FedMOON算法的复现与分析

沈书鹏

摘要

摘要： 联邦学习使多方能够协作训练机器学习模型，而无需传输本地数据。联邦学习的一个关键挑战是处理各方本地数据分布的异构性。本报告详细描述了对原论文中模型对比联邦学习MOON框架的复现过程及其对比试验和结果分析。通过深入分析原论文的方法论和实验设计，本报告首先将作者的预实验复现，印证作者的初步想法，重现代码的关键步骤、所遇到的挑战以及最终结果。此外，还对原论文的贡献和限制进行了评估，提出了可能的改进方向。

关键词：联邦学习；数据异构；对比学习

1 引言

由于日益增长的隐私问题和数据保护法规 [1]，各方无法将其私人数据发送到集中式服务器来训练模型。作为解决方案，引入了联邦学习 [2]，这是一种使多方能够共同学习机器学习模型而无需交换本地数据的方法。在原论文中，提出了一种新的方法来解决非独立同分布（non-IID）问题，这是从直观的观察和一个新颖的角度出发的。核心观点是，在整个数据集上训练的全局模型能够比在偏斜的子集上训练的局部模型学习到更好的表征信息。为了实现这一目标，提出了一种称为模型对比学习MOON [3]的方法。 MOON的主要思想是通过最大化当前局部模型学习到的表示与全局模型学习到的表征信息之间的一致性，从而纠正局部更新。这种方法的目的是改善局部模型的学习质量，尤其是在数据分布不均匀的情况下。

2 相关工作

2.1 联邦学习

联邦学习在本地模型在设备上独立训练，然后只分享模型参数或更新，而不是原始数据。联邦学习中的一个关键挑战是不同参与方数据分布的异构性，表现为数据的非独立同分布。非独立同分布（non-IID）是指在机器学习和特别是联邦学习中，数据在不同节点或参与者之间的分布不是统一或相同的。图 1中展示了CIFAR-10数据集N-IID下的数据分布，由于数据的差异性，与联邦学习框架的局限，局部模型与全局模型将会导致偏移。图 2中展示了在N-IID下的模型偏移，如何处理这种数据异构导致的模型偏移，成为联邦学习的一大热点。

FedAvg [4]是经典的联邦学习算法。其步骤为：服务器向各方发送全局模型；各方执行随机梯度下降（SGD）以在本地更新他们的模型；本地模型被发送到中央服务器；服务器对

模型权重进行平均，生成用于下一轮训练的全局模型。FedProx [5]在本地训练期间在目标中引入了近端项。近端项是根据当前全局模型和局部模型之间的 2norm 距离计算的。因此，局部模型更新在局部训练期间受到近端项的限制。SCAFFOLD [6]通过引入控制变量来纠正局部更新。与训练模型一样，控制变量也在本地训练期间由各方更新。局部控制变量和全局控制变量之间的差异用于校正局部训练中的梯度。FedMA [7]利用贝叶斯非参数方法以分层方式匹配和平均权重。FedAvgM [8]在更新服务器上的全局模型时应用动量。

