Clothing-Change Feature Augmentation for Person Re-Identification

Ke Han[1,2] Shaogang Gong[3] Yan Huang[1,2] Liang Wang[1,2] Tieniu Tan[1,2,4]
[1] Institute of Automation, Chinese Academy of Sciences (CASIA)
[2] University of Chinese Academy of Sciences (UCAS)
[3] Queen Mary University of London (QMUL) [4] Nanjing University
ke.han@cripac.ia.ac.cn, s.gong@qmul.ac.uk, {yhuang, wangliang, tnt}@nlpr.ia.ac.cn,

Abstract

Clothing-change person re-identification (CC Re-ID) aims to match the same person who changes clothes across cameras. Current methods are usually limited by the insufficient number and variation of clothing in training data, e.g. *each person only has 2 outfits in the PRCC dataset. In this work, we propose a novel Clothing-Change Feature Augmentation (CCFA) model for CC Re-ID to largely expand clothing-change data in the feature space rather than visual image space. It automatically models the feature distribution expansion that reflects a person's clothing colour and texture variations to augment model training. Specifically, to formulate meaningful clothing variations in the feature space, our method first estimates a clothing-change normal distribution with intra-ID cross-clothing variances. Then an augmentation generator learns to follow the estimated distribution to augment plausible clothing-change features. The augmented features are guaranteed to maximise the change of clothing and minimise the change of identity properties by adversarial learning to assure the effectiveness. Such augmentation is performed iteratively with an ID-correlated augmentation strategy to increase intra-ID clothing variations and reduce inter-ID clothing variations, enforcing the Re-ID model to learn clothing-independent features inherently. Extensive experiments demonstrate the effectiveness of our method with state-of-the-art results on CC Re-ID datasets.*

1. Introduction

Person re-identification (Re-ID) aims to match images of the same person across different locations over time. In early Re-ID methods, the target person was assumed to move within a short span of time and space, wearing the same clothes appearing in different camera views. Therefore most methods [20, 26, 28, 40, 41] leverage clothing information to identify persons. However, they cannot cope with the clothing-change situations, *e.g.* when a person

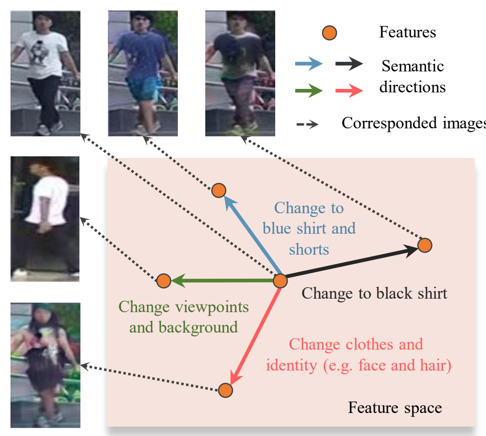


Figure 1. Transforming features towards specific directions can change specific semantics [37], *e.g.* clothing colours or textures, background, viewpoints and identities.

changes clothes significantly over a few days. To overcome this problem, more recent research has made an effort to consider clothing-change in model training and test, known as clothing-change person re-identification (CC Re-ID).

Current methods for CC Re-ID focus mostly on modeling clothing-independent identity information. They can be broadly divided into two categories: one-modality and multi-modality based methods. The one-modality methods learn clothing-independent representations solely from RGB images [7, 15]. The multi-modality methods exploit other auxiliary information, such as human silhouettes [13, 17], parsing masks [22, 27], keypoints [3, 25], 3D shape [2, 9], to help capture clothing-independent information like facial features and/or body shape characteristics.

However, existing methods' robustness to clothing variations is limited by the quite limited number and diversity of clothing in training data. For example, each person only has 2 outfits in the PRCC [39] dataset, which is insufficient to train a very robust Re-ID model. One direct approach is to synthesise more images of different clothes for each person using a generative model to augment training data. However, such direct data augmentation in the image space dramatically increases computational time and storage space, and its effectiveness on model generalisation is also not directly measurable. Moreover, due to the complexity of image synthesis, current generative Re-ID methods [21, 42] typically only model the exchange of clothes between two persons, and cannot generate plausible new clothes to expand the clothing-change library more freely.

To tackle the above problems, we propose a novel Clothing-Change Feature Augmentation (CCFA) model for CC Re-ID by augmenting implicitly clothing-change data in the feature space rather than image space. It aims to explore the plausible feature distribution expansion that reflects meaningful clothing colour and texture variations on a person' s appearance. Such feature augmentation is not only computationally more efficient, but also can expand significantly more new clothes that do not exist in the dataset in order to increase clothing-change variations in model training. Our work is motivated by recent findings that there exist many semantic directions in the deep feature space [36, 37]. Transforming a feature representation along specific directions can result in a representation corresponding to another image data sample of different semantics. For example in Fig.1, features can be transformed towards some directions to change information of clothing colours and textures, such as the blue shirt and shorts. We wish to explore such characteristics in CC Re-ID model training.

However, it remains challenging to properly implement clothing-change augmentation in the feature space. First, there are many semantic directions of feature expansion irrelevant to clothing, *e.g.* viewpoints and background in Fig.1, and there are also many semantically meaningless directions. It is nontrivial to find out meaningful clothing-change directions in order to maximise the diversity of clothing-change to benefit CC Re-ID model training. Critically, we do not have annotations for these feature augmentation directions. Second, changing clothes may damage the identity property, *i.e.* person-specific unique characteristics. For example, the red direction in Fig.1 changes both clothes and identity properties like the face and hair style, and makes the man changed to a different woman, causing a meaningless augmentation for the man. It is significant that a person's intrinsic identity property is maintained during feature augmentation in order to make meaningful clothing-change augmentation.

To address these challenges, we formulate *clothing-change ID-unchange feature augmentation learning* in our model. Specifically, our model first includes a clothing-change covariance estimation method to discover semantically meaningful clothing-change directions in the feature space. It statistically aggregates the intra-ID cross-clothing variances into a zero-mean multi-variate normal distribution, from which new plausible clothing variations can be formulated automatically. Then an augmentation generator is trained to generate feature augmentation. It is guaranteed not only to satisfy the estimated clothing-change directions (normal distribution), but also to maximise clothing-change and meanwhile minimise identity-change by adversarial learning to assure the effectiveness of augmentation.

Given this generator, we can iteratively augment clothing information on person features of each sample to expand clothing-change data in model training. To exploit the augmentation more efficiently, we further propose an ID-correlated augmentation strategy. Instead of augmenting each sample independently, we perform different augmentation for the samples of the same person, and the same augmentation for the samples of different persons in each mini-batch. This increases intra-ID clothing variations and reduces inter-ID clothing variations, enforcing the Re-ID model to automatically discover each person's clothing-independent unique (implicit identity) information more fully. Extensive experiments demonstrate that our method improves the model's accuracy significantly on CC Re-ID datasets PRCC [39] and LTCC [25].

