


Federated Learning-Based Face Recognition: Methods, Challenges and Future Prospects

Xiaoying Yang ^a

Information and Computing Science, South China University of Technology, Guangzhou, China

Keywords: Federated Learning, Face Recognition, Privacy Preserving.

Abstract: With the rapid development of face recognition technology, federated learning has become a widely used method for face recognition due to its distributed collaboration and privacy-preserving properties. This paper systematically introduces the existing work on federated learning for face recognition to provide a referenceable overview for research in this area. In order to understand the function of federated learning in different application scenarios of face recognition, this paper discusses the implementation of different models in detail, dissects the representative models of federated learning in solving the three aspects of privacy-preserving improvement, gradient correction, and small-sample image recognition, and sorts out and explains the working principle of the models to elucidate the advantages of them in applications. Then, the current challenges of federated learning for face recognition are presented, pointing out that the current issues of data heterogeneity, applicability expansion, and interpretability still need to be further researched and improved, and possible solutions for the future are proposed.

1 INTRODUCTION


Facial recognition technology represents a prominent area of interest within the domain of computer vision research. As society evolves and scientific advancements are made, the integration of facial recognition with artificial intelligence methodologies has found extensive applications across multiple sectors. For example, security monitoring, cell phone unlocking and online identity verification. Today, artificial intelligence is evolving rapidly in many fields (Liu, 2023; Qiu, 2024; Zhao, 2023). Many researchers are dedicated to the field of face recognition studies, carrying out numerous related works and producing algorithms that can implement face recognition (e.g., Label Distribution Learning (Chen, 2020)). In order to better utilize the face dataset to achieve the desired functionality, it can be noted that there are quite a number of models that are not only concerned with the accuracy of model training, but also more concerned with the privacy issues of acquiring face data.

For traditional centralized data processing and model training approaches, it is often necessary to collect a large number of face datasets. However,

accessing and sharing face datasets has become exceptionally challenging due to increasing concerns about data privacy and legal restrictions (Woubie, 2024). Also, it may face the risk of personal privacy leakage during the training process.

In this situation, federated learning, an emerging learning paradigm for privacy protection, provides a viable way to address the challenge. It allows model training on distributed devices without sharing sensitive data. This approach not only protects individual privacy, but also allows the use of large-scale decentralized data to train more robust and accurate face recognition models. By assigning computational tasks and model parameters to multiple nodes for collaborative training, distributed learning not only improves the efficiency of the system, but also enhances the robustness and generalization ability of the model. As a result, the application of federated learning in the field of face recognition has gained much attention as one of the important tools to address privacy and security issues.

In current research, a number of researchers have chosen to use federated learning to solve and improve the face recognition problem. For instance, a new approach in federated learning to enhance face

^a <https://orcid.org/0009-0005-2234-6154>

recognition involves creating privacy-agnostic clusters. Meng et al. introduced the PrivacyFace framework, which significantly boosts the joint learning's face recognition performance by transmitting privacy-neutral auxiliary data among clients (Meng, 2022). Additionally, Niu et al. developed a framework called FedGC, aimed at enhancing privacy in federated learning for face recognition. They also introduced novel concepts for gradient correction: a softmax-based regularizer designed to adjust the gradient of class embeddings by accurately integrating cross-client gradient contributions (Niu, 2022).

In this paper, a comprehensive overview of these latest technologies in the field of face recognition is dedicated, because of the importance of these technologies in promoting privacy and the breakthroughs they have made in recent years. The content of this paper focuses not only on the technologies themselves, but also on the role they play in advancing the field of face recognition. By deeply studying and analyzing the principles, application scenarios, and achievements of these technologies, the latest technologies in the field of face recognition are presented and some possible challenges for future development are identified.

