A review of federated learning algorithms in image classification

Xingyi Qu

Jinan-Birmingham Joint Institute, Jinan University, Guangzhou, China

qxy2003@stu2021.jnu.edu.cn

Abstract. Image classification is one of the most popular applications of machine learning. It has shown its potential in fields like healthcare, auto-driving and face recognition. Federated learning (FL) emerged in 2017, creating a major innovation to the field. The new structure brings new possibilities, but create new challenges such as data heterogeneity, privacy leakage and communication burdens in parameter updating. The paper solves the problem that there are relatively few papers providing a complete analysis relating to the application of FL in image classification. The paper contributes to describe the new challenges of FL in image classification and show the related reasons behind respectively, then analyze the current state-of-the-art algorithms designed for solving the challenges and improving image model performance by discussing the basic ideas and steps of algorithms and showing their pros and cons. The paper further more contributes to compare the performance of each model in accuracy and communication speed, and outline several possible directions for future advancement.

Keywords: federated learning, image classification, application, directions

1. Introduction

Image classification refers to the automatic process of categorization of images into multiple predefined groups [1]. It is one of the most vital and quick developing fields of computer vision and has a wide range of applications, including face recognition, autonomous driving and healthcare. Devices in these fields across the world generates or receives tremendous amounts of vision data, creating invaluable resources and chances for machine learning researchers.

However, some real-world applications of image classifications causes several more challenges to traditional machine learning solutions, which are: 1)Local models trained by different datasets will cause convergence problem of global model, limiting the capacity of the final model, 2) Data are sometimes private sensitive when they are collected in medical settings or in private automobiles, 3) uploading local data to central server and downloading complete model are time-consuming and add heavy burden to communication networks. These difficulties require novel training structures to combine with effective algorithms.

Federated learning (FL) was released by McMahan H B [2], et al in 2016 and had become one of the hottest research directions in the machine learning field. FL allows clients across different locations to train collectively, only updating their own parameters to form a global model. This novel structure provides exact solutions for the challenges mentioned above, as 1) it brings model to datasets instead of taking them to models [3], solving issues like the local devices have no enough calculation power to train image data, 2) it has distinct privacy advantages compared to classical centralized training methods since it only passes model parameters between models, making training medical image data a potential choice for healthcare institutions.

FL has already shown its potential in numerous fields where centralized training is not the most suitable solution, which is common in image classification involving privacy sensitive data. However, there is a major problem in the current research field of FL. Though numerous researchers had proposed algorithms to promote innovation and applications, there are few papers concerning the summary of the FL algorithms in the application of image classification. This paper fills up this critical blank in the research field by offering a complete analysis related to the current challenge and the state-of-the-art solutions.

This review contributes to provide a comprehensive analysis of the current advancements of the field by providing inspection to some of the state-of-the-art algorithms. The essay will guide readers through the creative design of algorithms, reveal results respectively and evaluate the pros and cons between them. Additionally, the paper will show the potential improvement within the current mainstream design and scrutinize the future prospects of FL applications in image classification.

2. Challenges and Method analysis in image classification

2.1. Challenges

In a typical image classification application with FL network, there are mainly three urgent challenges that mainstream methods focus to solve:

2.1.1. Data Heterogeneity: Heterogeneity causes local models deviate from each other, causing convergence problem of the global model. In trials of combining FL with image classification applications, researchers can easily meet these problems. Data collected are usually non-IID as each individual institute might use different image labels and adopt different image collection methods, getting images with various parameters. Moreover, data collected might be from different sources, making local models deviating further from each other. Muhammad Imran et al. points out that this kind of issue introduces troubles to the optimization of hyper-parameters and adversely affect productivity of training [4].

2.1.2. Privacy Leakage: Privacy leakage is another more essential problem, especially in analyzing medical images. In fields like autonomous driving and healthcare, image data are privacy sensitive in nature since they are usually connected with information closely related to users such as daily commute route or medial images, requiring further protection measures in model design. Although FL made an even further step towards privacy protection as it does not require data uploading, steps of communication of parameters could still bring hidden risks to training as it could reveal sensitive information to a third-party [5].

2.1.3. Communication Cost: Communication cost cannot be ignored, as FL networks usually involve multiple devices which are needed to be updated every round. Updating a large number of local models, such as those formed by image sensors and processors of millions of automobiles, adds great pressure to the network and demands improved communication structure in FL. Moreover, due to the differences in the network environment and local hardware calculation power, it is expectable to find it difficult to synchronize parameters in a large group of devices.

