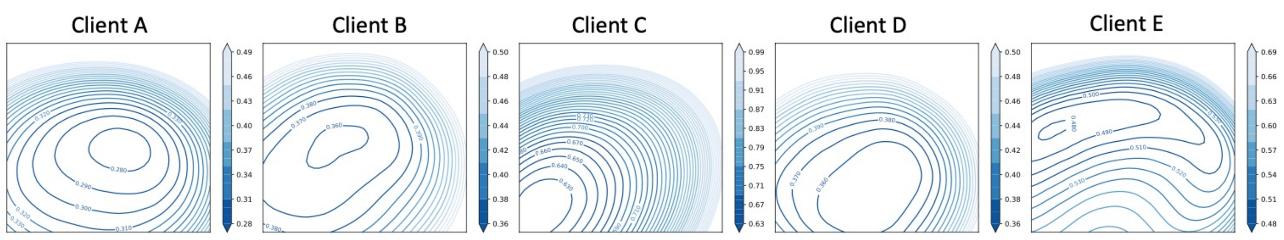
Litjens G, Bandi P, Ehteshami Bejnordi B, Geessink O, Balkenhol M, Bult P, Halilovic A, Hermsen M, van de Loo R, Vogels R, Manson QF, Stathonikos N, Baidoshvili A, van Diest P, Wauters C, van Dijk M, van der Laak J. 1399 H&E-stained sentinel lymph node sections of breast cancer patients: the CAMELYON dataset. Gigascience. 2018 Jun 1;7(6):giy065.

Feature Drfts caused by Non-IID



Main Reason

The distribution of data on different clients is quite inconsistent



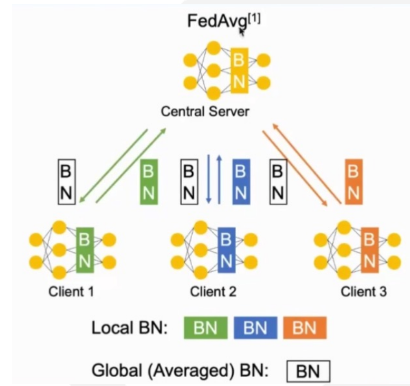
Loss landscape visualization of 5 clients in a multi-source medical image dataset (**Camelyon17**) with FedAvg algorithm, showing great heterogeneity

Jiang, Meirui, Zirui Wang, and Qi Dou. "Harmofl: Harmonizing local and global drifts in federated learning on heterogeneous medical images." *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 36. No. 1. 2022.

Existing Works (FL on Non-IIDs)

- **FedBN**

  Li, Xiaoxiao, et al. "Fedbn: Federated learning on non-iid features via **<u>local</u>** batch normalization." **ICLR 2021.**

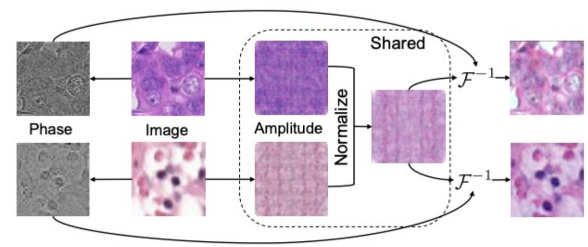


$$f^*(\mathbf{x};\mathbf{V},\gamma,\mathbf{c}) = \frac{1}{\sqrt{m}}\sum_{k=1}^{m} c_k \sum_{i=1}^{N} \sigma\left(\gamma_{k,i}\cdot\frac{\mathbf{v}_k^\top \mathbf{x}}{\|\mathbf{v}_k\|_{\mathbf{S}_i}}\right)\cdot \mathbb{1}\{\mathbf{x}\in \text{client } i\}$$

- **HarmoFL**

  Jiang, et al. “Harmofl: Harmonizing **<u>local</u>** and global drifts in federated learning on heterogeneous medical images." ***AAAI* 2022**.

$$\min_{\theta}\left[F(\theta) := \sum_{i=1}^{N} p_i F_i(\theta+\delta,\overline{\mathcal{D}_i})\right]$$

