

# Positive Style Accumulation: A Style Screening and Continuous Utilization Framework for Federated DG-ReID

Xin Xu

Wuhan University of  
Science and Technology  
Wuhan, China  
xuxin@wust.edu.cn

Wenke Huang\*

Wuhan University  
Wuhan, China  
wenkehuang@whu.edu.cn

Chaoyue Ren

Wuhan University of  
Science and Technology  
Wuhan, China  
renchaoyue@wust.edu.cn

Bin Yang

Wuhan University  
Wuhan, China  
yangbin\_cv@whu.edu.cn

Wei Liu\*

Wuhan University of  
Science and Technology  
Wuhan, China  
liuwei@wust.edu.cn

Zhixi Yu

Wuhan University of  
Science and Technology  
Wuhan, China  
yuzhixi@wust.edu.cn

Kui Jiang

Harbin Institute of Technology  
Harbin, China  
jiangkui@hit.edu.cn

## Abstract

The Federated Domain Generalization for Person re-identification (FedDG-ReID) aims to learn a global server model that can be effectively generalized to source and target domains through distributed source domain data. Existing methods mainly improve the diversity of samples through style transformation, which to some extent enhances the generalization performance of the model. However, we discover that **not all styles contribute to the generalization performance**. Therefore, we define styles that are **beneficial/harmful** to the model's generalization performance as **positive/negative styles**. Based on this, new issues arise: **How to effectively screen and continuously utilize the positive styles**. To solve these problems, we propose a Style Screening and Continuous Utilization (SSCU) framework. Firstly, we design a Generalization Gain-guided Dynamic Style Memory (GGDSM) for each client model to screen and accumulate generated positive styles. Specifically, the memory maintains a prototype initialized from raw data for each category, then screens positive styles that enhance the global model during training, and updates these positive styles into the memory using a momentum-based approach. Meanwhile, we propose a style memory recognition loss to fully leverage the positive styles memorized by GGDSM. Furthermore, we propose a Collaborative Style Training (CST) strategy to make full use of positive styles. Unlike traditional learning strategies, our approach

leverages both newly generated styles and the accumulated positive styles stored in memory to train client models on two distinct branches. This training strategy is designed to effectively promote the rapid acquisition of new styles by the client models, ensuring that they can quickly adapt to and integrate novel stylistic variations. Simultaneously, this strategy guarantees the continuous and thorough utilization of positive styles, which is highly beneficial for the model's generalization performance. Extensive experimental results demonstrate that our method outperforms existing methods in both the source domain and the target domain.

## CCS Concepts

- Information systems → Information retrieval.

## Keywords

Federated DG-ReID, Negative Style, Positive Style Memory

### ACM Reference Format:

Xin Xu, Chaoyue Ren, Wei Liu, Wenke Huang, Bin Yang, Zhixi Yu, and Kui Jiang. 2025. Positive Style Accumulation: A Style Screening and Continuous Utilization Framework for Federated DG-ReID. In *Proceedings of the 33rd ACM International Conference on Multimedia (MM '25), October 27–31, 2025, Dublin, Ireland*. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3746027.3755549>

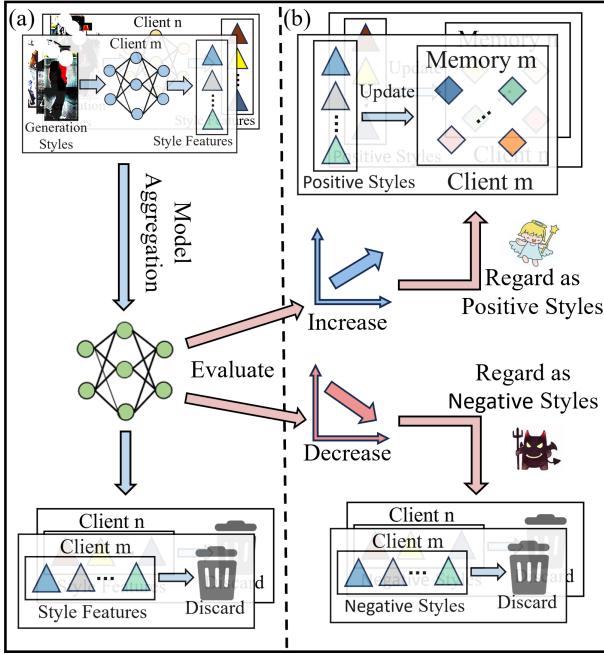
\*Corresponding author

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.  
MM '25, Dublin, Ireland.

© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
ACM ISBN 979-8-4007-2035-2/2025/10  
<https://doi.org/10.1145/3746027.3755549>

## 1 Introduction

In recent years, Person re-identification (ReID) has garnered significant research attention, aiming to achieve accurate cross-camera recognition of the same individual. With the success of deep learning, numerous high-performance ReID methods have been proposed [5, 6, 23, 32, 35, 40–42, 49, 52]. However, constrained by limitations in data collection and the complexity of real-world scenarios, these methods often underperform when deployed to unseen domains. To address this, recent efforts have focused on Domain Generalization for Person re-identification (DG-ReID) [3, 4, 11, 30, 38, 46]. DG-ReID aims to train models on multiple source



**Figure 1: The differences between SSCU and existing methods are as follows:** (a) Traditional methods discard the generated styles immediately after each round of training. (b) SSCU evaluates the global model at the end of each round of training, and if the model's performance has improved, the generated styles from the current round are regarded as positive styles and updated into the memory; otherwise, they are considered negative styles and discarded directly.

domains and test on unseen target domains, thereby enhancing robust cross-domain generalization. However, existing approaches rely on large-scale centralized labeled datasets, which often raise critical data privacy concerns in practical applications [36, 44].

To solve this issue, federated learning has been introduced into DG-ReID [19, 37, 48], termed Federated Domain Generalization for Person re-identification (FedDG-ReID). Federated learning is a distributed machine learning framework [2, 10, 25] that facilitates knowledge sharing through a cross-device/cross-institution collaborative training paradigm while rigorously safeguarding data privacy. The technology embodies the core principle of "moving models, not data," where participants solely exchange updates to model parameters while retaining raw data locally, thereby effectively addressing the dual challenges of data silos and privacy leakage. However, due to the limited amount of data that each client can access, and the significant heterogeneity between the data of different clients, traditional centralized generalization strategies cannot be directly applied. Existing methods predominantly focus on generating synthetic data via style transfer to simulate unseen domains [8, 13, 27, 31, 34, 39, 51]. However, as shown in Fig. 1, these methods **overlook that not all styles contribute to the model's generalization performance**, thereby lacking the capability to screen and utilize positive styles, which is crucial for the model's generalization performance in both the source and target domains.

