Decentralised Learning from Independent Multi-Domain Labels for Person Re-Identification

Guile Wu and Shaogang Gong
Queen Mary University of London
{guile.wu, s.gong}@qmul.ac.uk

***Abstract*—Deep learning has been successful for many computer vision tasks due to the availability of shared and centralised large-scale training data. However, increasing awareness of privacy concerns poses new challenges to deep learning, especially for human subject related recognition such as person re-identification (Re-ID). In this work, we solve the Re-ID problem by decentralised learning from non-shared private training data distributed at multiple user sites of independent multi-domain label spaces. We propose a novel paradigm called Federated Person Re-Identification (FedReID) to construct a generalisable global model (a central server) by simultaneously learning with multiple privacy-preserved local models (local clients). Specifically, each local client receives global model updates from the server and trains a local model using its local data independent from all the other clients. Then, the central server aggregates transferrable local model updates to construct a generalisable global feature embedding model without accessing local data so to preserve local privacy. This client-server collaborative learning process is iteratively performed under privacy control, enabling FedReID to realise decentralised learning without sharing distributed data nor collecting any centralised data. Extensive experiments on ten Re-ID benchmarks show that FedReID achieves compelling generalisation performance beyond any locally trained models without using shared training data, whilst inherently protects the privacy of each local client. This is uniquely advantageous over contemporary Re-ID methods.**

I. Introduction

In recent years, deep neural network learning has achieved incredible success in many computer vision tasks. However, it relies heavily upon the assumption that a large volume of data can be collected from source domains and stored on a centralised database for model training. Despite the current significant focus on centralised data centres to facilitate big data machine learning drawing from shared data collections, the world is moving increasingly towards localised and private distributed data analysis at-the-edge. This differs inherently from the current assumption of ever-increasing availability of centralised data and poses new challenges to deep learning, especially for human subject related recognition such as person re-identification [1].

Person re-identification (Re-ID) on urban streets at city-wide scales is useful in smart city design (e.g. population flow management) and for public safety (e.g. find a missing person) [2]–[4]. Existing Re-ID methods follow either supervised learning by collecting large-scale datasets for model training [2], [5] or unsupervised learning by assembling both labelled source domain data for model initialisation and unlabelled target domain data for model fine-tuning [6], [7].

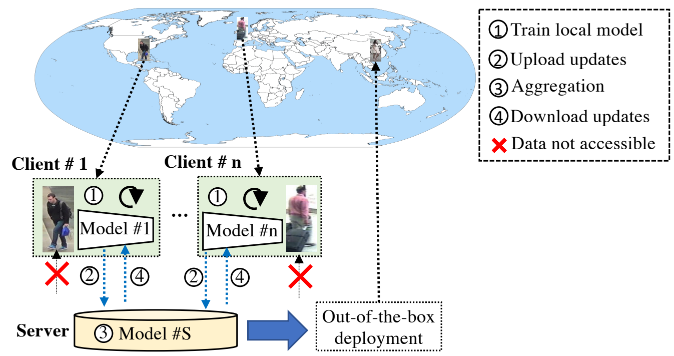


Fig. 1. An illustration of Federated Person Re-Identification (FedReID) for decentralised learning. A client refers to a user site with a local model and private local data, while a server refers to a centralised global model. Each client uses private local data to train a local model, while a server aggregates local model updates to construct a global model without accessing local data. This client-server collaborative learning process is iteratively performed to learn a global feature embedding model for out-of-the-box deployment with privacy protection.

Although these methods have achieved promising results, they are based on a *centralised learning* paradigm, which is inherently flawed when source datasets cannot be shared in a centralised training protocol due to privacy protection. This requires a new Re-ID paradigm for learning a generalisable global model with *distributed collections of non-shared data from independent multi-domain label spaces*.

In this work, we propose a fundamentally novel paradigm called Federated Person Re-Identification (FedReID) for decentralised model learning from distributed non-sharing data of independent label spaces. We construct a generalisable global Re-ID model (a centralised server) by distributed collaborative learning of multiple local models (localised and private clients) without sharing local training data nor collecting any centralised data. As illustrated in Fig. 1, different cities around the world can play the roles of localised clients. Each client receives global model updates from the central server and trains a local model using its own set of private non-shared data. Then, the central server aggregates local model updates to construct a generalisable model without accessing local data. This client-server collaborative learning process is iteratively performed, enabling FedReID to learn a generalisable global model from decentralised data with privacy protection. For deployment, the generalisable global Re-ID model from the server can be deployed directly without using additional centralised data for fine-tuning.

Our *contributions* are: **(1)** For the first time, we introduce decentralised model learning from distributed non-sharing data of independent multi-domain labels for person Re-ID. This study can potentially benefit other computer vision tasks that require decentralised model learning on distributed non-sharing data with privacy protection. **(2)** We propose a new paradigm called Federated Person Re-Identification (FedReID). Our approach explores the principle of federated learning [8], but is uniquely designed for decentralised Re-ID by reformulating the iterative client-server collaboration mechanism. In each local client, in addition to a local client model which consists of a feature embedding network for visual feature extraction and a mapping network for classification, we further use a local expert to regularise the training process of the local client model. **(3)** Extensive experiments on 10 Re-ID benchmarks show that FedReID can both protect local data privacy and achieve compelling generalisation performance, which is uniquely advantageous over contemporary Re-ID methods that assume shared centralised training data without privacy protection.

II. Related Work

Person Re-Identification. Learning generic feature representations is attractive for Re-ID real-world applications across domains. Conventional supervised Re-ID [2], [5], [9] relies heavily on centralised labelled training data in each target domain, whilst cross-domain unsupervised Re-ID [6], [7] relies on the availability of centralised training data from source domains for model initialisation and unlabelled data from target domains for model fine-tuning, so they are impractical for out-of-the-box deployments. More importantly, the centralised learning paradigm may not be feasible in practice when training data cannot be shared to a centralised training process due to privacy restrictions. Recently, domain generalised Re-ID is proposed to learn a generic feature embedding model by learning with a domain-invariant mapping network [10] or domain-specific knowledge selection [11]. However, these methods still require a centralised training process by assembling a large pool of data from multi-domain datasets. Different from all existing Re-ID methods, our FedReID has a fundamentally new decentralised learning paradigm for optimising a generalised Re-ID model through collaborative learning by communicating knowledge among the central server model and the local client models. Each client learns independently on distributed local private data, while the server uses local model updates to construct a global model without accessing local data nor collecting any centralised data, so FedReID embraces inherently privacy protection.