2 METHOD

2.1 Introduction of Federated Learning

Federated learning is a decentralized machine learning approach that allows each device to collaborate on building shared global models without directly sharing local data. Federated learning adheres to the two main ideas of local computation and model transfer (Zhang, 2021), which means that the data is kept at the edge client and bringing the model training to the edge. Such an approach reduces some of the system privacy risks and costs associated with traditional centralized machine learning approaches. The most common federated learning algorithm is Federated Averaging (FedAvg), which aggregates the updates using a weighted average (Solomon, 2024), effectively addresses the issues of protecting data privacy, ensuring security, and improving model performance.

The main workflow of federated learning shown in Figure 1 is as follows:

- 1) A global model is initialized either randomly or with a pre-trained model.
- 2) The aggregating server sends the current version of the global model to the available

clients.

- 3) The clients train the global model on their local data for a few iterations and send the model updates back.
- 4) The server aggregates the model updates to update the global model.
- 5) Steps 2-4 are repeated until the global model converges or achieves the desired performance.

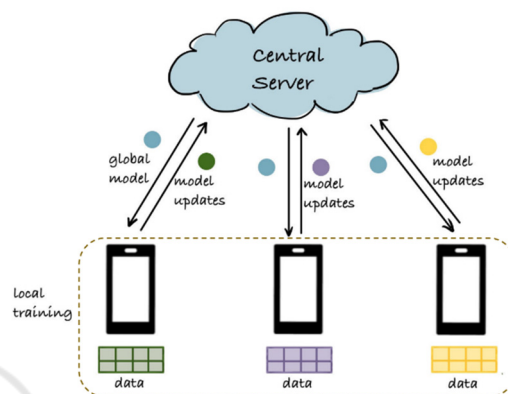


Figure 1: workflow of federated learning (Photo/Picture credit: Original).

2.2 Privacy Preserving Improvement

2.2.1 PrivacyFace

To address the privacy-utility paradox, Meng et al. proposed the PrivacyFace framework, which adds auxiliary information to the privacy agnostic information passed between clients. Firstly, a practical Differential Private Local Clustering (DPLC) mechanism is proposed to extract purified clusters from local class centers and broadcast the purified information of local classes to the world. Second, according to the principle of FedAvg, the server iteratively collects the feature extractor and privacy-independent clustering information from the client, then averages the parameters and sends them back to the client. In the local optimization phase, the consensus-aware loss sends information to each client to help each client train more discriminative features, which facilitates client alignment (Meng, 2022).

2.2.2 Privacy Projector for AIoT

To expedite the training convergence on AIoT devices, the server dispatches a pre-trained facial recognition model to each endpoint device. Upon receipt, each device supplements the model with a private projector. As training commences, every client modifies the entire model's parameters using

their local data to optimize the specified local objective. This processing is the key design of the privacy-preserving approach proposed by Ding et al. Specifically, the entire model now consists of a public module that really cares and a private module, and only the gradients (or model parameters) of the public module are uploaded to the server for averaging. In addition, the combined classification, also known as the private module, will act as a projector to help evaluate the quality of the extracted features (Ding, 2022).

2.3 Gradient Processing

2.3.1 FedGC

Niu et al. constructed a new training strategy to train models using privately dispersed non-IID face data and proposed the gradient-corrected joint averaging (FedGC) algorithm to address the local optimization problem due to the lack of a face-specific softmax-based loss function. They correct the gradient from a new perspective of backpropagation and introduce a cross-client gradient to ensure that the network is updated in the standard softmax direction. FedGC combines local optimization for softmax regular injection with inter-client optimization and is a privacy-preserving co-learning framework that ensures that each client has fully available private class embeddings (Niu, 2022).

2.3.2 Inverting Gradient Attack Combined with GAN Network

Liu et al. improved the traditional gradient leakage attack for face data recovery. First, for the face generation adversarial network (GAN) for face recognition task in a joint learning scenario, the image generation is restricted, and the virtual face image generated by the generative network is used instead of the initial input for constraints. In addition, round optimization is used, where each optimization is only for one of the images and one round of update is added when multiple images are updated. By these methods, the optimal gradient descent direction for attack network model training is specified to avoid falling into local minima, which improves the immunity of the attack and allows more photos to be recovered (Liu, 2021).