In this paper, we propose a **Style Screening and Continuous Utilization (SSCU)** framework to address the previously outlined issues of **positive style selection, memory, and continuous utilization**, achieving robust cross-domain generalization while ensuring privacy preservation. Specifically, we design a **Generalization Gain-guided Dynamic Style Memory (GGDSM)** for each client to enable selection and cross-round accumulation of positive styles. (1) Initialization: To build a robust identity-discriminative feature representation for each person identity, we perform clustering and averaging of all training data based on identity before the official training starts on each client. Category prototypes emphasize consistency within classes and differences between classes, and can be used to guide the model to learn more discriminative feature representations. (2) Positive style selection: At the end of each training round on the client, we evaluate the optimization effect of the generated styles on the global model. Based on the evaluation results, we determine whether these generated styles are positive for the model update, and if they are, we update them to the memory for continuous utilization; otherwise, we consider them to be negative styles and discard them directly. (3) Update strategy: Category prototypes are updated via momentum-based integration, ensuring stable incorporation of new styles while preserving previously memorized positive patterns. This helps the model progressively learn domain-invariant feature extraction capabilities. Furthermore, in order to realize the full use of the style in the memory, we propose a **Collaborative Style Training (CST)** training strategy comprising two parallel training branches: (1) New style adaptation branch: In each iteration, new stylized data is randomly generated, and features are extracted using the client-global model for loss calculation. The client-global model downloaded from the server possesses better generalization knowledge, making it more suitable for rapidly learning new style changes within a short period of time. (2) Positive style continuous utilization branch: In this branch, the client-local model and the client-global model are trained using the original images, and then optimized using a loss function based on the dynamic style memory. Since the category prototypes stored in the memory remember all the positive styles from previous rounds, this branch allows the model to continuously make use of them.

Our main contributions can be summarized as follows:

- **Empirical Contribution.** We discover that not all styles generated through style transformation methods contribute to the improvement of model generalization performance. Some styles may introduce invalid data, which is instead detrimental to model optimization.
- **Framework Contribution.** We propose a style screening and continuous utilization framework that effectively screens, memorizes, and continuously utilizes generative styles beneficial to model generalization performance with minimal additional overhead.
- **Technical Contribution.** We propose GGDSM and CST. GGDSM screens and memorizes styles that are positive for model generalization, while CST leverages both newly generated styles and the accumulated positive styles stored in memory to train client models. This enables the models to quickly adapt to new styles and continuously utilize positive

styles, resulting in significant improvements in generalization performance in both the source and target domains.

## 2 Related work

### 2.1 Domain Generalization for Person Re-identification

Recently, Domain Generalization for person re-identification have gained widespread attention. The goal is to enable models to learn from one or multiple source domains and generalize effectively to unseen domains. The key to domain generalization methods lies in enabling models to learn robust domain-invariant feature representations. Existing approaches achieve this goal from various perspectives, including optimizing training strategies [1, 15, 20, 21, 45], leveraging causal mechanisms [24, 26], and employing data augmentation [13, 18, 27, 31, 34, 39, 51]. Recent methods mainly seek to improve DG performance through three aspects: meta-learning, feature disentanglement, and data augmentation.

Methods based on meta-learning aim to enhance model generalization by simulating realistic training-test domain shifts. For example, Choi et al. [3] introduced MetaBIN, which combines Batch Normalization (BN) and Instance Normalization (IN) through a meta-learning training strategy and a set of learnable balancing parameters. Dai et al. [4] proposed a Relation-aware Mixture of Experts (RaMoE) method, integrating meta-learning into a novel mixture-of-experts paradigm via an efficient voting-based fusion mechanism to dynamically aggregate multi-source domain features.

Some methods adopt feature disentanglement strategies, focusing on disentangling different attribute features of pedestrians to help models learn more discriminative identity features. Jin et al. [12] proposed a Style Normalization and Recovery (SNR) module, which eliminates the influence of style features via IN layers while extracting and restoring identity-discriminative features filtered by IN. Zhang et al. [46] designed an Adaptive Cross-domain Learning (ACL) framework, which introduces multiple parallel feature embedding networks to capture domain-invariant and domain-specific features separately and adaptively aggregates them using domain-aware adapters to mitigate cross-domain interference. Zhang et al. [50] developed a Disentangled Invariant Representation (DIR) framework, constructing a Structural Causal Model (SCM) between identity-specific and domain-specific factors to eliminate spurious correlations and filter domain-related information.

Regarding data augmentation, recent studies have demonstrated that diversifying image styles (e.g., background environments, lighting conditions, camera positions) significantly enhances model generalization. Zhou et al. [51] proposed MixStyle, which generates new styles by linearly combining style feature statistics from different source domains. Nuriel et al. [29] introduced Permuted adaIN (pAdaIN), exchanging feature statistics within a batch to reduce network reliance on global image statistics and enhance the utilization of shape and local image cues.

### 2.2 Federated Domain Generalization for Person Re-identification

Federated learning is a distributed machine learning technique that enables multiple devices to collaborate while ensuring privacy

protection [7, 9]. McMahan et al. [28] proposed the first federated learning algorithm, FedAvg, which aggregates locally trained client models on the server via averaging and redistributes the aggregated model to clients for further training.

In recent years, as privacy concerns have grown in domain generalization tasks, integrating federated learning into ReID frameworks has emerged as a promising research direction. Although FedDG-ReID shares the same objective as DG tasks, traditional DG methods, which rely on large-scale centralized data for training, struggle to adapt directly under privacy-constrained decentralized data architectures. Existing mainstream FedDG-ReID approaches primarily address this challenge through three directions: client model training strategies, server model aggregation strategies, and local training data diversification. For example, Zhuang et al. [53] combined knowledge distillation to transfer client model knowledge to the server with dynamic weight adjustment methods, which dynamically adjust model aggregation weights based on client model variations. Wu et al. [37] introduced a local expert model for each client to enrich local knowledge acquisition and aggregated only feature embedding networks during server updates to preserve local classification knowledge. Zhang et al. [47] optimized local models via proximal and feature regularization terms to improve local training accuracy, while using cosine similarity of backbone features to determine global aggregation weights for each local model, ensuring global convergence. Regarding local training data diversification: Yang et al. [43] designed a Style Transfer Model (STM) for each client to generate new styles for local data, with loss functions constraining style diversity and authenticity. Liu et al. [22] simulated unseen domains via a domain compensation network (DCN), exposing models to broader data distributions during training.

