Robust Federated Learning with Noisy and Heterogeneous Clients
Supplementary Material

Xiuwen Fang[1], Mang Ye[1,2*]

[1]National Engineering Research Center for Multimedia Software, Institute of Artificial Intelligence, Hubei Key Laboratory of Multimedia and Network Communication Engineering, School of Computer Science, Wuhan University, Wuhan, China

[2] Hubei Luojia Laboratory, Wuhan, China

https://github.com/FangXiuwen/Robust_FL

Rationale Behind CCR

In this section, we demonstrate the fundamental principles behind CCR, *i.e.*, the loss can reflect the label quality of private dataset, and the loss drop rate indicates the learning efficiency of local model. We evaluate the CE loss of adding 20% symflip noise on different clients as shown in Fig. 1. Four different networks, ResNet10, ResNet12, ShuffleNet, and Mobilenetv2, are set to four clients respectively. Moreover, one of them is a noisy client, containing 20% of the symflip label noise.

*Corresponding Author: Mang Ye (yemang@whu.edu.cn)

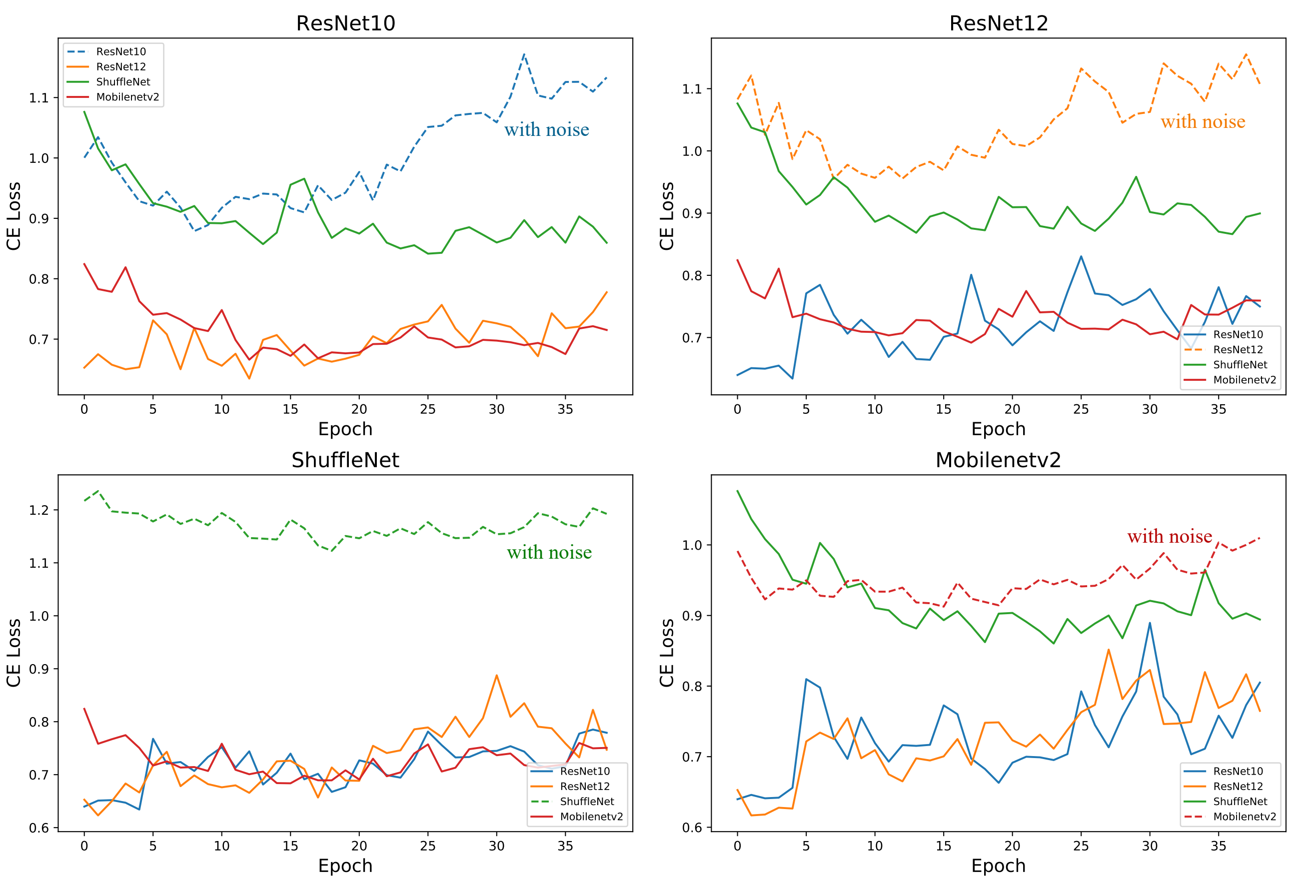

Figure 1. Evaluation of four scenarios under 20% symflip noise, and dotted lines represent the learning curves under lable noise, while solid lines denote the the learning curves with clean labels. For example, the upper left figure represents adding 20% symflip label noise to the client with ResNet10 as the local model and the other three clients all contain clean private datasets.

1) Lable Quality. From Fig. 1 we can observe that the existence of noisy clients in the federated learning system leads to a serious negative impact on the convergence of all models. With the iteration of training, the performance of local model on the noisy client will gradually decrease. Meanwhile, other clean clients will also be affected to varying degrees. Compared to the clean clients, the loss on the noisy client is much larger, which is calculated between the predictive output of the local model on the private noisy dataset and the given label. Especially when the local model of the noisy client is ShuffleNet, we can see that the loss of the noise client is higher than 1.1, while the loss of other clean clients is mostly concentrated below 0.9. This observation shows that the smaller the loss calculated by local model on private dataset, the higher the client's label quality. This motivates our design, in which the label quality of a client can be well estimated by loss value.

2) Learning Efficiency. Fig. 1 shows that the drop rate of loss on the noisy client is significantly slower than that on clean clients. In particular, when the noisy client is ResNet10, we can observe that as the training rounds increase, the loss drop rate on the noisy client is negative and significantly smaller than on the clean clients. We analyze the reason for this phenomenon is the low learning efficiency of the local model on the noisy client, which further leads to a low loss drop rate. Then the loss drop rate is measured to quantify the learning efficiency.