# Chapter 1

# Introduction

Person searches have received a lot of interest in a variety of contexts, including private amusement and public safety. However, it might be difficult to identify individual person objects in a group of identified person objects in gallery photographs. Because of the presence of different viewpoint [1],[2], varying low image resolutions [3],[4], illumination change [5], unconstrained poses [6],[7],[8], occlusions [9],[10], heterogeneous modalities [11],[12], complex camera environments, background clutter [13], unreliable bounding box generations, and so on, Re-ID is a difficult task. As a result, there are many variables and ambiguity. Furthermore, for practical model deployment, the dynamically updated camera network [14], [15], large scale gallery with efficient retrieval [16], group uncertainty [17], significant domain shift [18], unanticipated testing scenarios [19], incremental model updating [20], and changing clothes [21] all add to the difficulties. Because of these difficulties, Re-ID remains an unresolved topic. The major goal of the current approaches to person search is to use manually annotated identification labels to train deep networks to include a unique characteristic for each person's identity in a supervised way. Human labeling takes a significant amount of time and effort, often results in incorrect annotation, and reduces the overall effectiveness of person search. A few efforts have been made to give these unlabeled people already valid labels. They still use the supervised learning paradigm, and how well they do is mostly based on the labeled data. In this paper, we define a new person search issue without a named person. It identifies and suggests a context. -aware clustering -trained end-to-end network techniques for assembling people who are thought to have the same identity [61].

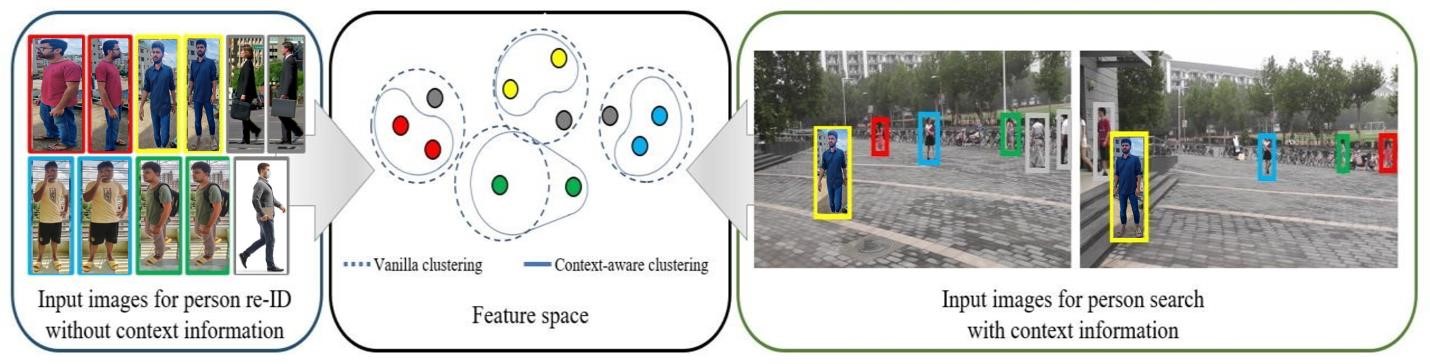


Figure 1: Comparison of the suggested feature-based clustering with situational priors and the traditional feature-based clustering.

We specifically just use context features of co-appearance, which states that many people who appear in one picture are very likely to appear subsequently in other photographs, and uniqueness, which states that only one person in each gallery image has the same identity as the supplied query individual [60]. Using the context, Figure 1 conceptually illustrates the effectiveness of the suggested clustering strategy [57]. The samples within a certain feature distance are compared to a vanilla clustering algorithm, which groups the samples into the same identity. But the vanilla clustering technique predicts the identities of the people in the gray areas since they resemble other people in appearance. The suggested clustering algorithm eliminates them based on their distinctiveness [56]. Additionally, although the people in the green boxes are not grouped together by the vanilla clustering algorithm as a result of the suggested approach is to place two green spots with a significant feature distance in the feature space between them. In accordance with the coappearance property, such strong positive samples are discovered [36].

## **1.1 Background**

Person Re-ID is critical in an agile surveillance scheme with substantial research gist and functional significance, because of the pressing requirement for people’s safety and the quantity of tracking equipment. Through the emergence of deep neural networks and the growing requirement for smart surveillance footage, deep neural networks have aroused strong interest in the computer vision sector [61][70]. Owing to the existence of multiple views, varying minimal resolutions, lighting shifts, arbitrary attitudes, occlusions, diverse formats, dynamic camera settings, background congestion, and inadequate class label formation, re-ID is a tough process [72][60]. Therefore, there are a variety of variations and ambiguities. The complex modified camera infrastructure, wide-scale collection with productive retrieval community instability, major domain change, unknown testing situations, gradual model updating, and even changing garments all add to the difficulties of realistic model deployment [75][85]. Because of these difficulties, re-ID remains an unresolved issue. Approaches have focused mostly on custom image generation using body shapes or distance metric modelling. Due to the advancement of deep learning, Person re-ID has demonstrated excellence on regularly cited standards [68]. After all, there is indeed a significant difference between research-based scenarios and real-world implementations. This inspires let us to perform a detailed survey, establish a robust benchmark for various re-ID procedures and consider some probable new paths [90].

## **1.2 Motivation**

Person re-identification (Re-Id) serves a crucial part in sustaining social and public security by efficiently searching, locating, and tracking the target person throughout surveillance camera networks [76][85]. When authorities can track a person's movements within a certain area, it makes it harder for people to do things that are unethical or illegal [68]. In our nation, the number of terrorists is continually rising. We must examine each case separately if we want to reduce the level of terrorism in our nation. It is important to start by identifying the areas with a higher crime rate or areas with more crime. It could be at places like banks, malls, hospitals, and other buildings. Crime will be fairly simple to manage if we can increase security in these areas. For greater security, we may deploy "person identification" technology at certain locations [111][112]. By doing this, we may recognize the perpetrator by looking at his or her attire, face, eyes, nose, and other physical characteristics [51][113]. This security system is already in use worldwide with CCTV cameras [110] [109]. By using this technology in all potential crime hotspots, we aim to be able to lower crime rates. For increased protection, we may deploy "person identification" technology at certain locations. By doing this, we may recognize the perpetrator by looking at his or her attire, face, eyes, nose, and other physical characteristics [55]. This security system is already in use worldwide with CCTV cameras. By using this technology in all potential crime hotspots, we aim to be able to lower crime rates.

## **1.3 Objective**

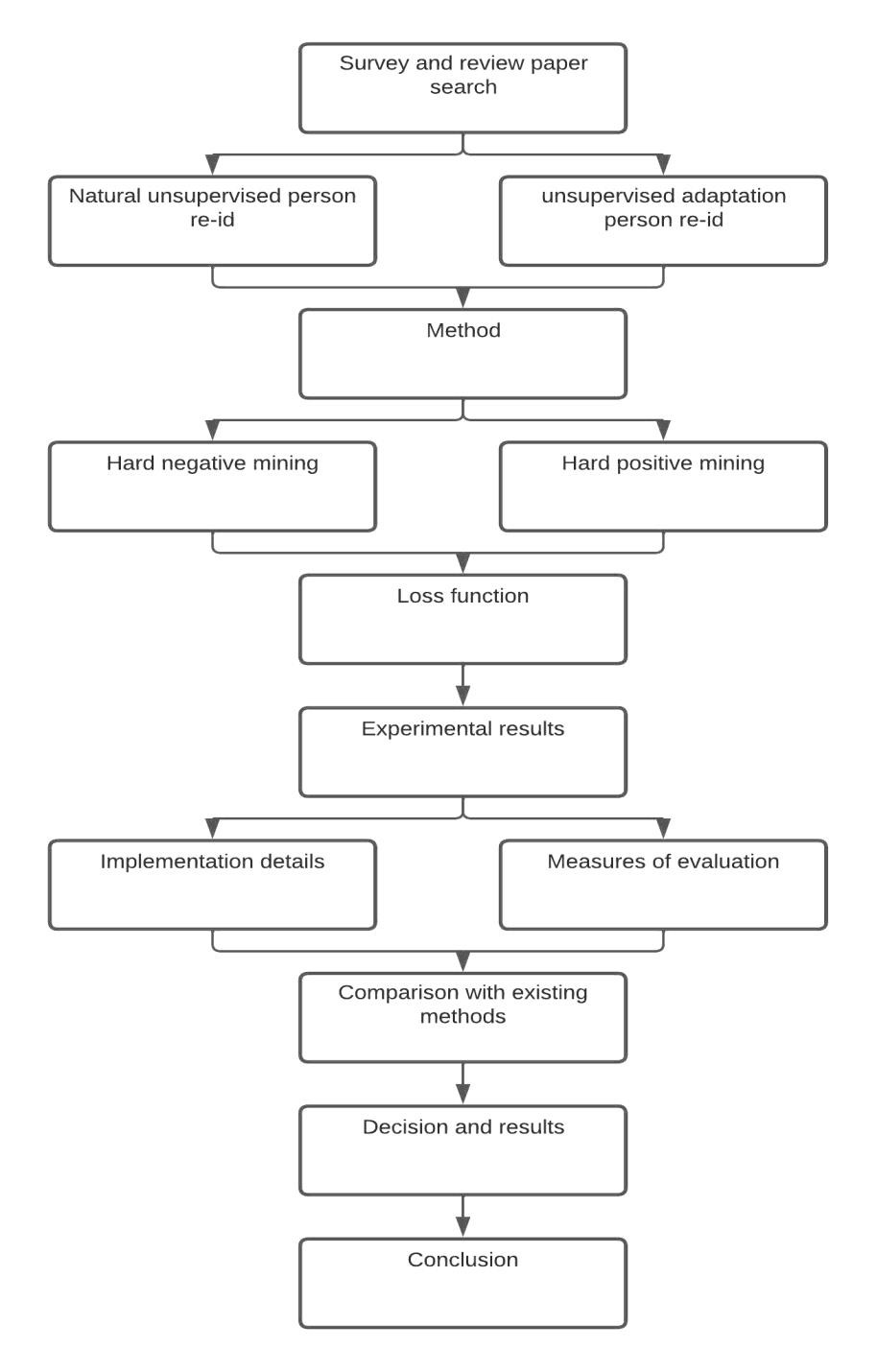
The objective of Person Re-Identification is to compare images taken at different times and locations by the same or separate cameras to see if the same person appears in both. The user's question might be communicated in many different ways, such as via an image, a video pattern, or even just a textual explanation.