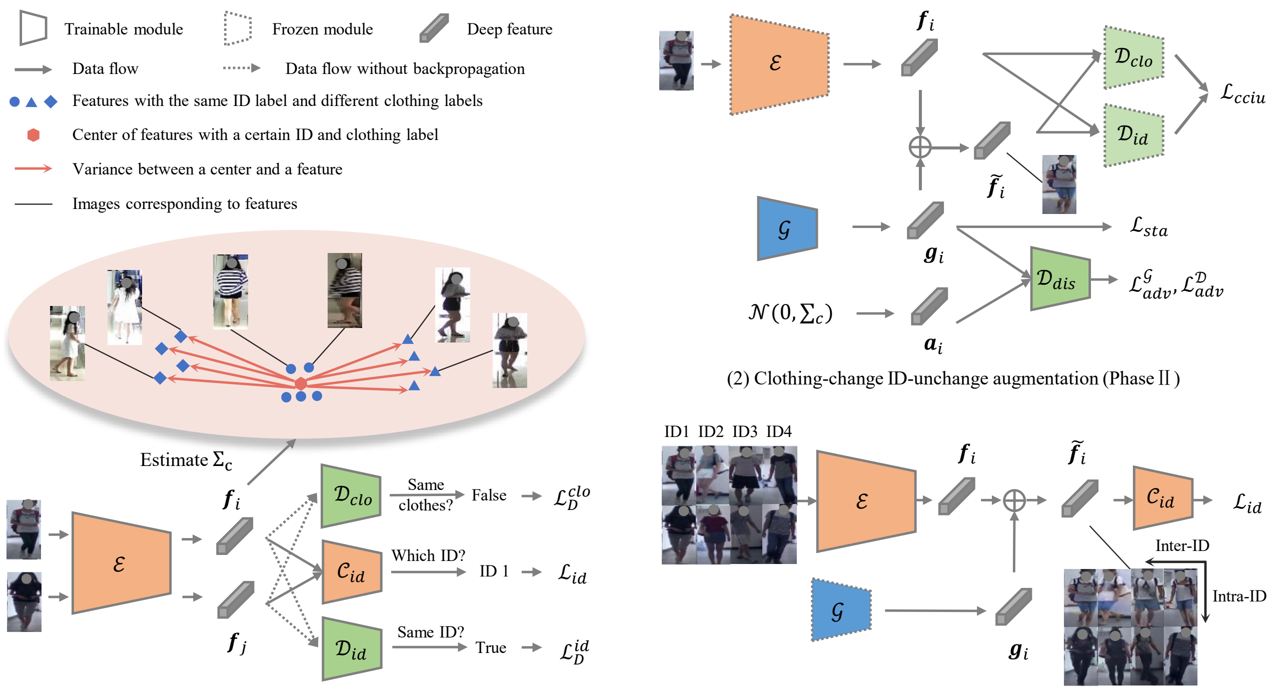
We summarise the contributions of this work as follows.

- For the first time we propose a CCFA model for CC Re-ID to implicitly augment clothing-change data in the feature space, by maximising clothing-change whilst minimising identity-change for person features.
- We present a clothing-change covariance estimation method to formulate clothing-change semantic directions of feature distribution expansion, and introduce an augmentation generator to implement the clothing-change ID-unchange augmentation.
- An ID-correlated augmentation strategy is proposed to increase intra-ID clothing variations and simultaneously to reduce inter-ID clothing variations, explicitly enforcing the Re-ID model to explore clothing-independent information more fully.
- Our method improves the model's robustness to clothing variations and achieves state-of-the-art results.

2. Related Work

Standard Person Re-Identification. Many standard Re-ID methods have been proposed to cope with various challenges, such as misalignment [32, 35], background bias [14, 29], occlusion [23, 31, 38], cross-resolution [8, 10], domain adaption [16, 30, 44], domain generalisation [4, 11, 24], etc. Due to primarily relying on clothing information to identify persons, these methods suffer severe performance drop when the target person changes clothes.

Clothing-Change Person Re-Identification. Clothing-change person Re-ID methods aim to learn clothing-independent person features to handle clothing variation-



(1) Clothing-change covariance estimation (Phase Ⅰ)

(2) Clothing-change ID-unchange augmentation (Phase Ⅱ)

(3) Training with the ID-correlated augmentation strategy (Phase Ⅲ)

Figure 2. A Clothing-Change Feature Augmentation (CCFA) Re-ID model consists of: (1) Estimating a clothing-change covariance vector Σ_c to formulate clothing variations in the feature space; (2) Following the normal distribution $\mathcal{N}(0, \Sigma_c)$ to learn clothing-change ID-unchange augmentation; (3) Training the Re-ID model with the ID-correlated augmentation strategy, *i.e.* different intra-ID augmentation and same inter-ID augmentation, to learn clothing-independent features.

s. These methods are divided into two classes, *i.e.* multi-modality and one-modality based methods. The multi-modality methods exploit other modalities as auxiliary information to RGB images for feature learning. Qian *et al.* [25] present a clothing-elimination shape-distillation model with human keypoints. Yang *et al.* [39] use human sketches to extract and aggregate body shape features under multiple angles. Similarly, the works in [13] and [17] transfer body shape knowledge extracted from human silhouettes to appearance features. Shu *et al.* [27] employ human parsing masks to divide human body parts and randomly exchange pixels of clothes among persons for data augmentation. In addition, human 3D body shape information is also modeled by exploring the depth information [9] or projecting into keypoints and silhouettes in a self-supervised manner [2]. However, these multi-modality methods do not make full use of the underlying clothing-irrelevant identity information in original RGB images, and are also potentially subject to the estimation quality of auxiliary modalities. To tackle the problem, the one-modality methods [7, 15] are proposed to directly mine clothing-independent features from RGB images without other modalities.

The above methods only learn identity features from the limited clothing-change patterns in existing training data. Differently, our feature augmentation method can massively expand new clothing-change data in model training to enhance clothing-independent feature learning effectively.

Semantic Transformations of Deep Features. Recent studies find that deep features learnt by convolutional networks are generally well linearised and potentially exhibit semantic transformation directions [1, 33]. Translating deep features along certain directions can change the semantic content of features, which has been exploited for face editing [34], changing labels [6], feature augmentation [19, 37], etc. Our method is motivated by feature augmentation methods but different in two main aspects. First, current methods [19, 37] blend different semantic transformation directions together and cannot model specific ones. In contrast, our method is specifically designed to address the challenge of how to formulate clothing-change and ID-unchange directions to benefit CC Re-ID most. Second, unlike other methods using statistical methods to roughly estimate plausible semantic directions, our method guarantees the semantic effectiveness of feature augmentation by adversarial learning in a learnable manner.

3. Method

In this work, we introduce a Clothing-Change Feature Augmentation (CCFA) model for CC Re-ID by optimising

Re-ID model training with feature distribution expansion, to improve the robustness to clothing variations without the need of collecting additional clothing-change training data.

