Federated Learning and Differential Privacy for Medical Image Analysis

Mohammed Adnan, Jesse Cresswell, Shivam Kalra, Graham Taylor, Hamid Tizhoosh

UNIVERSITY OF WATERLOO — UNIVERSITY of GUELPH — VECTOR INSTITUTE FOR ARTIFICIAL INTELLIGENCE — layer6

The paper in 30 seconds:

Question 1: Can Federated Learning (FL) be used when hospitals have heterogenous and non-IID data?

Question 2: Can Differentially Private FL achieve comparable performance with strong privacy guarantee?

Yes, we empirically demonstrate that Federated Learning can be used for sharing and analyzing complex medical images.
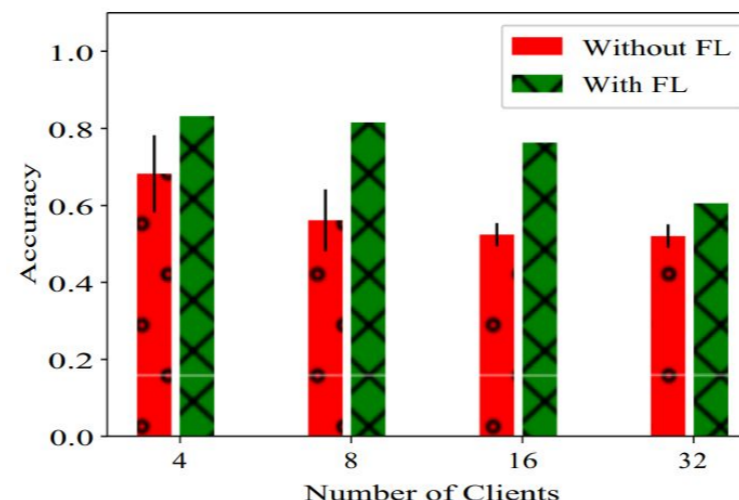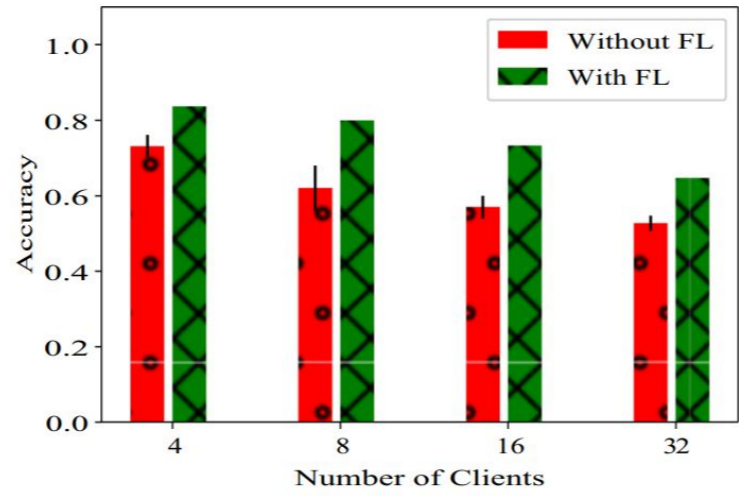
Federated Learning

1. Federated learning algorithms learn from decentralized data distributed across various client devices, in contrast to conventional learning algorithms

$$\omega^* = \frac{\omega_1 + \omega_2 + \omega_3 + \omega_4}{4}$$
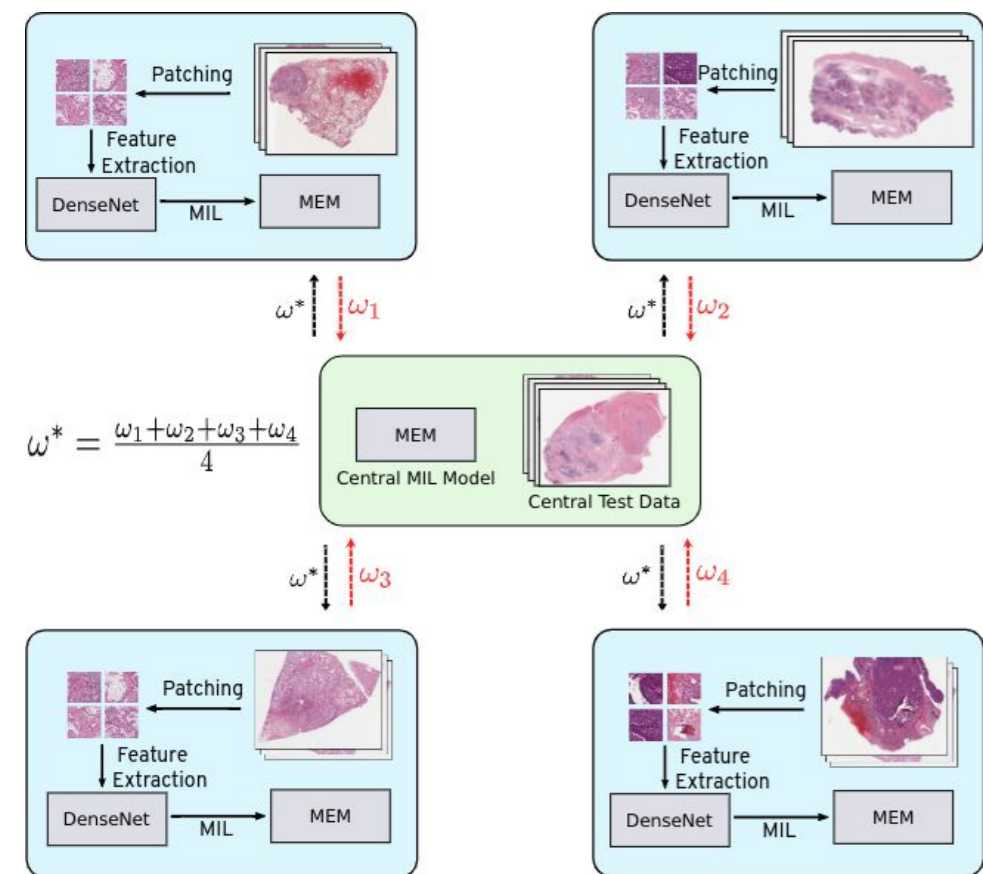
Experiment Series 1

We study the effect of IID/non-IID data & number of clients on FL performance.

(a) IID Data

(b) Non-IID Data

Experiment Series 2

1. In the second experiment series, we considered the effect of distributional differences from different source hospitals, and a requirement to preserve privacy.
2. We use data from actual hospitals to create client data and perform external validation.

Source hospital	Non-collaborative training		DP-FL training		FL training	
	Test	External	Test	External	Test	External
International Genomics Consortium	0.654	0.631				
Indivumed	0.648	0.556	0.823±0.01	0.707±0.01	0.823±0.01	0.741±0.01
Asterand	0.709	0.701				
John Hopkins	0.681	0.600				

References:

1. M. Adnan, S. Kalra, J. Cresswell, G. Taylor, H.Tizhoosh, Federated learning and differential privacy for medical image analysis, Nature Scientific Reports.
2. Poster template: Rashidinejad et al. at NeurIPS 2021