Assessing the Implications of Data Heterogeneity on Privacy-Enhanced Federated Learning: A Comprehensive Examination Using CIFAR-10

Phoebe Joanne Go
Department of Electronics and Computer Engineering
De La Salle University
City of Manila, Philippines
phoebe_go@dlsu.edu.ph

Victor Calinao Jr.
Department of Electronics and Computer Engineering
De La Salle University
City of Manila, Philippines
victor_calinaojr@dlsu.edu.ph

Melchizedek Alipio
Department of Electronics and Computer Engineering
De La Salle University
City of Manila, Philippines
melchizedek.alipio@dlsu.edu.ph

Abstract*— In the context of an increasingly digital society with pressing privacy concerns, our research investigates the effectiveness of privacy-preserving artificial intelligence solutions like Federated Learning. This study focuses on three main areas: Federated and Centralized Learning applications, the influence of data heterogeneity on client data accuracy, and the evaluation of contemporary federated algorithms in scenarios of extreme heterogeneity. Federated Learning, using the FedAvg algorithm, demonstrated superior testing accuracy (88.54%) over Centralized Learning (87.98%) on the non-heterogeneous CIFAR-10 dataset, indicating its potential as an efficient, privacy-preserving solution for various machine learning applications. Additionally, our findings highlight an inverse relationship between data heterogeneity and Federated Learning model accuracy, underscoring the need for strategies to mitigate this challenge and boost model performance. Upon evaluating several federated learning algorithms under high data heterogeneity (alpha=1.0), SCAFFOLD and FedOpt outperformed FedAvg and FedProx, demonstrating the significance of algorithm design in addressing data heterogeneity. SCAFFOLD and FedOpt showcased greater communication efficiency, attributed to their faster convergence and fewer required communication rounds. This study offers invaluable insights into addressing data heterogeneity, improving communication efficiency, and enhancing federated learning's performance and applicability in real-world scenarios, thereby furthering privacy-preserving artificial intelligence research.

Keywords—federated learning, data heterogeneity, cifar-10, FedAvg, FedOpt, FedProx, SCAFFOLD

I. INTRODUCTION

The increasing universality of digital technology and the proliferation of data have led to a growing concern about data privacy and security. Traditional centralized data control and storage architectures expose sensitive information to potential breaches, posing risks to users' privacy [1]. With the increasing prevalence of the Internet of Things (IoT) and the subsequent generation of vast amounts of data, the need for effective privacy preservation techniques has become more urgent than ever [2]. The advent of federated learning (FL) presents a potential resolution to prevailing privacy concerns by enabling decentralized, collaborative learning among multiple devices or organizations without sharing raw data [3].

FL has been successfully applied in various domains, including query suggestions for Google Keyboard [3], electronic health record (EHR) analysis [4], credit card fraud detection [5], and predicting patient outcomes for COVID-19 [6]. However, one critical challenge in FL is data heterogeneity, the differences in data distribution (e.g. numeric, categorical, text-based, etc.) across the devices or servers participating in the learning process, which arises when the data distribution across clients or devices is non-identically and independently distributed (non-IID) [11]. Data heterogeneity can negatively impact the convergence and generalization of FL models, affecting the model's performance [8],[9].

In applications like smartphone apps predicting user behavior, unique user data causes data heterogeneity. Federated learning trains models locally on each phone, and only model updates are shared with a central server. The process iterates until the global model performs optimally. For significant data deviations, algorithms like SCAFFOLD handle this heterogeneity, guaranteeing accurate and reliable federated learning.

Recent efforts have been made to assess and address data heterogeneity in FL. Several benchmarks and frameworks have been proposed to evaluate the impact of data heterogeneity on FL [9],[10]. Researchers have also explored techniques to improve the efficiency and robustness of FL in the presence of heterogeneous data, such as personalized federated learning [12],[17], communication-efficient federated learning [13],[14],[15],[16], and federated multi-task learning [21]. However, there is a need for a comprehensive examination of the implications of data heterogeneity on privacy-enhanced federated learning, especially with real-world datasets.

This paper contributes the following:

a) A comprehensive comparison of Federated Learning and Centralized Learning using the widely used image classification dataset, CIFAR-10.

b) An in-depth analysis of the impact of data heterogeneity on the performance and evaluation of several state-of-the-art FL algorithms (FedAvg, FedOpt, FedProx, and SCAFFOLD) under extreme heterogeneity conditions.

c) Provides valuable understanding of data heterogeneity issues in federated learning environments, sheds light on the practicality and constraints of various algorithms and suggests potential research paths for federated learning.