Fig. 2 summarises our method which includes three training phases. (1) To formulate semantically meaningful clothing variations in a feature space, we model a zero-mean multi-variate normal distribution $\mathcal{N}(0, \Sigma_c)$, and statistically estimate a clothing-change covariance vector Σ_c (Section 3.2). (2) Following the estimated distribution, an augmentation generator is deployed for synthesising both clothing-change and ID-unchange feature augmentation to expand the feature space (Section 3.3). (3) The Re-ID model is then trained by an ID-correlated augmentation strategy to maximise the learning of clothing-independent features automatically (Section 3.4).

Let us define some notations at first. Suppose a Re-ID feature extractor $\mathcal{E}$ is trained on a CC Re-ID dataset $\mathcal{S} = \{(\boldsymbol{x}_i, y_i, c_i)\}_{i=1}^{N}$, where $\boldsymbol{x_i}$ is a random image with an identity label y_i and a clothing label c_i, which are provided in the dataset. The feature vector of $\boldsymbol{x}_i$ is denoted as $\boldsymbol{f}_i = \mathcal{E}(\boldsymbol{x}_i)$ and extracted by the Re-ID feature extractor $\mathcal{E}$.

3.1. Revisit Feature Augmentation

Recent studies suggest that certain directions of feature expansion in a deep learning feature space correspond to some specific semantic transformations [33, 36]. Transforming a feature representation along specific directions can result in a representation corresponding to another image sample of different semantics. This motivates some works [19, 37] to augment training data in the feature space. The feature augmentation is usually formulated as

$$\tilde{\boldsymbol{f}}_i = \boldsymbol{f}_i + \lambda \cdot \boldsymbol{a}_i, where \;\; \boldsymbol{a}_i \sim \mathcal{N}(0, \Sigma_c), \tag{1}$$

where $\tilde{\boldsymbol{f}}_i$ is the augmented feature vector from $\boldsymbol{f}_i$. An augmentation vector $\boldsymbol{a}_i$ is modeled as a random sampling from a zero-mean multi-variate normal distribution $\mathcal{N}$, Σ_c is a covariance vector, and λ is a weight factor.

Mathematically, Σ_c is supposed to be a covariance matrix. But in this work, we simplify and formulate Σ_c as a vector $(\sigma_1^2, \cdots, \sigma_n^2)$, which is the variance of feature vectors of samples. This covariance vector is equivalent to a diagonal covariance matrix of $(\sigma_1^2, \cdots, \sigma_n^2)$ with non-diagonal elements being zero. Following [37], the feature vector $\boldsymbol{f}_i$ is assumed to be an n-dim independent random variable, so that we only need to compute n diagonal elements (*i.e.* variances) of the covariance matrix and set the non-diagonal elements (*i.e.* covariances) to zero. This avoids computing full covariances of n dimensions due to enormous costs.

Existing methods [19, 37] have studied how to construct the covariance vector Σ_c to formulate plausible feature augmentation. They usually blend various semantic directions together without distinguishing specific ones to increase diverse variations of data. From a different perspective with a different objective, we wish to distinguish explicitly the clothing-change directions from other factors such as viewpoints or background (Fig. 1). Our objective is to maximise the diversity of clothing-change in Re-ID model training without increasing labelled training data. To that end, we introduce a clothing-change covariance estimation method.

3.2. Clothing-Change Covariance Estimation

To formulate clothing-change semantic directions, we exploit the cross-clothing variances of each person's features to estimate a clothing-change covariance vector. As shown in Fig. 2.(1), in the feature space we compute some "clothing centers", which are the average representations of each outfit of each person. For example, the clothing center $\boldsymbol{\mu}_{y_i}^{c_i}$ is the average of all the feature samples with the identity label y_i and the clothing label c_i. Then we compute the cross-clothing variances, which are variances between a clothing center and the feature samples with the same identity label but different clothing labels in each mini-batch, and incrementally aggregate them into a covariance vector by moving average. This process is formulated as

$$(\Sigma_c')_t = \mathbb{E}[\big((\boldsymbol{\mu}_t)_{y_i}^{c_i} - \boldsymbol{f}_{y_i}^{c_j}\big)^2] \;\; (i \neq j), \tag{2}$$

where $(\boldsymbol{\mu}_t)_{y_i}^{c_i}$ is the clothing center with the identity label y_i and clothing label c_i at the t-th step (when t mini-batches have run). $\boldsymbol{f}_{y_i}^{c_j}$ is a feature sample with the identity label y_i and clothing label c_j in the t-th mini-batch. $(\Sigma_c')_t$ is the covariance vector computed in the t-th mini-batch, which is used to update the covariance vector Σ_c by moving average. By statistically aggregating the intra-ID cross-clothing variances, Σ_c captures the clothing-change semantic directions of person features from the training data.

Identity classifier $\mathcal{C}_{id}$. To enable the Re-ID module $\mathcal{E}$ to have a basic ability of extracting discriminative person features, we feed features to an identity classifier $\mathcal{C}_{id}$ to predict the identity labels of features subject to both cross-entropy and triplet losses. That is, $\mathcal{C}_{id}$ is composed of a fully-connected layer and supervised by the identity loss $\mathcal{L}_{id}$:

$$\mathcal{L}_{id} = \mathcal{L}_{ce} + \alpha \cdot \mathcal{L}_{tri}, \tag{3}$$

where $\mathcal{L}_{ce}$ and $\mathcal{L}_{tri}$ are respectively the cross-entropy and triplet loss, and α is a weight factor.

Clothing discriminator $\mathcal{D}_{clo}$ & Identity discriminator $\mathcal{D}_{id}$. In addition to the identity classifier, we also pre-train a clothing discriminator and an identity discriminator, which will be used in Section 3.3 to learn clothing-change and ID-unchange augmentation.

Specifically, these two discriminators are binary classifiers for deciding whether a pair of feature vectors have the same clothing/identity label. $\mathcal{D}_{clo}$ and $\mathcal{D}_{id}$ are learnt by the clothing discriminator loss $\mathcal{L}_D^{clo}$ and identity discriminator loss $\mathcal{L}_D^{id}$ respectively, which are

$$\mathcal{L}_D^{clo} = -\mathbb{E}[\mathbb{1}\cdot\log(\mathcal{D}_{clo}(\boldsymbol{f}_i, \boldsymbol{f}_j))+(1-\mathbb{1})\cdot\log(1-\mathcal{D}_{clo}(\boldsymbol{f}_i, \boldsymbol{f}_j))], \tag{4}$$

$$\mathcal{L}_D^{id} = -\mathbb{E}[\mathbb{1}\cdot\log(\mathcal{D}_{id}(\boldsymbol{f}_i, \boldsymbol{f}_j))+(1-\mathbb{1})\cdot\log(1-\mathcal{D}_{id}(\boldsymbol{f}_i, \boldsymbol{f}_j))], \tag{5}$$

where $\mathbb{1}=1$ if $\boldsymbol{f}_i$ and $\boldsymbol{f}_j$ have the same clothing label in Eq. (4) or identity label in Eq. (5). Otherwise $\mathbb{1}=0$. As shown in Fig. 2.(1), the backpropagation from two discriminators to the feature extractor $\mathcal{E}$ is cut off, which ensures that the discriminators do not perturb the discriminative features learnt by $\mathcal{E}$ via the identity loss $\mathcal{L}_{id}$.