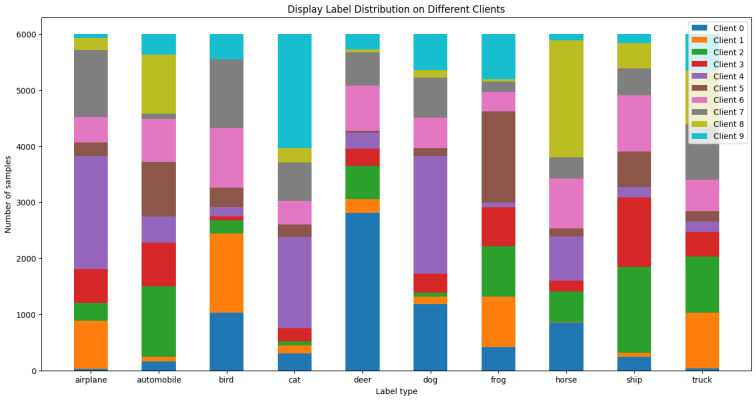

图 1. 数据异构分布图

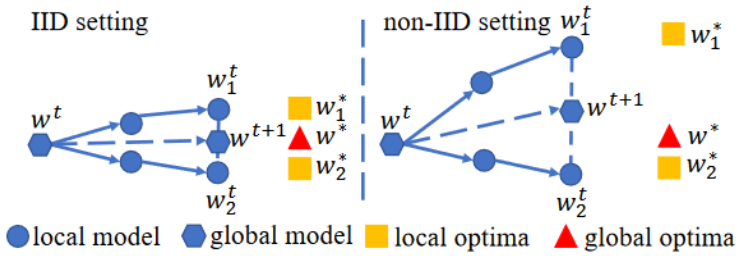

图 2. N-IID下的模型偏移

2.2 对比学习

对比学习是最近的一个热门研究方向，它试图从未标记的数据中学习良好的数据表征。它的核心思想是通过比较不同数据样本来学习数据表征。在这种方法中，模型被训练以区分正样本和负样本的数据对。通过这种方式，模型能够学习到更加丰富和区分性的特征表示。典型的对比学习框架是SimCLR [9]。图 3中展示了SimCLR框架中的NT-Xent Loss。在本文中，基于一个直观的想法，全局模型相较于本地模型能够提取更好的特征表示，因此提出模型对比学习来比较不同的模型，从中学习表征信息。

$$l_{i,j} = -\log \frac{\exp(\text{sim}(x_i, x_j)/\tau)}{\sum_{k=1}^{2N} \mathbb{I}_{[k \neq i]} \exp(\text{sim}(x_i, x_k)/\tau)}$$

图 3. NT-Xent Loss

3 本文方法

本文对作者预实验与MOON算法进行复现，因此分别介绍其预实验内容与MOON算法基本框架。

3.1 预实验

MOON基于一个直观的想法：在整个数据集上训练的模型能够比在倾斜子集上训练的模型提取更好的特征表示。具体来说，我们首先在CIFAR10上训练CNN模型。我们使用t-SNE [10]来可视化测试数据集中图像的隐藏向量. 图 4中展示了CIFAR-10上隐藏向量的t-SNE可视化。具体的实现方法是，以不平衡的方式将数据集划分为10个子集，并在每个子集上训练CNN 模型，测试数据集图像的隐藏向量。其中分别为初始化的CNN模型，训练后的CNN模型，FedAvg下Local的CNN模型与FedAvg下Global的CNN模型。我们可以观察到，与Local模型相比，Global模型具有相同类别的点更发散。局部训练阶段甚至由于局部数据分布偏斜导致模型学习到较差的表征，这进一步验证了全局模型应该能够学习到比局部模型更好的特征表示，局部更新存在模型漂移。