The main objectives of this project are

1. To identify the research gap or limitations of the existing work regarding person Re dentification.
2. To solve the following issues:
   1. Loss of identity, verification loss, and triplet loss.
   2. Inadequate resolution and noise.
3. Unlike previous systems, our system can detect sable-dressed person more efficiently.
4. To increase the processing speed and efficiency.
   * + Identify a person of concern through several cameras that are not conflicting
     + Decide if a question person has shown in some other location over different moments recorded by a different video,
     + Explore the identity of interested individuals on the same camera captured in different time snaps.
     + Represent the query individual by a photo, a visual clip, or simply a summary.

**1.4 Overview**

Our approach also aims to reduce cross-camera scene variance, which would be essential in unsupervised Re-ID. We investigate the overall relationship for the pairwise resemblance between intra-camera and cross-camera alignment, which is overlooked in current methods, rather than Camera-to-Camera alignment [114] [13]. With the aid of prior intra-camera matching experience, we seek to acquire reliable pairwise resemblance patterns for combining intra-camera and cross camera [115].

Due to intrinsic restrictions such as different resolution photos, lighting fluctuation, perspective variations, etc. individual Re-ID has remained a difficult work thus far [116]. The issue has gotten a lot of attention in recent times, and several different person Re-ID systems have already been developed. A review of existing methods for person Re-ID is offered in this work [117].

**Chapter 2**

## **1. Literature Review**

Person re-identification has grown rapidly in recent years, from feature design [46] to proximity metrics learning [58] to end-to-end deep learning [59]. The supervised models perform well when there are many labelled data, but the high cost of labelling limits their scalability. Re-identification of an individual without supervision [120][121]. Because it would be impractical to tag a huge number of identities to every specific case, lowering the costs of labelling, person re-identification has recently received greater attention [118]. The majority of the research focuses on using unsupervised learning to learn from unlabeled data for Re-ID [122]. The majority of complex unsupervised approaches for transfer learning or acquiring a previous understanding of Re-ID depend on source data from other scenes. [58] and [60] Using source data for pretraining and clustering and fine-tuning to learn from unlabeled goal data. [36] Learns how to use attribute marks to pass information. [62] Associates track videos through cameras to understand. These approaches use unlabeled data in a variety of ways, and the majority of them directly or implicitly reduce cross-camera scene variance. [35] Camera-to-Camera synchronization, whether at the feature level or even across the scene [63], is how they treat cross-camera scene variation. Semi-supervised learning attempts to learn a task from one or a small number of training instances [64] and some activities of one picture individual Re-ID exist [65][66][67][68]. In [88] Bak et al. apply quantitative instructional methodologies in one shot photograph Re-ID for such a pair of photographs that may be separated by color components plus texture. Wu et al [40] suggest a gradual sampling approach for one-shot video-based Re-ID that gradually predicts accurate pseudo labels and updates the deep model. On the other hand, the previous one-shot Re-ID does not sound right [119].

Person Re-Identification (Re-ID) tries to find a fit for a user query candidate in a collection of images from wide-angle channels that do not intersect [69][47]. Great strides have been made in fully-supervised person Re-ID owing to the strong deep Convolutional Neural Network (CNN) [39][25][73][74]. Unsupervised person Re-ID [75][76][77][78][79][53] i.e., learning with annotated source information and unlabeled test set, or solely relying on unlabeled aim training data, is becoming more popular as a way to avoid the cost of pricey individual ID labelling [123].

Conventional unsupervised person Re-ID works can be grouped into three types: a) utilizing domain adjustment to associate feature dispersion for both input and the output, b) using a Generative Adversarial Network (GAN) to process the image style transfer while preserving identity annotations on input vectors [76][81][82][83], and c) creating pseudo labels on intended realms for training by assigning similar images with similar labels via KNN search, clustering, etc[4][1][3][80][87]. Unsupervised person Re-ID is defined as a transfer learning task that uses labelled data from source areas during the first two categories. Generating pseudo-labels allows Re-ID systems to be trained in a fully unsupervised way, allowing for greater freedom [125].

The majority of pseudo-labels prediction algorithms work on the same principle, which is to compute example commonalities first, then allocate related samples found by grouping or KNN with identical tags [124]. The Re-ID reliability is mostly determined by the estimated sample resemblance throughout this phase. To produce significant pseudo-labels, examples from the same personality should be more comparable than samples from various ids. Furthermore, the unsupervised person Re-ID option makes the learning process of reliable sample similarities challenging, particularly for samples from various cameras [126] . Multi-cameras with varying characteristics and situations, for example, can be used to record each identity. These factors have the potential to substantially alter the identity's appearance [127]. To put it another way, the domain gap between cameras makes it impossible to distinguish between samples of the same identity and to optimize intra-class feature similarity [128][129].

Table 1. Summary of literature review

# Publication/author Focus Finding

|  |  |  |
| --- | --- | --- |
| Salman Khan & others [131] | Dispersion based Clustering (DBC) approach. | This study proposes a simple but effective clustering technique for unsupervised person re-identification. Dispersion, a foundational notion in statistics, is examined to provide robust clustering criteria. |
| Mang Ye & others [88] | Metrics for  Single-/CrossModality | They divide the many elements necessary in creating a person Re-ID technology into closed-world and open-world scenarios. When using the |

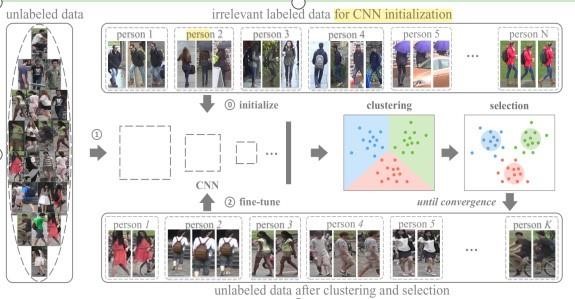
|  |  |  |
| --- | --- | --- |
|  | Evaluation and  Baseline | well-studied closed-world environment, a number of survey assumptions are usually made. This has been done with great success using deep learning techniques on many datasets[202]. |
| De Cheng & others [132] | Hybrid dynamic clusteringbased approach. | It turns the unsupervised Re-Id problem into a framework for dynamic contrastive learning from the local to the global level and selfsupervised distillation. |
| Haoxuanye Ji & others [133] | Meta Pairwise  Relationship  Distillation  (MPRD) method. | In this research, a (CNN) and (GCN) are used. The GCN utilizes the currently selected features by the CNN to calculate the pseudo labels of sample pairs, and the CNN learns good characteristics by including high-fidelity positive and negative sample pairs enforced by the GCN. |
| Zilong Ji & others [134] | VA (voxel attention) and  2-stage clustering (TC). | The (ADTC) method is suggested in this research as a means of resolving intrinsic issues. This approach specifically comprises of two approaches. First, they utilize an unsupervised attention kernel to transfer learned information from the image background to the pedestrian foreground. Second, to facilitate learning, it divides the clustering process into two steps[203]. |
| Takashi Isobe & others [135] | CCL, PDA, and  Fourier  Augmentation. | The issue of person re-ID with unsupervised domain adaptation is raised where annotations are known for the source task but not the destination. |

|  |  |  |
| --- | --- | --- |
| Xuanyu He & others [38] | Based on stateof-the-art  person Re-ID models to their dataset | In this research, they provide an image-based person re-id dataset that  was gathered from five nonoverlapping camera viewpoints in Dublin, Ireland's sizable and crowded airport. Other image-based person reidentification datasets can't help us with our study on how to combine visual and temporal data. |
| Xue Li & others [] | Style separation and the  contrastive learning (CA-U re-ID) method in detail. | Images of the same thing taken by different cameras that don't look at it from the same angle can be very useful for intelligence surveillance and public safety. |
| Yutian Lin & others [136] | Based on  (BUC) approach to jointly optimize a (CNN). | It uses a diversity control parameter in the bottom-up clustering procedure to make sure that the amount of data in each cluster is the same. Therefore, that model successfully strikes a balance between similarity and variety. |
| Jingya Wang & others [36] | Used transferable  Joint Deep  Learning approach. | This study demonstrates (TJ-AIDL), which permits the simultaneous development of an attribute-semantic and identity-discriminative feature representation space transferable to any new (unseen) target domain for re-id tasks without the requirement of fresh labeled training data from the desired domain. |
| Geng & others [89] | learning local relations among parts of pedestrian images | Aims to match the given person in the different scenarios |
| Mang Ye & others [42] | Recognizing a suspect using several, nonoverlapping video feeds | Develop an AGW benchmark and a proper testing metric (mINP) for ReID people. |
| Nambiar et al. [91] | Gait-based  Person Re-ID | The fair analogy has been challenging due to the absence of a reference dataset which may challenge the algorithms with a large variety of conceivable circumstances. |
| Zhengxu Yu et al [92] | Feature Learning | semi-supervised (AIFL) framework and generate cloth altering pictures as per the intended cloth embedding. |
| Hong-Xing Yu et al [93] | cross-view clusteringbased asymmetric distance metric | Clustering-based Asymmetric Metric Learning (CAMEL). |

## **2.2 Natural unsupervised studying person re-ID**

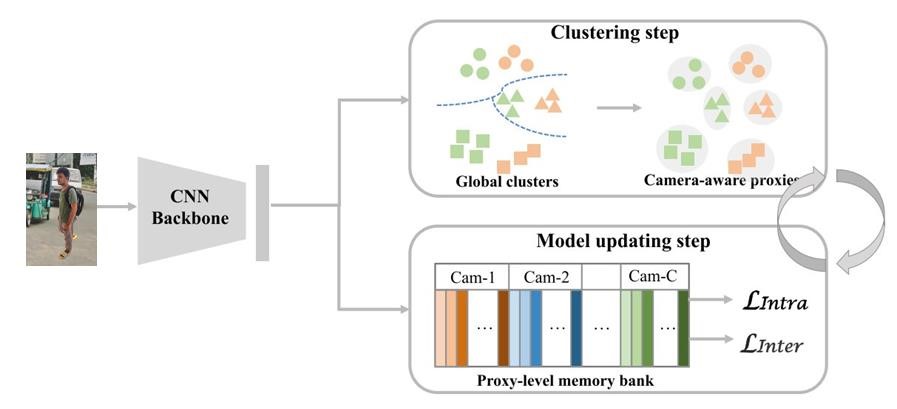
**2.1.1 Lack of annotations** by using highly labeled data and the achievement of deep networks, state-of-the-art supervised techniques have achieved remarkable success in person re-id. Hehe Fan [192] has nonetheless suggested the first mature unsupervised learning person re-id in the unsupervised field. He offers a reliable USL baseline in his study. The starting point he mentioned can be feature extraction is the first of three steps that are described. Unlabeled data are entered into the backbone. (Resnet50 [2]) and then deploy the clustering technique using the extracted characteristics. Selecting the benign samples is the next stage. Its angle with regards to the center is smaller than a specified parameter. The final step is to adjust the spine with the benign samples chosen in the subsequent stage. On existing datasets, the suggested strategy didn't produce very impressive results. But it offers others a strong framework to develop on.