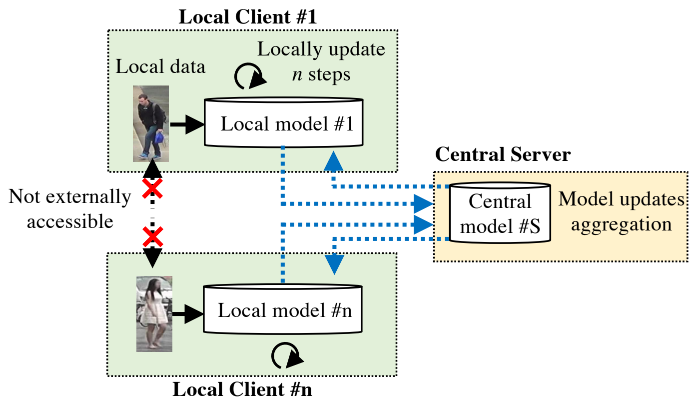
Federated Learning. Federated learning [8], [12]–[14] is a recently proposed machine learning technique that allows local users to collaboratively train a centralised model without sharing local data. Conventional federated learning aims at learning a shared model with decentralised data for the same class label space and reducing communication cost. For example, in [12], McMahan *et al.* introduced Federated Stochastic Gradient Descent (FedSGD) and Federated Average (FedAVG) to iteratively aggregate a shared model by averaging local updates. Our FedReID shares the merit of federated learning but requires a fundamentally different formulation to facilitate the generalisation of a global model for Re-ID. In person Re-ID, each local domain is independent (non-overlapping) from the other domains with totally different person populations (ID space) from different locations/cities, resulting in discrepancies in ID space and context. Thus, we need to learn simultaneously the non-sharing local knowledge in each local client and the latently shared generalisable knowledge in the central server. To this end, in FedReID, each client consists of a feature embedding network for visual feature extraction and a mapping network for classification which is decoupled from the central aggregation process, while the server constructs a generalisable global feature embedding model using updates of local models. Besides, in each local client, we additionally use a local expert to regularise the training process of the local client model to improve the generalisation performance.

Distributed Deep Learning. FedReID differs significantly from distributed deep learning [15]–[17]. Distributed deep learning aims at training very large scale deep networks (over billions of parameters) using massive hardware involving tens of thousands of CPU/GPU cores with parallel distributed computation (either model parallelism or data parallelism), with shared large training data. In contrast, FedReID considers the problem of optimising a generalisable global model by asynchronous knowledge aggregation from multi-domain locally learned models without centrally sharing training data.

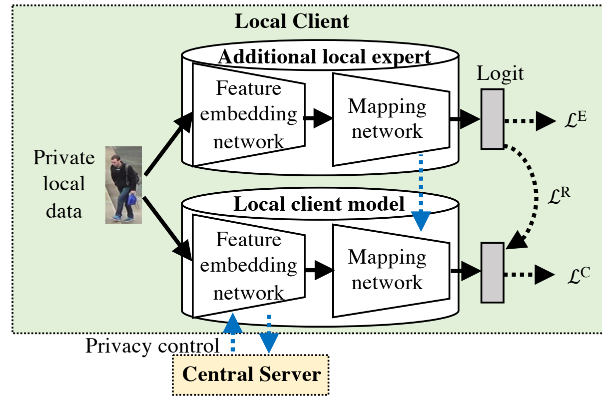
Private Deep Learning. Private deep learning [18], [19] aims at constructing privacy-preserving models and preventing the model from inverse attack [20], [21]. A popular solution [19] is to use knowledge distillation to transfer private knowledge from multiple teacher ensembles or a cumbersome teacher model to a public student model with restricted distillation on training data. In contrast, FedReID does not use any centralised training data (labelled or unlabelled) for model aggregation. Privacy is implemented intrinsically in FedReID by decentralised model training through iterative client-server collaborative learning by asynchronous knowledge aggregation, without central (server) data sharing in model updates.

III. Methodology

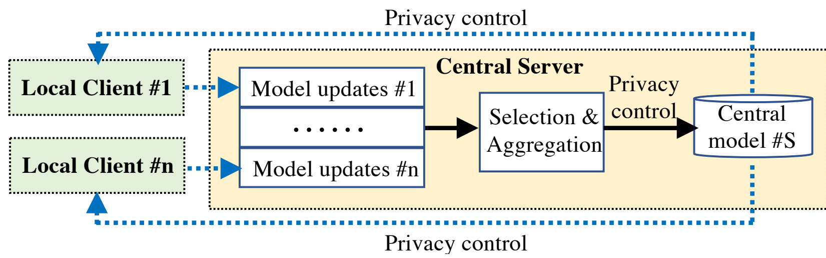
Overview. In this work, we investigate decentralised person Re-ID, a new problem to Re-ID which aims at optimising a generalised model via decentralised learning from independent multi-domain label spaces without assembling local private data. Suppose there are N private datasets captured from different locations that cannot be shared for model training due to privacy protection, *i.e.* there are N localised clients. As shown in Fig. 2(a), each client updates a local model with t_{max} steps separately using its own private data and uploads the model updates to a centralised server. The central server aggregates local model updates to construct a global model and transmits global model updates to each client. This client-server collaborative learning process is iteratively



(a) Iterative client-server collaborative learning.



(b) Local client model optimisation.



(c) Central server model update.

Fig. 2. An overview of the proposed Federated Person Re-Identification (FedReID).

processed, enabling FedReID to learn from decentralised data with privacy protection.