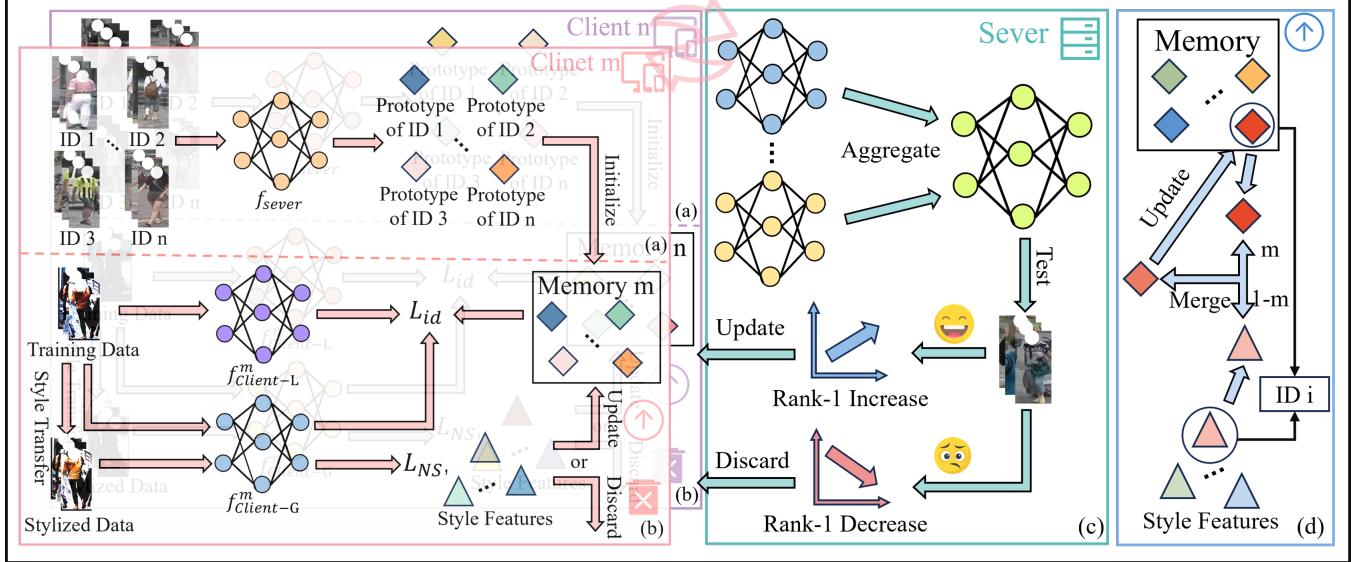
## 3 Methodology

For FedDG-ReID, we provide  $K$  source domains  $D_{\text{source}} = \{D_k\}_{k=1}^K$ . During training, there is no interaction among the source domains. The goal of the training task is to obtain a global model that can be generalized. During the testing phase, the global model is tested directly in the source or target domain  $D_T$ .

As shown in Fig. 2, our method mainly consists of two parts: (1) generalization gain-guided dynamic style memory (Section 3.1) and (2) collaborative style training strategy (Section 3.2). The collaborative training process under the federated framework mainly consists of five steps: Step (1) Local training: Use specific training strategies to optimize the model located on the client; Step (2) Model upload: Upload the specific model optimized by local training to the server for aggregation; Step (3) Model aggregation: Use specific methods to aggregate the models uploaded by each client to obtain a global model; Step (4) Model redistribution: Redistribute the new global model to each client, and continue to iterate steps (1) to (4) until convergence; Step (5) Test the final global model on the target domain. Our method mainly works in step (1).

### 3.1 Generalization Gain-Guided Dynamic Style Memory

Style diversification methods can effectively help models learn more discriminative features. Although existing work has demonstrated their effectiveness in the FedDG-ReID task, these methods



**Figure 2: Illustration of our SSCU method.** (a) **Memory Initialization.** Before the training begins, clients use the pretrained global model downloaded from the server to extract features, obtaining the category prototypes for each identity. (b) Our CST strategy. Specifically, the figure illustrates the training process of both the new style adaptation branch and the positive style continuous utilization branch, as well as the generation of style features used for updating the memory. (c) Subsequent learning of our method. After local training concludes, the server aggregates the client-global models uploaded by each client and evaluates them. The evaluation results decide whether to memorize the generated styles of this round. (d) We present the details of the memory update. New style features are updated to the category prototypes corresponding to their identities in a momentum-based manner.

overlook the potential issue in practical applications that not all generated styles are beneficial to the gains in model generalization performance. Consequently, the model may fail to form long-term memory of positive styles and filter out negative styles. Considering the aforementioned issues, we propose a generalization gain-guided dynamic style memory. We maintain a separate memory for each client, storing the category prototype of each identity, with no interaction between the memories of different clients. Using category prototypes as classification features, which can simplify feature representation and improve classification efficiency, is highly suitable for complex open-set scenarios like ReID.

**Memory initialization.** As shown in Fig. 2(a), we use a pretrained global model to extract features from all training samples, and then aggregate these features by identity to calculate the average, obtaining the category prototype to each identity.

$$M_k^i = \frac{1}{n_k^i} \sum_{x \in X_k^i} f_{\text{pretrain}}(x), \quad (1)$$

where  $X_k^i$  is the set of all images with identity  $i$  in client  $k$ , with a total number of  $n_k^i$ , and the category prototype is  $M_k^i$ .

**Positive style screening and memory update.** We design an update mechanism for the memory that is guided by the generalization performance gain of the global model, in order to selectively preserve positive generated styles and filter out negative ones. Specifically, after each training round, We assess the global model's performance, with Rank-1 serving as the key indicator

of generalization capability. We then compare these evaluation outcomes with those from the preceding round. If the styles produced in the current round have helped to enhance the model's generalization performance, we regard them as positive styles and update them to the memory; otherwise, we regard them as negative styles and discard them directly. We update the memory using a momentum-based approach, which operates as follows:

$$\hat{x}_k^i = f_T^k(x_k^i), \quad (2)$$

$$M_k^i \leftarrow (1 - m)M_k^i + m \sum_{x \in B_k^i} \|f_{\text{Client-G}}^k(\hat{x}_k^i)\|_2, \quad (3)$$

where  $B_k^i$  represents all images of identity  $i$  in the mini-batch training data at the  $k$ -th client,  $f_{\text{Client-G}}^k$  is the client-global model of the  $k$ -th client,  $f_T^k$  is the style transformation model of the  $k$ -th client,  $\|\cdot\|_2$  indicates the  $L_2$  normalization operation, and  $m \in [0, 1]$  is a momentum coefficient used to control the update rate of category prototypes. By allowing category prototypes to continuously accumulate positive style features through a momentum-based approach over multiple rounds of training, we can not only achieve sustained cumulative memory of positive styles but also implicitly expand the training samples, so that we can also alleviate the challenges posed by limited training data to a certain extent.

**Style memory recognition loss.** In the task of FedDG-ReID, cross-entropy loss is often used to help the model learn to distinguish between different individuals. It features low computational complexity, ease of implementation, and optimization. Existing methods have also demonstrated their rapid convergence, efficient training, and strong application compatibility. However, due to the characteristic of cross-entropy loss that overly focuses on the correctness of identity classification, the trained models tend to overfit to specific styles and struggle to adapt to data heterogeneity. Considering these drawbacks, we have designed a style memory recognition loss. This is achieved by maximizing the similarity between features and prototypes of the same category, and minimizing the similarity with other categories, thus helping the model learn features that are more robust to noise and style variations. Therefore, by applying cross-entropy loss and recognition loss at different training scenes, the strengths of both methods can complement each other, aiding the model in achieving better generalization capability. The implementation of the loss function we designed is as follows:

Given an image, we pass it into the model  $f$  that needs to be optimized to perform forward propagation and obtain features, and then calculate the similarity between it and each prototype in the memory. For a source domain  $D_k$  with  $P_k$  identities, the specific calculation formula is as follows:

$$L_{id}(f, x_k^i, M_k^i) = -\log \frac{\exp(\|f(x_k^i)\|_2^T M_k^i / \tau)}{\sum_{n=1}^{P_k} \exp(\|f(x_k^i)\|_2^T M_k^n / \tau)}, \quad (4)$$

among them,  $x_k^i$  is an image with identity  $i$  in domain  $D_k$ ,  $\tau$  is the temperature factor,  $\exp(\cdot)$  denotes the operation of raising  $e$  to a power, and the similarity between the extracted features and the category prototypes is calculated using the dot product.