In this paper, section II explores the current algorithms like FedAvg, FedOpt, FedProx, and SCAFFOLD; section III describes their experimental design, including dataset selection and heterogeneity manipulation and measurement methods; section IV showcases findings, such as data heterogeneity's effect on model efficacy and a comparative analysis of various federated learning algorithms; and section V summarizes essential insights and stresses the need to account for data heterogeneity in federated learning algorithms.

II. Review of Related Literature

Data heterogeneity is a significant challenge in FL, as it can adversely affect the convergence and generalization of FL models [7],[8],[9]. Several researchers have proposed techniques to address data heterogeneity in FL, including personalized federated learning [12],[17], communication-efficient federated learning [13],[14],[15],[16], and federated multi-task learning [21].

Personalized federated learning has been proposed to tailor the learning process to each client, considering their unique data distributions [12],[17]. Yang et al. [12] provided a comprehensive overview of federated learning concepts and applications, highlighting the challenges of non-IID data and the need for personalized learning approaches. Fallah et al. [17] introduced a meta-learning approach for personalized federated learning, enabling the learning algorithm to adapt to the unique characteristics of each client's data distribution.

Communication-efficient federated learning techniques aim to minimize the communication overhead existing between the engaged clients and the primary server. [13],[14],[15],[16]. McMahan et al. [16] proposed a communication-efficient learning approach for deep networks based on decentralized data, which significantly reduced the quantity of information transmitted between clients and servers. Li and Song [13] developed a privacy-preserving, communication-efficient federated multi-armed bandits algorithm, ensuring both efficient learning and privacy preservation.

Federated multi-task learning (FML) is another approach to tackle data heterogeneity by learning multiple related tasks simultaneously [21]. Smith et al. [21] introduced an FML framework that allows clients to learn multiple tasks concurrently while sharing a global model. This methodology has demonstrated enhancements in FL model performance amidst data heterogeneity, concurrently maintaining data confidentiality and security.

Researchers have also explored various techniques to mitigate the impact of data heterogeneity in FL through data augmentation [8], redefining data heterogeneity [9], and analyzing the convergence of FL algorithms on non-IID data [25]. De Luca et al. [8] address the overfitting, limited training data and data heterogeneity using data augmentation. Morafah et al. [9] proposed a notion of data heterogeneity called "data diversity," standard benchmark called "FedHetero," and a novel FL approach called "Semantic-Aware Federated Learning (SAFL)" to address semantic differences across devices and assess the impact of data heterogeneity on FL.

FL has been successfully applied to various domains that require privacy preservation, including healthcare [4],[6],[22],[23], finance [5], and language modeling [22]. For instance, Chen et al. [22] developed and validated a federated learning model for the International Classification of Diseases, 10th Revision (ICD-10) classification using deep contextualized language models, demonstrating the effectiveness of FL in preserving privacy while achieving high classification accuracy. Li et al. [23] presented a multi-site functional magnetic resonance imaging (fMRI) assessment utilizing privacy-maintaining federated learning, in tandem with domain adaptation, showcasing the potential of FL in collaborative medical research.

In the context of privacy-enhanced FL, researchers have investigated various techniques to ensure the robustness and privacy of the learning process, such as Byzantine-robust aggregation schemes [19], loss-aware weight quantization [20], and federated optimization in heterogeneous networks [18]. Li et al. [19] conducted an experimental study on Byzantine-robust aggregation schemes, demonstrating their effectiveness in providing robustness against adversarial clients in FL. Hou and Kwok [20] proposed a loss-aware weight quantization method for deep networks to improve the communication efficiency while maintaining model accuracy. Li et al. [18] explored federated optimization in heterogeneous networks, addressing the challenges of clients with different computational capabilities, network bandwidths, and data distributions.