Based on these challenges, this section aims to give detailed explanations to algorithms realizing advancement in solving these critical issues.

2.2. Methods for solving data heterogeneity.

In order to tackle the non-IID problem in the field of image classification, researchers had designed various algorithms. General ideas are to improve the logic of model selection and suppress the local model drifting while minimizing the computational cost. Researchers have tried to realize it by combining classic machine learning methods with federated learning structures or applying novel changes to the original mathematical formula. The paper will discuss some of the algorithms based on personalized learning, contrastive learning and attention algorithm.

In 2020, Canh T. Dinh et al. [6] proposed a novel algorithm called Personalized Federated Learning with Moreau Envelops (pFedMe). Based on the original FL structure, researchers introduced the principle of personalization into the classic Federated Average (FedAvg) algorithm, formulating the problem as a bi-level problem. The bi-level problem comes from the following consideration: the global model is now found by utilizing the data aggregation from numerous clients at the outer level, while local model is optimized according to its own data distribution and is limited within a bounded distance from the global mode at the inner level.

Unlike the traditional FL method, pFedMe add a new regularization parameter λ to form a loss function with l_2-norm called Moreau Envelops for each client. Larger λ could benefit clients with capricious data from sufficient data regression, improving the effectiveness of local models, while smaller λ is capable of helping clients with abundant data focus on personalization. Overall, the general idea is to allow each client to pursue the optimal model with different directions, but not deviate far from the “reference model”, to which every client contributes.

There are still some disadvantages for pFedMe. Although it managed to solve the non-IID problem effectively, it does not include processes to reveal the best suited hyper-parameters, requiring further training loops to adapt them to specific settings. The Moreover, the improvement made by pFedMe is not comprehensive, for some datasets pFedMe only performs slightly better than other old algorithms, which means that in some real-world applications, the benefit of pFedMe is covered by the extra cost in training.

Qinbin Li et al. released Model-Contrastive Learning (MOON) in 2021 [7]. The algorithm is constructed in two levels. Figure 1 shows the general structure of MOON. In the global level, the central server tries to learn a model from each updated local model parameters, aggregating the new global model from weighted averaging based on the size of local dataset. In the local level, for each training input, the model extracts the characterization of the input from the current global model, the characterization of the input from the local model in the previous round and that from the local model being updated. Hence, the algorithm will decrease the distance between the representation learned by the local model and that of global model, while increase the distance between the local model and its own counterpart of previous iterations. The concept of “contrastive learning” is embodied by that the algorithms tries to compare these defined representations.

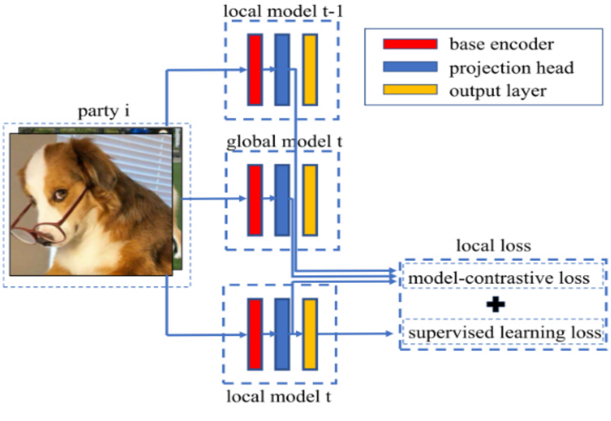

Figure 1. Flowchart showing algorithm structure of MOON [7]

Comparing with the classic method, MOON introduces new contrastive learning to FL structure. Moreover, instead of applying traditional method of learning visual data by comparing the representations of different images [8], MOON creatively adopts process of model contrast. By differentiating the parameters of local models with that of global model, along with measuring the difference between local models in each iteration, MOON achieves high efficiency in aggregating the global model.

Shanshan Jiang et al. proposed an original algorithm based on multi-head attention algorithm (M-FedAvg) in 2023 [9]. Past mainstream FL algorithm consider neither correlation between features, nor the data difference coming from the reasonable personalization of each client, leaving new possibilities. The new design tries to solve non-IID problems by enhancing the typical structure of FL by means of attention mechanisms in both ends. In the local level, a multi-head attention algorithm is introduced to learn the correlation between local features and improve the personalized degree of local parameters. In the global level, the researchers succeeded in combining an especially improved fusion framework of FL with the multi-head attention mechanism. Figure 2 shows how the attention layer is added to the complete neural network layers.