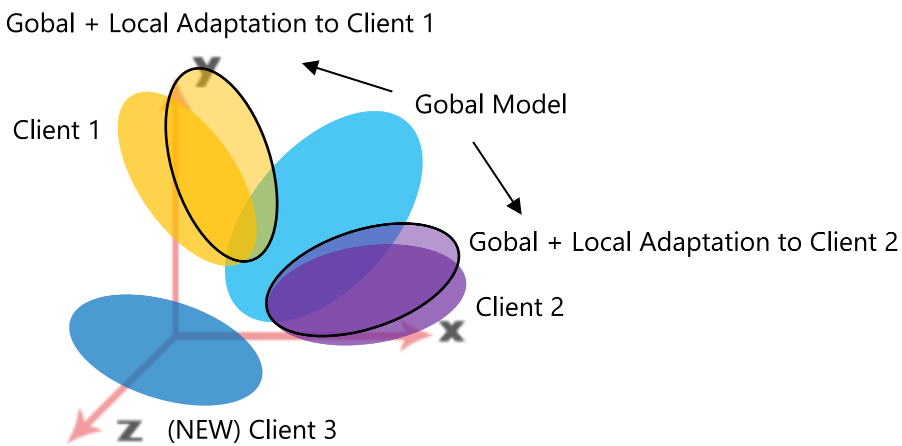


- **LC-Fed**

  Wang, Jiacheng, et al. “Personalizing Federated Medical Image Segmentation via **<u>Local</u>** Calibration.” ***ECCV* 2022**.

Shortcomgings of Above Methods

- **Lack of Generalization (Local Adaptation)**



Deficiencies that need to be improved

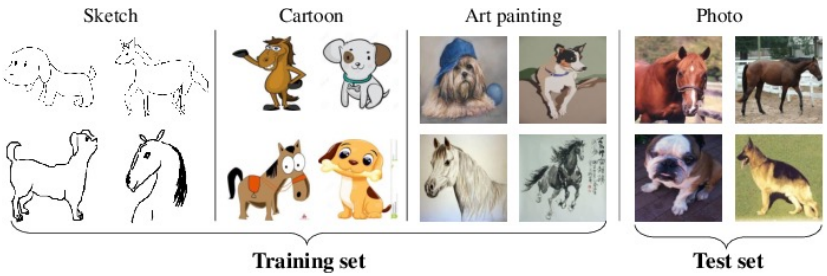
Obviously, while the global model being adapted to Client 1, it will hardly perform well in Client 2 or any other new clients. **In other words, the methods with local adaptation will lose the ability to generalize to non-local area.**

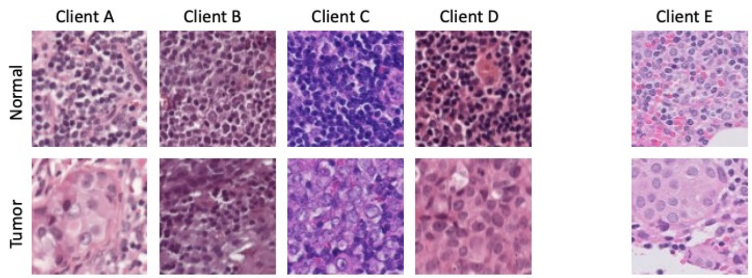
Motivation

(For AI Research Contribution) We want to design a federated method **without any client-specific components** to achieve a higher generalization performance.
(For Social Contribution) To find a way to train a federated medical imaging model that can **benefit all patients** instead of only the patients from certain parting hospital.

Domain Adaptation

- **In traditional centralized learning, there is a class of methods for solving sample feature heterogeneity called Domain Adaptation (DA).**





Similar Task

Main Challenge

Due to restrictions on access to cross-domain data in FL, divergence between domains (clients) is hard to be computed.