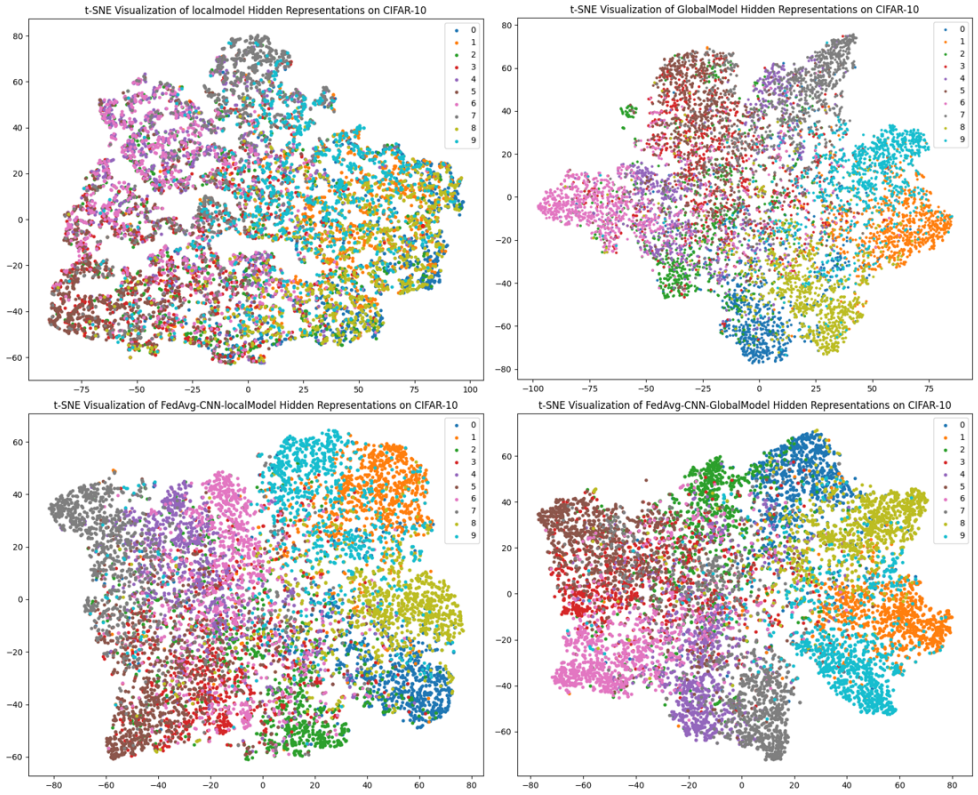

图 4. CIFAR-10上隐藏向量的t-SNE可视化

基于以上直觉，作者提出MOON。仅在本地训练阶段引入轻量级但新颖的修改。由于局部训练总是存在漂移，并且全局模型比局部模型学习到更好的表征信息，因此MOON的目标是减少局部模型与全局模型之间的表征距离，并增加局部模型之间表征距离。

3.2 MOON

MOON即为模型对比联邦学习，MOON在模型层面创新性地应用了对比学习原理，它的重点是减少局部模型和全局模型学习的表征之间的差异。仿照对比学习经典框架SimCLR提出MOON框架，如图 5所示。

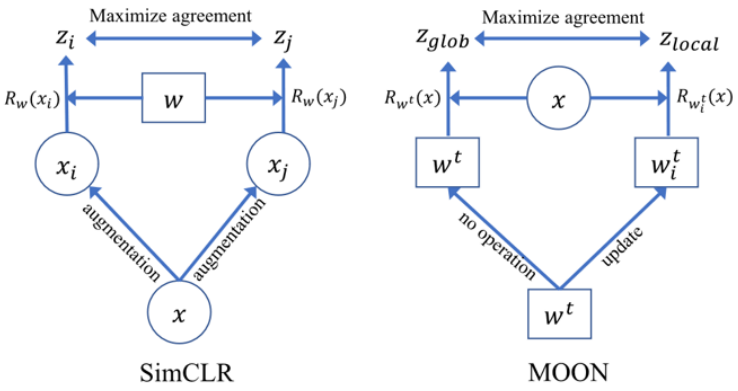

图 5. SimCLR与MOON结构对比

SimCLR与MOON的比较。这里x表示图像，w表示模型，R表示计算表征的函数。 SimCLR最大化同一图像的不同表征之间的一致性，而MOON最大化小批量上局部模型和全局模型表征的一致性。

MOON和传统对比学习之间的主要区别：MOON目前用于联邦环境中的监督学习，而对比学习则用于集中环境中的无监督学习。 MOON受到对比学习的启发，是一种新的学习方法，用于处理联邦学习中各方之间的非独立同分布数据分布。参考前文NT-Xent Loss接下来介绍MOON Loss。