3.3. Clothing-Change ID-Unchange Augmentation

Given the estimated covariance vector Σ_c, previous works [19, 37] deploy Eq. (1) to perform feature augmentation. However, since Σ_c is estimated statistically as an approximation, directly augmenting features by $\mathcal{N}(0, \Sigma_c)$ cannot always produce semantically meaningful feature augmentation. The meaningful augmentation should meet two criteria: 1) the clothing information in person features is changed; 2) the identity property cannot be changed, *e.g.* to avoid the red direction in Fig. 1, so that the identity label remains valid. To implement such augmentation, we introduce a *clothing-change ID-unchange augmentation learning* method, composed of four components: (1) an augmentation generator $\mathcal{G}$ generates feature augmentation and (2) a distribution discriminator $\mathcal{D}_{dis}$ ensures the augmented features follow the estimated clothing-change directions (normal distribution); (3) a clothing discriminator and (4) an identity discriminator guarantee the augmentation is clothing-change and ID-unchange, respectively.

Clothing-change ID-unchange loss $\mathcal{L}_{cciu}$. Given a random vector $\boldsymbol{z}_i$ as an input, $\mathcal{G}$ generates an augmentation vector $\boldsymbol{g}_i = \mathcal{G}(\boldsymbol{z}_i)$ for the feature vector $\boldsymbol{f}_i$. To guarantee that $\boldsymbol{g}_i$ meets the above two criteria, we present a clothing-change ID-unchange loss $\mathcal{L}_{cciu}$ formulated as

$$\mathcal{L}_{cciu} = -\mathbb{E}[\log(1-\mathcal{D}_{clo}(\boldsymbol{f}_i, \tilde{\boldsymbol{f}}_i))+\log(\mathcal{D}_{id}(\boldsymbol{f}_i, \tilde{\boldsymbol{f}}_i)], \tag{6}$$

where $\tilde{\boldsymbol{f}}_i = \boldsymbol{f}_i + \lambda \cdot \boldsymbol{g}_i$ indicates the augmented feature by $\mathcal{G}$. $\mathcal{L}_{cciu}$ pushes $\mathcal{G}$ to generate the augmentation vector $\boldsymbol{g}_i$ that maximises clothing-change whilst minimising identity-change. $\mathcal{D}_{clo}$ and $\mathcal{D}_{id}$ are pre-trained in Section 3.2 and the parameters are frozen here. Otherwise $\mathcal{D}_{clo}$ and $\mathcal{D}_{id}$ would ultimately output constant 0 and 1 respectively via $\mathcal{L}_{cciu}$, degenerating their ability of discriminating features.

Adversarial loss $\mathcal{L}_{adv}^{\mathcal{G}}$ and $\mathcal{L}_{adv}^{\mathcal{D}}$. The augmentation vector $\boldsymbol{g}_i$ is supposed to approximately satisfy the empirical clothing-change distribution from the training data, *i.e.* $\boldsymbol{g}_i \simeq \mathcal{N}(0, \Sigma_c)$. To ensure that, we supervise the distribution of $\boldsymbol{g}_i$ by an distribution discriminator $\mathcal{D}_{dis}$ where $\mathcal{D}_{dis}$ distinguishes an augmentation vector either drawn from the normal distribution $\mathcal{N}$ or generated by the generator $\mathcal{G}$.

(1) $\boldsymbol{f}_i$ (2) $\tilde{\boldsymbol{f}}_i$ with ID-uncorrelated augmentation (3) $\tilde{\boldsymbol{f}}_i$ with ID-correlated augmentation

Figure 3. Images corresponding to original features $\boldsymbol{f}_i$ and augmented features $\tilde{\boldsymbol{f}}_i$ on PRCC, visualised by the generative model [42]. Vertically aligned examples have the same identity label, and horizontally aligned examples have different identity labels.

$\mathcal{D}_{dis}$ is optimised by an adversarial loss $\mathcal{L}_{adv}^{\mathcal{D}}$:

$$\mathcal{L}_{adv}^{\mathcal{D}} = -\mathbb{E}[\log(\mathcal{D}_{dis}(\boldsymbol{a}_i)) + \log(1 - \mathcal{D}_{dis}(\boldsymbol{g}_i))], \tag{7}$$

where $\boldsymbol{a}_i$ and $\boldsymbol{g}_i$ indicate the augmentation vector is from $\mathcal{N}$ and $\mathcal{G}$, respectively. Accordingly, the adversarial loss $\mathcal{L}_{adv}^{\mathcal{G}}$ is employed on the generator $\mathcal{G}$ to make $\mathcal{G}$ generate the normal-like distribution to confuse $\mathcal{D}_{dis}$:

$$\mathcal{L}_{adv}^{\mathcal{G}} = \mathbb{E}[-\log(\mathcal{D}_{dis}(\boldsymbol{g}_i))]. \tag{8}$$

Statistic loss $\mathcal{L}_{sta}$. We further make sure that $\boldsymbol{g}_i$ can approximate $\mathcal{N}(0, \Sigma_c)$ by supervising the statistics. Specifically, $\boldsymbol{g}_i$ should have zero as prior mean and Σ_c as prior variance, which is guaranteed by a statistic loss $\mathcal{L}_{sta}$:

$$\mathcal{L}_{sta} = \|\mathbb{E}[\boldsymbol{g}_i]\|_1 + \|\mathbb{E}[(\mathbb{E}[\boldsymbol{g}_i] - \boldsymbol{g}_i)^2] - \Sigma_c\|_1, \tag{9}$$

where $\mathbb{E}[\boldsymbol{g}_i]$ and $\mathbb{E}[(\mathbb{E}[\boldsymbol{g}_i] - \boldsymbol{g}_i)^2]$ are the statistic mean and variance of $\boldsymbol{g}_i$, respectively.

Total loss $\mathcal{L}_{\mathcal{G}}$. The generator $\mathcal{G}$ is jointly trained with the weighted sum of the losses $\mathcal{L}_{adv}^{\mathcal{G}}$, $\mathcal{L}_{sta}$ and $\mathcal{L}_{cciu}$, which is formulated as

$$\mathcal{L}_{\mathcal{G}} = \mathcal{L}_{adv}^{\mathcal{G}} + \gamma_1 \cdot \mathcal{L}_{sta} + \gamma_2 \cdot \mathcal{L}_{cciu}. \tag{10}$$

It ensures that $\mathcal{G}$ generates the augmentation vector $\boldsymbol{g}_i$ that not only satisfies the estimated clothing-change normal distribution $\mathcal{N}(0, \Sigma_c)$, but also maximises clothing-change and minimises identity-change in person features.