To be precise, the algorithm alters the traditional setting by adding new attention layers to the classic neural network adopted by local model. After passing through initial layers and getting the feature map of data, the algorithm connects each feature with a designated amount of parallel attention heads to obtain image features, which are transformed to calculate attention. While in the global model, the researchers designed a fitting structure combining personalized parameter with weighted average. The parameter is derived from the distance between pre-trained global model and current local models. The overall model is improved in performance of solving data heterogeneity as attention mechanism finds the correlation between each local clients and generates a robust model work well in different types of datasets.

The model creatively combined attention mechanism with local model training, enhancing the vital information of data and reducing the useless information. It could be further improved, however, by introducing attention mechanism to global model aggregation, since it only performs weighted averaging. Adding methods to effectively selecting method while abandoning some of the poorest model could be considered.

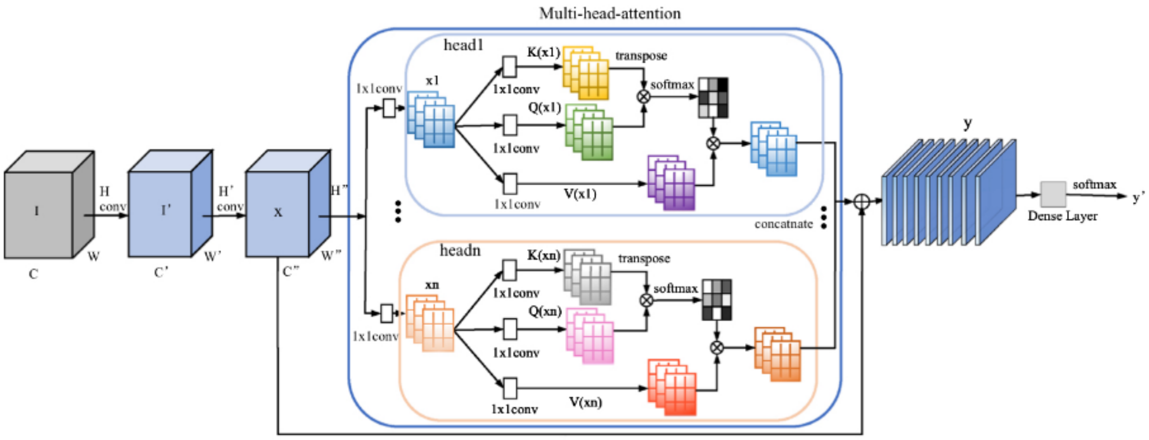

Figure 2. The Local model multi-head attention mechanism diagram[9]

2.3. *Methods designed for enhanced privacy protection*

In some of the applications of image classification, privacy is the major concern, especially in healthcare and auto-driving. Leakage of these information might cause serious legal problems and lead to enormous losses. Even if the formation of FL greatly decreases the possibilities of privacy leakage, attackers could still breach the system using multiple techniques, stealing the privacy information from the model uploading sequence [10]. In FL background, there are two types of attack forms: feature inference attacks and label inference attacks [11]. Therefore, new algorithms are urgently needed to shield the FL network from attack sources. Researchers had already proposed multiple adjustments to the training system. The paper here introduces the following methods: 1) Adaptive Obfuscation, 2) Differential Privacy.

Hanlin Gu et al. proposed a novel algorithm called FedPass combining FL structure with Adaptive Obfuscation (AO) in 2023 [11]. Comparing with the past methods suffering from limited privacy conservation and loss of information, the authors introduce adaptive design to find balance between obfuscation and performance. A series of randomly generated private passports are implemented in both local and global models to hinder attackers from accessing features of data. Researchers realized this process by adding an extra passport layer into the neural network, blurring the essential information. The encrypted parameters are the send to the central server to aggregate a global model, which is encoded again and is used to calculate and distribute parameters to each local models. The training ends when no further improvement could be made. Figure 3 depicts how passports are generated and distributed into models.

The new algorithm comes with multiple advantages. The obfuscation is trained in accordance with the optimization of model parameters and the less adverse way of encryption is chosen. Passports in global and local models prevent attackers from recovering data feature and inferring data labels as they suffer from a non-zero recovery error. The calculation process is efficient as well since the algorithm does not run any computationally intensive sequences. All of the advantages above make it a potential choice for training.