- **In mainstream works of DA, most of the methods aim to find a domain-invariant feature extraction ϕ, by minimizing the discrepancy between different domains.**

Minimization Objective

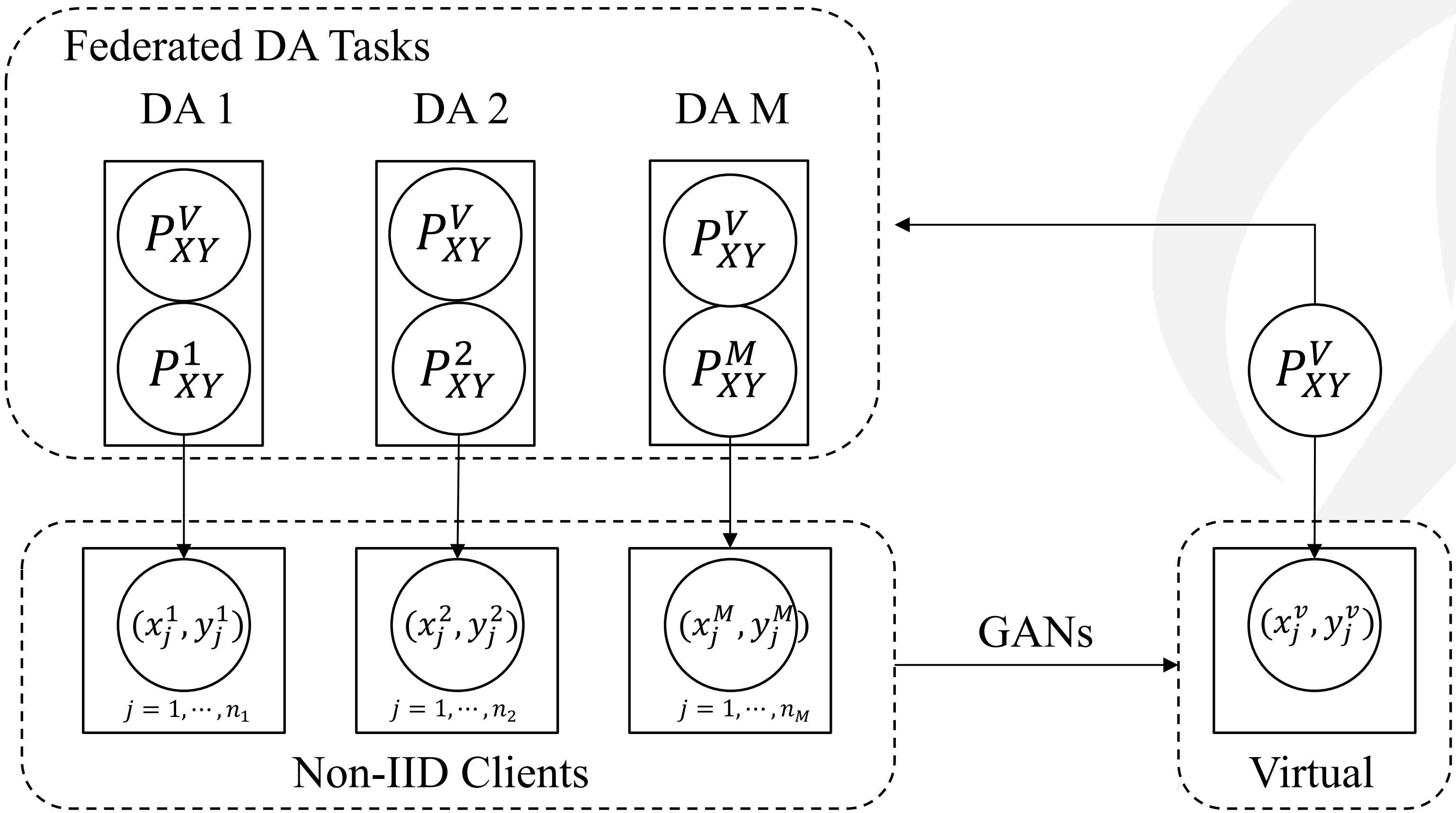
$$\varepsilon^t(h) < \boxed{\varepsilon^s(h)} + \boxed{d_H(P_X^s, P_X^t)} + \boxed{\lambda_H}$$