III. Implementation

TABLE I
PRESENTS THE IMPLEMENTATION ENVIRONMENT FOR THE FEDERATED OPTIMIZATION ALGORITHMS (FEDAVG, FEDOPT, FEDPROX, AND SCAFFOLD) TESTED ON THE CIFAR-10 DATASET

Hardware	ASUS FX505DY AMD Ryzen 5 Laptop
Software	Windows 11 (64-bit) Home
	Python 3.8
	TensorFlow 2.9
	NVFlare 2.2
Parameters	Dataset: CIFAR-10 Federated Optimization: FedAvg, FedOpt, FedProx, SCAFFOLD

A. Baseline vs Federated Learning

Wang et al. [24] introduced nonparametric techniques to address data heterogeneity by employing Dirichlet sampling. In the current implementation, this method is utilized to generate heterogeneous data partitions from the CIFAR10 dataset. The degree of heterogeneity is denoted by the alpha parameter, which ranges from 0 to 1. An alpha value of 1 signifies a completely homogeneous data distribution, while an alpha of 0 indicates an extreme heterogeneous partitioning of the dataset.

TABLE II
PARAMETER I. CENTRALIZED LEARNING VS FEDERATED LEARNING

Centralized Learning Executes	Edge Client	1
	Alpha	1
	Local Epochs	25
	Total Updates	25
Federated Learning Executes	Edge Clients	8
	Alpha	1
	Communication Rounds	50
	Local Epochs	4

	Total Updates	25

Table II presents the parameter configurations for the comparative study of Centralized Learning (CL) and Federated Learning (FL) implementations. The CL method incorporates a single client containing the entire CIFAR10 dataset in its edge data storage, exhibiting a purely homogeneous data distribution. It should be noted that the alpha value for Centralized Learning is consistently set at 1. Conversely, the FL approach involves 8 clients and employs the FedAvg optimization algorithm with a homogeneous database partitioning scheme. To establish a valid benchmark comparison between the federated learning clients, all of them were assigned identical local epoch parameters.

Effect of Data Heterogeneity

Data heterogeneity refers to the discrepancies or variations in data distributed among decentralized devices or clients. It can considerably affect the performance of a federated model, as it may impact the quality and uniformity of the model updates aggregated from participating clients.

TABLE III
EFFECT OF DATA HETEROGENEITY

FedAvg Executes	Edge Clients	8
	Alpha	1.0, 0.5, 0.3, 0.1
	Communication Rounds	50
	Local Epochs	4
	Total Updates	25

Table III examines the performance of FedAvg optimizer under varying alpha values, highlighting the importance of assessing data heterogeneity in federated learning.

B. *FL Algorithms Assessments in Extreme Heteregoneity*

Federated Averaging [25] (FedAvg), Federated Optimization [26] (FedOpt), Federated Proximal [18] (FedProx), and Self-Adjusting Federated Learning through Communication-Efficient LOcal Descent [27] (SCAFFOLD) were implemented under identical setups and hyperparameters in a context of extreme data heterogeneity.

TABLE IV
FEDAVG VS FEDOPT VS FEDPROX VS SCAFFOLD

FedAvg, FedOpt, FedProx, and SCAFFOLD Executes	Edge Clients	8
	Alpha	0.1
	Communication Rounds	50
	Local Epochs	4
	Total Updates	25

Federated Averaging (FedAvg) is a widely-used optimization algorithm in the domain of federated learning. Introduced by McMahan et al. [16], it has become a popular choice for training machine learning models in a decentralized fashion, while maintaining data privacy and minimizing communication overhead.

Essentially, the algorithm allows a model to be trained across multiple decentralized nodes, with each node learning from its local data and subsequently sharing its local model updates with a central server. This local model training significantly reduces the need for data exchange over the network, which greatly enhances computational efficiency and decreases communication costs.

The privacy-preserving nature of the FedAvg algorithm stems from this decentralized approach to learning. By ensuring that raw data never leaves its original node, and only model updates are communicated to the central server, the algorithm effectively safeguards individual data privacy. This feature becomes exceedingly valuable in scenarios involving sensitive data, such as healthcare or financial data.

FedAvg has been implemented in various studies to address different challenges in FL. For example, Zhao et al. [11] investigated the effects of non-IID data in a heterogeneous setting and proposed strategies to overcome the challenges posed by such data distributions. In another study, Sattler et al. [28] explored the combination of knowledge distillation and federated learning with the aim of enhancing the global model's efficacy. These studies, among others, have demonstrated the versatility and effectiveness of the FedAvg algorithm in diverse scenarios.

It suggests that the choice of hyperparameters, the handling of data heterogeneity, and the communication efficiency between clients and the central server play significant roles in determining the performance of the federated model.