### 3.2 Collaborative Style Training Strategy

Unlike traditional learning strategies, we propose a novel collaborative training strategy for new and old styles, aiming to achieve rapid adaptation to new styles and continuous utilization of positive styles. As shown in Fig. 2 (b), our strategy consists of two branches: (1) New style adaptation branch: This branch adjusts the style statistics of the original images to generate diverse data, and then uses the client-global model for rapid learning and adaptation. (2) Positive style continuous utilization branch: This branch focuses on continuously learning and utilizing verified positive styles. In each iteration, we leverage the memory mechanism's ability to remember and reinforce the positive styles learned. This helps the model maintain a rich style knowledge system during long-term training, thereby achieving more robust and sustained generalization performance improvement.

**The new style adaptation branch.** To ensure that the client models can still achieve sufficient generalization ability with limited data, it is essential for them to continuously learn new styles. We refer to the stylization method proposed in [43] and input the original image  $x_k^i$  into the stylization network at the beginning of each iteration to generate a new stylized image  $\hat{x}_k^i$  like Eq. 2. Then, we download the global model from the server to the client, referred to as the client-global model  $f_{Client-G}$ . Compared to the

---

**Algorithm 1** : SSCU

---

**Require:** Client set  $K$ , sample set of identity  $i$  at client  $k$ :  $X_k^i = [X_k^{(i,1)}, \dots, X_k^{(i,n_k^i)}]$ , the total number of pedestrian identities on client  $k$ :  $P_k$ , the client-local model of a client  $k$ :  $f_k^{Client-L}$  and style transformation model:  $f_T^k$ , server model  $f_{server}$ , communication rounds  $E$

```

1: function INITIALIZE( $X_k, M_k, f$ )
2:   for  $id = 1, 2, \dots, P_k$  do
3:      $M_k^{id} \leftarrow \frac{1}{n_k^{id}} \sum_{n=1}^{n_k^{id}} \|f(X_k^{(id,n)})\|_2$  in Eq. 1
4:   end for
5:   return  $M_k$ 
6: end function
7: function UPDATE( $M_k, F_k$ )
8:   for  $id = 1, 2, \dots, P_k$  do
9:      $M_k^{id} \leftarrow (1 - m)M_k^{id} + mF_k^{id}$  in Eq. 3
10:     $\triangleright F_k^{id}$  is  $L_2$  normalized
11:   end for
12:   return  $M_k$ 
13: end function
14: function CST( $f_{Client-G}^k, f_{Client-L}^k, f_T^k, X_k, M_k$ )
15:    $B_k \leftarrow$  Data sampling( $X_k$ )  $\triangleright$  Common random sampling
16:    $\hat{B}_k \leftarrow f_T^k(B_k)$  in Eq. 2
17:    $f_{Client-G}^k \leftarrow L_{NS}(f_{Client-G}^k, \hat{B}_k)$  in Eq. 5
18:    $f_{Client-G}^k, f_{Client-L}^k \leftarrow L_{PS}($ 
19:    $f_{Client-G}^k, f_{Client-L}^k, B_k, M_k)$  in Eq. 8
20:    $\hat{F}_k = \|f_{Client-G}^k(\hat{B}_k)\|_2$   $\triangleright \hat{F}_k$  is style features
21:   return  $f_{Client-G}^k, f_{Client-L}^k, \hat{F}_k$ 
22: end function
23: for  $k \in K$  in parallel do
24:    $M_k \leftarrow$  INITIALIZE( $X_k, M_k, f_{server}$ )
25: end for
26: for  $e = 1, 2, \dots, E$  do
27:   for  $k \in K$  in parallel do
28:      $f_{Client-G}^k \leftarrow$  Distribute( $f_{server}$ )
29:      $\triangleright$  Also known as server model download
30:      $f_{Client-G}^k, f_{Client-L}^k, \hat{F}_k \leftarrow$  CST(
31:        $f_{Client-G}^k, f_{Client-L}^k, f_T^k, X_k, M_k$ )
32:   end for
33:    $f_{server} \leftarrow \sum_{k=1}^K \frac{N_k}{N} f_{Client-G}^k$   $\triangleright$  Evaluate after aggregation
34:   if  $e = 0$  or  $f_{server}$ 's Rank-1 increase then
35:      $\triangleright$  Compare with the last round
36:     for  $k \in K$  in parallel do
37:        $M_k \leftarrow$  UPDATE( $M_k, \hat{F}_k$ )
38:     end for
39:   end if
40: end for
41: return  $f_{server}$ 

```

---

client-local model, the client-global model possesses richer global generalization knowledge. Coupled with the rapid convergence characteristic of the cross-entropy loss, it can achieve faster learning

and adaptation to new styles. Meanwhile, we also employ the widely used triplet loss. Therefore, the training of the new style adaptation branch on the  $k$ -th client can be expressed as:

$$\begin{aligned} L_{NS}^k &= L_{CE}(cls_k(f_{Client-G}^k(x_k^i)), y_k^i) \\ &\quad + L_{Tri}(f_{Client-G}^k(\hat{x}_k^i, y_k^i)), \end{aligned} \quad (5)$$

among them,  $cls_k$  is a classifier maintained by the client  $k$  that possesses local knowledge, used for precise classification,  $y_k^i$  is the label corresponding to  $x_k^i$ ,  $L_{CE}$  stands for calculation of cross entropy and  $L_{Tri}$  stands for calculation of triplet loss.

**Positive style continuous utilization branch.** Existing learning strategies do not take into account the issue that not all styles contribute to the improvement of model generalization performance, and therefore cannot achieve continuous and effective utilization of positive styles while learning new ones. To address this issue, we design a continuous utilization branch for positive styles, coupled with the previously introduced dynamic style memory, to enable effective training of client models using positive styles continuously. Specifically, in this branch, both the client-local model and the client-global model are trained using the untransformed original images  $x_k^i$  as input, and optimized using the style memory recognition loss introduced earlier (Eq. 4):

$$L_{PS\_L}^k(f_{Client-L}^k, x_k^i, M_k^i) = L_{id}(f_{Client-L}^k, x_k^i, M_k^i), \quad (6)$$

$$L_{PS\_G}^k(f_{Client-G}^k, x_k^i, M_k^i) = L_{id}(f_{Client-G}^k, x_k^i, M_k^i), \quad (7)$$

$$\begin{aligned} L_{PS}^k(f_{Client-L}^k, f_{Client-G}^k, x_k^i, M_k^i) &= L_{PS\_L}^k(f_{Client-L}^k, x_k^i, M_k^i) \\ &\quad + L_{PS\_G}^k(f_{Client-G}^k, x_k^i, M_k^i), \end{aligned} \quad (8)$$

where  $f_{Client-L}^k$  is the client-local model of client  $k$ . By applying recognition loss to the client-local model, it helps to maintain specific domain knowledge, thereby assisting the style transformation model in generating higher-quality data [43]. Meanwhile, applying recognition loss to the client-global model enables continuous utilization and reinforcement of positive styles that have been quickly adapted, further enriching the model's generalization knowledge.