As shown in Fig. 2(b), there are a local client model and a local expert, which are trained together to improve the performance of the client. Specifically, in the i-th client model ($i \in N$), we construct a feature embedding network $\phi(\omega^f_{i,t,k}; x)$ to extract feature representations $\mathcal{V}_i = \{v_{i,j}\}^L_{j=1}$ of person images (there are L person images and I_i identities) in a local dataset $\mathcal{X}_i = \{x_{i,j}\}^L_{j=1}$:

$$v_{i,j} = \phi(\omega^f_{i,t,k}; x_{i,j}) \tag{1}$$

where $\omega^f_{i,t,k}$ are model parameters the i-th feature embedding network at the t-th local step at the k-th global communication epoch. And then, we employ a mapping network $\delta(\omega^c_{i,t,k}; v)$ for classification:

$$d_{i,j} = \delta(\omega^c_{i,t,k}; v_{i,j}) \tag{2}$$

where $d_{i,j}$ is the logit of the j-th sample in the i-th local client, $\omega^c_{i,t,k}$ are model parameters of the i-th mapping network at the t-th local step at the k-th global epoch. Meanwhile, we use an additional local expert $(\{\phi(\omega^{Ef}_{i,t,k}; x), \delta(\omega^{Ec}_{i,t,k}; v)\})$ which is learned with local knowledge of each client (domain) and helps the updated local client model to learn richer knowledge [1]. Thus, the optimisation objective $\mathcal{L}_i$ of the i-th client is formulated as:

$$\mathcal{L}_i = \mathcal{L}^C_i + \mathcal{L}^E_i + \mathcal{L}^R_i \tag{3}$$

[1]Since the local expert will not be used for the bidirectional client-server knowledge communication, we use "local model" to refer to a local client model which is used for the client-server collaboration and use "local expert" to refer to the additional local expert in each client.

where $\mathcal{L}^C_i$ is the identity classification loss of the i-th local client model, $\mathcal{L}^E_i$ is the identity classification loss of the i-th local expert, and $\mathcal{L}^R_i$ is the local expert regularisation from the i-th local expert to the i-th local client model.

As shown in Fig. 2(c), the central server does not use any centralised data for model optimisation. Instead, it selects and aggregates model updates from local clients to construct a server model $\sigma(\theta_k, x)$ without accessing local private data, where θ_k are model parameters of the server model at the k-th global epoch.

In deployment, the global feature embedding model in the central server is directly used for Re-ID matching with a generic distance metric (*e.g.* $L2$ distance).

Client-Server Iterative Updates. An intuitive idea for implementing decentralised learning from multiple user sites is to average multiple trained local models in parameter space to generate a global model. However, this could lead to an arbitrarily bad model [22]. Recent research in federated learning [8], [12] shows that local client models and a central server model can be iteratively updated for distributed model learning. Suppose the i-th client is optimised using SGD with a learning rate η, then its model parameters $\omega_{i,t+1,k}$ at the $(t+1)$-th local step are updated as:

$$\omega_{i,t+1,k} = \omega_{i,t,k} - \eta \nabla \mathcal{G}_{i,t+1,k} \tag{4}$$

where $\nabla \mathcal{G}_{i,t+1,k}$ is the set of average gradients of the i-th client at the $(t+1)$-th local step at the k-th global epoch. After t_{max} steps for local model updates in the clients, the server *randomly selects C-fraction* ($C \in [0, 1]$) local clients

N_C (here N_C is the set of selected clients) for the server model parameters θ_k aggregation:

$$\theta_{k+1} = \frac{1}{\lceil C \cdot N \rceil} \sum_{m \in N_C} \omega_{m,t_{max},k} \tag{5}$$

where $1 \leq \lceil C \cdot N \rceil \leq N$, $\omega_{m,t_{max},k}$ are model parameters of the $m\text{-}th$ client at the $t_{max}\text{-}th$ local step of the $k\text{-}th$ global epoch. Then, in turn, each client receives θ_k to update the local model:

$$\omega_{i,0,k+1} = \theta_{k+1} \tag{6}$$

where $\omega_{i,0,k+1}$ are the model parameters of the $i\text{-}th$ client at the initial (t=0) step of the $k\text{-}th$ global epoch. In this way, the local clients and the server are iteratively updated for k_{max} global epochs, and finally we can obtain a global model in the central server for deployment.

FedReID Client-Server Collaboration. In conventional federated learning, all model parameters in the selected client models (including feature embedding layers and classification layers) are used to update the centralised server model (Eq. (5)). However, in decentralised Re-ID, aggregating all model parameters can lead to performance degradation in both local and global models, because each local dataset is usually captured in different locations where person ID space and context are non-overlapping. To optimise a centralised model across different domains, we reformulate federated learning to simultaneously consider the non-sharing local knowledge in each local client and the latently shared generalisable knowledge in the central server.

Specifically, we decouple $\omega^c_{i,t,k}$ (the mapping network) from the aggregation in Eqs. (5) and (6) to preserve local classification knowledge in each client, and aggregate $\omega^f_{i,t,k}$ (feature embedding network) to construct a generalisable feature embedding model for deployment. Starting the feature embedding network of each local client model from the same initialisation, we can accumulate updates of multiple local feature embedding networks to find wider optima in the parameter space of a global model. Thus, in each local client, Eq. (6) is reformulated as:

$$\{\omega^f_{i,0,k+1}, \omega^c_{i,0,k+1}\} = \{\theta^f_{k+1}, \omega^c_{i,t_{max},k}\} \tag{7}$$

Since local data in different clients are from different domains, $\omega^c_{i,t,k}$ in each client corresponds to classification knowledge for different domains. Thus, in the central server, $\omega^c_{i,t,k}$ does not need to be aggregated. Besides, since the feature embedding network of each local client starts from the same initialisation (Eq. (7)), accumulating local gradient updates of feature embedding networks corresponds to aggregating model parameters in the feature embedding space, therefore, Eq. (5) can be formulated as:

$$\begin{aligned} \theta^f_{k+1} &= \theta^f_k - \frac{\eta}{\lceil C \cdot N \rceil} \sum_{m \in N_C} \sum_{t=1}^{t_{max}} \nabla \mathcal{G}^f_{m,t,k} \\ &= \frac{1}{\lceil C \cdot N \rceil} \sum_{m \in N_C} (\theta^f_k - \eta \sum_{t=1}^{t_{max}} \nabla \mathcal{G}^f_{m,t,k}) \\ &= \frac{1}{\lceil C \cdot N \rceil} \sum_{m \in N_C} \omega^f_{m,t_{max},k} \end{aligned} \tag{8}$$

where θ^f_k is model parameters of the feature embedding network of the central server model, $\nabla \mathcal{G}^f_{m,t,k}$ is the set of average gradient updates of the $m\text{-}th$ local feature embedding network at the $k\text{-}th$ global epoch.