Constant

Wang, Jindong, et al. "Generalizing to unseen domains: A survey on domain generalization." *IEEE Transactions on Knowledge and Data Engineering* (2022).
Litjens G, Bandi P, Ehteshami Bejnordi B, Geessink O, Balkenhol M, Bult P, Halilovic A, Hermsen M, van de Loo R, Vogels R, Manson QF, Stathonikos N, Baidoshvili A, van Diest P, Wauters C, van Dijk M, van der Laak J. 1399 H&E-stained sentinel lymph node

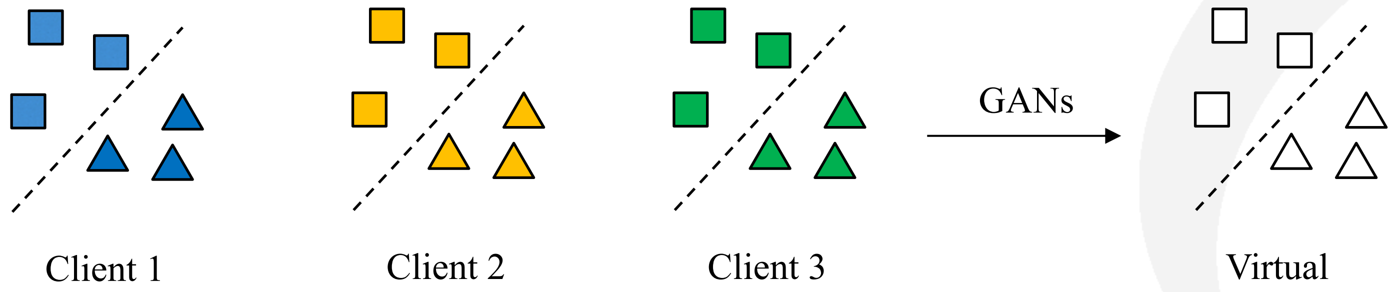
Our Ideas: Shared Virtual Dataset

- **Use GANs to create an virtual dataset as a common source for each client to perform federated domain adaptation task.**

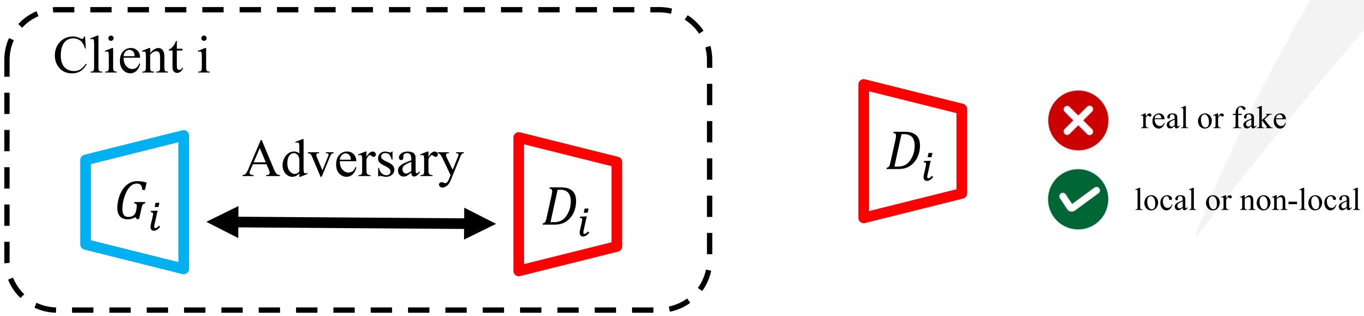


Non-IID Issue of GANs

- **Due to that we consider the virtual dataset as the source domain, so that we hope the virtual dataset can preserve as many global homogeneous features as possible and not contain any client-specific features.**