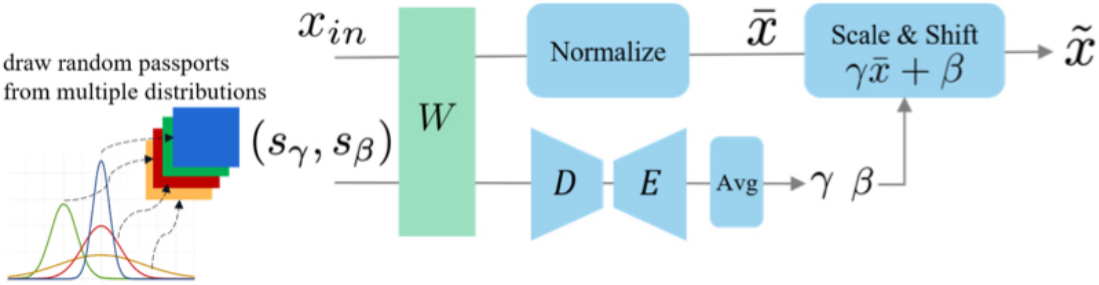

Figure 3. Realizing AO by adding passport layers into neural network layers[11]

Wei et al. proposed a new encryption algorithm based on differential privacy (DP) called Noising before Model Aggregation FL (NbAFL) in 2019 [12]. Adding artificial noise to local model parameters before sending the to central servers could help stop attackers from analyzing differences between parameters of each iteration to get private data. By systematically proving that the algorithm could satisfy DP with certain protection level and showing the convergence bound of the loss function under the properly defined variances of noise, researchers show the existence of the tradeoff between performance and protection ability and illustrate the optimal communication rounds to achieve a global model with that particular tradeoff.

One of the major advantages of the essay is that it gives complete proof on the convergence behavior of FL with privacy-preserving noise perturbation. It is cutting-edge as they fill up the empty of past essays. These rigorous analytical results will be the cornerstone of designs of FL algorithm with added noise in the future. Researchers also shows that by carefully selecting the particular noise level and the number of active local model, it is possible to reach similar results with the de-noised setting, manifesting the high potential of the model in the field of application.

Badih Ghazi introduced another improvement to FL structure by combining the method of randomly shuffling models with differential privacy [13]. The major difference between it and the AO is that, instead of adding noise, it blurs local models by using a conceptual “invisibility cloak” that shuffling local parameters, making local parameters almost identical to random noise while keeping zero distortion on the sum. In local models, parameters are reorganized with the help of “cloak” and are transmitted as noise to the central server, where they are decrypted for the global model aggregation. Figure 4 is provided to explain the overall principle.

One of the most vital advancements is that based on a series of past algorithms, the new method successfully further reduced the aggregation error and the amount of communication error which increase only polylogarithmically in n. It proved that the shuffled model design is a fertile middle ground between DP and multiple party computation. A few more problems are still to be solved for this algorithm. The protocol might shuffle models and gets exactly the same arrangement as before, failing to encrypt the parameters. Moreover, it is unclear how many messages are needed to achieve DP without the particular cost, providing further ground for improvements.

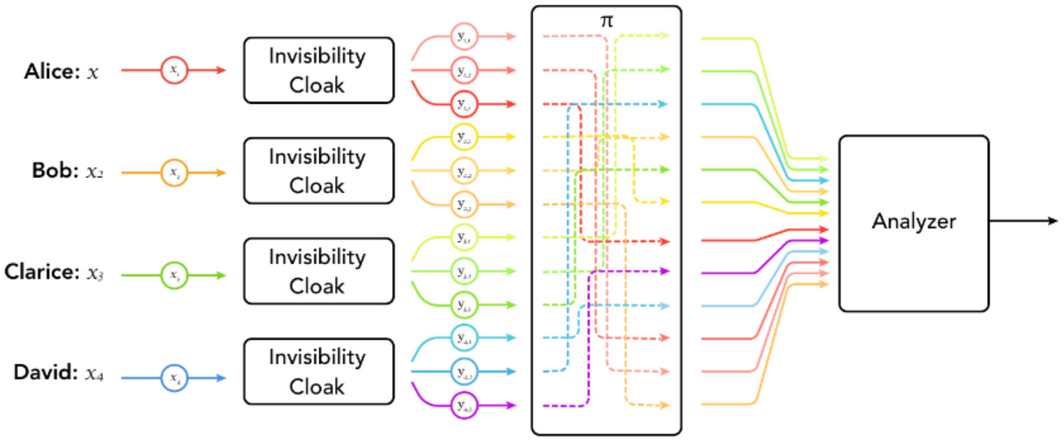

Figure 4. The principle if invisibility cloak and shuffling in the algorithm [13]