Optimisation Objective. In FedReID, the central server does not use any centralised data for model fine-tuning, so its optimisation process is the selection and aggregation process as formulated in Eq. (8). In each local client, as shown in Fig. 2(b) and Eq. (7), the local client model receives global model updates from a central model. Then, we use a cross-entropy loss to learn classification knowledge:

$$\mathcal{L}^C_i = -\sum_{j=1}^{I_i} y_{i,j} log \frac{exp(d_{i,j})}{\sum_{b=1}^{I_i} exp(d_{i,b})} \tag{9}$$

where $y_{i,j}$ is the ground-truth label. To further improve generalisation of the local client model, we use a local expert to help the local client model to learn richer knowledge via knowledge distillation [23], which also potentially facilitates the aggregation in the global model. Specifically, the local expert with model parameters $\{\omega^{Ef}_{i,0,k+1}, \omega^{Ec}_{i,0,k+1}\}$ is initialised with Eq. (10) and also optimised with a cross-entropy loss with Eq. (11):

$$\{\omega^{Ef}_{i,0,k+1}, \omega^{Ec}_{i,0,k+1}\} = \{\omega^f_{i,t_{max},k}, \omega^c_{i,t_{max},k}\} \tag{10}$$

$$\mathcal{L}^E_i = -\sum_{j=1}^{I_i} y_{i,j} log \frac{exp(d'_{i,j})}{\sum_{b=1}^{I_i} exp(d'_{i,b})} \tag{11}$$

where $d'_{i,j}$ is the logit computed by Eqs. (1) and (2) with model parameters of the local expert. As shown in Fig. 2(b), Eqs. (7) and (10), the difference between the local client model and the local expert is that the former receives a global model for starting feature embedding network with the same initialisation among clients, while the latter utilises the latest local feature embedding network as an initialisation. Since the local expert is only learned with local data without receiving external updates, it performs better than the global model on the local domain but worse than the global model on the other domains (likely overfit locally). Thus, it can be used as a regularisation to help the updated local client model to learn richer knowledge on a specific local domain. We feed the same input to the two models but with separate random augmentation and compute soft probability distributions for the local client model ($\mathcal{P}_{i,j}$) and the local expert ($\mathcal{Q}_{i,j}$) as:

$$\mathcal{P}_{i,j} = \frac{exp(d_{i,j}/T)}{\sum_{b=1}^{I_i} exp(d_{i,b}/T)}, \mathcal{Q}_{i,j} = \frac{exp(d'_{i,j}/T)}{\sum_{b=1}^{I_i} exp(d'_{i,b}/T)} \tag{12}$$

where T is a temperature [23]. The local expert regularisation is therefore defined as the KL-divergence $\mathcal{L}^R_i$ between $\mathcal{P}_{i,j}$ and $\mathcal{Q}_{i,j}$:

$$\mathcal{L}^R_i = T^2 \sum_{j=1}^{I_i} \mathcal{Q}_{i,j} \cdot log \frac{\mathcal{Q}_{i,j}}{\mathcal{P}_{i,j}} \tag{13}$$

In summary, each local client is optimised with Eq. (3) using local private data independent from other clients. We summarise the training process in Algorithm 1.

Algorithm 1 Federated Person Re-Identification (FedReID).

Intialise: Local datasets $\mathcal{X}_i(i{=}1,...N)$, client number N, selected client fraction C, local client models $\{\phi(\omega^f_{i,t,k};x), \delta(\omega^c_{i,t,k};v)\}$, local experts $\{\phi(\omega^{Ef}_{i,t,k};x), \delta(\omega^{Ec}_{i,t,k};v)\}$, a global model of the central server $\sigma(\theta_k,x)$.

for $k = 1 \to k_{max}$ **do** /* *Global communication* */
    /* Local Client Update */
    for $i \to N$ **do** /* *i-th client* */
        Update local client model with Eq. (15)
        Update local expert with Eq. (10)
        for $t = 1 \to t_{max}$ **do** /* Local steps */
            Compute features of local data (Eq. (1))
            Compute logits of local data (Eq. (2))
            Compute classification loss $\mathcal{L}^C_i$ (Eq.(9))
            Compute classification loss $\mathcal{L}^E_i$ (Eq.(11))
            Compute expert regularisation $\mathcal{L}^R_i$ (Eq. (13))
            Backward to update local models Eq. (3)
        end for
    end for
    /* Central Server Update */
    Randomly select C-fraction clients
    Update the central server with Eq. (14)
end for
Output: A generalised global feature embedding model in the central server

Privacy Protection. In FedReID, local sensitive data are inherently protected by the decentralised learning process. To further protect sensitive data from inverse attack [20], we use a white noise [24] to hide the contributions of each client in Eq. (8):

$$\theta^f_{k+1} = \frac{1}{\lceil C{\cdot}N \rceil} \sum_{m \in N_C} \omega^f_{m,t_{max},k} + \beta\mathcal{N}(0,1) \tag{14}$$

where $\mathcal{N}(0,1)$ is the white noise matrices with mean 0 and variance 1. $\beta \in [0,1]$ is a scale factor to control the balance between privacy-preserving and Re-ID accuracy. When $\beta = 0$, the white noise is removed from the aggregation. Moreover, in the client-server collaboration, we can further hide the collaboration information in Eq. (7) as:

$$\{\omega^f_{i,0,k+1}, \omega^c_{i,0,k+1}\} = \{\theta^f_{k+1} + \beta\mathcal{N}(0,1), \omega^c_{i,t_{max},k}\} \tag{15}$$