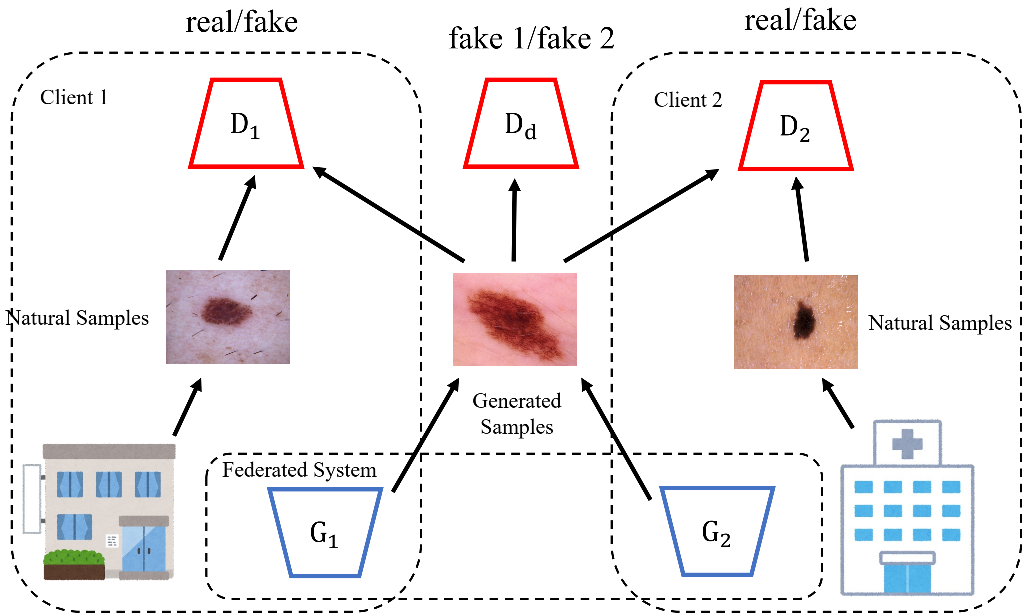


- **However in Non-IID distribution, the discriminator from different clients will learn different understanding of the real samples.**



Virtual Homogeneous Generation

- **The key to going over the bottleneck caused by Non-IID data is avoiding the client-specific tendencies of client discriminators. Based on this idea, we propose a distributed GAN architecture with a common global adversary.**



For client discriminators:

$$V(G_i, D_i) = \mathbb{E}_{z\sim P_z}[\log(1 - D_i(G_i(z)))] + \mathbb{E}_{X\sim P_d}[\log(D_i(x))]$$

For global adversary:

$$R(D_d) = \mathbb{E}_{z\sim p_z}[\log D_d^i G_i(z)]$$

The complete objective function:

