CL-FedFR: Curriculum Learning for Federated Face Recognition

Devilliers Caleb Dube[1] [a], Çiğdem Eroğlu Erdem[2,3] [b] and Ömer Korçak[2] [c]

[1]*Department of Electrical and Electronics Engineering, Boğaziçi University, Istanbul, Turkey*

[2]*Department of Computer Engineering, Faculty of Engineering, Marmara University, Istanbul, Turkey*

[3]*Department of Electrical and Electronics Engineering, Özyeğin University, Istanbul, Turkey*

Keywords: Curriculum Learning, Deep Learning, Face Recognition, Federated Learning, Privacy.

Abstract: Face recognition (FR) has been significantly enhanced by the advent and continuous improvement of deep learning algorithms and accessibility of large datasets. However, privacy concerns raised by using and distributing face image datasets have emerged as a significant barrier to the deployment of centralized machine learning algorithms. Recently, federated learning (FL) has gained popularity since the private data at edge devices (clients) does not need to be shared to train a model. FL also continues to drive FR research toward decentralization. In this paper, we propose novel data-based and client-based curriculum learning (CL) approaches for federated FR intending to improve the performance of generic and client-specific personalized models. The data-based curriculum utilizes head pose angles as the difficulty measure and feeds the images from "easy" to "difficult" during training, which resembles the way humans learn. Client-based curriculum chooses "easy clients" based on performance during the initial rounds of training and includes more "difficult clients" at later rounds. To the best of our knowledge, this is the first paper to explore CL for FR in a FL setting. We evaluate the proposed algorithm on MS-Celeb-1M and IJB-C datasets and the results show an improved performance when CL is utilized during training.

1 INTRODUCTION

The utilization of face recognition (FR) technology has experienced a significant increase in the past few years, with most common uses in smartphones (Baqeel and Saeed, 2019), security (Kumar et al., 2019), access control (Kortli et al., 2020), surveillance (Jose et al., 2019), and border control (Damer et al., 2020). The fundamental concept behind FR is the identification of distinctive patterns in the facial features of an individual. These features include the distance between the eyes, nose, and mouth, as well as the structure of the cheekbones and jawline (Meena and Sharan, 2016; Elmahmudi and Ugail, 2018; Oloyede et al., 2020). Deep learning frameworks, particularly those based on convolutional neural networks (CNNs) have proven to be capable of learning these essential features from massive amounts of data with high generalization capability (Almabdy and Elrefaei, 2019). Hence, they dominate the state of the art techniques as shown in comprehensive surveys (Guo and Zhang, 2019; Taskiran et al., 2020).

Curriculum learning (CL) is a machine learning technique that resembles the learning steps used by humans which is based on beginning the learning process with easier data (or concepts) and gradually progressing to harder concepts (Soviany et al., 2022). In (Wang et al., 2021), the authors investigated whether all machine learning algorithms can really benefit from CL. They argued that although some applications may experience improved performance, the benefits of CL are not universal. Nonetheless, in some applications including computer vision and natural language processing, it has been shown that CL can improve the generalization ability of the models and also enhances the convergence rate (Jiang et al., 2014; Platanios et al., 2019; Nagatsuka et al., 2023; Sinha et al., 2020). In (Büyüktaş et al., 2021; Yang et al., 2023), CL algorithms for FR have been proposed and the results show that CL provides significant improvements in performance. However, their approaches suffer from privacy concerns arising from the use of centralized models.

Privacy concerns, coupled with power limitations, and network latency due to constant data transfer be-

[a] https://orcid.org/0000-0002-3871-9489
[b] https://orcid.org/0000-0002-9264-5652
[c] https://orcid.org/0000-0003-4419-556X

Dube, D., Erdem, Ç. and Korçak, Ö.
CL-FedFR: Curriculum Learning for Federated Face Recognition.
DOI: 10.5220/0012574000003660

In *Proceedings of the 19th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2024) - Volume 2: VISAPP*, pages 845-852
ISBN: 978-989-758-679-8; ISSN: 2184-4321

tween the central server and local device led to the inception of federated learning (FL) by Google in 2016 (AbdulRahman et al., 2020; Li et al., 2020a). FL is a machine learning technique that utilizes data stored on edge devices to train a model and shares only the model parameters with the server for aggregation (Li et al., 2020b; McMahan et al., 2017). In natural language processing, FL has been employed in mobile phones for next-word prediction (Hard et al., 2018; Stremmel and Singh, 2021), keyboard search suggestion (Yang et al., 2018), and emoji prediction (Ramaswamy et al., 2019). FL has garnered significant attention for its potential to revolutionize the healthcare sector because of the privacy requirement of medical records (Antunes et al., 2022). Brisimi et al. (Brisimi et al., 2018) utilize FL for hospitalization prediction using cardiovascular data and they conclude that their distributed technique provides faster convergence compared to centralized methods.