**2.2.2 Label noise** issue of insufficient annotation is effectively resolved by the prior unsupervised learning person re-id [148]. Efficiency and speed, however, continue to be major issues. Learning discriminative features without actual labels is a pressing problem. The solution to the issue is to first give each person's picture a particular label before switching to a multi-label classification system utilizing the model for label prediction [194]. Additionally, they put forward the multi-label categorization loss based on memory, which employs an additional category and blends single-label and multi-label classification in its design. They succeed in resolving the problem.



## Figure 2. Early purely unsupervised learning [2]

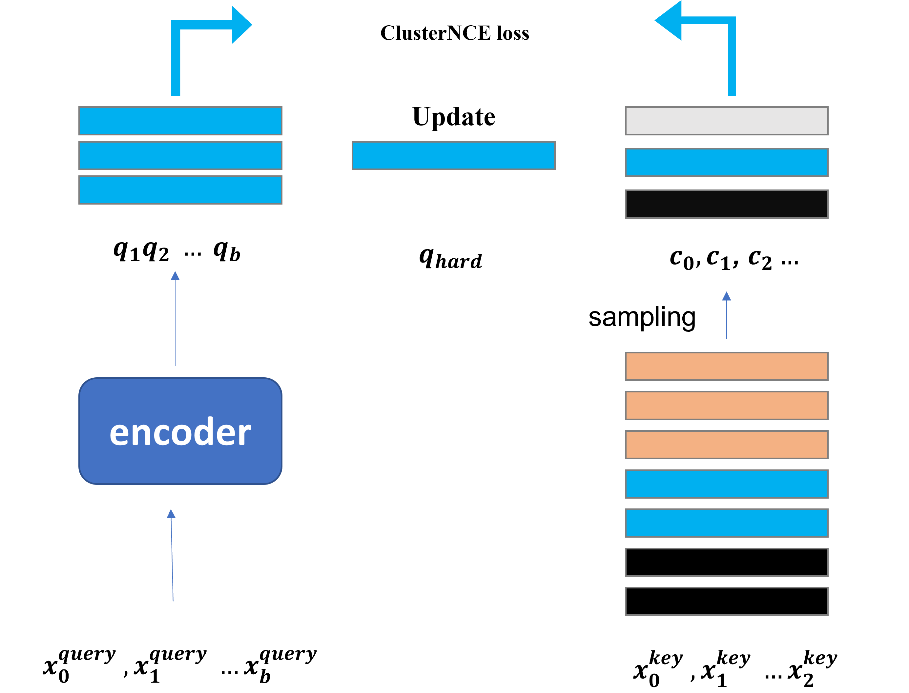
**2.2.3 Disparity in distribution** the important factor in the poor performance is that the distribution disparity between the data gathered by various cameras was created by the disregard of the difference brought on by the change in camera perspectives and design differences of two different cameras [147]. Most clustering-based techniques, such as BUC [195], operate camera-independently, retaining id-to-id similarity while potentially neglecting intra-id variation. Menglin Wang et al. [196] gave up utilizing camera-agnostic techniques to address this issue. They split up a single cluster into many proxies, each of which represents individual images captured by the same camera [130]. The noise and intra-id variation brought on by a separate camera are significantly reduced by this technique. Compared to the earlier studies, they have produced excellent results.



## Figure 3. Camera-aware methods [197]

Shiyu Xuan et al. [197] suggested an alternative approach to address the issue. They observed that the inconsistency had not been addressed in earlier publications. They divide the sample similarity calculation into two parts, intra-camera computation, to strengthen their technique. The intracameral computation may make use of the CNN features for similarity computation, resulting in labels that are more trustworthy. They are categorized in the subsequent step. The scores for each item's categorization on multiple cameras are taken into account as a new feature vector in the second step. With this new function, labels are made that last longer and the difference in distribution between cameras is reduced. With this strategy, they get rank-1 accuracy on the Market1501 dataset, which is also a great result.

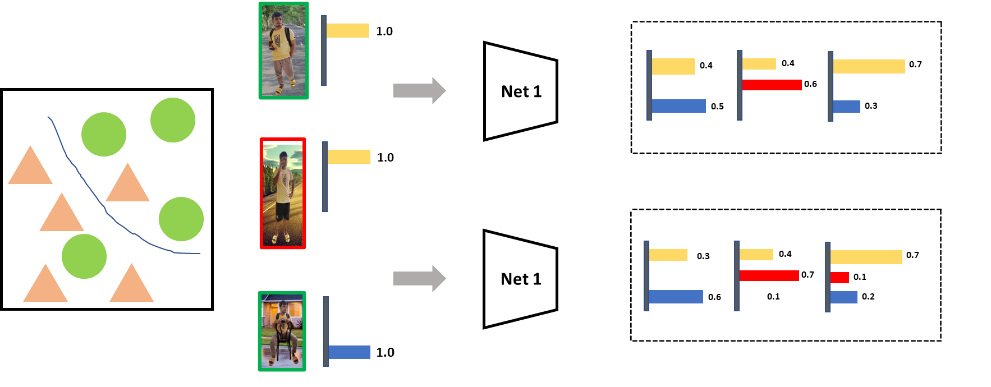
**2.2.4 A difference in cluster size** may cause an update imbalance. In earlier work, the cuttingedge methods calculated multi-label classification damage with each instance while storing each item in the memory dictionary has a similarity matrix. Such a baseline primarily has two drawbacks [198]. First, the updating process is unstable, which causes the distribution of data sets to be skewed. Second, the pseudo-labels in the dictionary are not best represented by the center of each cluster. They employ a single feature vector for each cluster in order to solve the issue, and throughout the training phase, the hard query instance feature makes changes to the characteristic, which is the query that is the most dissimilar to the cluster. The above issues will be resolved in this way. The most effective unsupervised re-id approach is cluster-contrast [35].



## Figure 4. Clustering boosted method [199]

### 2.3 unsupervised adaptation person re-ID

The majority of UDA methods are created in a closed-world environment, assuming that both the source domain and the target domain have the same classes [149]. The two different domains in the open-world scenario are entirely different, hence the assumption is false. Persona with Unsupervised Domain Adaptation Re-ID still has several issues [50]. The domain difference is the major issue, and the issues listed below are a subset of the main issue.



## Figure 5. Camera style transfer [10]

**2.3.1 Intricate set designs and varying lighting** This is an immediate issue that requires inquiry. GAN-based transfer work gained a lot of traction in 2018, leading to the widespread usage of GAN to produce style-transfer pictures. Since the original GAN can only produce random pictures, PTGAN or CycleGAN have been extensively used in different projects. PTGAN, which Longhui Wei [8] designed for person-re-id, produced excellent results. They want to demonstrate that a style can be transferred from a training domain to a test domain in their experiment. The first deals with the issue, including backdrops, lighting, and resolutions. Style transfer and maintaining one's individuality are the two objectives of the PTGAN. Before learning how to transfer styles, you have to learn how to map styles between different person datasets. In order to accomplish the two aims, the process of person identity preservation is to preserve the identity of one person after transfer. They use this approach to resolve the issue as a consequence. Additionally, they provide MSMT17, a brand-new dataset that is essential for re-id. The enriched training set, which was made by Zhun Zhong [200] by putting labeled training pictures into each camera and putting them together with the original training samples using CycleGAN, may make training more interesting.

**2.3.2 In domain adaptation person re-id-** the unavoidable label noise based on the clustering process remains a significant issue. The person re-identification work differs from other domain adaptation jobs in that it is uncertain how many labels are present in the target domain [201]. They suggest softly modifying the noisy pseudo labels in the selected domain to address the issue and lessen its consequences. They emerge with a brand-new framework they call Mutual MeanTeaching. Additionally, they give up the conventional triplet loss, which is ineffective for labels with delicate improvements. Instead, a new soft max triplet damage is made to go with the soft, polished label in the training section. In doing so, the new model successfully completes the unsupervised operations in Market-to-Duke, Duke-to-Market, and MSMT-to-Market domain adaption, and Duke-to-MSMT with notable outcomes of 14.4 percent, 18.2 percent, 13.1 percent, and 16.4 percent map.

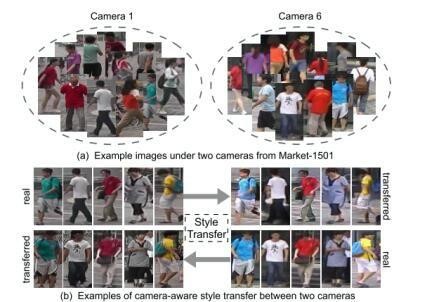


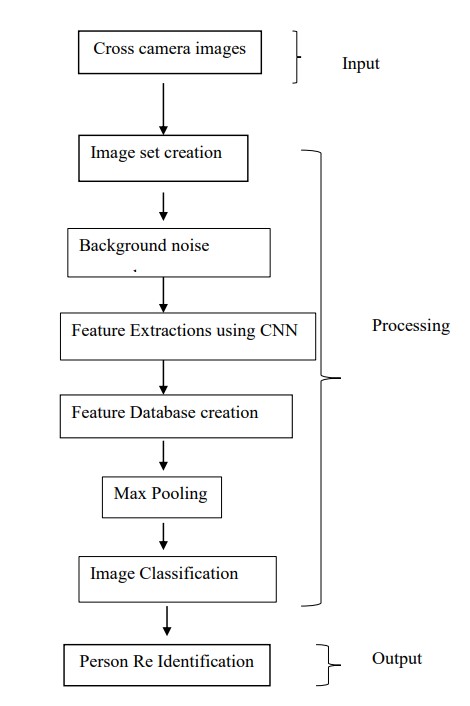
Figure 6. Mutual Mean-teaching [9]

# Chapter 3

## **3. Person Re-Id**

Person re-identification (Re-ID) is a human-centric AI system that searches a huge number of films for people of interest. It helps with tasks like looking for video sequences relating to a celebrity of attention from a Television show, discovering a missing child in a town centre using CCTV footage, and identifying offenders in video surveillance systems, all of which involve tedious video gazing. Its quality and efficiency help to speed up the footage analysis process. In general Re-ID, we have achieved great progress in recent years.