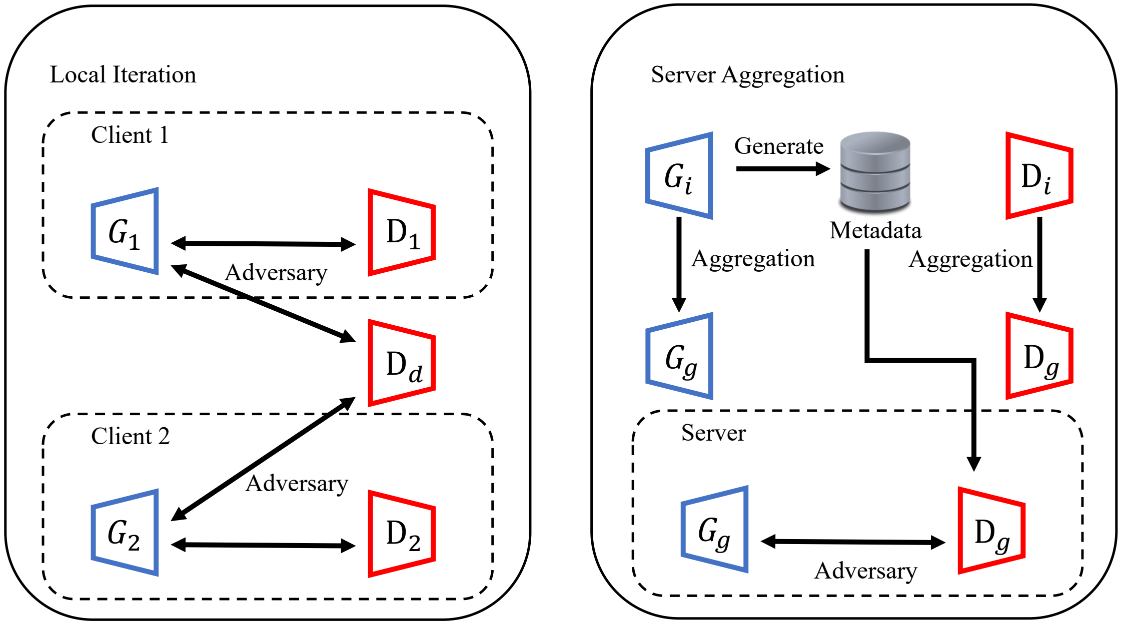
$$\sum_{i=1}^{K} \overbrace{V(G_i, D_i)}^{reality} + \overbrace{\lambda R(D_d)}^{generalization}$$

Optimization:

$$\min_{\{G_i\}_{i=1}^K} \max_{D_d} \max_{\{D_i\}_{i=1}^K} \sum_{i=1}^{K} V(G_i, D_i) + \lambda R(D_d)$$

Metadata Retraining

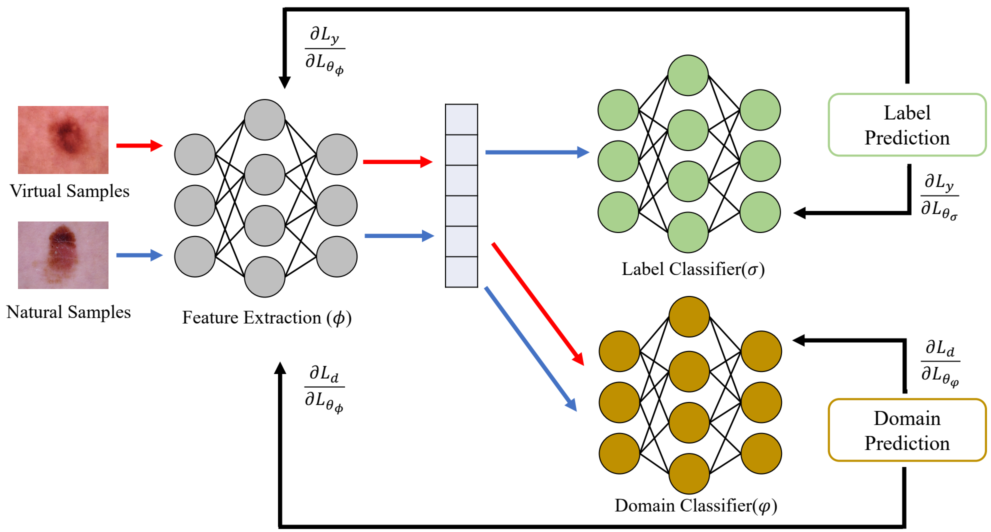
- **To further achieve unbiased, at each round of aggregation, the server also collects metadata produced by each client generator. We update the federated model with aggregated client parameters and retrain this model towards generalization with all client generations (metadata) to obtain a bias-free model.**



Adversarial Domain Adaptation

- **Our approach aims to find a feature extraction ϕ, to obtain features that cannot disseminate between the target domain (the real) and source domain (the fake), Among which we can assure that the obtained feature containing no client-specific information due to that the generated contains only characteristics of commonality.**

- **We implement our generalization goal with a parallel adversarial classifier for domain prediction.**