In computer vision, several decentralized FL techniques have been proposed (He et al., 2021; Kairouz et al., 2021; McMahan et al., 2017). However, these methods cannot be directly applied to FR because local clients have distinct classes, which calls for a model architecture with different parameters across the clients. To address this problem, Aggarwal et al. (Aggarwal et al., 2021) proposed FedFace, a FL model for FR. They only considered the scenario with a single identity (ID) per client, but in real situations, local devices could contain several IDs. Furthermore, edge devices share their ID proxies with the central server, resulting in privacy concerns as this information could be used to reconstruct the original images (Liu et al., 2022).

Recently, Vahidian et al. (Vahidian et al., 2023) performed a study of the benefits of ordered learning in a federated environment. They performed extensive experiments using object recognition datasets and concluded that CL can provide performance improvement on the global model. However, they did not perform evaluation on local models to investigate the impact of CL on local training. Also, as previous studies in (Wang et al., 2021) have shown that the advantages of CL cannot be generalized, the impact of CL across different fields has to be investigated. In our study, we seek to bridge the gap by analyzing the efficacy of CL on local models and by introducing CL to federated FR.

Liu et al. (Liu et al., 2022) presented FedFR framework, a FL based approach to address the drawbacks of prior research in federated FR. Their main objective was to enhance user privacy while improving both personalized and generic FR. Personalized FR is performed on the local clients whereas generic FR is performed on the global data. They introduced a decoupled feature customization module to collaboratively optimize personalized models. This module helps in obtaining an optimal personalized FR model for each of the local clients. However, since the face images in the datasets are randomly arranged, optimizing the objective function during training may not result in optimal convergence.

In this work, we combine the advantages of CL (Büyüktaş et al., 2021) and FL (Liu et al., 2022) for FR to further improve both the generic and personalized FR performance. We can summarize the contributions of this paper as follows:

- We introduce data-based CL to the FedFR framework based on head pose angles.
- We propose to apply client-based CL to FedFR during training.
- We also show that combining data-based and client-based curricula provides better generic FR performance than just using client-based CL.

2 BACKGROUND: FedFR

In this section, basic information about the FedFR framework (Liu et al., 2022) is provided, which is improved by using the proposed CL approaches as described in the next section.

In FedFR framework, there are C clients and a central server with the initial global FR model Θ_g^0 and the global class embeddings trained using a large global (public) dataset D_g, which contains N_g images from K_g IDs. This dataset is used to pretrain the initial global model and part of it is shared to the local clients during training to prevent overfitting and address the problems of heterogeneous clients as explained in (Zhao et al., 2018). Each client i initializes its local model as $\Theta_{l(i)}^0 = \Theta_g^0$ and has a local dataset containing $N_{l(i)}$ images from $K_{l(i)}$ distinct IDs, which is neither shared with other clients nor with the server. The goal is to optimize both the model Θ_g for generic face representation and $\Theta_{l(i)}$ for personalized client customization while preserving the privacy of the local client IDs.

Note that the ID distributions on each client are different, that is, the data is not independent and identically distributed (non-IID). At client i, the framework uses: i) a hard negative sampling stage to select the most critical data from the global dataset to reduce the computations, ii) contrastive regularization to limit the deviation of the local model from the global model, and iii) decoupled feature customization to learn a customized feature space optimized for

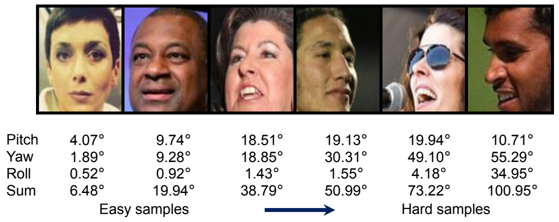

Figure 1: Face image samples from the MS-Celeb-1M dataset with their respective absolute head pose angles and the sum estimated using Openface 2.2.0 toolkit.

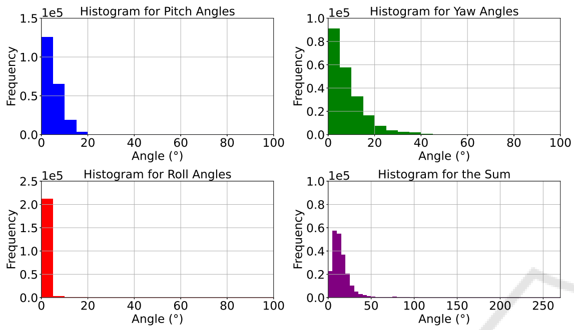

Figure 2: Histograms of absolute pitch, yaw, roll angles, and their sum for images in the selected subset of the MS-Celeb-1M dataset.

recognizing the IDs at the client.