IV. Experiments

A. Datasets and Settings

Datasets. We used four large-scale Re-ID datasets (DukeMTMC-ReID [25], Market-1501 [26], CUHK03-NP [27], [28] and MSMT17 [29]) as non-shared local datasets in four client sites. Each of the four local clients didn't share its training dataset with other clients nor the server. This is significantly different from other generalised Re-ID methods [10], where FedReID is trained with decentralised data, while existing methods are trained with centralised data. The FedReID model was then tested on five smaller Re-ID datasets (VIPeR [30], iLIDS [31], 3DPeS [32], CAVIAR [33] and GRID [34]), plus a large-scale Re-ID dataset (CUHK-SYSU person search [35]) as new unseen target domains for out-of-the-box deployment tests. For the CUHK-SYSU test, we used the ground-truth person bounding box annotation from the dataset for Re-ID test, of which there are 2900 query persons and each person contains at least one image in the gallery (both query and gallery sets are fixed removing distractors in the variable gallery sets). Following previous studies [10], on small Re-ID datasets, we did random half splits to generate 10 training/testing splits. In each split, we randomly selected one image of each test identity as the query while the others as the gallery for evaluation. The dataset statistics are summarised in Table I. Besides, we employed CIFAR-10 [36] for federated formulation generalisation analysis on image classification.

TABLE I
THE RE-ID DATASET STATISTICS. 'ID': NUMBER OF IDENTITIES; 'IMG': NUMBER OF IMAGES.

Types	Benchmarks	Train ID	Train Img	Test ID	Test Img
Local Clients Training	Duke	702	16522	-	-
	Market	751	12936	-	-
	CUHK03	767	7365	-	-
	MSMT17	1041	30248	-	-
New Domains Testing	VIPeR	-	-	316	632
	iLIDS	-	-	60	120
	3DPeS	-	-	96	192
	CAVIAR	-	-	36	72
	GRID	-	-	125	1025
	CUHK-SYSU	-	-	2900	8347

Evaluation Metrics. We used Rank-1 (R1) accuracy and mean Average Precision (mAP) for Re-ID performance evaluation. Note that, FedReID is uniquely designed to learn a generalised model from distributed datasets with privacy protection, not to achieve the best accuracy.

Implementation Details. In our design, the feature embedding network is ResNet-50 [37] (pretrained on ImageNet), while the mapping network consists of two fully connected layers, in which the first fully connected layer follows by a batch normalization layer, a ReLU layer and Dropout. By default, we set the number of local clients N=4, selected client fraction C=1.0 , and privacy control parameter β=0. These hyperparameters can be determined by different application requirements. We empirically set batch size to 32, maximum global communication epochs k_{max}=100, maximum local steps t_{max}=1, and T=3 in Eq. (12). We used SGD as the optimiser with Nesterov momentum 0.9 and weight decay $5e$-4. The learning rates were set to 0.01 for embedding networks and 0.1 for mapping networks, which decayed by 0.1 every 40 epochs.

B. Federated Formulation Generalisation Analysis

To evaluate the generalisation of the reformulated paradigm of FedReID, we compared FedReID with FedSGD [12], FedAVG [12], and FedATT [14]. When adapting FedSGD, FedAVG and FedATT for Re-ID, we set the last classification layer with the maximal identity number among local datasets. When adapting FedReID for image classification, we aggregate all layers (including classifiers) during the client-server

TABLE II
EVALUATING GENERALISATION OF FEDERATED FORMULATIONS. TOP-1/R1 ACCURACIES ARE REPORTED. FEDSGD MEANS SETTING C=1.0 AND t_{max}=1.

Methods	CIFAR-10	VIPeR	iLIDS
FedSGD	90.72±0.04	41.0	65.2
FedAVG (C=0.5)	93.04±0.19	41.5	65.3
FedAVG (C=1.0)	93.21±0.08	41.0	65.2
FedATT (C=0.5)	92.86±0.12	38.3	65.3
FedATT (C=1.0)	92.97±0.21	40.2	61.3
FedReID (C=0.5)	93.14±0.14	45.3	68.3
FedReID (C=1.0)	**93.31±0.11**	**46.2**	**69.7**
Baseline (Centralised joint-train)	93.58±0.14	44.6	65.5

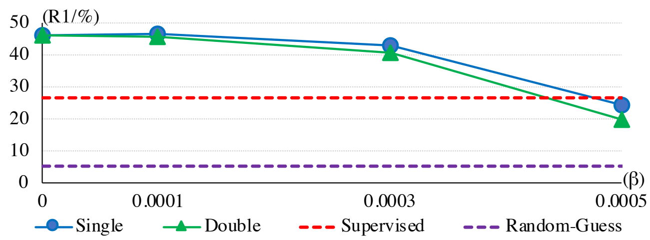


Fig. 3. Evaluating the privacy control parameter β on VIPeR. 'Single': use β in Eq. (14), 'Double': use β in both Eqs. (14) and (15), 'Random-Guess': initialise the model with ImageNet pretrained parameters without training.

collaboration. We employed ResNet-32 for experiments and set t_{max}=5 on CIFAR-10. As shown in Table II, FedReID achieves better generalisation than the other federated variants. When training and testing on the same domain for image classification (CIFAR-10), FedReID performs closely to the centralised joint-training, but FedReID only accesses to local model updates with privacy protection whilst joint-training assembles together all local data without privacy concerns. When testing out-of-the-box on unseen new domains for decentralised Re-ID (VIPeR and iLIDS), FedReID performs even better than joint-training, which can be attributed to: (1) FedReID accumulate multiple local model updates to find wider optima in parameter space of a global model; (2) With the client-server collaboration, FedReID learns more soft distribution knowledge. Therefore, FedReID could get better generalisation for an unseen new Re-ID domain while joint-training is the upper bound for testing on a source client.