Algorithm 1: FedAvg Algorithm

```
Input: global_model, clients_data, num_comm_rounds, local_epochs
Output: global_model

Initialize num_clients ← length(clients_data)
for round ← 1 to num_comm_rounds do
    Initialize clients_weights ← empty list
    for client_idx ← 1 to num_clients do
        local_model_weights ← TrainLocalModel(global_model, clients_data[client_idx], local_epochs)
        Append (length(clients_data[client_idx]), local_model_weights) to clients_weights
    end for
    aggregated_weights ← AggregateModelWeights(clients_weights)
    Set global_model weights to aggregated_weights
end for
Return global_model
```

Federated Optimization (FedOpt) is an advanced optimization algorithm that aims to improve the performance and communication efficiency of decentralized machine learning models. It extends the Federated Averaging (FedAvg) algorithm by incorporating techniques such as adaptive learning rates and momentum-based optimization, which enable more efficient and robust model convergence.

FedOpt has been used in numerous studies to tackle FL challenges, such as, Li et al. [29] presented FedOpt and showed its enhanced performance over FedAvg. Additionally, Caldas et al. [10] studied how adaptive learning rates affect federated learning algorithms, including FedOpt.

These studies highlight the value of adaptive learning rates and momentum-based optimization in improving the convergence of federated models. Such techniques enable FedOpt to adjust to each client's local data traits, streamline model updates, lessen communication overhead, and boost overall performance.

Algorithm 2: FedOpt Algorithm

```
Input: global_model, clients_data, num_comm_rounds, local_epochs
Output: global_model

Initialize num_clients ← length(clients_data)
Initialize global_velocity ← zero tensor with the same shape as global_model.trainable_variables
for round ← 1 to num_comm_rounds do
```

```
    Initialize clients_weights_velocities ← empty list
    for client_idx ← 1 to num_clients do
        (local_model_weights, local_velocity) ←
TrainLocalModel(global_model, clients_data[client_idx], local_epochs,
global_velocity)
        Append (length(clients_data[client_idx]), local_model_weights,
local_velocity) to clients_weights_velocities
    end for
    (aggregated_weights, aggregated_velocities) ←
AggregateModelWeightsVelocities(clients_weights_velocities)
    Set global_model weights to aggregated_weights
    Set global_velocity to aggregated_velocities
end for
Return global_model
```

The FedOpt algorithm optimizes a global model through iterative communication rounds by aggregating local model updates from participating clients. Clients train their local models using adaptive learning rates and momentum-based optimization, improving convergence properties. The server updates the global model and global velocity tensor, ensuring efficient and effective federated learning.

The Federated Proximal (FedProx), is an optimization algorithm designed for FL systems with non-IID data and heterogeneous client participation. The algorithm builds on FedAvg by introducing a proximal term that mitigates the impact of stragglers and clients with poor-quality updates.

This was proposed by Sahu et al. in their paper titled "On the Convergence of Federated Optimization in Heterogeneous Networks" [25]. The authors demonstrated that FedProx is well-suited for heterogeneous FL settings involving non-IID data, and compared to FedAvg, offers better convergence properties. By incorporating a proximal term, the algorithm penalizes weight updates that deviate substantially from the global model, thus improving convergence properties. It showcases its adaptability and robustness in a variety of scenarios. They conducted experiments using CIFAR-10 and FEMNIST.

Algorithm 3: FedProx Algorithm

```
Input: global_model, clients_data, num_comm_rounds, local_epochs,
learning_rate, mu
Output: global_model

Initialize num_clients ← length(clients_data)
for round ← 1 to num_comm_rounds do
Initialize clients_weights ← empty list
Initialize clients_sizes ← empty list
for client_idx ← 1 to num_clients do
local_model_weights ← ClientUpdate(global_model,
clients_data[client_idx], local_epochs, learning_rate, mu)
Append local_model_weights to clients_weights
Append length(clients_data[client_idx]) to clients_sizes
end for
aggregated_weights ← AggregateModelWeights(clients_weights,
clients_sizes)
Set global_model weights to aggregated_weights
end for
Return global_model
```

The FedProx algorithm advances global model optimization in federated learning by repeatedly collecting and merging local model updates from clients. It uses proximal regularization to manage infrequent updates in non-IID environments, improving efficiency, effectiveness, and convergence of federated learning.

SCAFFOLD (Self-Adjusting Federated Learning through Communication-Efficient LOcal Descent) is a communication-efficient federated learning algorithm that adapts to the heterogeneity in the data distribution and client participation. It is designed to address the challenges of heterogeneous data and stragglers, while maintaining communication efficiency.

