5X-07

モデルのパラメータの信頼度を考慮したFederated Learning

蒐場　涼介†　　渡辺　裕‡
早稲田大学　基幹理工学部†

1. まえがき

Federated Learning (FL) は，複数のクライアントがデータを共有せずに，協調して機械学習モデルを訓練する分散型学習手法である．この手法は，各クライアントにおいて学習させたモデルパラメータのみをサーバで集約するため，プライバシ情報の保護を実現できる．しかしこのデータ保持の制約に起因して，クライアント間のデータ分布が不均一な場合に，モデルの学習が進みにくいことが課題となっている．そこで本稿では，この課題に対処するため，新たなモデルの集約手法を提案する．各クライアントのパラメータの情報量に基づき，それらのパラメータが共有部の更新に与える貢献度を調整する．Cifar10とCifar100を用いた分類タスクでの実験により，本手法が特定のケースにおいて従来の手法よりも有効であることを示す．

2. 関連手法

2.1 Personalized Federated Learning

Personalized Federated Learning (PFL) [1]は，クライアント間のデータ分布が不均一なケースに対処するためのFLの手法である．従来のFLではモデル全体のパラメータを共有し学習するが，PFLではモデルの一部を共有せず，専有部として学習する．まず，各クライアントにおけるモデルのパラメータW_iをθとw_iに分割する．θは共有部，w_iは専有部，$i \in \{1, \dots, M\}$は各クライアントの番号，Mはクライアントの数を表す．PFLの学習ステップは以下の2つに分けられる．

- **クライアントの学習**：サーバから共有部のパラメータθ^tを受信し，クライアントiのパラメータを$W_i^t = (\theta^t, w_i^t)$で初期化する．ローカルで指定のエポック数$E_{local}$学習し，更新されたパラメータ$W_i^{t+1} = (\theta_i^t, w_i^{t+1})$を得る．
- **モデルの集約**：各クライアントでの学習前後のパラメータの差分$\Delta\theta_i^t = \theta_i^t - \theta^t$を集約し，新しい共有部のパラメータ$\theta^{t+1}$を作成する．そして，再びクライアントに配布し，クライアントで学習させる．

Federated Learning Considering the Confidence Score of Model Parameters
†Ryosuke Nutaba, †Hiroshi Watanabe, †Waseda University

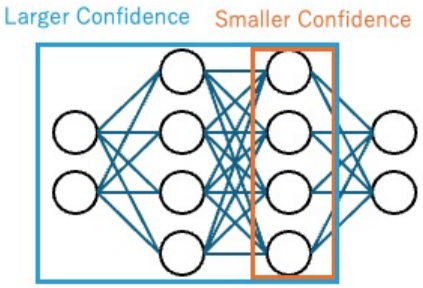

図1　提案手法概要図

ただし，tは$t \in \{1, \dots, E_{global}\}$番目の学習ラウンドを表し，$E_{global}$は共有部を更新する回数を表す．以上により，クライアント間のデータ分布の差異による学習進行の阻害を軽減できる．

2.2 FedAS

FedAS[2]はPFLにおけるクライアント内の不整合とクライアント間の不整合を解決するための手法である．前者の問題は，クライアントの学習ステップの開始時，θ^{t+1}とw_i^tが一貫していない状態を指す．後者の問題は，クライアントごとの学習の進行度合いが不均一な状態を指す．FedASでは，前者の問題に対し，クライアントの学習ステップの最初にθ^{t+1}をθ^tに近づけるParameter-Alignment (PA)を提案し，後者の問題に対し，モデルの集約のステップにて，各クライアントのモデルのパラメータにFisher Information Matrix (FIM) [3]を用いて重み付けるClient-Synchronization (CS)を提案している．

3. 提案手法

FedASでは，学習の進行を妨げる質の悪いパラメータを，FIMを用いて効率的に学習から取り除くが，FIMの小さいパラメータの貢献度をすべての層で一律で削減してしまっているため，データ分布の差異に敏感な入力層では，質の悪いパラメータの影響が依然として大きい．そこで本稿では，クライアントから収集したモデルのパラメータ$\Delta\theta_i^t$の信頼度をFIMで評価し，信頼度の大きさに応じて，グローバルモデルの各層ごとに，更新に用いるクライアントとその貢献度を変更する．