### 3.3 Subsequent Learning

**Model upload and aggregation.** After completing local training, each client uploads the feature extractor part of its client-global model to the central server. A data-volume-weighted aggregation strategy is employed for fusion, ensuring that clients with more abundant data have a greater “say” in the global model aggregation.

**Continuous training and final testing.** The clients and the server iteratively perform model download, local training, model aggregation, and model redistribution until training convergence. Our entire methodological process is illustrated in Alg. 1.

## 4 Experiments

### 4.1 Experimental Settings

**Datasets.** We conduct experiments on four large-scale datasets: CUHK02, CUHK03, MSMT17, and Market1501. For simplicity, we refer to them as C2 (CUHK02), C3 (CUHK03), MS (MSMT17), and

M (Market1501). In order to make a comprehensive comparison of the generalization performance on the source and target domains with existing works, we adopt three evaluation schemes:

*Protocol-1:* Under this Protocol, we adopt a leave-one-out strategy, where each time one dataset is selected from the four datasets mentioned above as the test set, all the remaining datasets are used as the training set.

*Protocol-2:* This protocol is a supplement to *Protocol-1*. For the two cases in *Protocol-1* where MS is used as the training dataset, we reduce the number of source domains by one to test the generalization performance of our method with fewer source domains(MS is always used for training).

*Protocol-3:* In this protocol, we choose C2, C3, and M as the training datasets, and separately use the test data from one of these datasets to test on the source domains.

**Implementation Details.** We treat each source domain as a client, using the same network as DACS [43] as our backbone. The number of training epochs is set to 60, the batch size is set to 64, the initial learning rate is set to 1e-3, weight decay is set to 5e-4, momentum is set to 0.9, and MultiStepLR is used as the learning rate scheduler with milestones set at the 20th and 40th epochs. We use python (3.9) and PyTorch (2.1.0) to train on two Nvidia GeForce RTX-2080Ti GPUs.

### 4.2 Comparisons With State-of-the-art Methods

We conduct comparisons with SOTA methods under three different protocols to demonstrate the effectiveness of our approach.

*Protocol-1:* As shown in Tab. 1, the methods we compare can be categorized into five classes: (1) First are the classical federated learning algorithms, such as SCAFFOLD [14], FedProx [17] and MOON [16]. (2) Next are domain generalization methods based on styling approaches, with the most classical being MixStyle [51] and CrossStyle [33], which we deploy directly on the clients to generate diverse data for local training. (3) Following this are the representative federated ReID methods, such as FedPav [53] and FedReID [37]. (4) Subsequently, there are single-source domain generalization ReID methods without privacy constraints, such as SNR [12], where we deploy each module after every layer of ResNet. (5) Finally, there are federated domain generalization ReID methods, such as DACS [43]. The experimental data shows that our method achieves state-of-the-art performance. Specifically, for “MS+C2+C3→M”, the mAP reaches 39.2%, and Rank-1 reaches 66.4%, which is at least  $\uparrow 2.9\%$  and  $\uparrow 5.2\%$  ahead of other methods respectively. For “M+C2+C3→MS”, the mAP reaches 11.9%, and Rank-1 reaches 32.3%, which is at least  $\uparrow 1.5\%$  and  $\uparrow 4.8\%$  ahead of other methods respectively. For “M+C2+MS→C3”, the mAP reaches 32.8%, which is at least  $\uparrow 2.1\%$  higher than other methods, and Rank-1 reaches 34.1%, on par with the best-performing method.

*Protocol-2:* As shown in Tab. 2, for settings “MS+C2+C3→M” and “M+C2+MS→C3”, we respectively remove C2, C3, and M from the source domains to compare our method with the state-of-the-art methods under the condition of having fewer source domain data. We compare our method with those that still perform excellently in this field, such as the FedPav [53], FedReID [37], and DACS [43]. It can be observed that even in the face of the challenge of reduced source domain data, our method still outperforms the SOTAs.

**Table 1: Comparison of different methods under *Protocol-1*. M: Market1501, C2: CUHK02, C3: CUHK03, MS: MSMT17. Average represents the average performance over three unseen domains.**

Category	Methods	Reference	MS+C2+C3→M		M+C2+C3→MS		MS+C2+M→C3		Average	
			mAP	rank-1	mAP	rank-1	mAP	rank-1	mAP	rank-1
Federated Learning	SCAFFOLD[14]	ICML 2020	26.0	50.5	5.3	15.8	22.9	26.0	18.1	30.8
	MOON[16]	CVPR 2021	26.8	51.1	4.8	14.5	20.9	22.5	17.5	29.4
Domain Generalization	FedProx[17]	MLSys 2020	29.3	53.8	5.8	17.4	19.1	17.7	18.1	29.6
	MixStyle[51]	ICLR 2021	31.2	53.5	5.5	16.0	28.6	31.5	21.8	33.7
Federated-ReID	CrossStyle[33]	ICCV 2021	35.5	59.6	4.6	14.0	27.8	28.0	22.6	33.9
	FedReID[37]	AAAI 2021	30.1	53.7	4.5	13.7	26.4	26.5	20.3	31.3
DG-ReID	FedPav[53]	MM 2020	25.4	49.4	5.2	15.5	22.5	24.3	17.7	29.7
DG-ReID	SNR[12]	CVPR 2020	32.7	59.4	5.1	15.3	28.5	30.0	22.1	34.9
FedDG-ReID	DACS[43]	AAAI 2024	36.3	61.2	10.4	27.5	30.7	<b>34.1</b>	25.8	40.9
	<b>SSCU (ours)</b>	This Paper	<b>39.5</b>	<b>66.4</b>	<b>11.9</b>	<b>32.3</b>	<b>32.8</b>	<b>34.1</b>	<b>28.1</b>	<b>44.3</b>

**Table 2: Comparison of different methods under *Protocol-2*. Under this protocol, we use fewer source domains to further compare generalization performance.**

Methods	MS+C3→M		MS+C2→M		MS+C2+C3→M	
	mAP	rank-1	mAP	rank-1	mAP	rank-1
FedPav	27.5	51.5	24.8	48.5	25.4	49.4
FedReID	31.0	55.0	28.1	52.4	30.1	53.7
DACS	33.2	58.1	30.3	56.3	36.3	61.2
<b>SSCU (ours)</b>	<b>36.7</b>	<b>62.8</b>	<b>34.8</b>	<b>62.7</b>	<b>39.5</b>	<b>66.4</b>
Methods	MS+M→C3		MS+C2→C3		MS+C2+M→C3	
	mAP	rank-1	mAP	rank-1	mAP	rank-1
FedPav	15.2	14.1	17.3	17.0	22.5	24.3
FedReID	16.1	15.3	21.8	20.4	26.4	26.5
DACS	18.2	17.7	22.9	23.5	30.7	<b>34.1</b>
<b>SSCU (ours)</b>	<b>20.9</b>	<b>20.8</b>	<b>27.1</b>	<b>29.3</b>	<b>32.8</b>	<b>34.1</b>