At each communication round t, not only the model parameters, $\Theta^t_{l(i)}$ but also the learned class embeddings $\Phi^t_{l(i)}$ related to the K_g global IDs are sent to the server. It is important to note that the local class embeddings are not shared with the server as this information can be used to reconstruct the images, resulting in privacy concerns.

3 METHODOLOGY

In this section, we provide the details of the proposed data and client-based CL approaches for federated FR (CL-FedFR). We also give the details of the overall algorithm.

3.1 Data-Based Curriculum Design Based on Head Pose Angles

Face image datasets contain numerous images with diverse factors such as head pose, illumination, and resolution which affect the FR performance. Within this scope, Dutta et al. (Dutta et al., 2012) investigated the importance of image quality using a view based FR technique. They observed that the head pose had the most impact on the FR performance. Furthermore, they found out that the head pose determines the im-

Inputs: C clients each with $N_{l(i)}$ images from $K_{l(i)}$ disjoint identities; Number of local epochs E; Pre-trained global model Θ^0_g and class embeddings Φ^0_g.
Each client orders its own dataset into n subsets of increasing difficulty
$D_{l(i)} = \left\{D^j_{l(i)}\right\}^n_{j=1}$
Output: Optimal global model Θ^*_g
Server executes:
for each round $t = 0, \ldots, T-1$
  share Θ^t_g and Φ^t_g with clients
  Client training:
  for each client i **in parallel do**
    $\Theta^t_{l(i)}, \Phi^t_{l(i)} \leftarrow \text{ModelUpdateWithCL}(i, \Theta^t_g, \Phi^t_g)$
  $\Theta^{t+1}_g = \frac{1}{N} \sum_{i\in[C]} N_{l(i)} \cdot \Theta^t_{l(i)}$
  $\Phi^{t+1}_g = \frac{1}{N} \sum_{i\in[C]} N_{l(i)} \cdot \Phi^t_{l(i)}$
def ModelUpdateWithCL$\left(i, \Theta^t_{l(i)}, \Phi^t_{l(i)}\right)$**:**
  $D^{\text{train}}_{l(i)} = \emptyset$
  for $j = 1 : n$ **do**
    $D^{\text{train}}_{l(i)} = D^{\text{train}}_{l(i)} \cup D^j_{l(i)}$
    for $e = 1 : E$ **do**
      fine-tune$\left(\Theta^t_{l(i)}, \Phi^t_{l(i)}, D^{\text{train}}_{l(i)}\right)$
    end
    return $\Theta^t_{l(i)}, \Phi^t_{l(i)}$ to server
  end

Algorithm 1: Proposed data-based curriculum learning for federated face recognition.

pact that other image quality factors have on the FR performance. Therefore, recent works in CL for FR have used the head pose as their difficulty measure (Büyüktaş et al., 2021; Yang et al., 2023). Accordingly, we use the sum of absolute pitch, yaw, and roll head pose angles to order the data from easy to hard as shown in Figure 1.

We utilize a subset of the MS-Celeb-1M (Guo et al., 2016) dataset for training our model. For each ID, we use Openface 2.2.0 toolkit, an updated version of Openface 2.0 (Baltrusaitis et al., 2018), to estimate the head pose angles. Figure 2 shows the histograms of the absolute head pose angles and the sum. The absolute yaw angles show the most diversity whereas