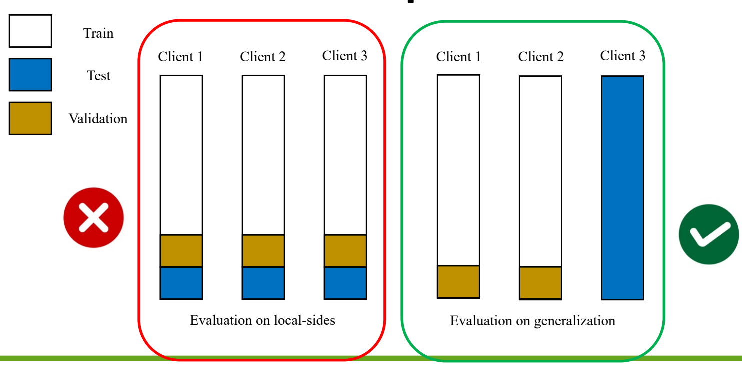


The HAM10000 Dataset

- **We use a famous challenging public skin lesion dataset HAM10000 (Human Against Machine with 10000 training images).**

Source	License	Number of samples each category						
		akiec	bcc	bkl	df	mel	nv	vasc
Rosendahl	CC BY-NC 4.0	295	296	490	30	342	803	3
ViDIR Legacy	CC BY-NC 4.0	0	5	10	4	67	350	3
ViDIR Current	CC BY-NC 4.0	32	211	475	51	680	1832	82
ViDIR Molemax	CC BY-NC 4.0	0	2	124	30	24	3720	54

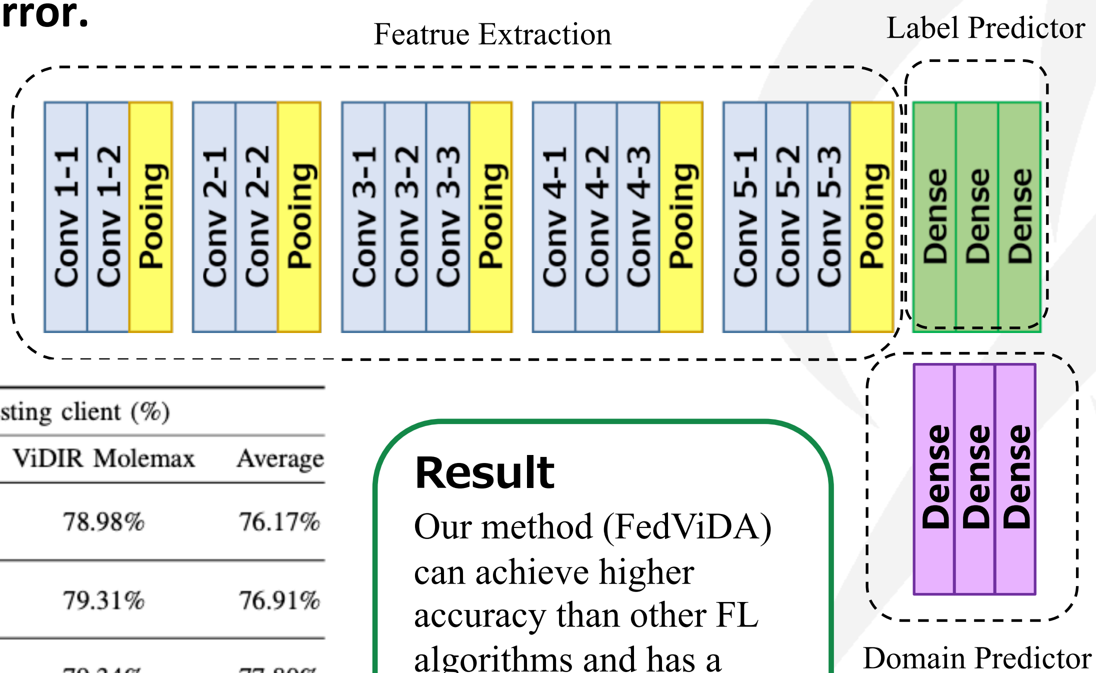
- **Considering that the samples of three of these categories are too few, we only use the four most populated. And divide them into 2 classes.**
  - Class 0: Sample of *nv* and *bkl*, considered as benign lesions
  - Class 1: Sample of *mel* and *bcc,* strongly associated with potential skin cancer.
- **We want to emphasize the generalization performance of the model, the test set will not be split from the training clients.**



Tschandl, P., Rosendahl, C. & Kittler, H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci Data* **5**, 180161 (2018).