2.4 Few-Shot Face Recognition

2.4.1 FedAffect

FedAffect, proposed by Shome et al. is a small sample

size federated learning framework that learns from a small number of labeled FER images dispersed across user devices. In each round of learning, a small number of labeled private facial expression data samples are used to train local models, and then the weights of all the local models are aggregated to a central server to obtain a globally optimal model. In addition, since user devices are the source of a large amount of unlabeled data, a self-supervised approach based on collaborative learning is designed to isolate and update a network of feature extractors on unlabeled private facial data to learn robust and diverse facial representations (Shome, 2021).

2.4.2 FedFace

In the extreme case where there is only one recognized face image per mobile device, Aggarwal et al. propose a federated learning framework called FedFace to improve the performance of CosFace, a pre-trained face recognition system, in order to protect privacy by reducing face data aggregation. FedFace utilizes multiple face images on a client to learn an accurate generalized face recognition model, which has face images stored on each mobile client that are not shared with any device, to collaboratively utilize other human face data on client nodes. Class embeddings are initialized using an average feature initialization scheme, and an extended regularizer is used to ensure that class embeddings are well separated (Aggarwal, 2021).

3 DISCUSSIONS

Despite the many advantages of federated learning for privacy preservation and data security in face recognition, there are many challenges such as data heterogeneity, applicability, interpretability, communication and computational overhead.

In federated learning, there may be differences in data distribution between different devices or diverse data sources, which is known as data heterogeneity. In face recognition, performance and generalization of federated learning models can be impacted by data heterogeneity arising from various device-specific factors. These include changes in appearance, aging, pose, lighting intensity variations, and more broadly, facial expressions, missing data, use of cosmetics, and occlusions. Each of these elements can influence how effectively the model learns and generalizes across different settings. The need for future research on dynamic face heterogeneous data is also expressed in FedAffect by Shome et al. (Shome, 2021) and

FedFace by Aggarwal et al. (Aggarwal, 2021). To address such problems, the use of 3D sensors is a possible solution, and its recent development has been demonstrated to overcome the main limitations of 2D face recognition techniques, where the geometric information provided by 3D face data can significantly improve the accuracy of face recognition under unfavorable acquisition conditions. However, the lack of 3D face recognition databases hinders the development of deep learning-based approaches and requires further research in the future (Adjabi, 2020).

Also, face recognition has been successfully used in many user-collaborative applications, but a recognition without application-scenario limitations remains a worthy goal of the work. In practice, it is challenging to collect and label enough samples of the countless scenarios in the real world. A promising solution is to first learn generic models and then transfer them to application-specific scenarios (Wang, 2021). It is hoped that the applicability of federated learning face recognition can be addressed through, for example, transfer learning.

Moreover, federated learning suffers from a number of interpretability drawbacks since the models are trained on local devices rather than sending datasets to a centralized server. In federated learning systems, the results of the models are often difficult to understand and do not help to understand the contribution of each user and provide an objective opinion on incentive strategies within the federated learning system. In addition, it can also affect the ability of domain experts to understand the relationship between the data in key domains (e.g., healthcare and finance) and the final trained model (Liu, 2022). The Shapley value, which is used to identify which features are the main drivers of the model's predicted results, helps to improve the interpretability and credibility of the model, whereas it focuses on the vertical federated learning (Ghorbani, 2019). Alternatively, Gradient-weighted Class Activation Mapping (Grad-CAM), which is used to generate "visual explanations" for decisions from large-scale Convolutional Neural Network (CNN)-based models, is a solution that makes modeling potentially more transparent (Selvaraju, 2017).

4 CONCLUSIONS

In this work, a comprehensive review of federated learning for face recognition is presented. First, a brief workflow of federated learning is introduced.