C. Privacy Protection Analysis

The inherent defensive ability of FedReID is given by decentralised learning and model aggregation. Besides, β in Eqs. (14) and (15) can further control privacy protection. From Fig. 3, we can see that: (1) R1 accuracy of FedReID gradually decreases when β increases, but FedReID achieves significantly better accuracy than the Random-Guess; (2) Single β protection performs slightly better than double protection, but the double one protects more information; (3) When $\beta = 0.0005$, R1 accuracy of FedReID remains close to the overfitting local supervised method, which indicates the compromise of accuracy and privacy.

TABLE III
COMPARISON WITH INDIVIDUAL CLIENTS AND ENSEMBLES.

Settings	Methods	VIPeR (R1)	iLIDS (R1)
Individuals	Client (Duke)	25.0	56.2
	Client (Market)	26.1	48.0
	Client (CUHK03)	21.6	41.0
	Client (MSMT)	27.3	60.5
Ensembles	Feat-Concatenation	29.4	56.2
	Parameter-Average	19.9	41.8
Decentralised	FedReID	**46.2**	**69.7**
Centralised	Baseline (Joint-train)	44.6	65.5

D. Comparison with Individuals and Ensembles

We separately trained the baseline model on four localised datasets as the individuals and used feature concatenation and model parameter average as the ensembles. As shown in Table III: (1) FedReID significantly outperforms the individuals and the ensembles, which shows that the collaboration between the localised clients and the centralised server facilitates holistic optimisation, enabling FedReID to construct a better generalisable model with privacy protection; (2) Compared with the centralised joint-training baseline, FedReID achieves competitive R1 accuracies, demonstrating its effectiveness; (3) Averaging multiple trained local models in parameter space leads to an arbitrarily bad model.

E. Generalised Re-ID Performance Evaluation

FedReID is designed uniquely for protecting local client privacy by learning a generalisable model without centralised sharing of training data, whilst no existing Re-ID methods explicitly integrate privacy protection requirements into their designs. To further show the generalisation of the central model, we evaluated FedReID against some contemporary Re-ID methods. We employed out-of-the-box evaluation without using any training data on target domains.

Results on Smaller Benchmarks. We compared FedReID with: (1) four unsupervised cross-domain fine-tuning methods (TJAIDL [38], DSTML [39], UMDL [40], PAUL [7]), and (2) five unsupervised generalisation methods (SyRI [41], SSDAL [42], MLDG [43], CrossGrad [44], DIMN [10]). As shown in Table IV, FedReID performs competitively against the contemporary methods, which shows the effectiveness of the generalised global model for out-of-the-box deployment with privacy protection. For example, FedReID achieves 46.5% R1 on VIPeR and 69.7% R1 on iLIDS, which are the second-best and close to DIMN [10] which assembles all training data together without data privacy protection. Note that, FedReID is uniquely designed to learn a global model from distributed clients with privacy protection, so we didn't elaborately design the backbone to surpass the accuracy of contemporary methods without privacy protection, but FedReID is compatible with state-of-the-art techniques, such as saliency detection and style transfer [2], [5], [45].

Results on Large Benchmark. To further evaluate FedReID on a large-scale target domain, we used the Re-ID subset of CUHK-SYSU person search dataset, which has distinctively different scene context to most other Re-ID datasets above. As

TABLE IV
GENERALISED RE-ID PERFORMANCE EVALUATION ON VIPER, ILIDS, 3DPES, CAVIAR AND GRID. R1 ACCURACY. †: RE-ID DOMAIN GENERALISATION RESULTS.

Settings	Methods	VIPeR	iLIDS	3DPeS	CAVIAR	GRID
w/o privacy (Cross-domain fine-tune)	TJAIDL	38.5	-	-	-	-
	DSTML	8.6	33.4	32.5	28.2	-
	UMDL	31.5	49.3	-	41.6	-
	PAUL	45.2	-	-	-	-
w/o privacy (Centralised generalised)	SyRI	43.0	56.5	-	-	-
	SSDAL	43.5	-	-	-	22.4
	MLDG†	23.5	53.8	-	-	15.8
	CrossGrad†	20.9	49.7	-	-	9.0
	DIMN	51.2	70.2	-	-	29.3
w/ privacy (Decentralised)	FedReID	46.2	69.7	67.0	45.6	24.2

TABLE V
EVALUATION ON CUHK-SYSU PERSON SEARCH SUBSET FOR RE-ID. ⋆: EXPERIMENTS USING GROUND-TRUTH PERSON IMAGES AND A GALLERY SIZE OF 100 IMAGES PER QUERY.

Methods	mAP	R1	R5	R10
DSIFT⋆+Euclidean	41.1	45.9	-	-
BoW⋆+Cosine	62.5	67.2	-	-
DLDP⋆	74.0	76.7	-	-
FedReID	80.4	83.4	90.3	92.4
Baseline (Centralised joint-train)	74.7	77.4	87.2	90.1

TABLE VI
SOURCE LABELLED CLIENT TESTS. THE CENTRAL MODEL OF FEDREID ONLY AGGREGATES LOCAL MODEL UPDATES *without accessing local data*.

Methods	Market		Duke	
	R1	mAP	R1	mAP
Baseline (local supervised)	88.3	71.4	77.3	58.9
FedReID (direct w/o access local data)	80.2	60.1	68.0	52.1
FedReID (supervised w/ local data&distillation)	90.3	76.2	77.4	60.7
DPR (unsupervised multi-domain distillation)	61.5	33.5	48.4	29.4
DPR (semi-supervised multi-domain distillation)	63.7	35.4	57.4	36.7

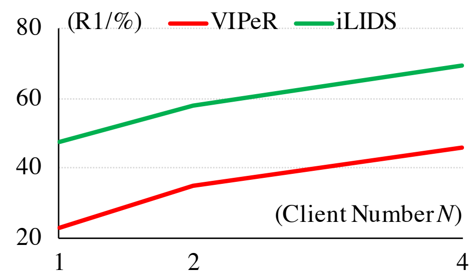


Fig. 4. Evaluating client numbers on VIPeR and iLIDS.