Evaluation

- **We use a deep neural network of VGG-16 as our computing model for clients and train it over 100 epochs with a pre-defined aggregation frequency. We use the cross-entropy function to calculate the loss of this binary classification error.**

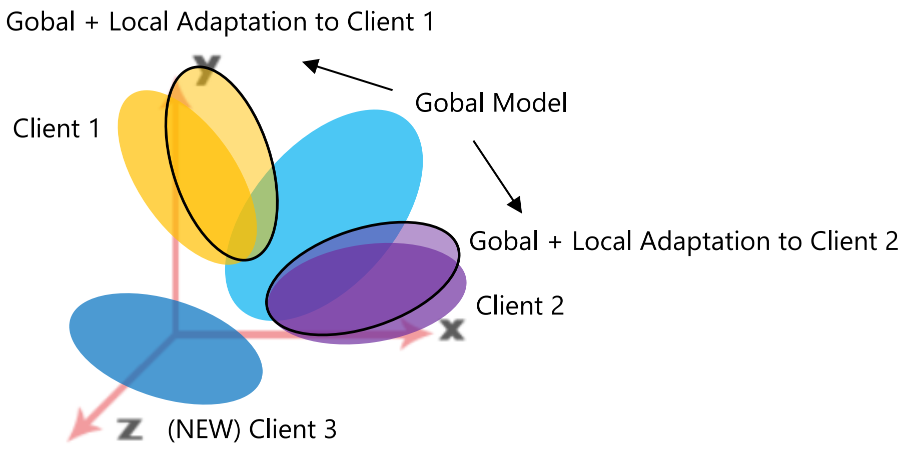


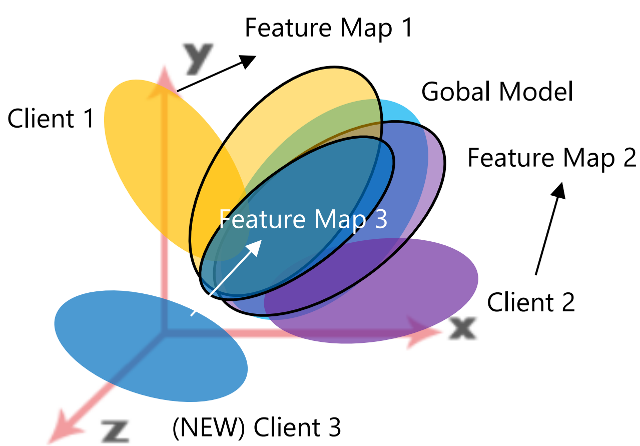
Methods	Accuracy on the testing client (%)			
	Rosendahl	ViDIR Current	ViDIR Molemax	Average
FedAvg (PMLR2017)	78.91%	70.62%	78.98%	76.17%
FedProx (MLSys2020)	78.52%	72.9%	79.31%	76.91%
FedNova (NeurIPS2020)	77.92%	76.14%	79.34%	77.80%
FedViDA (Ours)	**80.22%**	**80.17%**	**83.01%**	**81.13%**

Result

Our method (FedViDA) can achieve higher accuracy than other FL algorithms and has a very stable performance.

Discussion: Challenge & Opportunity





Federated personalized learning (Mainstream)

- **High local performance**
- Weak generalized performance
- Weak robustness

Learning commonality, Adapting personality.

Federated generalized learning (Ours)

- Acceptable local performance
- **Higher generalized performance**
- **Higher robustness**

Learning commonality, Removing personality.

Thank You for Your Attention