**Table 3: Comparison of different methods on *Protocol-3*. Under this protocol, we focus on the model’s identification performance on the source domains, the most basic yet often overlooked capability by existing methods.**

Methods	M+C2+C3 → M		M+C2+C3 → C2		M+C2+C3 → C3	
	mAP	rank-1	mAP	rank-1	mAP	rank-1
<b>FedProx</b>	61.0	80.4	66.8	65.5	24.2	23.9
<b>FedPav</b>	53.9	76.0	59.7	56.3	19.6	19.6
<b>FedReID</b>	71.8	87.6	82.9	82.8	44.0	44.9
<b>DACS</b>	72.1	88.2	84.5	83.4	47.4	50.1
<b>SSCU (ours)</b>	<b>73.0</b>	<b>88.7</b>	<b>84.9</b>	<b>83.9</b>	<b>50.4</b>	<b>53.2</b>

*Protocol-3:* As shown in Tab. 3, we deployed the models trained under the source domain setting of *Protocol-1* back to each source domain to test the recognition capability of our method on the source domains and also compared it with the state-of-the-art methods. The experimental results show that our method not only achieves

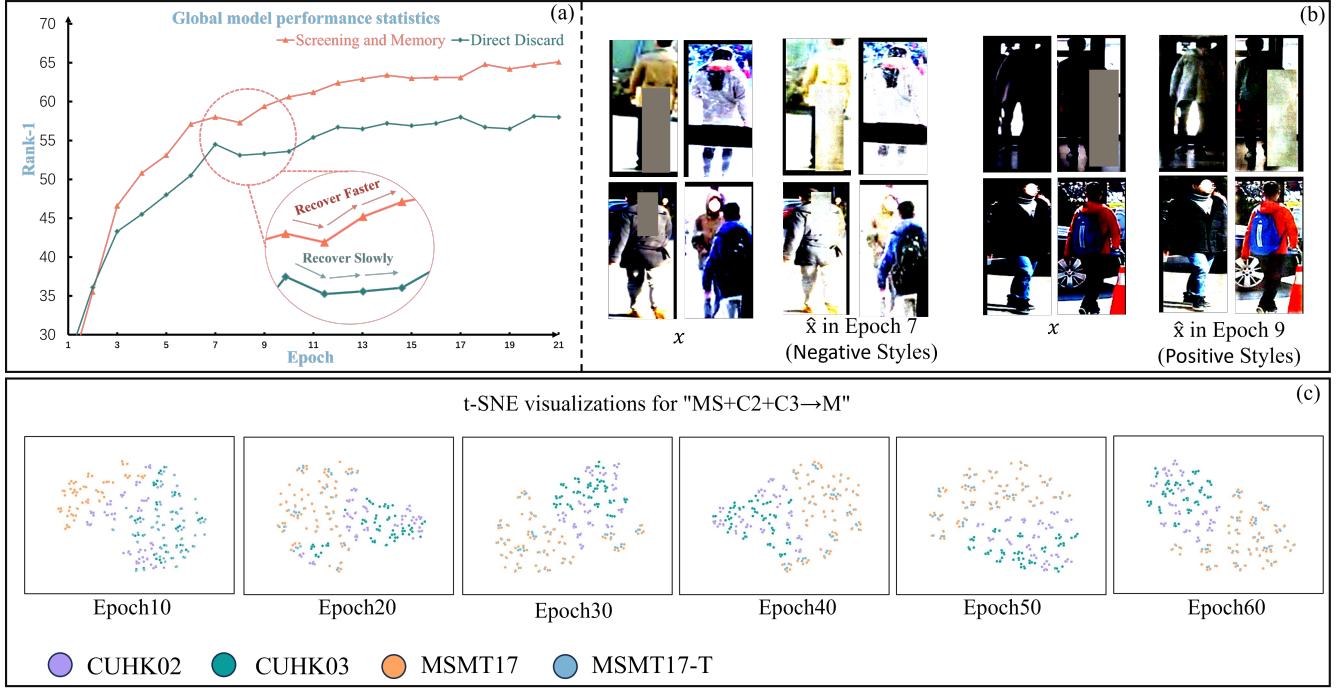
**Table 4: Ablation study on new style adaptation branch (NSA) and positive style continuous utilization branch (PSCU).**

Attributes	MS+C2 +C3→M		M+C2 +C3→MS		
	NSA	PSCU	mAP	rank-1	mAP
×	×	20.8	51.9	6.1	20.7
×	✓	33.6	61.3	7.5	22.6
✓	✗	36.2	63.9	7.9	24.7
✓	✓	39.5	66.4	11.9	32.3

excellent generalization performance but also ensures the optimal recognition performance on the source domains.

### 4.3 Ablation Study

**Effectiveness of the NSA and PSCU.** As previously mentioned, our CST strategy consists of two branches: the New Style Adaptation (NSA) branch and the Positive Style Continuous Utilization (PSCU) branch. We explored the effects of these two branches separately to demonstrate the effectiveness of our strategy. “Baseline” indicates that we directly trained the client-global model using the original images and cross-entropy loss during the local training phase. “Baseline+NSA” means that we trained both the client-local and client-global models using the original images, while also training the client-global model with style-transformed images, but without introducing the memory mechanism. Therefore, the classification loss used was the cross-entropy loss function. “Baseline+PSCU” indicates that for each round of generated style images, they were directly memorized in the memory without being used for the training of the client-global model in the current round. The classification loss was our proposed style memory identification



**Figure 3: (a) Performance statistics of the method that directly discards generated styles and the method that screens and memorizes styles. (b) Display of positive and negative styles during the training process. (c)  $t$ -SNE visualization of “MS+C2+C3→M”. MSMT17-T: Style data generated from the MSMT17 dataset.**

loss. "Baseline+NSA+PSCU" is the method we introduced earlier. All settings were the same as our basic training phase. As shown in the Tab. 4, our training pipeline shows significant improvement. Both branches contribute to the improvement of the model's generalization performance, and the best results are achieved when they are used in conjunction.

#### 4.4 Visualization

During the implementation of the "MS+C2+C3→M" experiment, we statistically analyze the performance changes of our method when applying the screening and memorizing strategy (introduced in the previous text) for styles and the direct discarding method (used in existing approaches), and plot the results as line graphs for a better understanding of our solution. As shown in Fig. 3(a), both methods experienced a severe performance drop after the completion of the seventh round of training. However, due to our screening and memorizing strategy, the model could continuously benefit from positive styles, thus recovering more quickly from the impact of the performance decline. In Fig. 3(b), we present the negative styles generated in the seventh round, where it can be observed that the transformed images have lost a significant amount of detail, which is detrimental to model optimization. Correspondingly, we showcase the positive styles generated in the ninth round, where it is clearly evident that these images retain detailed information while undergoing style transformation, and following the ninth round of training for both methods, there was an enhancement in their performance. In the  $t$ -SNE visualization shown in Fig. 3(c),

it can be clearly observed that in the initial stage of training, the feature points of different datasets are distributed relatively dispersely in the feature space, without forming distinct distinctions. As the training progresses step by step, the feature points gradually gather towards the centers of their respective datasets, forming tighter and more defined clusters, and gradually stabilize. It is worth noting that during the analysis of the CHUK02 and CUHK03, it was found that these two datasets have a certain similarity in content, resulting in some features being distributed relatively close.