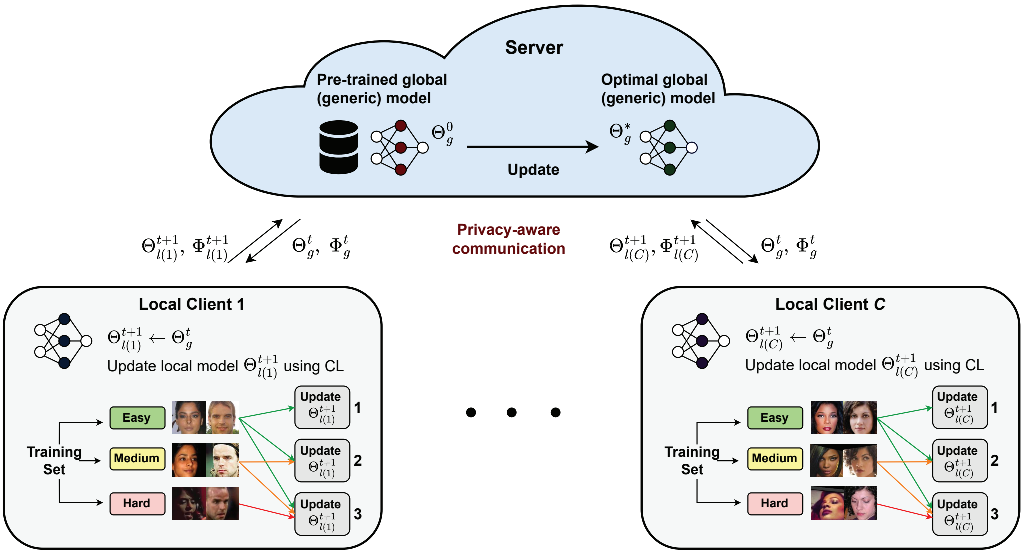

Figure 3: Proposed curriculum learning for federated face recognition framework. At each client, the first training is conducted with just the easy subsets of local datasets. Then, the experiments are repeated with the easy subset augmented with more difficult subsets using the optimal model of the previous training as the initial model for the subsequent training stages.

almost all absolute roll angles are below 10°. Just above 1% of the images have a sum of absolute head pose angles greater than 50°. We split the images at each client into different subsets of difficulty ranging from easy to hard based on the absolute sum of head pose angles. We experiment with different splitting strategies which are explained in Section 4.

3.2 Client-Based Curriculum Design Based on Performance

We further investigate the effect of a client-based split on the model performance. Firstly, we train the model for a few communication rounds and then perform personalized evaluation. The clients are then ordered based on their personalized evaluation results and fed to the FedFR model starting with the "easy" high performing top half and gradually introducing the "difficult" low performing ones.

3.3 Proposed Algorithm

In our setup, we propose to employ CL to train the FedFR model (Liu et al., 2022) as shown in Figure 3. The advantages of CL, such as the capability to enhance convergence and improve performance in particular scenarios, served as the inspiration to this approach. In the first part of the training process, the model is fine-tuned using only the easy subsets from all the local clients for the data-based curriculum and/or only the easy clients for the client-based curriculum. Then, the easy subset is augmented with the medium difficulty subset for the second part of the training. The optimal global model obtained from the first part of the training is used as the initial backbone model for the second part of the training. The training is repeated for each of the curriculum sets with gradual increase in difficulty until the entire dataset is used. The summary of our proposed data-based CL-FedFR approach is presented in Algorithm 1. Similar steps are followed for the client-based CL-FedFR approach but with clients as the difficulty measure instead of the images.

4 EXPERIMENTS

4.1 Experimental Setup

Similar to (Liu et al., 2022), we use a subset of the MS-Celeb-1M (Guo et al., 2016) dataset which consists of 10,000 IDs for training and evaluating the personalized models. We employ a 64-layer CNN architecture in (Liu et al., 2017) as the initial global model. In our FL setup, the number of communication rounds is $T = 20$, the number of local epochs is $E = 10$, and the learning rate is 0.001. The rest of the hyper-parameter settings are the same as in FedFR.

In order to make our experimental results comparable with FedFR, we similarly select 6,000 IDs from the MS-Celeb-1M subset and use it as the global dataset for pretraining the initial global model. Part of this global dataset is also shared with the local clients using the aforementioned hard negative sampling strategy. The global dataset consists of between 60 and 80 images per ID to give a total of $420,671$ images. For local training, we distribute the remaining $4,000$ IDs to $C = 40$ clients each with $K_l = 100$ distinct IDs. Each of these IDs contain between 50 and 60 images to give a total of $215,144$ images. In each local client, 40 images per ID are reserved for personalized evaluation in the final training stage, and the remaining images, together with the shared images from the global dataset are used for training the local models. We perform all the experiments on a server with Intel® Xeon(R) W-2255 CPU @ 3.70GHz × 20 and NVIDIA RTX A6000 graphics card.

We perform experiments under 5 different settings using the same MS-Celeb-1M subset for the final training stage. For example, in setting CL-FedFR-P2 defined in section 4.1.1, each local model is trained with the easy local dataset first. The model is then further trained on all the local images and the images obtained from the global dataset, starting with the global model obtained from the training with the easy subset as the initial model. Then, the reserved 40 images per ID are used for personalized evaluation after all the training stages have been completed. For generic evaluation, we use the IJB-C (Maze et al., 2018) dataset, an extension of the IJB-A, which contains $3,531$ IDs with about $138,000$ face images and $11,000$ face videos.