Person Re-ID in this study, except as otherwise noted, refers to the pedestrian recall challenge over several security cameras from a vision-based view. Developing an individual's Re-ID scheme for a particular situation involves five main steps. (as depicted in Fig. ):



## Figure 1 Fundamental steps for Person Re-Identification

1. Step 1: **Gathering Raw Data**: The basic necessity of practical video investigation is to acquire raw footage information from CCTV. These images are generally placed in various locations and various situations [94]. Most likely, there is a lot of intricate and loud background clutter in this actual data.
2. Step 2: **Bounding Box Creation:** From the raw video data, extract the bounding boxes that include the person images. In most large-scale applications, manually cropping all of the human photos is unfeasible. Person detection [95][96] or tracking techniques [97][98] are commonly used to obtain bounding boxes.
3. Step-3: **Annotating the cross-camera labels using the training data:** Learning input tagging is usually necessary for discriminative Re-ID simulation because of substantial cross-camera differences. Because of the substantial domain change [99], we must frequently annotate the training data in each new situation.
4. Step 4: **Pattern Training:** Using the previously annotated human images/videos, train a discriminative and robust Re-ID model. Developing a Re-ID system begins with this paradigm, which has been extensively researched in the literature. To deal with the myriad problems extensive models have been created, focusing on feature representation learning [100][101], distance metric learning [102][103] or their mixtures.
5. Step 5: **Person Retrieval**: It's done during the testing period that pedestrians are retrieved We derive local features from an individual (query) and an art collection using the Re-ID model depending on the previous phase. The calculated query-to-gallery similarity is sorted to produce a retrieved ranking list. Some systems have also looked toward optimizing ranking to improve retrieval performance [104][105].

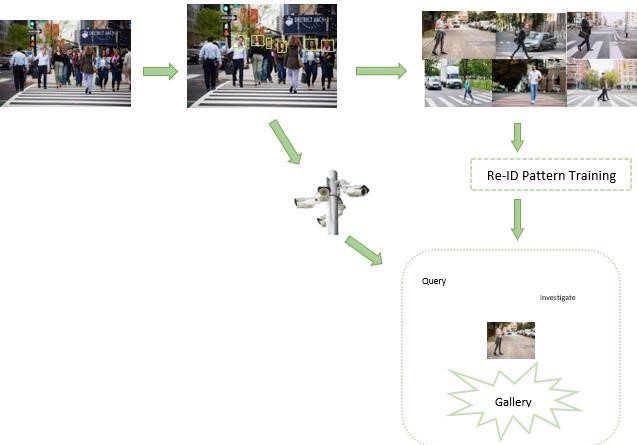


Figure 7 the five essential phases involved in building a realistic person Re-ID mechanism: 1) Gathering Raw Data, 2) Bounding Box Creation, 3) Annotating the cross-camera labels using the training data, 4) Pattern Training, and 5) Pedestrian Retrieval

### 3.1 Closed-World Person Re-Id

This section summarizes Re-ID for closed-world individuals. In this context, the following assumptions are frequently made: Participants are illustrated by embeddings, with about the same belonging taking up the large percentage of the input vector region; 3) The learning has plenty of annotated training data for monitored discriminative Re-ID model learning; 4) The captions are correct; 5) The request person should appear in the photo album. Feature representation learning, deep metric learning, and ranking optimization all fall under the umbrella of a closed-world Re-ID system. Feature representation learning focuses on establishing feature creation procedures, while deep metric learning employs a variety of loss functions and sampling tactics to design training objectives. The next part gives an overview of the datasets as well as an in-depth analysis.

**3.2 Feature Learning**

Feature Representation Learning is a form of learning that seeks to improve feature creation techniques. Utilizing photos of certain people wearing diverse outfits, [92] introduce the semi-supervised Apparel-invariant Feature Learning (AIFL) framework towards training an apparel invariant walker image. It is costly to train fully paired labelled data. To get around this issue, [58] suggested a patch-based unsupervised learning system that uses patches rather than entire images to train discriminative features. Learning a discriminative model is made possible by exploiting the resemblance between patches. [48] build a PatchNet to choose patches from feature space and then train discriminative features for all those patches. Also, describe an autonomous patch-based robust feature training loss continue providing excellent advice for the PatchNet to learn discriminative patch features on unlabeled datasets. Furthermore, we create an image-level feature learning loss that uses all of the patch features from the same picture to provide image-level instruction towards the PatchNet. To aid in the unsupervised re-identification of individuals, [107] provide the Cross-camera Erased Feature Learning (CEFL) framework, which takes advantage of the cross-camera visual feature aspect as well as the regional information to train discriminative features. Strategies for unsupervised re-identification of individuals typically use pseudo-labels generated through clustering. However, the strength of the learnt attributes, which are disproportionately controlled by the colours of photos, has a significant impact on clustering results. [108] try to censor the detrimental controlling impact of colours in addition to learning more useful features in unsupervised person Re-ID. For the unconstrained person, Re-ID, [98] offer a Cluster-guided Asymmetric Contrastive Learning (CACL) solution, whereby the grouping outcome is used to direct feature learning in an appropriately constructed asymmetric contrastive learning framework. All object and node contrastive learning are used in CACL to aid the siamese net in accumulating discriminant characteristics concerning the clumping outcome inside as well as between distinct data enhancement view perspectives, accordingly.

### 3.2 Re-id methodology

Using person identity labels, the traditional supervised person techniques train the Quick R-CNN [11]and a re-id header in the featured phase. In this paper, we provide a unique person search problem with existing bounding box labels. Identifying a person, but the labels for the person's integrity and identity are not provided for person re-id. We may use the context of the un - supervised person re-id difficulty to our advantage. Solve this issue by identifying the individuals depicted in a gallery photo. Figure 2 depicts the overall operations of the challenging clustering modules-containing mining in strong positive conditions (SPC) and mining in strong negative conditions (SNC) (SPC) technique.

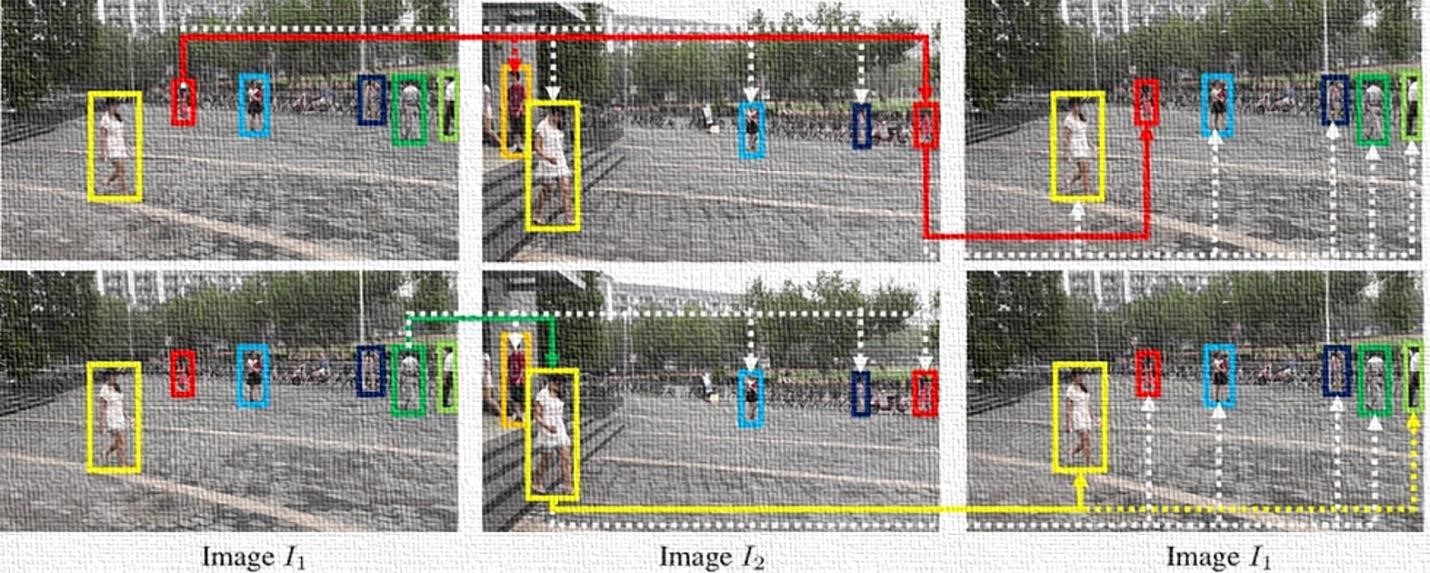


Figure 8: Acts of SNC that rely on individuality. Similarity between two people's physical characteristics is shown by the white and colored lines. Each set of solid lines indicates a high degree of similarity between the corresponding features.

#### 3.3 Mining in Strong Negative Conditions

Presently Unsupervised Individual Re-id algorithms [12], [13] may provide confined results on actual datasets due to the fact that each person identity contains different quantities of samples, such as in the person search dataset PRW. For unsupervised re-identification of individuals, we propose a new clustering approach in order to successfully decrease strong negative samples, strong negative Condition (SNC) makes use of the uniqueness aspect of person search [153][151].