2.4. Methods solving communication burdens

A typical difficulty in the application of FL is the overwhelming communication cost. In a typical image classification application, communication problem rises as large training networks of FL might contains numerous devices. For instance, in some special situations like photo classification in mobile phone apps, programmers need to utilize millions of devices for local training process and final global aggregation.

Md Zarif Hossain et al. proposed an innovative improving scheme that combining the classic FL algorithm with an automatic hyper-parameter tuning method called AVO [14]. It is a multi-staged optimization method simulating the hunting process of African vultures, represented by a concurrent training process of hyper-parameter tuning. An initial population of vultures is created to represent solutions and venture through the problem space for the optimal one, changing their strategies based on the individual experience and interactions with other vultures. The best vulture’s final parameters are selected after reaching pre-defined iterations or a sufficient solution is calculated. Figure 5 displays the structure of FedAVO.

The method brings new possibilities as the selection of hyper-parameters is a relatively new research direction, filling up the empty in serval aspects of FL. Past algorithms focus on choosing the best hyper-parameters excessively, ignoring the heavy communication burden. AVO bring various advantages to FL solution. It is proved that AVO could effectively combine with multiple classical FL algorithms [14], greatly reducing communication rounds and improve accuracy.

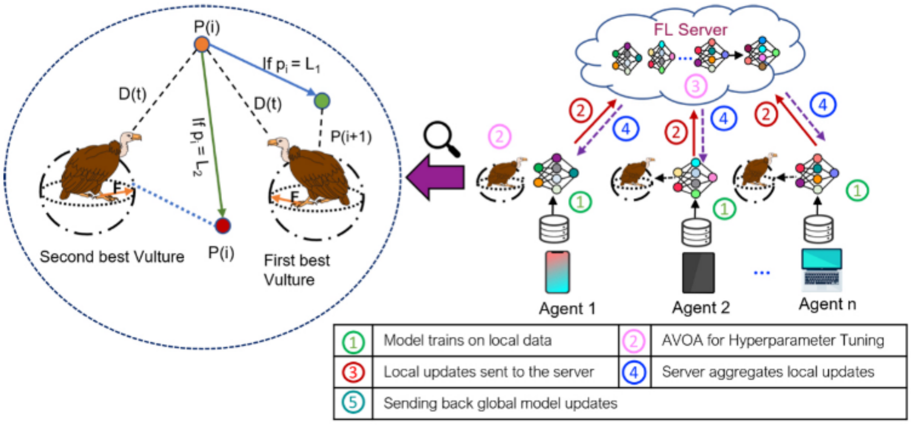
P(i)
D(t)
If p_i = L_1
D(t)
P(i+1)
If p_i = L_2
F
F
Second best Vulture
P(i)
First best Vulture
FL Server
Agent 1
Agent 2
...
Agent n
① Model trains on local data
② AVOA for Hyperparameter Tuning
③ Local updates sent to the server
④ Server aggregates local updates
⑤ Sending back global model updates

Figure 5. The system design and steps of AVO [14]

Chun-Chih Kuo et al. proposed an original method called Deep Gradient Compression with Global Momentum Fusion (DGCwGMF), which reduces communication overheads between local clients and the central server [15]. The researchers begin with analyzing two main communication overheads with the existing method and show that some methods based on momentum could bring extra communication overheads. Then an algorithm is proposed to minimize the communication overheads between local clients. The algorithm is designed as an improvement on the DGC. With GMF, the long-term momentum direction is kept while the parameter gradient is compressed, finding a configurable trade-off between the local gradient and the global momentum. DGCwGMF have showed that it provides decent performance while having 20.4% fewer communication overheads than DGC in the image classification tasks. The algorithm provides the best performance in some specific setting with low compression rates, making it a potential choice in the research field.