3.4. ID-Correlated Augmentation Strategy

Having trained the augmentation generator $\mathcal{G}$, we resume to train the Re-ID feature extractor $\mathcal{E}$ using $\mathcal{G}$ to augment clothing features iteratively. A direct augmentation strategy is to perform clothing-change augmentation on each feature sample independently, *i.e.* each sample is augmented into different clothes randomly, as visualised in Fig. 3.(2). To exploit the augmentation more efficiently for CC Re-ID, we propose an ID-correlated augmentation strategy. We show the implementation in Algorithm 1 (Phase III) and visualisation explanation in Fig. 3.(3). Specifically, we perform

Algorithm 1 Training procedure of the CCFA model

Input: $M \cdot J$ images, where M and J are the numbers of persons and images per person in a mini-batch.

for *each mini-batch* **do**
  Extract features by $\mathcal{E}$.
  if ***Phase I*** **then**
    Estimate covariance vector Σ_c via Eq. (2).
    Optimise $\mathcal{E}$ and $\mathcal{C}_{id}$ via Eq. (3), $\mathcal{D}_{clo}$ via Eq. (4), and $\mathcal{D}_{id}$ via Eq. (5).
  else if ***Phase II*** **then**
    Freeze $\mathcal{E}$, $\mathcal{D}_{clo}$ and $\mathcal{D}_{id}$.
    Optimise $\mathcal{G}$ via Eq. (10) and $\mathcal{D}_{dis}$ via Eq. (7).
  else if ***Phase III*** **then**
    Freeze $\mathcal{G}$.
    for j in range(J) **do**
      $\mathcal{G}$ generates a random augmentation vector $\boldsymbol{g}_j$.
      Perform ID-correlated augmentation:
      $\tilde{\boldsymbol{f}}_{1j} = \boldsymbol{f}_{1j} + \lambda \cdot \boldsymbol{g}_j, \cdots, \tilde{\boldsymbol{f}}_{Mj} = \boldsymbol{f}_{Mj} + \lambda \cdot \boldsymbol{g}_j$,
      where $\boldsymbol{f}_{mj}$ is the feature vector of the j-th sample of the m-th person in a mini-batch, and $\tilde{\boldsymbol{f}}_{mj}$ is the augmented feature vector.
    end
    Optimise $\mathcal{E}$ and $\mathcal{C}_{id}$ via Eq. (3).

end

different augmentation on the samples of the **same** person (vertically aligned examples), and the **same** augmentation on the samples of **different** persons (horizontally aligned examples) in a mini-batch. Here the same augmentation means using the same augmentation vector $\boldsymbol{g}_i$. For example in Fig. 3.(3), the images on the first line are all augmented into black shirts and blue shorts, and the images on the second line are all augmented into green shirts and black pants. Through this way, the intra-ID clothing variations are increased and the inter-ID clothing variations are reduced. This prevents some specific clothing patterns overfitting to a specific person and enforces the Re-ID model to explore clothing-independent unique identity characteristics of each person to distinguish persons.

As shown in Fig. 2.(3), the Re-ID feature extractor $\mathcal{E}$ and identity classifier $\mathcal{C}_{id}$ pre-trained in Phase I proceed to be trained with feature augmentation and optimised by the identity loss $\mathcal{L}_{id}$ in Eq. (3). This enables $\mathcal{E}$ to extract person features robust to the augmented clothing variations.

3.5. Training and Test

The training procedure includes three phases, which are summarised in Algorithm 1. We use the optimised $\mathcal{E}$ to extract features when test, and the Euclidean distance to measure the similarities between probe and gallery images.

4. Experiments

4.1. Datasets and Protocols

We conduct experiments on publicly available CC Re-ID datasets PRCC [39] and LTCC [25]. PRCC contains 33,698 images of 221 persons captured by 3 cameras, and each person has 2 outfits. LTCC has 17,119 images of 152 persons, and each person is captured by at least 2 cameras, with the number of outfits ranging from 2 to 14.

Following the established evaluation protocol [7, 25, 39], three test modes are defined as follows: 1) **clothing-change (CC) mode** (only clothing-change ground-truth samples are used to calculate accuracy), 2) **same-clothing (SC) mode** (only clothing-consistent ground-truth samples are used to calculate accuracy), and 3) **general mode** (both clothing-change and clothing-consistent ground-truth samples are used to calculate accuracy). We use the average cumulative match characteristic and report accuracies in CC and SC modes on PRCC, CC and general modes on LTCC.

4.2. Implementation Details

The Re-ID feature extractor $\mathcal{E}$ is a ResNet50 [12] backbone pre-trained on ImageNet [5]. The structures of other modules are introduced in the supplementary material. All the images are resized to 384×128. Random horizontal flipping and random erasing [43] are used for data augmentation. The Adam optimiser [18] ($\beta_1 = 0.9$ and $\beta_2 = 0.999$) is adopted. The learning rates of $\mathcal{E}$ and $\mathcal{G}$ are 0.001, the learning rates of $\mathcal{C}_{id}$, $\mathcal{D}_{clo}$, $\mathcal{D}_{id}$ and $\mathcal{D}_{dis}$ are all 0.01. We set the numbers of training epochs in phase I, II and III as 30, 30 and 80, respectively. The hyper-parameters are set as follows. The weight factor α=1, γ_1=1, γ_2=0.1. The margin parameter of the triplet loss $\mathcal{L}_{tri}$ is 0.3. λ is uniformly sampled from the range 10±0.5 (*i.e.* [9.5, 10.5]) in each mini-batch. We train our model on 2 NVIDIA Titan Xp GPUs, with 16 randomly sampled persons and 8 images per person in each mini-batch.

4.3. Comparison with State-of-the-Art Methods

CC Re-ID methods. We compare our CCFA model with state-of-the-art CC Re-ID models on PRCC and LTCC in Table 1. On LTCC, CCFA surpasses all the other methods obviously in both CC and general test modes. On PRCC, in the CC mode CCFA achieves comparable accuracies to the first-rank model SPS [27]. In the SC mode, the 98.7% mAP and 99.6% rank 1 accuracies of CCFA are already close to saturation. A slight disadvantage to CAL [7] is because CCFA aims to learn clothing-independent features but there are only clothing-consistent ground-truth samples in this mode.

Our method has two major advantages over other CC Re-ID methods. First, other methods only learn representations from the existing clothing-change data in the dataset. Our method can synthesise abundant semantically

Table 1. Evaluations on the PRCC and LTCC datasets (%), where "sketch", "pose", "sil.", "parsing" and "3D" denote respectively contour sketches, keypoints, silhouettes, human parsing and 3D shape information. Bold and underlined numbers are the top two scores.