Let G = {*I*1*, I*2*,* ···, IN} be a collection of pictures, with N being the total number of images, and let X*l* stand for the list of *xlj*’s identified in *Il*. Assume we discover a result from the inquiry's positive sample in the *i*-th picture *Il* with *l* 6= k, and that the i-th individual *xik* in the k-th image *Ik* was chosen at random. First, we choose the people in Il who are candidates because their feature similarities to the query individual are greater than certain threshold *δ*. finally start with the positive samples set as

 (1) where f*lj* is the *xlj* feature vector from the recollection of features, and s(f,g) represents the connection between f and g features. It is possible to assemble a collection of affirmative samples for the query *xik* by collecting  across all the gallery photos and using the union 

. where s(f,g) is the connection between two features of f and g and f*lj* is the feature vector of *xlj* in certain characteristics. The gallery photos'  values could be aggregated, and their union

 might be used as a positive sample set for the query *xik*

***xlj\**** *= argmax s(fki ,flj*)*,* (2)  *xlj*∈*Cˆl (xik)*

which is most like the query *xik*, is kept in , and the other samples are culled from . As a second measure to ensure that the generated candidate *xlj*∗ fulfills the cycle consistency to *xik*, we use the backward matching of WTA method from *Il* to *Ik* to further enhance the precision of the clustering. If *xik*  is substituted for  in the original question, then *xik* and *xlj*∗ are considered to be identical and placed in the same set. If *xlj*∗ is also present in criteria is not met, and *xlj*∗ must be removed. If anybody who matches the description of *xik*. A set satisfying the *xi* as been found at last

0

20

40

60

80

100

120

1

2

3

≥5

0

20

40

60

80

100

2

3

4

5

6

7

≥8

(a) PRW (b) SYSU-CUHK

Figure 4: Ratios of picture sets that have one genuine positive pair of people (orange) to those that consist of numerous (blue), based on the accuracy of the matching. Examples of such datasets are (a) PRW and (b) SYSU-CUHK

In Figure 3, we can see how the suggested SNC operates, with the person's locations shown by bounding boxes of various colors, reflecting their individual identities. The first row represents the selection of a query, which in this case is the individual shown by the red box in I1. Since both people in the orange and red boxes are wearing red, and because there is a degree of feature similarity that is much more than what is required for the query, I2 labels them as potential positive samples. On the other hand, the right person is selected from the red box in I2 by applying WTA to establish the individual with the highest degree of similarity to the query. When it is determined, through the process of backward matching, that the individual from the initial query in I1 has the most in common with the person whose information is contained within the red box in I2, the individuals who have been found to have the most in common are referred to as positive pairs. Nevertheless, as shown in the second row of Figure 3, when the person encased in green boxes in I1 is used as a query, the person encased in yellow boxes in I2 is recognized by [193] despite the fact that this is not a positive class sample because the person encased in green boxes in I2 does not exist. Because the person in the yellow box, rather than the original query individual in the green box, was identified from I1 using the backward matching process, this strong negative sample has been removed from the clusters. The mining SNC cannot be applied to the unsupervised person re-id framework since there is no contextual information available for the connected gallery photographs; however, it might be used in the person inquiry framework.

#### 3.4 Mining in Strong Positive Conditions

Unsupervised clustering using the uniqueness-based SPC also misses some true positive data. We also suggest a hard positive condition (SPC) method that identifies difficult positive examples with little similarity to the query in terms of features. Depending on the co-appearance property, which states that several people in a picture are possible occurrence similarly in other pictures, we use the contextual data of the nearby individuals to the query who appeared in the same picture. Figure 4 shows our investigation of this characteristic using the PRW and CUHK-SYSU datasets. In order to classify all the sets of images that contain at least one true positive pair of people, we gather all the sets of images and divide them into various categories based on the corresponding to potential, or the smallest number of people in each between the two photos that make it up each duo[150][152].

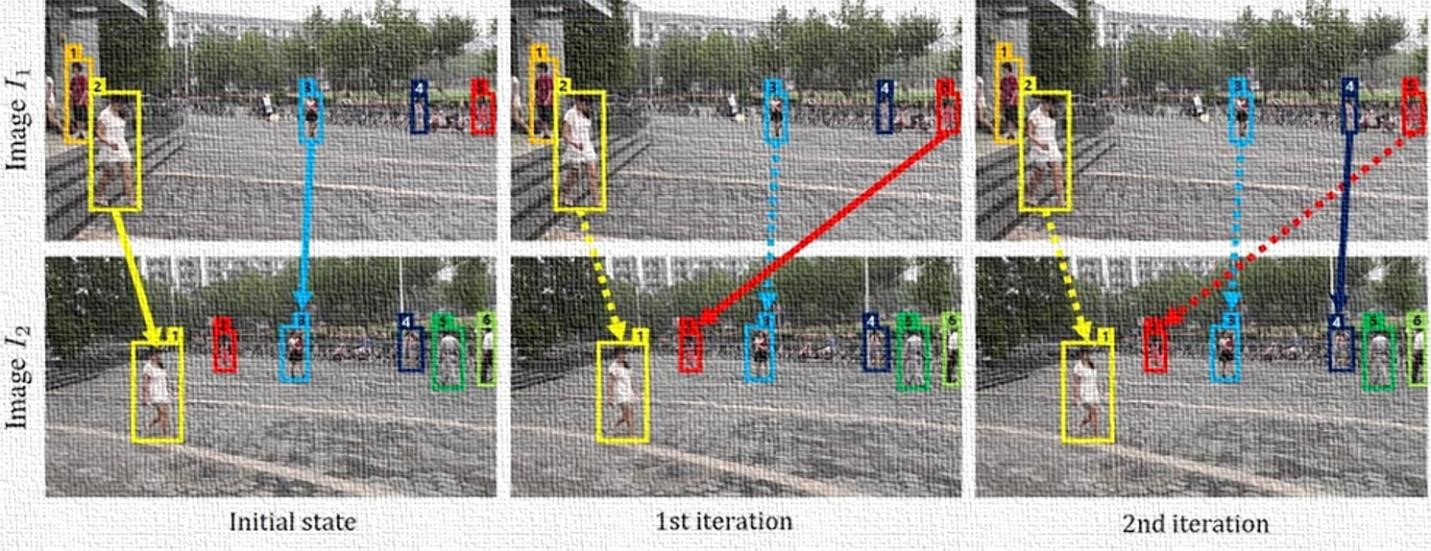


Figure 9: Performance characteristics of the high-performance computer. Solid arrows indicate freshly discovered positive couples at this iteration, whereas dotted arrows link those already discovered in earlier stages.

For each category, we compare the frequency with which two photos include a single real positive match of persons where the co-appearance criteria are not met (red bar) with the frequency with which numerous real positive pairings of people where the feature is fulfilled (blue bar). Figure 4 displays the statistical findings demonstrate that as more people appear in an image, the number of multiple T positive pairings where the condition of co-appearance is not satisfied (red bar) and numerous strong positive pairings that satisfy the co-appearance property (blue bar). Figure 4 displays the statistical findings that the quantity of several genuine positive pairings increases as the number of people in a picture grows; this is notably true for the CUHK-SYSU dataset. Just at t-th iteration, we calculate *A*(*t*)(*k,l*) between *Ik* and *Il* by taking into account the clustering outcomes of the nearby people provided by

*A*(*t*)(*k,l*)= ∑ *s*(*fki ,flj),* (3)  *x*

Where, C*l*(*t*)(*x*) represents the x-cluster at the t-th iteration of the l-th picture. Keep in mind that *A*(*t*)(*k,l*) is derived from the most recent clusters of C*l*(*t*)(*x*) and grows when more people in *Ik* are paired with people in *Il* who have more commonalities with them. Then, in the next iteration, we adjust the initial feature similarity of  *s*(f*ki ,*f*lj*) to *s*(*t*+1)(f*ki ,*f*lj*) for all the pairs of *xik*  X*k* and *xlj*  X*l*, as

 (4) where *β* signifying a load to adjust the contribution of empirically set to 0*.*1. In order to improve the clustering results, we once more apply the SNC between *Ik* and *Il* using the updated feature similarities. Because the feature matches are raised by the same amount, non-empty clusters are preserved. However, by exploiting the increasing feature matches to find strong positive samples, we can determine whether empty clusters include new components. In this study, SPC iterates a maximum of three times.

Figure 5 presents two pictures, *I*1 and *I*2, to which the SPC is applied in a cyclical fashion. Annotations of the names of the individuals are placed above the bounding boxes. The two true positive pairs of and are assumed to be properly clustered from the outset, but the other two true positive pairings of and are overlooked to be successfully clustered owing to very low feature similarities. However, because the feature similarity *x*51 and *x*22 and *A*(0)(1*,*2) becomes greater than the threshold *δ* by the SPC of *x*21 and *x*31 linked with the people of *x*51andnearby to is freshly discovered as a positive pair *A*(1)(1*,*2),



,

*s*



At the next iteration, the newly formed pair (16) is identified as a positive one because the extra positive pair sraises the SPC to ,which in turn raises the feature similarity (15) by the same amount.

**3.5 Loss Function**

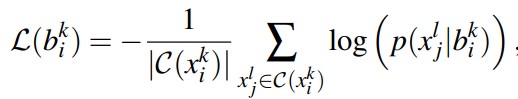
The scale invariant store contains the adjusted attributes of all person objects in gallery photographs. After storing grades in a scale-invariant way, each feature vector f*ki*  R256 for *xik*  X*k* starts out as a null vector and is then changed by

f*ki* ← 1(f*ki* +g*ki* )*,* (5)  *Z*

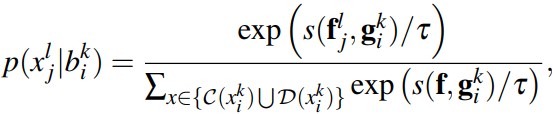
Wherein g*ki*  R256 is the feature vector obtained from a boundary box *bki*  that was identified and physically overlapped with the predicted bounding frame of *xik* in *Ik*; Z is the normalization factor, making is the result. The recognition loss and the re-ID loss constitute the total loss function, and we employ the stronger R-CNN [11] recognition loss. We suggest a re-id loss for the discovered boundary box *bki*  supplied by

; and



 (6)

that maximizes the chance and promotes *bki* to be identical to *xlj*  C(*xik*)

 (7)

Wherein is a heat capacity that has been experimentally fixed to 0.1, and D(*xik*) is the fixed of the negative instances in C*c*(*xik*) that have top 1% matches to *xik*. Be aware that the decrease in equation (7) mostly comprises problematic negative examples with high matches to equation *xik* and ignores negative samples with fairly low matches to *xik* . So, the suggested re-ID loss makes it more likely that samples will be good, while making it less likely that samples will be hard to reject.

Furthermore, the chances of *p*(*xlj*|*bki* ) is increased by  for all *xlj*  C(*xik*) datasets of *xik* that are grouped to be positive. In the end, the system works well because HNC and HPC work around each other to include most of the positive cases in C(*xik*) while reliably keeping even difficult negative cases out of C(*xik*).