One of the foundational studies on SCAFFOLD is by Karimireddy et al. [27], which introduced the algorithm and evaluated its performance on several benchmark datasets, including MNIST, CIFAR-10, and Shakespeare. The results demonstrated that SCAFFOLD outperforms other FL algorithms in terms of communication efficiency and model convergence. Furthermore, the algorithm has shown improved generalization performance, particularly in the presence of non-IID data and stragglers.

Non-IID data refers to the uneven data distribution across participating clients or nodes, due to varied user behavior, regional specifics, and device capabilities. This creates data heterogeneity, a contrast to IID data, where each data point is independently drawn with equal probability.

The SCAFFOLD algorithm excelled because it includes a correction term in client model updates to lessen data heterogeneity's impact, aiding in better model convergence and accuracy in highly heterogeneous settings. It is employing local control variates, reducing update variance, and adapting to diverse data distributions and straggling clients or clients with infrequent updates. Its design minimizes communication overhead by accelerating convergence and requiring fewer rounds.

Algorithm 4: SCAFFOLD Algorithm

```
Input: global_model, clients_data, num_comm_rounds, local_epochs,
learning_rate
Output: global_model

Initialize control_variates ← zero tensors with the same shape as
global_model.trainable_variables
for comm_round ← 1 to num_comm_rounds do
    Initialize clients_gradients ← empty list
    for client_idx ← 1 to num_clients do
        Initialize local_model ← copy of global_model
        for epoch ← 1 to local_epochs do
            gradients ← LocalUpdate(local_model,
clients_data[client_idx], control_variates, learning_rate)
            Append gradients to clients_gradients
        end for
    end for
    aggregated_gradients ← AggregateGradients(clients_gradients)
    Update global_model.trainable_variables with
aggregated_gradients
    Update control_variates with aggregated_gradients and
clients_gradients
end for
Return global_model
```

The performance of the federated model is heavily affected by factors such as communication rounds, local epochs, and client data heterogeneity.

IV. Results and Discussions

This section encompasses three key aspects of the author's contribution: (1) implementing Federated and Centralized Learning with corresponding communication rounds and local epochs, (2) analyzing the effect of data heterogeneity among clients on accuracy performance, and (3) evaluating state-of-the-art federated algorithms under extreme heterogeneity conditions.

A. Baseline Learning vs Federated Learning

Figure 1 compares the testing accuracy of Centralized Learning and Federated Learning using FedAvg algorithm. In conclusion, our study reveals that Federated Learning, achieving 88.54% accuracy, outperforms Centralized Learning with 87.98% accuracy on the non-heterogeneous CIFAR-10 dataset. This finding highlights the effectiveness of Federated Learning in addressing the challenges posed by Centralized Learning while maintaining competitive accuracy levels. As the demand for privacy-preserving, decentralized data processing methods increases, Federated Learning emerges as a valuable alternative for various machine learning applications. It is essential for researchers and practitioners to continue exploring and optimizing Federated Learning techniques to further improve their performance and applicability across diverse real-world scenarios.

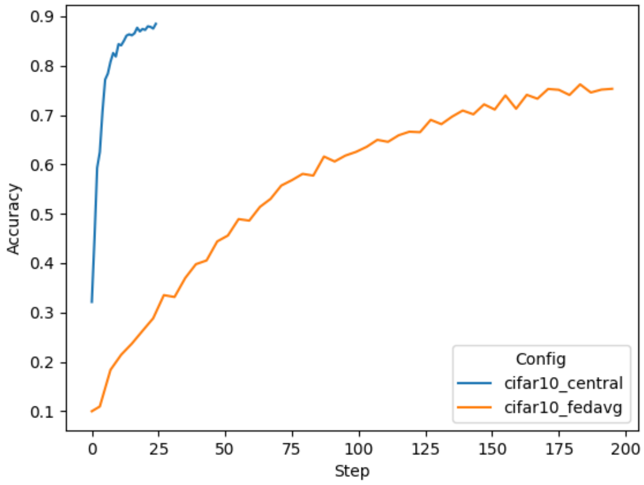

Fig. 1. Centralized learning (CL) vs Federated learning (FL)

B. *Effect of Data Heterogeneity*

In Figure 2, the test result reveals a clear inverse relationship between data heterogeneity and model accuracy in Federated Learning. As the alpha value decreases, representing increased data heterogeneity, the model's accuracy declines accordingly: alpha=1.0 (88.54%), alpha=0.5 (86.33%), alpha=0.3 (83.50%), and alpha=0.1 (77.33%). These findings emphasize the challenges posed by data heterogeneity in federated learning environments and the importance of addressing these challenges to improve model performance.