Method Type	Method	Modality	PRCC [39]				LTCC [25]			
			CC Mode		SC Mode		CC Mode		General Mode	
			mAP	Rank 1	mAP	Rank 1	mAP	Rank 1	mAP	Rank 1
Feature Augmentation	SFA [19]	RGB	47.8	49.6	94.8	98.3	11.8	34.8	33.6	61.7
	IDSA [37]	RGB	49.1	50.2	95.6	98.6	12.2	34.2	33.9	64.6
CC Re-ID	CESD [25]	RGB+pose	-	-	-	-	12.4	26.1	34.3	71.4
	SPT+ASE [39]	Sketch	-	34.4	-	64.2	-	-	-	-
	3DSL [2]	RGB+pose+sil.+3D	51.3	-	-	-	14.8	31.2	-	-
	FSAM [13]	RGB+pose+sil.	-	54.5	-	98.8	16.2	38.5	35.4	73.2
	SPS [27]	RGB+parsing	57.2	**62.8**	96.7	99.5	16.7	42.1	37.6	70.9
	RCSANet [15]	RGB	50.2	48.6	97.2	**100**	-	-	-	-
	GI-ReID [17]	RGB+sil.	37.5	-	-	-	10.4	23.7	29.4	63.2
	CAL [7]	RGB	55.8	55.2	**99.8**	**100**	18.0	40.1	40.8	74.2
	CCFA (Phase I)	RGB	47.5	48.1	95.3	98.0	11.4	33.8	30.6	65.7
	CCFA (Phase III)	RGB	**58.4**	61.2	98.7	99.6	**22.1**	**45.3**	**42.5**	**75.8**

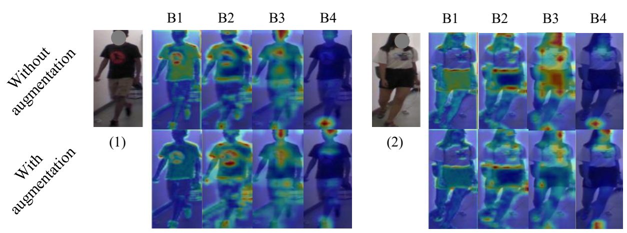


Figure 4. Activated feature maps on PRCC in Phase I and III. B1, B2, B3 and B4 indicate that feature maps are extracted from four blocks respectively in the ResNet50 backbone of the CCFA model.

meaningful new clothing variations in the feature space to augment model training without the need of more labelled training data, and effectively improve the Re-ID model's robustness to clothing variations. Second, most methods [2, 13, 17, 25, 27, 39] need auxiliary training data of other modalities, such as human parsing, keypoints, silhouettes, etc. In contrast, our model only uses RGB images, which is not only computationally less expensive, but also free of estimation errors from the auxiliary modalities.

Feature augmentation methods. To our best knowledge, feature augmentation has not been investigated in clothing-change Re-ID. We compare two feature augmentation models SFA [19] and IDSA [37] designed for the classification task in Table 1. CCFA outperforms them by a big margin on both PRCC and LTCC, due to the following two reasons. First, the two methods blend different semantic augmentation directions together, while our method specifically models clothing-change ID-unchange directions to maximise the diversity of clothing to benefit CC Re-ID most. Second, unlike them only estimating covariance with rough statistic methods, our method formulates an learnable generator to assure the effectiveness of augmentation.

4.4. Ablation Study

Phase I (without augmentation) v.s. Phase III (with augmentation). ***1) Accuracy.*** Table 1 lists the accuracies of our CCFA model in Phase I and Phase III. CCFA (Phase I) is trained in Phase I without feature augmentation and regarded as the baseline. CCFA (Phase III) is further trained from CCFA (Phase I) in Phase III with feature augmentation, and improves the accuracies significantly over the baseline on both PRCC and LTCC, verifying the effectiveness of feature augmentation.

2) Activated feature maps. Fig. 4 compares activated feature maps of two examples in Phase I and III. On the B1 (block 1) feature maps, Phase III weakens the activations on the clothing regions, such as the shirt in the example A and skirt in the example B. On the B4 (block 4) feature maps, activations on faces and shoulders are strengthened obviously in Phase III. This indicates that feature augmentation makes the model focus more on clothing-independent identity characteristics.

Table 2. Rank 1 evaluated by generating augmentation vectors from the normal distribution $\mathcal{N}$ or generator $\mathcal{G}$ on PRCC (%).

Augmentation Generation	CC Mode	SC Mode
Baseline (w/o augmentation)	48.1	98.0
With $\mathcal{N}(0, \Sigma_c)$	57.6	99.2
With $\mathcal{G}$ (ours)	**61.2**	**99.6**

Clothing-change covariance Σ_c & Augmentation generator $\mathcal{G}$. We demonstrate the effectiveness of the estimated clothing-change covariance vector Σ_c and augmentation generator $\mathcal{G}$ by experimental results in Table 2. The "baseline" does not perform feature augmentation. "With $\mathcal{N}(0, \Sigma_c)$" indicates performing feature augmentation with augmentation vectors from $\mathcal{N}(0, \Sigma_c)$, which achieves the much higher accuracy than the baseline, showing Σ_c indeed formulates some effective feature augmentation. "With $\mathcal{G}$" uses the generator $\mathcal{G}$ to generate augmentation vectors, and further improves the accuracy by 3.6% and 0.4% in the CC and SC mode, respectively. Although also basically satisfying $\mathcal{N}(0, \Sigma_c)$, $\mathcal{G}$ further strengthens the effectiveness

Table 3. Rank 1 evaluated by ID-uncorrelated and ID-correlated augmentation strategies on PRCC (%).

Augmentation Strategy	CC Mode	SC Mode
Baseline	48.1	98.0
ID-uncorrelated	55.5	98.6
ID-correlated (ours)	**61.2**	**99.6**

Table 4. Evaluations of loss functions in Phase II on PRCC. The shown results are rank 1 scores (%).

$\mathcal{L}_{cciu}$	$\mathcal{L}_{sta}$	$\mathcal{L}_{adv}^{\mathcal{G}}$	$\mathcal{L}_{adv}^{\mathcal{D}}$	CC Mode	SC Mode
✗	✓	✓	✓	57.0	99.1
Use $\mathcal{D}_{clo}$ term only	✓	✓	✓	59.3	99.5
Use $\mathcal{D}_{id}$ term only	✓	✓	✓	58.4	99.2
✓	✗	✓	✓	56.7	96.3
✓	✓	✗	✓	37.4	71.7
✓	✓	✓	✗	44.5	82.5
✓	✓	✗	✗	49.7	92.3
✓	✓	✓	✓	**61.2**	**99.6**

of augmentation by semantically guaranteeing the clothing-change and ID-unchange augmentation.