**3.6 Loss of verification**

It uses a contrastive loss [ 171], [172] or a binary verification loss [173][174] to maximize the pairwise connection. There are various variations, such as [175]'s pairwise comparison using ranking SVM.

Binary verification [176] distinguishes between the positive and negative aspects of a pair of input images. The differential characteristic is classified as positive or negative by the verification network. To increase performance, verification is frequently paired with identity loss[177].

Loss of three triplets. It addresses Re-ID model training as a retrieval ranking problem. The fundamental idea is to reduce the distance between the favorable and unfavorable pairings by a specific amount [178]. If we directly optimize the above loss function, the huge fraction of easy triplets will dominate the training process, resulting in low discriminability. Various informative triplet mining algorithms have been developed to address this issue [179],[180]. The main concept is to pick the three most informative triplets [181],[182]. In particular,[183] introduces mild positive mining with a weight constraint, which directly optimizes the feature difference. Harmans et al. [184] show that discriminative Re-ID model learning benefits from online hardest positive and negative mining inside each training batch. For informative triplet mining, some methods investigated the point-to-set similarity strategy[185]. A soft hard-mining strategy improves resilience against outlier samples. A quadruplet deep network is constructed in [186] to further enrich the triplet supervision, with each quadruplet including one anchor sample, one positive sample, and two mined negative samples. The quadruplets are created using online hard negative mining with a margin. Smaller intra-class variation and bigger inter-class variation arise from optimizing the quadruplet connection. One of the most popular strategies for deep Re-ID model learning is to combine triplet loss and identity loss [187],[188]. For discriminative feature representation learning, these two components are mutually helpful.

Loss of OIM. A memory bank approach is used to construct an Online Instance Matching (OIM) loss [189] ,[190] in addition to the three-loss functions stated above. Denominator [191] is derived from an unlabeled ID feature set that has been remembered, which explains a large number of IDs that are not focused. Unsupervised domain adaptive Re-ID [192] uses this memory method as well.

# Chapter 4

## **4. Dataset and Experimental Results**

**4. Dataset**

**CUHK03.** The collection, consisting of 13164 photos of 1360 IDs, has been one of the greatest people ReID collections. Every passenger is filmed by a camera module, as well as all IDs are derived through six camera positions. There are two options in this collection of data. One is labelled instantly by a sensor, while another is personally evaluated by people. The earlier is closer to real-world situations than the latter [162].

**GRID.** The data was acquired from a subway tunnel using eight non-overlapping field-of-view lenses[163]. This dataset's photos feature a lot of lighting fluctuations and are low contrast. The 250 sets of human photographs inside this dataset were done from two different camera angles, whereas the remaining 775 photos were captured by a single camera.

**ViPER.** Each one of the 632 IDs throughout the ViPER dataset [164] includes two photos obtained from separate non-overlapping camera systems. The dataset contains 28 different perspective angle pairs including eight identical viewpoint orientations, with considerable differences in environment, illumination changes, and perspective. As a result, it is gathered to verify the perspective invariant individual ReID model. All of the photos in the collection have been cropped to a size of 128 x 48 pixels. Among the most difficult datasets for computerized individual ReID was this.

**CUHK01**. [165] is made up of 3884 photos of 971 walkers from two security camera perspectives. One camera records a pedestrian's rear or frontal sight, while the other one catches the pedestrian's profile views.

**DukeMTMC-Reid.** The DukeMTMC-Reid is a human ID that is founded on a picture. The ReID collection is a part of the DukeMTMC database. It is made up of 36411 walker photos from eight high surveillance cameras that belong to 1812 different people. 1404 of the 1812 people were captured by more than two camera views, while the rest were considered distraction technique identifications.

**Market-1501**. the dataset includes 32643 squares with tags for about 1501 individuals. Two to a maximum of six camera systems are stationed next to the store to monitor every shopper. The Deformable Part Model (DPM) sensor can detect the movements of individuals.

**Partial-Reid.** This collection [166] was produced specifically for partial human ReID and has 600 photos of 60 walkers, each with five partial and real images. These photographs were taken from various backdrops, perspectives, and types of occlusions.

**Airport.** The data were acquired from six cams in an airport terminal indoor competence monitoring system. It comprises 39902 bounding box pictures of 9651 persons, with a mean of 3.13 pictures per person. The frames range in size from 54130 to 166403. 1382 of the 9651 people have been connected including at least 2 different camera systems. The sources [167] [168] further provide precise information on this data. To test the models, video (iLIDS-Video, MARS, PRID2011 [75]) and image (CUHK03 [72], Market-1501 [169], DukeMTMC [170]) based person re-id performance monitoring statistics were used. These datasets were usually assessed individually in earlier studies. As recent large-scale image-based re-id datasets were often built by sampling human anchor boxes from videos, we believe that these image datasets exhibit similar properties to those video-based datasets.

Re-identification Across indoor-outdoor Dataset (RAiD) [161]: It comprises 43 IDs in 6920 bounding boxes that were recorded by 4 cameras. The cameras are divided into four groups, the first two of which are inside and the last two outside. Apparently, the photos comprise of quite huge light fluctuations because of interior and outdoor circumstances.

**4.1 Experimental Results**

For person searching, the CUHK-SYSU [205] and PRW [206] benchmarks have been frequently employed. The CUHK-SYSU dataset contains 11,206 photos for training and 6,978 images for testing, and it includes 96,143 bounding boxes with 8,432 identities to represent specific people. There are 2900 practice inquiries and 300 practice picture galleries available. Training and testing pictures make up the PRW dataset, which totals 5,704 in total. It is composed of 6,112 individual shots taken from 6 predetermined vantage points. There are 43,110 person-bounding boxes available, covering 932 unique identities, and 2,057 sample queries are provided. While PRW It doesn't provide enough information to define galleries for person-matching searches. When it comes to handling inquiries, we split galleries into two distinct varieties: the standard gallery and the multi-view gallery [47]. Throughout this paper.



Figure 10: Clustering outcomes (a) Query persons, (b) SPC, and (c) SPC+SNC combined

|  |  |  |  |
| --- | --- | --- | --- |
|  | w/o SNC | with SNC | SNC gain |
| **w/o SPC** | 27.68/59.35 | 32.87/62.68 | +5.19/+3.51 |
| **With SPC** | 28.01/60.28 | 36.61/64.85 | +8.60/+4.57 |
| **SPC gain** | +0.33/+0.93 | +3.74/+1.99 | +8.93/+5.50 |

Table 2: study of the effects of SNC and SPC on elimination. The result is displayed in every cell as a Top-1 level and map, respectively.

The usual gallery excludes the picture where the query was detected and includes all the other test photos, but does not display any of the photographs with unlabeled people. As with many already existing solutions, this should include a gallery [52]. The multi-view gallery presents a significant challenge due to its inclusion of all unlabeled individuals and exclusion of duplicate test photographs acquired using the same camera view as the query image [46]. Even though there are identification labels on both samples, it's important to note that we don't use them when teaching the system to work in an environment without identification labels [53].

**4.1.1 Implementation details** for image processing in Figure 2, we use the ResNet50 [207] [13] stem network, which includes the "conv1" through "conv4" blocks, and the "conv5" block as a submission heading. Every encoder of the re-id header is described as a fully connected layer to enable it to condense the high-dimensional features transmitted from the base station to an input vector with a size of 128. We construct the proposal system with the same hyper-parameters as for the SYSU-CUHK and PRW samples, except that we change the number of instruction periods and the proportional gain decline [156][153]. The development strategy is empirically supported by the findings of the ablation study and other major studies. The supplemental information offers more details.

**4.1.2 Measures of evaluation** The Maximum of Average Precision (map) and Top-k Factor are the two measurements that are typically used to assess the effectiveness of person searches. By calculating the average area under the precision-recall curves, map takes into account both the precision and the recall of the anticipated results. A Top-k score determines whether or not there is at least one real positive sample among the anticipated top-k candidates that best fits a given query.

### 4.2 Comparison with Existing Methods

The suggested method is contrasted with existing supervised person lookup techniques such as expansions of successful unsupervised person re-id techniques, which can be trained without person identification tags or any prior knowledge, accordingly. A quicker R-CNN network that has been programmed for person detection comes before every re-id network as an extender [204]. In Table 2, the following methodology is quantitatively compared to six splitting person approaches [ [208], [209], [210], [20], [21], [22]] from CUHK-SYSU and the standard PRW gallery. Table 3 compares the suggested move's efficacy to that of the two different unsupervised state-ofthe-art re-id approaches [23] [24], as measured in the CUHK-SYSU and both PRW galleries.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Method | MGTS [1] | | DMRN  et [19] | SeqNet [21] | AlignPS [22] | PGA [26] | OR [23] | **Proposed** |
| Supervised | Yes | | Yes | Yes | Yes | Yes | Yes | No |
| CUHKSYSU | map | 83.3 | 93.2 | 94.8 | 94.0 | 92.3 | 93.2 | **81.1** |
| Top-1 | 84.2 | 94.2 | 95.7 | 94.5 | 94.7 | 93.8 | **83.2** |
| PRW | map | 32.8 | 46.9 | 47.6 | 46.1 | 44.2 | 52.3 | **41.7** |
| Top-1 | 72.1 | 83.3 | 87.6 | 82.1 | 85..2 | 71.5 | **86.0** |

Table 3: Comparative analysis of the normal PR gallery and the supervised person selection techniques now in use on CUHK-SYSU.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Method** | | DET+BU  C [14] | DET+ML  C [24] | **Propose d** | DET+BU  C [14] | DET+ML  C [24] | **Propose d** |
| CUHK  -SYSU | map | 74.8 | 69.2 | **81.1** | - | - | **-** |
| Top-1 | 77.4 | 73.7 | **83.2** | - | - | **-** |
| PRW | map | 26.0 | 25.4 | **41.7** | 18.6 | 17.1 | **36.6** |
| Top-1 | 83.6 | 84.7 | **86.0** | 53.0 | 50.8 | **64.9** |
| Galler y |  | **Regular** |  |  | **Multi-view** |  |

Table 4: The expanded non-supervised person Re-ID techniques, the required value of the normal and multi-view galleries for PRW, as well as CUHK-SYSU.



Figure 11: Cases that present difficulties for the recommended unsupervised clustering approach A query figure is depicted in the first picture from the left, and the found individuals are marked in green and red, respectively, for true and false matches.