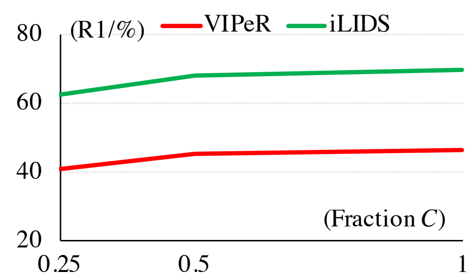


Fig. 5. Evaluating client fraction on VIPeR and iLIDS.

shown in Table V, FedReID achieves competitive performance compared with other unsupervised methods (DSIFT [46], BoW [26] and DLDP [47]) and the centralised join-training baseline, which shows the generalisation of the global model of FedReID for deployment on large-scale Re-ID.

F. Further Analysis and Discussion

Source Client Tests. In addition to out-of-the-box deployment, we also evaluated FedReID on source clients which contain private local data. As shown in Table VI, although FedReID (direct) only aggregates local model updates, it still achieves competitive performance, especially when compared with DPR [48], a state-of-the-art multi-domain distillation Re-ID method, which fine-tunes a model with local unlabelled data (unsupervised) or with partial local labelled data (semi-supervised). We also tested FedReID (supervised) by using local labelled data for model learning with the global model for knowledge distillation. We can see that FedReID (supervised) improves FedReID (direct) and outperforms Baseline (local supervised), which shows the effectiveness of FedReID for supervised deployment.

Client Number N**.** Fig. 4 compares central server aggregation with different numbers of local clients, where N=1, 2 and 4 denote Market, Market+Duke and Market +Duke+Cuhk03+Msmt as clients, respectively. We can see that that collaboration of multi-domain clients is good for learning more generalisable knowledge in the central server.

Client Fraction C**.** Fig. 5 compares FedReID with different client fractions C. We can see that updating with an arbitrary client (C=0.25) is inferior to aggregating multiple clients, whilst aggregating all clients (C=1.0) performs slightly better than aggregating randomly selected clients (C=0.5), but random selection protects more data privacy.

Client Local Step t_{max}**.** Fig. 6 compares FedReID with different client local steps which potentially promote communication efficiency. Overall, the performance of FedReID gradually decreases when t_{max} increases due to the accumulation of biases in each local client.

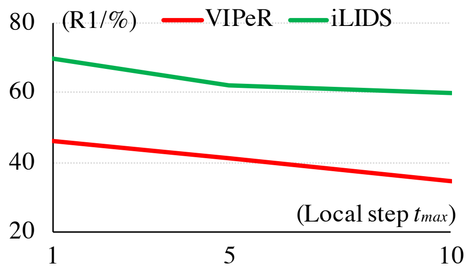


Fig. 6. Evaluating client local steps on VIPeR and iLIDS.

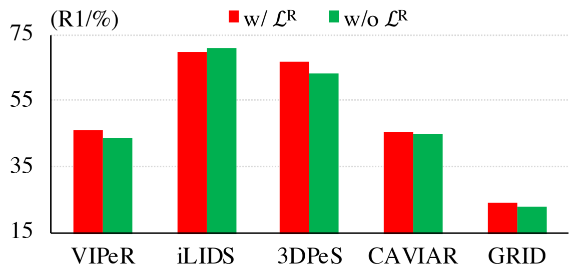


Fig. 7. Evaluating expert regularisation on VIPeR, iLIDS, 3DPeS, CAVIAR, and GRID.

Expert Regularisation $\mathcal{L}_i^R$. Fig. 7 shows the evaluation on the regularisation of local experts in FedReID. We can see that with the expert regularisation, FedReID gets better generalisation performance, which shows that the expert regularisation provides richer knowledge to facilitate the improvement of FedReID.

V. Conclusion

In this work, we introduced decentralised learning from independent multi-domain label spaces for person Re-ID and proposed a fundamentally novel Re-ID paradigm called Federated Person Re-Identification (FedReID). We trained multiple local models at different user sites (clients) independently using non-shared private local data and aggregated local model updates to construct a generalisable global model as a central server without accessing local data. We iteratively performed this client-server collaborative learning process under privacy control, so that we can get a generalisable global feature embedding model for out-of-the-box deployment with privacy protection. Extensive experiments on ten Re-ID benchmarks show that FedReID can both protect local data privacy and achieve compelling generalisation performance, which is uniquely advantageous over contemporary Re-ID methods that assume shared centralised training data without privacy protection.