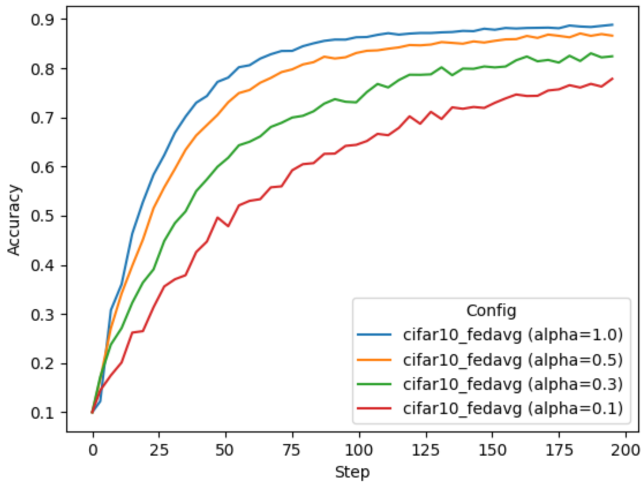

Fig. 2. FedAvg on 4 different data heterogeneity

C. *FL Algorithms Assessments in Extreme Heteregoneity*

In Figure 3, the evaluation of various federated learning algorithms in the context of high data heterogeneity (alpha=1.0) reveals a distinct performance hierarchy, it highlights the superior performance of SCAFFOLD (82.22%) and FedOpt (80.13%) over FedAvg (77.33%) and FedProx (76.15%). The improved convergence rates of SCAFFOLD and FedOpt can be attributed to SCAFFOLD's use of a correction term when updating client models and FedOpt's employment of SGD with momentum for global model updates.

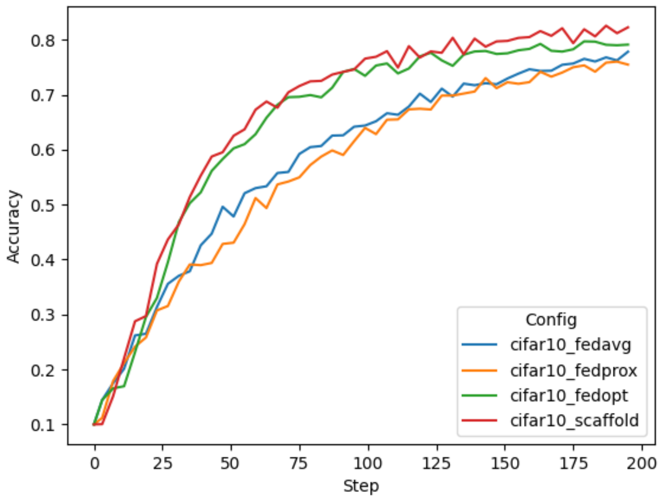

Fig. 3. FedAvg vs FedOpt vs FedProx vs SCAFFOLD

This study significantly advances the understanding of data heterogeneity in federated learning settings by investigating the performance of FedAvg, FedOpt, FedProx, and SCAFFOLD. It reveals strengths and weaknesses, deepening understanding of real-world applicability and limitations.

Under extreme heterogeneity conditions, SCAFFOLD emerges as the most efficient performer due to its design that addresses non-IID data, making it particularly adept at handling heterogeneous data in Federated Learning (FL). Both SCAFFOLD and FedOpt demonstrate superior communication efficiency by converging faster and requiring fewer communication rounds, in contrast to FedAvg and FedProx, which necessitate more rounds for convergence, thereby affecting their efficiency. Additionally, the robustness of SCAFFOLD and FedOpt to local updates is enhanced by the local control variates of SCAFFOLD and the use of Stochastic Gradient Descent (SGD) with momentum by FedOpt. This is a marked difference from FedAvg and FedProx, which may exhibit heightened sensitivity to local updates, consequently affecting their performance in heterogeneous settings. Ultimately, the enhanced scalability of SCAFFOLD and FedOpt is attributable to their capacity to manage heterogeneous data and achieve faster convergence.

V. Conclusion

Our research centered on three main areas: (1) the application of Federated and Centralized Learning with appropriate communication rounds and local epochs, (2) the investigation of data heterogeneity's impact on the accuracy of client data, and (3) the assessment of modern federated algorithms under severe heterogeneity conditions.