For the data-based curriculum, we first estimate the pitch, yaw, and roll angles using OpenFace 2.2.0 and then sort the images with respect to the sum of the absolute values of these angles in ascending order. Next, we segment this ordered dataset into different numbers of subsets of increasing difficulty. The aim of the various number of difficulty levels is to investigate the impact of a splitting criterion on the model performance. We first employ CL on just the local dataset and then apply CL on both the local and global datasets. Moreover, we perform a client-based split considering the performance of the clients after training for a few communication rounds. Finally, we combine data-based and client-based curricula. The description of the curriculum sets and their respective subsets are given in the following sections.

4.1.1 Applying CL on Local Dataset

We apply three different curriculum settings on the local dataset as described below:

- **CL-FedFR-P2:** CL-FedFR based on dual percentage-wise split.
  - *Easy subset*: First 50% of ordered local images.
  - *Hard subset*: Last 50% of ordered local images.
- **CL-FedFR-P3:** CL-FedFR based on ternary percentage-wise split.
  - *Easy subset*: First 33% of ordered local images.
  - *Medium subset*: Images in rank between 33% and 67%.
  - *Hard subset*: Last 33% of ordered local images.
- **CL-FedFR-P4:** CL-FedFR based on quadrant percentage-wise split.
  - *Easy subset*: First 25% of ordered local images.
  - *Medium subset 1*: Images in rank between 25% and 50%.
  - *Medium subset 2*: Images in rank between 50% and 75%.
  - *Hard subset*: Last 25% of ordered local images.

4.1.2 Applying CL on Local and Global Datasets

- **CL-FedFR-P3-G:** CL-FedFR based on ternary percentage-wise split of both local and global datasets.
  - *Easy subset*: First 33% of the ordered images per dataset.
  - *Medium subset*: Images in rank between 33% and 67%.
  - *Hard subset*: Last 33% of the ordered images per dataset.

4.1.3 Client-Based Curriculum

- **CL-FedFR-C:** CL-FedFR based on client performance. The model is trained for the first 5 rounds. Then, it is evaluated and the clients are sorted in descending order based on their personalized evaluation results. Thereafter, it is trained for 15 rounds using the easy subset. Finally, the easy subset is augmented with the hard subset and the model is trained for 20 rounds.
  - *Easy subset*: Top performing 20 clients.
  - *Hard subset*: Bottom performing 20 clients.

4.1.4 Client-Based Curriculum and CL on Local Dataset

- **CL-FedFR-P2-C:** CL-FedFR based on dual percentage-wise split of CL-FedFR-C subsets.
  - *Easy subset 1*: First 50% of the images of the ordered top performing 20 clients.
  - *Easy subset 2*: Last 50% of the images of the ordered top performing 20 clients.
  - *Hard subset 1*: First 50% of the images of the ordered bottom performing 20 clients.
  - *Hard subset 2*: Last 50% of the images of the ordered bottom performing 20 clients.

Table 1: Personalized and generic face verification and identification results given in % using 40 clients, each with 100 IDs in a federated setting. The best method and result for each evaluation protocol is in bold.

Method	Personalized Evaluation (MS-Celeb-1M)				Generic Evaluation (IJB-C)			
	Verification		Identification		Verification		Identification	
	1:1 TAR @ FAR		1:N TPIR @ FPIR		1:1 TAR @ FAR		1:N TPIR @ FPIR	
	$1e-6$	$1e-5$	$1e-5$	$1e-4$	$1e-5$	$1e-4$	$1e-2$	$1e-1$
FedFR (Liu et al., 2022)	88.32	95.46	95.17	97.94	77.60	85.21	73.60	81.27
Yu et al. (2020)	75.82	87.65	89.50	94.67	-	-	-	-
CL-FedFR-P2	90.53	96.19	96.46	98.34	78.00	85.57	**74.10**	**82.12**
CL-FedFR-P3	90.40	96.15	96.77	98.51	77.84	85.56	73.94	82.00
CL-FedFR-P4	83.81	96.12	79.74	83.32	78.11	85.61	73.86	82.08
CL-FedFR-P3-G	90.48	96.20	**96.90**	**98.52**	78.01	85.48	73.85	81.77
CL-FedFR-C	**90.65**	**96.31**	96.65	**98.52**	77.54	85.16	73.48	81.66
CL-FedFR-P2-C	90.13	96.05	96.21	98.37	77.56	85.31	73.70	81.64
AntiCL-FedFR-P2	77.76	87.69	79.91	92.00	**78.23**	**85.80**	74.03	82.02

4.1.5 Anti-CL on Local Dataset

- **AntiCL-FedFR-P2:** CL-FedFR based on dual percentage-wise split with training starting from "hard" to "easy" subset.
  - *Hard subset*: Last 50% of ordered local images.
  - *Easy subset*: First 50% of ordered local images.