Even without the labels of person identification, the quantifiable performance of the recommended strategy, as shown in Table 2, achieves more than 85% of the high spots among the most sophisticated supervised systems on both datasets in terms of map and Top-1. Additionally, the recommended approach outperforms the expanded unsupervised person re-id approaches in both scenarios. As shown in Table 3, where the bottom block presents the outcomes assessed in the challenging multi-view gallery of PRW, the suggested technique, which received the maximum map level of 36.6 as opposed to the second-best approach, which received 18.6 map rating, virtually enhances the results twice [55].

It must be remarked that the suggested approach and the extenders of BUC and MLC create person groupings mainly by utilizing the discovered persons' performances in unsupervised cases [48] [49]. As an outcome, they commonly deliver non-negligible similarities among various individuals with similar characteristics, especially in the PRW test set, which also consists of many people wearing similar outfits, as seen in Figure 8. We achieve good Top-1 ratings in these circumstances, but only mediocre map ratings. Because of the size limitations, the qualitative research data is included in the supporting information.

# Chapter 5

## **5. Conclusion and future work**

In this analysis of recent re-id improvements concentrates specifically on non-supervised techniques that deal with the problem of Scalability of data [142][145]. Despite the impressive advances made in recent years, there are still substantial issues associated with person re-id, including domain specific information deficiency, camera-invariant person feature learning, identity label estimation, and discriminative person feature learning [138][140]. A few promising new research directions are briefly discussed.

**5.1.1 Unsupervised end-to-end person re-identification** despite the fact that most person re-id works employ hand-annotated or detected bounding boxes to train the re-id, it is more practical to learn a person re-id model end-to-end in which raw images are used directly to train the re-id model. The idea of integrating person to person enquiry and person re-id into one procedure was put out by [37]. Different supervised end-to-end methods for re-identifying individuals have been developed since then. The use of large unlabeled pictures to develop re-id models in a trained and tested end-to-end fashion, which is more useful in world-at-large use, is currently lacking [44] [45]. As a result, we believe that unsupervised end-to-end person re-identification presents a significant research opportunity, particularly when using evolutionary vision transformers [31] [32] [33] [34].

**5.1.2 Personality re-identification without supervision** that may be used in general Unsupervised cross-domain person re-id models rely on the assumption that the source and destination domains have a lot in common with regard to data. The knowledge acquired in the source domain may therefore be applied to the target domain [137][139][141]. However, since each target area must go through this extensive adaptation process, its application to arbitrary domains is limited [28]. The success of unsupervised cross-domain person re-id models is contingent on the quality of the connections between the source and destination domains. When the source domain data is sufficiently diverse to cover the target domain, the pre-trained re-id model on the source domain may be able to deploy to the target domain without the need for domain adaptation [41] [42] [43]. In this case, if the target domain data is sufficiently varied to cover the source domain, the pretrained re-id model on the source domain may be able to deploy to the target domain without domain adaptation. Specifically, the person re-id that may be used in a generalized fashion without supervision is achieved. In order to be effective, a re-Id model has to be able to generalize to new conditions with ease [133][134]. Therefore, it is necessary to provide the model with sufficient synthesized and annotated data covering a wide range of photometric and geometrical variables for training. Both DIMN [27] and OsNet [30] attempted to train generic representations of human characteristics using publicly labeled data. In the hope that they will work in the unseen domains, person representations are averaged among several representations [159][154][155]. However, because the data distributions in the target domains were not explored, these models didn't perform well. The major obstacle to gathering enough labeled data to train models that are well-generalizable to uncharted areas is labeling costs. Unlabeled data, on the other hand, is easy to gather and will cover more variances for the purpose of re-id [40]. Therefore, unsupervised generalizable person reidentification is an encouraging route to go in since it can be implemented at no cost in any area [146].

**5.1.3 Unsupervised cross-domain person re-identification using text** person re-identification [26] seeks to identify a target individual for a given textual description from a collection of photos or videos. When the target person's textual description is accessible but the query image is not, text-based re-id is useful. In a situation of emergency like the search for a missing person, it is extremely valuable. It is difficult to pre-train a text-based re-id model using annotated data and adapt it to unlabeled domains, while it is necessary to train a text-based re-id model for each environment [38]. The large amount of unlabeled data can be used to build person feature representations for the target domains, similar to the person who is based on images and is unsupervised across domains re-id. For text-based domain adaptation, the adaptation method must align word-level properties between the domains of origin and destination rather than anticipate global pseudo-labels. The attention mechanism and different distribution distance losses can help you learn about how features are represented [39].

### 5.2 Conclusion

In this research, we examine various current unsupervised re-identification techniques and address some existing issues with Person Re-ID. We divide the unsupervised person Re-ID issue into two categories: simply unsupervised learning person re-id (USL) and unsupervised domain adaptation person re-id (UDL) (UDA) [158][160]. We discovered that the performance of the existing USL and UDA approaches is identical. We explored contextual features of the person search framework and presented two unsupervised clustering approaches for classifying unlabeled person samples: the uniqueness-based SNM and the coappearance-based SPM. We performed the comparison tests [37]. According to experimental results, the recommended strategy performs better than extended unsupervised person search techniques while delivering performance that is equal to modern, state of-the-art supervised person search methods re-id techniques [54].

# References

1. H. F. . L. Z. and Y. Y. , Unsupervised Person Re-identification:Clustering and Fine-tuning, arxiv.org, 29 Jun 2017.
2. K. H. X. Z. S. R. and J. S. , "Deep Residual Learning for Image Recognition," 2016 IEEE Conference.
3. D. W. and S. Z. , "Unsupervised Person Re-Identification via Multi-Label Classification," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020.
4. Y. L. X. D. L. Z. Y. Y. and Y. Y. , "A Bottom-Up Clustering Approach to Unsupervised Person Re-Identification," 2019-07-17 Proceedings of the AAAI Conference on Artificial Intelligence.
5. M. W. B. L. J. H. X. G. and X.-S. H. , Camera-aware Proxies for Unsupervised Person Re-Identification, arXivLabs, 19 Dec 2020.
6. S. Z. and S. X. , "Intra-Inter Camera Similarity for Unsupervised Person Re-Identification," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021.Deep Learning for Person Re-identification: A Survey and Outlook.
7. G. W. Z. D. W. Y. X. L. P. T. and S. Z., "Cluster Contrast for Unsupervised Person Re-Identification," 22 Mar 2021.
8. L. e. a. Wei, "Person transfer gan to bridge domain gap for person re-identification," San Antonio, Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.
9. G. Y. D. C. and H. L. , "Mutual mean-teaching: Pseudo label refinery for unsupervised domain adaptation on person re-identification.," (2020)..
10. . Z. and Z. e. a. , "Invariance matters: Exemplar memory for domain adaptive person reidentification," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019.Deep Learning for Person Re-identification: A Survey and Outlook..
11. S. Ren, K. H. R. G. and J. S. , Faster r-cnn: Towards realtime object detection with region proposal networks. In Advances in Neural Information Processing Systems, 2015.

1. Y. L. and Y. Y. , A Bottom-Up Clustering Approach to Unsupervised Person Re-Identification, PKP Publishing Services Network, In Proceedings of the AAAI Conference 2019.
2. Y. L. L. X. Y. W. and C. Y. , "Unsupervised Person Re-identification via Softened Similarity Learning," In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2020..
3. T. X. S. L. B. W. L. L. and X. W. , "Joint Detection and Identification Feature Learning for Person Search," 2017 2nd International Conference on Electrical & Electronic Engineering (ICEEE).
4. . L. Z. H. Z. S. S. M. C. Y. Y. and Q. , "Person re-identification in the wild," In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017.
5. K. H. X. Z. S. R. and a. J. S. , "Deep residual learning for image recognition.," In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016..
6. J. D. P. Z. H. L. and H. W. , Dynamic imposter based online instance matching for person search, Dalian 116023, China: ScienceDirect, April 2020.
7. C. H. Z. Z. C. G. . N. S. and a. Y. Y. , "Decoupled and memory-reinforced networks Towards effective feature learning for one-step person search.," In Proceedings of the AAAI Conference on Artificial Intelligence, 2021.
8. H. K. S. J. . I.-J. K. and a. K. S. , " Prototype-guided saliency feature learning for person search.," In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2021.
9. Z. L. and D. M. , " Sequential end-to-end network for efficient person search.," In Proceedings of the AAAI Conference on Artificial Intelligence, 2021..
10. Y. Y. . J. L. J. Q. S. B. S. L. . L. L. F. Z. and . L. S. , "Anchor-free person search.," In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2021.
11. H. Y. and C. X. , "Joint person objectness and repulsion for person search," IEEE Transactions on Image Processing, 2021.
12. W. S. H. L. F. M. and W. H. , "Instance Enhancing Loss: Deep Identity-Sensitive Feature Embedding for Person Search," 2018 25th IEEE International Conference on Image Processing (ICIP).