In 2022, Yuzhu Mao et al. designed a new algorithm based on the method called Adaptive Quantized Gradient (AQG) [16]. Compared with the original method of assign gradients to fixed quantization bits, the new method utilizes adaptive quantization, which changes the quantization level according to the update rounds of the local model. The rationale of AQG is that the inner precision selection standards employ the interior property of heterogeneity of local update to reduce the unneeded transmission cost, accelerating the communication speed.

AQG brings new possibilities to FL, as it outperforms current mainstream method in terms of the summation of transmission bits, achieving greater transmission reduction in non-IID settings while keeping desired convergence properties. It is also possible to combine it with pre-existing algorithms, bringing new vitality to the old frameworks of FL.

3. Performance analysis under different view points

The paper will analyze the performances of algorithms based on the different types of problems they aim to solve. All of these algorithms use typical image test set such as MNIST and CIFAR-10, which creates perfect preconditions for effective comparation between algorithms.

Since different types of datasets are adopted in the original papers and each algorithm has its own unique focus and hyper-parameter settings in the testing stage, not all algorithms will be shown in the table. The first comparison focuses on the accuracy. The table 1 is listed to compare the performance of algorithms on different datasets, whose data are collected from research essays. Only the best performance result shown in the essay will be given.

Table 1. The performance of algorithms on different datasets

MNIST		CIFAR-10	
Algorithm	Accuracy	Algorithm	Accuracy
pFedMe	95.6%	FedAvg	66.3%
M-FedAvg	98.6%	MOON	69.1%
NbAFL	87.8%	FedAVO	66.3%
FedAVO	99.7%	DGCwGMF	80.6%
FedAvg	98.6%	--	--

From the tables above, one could see that in the most ideal settings, algorithms could achieve high performance, getting relatively high accuracy results. To apply algorithms effectively, one needs to find the suitable one for the particular application, which requires further experiments in real life.

The results related to privacy protection are discussed and the main focus will be the protection effect. Researchers designed two attack methods for FedPass. In the Feature Reconstruction Attack, FedPass achieves the best performance among other algorithms, receiving a main task accuracy of 0.91. The model performance is almost lossless as the model recovers most of the information. In the tests of NbAFL, the model performance is related to the pre-defined privacy guarantees. By setting lower standards for privacy protection, it is feasible to get models with higher performance. It is possible to protect privacy effective as long as using accuracy as sacrifice is acceptable. There are a 15% accuracy gap between the model with the strongest protection level and the model with non-private approach. It is disappointing that no real experiment is run on the new structure of Shuffling Models. A series of strict proof are still provided to show the correctness of the new structure and the related possible improvement. It is still a fertile ground requiring further learning.

A series of results about accelerating communication are analyzed by mainly showing the reduction on the communication rounds. FedAVO outperformed multiple powerful algorithms in terms of accuracy while requiring less time for training. It means that to reach a particular performance level, FedAVO only requires fewer communication rounds. DGCwGMF achieves time saving target by compressing parameters. The results showed that the algorithm saves 20.4% of communication rounds compared with previous DGC method while maintaining the accuracy level. AQG performs well in treating both IID and non-IID data with effective accelerating results while reaching low loss values. Moreover, AQG benefits more from non-IID settings, which is consistent with expectations as AQG is designed to utilize the data heterogeneity.

4. Future developments

Federated Learning has become the promising land for researchers in the machine learning field. It has already revolutionized the way of large-scale distributed learning and created new possibilities for future model designers. In image classification field, as the demands of a robust model rises and privacy protection regulations tighten, FL offered compatible solutions and has achieved great accomplishment with its decentralized approach. In autonomous driving, healthcare and facial recognition, FL is able to utilize vast amount of data and create accurate classification models, helping doctors treating diseases and engineers improve driving algorithms while protecting privacy from possible leakage.

FL is expected to be further combined with classic machine learning method to further fulfill requirements of image classification. Techniques such as attention mechanisms, neural networks and adaptive client selection, with the help of FL structure, will solve problems like data heterogeneity and improve the final global model. Additionally, finding innovative way of integrating privacy protection method like Differential Privacy will further improve the protection ability of FL. Moreover, advancements in parameter communication protocols will reduce the pressure on network, which is critical for places lack of stable connections, creating precious chances for institutions like backward health institutions. By researching more advanced algorithms for model compression, it is possible to bring advantages of large models and rich datasets existed in developed regions to the most remote places.

FL is a relatively nascent field and researchers are in a critical time to shape and define future directions. More problems still await in real applications, bringing still more challenges. It is vital to unite broader research communities to tackle challenges, improve existing implementations and find brand-new methods for FL.