4.2 Evaluation Protocols

We perform generic and personalized evaluation of the models, with the generic evaluation being done on the global model in the server and the personalized being performed on the local models in each of the clients. We follow the commonly used IJB-C evaluation protocol for generic face recognition by performing 1:1 face verification protocol and 1:N face identification protocol. We use the true acceptance rates (TAR) at various false acceptance rates (FAR) for 1:1 face verification protocol and true positive identification rates (TPIR) at different false positive identification rates (FPIR) for 1:N face identification protocol.

For personalized evaluation, we follow similar protocols as performed in FedFR for fair comparison. For 1:1 face verification protocol, we first determine a list of positive and negative pairs just as in the IJB-C protocol. Then in each of the local clients, we formulate authentic matches from local IDs and create imposter matches by pairing one local ID with a random ID from a different client. The reported TAR values are the average TAR values from the 40 clients. For 1:N face identification protocol, we simulate a login experience on a local client. The images of each ID are combined to form the gallery features and the probe features are the images from all the clients.

4.3 Face Verification and Identification Results

The personalized and generic FR results are presented in Table 1. The personalized evaluation is performed on each of the 40 local models, $\Theta_{l(i)}$ and the average results are reported. The generic evaluation is performed on the global model, Θ_g which is at the server.

We compare our personalized evaluation results with FedFR (Liu et al., 2022) and a personalized framework proposed by Yu et al. (Yu et al., 2020). In (Yu et al., 2020), they evaluated three local adaptation techniques for federated models: fine-tuning, multi-task learning, and knowledge distillation. Their best results which are reported in this paper were obtained using knowledge distillation technique. It can be seen that our approach enhances the model performance as all the best results were recorded after applying CL. For generic evaluation, we compare our approach with FedFR and similarly, a slight improvement in performance is observed.

In our CL approach, the best personalized FR results were generally obtained using CL-FedFR-C whereas using CL-FedFR-P4 offers the worst results. Despite a significant improvement on the personalized evaluation results, using CL-FedFR-C is not beneficial for generic FR. CL-FedFR-P2-C improves generic FR for CL-FedFR-C but provides slightly lower personalized evaluation results. Similarly, using CL-FedFR-P4 offers a notable improvement on the generic evaluation, however, a significant decrease in personalized evaluation in comparison to FedFR. This justifies the conclusion of Wang et al. (Wang et al., 2021) that although CL offers analytical benefits such as enhanced convergence, some curriculum designs and applications may not necessarily provide improved performance. Therefore, this shows that improved personalized performance does not directly imply improved generic performance.

We also conduct anti-CL experiments to further justify the benefits of CL-FedFR on both local and global data. The results from AntiCL-FedFR-P2 show a significant reduction in personalized FR, however, an improvement in generic FR. These results show a tradeoff between personalized and generic FR. Nevertheless, since there is a major performance drop in personalized FR, anti-CL cannot be selected as an optimal design. Consequently, research is directed toward designing a curriculum that offers optimal improvement on both local and global data.

CL-FedFR-P2, CL-FedFR-P3, and CL-FedFR-P3-G produced an average performance increase of 1.16%, 1.24%, and 1.30% for personalized evaluation, respectively, and 0.53%, 0.42%, and 0.36% for generic evaluation, respectively, in comparison to FedFR. The trend in these % increase values further show a trade-off in performance between generic and personalized performance.

CL-FedFR-P2 has two splits, CL-FedFR-P3-G has three splits but uses CL on the global dataset, and CL-FedFR-P3 has three splits and uses all the shared images from the global dataset for each split during training. Therefore, the training time of CL-FedFR-P3 is inherently the longest amongst these three designs and cannot be selected in favor of the other two designs. Moreover, since most of the curricula provide a notable improvement in personalized evaluation, the choice of the best curriculum design becomes biased toward one that significantly improves generic evaluation. As such, we choose CL-FedFR-P2 in favor of CL-FedFR-P3-G to be the optimal CL design in our approach.