3.3 损失函数MOON Loss定义

根据对比学习的核心思想，提出模型对比损失与监督学习损失。其中监督学习损失计算本地数据与预测数据之间的差异，也就是传统的损失函数。而模型对比损失则是由两个部分构成，一部分是计算本地模型与全局模型的差异，另一部分是计算本地模型与先前的本地模型的差异。从前文我们得知，本地模型的表征信息应当远离先前的模型，而靠近全局模型的表征信息。我们通过余弦相似度来计算参数相似度，要注意的是，在联邦学习中，一般传递的参数为更新梯度，因此要先计算出梯度更新后的模型参数然后再计算模型对比损失。这与NT-Xent Loss有相似的思想，此处添加参数μ来控制对比损失的影响。如图 6所示。

$$\ell_{\text{con}} = -\log \frac{\exp(\text{sim}(z, z_{\text{glob}})/\tau)}{\exp(\text{sim}(z, z_{\text{glob}})/\tau) + \exp(\text{sim}(z, z_{\text{prev}})/\tau)}$$

$$Model - Contrastive\ Loss$$

$$\ell = \ell_{\text{sup}}\left(w_i^t; (x, y)\right) + \mu \ell_{\text{con}}\left(w_i^t; w_i^{t-1}; w^t; x\right)$$

$$MOON\ \text{Local}\ Loss$$

图 6. MOON Loss function

4 复现细节

原论文的对经典算法FedAvg的改动只有本地模型的更新策略，即限定本地模型偏移的更新策略。这是为了让本地模型向能更好表征全局信息的全局模型靠近。与前文提到所提到的方法内容一致，本文着重复现文章的预实验与MOON框架，接下来将详细介绍它们的实现。

4.1 与已有开源代码对比

原论文中并没有给出预实验的具体代码，此部分由本人按照作者的思路实现，细节与原论文方法可能有所不同，但其实验结果相似。MOON算法使用了作者的代码与结构，但作者的代码似乎有些问题，CPU的占用率会非常高。

```
# 下载和加载CIFAR-10数据集
testset = torchvision.datasets.CIFAR10(root='./data', train=False, download=True, transform=transform)
testloader = torch.utils.data.DataLoader(testset, batch_size=100, shuffle=False, num_workers=2)

# 加载预训练的模型
model = ModelFedCon(base_model = 'simple-cnn', out_dim = 256, n_classes = 10)
model.load_state_dict(torch.load('D:\\Code\\Python\\MOON\\MOON_1\\models\\fedavg\\globalmodelexperiment_log-2023-12-06-1530-43.pth'))
model = nn.Sequential(*list(model.children())[:-1])
model.eval()

# 遍历测试数据集，获取模型的隐藏层表征
all_hidden_representations = []
all_labels = []

with torch.no_grad():
    for images, labels in testloader:
        hidden_representations = model(images)
        all_hidden_representations.append(hidden_representations.view(hidden_representations.size(0), -1).numpy())
        all_labels.append(labels.numpy())

# 将列表转换为numpy数组
all_hidden_representations = np.concatenate(all_hidden_representations, axis=0)
all_labels = np.concatenate(all_labels, axis=0)

# 使用t-SNE进行降维
tsne = TSNE(n_components=2, random_state=42)
reduced_representation = tsne.fit_transform(all_hidden_representations)

# 可视化t-SNE降维后的结果
plt.figure(figsize=(10, 8))
for i in range(10):
    plt.scatter(reduced_representation[all_labels == i, 0], reduced_representation[all_labels == i, 1], label=str(i), s=10)

plt.title('t-SNE Visualization of FedAvg-CNN-GlobalModel Hidden Representations on CIFAR-10')
plt.legend()
plt.show()
```

图 7. 隐藏层可视化实现

4.2 实验环境搭建

本文利用sklearn库实现t-SNE降维可视化；使用Pytorch库实现联邦学习基本框架；运行环境为CPU:Intel(R) Xeon(R) Gold 6154 与 GPU:GTX Titan X；

4.3 创新点

本次实验创新点相对较少，主要着重与对于现有工作的复现。重现了作者的预实验：抽取了4种隐藏层向量进行可视化，数据集的表征向量，训练后的CNN的隐藏层向量，FedAvg中本地模型的隐藏层向量，FedAvg中全局模型的隐藏层向量，实现逻辑再4.1节图 7中，其实验结果在3.1节图 4中，此部分为独立完成。而MOON算法对作者的开源代码进行调试与修改，尝试多卡并行与效率优化。