References

[1] S. Gong, M. Cristani, S. Yan, and C. C. Loy, *Person Re-Identification*. Springer, 2014.
[2] C. Liu, X. Chang, and Y.-D. Shen, "Unity style transfer for person re-identification," in *CVPR*, 2020.
[3] Q. Dong, S. Gong, and X. Zhu, "Person search by text attribute query as zero-shot learning," in *ICCV*, 2019.
[4] G. Wu, X. Zhu, and S. Gong, "Tracklet self-supervised learning for unsupervised person re-identification," in *AAAI*, 2020.
[5] X. Chen, C. Fu, Y. Zhao, F. Zheng, J. Song, R. Ji, and Y. Yang, "Salience-guided cascaded suppression network for person re-identification," in *CVPR*, 2020.
[6] D. Wang and S. Zhang, "Unsupervised person re-identification via multi-label classification," in *CVPR*, 2020.
[7] Q. Yang, H.-X. Yu, A. Wu, and W.-S. Zheng, "Patch-based discriminative feature learning for unsupervised person re-identification," in *CVPR*, 2019.
[8] J. Konečný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated learning: Strategies for improving communication efficiency," *arXiv preprint arXiv:1610.05492*, 2016.
[9] G. Wu, X. Zhu, and S. Gong, "Person re-identification by ranking ensemble representations," in *ICIP*, 2019.
[10] J. Song, Y. Yang, Y.-Z. Song, T. Xiang, and T. M. Hospedales, "Generalizable person re-identification by domain-invariant mapping network," in *CVPR*, 2019.
[11] T. Xiao, H. Li, W. Ouyang, and X. Wang, "Learning deep feature representations with domain guided dropout for person re-identification," in *CVPR*, 2016.
[12] H. B. McMahan, E. Moore, D. Ramage, S. Hampson *et al.*, "Communication-efficient learning of deep networks from decentralized data," in *AISTATS*, 2017.
[13] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," *ACM TIST*, 2019.
[14] S. Ji, S. Pan, G. Long, X. Li, J. Jiang, and Z. Huang, "Learning private neural language modeling with attentive aggregation," in *IJCNN*, 2019.
[15] J. L. McClelland and D. E. Rumelhart, *Explorations in parallel distributed processing: A handbook of models, programs, and exercises*. MIT press, 1989.
[16] J. Dean, G. Corrado, R. Monga, K. Chen, M. Devin, M. Mao, M. Ranzato, A. Senior, P. Tucker, K. Yang *et al.*, "Large scale distributed deep networks," in *NeurIPS*, 2012.
[17] F. N. Iandola, M. W. Moskewicz, K. Ashraf, and K. Keutzer, "Firecaffe: near-linear acceleration of deep neural network training on compute clusters," in *CVPR*, 2016.
[18] N. Papernot, M. Abadi, U. Erlingsson, I. Goodfellow, and K. Talwar, "Semi-supervised knowledge transfer for deep learning from private training data," in *ICLR*, 2017.
[19] J. Wang, W. Bao, L. Sun, X. Zhu, B. Cao, and S. Y. Philip, "Private model compression via knowledge distillation," in *AAAI*, 2019.
[20] M. Fredrikson, S. Jha, and T. Ristenpart, "Model inversion attacks that exploit confidence information and basic countermeasures," in *ACM CCS*, 2015.
[21] N. Papernot, S. Song, I. Mironov, A. Raghunathan, K. Talwar, and Ú. Erlingsson, "Scalable private learning with pate," in *ICLR*, 2018.
[22] I. J. Goodfellow, O. Vinyals, and A. M. Saxe, "Qualitatively characterizing neural network optimization problems," in *ICLR*, 2015.
[23] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," *arXiv preprint arXiv:1503.02531*, 2015.
[24] R. C. Geyer, T. Klein, and M. Nabi, "Differentially private federated learning: A client level perspective," in *NeurIPS*, 2017.
[25] Z. Zheng, L. Zheng, and Y. Yang, "Unlabeled samples generated by gan improve the person re-identification baseline in vitro," in *ICCV*, 2017.
[26] L. Zheng, L. Shen, L. Tian, S. Wang, J. Wang, and Q. Tian, "Scalable person re-identification: A benchmark," in *ICCV*, 2015.
[27] W. Li, R. Zhao, T. Xiao, and X. Wang, "Deepreid: Deep filter pairing neural network for person re-identification," in *CVPR*, 2014.
[28] Z. Zhong, L. Zheng, D. Cao, and S. Li, "Re-ranking person re-identification with k-reciprocal encoding," in *CVPR*, 2017.
[29] L. Wei, S. Zhang, W. Gao, and Q. Tian, "Person transfer gan to bridge domain gap for person re-identification," in *CVPR*, 2018.
[30] D. Gray and H. Tao, "Viewpoint invariant pedestrian recognition with an ensemble of localized features," in *ECCV*, 2008.
[31] W.-S. Zheng, S. Gong, and T. Xiang, "Associating groups of people," in *BMVC*, 2009.
[32] D. Baltieri, R. Vezzani, and R. Cucchiara, "3DPeS: 3D people dataset for surveillance and forensics," in *ACM J-HGBU*, 2011.
[33] D. S. Cheng, M. Cristani, M. Stoppa, L. Bazzani, and V. Murino, "Custom pictorial structures for re-identification," in *BMVC*, 2011.
[34] C. C. Loy, C. Liu, and S. Gong, "Person re-identification by manifold ranking," in *ICIP*, 2013.
[35] T. Xiao, S. Li, B. Wang, L. Lin, and X. Wang, "Joint detection and identification feature learning for person search," in *CVPR*, 2017.
[36] A. Krizhevsky and G. Hinton, "Learning multiple layers of features from tiny images," Tech. Rep., 2009.

[37] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *CVPR*, 2016.

[38] J. Wang, X. Zhu, S. Gong, and W. Li, “Transferable joint attribute-identity deep learning for unsupervised person re-identification,” in *CVPR*, 2018.

[39] J. Hu, J. Lu, and Y.-P. Tan, “Deep transfer metric learning,” in *CVPR*, 2015.

[40] P. Peng, T. Xiang, Y. Wang, M. Pontil, S. Gong, T. Huang, and Y. Tian, “Unsupervised cross-dataset transfer learning for person re-identification,” in *CVPR*, 2016.

[41] S. Bak, P. Carr, and J.-F. Lalonde, “Domain adaptation through synthesis for unsupervised person re-identification,” in *ECCV*, 2018.

[42] C. Su, S. Zhang, J. Xing, W. Gao, and Q. Tian, “Deep attributes driven multi-camera person re-identification,” in *ECCV*, 2016.

[43] D. Li, Y. Yang, Y.-Z. Song, and T. M. Hospedales, “Learning to generalize: Meta-learning for domain generalization,” in *AAAI*, 2018.

[44] S. Shankar, V. Piratla, S. Chakrabarti, S. Chaudhuri, P. Jyothi, and S. Sarawagi, “Generalizing across domains via cross-gradient training,” in *ICLR*, 2018.

[45] X. Jin, C. Lan, W. Zeng, Z. Chen, and L. Zhang, “Style normalization and restitution for generalizable person re-identification,” in *CVPR*, 2020.

[46] R. Zhao, W. Ouyang, and X. Wang, “Unsupervised salience learning for person re-identification,” in *CVPR*, 2013.

[47] A. Schumann, S. Gong, and T. Schuchert, “Deep learning prototype domains for person re-identification,” in *ICIP*, 2017.

[48] A. Wu, W.-S. Zheng, X. Guo, and J.-H. Lai, “Distilled person re-identification: Towards a more scalable system,” in *CVPR*, 2019.