ID-correlated augmentation. We compare two feature augmentation strategies in Phase III, *i.e.* ID-uncorrelated and ID-correlated strategies. As shown in Fig. 3, ID-uncorrelated means randomly augmenting all the images independently to different clothes in a mini-batch, while ID-correlated means performing different augmentation for a person's different images and sharing it among persons. The experimental results in Table 3 suggest that ID-uncorrelated augmentation can improve the rank 1 accuracy over the baseline, benefiting from expanding the diversity of clothing. ID-correlated augmentation further exceeds ID-uncorrelated augmentation by 5.7% and 1.0% in the CC and SC mode, respectively. It explicitly increases intra-ID clothing variations and reduces inter-ID clothing variations, which can effectively prevent the model relying on specific clothing information to identify persons and help learn clothing-independent unique identity characteristics.

Loss functions. Table 4 lists the ablation study of each loss function. 1) Removing $\mathcal{L}_{cciu}$ or only leaving the $\mathcal{D}_{clo}$/$\mathcal{D}_{id}$ term of $\mathcal{L}_{cciu}$ in Eq. (6) degenerates the accuracy to different degrees. This suggests that maximising clothing-change and minimising identity-change are both nontrivial for meaningful clothing-change feature augmentation. 2) Turning off $\mathcal{L}_{sta}$ or $\mathcal{L}_{adv}^{\mathcal{G}}$ decreases the accuracy, because it makes $\mathcal{G}$ not fit the estimated normal distribution well. Using $\mathcal{L}_{adv}^{\mathcal{G}}$ solely cannot guarantee that the generated distribution has the prior mean of 0 and variance of Σ_c. With the statistic loss solely, the generated distribution does not fit the normal distribution $\mathcal{N}$, as many other distributions have such statistics. 3) Taking off $\mathcal{L}_{adv}^{\mathcal{D}}$ also results in severe performance drop. Similar to $\mathcal{L}_{adv}^{\mathcal{G}}$, $\mathcal{L}_{adv}^{\mathcal{D}}$ is essential to supervise $\mathcal{G}$ to fit the estimated normal distribution by adversarial learning. 4) Considering that only removing $\mathcal{L}_{adv}^{\mathcal{G}}$ or $\mathcal{L}_{adv}^{\mathcal{D}}$ may lead to degradation of the generator itself, we also remove them together. This undoubtedly reduces the accuracy due to disenabling the adversarial learning.

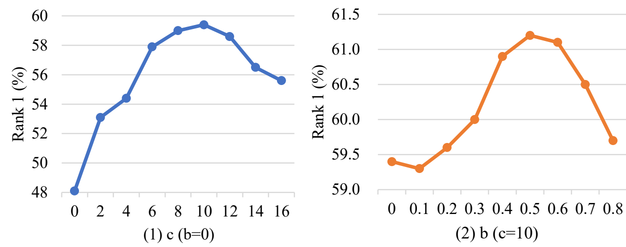


Figure 5. Effect of the weight factor λ on PRCC in the CC mode. λ is uniformly sampled from the range $c \pm b$ (*i.e.* $[c-b, c+b]$).

Weight factor λ. λ is the weight factor controlling the balance between the original feature vector $\boldsymbol{f}_i$ and the augmentation vector $\boldsymbol{g}_i$. We uniformly sample λ from a range $c \pm b$, (*i.e.* $[c-b, c+b]$) in each mini-batch instead of setting it to a fixed number, so that the feature augmentation can be scalable relatively. As shown in Fig. 5, at first we adjust the center c of the range by setting b to 0, and find that the rank 1 accuracy reaches the top when c=10. Then we fine-tune the range around 10 by adjusting the bias b of the range and fixing c at 10. The highest rank 1 accuracy is achieved when b=0.5, so the range is finally set to 10 ± 0.5.

5. Conclusion

To our best knowledge, for the first time we have formulated a CC Re-ID model by feature distribution expansion to implicitly synthesise semantically meaningful clothing-change augmentation in the feature space. First, we estimated statistically a clothing-change covariance from the training data to formulate clothing-change semantic directions in the feature space. Second, an augmentation generator was introduced to simulate the estimated distribution and further guarantee the clothing-change and ID-unchange augmentation by adversarial learning. Finally, we introduced an ID-correlated strategy to improve the efficiency of feature augmentation. Overall, our method has significantly enhanced Re-ID model robustness to clothing variations.

Acknowledge

This work was jointly supported by National Key R&D Program of China (2022ZD0117900), National Natural Science Foundation of China (62236010, 62276261, and 61721004), Key Research Program of Frontier Sciences CAS Grant No.ZDBS-LYJSC032, the Alan Turing Institute Turing Fellowship, Vision Semantics, and Veritone.

References

[1] Yoshua Bengio, Grégoire Mesnil, Yann Dauphin, and Salah Rifai. Better mixing via deep representations. In *ICML*, 2013. 3

[2] Jiaxing Chen, Xinyang Jiang, Fudong Wang, Jun Zhang, Feng Zheng, Xing Sun, and Wei-Shi Zheng. Learning 3d shape feature for texture-insensitive person re-identification. In *CVPR*, 2021. 1, 3, 7

[3] Zerui Chen, Yan Huang, Hongyuan Yu, Bin Xue, Ke Han, Yiru Guo, and Liang Wang. Towards part-aware monocular 3d human pose estimation: An architecture search approach. In *ECCV*, 2020. 1

[4] Seokeon Choi, Taekyung Kim, Minki Jeong, Hyoungseob Park, and Changick Kim. Meta batch-instance normalization for generalizable person re-identification. In *CVPR*, 2021. 2

[5] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *CVPR*, 2009. 6

[6] Jacob R Gardner, Paul Upchurch, Matt J Kusner, Yixuan Li, Kilian Q Weinberger, Kavita Bala, and John E Hopcroft. Deep manifold traversal: Changing labels with convolutional features. *arXiv preprint arXiv:1511.06421*, 2015. 3

[7] Xinqian Gu, Hong Chang, Bingpeng Ma, Shutao Bai, Shiguang Shan, and Xilin Chen. Clothes-changing person re-identification with rgb modality only. In *CVPR*, 2022. 1, 3, 6, 7

[8] Ke Han, Yan Huang, Zerui Chen, Liang Wang, and Tieniu Tan. Prediction and recovery for adaptive low-resolution person re-identification. In *ECCV*, 2020. 2

[9] Ke Han, Yan Huang, Shaogang Gong, Liang Wang, and Tieniu Tan. 3d shape temporal aggregation for video-based clothing-change person re-identification. In *ACCV*, 2022. 1, 3

[10] Ke Han, Yan Huang, Chunfeng Song, Liang Wang, and Tieniu Tan. Adaptive super-resolution for person re-identification with low-resolution images. *PR*, 2021. 2

[11] Ke Han, Chenyang Si, Yan Huang, Liang Wang, and Tieniu Tan. Generalizable person re-identification via self-supervised batch norm test-time adaption. In *AAAI*, 2022. 2

[12] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *CVPR*, 2016. 6

[13] Peixian Hong, Tao Wu, Ancong Wu, Xintong Han, and Wei-Shi Zheng. Fine-grained shape-appearance mutual learning for cloth-changing person re-identification. In *CVPR*, 2021. 1, 3, 7