まず，各クライアントiにおいて，学習終了時点でFIMを計算し，そのFIMのトレースの値α_i

表 1．Cifar10 の分類タスクにおける正解率

Method	$\beta = 0.1$			$\beta = 0.5$			$\beta = 1.0$		
	$P = 0.2$	$P = 0.6$	$P = 1.0$	$P = 0.2$	$P = 0.6$	$P = 1.0$	$P = 0.2$	$P = 0.6$	$P = 1.0$
FedAS	87.02	87.40	**87.24**	75.75	77.13	77.55	**69.60**	71.47	71.69
Ours	**87.25**	**87.53**	87.23	**75.90**	**77.22**	**77.91**	69.53	**71.74**	**71.94**

表 2．Cifar100 の分類タスクにおける正解率

Method	$\beta = 0.1$			$\beta = 0.5$			$\beta = 1.0$		
	$P = 0.2$	$P = 0.6$	$P = 1.0$	$P = 0.2$	$P = 0.6$	$P = 1.0$	$P = 0.2$	$P = 0.6$	$P = 1.0$
FedAS	54.19	56.69	57.54	36.20	38.38	**39.75**	29.07	32.20	32.40
Ours	**54.57**	**57.52**	**58.24**	**36.75**	**38.84**	39.25	**29.66**	**32.45**	**32.57**

を得る．θのj層目のパラメータθ_jの更新には，α_iの値が大きい順に，式(1)に示されるM_j個のクライアントを選択する．

$$M_j = \left\lfloor \frac{j}{N_\theta} \cdot M \right\rfloor. \tag{1}$$

ただし，θの層数をN_θ，$j \in \{1, \dots, N_\theta\}$とする．次に，利用する$M_j$ 個のクライアントのα_i を，式(2)で示されるように正規化し，式(3)を用いて層内のクライアントの貢献度を調整する．

$$\bar{\alpha}_{ij} = \frac{\alpha_i}{\Sigma_{k \in S_j} \alpha_k}, i \in S_j. \tag{2}$$

$$\theta_j^{t+1} = \theta_j^t + \sum_{i \in S_j} \bar{\alpha}_{ij} \cdot \Delta\theta_{ij}^t. \tag{3}$$

ここで，S_jは選択されたj層目のM_j個のクライアントの集合，$\bar{\alpha}_{ij}$はj層目でのクライアントiの貢献度，θ_{ij}はクライアントiのj層目のパラメータを表す．この重み付けによって，入力に近い層ほど，信頼度の小さいパラメータの比重が小さくなる．

4. 実験

4.1 実験設定

Cifar10[4], Cifar100[4]データセットを用いた画像分類タスクにて提案手法の評価を行う．データの分割には，パラメータβに従うディリクレ分布を用いる．βは大きいほどデータの分布の偏りが大きくなる．モデルには 4 層の CNN を用いる．最初の 3 層を共有部，最終層を専有部とする．Cifar10 では出力層を 10，Cifar100 では出力層を 100 に設定する．

クライアント数Mは20，グローバルエポック数E_{global}は 40，ローカルエポック数E_{local}は 5 と設定する．また，クライアントのうち，1 回の共有部の更新に参加する割合をP として導入し，Pに応じて 1 回の共有部の更新に参加するクライアントをランダムに決定する．また，クライアント内の不整合に対処するため，FedAS の PA を用いる．

4.2 実験結果

Cifar10 と Cifar100 の分類タスクにおいて，$\beta \in \{0.1, 0.5, 1.0\}$，$P \in \{0.2, 0.6, 1.0\}$ に設定したときの，従来手法である FedAS と提案手法の正解率の値をそれぞれ表 1，表 2 に示す．表 1 と表 2 より，多くのケースで，本手法が従来手法よりも有効であるとわかる．また，Cifar10 に比べ，Cifar100 での精度の上昇幅が大きいことがわかる．この原因の 1 つとして，クラス数が多い Cifar100 ではデータ分布の差異が大きく，それに伴い質の低いパラメータが多く含まれていたが，本手法では，入力層付近において質の悪いパラメータを効果的に除去できたことが考えられる．

5. むすび

本稿では，モデルパラメータの FIM を信頼度として評価し，信頼度に応じてグローバルモデルの更新に関与する度合いを決定する手法を提案した．Cifar10 と Cifar100 での分類タスクでの実験の結果，従来手法に比べて，分類精度を向上することを示した．

参考文献

[1] Z. Qu *et al.*, “How to Prevent the Poor Performance Clients for Personalized Federated Learning?” Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 12167-12176.

[2] X. Yang *et al.*, “FedAS: Bridging Inconsistency in Personalized Federated Learning,” Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024, pp. 11986-11995.

[3] S. Amari, “Natural Gradient Works Efficiently in Learning,” Neural Computation, vol. 10, no. 2, pp. 251-276.

[4] A. Krizhevsky, “Learning multiple layers of features from tiny images,” 2009.