5 实验结果分析

此部分将展示MOON算法与联邦学习主要算法的数据指标。由于联邦学习训练时间非常漫长，因此只测试了CIFAR-10这个数据集。以下将分别对算法精度，通信效率与本地训练轮数对联邦学习N-IID下的影响来探讨分析。图8展示了N-IID数据分区策略。图9展示了MOON算法与其他算法的精度比较。

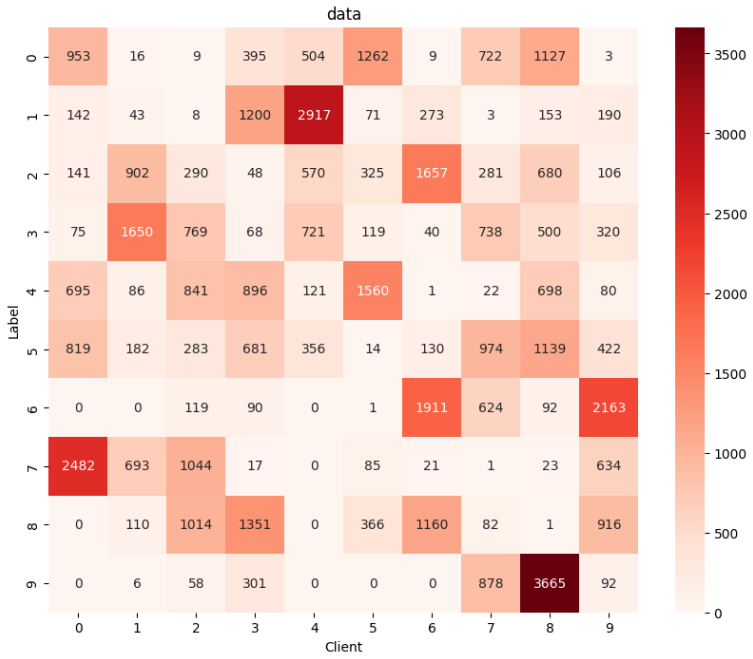

图 8. N-IID数据划分

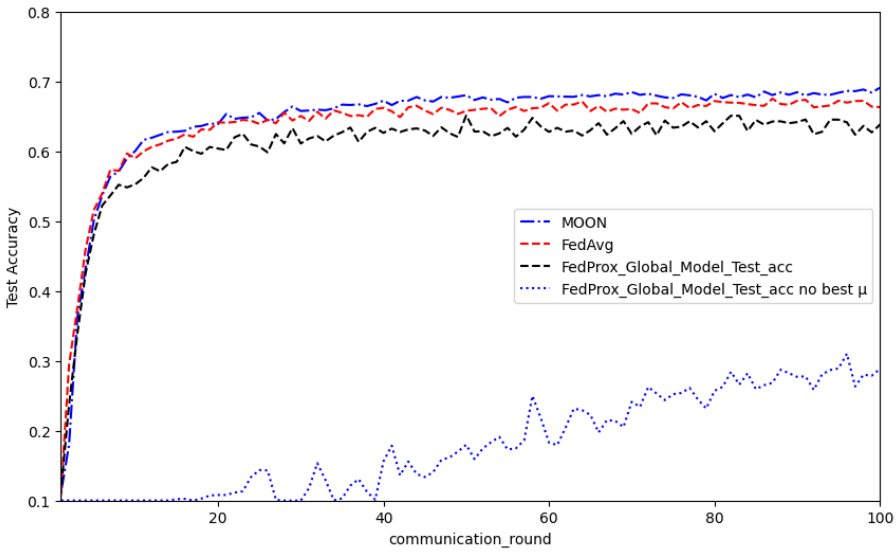

图 9. Accuracy Comparison

在N-IID环境设置下，对比不同的联邦学习方法。MOON算法的超参数使用作者给出的数值，我们可以观察到MOON确实是精度最高的方法，虽然对比FedAvg有提升，但效果并不是非常显著。对于FedProx，其中也有超参数需要设置，当μ没有设置为很小的值时，FedProx的