The study reveals that Federated Learning using the FedAvg algorithm surpasses Centralized Learning in testing accuracy (88.54% vs 87.98%) on the CIFAR-10 dataset, suggesting that Federated Learning could serve as an effective, privacy-preserving alternative for numerous machine learning applications.

We found an inverse correlation between data heterogeneity and model accuracy within Federated Learning, indicating the need for strategies to address this challenge and enhance model performance.

Furthermore, the evaluation of various federated learning algorithms under high data heterogeneity (alpha=1.0) indicates superior performance of SCAFFOLD (82.22%) and FedOpt (80.13%) over FedAvg (77.33%) and FedProx (76.15%). This highlights the importance of algorithm design in addressing data heterogeneity, with SCAFFOLD and FedOpt demonstrating better communication efficiency due to faster convergence and fewer communication rounds.

Future research should consider methods for preprocessing, partitioning, and data augmentation to boost performance in federated learning under heterogeneous data conditions. Furthermore, strategies for improved client selection, dynamic learning rates, and adaptable communication strategies could enhance the efficiency and applicability of federated learning across various scenarios. The development of adaptive data management techniques in federated learning could contribute to system performance enhancements, thus addressing the challenge of data heterogeneity.

Acknowledgment

The authors would like to acknowledge the support of the Department of Electronics and Computer Engineering of De La Salle University, Philippines.

Go, P. would like to acknowledge the support of DOST-ERDT, Republic of the Philippines and the Office of the Vice President for Research and Innovation (Conference Subsidy Graduate Students), De La Salle University, Philippines.

References

[1] P. Imrie and P. Bednar, “End User Effects of Centralized Data Control,” Nov. 2014, vol. 11. doi: 10.1007/978-3-319-23784-8_17.

[2] “Internet of Things and Privacy - Issues and Challenges - Office of the Victorian Information Commissioner.” https://ovic.vic.gov.au/privacy/resources-for-organisations/internet-of-things-and-privacy-issues-and-challenges/ (accessed Dec. 19, 2022).

[3] T. Yang et al., “Applied Federated Learning: Improving Google Keyboard Query Suggestions.” arXiv, Dec. 06, 2018. doi: 10.48550/arXiv.1812.02903.

[4] D. Liu, T. Miller, R. Sayeed, and K. D. Mandl, “FADL:Federated-Autonomous Deep Learning for Distributed Electronic Health Record.” arXiv, Dec. 02, 2018. doi: 10.48550/arXiv.1811.11400.

[5] W. Yang, Y. Zhang, K. Ye, L. Li, and C.-Z. Xu, “FFD: A Federated Learning Based Method for Credit Card Fraud Detection,” 2019, pp. 18–32. doi: 10.1007/978-3-030-23551-2_2.

[6] “Federated Learning of Electronic Health Records Improves Mortality Prediction in Patients Hospitalized with COVID-19 - PMC.” https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7430624/ (accessed Dec. 19, 2022).

[7] J. Kang, Z. Xiong, D. Niyato, Y. Zou, Y. Zhang, and M. Guizani, “Reliable Federated Learning for Mobile Networks,” IEEE Wirel. Commun., vol. 27, no. 2, pp. 72–80, Apr. 2020, doi: 10.1109/MWC.001.1900119.

[8] A. B. de Luca, G. Zhang, X. Chen, and Y. Yu, “Mitigating Data Heterogeneity in Federated Learning with Data Augmentation.” arXiv, Jun. 20, 2022. doi: 10.48550/arXiv.2206.09979.

[9] M. Morafah, S. Vahidian, C. Chen, M. Shah, and B. Lin, “Rethinking Data Heterogeneity in Federated Learning: Introducing a New Notion and Standard Benchmarks.” arXiv, Sep. 30, 2022. doi: 10.48550/arXiv.2209.15595.

[10] S. Caldas et al., “LEAF: A Benchmark for Federated Settings.” arXiv, Dec. 09, 2019. doi: 10.48550/arXiv.1812.01097.

[11] Y. Zhao, M. Li, L. Lai, N. Suda, D. Civin, and V. Chandra, “Federated Learning with Non-IID Data,” 2018, doi: 10.48550/arXiv.1806.00582.