5 CONCLUSION

In this paper, we proposed a novel CL for federated FR technique. We adopted the FedFR framework and applied CL with the objective of improving the personalized and generic FR. In our approach, we used a data-based curriculum based on head pose angles and a client-based curriculum based on the FR performance. In data-based CL, the training sets were arranged so that images with easy-to-recognize head poses were used first, followed by a gradual inclusion of those with difficult-to-recognize head poses. The experimental results using the MS-Celeb-1M and IJB-C datasets show improved model performance. While we can generally conclude that CL offers a notable benefit in federated FR, it is important to note that the choice of the curriculum has an impact on the performance. The future works in this research area can be directed toward identifying more discriminative ways of creating client-based curricula.

ACKNOWLEDGEMENTS

This research work was supported by the Scientific and Technological Research Council of Turkey (TÜBİTAK) under project EEAG-122E025.

REFERENCES

AbdulRahman, S., Tout, H., Ould-Slimane, H., Mourad, A., Talhi, C., and Guizani, M. (2020). A survey on federated learning: The journey from centralized to distributed on-site learning and beyond. *IEEE Internet of Things Journal*, 8(7):5476–5497.

Aggarwal, D., Zhou, J., and Jain, A. K. (2021). Fedface: Collaborative learning of face recognition model. In *2021 IEEE International Joint Conference on Biometrics (IJCB)*, pages 1–8. IEEE.

Almabdy, S. and Elrefaei, L. (2019). Deep convolutional neural network-based approaches for face recognition. *Applied Sciences*, 9(20):4397.

Antunes, R. S., André da Costa, C., Küderle, A., Yari, I. A., and Eskofier, B. (2022). Federated learning for healthcare: Systematic review and architecture proposal. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 13(4):1–23.

Baltrusaitis, T., Zadeh, A., Lim, Y. C., and Morency, L.-P. (2018). Openface 2.0: Facial behavior analysis toolkit. In *2018 13th IEEE international conference on automatic face & gesture recognition (FG 2018)*, pages 59–66. IEEE.

Baqeel, H. and Saeed, S. (2019). Face detection authentication on smartphones: End users usability assessment experiences. In *2019 International Conference on Computer and Information Sciences (ICCIS)*, pages 1–6.

Brisimi, T. S., Chen, R., Mela, T., Olshevsky, A., Paschalidis, I. C., and Shi, W. (2018). Federated learning of predictive models from federated electronic health records. *International journal of medical informatics*, 112:59–67.

Büyüktaş, B., Erdem, Ç. E., and Erdem, T. (2021). Curriculum learning for face recognition. In *2020 28th European Signal Processing Conference (EUSIPCO)*, pages 650–654. IEEE.

Damer, N., Grebe, J. H., Chen, C., Boutros, F., Kirchbuchner, F., and Kuijper, A. (2020). The effect of wearing a mask on face recognition performance: an exploratory study. In *2020 International Conference of the Biometrics Special Interest Group (BIOSIG)*, pages 1–6.

Dutta, A., Veldhuis, R., and Spreeuwers, L. (2012). The impact of image quality on the performance of face recognition. In *33rd WIC Symposium on Information Theory in the Benelux, Boekelo, The Netherlands*, pages 141–148.

Elmahmudi, A. and Ugail, H. (2018). Experiments on deep face recognition using partial faces. In *2018 Interna-*

tional Conference on Cyberworlds (CW), pages 357–362.

Guo, G. and Zhang, N. (2019). A survey on deep learning based face recognition. *Computer vision and image understanding*, 189:102805.

Guo, Y., Zhang, L., Hu, Y., He, X., and Gao, J. (2016). Ms-celeb-1m: A dataset and benchmark for large-scale face recognition. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part III 14*, pages 87–102. Springer.

Hard, A., Rao, K., Mathews, R., Ramaswamy, S., Beaufays, F., Augenstein, S., Eichner, H., Kiddon, C., and Ramage, D. (2018). Federated learning for mobile keyboard prediction. *arXiv preprint arXiv:1811.03604*.

He, C., Shah, A. D., Tang, Z., Sivashunmugam, D. F. N., Bhogaraju, K., Shimpi, M., Shen, L., Chu, X., Soltanolkotabi, M., and Avestimehr, S. (2021). Fedcv: a federated learning framework for diverse computer vision tasks. *arXiv preprint arXiv:2111.11066*.

Jiang, L., Meng, D., Yu, S.-I., Lan, Z., Shan, S., and Hauptmann, A. (2014). Self-paced learning with diversity. *Advances in neural information processing systems*, 27.

Jose, E., M., G., Haridas, M. T. P., and Supriya, M. (2019). Face recognition based surveillance system using facenet and mtcnn on jetson tx2. In *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*, pages 608–613.