Then, the improvement of federated learning for traditional face recognition techniques is shown through three sections: privacy improvement, gradient processing, and few-shot face recognition. Each section concentrates on the principles of algorithmic implementation of the model and has shown better results in their respective application areas. In addition, the current challenges of federated learning, which are the main obstacles to achieving more effective and widespread applications of face recognition are noted with possible solutions. However, this paper mainly focuses on the application of federated learning for face recognition aspects, and does not have a more in-depth study on specific algorithms for federated learning. In the future, it is hoped that this part can be added to form a more complete system.

REFERENCES

- Adjabi, I., Ouahabi, A., Benzaoui, A., & Taleb-Ahmed, A. 2020. Past, present, and future of face recognition: A review. *Electronics*, 9(8), 1188.
- Aggarwal, D., Zhou, J., & Jain, A. K. 2021. Fedface: Collaborative learning of face recognition model. In *2021 IEEE International Joint Conference on Biometrics (IJCB)* (pp. 1-8). IEEE.
- Chen, S., Wang, J., Chen, Y., Shi, Z., Geng, X., & Rui, Y. 2020. Label distribution learning on auxiliary label space graphs for facial expression recognition. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 13984-13993).
- Ding, Y., Wu, X., Li, Z., Wu, Z., Tan, S., Xu, Q., ... & Yang, Q. 2022. An efficient industrial federated learning framework for AIoT: a face recognition application. *arXiv preprint arXiv:2206.13398*.
- Ghorbani, A., & Zou, J. 2019. Data shapley: Equitable valuation of data for machine learning. In *International conference on machine learning* (pp. 2242-2251). PMLR.
- Liu, J., Huang, J., Zhou, Y., Li, X., Ji, S., Xiong, H., & Dou, D. 2022. From distributed machine learning to federated learning: A survey. *Knowledge and Information Systems*, 64(4), 885-917.
- Liu, Y., Xu, K., Cui, J., & Zheng, Q. 2021. Inverting Gradient Attack Combined with GAN Network in Federated Learning of Face Recognition. In *2021 IEEE 3rd International Conference on Frontiers Technology of Information and Computer (ICFTIC)* (pp. 317-325). IEEE.
- Liu, Y., Yang, H., & Wu, C. 2023. Unveiling patterns: A study on semi-supervised classification of strip surface defects. *IEEE Access*, 11, 119933-119946.
- Meng, Q., Zhou, F., Ren, H., Feng, T., Liu, G., & Lin, Y. 2022. Improving federated learning face recognition via

- privacy-agnostic clusters. arXiv preprint arXiv:2201.12467.
- Niu, Y., & Deng, W. 2022. Federated learning for face recognition with gradient correction. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 36, No. 2, pp. 1999-2007).
- Qiu, Y., Yang, Y., Lin, Z., Chen, P., Luo, Y., & Huang, W. 2020. Improved denoising autoencoder for maritime image denoising and semantic segmentation of USV. *China Communications*, 17(3), 46-57.
- Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. 2017. Grad-cam: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE international conference on computer vision (pp. 618-626).
- Shome, D., & Kar, T. 2021. FedAffect: Few-shot federated learning for facial expression recognition. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 4168-4175).
- Solomon, E., & Woubie, A. 2024. Federated Learning Method for Preserving Privacy in Face Recognition System. arXiv preprint arXiv:2403.05344.
- Wang, M., & Deng, W. 2021. Deep face recognition: A survey. *Neurocomputing*, 429, 215-244.
- Woubie, A., Solomon, E., & Attieh, J. 2024. Maintaining Privacy in Face Recognition using Federated Learning Method. *IEEE Access*.
- Zhang, C., Xie, Y., Bai, H., Yu, B., Li, W., & Gao, Y. 2021. A survey on federated learning. *Knowledge-Based Systems*, 216, 106775.
- Zhao, F., Yu, F., Trull, T., & Shang, Y. 2023. A new method using LLMs for keypoints generation in qualitative data analysis. In 2023 IEEE Conference on Artificial Intelligence (CAI) (pp. 333-334). IEEE.