[12] Q. Yang, Y. Liu, T. Chen, and Y. Tong, “Federated Machine Learning: Concept and Applications,” ACM Trans. Intell. Syst. Technol., vol. 10, no. 2, p. 12:1-12:19, Jan. 2019, doi: 10.1145/3298981.

[13] T. Li and L. Song, “Privacy-Preserving Communication-Efficient Federated Multi-Armed Bandits,” IEEE Journal on Selected Areas in Communications, vol. 40, no. 3, pp. 773–787, Mar. 2022, doi: 10.1109/JSAC.2022.3142374.

[14] F. Sattler, S. Wiedemann, K.-R. Müller, and W. Samek, “Robust and Communication-Efficient Federated Learning From Non-i.i.d. Data,” IEEE Transactions on Neural Networks and Learning Systems, vol. 31, no. 9, pp. 3400–3413, Sep. 2020, doi: 10.1109/TNNLS.2019.2944481.

[15] K. Bonawitz et al., “Towards Federated Learning at Scale: System Design.” arXiv, Mar. 22, 2019. doi: 10.48550/arXiv.1902.01046.

[16] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, “Communication-Efficient Learning of Deep Networks from Decentralized Data.” arXiv, Feb. 28, 2017. Accessed: Dec. 15, 2022. [Online]. Available: http://arxiv.org/abs/1602.05629

[17] A. Fallah, A. Mokhtari, and A. Ozdaglar, “Personalized Federated Learning: A Meta-Learning Approach.” arXiv, Oct. 22, 2020. doi: 10.48550/arXiv.2002.07948.

[18] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, “Federated Optimization in Heterogeneous Networks.” arXiv, Apr. 21, 2020. doi: 10.48550/arXiv.1812.06127.

[19] S. Li, C. H. E. Ngai, and T. Voigt, “An Experimental Study of Byzantine-Robust Aggregation Schemes in Federated Learning,” preprint, Apr. 2022. doi: 10.36227/techrxiv.19560325.v1.

[20] L. Hou and J. T. Kwok, “LOSS-AWARE WEIGHT QUANTIZATION OF DEEP NET-,” 2018.

[21] V. Smith, C.-K. Chiang, M. Sanjabi, and A. Talwalkar, “Federated Multi-Task Learning.” arXiv, Feb. 27, 2018. doi: 10.48550/arXiv.1705.10467.

[22] P.F. Chen et al., “Training a Deep Contextualized Language Model for International Classification of Diseases, 10th Revision Classification via Federated Learning: Model Development and Validation Study,” JMIR Medical Informatics, vol. 10, no. 11, p. e41342, Nov. 2022, doi: 10.2196/41342.

[23] X. Li, Y. Gu, N. Dvornek, L. Staib, P. Ventola, and J. S. Duncan, “Multi-site fMRI Analysis Using Privacy-preserving Federated Learning and Domain Adaptation: ABIDE Results.” arXiv, Dec. 06, 2020. doi: 10.48550/arXiv.2001.05647.

[24] H. Wang, M. Yurochkin, Y. Sun, D. Papailiopoulos, and Y. Khazaeni, “Federated Learning with Matched Averaging.” arXiv, Feb. 15, 2020. doi: 10.48550/arXiv.2002.06440.

[25] X. Li, K. Huang, W. Yang, S. Wang, and Z. Zhang, “On the Convergence of FedAvg on Non-IID Data.” arXiv, Jun. 25, 2020. Accessed: Dec. 15, 2022. [Online]. Available: http://arxiv.org/abs/1907.02189

[26] M. Asad, A. Moustafa, and T. Ito, “FedOpt: Towards Communication Efficiency and Privacy Preservation in Federated Learning,” Applied Sciences, vol. 10, no. 8, Art. no. 8, Jan. 2020, doi: 10.3390/app10082864.

[27] S. P. Karimireddy, S. Kale, M. Mohri, S. J. Reddi, S. U. Stich, and A. T. Suresh, "SCAFFOLD: Stochastic Controlled Averaging for Federated Learning." arXiv, Apr. 09, 2021. Accessed: Dec. 19, 2022. [Online]. Available: http://arxiv.org/abs/1910.0637

[28] F. Sattler, S. Wiedemann, K.-R. Müller, and W. Samek, "Robust and Communication-Efficient Federated Learning from Non-IID Data," IEEE Transactions on Neural Networks and Learning Systems, 2020.

[29] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated Optimization in Heterogeneous Networks," arXiv preprint arXiv:1812.06127v5, 2018