## 5 Conclusion

In this paper, we propose a Style Screening and Continuous Utilization (SSCU) framework for federated generalized person re-identification. This method not only achieves precise screening of positive styles but also ensures their maximum utilization through an efficient memory mechanism. Specifically, our approach first introduces a generalization gain-guided dynamic style memory for screening and memorization. To ensure the continuous utilization of these positive styles, we further design a style memory recognition loss function. Correspondingly, we devise a style collaborative training strategy to simultaneously learn newly generated styles and the positive styles stored in memory. The new style rapid adaptation branch facilitates quick adaptation to new style changes through the use of a client-global model. The positive style continuous utilization branch is responsible for fully leveraging the positive styles. Extensive experiments show the efficacy of our method.

## Acknowledgments

This work was supported by the National Nature Science Foundation of China (No. 62376201). This research was financially supported by funds from Key Laboratory of Social Computing and Cognitive Intelligence (Dalian University of Technology), Ministry of Education (No. SCCI2024YB02).

## References

- [1] Yogesh Balaji, Swami Sankaranarayanan, and Rama Chellappa. 2018. Metareg: Towards domain generalization using meta-regularization. *Advances in neural information processing systems* 31 (2018).
- [2] Neda Bugshan, Ibraheem Khalil, Mohammad Saidur Rahman, Mohammed Atiquzzaman, Xun Yi, and Shahriar Badsha. 2022. Toward trustworthy and privacy-preserving federated deep learning service framework for industrial internet of things. *IEEE Transactions on Industrial Informatics* 19, 2 (2022), 1535–1547.
- [3] Seocheon Choi, Taekyung Kim, Minki Jeong, Hyoungseob Park, and Changick Kim. 2021. Meta batch-instance normalization for generalizable person re-identification. In *Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition*. 3425–3435.
- [4] Yongxing Dai, Xiaotong Li, Jun Liu, Zekun Tong, and Ling-Yu Duan. 2021. Generalizable person re-identification with relevance-aware mixture of experts. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 16145–16154.
- [5] Zan Gao, Lishuai Gao, Hua Zhang, Zhiyong Cheng, Richang Hong, and Shengyong Chen. 2020. DCR: A unified framework for holistic/partial person ReID. *IEEE Transactions on Multimedia* 23 (2020), 3332–3345.
- [6] Zan Gao, Hongwei Wei, Weili Guan, Weizhi Nie, Meng Liu, and Meng Wang. 2022. Multigranular visual-semantic embedding for cloth-changing person re-identification. In *Proceedings of the 30th ACM international conference on multimedia*. 3703–3711.
- [7] Wenke Huang, Mang Ye, and Bo Du. 2022. Learn from others and be yourself in heterogeneous federated learning. In *CVPR*.
- [8] Wenke Huang, Mang Ye, Zekun Shi, and Bo Du. 2023. Generalizable Heterogeneous Federated Cross-Correlation and Instance Similarity Learning. *TPAMI* (2023).
- [9] Wenke Huang, Mang Ye, Zekun Shi, He Li, and Bo Du. 2023. Rethinking Federated Learning with Domain Shift: A Prototype View. In *CVPR*.
- [10] Wenke Huang, Mang Ye, Zekun Shi, Guancheng Wan, He Li, Bo Du, and Qiang Yang. 2024. A Federated Learning for Generalization, Robustness, Fairness: A Survey and Benchmark. *TPAMI* (2024).
- [11] Bingliang Jiao, Lingqiao Liu, Liying Gao, Guosheng Lin, Lu Yang, Shizhou Zhang, Peng Wang, and Yanning Zhang. 2022. Dynamically transformed instance normalization network for generalizable person re-identification. In *European conference on computer vision*. Springer, 285–301.
- [12] Xin Jin, Cuiling Lan, Wenjun Zeng, Zhibo Chen, and Li Zhang. 2020. Style normalization and restitution for generalizable person re-identification. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 3143–3152.
- [13] Juwon Kang, Sohyun Lee, Namyup Kim, and Suha Kwak. 2022. Style neophile: Constantly seeking novel styles for domain generalization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 7130–7140.
- [14] Sai Praneeth Karimireddy, Satyen Kale, Mehryar Mohri, Sashank Reddi, Sebastian Stich, and Ananda Theertha Suresh. 2020. Scaffold: Stochastic controlled averaging for federated learning. In *International conference on machine learning*. PMLR, 5132–5143.
- [15] Jin Kim, Jiyoung Lee, Jungin Park, Dongbo Min, and Kwanghoon Sohn. 2022. Pin the memory: Learning to generalize semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 4350–4360.
- [16] Qinbin Li, Bingsheng He, and Dawn Song. 2021. Model-contrastive federated learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 10713–10722.
- [17] Tian Li, Anit Kumar Sahu, Manzil Zaheer, Maziar Sanjabi, Ameet Talwalkar, and Virginia Smith. 2020. Federated optimization in heterogeneous networks. *Proceedings of Machine learning and systems* 2 (2020), 429–450.
- [18] Yiming Li, Shuo Shao, Yu He, Junfeng Guo, Tianwei Zhang, Zhan Qin, Pin-Yu Chen, Michael Backes, Philip Torr, Dacheng Tao, and Kui Ren. 2025. Rethinking Data Protection in the (Generative) Artificial Intelligence Era. *arXiv preprint arXiv:2507.03034* (2025).
- [19] Ying Li, Xingwei Wang, Rongfei Zeng, Praveen Kumar Donta, Ilir Murturi, Min Huang, and Schahram Dustdar. 2023. Federated domain generalization: A survey. *arXiv preprint arXiv:2306.01334* (2023).
- [20] Yiyi Li, Yongxin Yang, Wei Zhou, and Timothy Hospedales. 2019. Feature-critic networks for heterogeneous domain generalization. In *International conference on machine learning*. PMLR, 3915–3924.
- [21] Ke Liang, Lingyu Meng, Hao Li, Jun Wang, Long Lan, Miaomiao Li, Xinwang Liu, and Huaimin Wang. 2025. From Concrete to Abstract: Multi-view Clustering on Relational Knowledge. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2025), 1–18. doi:10.1109/TPAMI.2025.3582689
- [22] Fangyi Liu, Mang Ye, and Bo Du. 2024. Domain generalized federated learning for Person Re-identification. *Computer Vision and Image Understanding* 241 (2024), 103969.
- [23] Hao Luo, Youzhi Gu, Xingyu Liao, Shenqi Lai, and Wei Jiang. 2019. Bag of tricks and a strong baseline for deep person re-identification. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*. 0–0.
- [24] Fangrui Lv, Jian Liang, Shuang Li, Bin Zang, Chi Harold Liu, Ziteng Wang, and Di Liu. 2022. Causality inspired representation learning for domain generalization. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 8046–8056.
- [25] Yanbiao Ma, Wei Dai, Wenke Huang, and Jiayi Chen. 2025. Geometric Knowledge-Guided Localized Global Distribution Alignment for Federated Learning. In *CVPR*.
- [26] Divyat Mahajan, Shruti Tople, and Amit Sharma. 2021. Domain generalization using causal matching. In *International conference on machine learning*. PMLR, 7313–7324.
- [27] Massimiliano Mancini, Zeynep Akata, Elisa Ricci, and Barbara Caputo. 2020. Towards recognizing unseen categories in unseen domains. In *European Conference on Computer Vision*. Springer, 466–483.
- [28] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Agüera y Arcas. 2017. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*. PMLR, 1273–1282.
- [29] Oren Nuriel, Sagie Benaim, and Lior Wolf. 2021. Permuted adain: Reducing the bias towards global statistics in image classification. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 9482–9491.
- [30] Lei Qi, Lei Wang, Yinghuan Shi, and Xin Geng. 2022. A novel mix-normalization method for generalizable multi-source person re-identification. *IEEE Transactions on Multimedia* 25 (2022), 4856–4867.
- [31] Yichun Shi, Xiang Yu, Kihyuk Sohn, Manmohan Chandraker, and Anil K Jain. 2020. Towards universal representation learning for deep face recognition. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 6817–6826.
- [32] Yifan Sun, Liang Zheng, Yi Yang, Qi Tian, and Shengjin Wang. 2018. Beyond part models: Person retrieval with refined part pooling (and a strong convolutional baseline). In *Proceedings of the European conference on computer vision (ECCV)*. 480–496.
- [33] Zhiqiang Tang, Yunhe Gao, Yi Zhu, Zhi Zhang, Mu Li, and Dimitris N Metaxas. 2021. Crossnorm and selfnorm for generalization under distribution shifts. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 52–61.
- [34] Riccardo Volpi and Vittorio Murino. 2019. Addressing model vulnerability to distributional shifts over image transformation sets. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 7980–7989.
- [35] Haochen Wang, Jiayi Shen, Yongtuo Liu, Yan Gao, and Efstratios Gavves. 2022. Nformer: Robust person re-identification with neighbor transformer. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 7297–7307.
- [36] Yang Wang, Jinjia Peng, Huibing Wang, and Meng Wang. 2022. Progressive learning with multi-scale attention network for cross-domain vehicle re-identification. *Science China Information Sciences* 65, 6 (2022), 160103.
- [37] Guile Wu and Shaogang Gong. 2021. Decentralised learning from independent multi-domain labels for person re-identification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35. 2898–2906.
- [38] Boqiang Xu, Jian Liang, Lingxiao He, and Zhenan Sun. 2021. Meta: Mimicking embedding via others' aggregation for generalizable person re-identification. *arXiv preprint arXiv:2112.08684* (2021).
- [39] Chenggang Yan, Biao Gong, Yuxuan Wei, and Yue Gao. 2020. Deep multi-view enhancement hashing for image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43, 4 (2020), 1445–1451.
- [40] Chenggang Yan, Yiming Hao, Liang Li, Jian Yin, Anan Liu, Zhendong Mao, Zhenyu Chen, and Xingyu Gao. 2021. Task-adaptive attention for image captioning. *IEEE Transactions on Circuits and Systems for Video technology* 32, 1 (2021), 43–51.
- [41] Chenggang Yan, Lixuan Meng, Liang Li, Jiehua Zhang, Zhan Wang, Jian Yin, Jiyong Zhang, Yaoqi Sun, and Bolun Zheng. 2022. Age-invariant face recognition by multi-feature fusion and decomposition with self-attention. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* 18, 1s (2022), 1–18.
- [42] Cheng Yan, Guansong Pang, Xiao Bai, Changhong Liu, Xin Ning, Lin Gu, and Jun Zhou. 2021. Beyond triplet loss: person re-identification with fine-grained difference-aware pairwise loss. *IEEE Transactions on Multimedia* 24 (2021), 1665–1677.
- [43] Fengxiang Yang, Zhun Zhong, Zhiming Luo, Yifan He, Shaozi Li, and Nicu Sebe. 2024. Diversity-authenticity co-constrained stylization for federated domain