5. Conclusion

Federated learning has gained momentum since its release in 2017 and has become the common solutions for distributed machine learning. Image classification has significant function in various fields and provides great chances for FL since in some scenarios central training is costly or unacceptable. To promote image classification application in these fields and resolve related challenges, it is critical to research the effective combination of FL and image classification. In this article, an overview of the FL algorithms for image classification is provided. The paper analyzes challenges and problems exists in the process of applicating FL in image classification, introduces multiple algorithms by showing basic ideas and steps of them and discusses the pros and cons respectively. The performances of algorithms are compared based on the final model performance, showing the advances of FL in the real applications of image classification. Finally, the future development of applications is concluded in order to promote further innovation in FL and image classification. FL has proved itself to be a potential solution for treating images in fields like healthcare and autonomous driving and greater advances are expected.

References

[1] Wu M., Zhou J., Peng Y., Wang S., Zhang Y. Deep Learning for Image Classification: A Review. In: Su R., Zhang YD., Frangi AF. (Eds.), Proc. Int. Conf. Med. Imag. Comput.-Aid. Diagn. (MICAD). MICAD 2023, Lect. Notes Electr. Eng., Vol. 1166. Springer, Singapore.

[2] Konecný J., McMahan B., Ramage D., Richtárik P. Federated Optimization: Distributed Machine Learning for On-Device Intelligence. arXiv preprint arXiv:1610.02527 2023.

[3] Nampalle K., Singh P., Narayan U., Raman B. Vision Through the Veil: Differential Privacy in Federated Learning for Medical Image Classification. arXiv preprint arXiv:2306.17794, 2022.

[4] Mahlool D., Abed M. A Comprehensive Survey on Federated Learning: Concept and Applications. arXiv preprint arXiv:2201.09384 2022.

[5] Li T., Sahu A., Talwalkar A., Smith V. Federated Learning: Challenges, Methods, and Future Directions. IEEE Signal Process. Mag., 2020, 37(3), pp. 50-60.

[6] Dinh C., Tran N., Nguyen J. Personalized Federated Learning with Moreau Envelopes. In Adv. Neural Inf. Process. Syst. 2020, pp. 21394-21405.

[7] Li Q., He B., Song D. Model-Contrastive Federated Learning. In 2021 IEEE/CVF Conf. Comput. Vis. Pattern Recognit. 2021, pp. 10708-10717.

[8] McMahan B., Moore E., Ramage D., Hampson S., Aguera y Arcas BA. Communication-Efficient Learning of Deep Networks from Decentralized Data. In Int. Conf. Artif. Intell. Stat, 2021.

[9] Chen T., Kornblith S., Norouzi M., Hinton G. A Simple Framework for Contrastive Learning of Visual Representations. arXiv preprint arXiv:2002.05709, 2020.

[10] Jiang S., Lu M., Hu K., et al. Personalized Federated Learning Based on Multi-Head Attention Algorithm. Int. J. Mach. Learn. Cybern., 2023, 14, pp. 3783-3798.

[11] Liu B., Lv N., Guo YC., Li YW. Recent Advances on Federated Learning: A Systematic Survey. Neurocomputing, 2024, 597, pp. 128019.

[12] Gu H., Luo J., Kang Y., Fan L., Yang Q. FedPass: Privacy-Preserving Vertical Federated Deep Learning with Adaptive Obfuscation. In Int. Joint Conf. Artif. Intell 2023.

[13] Wei K., Li J., Ding M., Ma C., Yang H., Farokhi F., Jin S., Quek T., Poor HV. Federated Learning With Differential Privacy: Algorithms and Performance Analysis. IEEE Trans. Inf. Forens. Secur., 2020, 15, pp. 3454-3469.

[14] Ghazi B., Pagh R., Velingker A. Scalable and Differentially Private Distributed Aggregation in the Shuffled Model. arXiv preprint arXiv:1906.08320, 2019.

[15] Hossain M., Imteaj A. FedAVO: Improving Communication Efficiency in Federated Learning with African Vultures Optimizer. arXiv preprint arXiv:2305.01154, 2023.

[16] Kuo T. Lin C. Improving Federated Learning Communication Efficiency with Global Momentum Fusion for Gradient Compression Schemes. arXiv preprint arXiv:2211.09320, 2022.