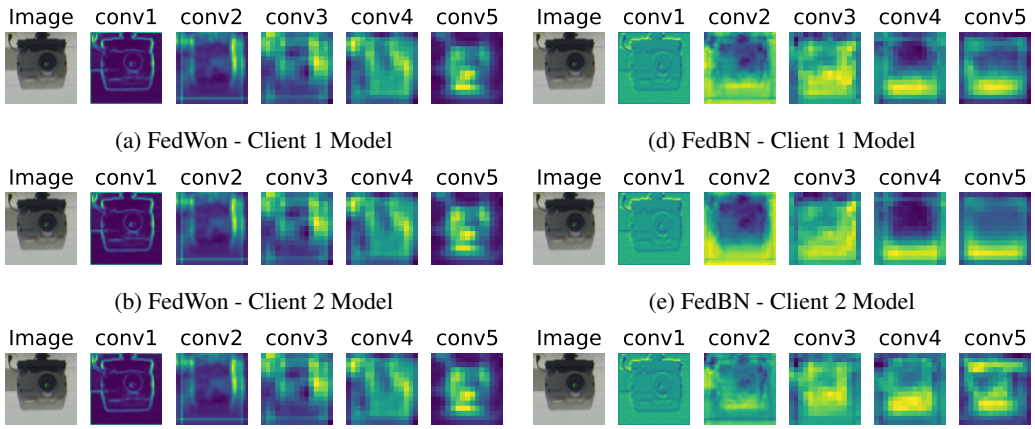
Is Normalization Indispensable for Multi-domain Federated Learning?

Anonymous Author(s)
Affiliation
Address
email



(a) FedWon - Client 1 Model

(b) FedWon - Client 2 Model

(c) FedWon - Server Global Model

(d) FedBN - Client 1 Model

(e) FedBN - Client 2 Model

(f) FedBN - Server Global Model

Figure 1: Visualization of feature maps of FedBN and our proposed FedWon on WebCam domain of the Office-Caltech-10 dataset.

Table 1: Comparison on average L2 Distance and cosine similarity of feature maps among client models and bewteen server and client models.

Methods		L2 Distance ↓	Cosine Similarity ↑
FedBN	Among Client Models	1.620	0.711
FedWon		**0.868**	**0.993**
FedBN	Bewteen Server and Client Models	1.574	0.752
FedWon		**0.870**	**0.998**

Figure 1 presents the visualization of feature maps obtained through different models of FedBN and our newly proposed FedWon. These feature maps are the output of each convolution layer in AlexNet on Office-Caltech-10 dataset. These models are two models from clients (client 1 and client 2) and one server global model. The feature maps are FedBN are rather distinct in different client models and especially in server and client model. In contrast, the feature maps of different models are quite similar in FedWon. Besides, Table 1 compares the average L2 Distance and cosine similarity of feature maps among client models and bewteen server and client models. FedWon achieves lower L2 distances and higher cosine similarities among feature maps. These results demonstrate that our proposed FedWon with scaled weight standardization can effectively mitigate the domain shifts among different domains.

Submitted to 37th Conference on Neural Information Processing Systems (NeurIPS 2023). Do not distribute.