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., Bonawitz, K., Charles, Z., Cormode, G., Cummings, R., et al. (2021). Advances and open problems in federated learning. *Foundations and Trends® in Machine Learning*, 14(1–2):1–210.

Kortli, Y., Jridi, M., Al Falou, A., and Atri, M. (2020). Face recognition systems: A survey. *Sensors*, 20(2):342.

Kumar, A., Kumar, P. S., and Agarwal, R. (2019). A face recognition method in the iot for security appliances in smart homes, offices and cities. In *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*, pages 964–968.

Li, L., Fan, Y., Tse, M., and Lin, K.-Y. (2020a). A review of applications in federated learning. *Computers & Industrial Engineering*, 149:106854.

Li, T., Sahu, A. K., Talwalkar, A., and Smith, V. (2020b). Federated learning: Challenges, methods, and future directions. *IEEE signal processing magazine*, 37(3):50–60.

Liu, C.-T., Wang, C.-Y., Chien, S.-Y., and Lai, S.-H. (2022). Fedfr: Joint optimization federated framework for generic and personalized face recognition. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 1656–1664.

Liu, W., Wen, Y., Yu, Z., Li, M., Raj, B., and Song, L. (2017). Sphereface: Deep hypersphere embedding for face recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 212–220.

Maze, B., Adams, J., Duncan, J. A., Kalka, N., Miller, T., Otto, C., Jain, A. K., Niggel, W. T., Anderson, J., Cheney, J., et al. (2018). Iarpa janus benchmark-c: Face dataset and protocol. In *2018 international conference on biometrics (ICB)*, pages 158–165. IEEE.

McMahan, B., Moore, E., Ramage, D., Hampson, S., and y Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*, pages 1273–1282. PMLR.

Meena, D. and Sharan, R. (2016). An approach to face detection and recognition. In *2016 International Conference on Recent Advances and Innovations in Engineering (ICRAIE)*, pages 1–6.

Nagatsuka, K., Broni-Bediako, C., and Atsumi, M. (2023). Length-based curriculum learning for efficient pretraining of language models. *New Generation Computing*, 41(1):109–134.

Oloyede, M. O., Hancke, G. P., and Myburgh, H. C. (2020). A review on face recognition systems: recent approaches and challenges. *Multimedia Tools and Applications*, 79:27891–27922.

Platanios, E. A., Stretcu, O., Neubig, G., Poczós, B., and Mitchell, T. (2019). Competence-based curriculum learning for neural machine translation. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 1162–1172.

Ramaswamy, S., Mathews, R., Rao, K., and Beaufays, F. (2019). Federated learning for emoji prediction in a mobile keyboard. *arXiv preprint arXiv:1906.04329*.

Sinha, S., Garg, A., and Larochelle, H. (2020). Curriculum by smoothing. *Advances in Neural Information Processing Systems*, 33:21653–21664.

Soviany, P., Ionescu, R. T., Rota, P., and Sebe, N. (2022). Curriculum learning: A survey. *International Journal of Computer Vision*, 130(6):1526–1565.

Stremmel, J. and Singh, A. (2021). Pretraining federated text models for next word prediction. In *Advances in Information and Communication: Proceedings of the 2021 Future of Information and Communication Conference (FICC), Volume 2*, pages 477–488. Springer.

Taskiran, M., Kahraman, N., and Erdem, C. E. (2020). Face recognition: Past, present and future (a review). *Digital Signal Processing*, 106:102809.

Vahidian, S., Kadaveru, S., Baek, W., Wang, W., Kungurtsev, V., Chen, C., Shah, M., and Lin, B. (2023). When do curricula work in federated learning? In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5084–5094.

Wang, X., Chen, Y., and Zhu, W. (2021). A survey on curriculum learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(9):4555–4576.

Yang, J., Wang, Z., Huang, B., Xiao, J., Liang, C., Han, Z., and Zou, H. (2023). Headpose-softmax: Head pose adaptive curriculum learning loss for deep face recognition. *Pattern Recognition*, 140:109552.

Yang, T., Andrew, G., Eichner, H., Sun, H., Li, W., Kong, N., Ramage, D., and Beaufays, F. (2018). Applied federated learning: Improving google keyboard query suggestions. *arXiv preprint arXiv:1812.02903*.

Yu, T., Bagdasaryan, E., and Shmatikov, V. (2020). Salvaging federated learning by local adaptation. *arXiv preprint arXiv:2002.04758*.

Zhao, Y., Li, M., Lai, L., Suda, N., Civin, D., and Chandra, V. (2018). Federated learning with non-iid data. *arXiv preprint arXiv:1806.00582*.