[14] Yan Huang, Qiang Wu, JingSong Xu, and Yi Zhong. Sbsgan: Suppression of inter-domain background shift for person re-identification. In *ICCV*, 2019. 2

[15] Yan Huang, Qiang Wu, JingSong Xu, Yi Zhong, and ZhaoXiang Zhang. Clothing status awareness for long-term person re-identification. In *ICCV*, 2021. 1, 3, 7

[16] Zhipeng Huang, Zhizheng Zhang, Cuiling Lan, Wenjun Zeng, Peng Chu, Quanzeng You, Jiang Wang, Zicheng Liu, and Zheng-Jun Zha. Lifelong unsupervised domain adaptive person re-identification with coordinated anti-forgetting and adaptation. In *CVPR*, 2022. 2

[17] Xin Jin, Tianyu He, Kecheng Zheng, Zhiheng Yin, Xu Shen, Zhen Huang, Ruoyu Feng, Jianqiang Huang, Zhibo Chen, and Xian-Sheng Hua. Cloth-changing person re-identification from a single image with gait prediction and regularization. In *CVPR*, 2022. 1, 3, 7

[18] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *ICLR*, 2015. 6

[19] Pan Li, Da Li, Wei Li, Shaogang Gong, Yanwei Fu, and Timothy M Hospedales. A simple feature augmentation for domain generalization. In *ICCV*, 2021. 3, 4, 5, 7

[20] Wei Li, Xiatian Zhu, and Shaogang Gong. Harmonious attention network for person re-identification. In *CVPR*, 2018. 1

[21] Yu-Jhe Li, Zhengyi Luo, Xinshuo Weng, and Kris M Kitani. Learning shape representations for clothing variations in person re-identification. *WACV*, 2020. 2

[22] Xiaodan Liang, Ke Gong, Xiaohui Shen, and Liang Lin. Look into person: Joint body parsing and pose estimation network and a new benchmark. *TPAMI*, 2019. 1

[23] Jiaxu Miao, Yu Wu, Ping Liu, Yuhang Ding, and Yi Yang. Pose-guided feature alignment for occluded person re-identification. In *ICCV*, 2019. 2

[24] Hao Ni, Jingkuan Song, Xiaopeng Luo, Feng Zheng, Wen Li, and Heng Tao Shen. Meta distribution alignment for generalizable person re-identification. In *CVPR*, 2022. 2

[25] Xuelin Qian, Wenxuan Wang, Li Zhang, Fangrui Zhu, Yanwei Fu, Tao Xiang, Yu-Gang Jiang, and Xiangyang Xue. Long-term cloth-changing person re-identification. *ACCV*, 2020. 1, 2, 3, 6, 7

[26] Yantao Shen, Hongsheng Li, Tong Xiao, Shuai Yi, Dapeng Chen, and Xiaogang Wang. Deep group-shuffling random walk for person re-identification. In *CVPR*, 2018. 1

[27] Xiujun Shu, Ge Li, Xiao Wang, Weijian Ruan, and Qi Tian. Semantic-guided pixel sampling for cloth-changing person re-identification. *IJIS*, 2021. 1, 3, 6, 7

[28] Jianlou Si, Honggang Zhang, Chun-Guang Li, Jason Kuen, Xiangfei Kong, Alex C Kot, and Gang Wang. Dual attention matching network for context-aware feature sequence based person re-identification. In *CVPR*, 2018. 1

[29] Chunfeng Song, Yan Huang, Wanli Ouyang, and Liang Wang. Mask-guided contrastive attention model for person re-identification. In *CVPR*, 2018. 2

[30] Jifei Song, Yongxin Yang, Yi-Zhe Song, Tao Xiang, and Timothy M Hospedales. Generalizable person re-identification by domain-invariant mapping network. In *CVPR*, 2019. 2

[31] Yifan Sun, Qin Xu, Yali Li, Chi Zhang, Yikang Li, Shengjin Wang, and Jian Sun. Perceive where to focus: Learning visibility-aware part-level features for partial person re-identification. In *CVPR*, 2019. 2

[32] Yifan Sun, Liang Zheng, Yi Yang, Qi Tian, and Shengjin Wang. Beyond part models: Person retrieval with refined part pooling (and a strong convolutional baseline). In *ECCV*, 2018. 2

[33] Paul Upchurch, Jacob Gardner, Geoff Pleiss, Robert Pless, Noah Snavely, Kavita Bala, and Kilian Weinberger. Deep feature interpolation for image content changes. In *CVPR*, 2017. 3, 4

[34] Paul Upchurch, Jacob Gardner, Geoff Pleiss, Robert Pless, Noah Snavely, Kavita Bala, and Kilian Weinberger. Deep feature interpolation for image content changes. In *CVPR*, 2017. 3

[35] Guanshuo Wang, Yufeng Yuan, Xiong Chen, Jiwei Li, and Xi Zhou. Learning discriminative features with multiple granularities for person re-identification. In *ACM MM*, 2018. 2

[36] Yulin Wang, Gao Huang, Shiji Song, Xuran Pan, Yitong Xia, and Cheng Wu. Regularizing deep networks with semantic data augmentation. *TPAMI*, 2021. 2, 4

[37] Yulin Wang, Xuran Pan, Shiji Song, Hong Zhang, Gao Huang, and Cheng Wu. Implicit semantic data augmentation for deep networks. *NeurIPS*, 2019. 1, 2, 3, 4, 5, 7

[38] Zhikang Wang, Feng Zhu, Shixiang Tang, Rui Zhao, Lihuo He, and Jiangning Song. Feature erasing and diffusion network for occluded person re-identification. In *CVPR*, 2022. 2

[39] Qize Yang, Ancong Wu, and Wei-Shi Zheng. Person re-identification by contour sketch under moderate clothing change. *TPAMI*, 2019. 1, 2, 3, 6, 7

[40] Haiyu Zhao, Maoqing Tian, Shuyang Sun, Jing Shao, Junjie Yan, Shuai Yi, Xiaogang Wang, and Xiaoou Tang. Spindle net: Person re-identification with human body region guided feature decomposition and fusion. In *CVPR*, 2017. 1

[41] Liang Zheng, Liyue Shen, Lu Tian, Shengjin Wang, Jingdong Wang, and Qi Tian. Scalable person re-identification: A benchmark. In *ICCV*, 2015. 1

[42] Zhedong Zheng, Xiaodong Yang, Zhiding Yu, Liang Zheng, Yi Yang, and Jan Kautz. Joint discriminative and generative learning for person re-identification. In *CVPR*, 2019. 2, 5

[43] Zhun Zhong, Liang Zheng, Guoliang Kang, Shaozi Li, and Yi Yang. Random erasing data augmentation. In *AAAI*, 2020. 6

[44] Zhun Zhong, Liang Zheng, Zhiming Luo, Shaozi Li, and Y-i Yang. Invariance matters: Exemplar memory for domain adaptive person re-identification. In *CVPR*, 2019. 2