收敛速度相当慢，并且没有参数优化的FedProx准确性很差，我们加入优化器进行优化，最终得到了一个差强人意的结果，但仍然与经典的FedAvg有所差距。但MOON算法的设置不仅仅在精度上得到提升，由于更新策略的改变，势必对收敛速度有所影响，接下来我们分析通信效率与本地训练轮数对精度的影响。

表 1. Communication Efficiency & Number of Local Epochsq

Method	CIFAR-10		EPOCHS				
	#rounds	speedup	1	5	10	20	40
FedAvg	100	1×	0.619	0.663	0.660	0.660	0.663
FedProx	52	1.9×	0.615	0.659	0.658	0.646	0.633
MOON	**27**	**3.7×**	**0.621**	**0.675**	**0.680**	**0.678**	**0.671**

以FedAvg训练100次通信轮数的正确率为基准，测试各种联邦学习算法使用的通信轮数，从上方的表中我们可以知道，FedProx（此处为参数优化后）在加入修正项后得到的训练轮数是比FedAvg要少的，但是它有些不稳定，最后得到的正确率反而比FedAvg要低。MOON的通信轮数则更好了，它有更好的通信效率。此外，实验研究了不同本地轮数下正确率的影响，可以发现每一种联邦学习算法在开始从零开始增加epochs时的本地模型的正确率都是增加的，但是随着epochs的增加，正确率反而下降了，我认为这是过拟合导致的。因为数据异构导致模型偏移，使得过拟合程度更加严重，但是由于MOON增加了本地限制，让本地模型向有更好表征信息的全局模型接近，它在各种epochs上的测试结果均优于其他算法。

6 总结与展望

MOON在N-IID下，解决数据标签漂移上有很好的表现，但是局限与同模型的网络体系，即无法在模型异构的场景下使用。其次文章的方法虽然从对比学习中得到思路，但是这个方法不应当只局限于图像数据集中，其他的数据也可以通过思考表征信息的聚合来达到模型对比联邦学习的目的。因此模型对比学习（MOON），这是一种简单而有效的联邦学习方法。MOON引入了一种新的学习概念，即模型级别的对比学习。实验表明，MOON在图像分类任务上比最先进的方法取得了一些改进。

参考文献

[1] Paul Voigt and Axel Von dem Bussche. The eu general data protection regulation (gdpr). *A Practical Guide, 1st Ed., Cham: Springer International Publishing*, 10(3152676):10–5555, 2017.

[2] Peter Kairouz, H Brendan McMahan, Brendan Avent, Aurélien Bellet, Mehdi Bennis, Arjun Nitin Bhagoji, Kallista Bonawitz, Zachary Charles, Graham Cormode, Rachel Cummings, et al. Advances and open problems in federated learning. *Foundations and Trends® in Machine Learning*, 14(1–2):1–210, 2021.

[3] Qinbin Li, Bingsheng He, and Dawn Song. Model-contrastive federated learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10713–10722, 2021.

[4] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*, pages 1273–1282. PMLR, 2017.

[5] Tian Li, Anit Kumar Sahu, Manzil Zaheer, Maziar Sanjabi, Ameet Talwalkar, and Virginia Smith. Federated optimization in heterogeneous networks. *Proceedings of Machine learning and systems*, 2:429–450, 2020.

[6] Sai Praneeth Karimireddy, Satyen Kale, Mehryar Mohri, Sashank Reddi, Sebastian Stich, and Ananda Theertha Suresh. Scaffold: Stochastic controlled averaging for federated learning. In *International conference on machine learning*, pages 5132–5143. PMLR, 2020.

[7] Hongyi Wang, Mikhail Yurochkin, Yuekai Sun, Dimitris Papailiopoulos, and Yasaman Khazaeni. Federated learning with matched averaging. *arXiv preprint arXiv:2002.06440*, 2020.

[8] Tzu-Ming Harry Hsu, Hang Qi, and Matthew Brown. Measuring the effects of non-identical data distribution for federated visual classification. *arXiv preprint arXiv:1909.06335*, 2019.

[9] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, pages 1597–1607. PMLR, 2020.

[10] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008.