- generalization in person re-identification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 38, 6477–6485.
- [44] Mang Ye, Jianbing Shen, Gaojie Lin, Tao Xiang, Ling Shao, and Steven CH Hoi. 2021. Deep learning for person re-identification: A survey and outlook. *IEEE transactions on pattern analysis and machine intelligence* 44, 6 (2021), 2872–2893.
- [45] Ji Zhang, Jingkuan Song, Lianli Gao, Ye Liu, and Heng Tao Shen. 2022. Progressive meta-learning with curriculum. *IEEE Transactions on Circuits and Systems for Video Technology* 32, 9 (2022), 5916–5930.
- [46] Pengyi Zhang, Huanzhang Dou, Yunlong Yu, and Xi Li. 2022. Adaptive cross-domain learning for generalizable person re-identification. In *European conference on computer vision*. Springer, 215–232.
- [47] Pengling Zhang, Huibin Yan, Wenhui Wu, and Shuoyao Wang. 2023. Improving federated person re-identification through feature-aware proximity and aggregation. In *Proceedings of the 31st ACM International Conference on Multimedia*, 2498–2506.
- [48] Ruipeng Zhang, Qinwei Xu, Jiangchao Yao, Ya Zhang, Qi Tian, and Yanfeng Wang. 2023. Federated domain generalization with generalization adjustment. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*.
- [49] Xinyu Zhang, Dongdong Li, Zhigang Wang, Jian Wang, Errui Ding, Javen Qinfeng Shi, Zhaoxiang Zhang, and Jingdong Wang. 2022. Implicit sample extension for unsupervised person re-identification. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 7369–7378.
- [50] Yi-Fan Zhang, Zhang Zhang, Da Li, Zhen Jia, Liang Wang, and Tieniu Tan. 2022. Learning domain invariant representations for generalizable person re-identification. *IEEE Transactions on Image Processing* 32 (2022), 509–523.
- [51] Kaiyang Zhou, Yongxin Yang, Yu Qiao, and Tao Xiang. 2021. Domain generalization with mixstyle. *arXiv preprint arXiv:2104.02008* (2021).
- [52] Haowei Zhu, Wenjing Ke, Dong Li, Ji Liu, Lu Tian, and Yi Shan. 2022. Dual cross-attention learning for fine-grained visual categorization and object re-identification. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 4692–4702.
- [53] Weiming Zhuang, Yonggang Wen, Xuesen Zhang, Xin Gan, Daiying Yin, Dongzhan Zhou, Shuai Zhang, and Shuai Yi. 2020. Performance optimization of federated person re-identification via benchmark analysis. In *Proceedings of the 28th ACM international conference on multimedia*. 955–963.