1. . D. W. and S. Z. , "Unsupervised person re-identification via multilabel classification," In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2020.
2. L. Q. L. W. J. H. L. Z. Y. S. and Y. G. , "A novel unsupervised camera-aware domain adaptation framework for person re-identification.," In Proceedings of the IEEE Conference on International Conference on Computer Vision, 2019..
3. P. Peng, T. Xiang, Y. Wang, M. Pontil, S. Gong and Huang, "Unsupervised cross-dataset transfer learning for," *In CVPR,* 2016.
4. W.-S. Z. H.-X. Y. S. G. J. L. A. Wu, *computer vision foundation,* no. Proceedings of the IEEE International Conference on Computer Vision (ICCV), p. pp. 5380–5389, 2017.
5. W.-S. Z. H.-X. Y. S. G. J. L. Ancong Wu, "RGB-Infrared Cross-Modality Person Re-Identification," *Proceedings of the IEEE International Conference on Computer Vision (ICCV),* pp. pp. 5380-5389, 2017.
6. Y. Y. Y.-Z. s. T. T. M. j. song, "person Re-Identification by domain-Invariant mapping Network," pp. pp. 719-728, 2019.
7. J. S. S. M. I. G. L. T. X. L. S. a. S. C. H. H. F. I. ,. Mang Ye, "Deep Learning for Person Reidentification: A Survey and Outlook," *IEEE,* vol. vol.44, pp. pp 2872 - 2893, 2020.
8. F. M. G. S. N. U. A. K. Z. N. Carion, "N. Carion, F. Mass End-to-end object detection with transformers, , Cham,2020, pp.," *Springer International Publishing,* pp. pp. 213-229, 2020.
9. D. T. B. C. N. D. T. D. B. K. L. X.-B. Nguyen, " Clusformer: A Transformer Based Clustering Approach to Unsupervised Large-Scale Face and Visual Landmark Recognition," p. pp. 10847–10856, 2021.
10. Y. L. Y. C. H. H. Y. W. Z. Z. S. L. B. G. Z. Liu, " Swin Transformer: Hierarchical Vision Transformer using Shifted Windows,," *arXiv:2103.14030 [cs]ArXiv: 2103.14030.,* (Aug. 2021).
11. X. Z. S. G. W. Li, "Harmonious Attention Network for Person Re-Identification," p. pp. 2285–2294., 2018.
12. A. W. W.-S. Z. H.-X. Yu, "Cross-View Asymmetric Metric Learning for Unsupervised Person Re-Identification," p. pp. 994–1002, 2017.
13. J. W. X. Z. S. G. W. Li1, "Transferable Joint Attribute-Identity Deep Learning for Unsupervised Person Re-Identification", 2018.
14. S. M. D. X. S. L. Haojie Liu, "SFANet: A Spectrum-aware Feature Augmentation Network for Visible-Infrared Person Re-Identification," *IEEE transactions on neural networks and learning systems,* no. Computer Science, 24 February 2021.
15. W. Z. ,. R. S. ,. Q. Z. ,. X. L. Xuanyu He1, "Take More Positives: An Empirical Study of Contrastive Learing in Unsupervised Person Re-Identification," no. Shandong University 2Hong Kong Baptist University 3Meituan, 2021.
16. X. Y. Z. Y. B. K. a. J. K. Yang Zou, "Joint disentangling and adaptation for cross-domain person re-identification," *In ECCV,* 2020.
17. L. Z. Z. L. S. L. a. Y. Y. Zhun Zhong, "Invariance matters: Exemplar memory for domain adaptive person re-identification," *In CVPR,* 2019.
18. Q. Y. S. L. M. J. R. J. a. Y. T. Yunpeng Zhai, "Multiple expert brainstorming for domain adaptive person re-identification," *In ECCV,* 2020.
19. J. C. C. S. a. M. Y. Xinyu Zhang, "Self-training with progressive augmentation for unsupervised cross-domain person re-identification," *In ICCV,* 2019.
20. L. Z. Y. Y. Q. T. a. S. Yifan Sun, "Beyond part models: Person retrieval with refined part pooling (and a strong convolutional baseline).," *In ECCV,* 2018.
21. X. S. a. L. Zheng., "Dissecting person re identification from the viewpoint of viewpoint.," *In CVPR,* 2019.
22. F. S. R. Z. R. C. C. T. Ergys Ristani, "Performance measures and a data set for multi-target, multi-camera tracking.," *In ECCV,,* 2016.
23. D. C. F. Z. R. Z. a. H. s. L. Yixiao Ge, "Self-paced contrastive learning with hybrid memory for domain adaptive object re-id.," *In NeurIPS, ,* 2020..
24. M. E. a. L. S. Davis., "Deep residual learning in the jpeg transform domain.," *In ICCV,,* 2019..
25. "Domain separation networks.," *In NeurIPS,* , 2016..

1. S. K. M. N. a. G. f. H. Ting Chen, "A simple framework for contrastive learning of visual representations.," *In ICML,,* 2020..
2. C. X. C. X. a. D. T. Yunhe Wang, "Packing convolutional neural networks in the frequency domain," *TPAMI,,* vol. 41(10), pp. pp- 2495–2510,, 2018..
3. W. Y. W. G. Z. Y. S. H. a. H. T. Fu Yang, "Self-similarity grouping: A simple unsupervised cross domain adaptation approach for person re-identification..," *In ICCV,* 2019..
4. Z. Z. Z. L. Y. C. L. S. L. a. N. S. Fengxiang Yang, " Joint noise-tolerant learning and meta camera shift adaptation for unsupervised person re-identification.," *In AAAI, ,* 2021.
5. N. Wojke and A. Bewley, “Deep cosine metric learning for person re-identification,” 2018, doi: 10.1109/WACV.2018.00087.
6. J. Garcia, N. Martinel, A. Gardel, I. Bravo, G. L. Foresti, and C. Micheloni, “Discriminant context information analysis for post-ranking person re-identification,” *IEEE Trans. Image Process.*, 2017, doi: 10.1109/TIP.2017.2652725.
7. H. Fan, L. Zheng, C. Yan, and Y. Yang, “Unsupervised Person Re-identification,” *ACM Trans. Multimed. Comput. Commun. Appl.*, 2018, doi: 10.1145/3243316.
8. G. Lisanti, I. Masi, A. D. Bagdanov, and A. Del Bimbo, “Person re-identification by iterative re-weighted sparse ranking,” *IEEE Trans. Pattern Anal. Mach. Intell.*, 2015, doi: 10.1109/TPAMI.2014.2369055.
9. R. Müller, S. Kornblith, and G. Hinton, “When does label smoothing help?,” 2019.
10. R. R. Varior, B. Shuai, J. Lu, D. Xu, and G. Wang, “A siamese long short-term memory architecture for human re-identification,” 2016, doi: 10.1007/978-3-319-46478-7\_9.
11. M. Li, X. Zhu, and S. Gong, “Unsupervised Person Re-identification by Deep Learning Tracklet Association,” 2018, doi: 10.1007/978-3-030-01225-0\_45.
12. S. Bąk, P. Carr, and J. F. Lalonde, “Domain Adaptation Through Synthesis for Unsupervised Person Re-identification,” 2018, doi: 10.1007/978-3-030-01261-8\_12.
13. L. Fei-Fei, R. Fergus, and P. Perona, “One-shot learning of object categories,” *IEEE Trans. Pattern Anal. Mach. Intell.*, 2006, doi: 10.1109/TPAMI.2006.79.
14. S. Bąk and P. Carr, “One-shot metric learning for person re-identification,” 2017, doi:

10.1109/CVPR.2017.171.

1. D. Figueira, L. Bazzani, H. Q. Minh, M. Cristani, A. Bernardino, and V. Murino, “Semisupervised multi-feature learning for person re-identification,” 2013, doi:

10.1109/AVSS.2013.6636625.

1. X. Liu, M. Song, D. Tao, X. Zhou, C. Chen, and J. Bu, “Semi-supervised coupled dictionary learning for person re-identification,” 2014, doi: 10.1109/CVPR.2014.454.
2. W. Chen, X. Chen, J. Zhang, and K. Huang, “Beyond triplet loss: A deep quadruplet network for person re-identification,” 2017, doi: 10.1109/CVPR.2017.145.

L. Zheng, L. Shen, L. Tian, S. Wang, J. Wang, and Q. Tian, “Scalable Person Re-identification : A Benchmark Scalable Person Re-identification : A Benchmark,” *IEEE Int. Conf. Comput. Vis.*, 2017.

1. Z. Zheng, L. Zheng, and Y. Yang, “A discriminatively learned CNN embedding for person reidentification,” *ACM Trans. Multimed. Comput. Commun. Appl.*, 2017, doi: 10.1145/3159171.
2. H. Luo *et al.*, “A Strong Baseline and Batch Normalization Neck for Deep Person ReIdentification,” *IEEE Trans. Multimed.*, 2020, doi: 10.1109/TMM.2019.2958756.
3. F. Wang, W. Zuo, L. Lin, D. Zhang, and L. Zhang, “Joint learning of single-image and cross-image representations for person re-identification,” 2016, doi: 10.1109/CVPR.2016.144.
4. X. Liu, S. Zhang, and M. Yang, “Self-Guided Hash Coding for Large-Scale Person Reidentification,” 2019, doi: 10.1109/MIPR.2019.00051.
5. L. Wei, X. Liu, J. Li, and S. Zhang, “VP-ReID: Vehicle and person re-identification system,” 2018, doi: 10.1145/3206025.3206086.
6. D. Chen, D. Xu, H. Li, N. Sebe, and X. Wang, “Group Consistent Similarity Learning via Deep CRF for Person Re-identification,” 2018, doi: 10.1109/CVPR.2018.00902.
7. Y. Suh, J. Wang, S. Tang, T. Mei, and K. M. Lee, “Part-aligned bilinear representations for person re-identification,” 2018, doi: 10.1007/978-3-030-01264-9\_25.
8. Y. Yuan, W. Chen, Y. Yang, and Z. Wang, “In defense of the triplet loss again: Learning robust person re-identification with fast approximated triplet loss and label distillation,” 2020, doi: 10.1109/CVPRW50498.2020.00185.
9. H.Shi *et al.*, “Embedding deep metric for person Re-identification: A study against large variations,” 2016, doi: 10.1007/978-3-319-46448-0\_44.
10. S. Zhou, J. Wang, J. Wang, Y. Gong, and N. Zheng, “Point to set similarity based deep feature learning for person re-identification,” 2017, doi: 10.1109/CVPR.2017.534.
11. Y. Fu *et al.*, “Self-similarity grouping: A simple unsupervised cross domain adaptation approach for person re-identification,” 2019, doi: 10.1109/ICCV.2019.00621.
12. L. Wei, S. Zhang, W. Gao, and Q. Tian, “Person Transfer GAN to Bridge Domain Gap for Person Re-identification,” 2018, doi: 10.1109/CVPR.2018.00016.
13. Z. Zhong, L. Zheng, Z. Zheng, S. Li, and Y. Yang, “Camera Style Adaptation for Person Re-identification,” 2018, doi: 10.1109/CVPR.2018.00541.
14. W. Deng, L. Zheng, Q. Ye, G. Kang, Y. Yang, and J. Jiao, “Image-Image Domain Adaptation with Preserved Self-Similarity and Domain-Dissimilarity for Person Reidentification,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 994–1003, 2018, doi: 10.1109/CVPR.2018.00110.
15. Y. Lin, X. Dong, L. Zheng, Y. Yan, and Y. Yang, “A bottom-up clustering approach to unsupervised person re-identification,” *33rd AAAI Conf. Artif. Intell. AAAI 2019, 31st Innov. Appl. Artif. Intell. Conf. IAAI 2019 9th AAAI Symp. Educ. Adv. Artif. Intell. EAAI 2019*, pp. 8738–8745, 2019, doi: 10.1609/aaai.v33i01.33018738.
16. H. Fan, L. Zheng, C. Yan, and Y. Yang, “Unsupervised person re-identification: Clustering and fine-tuning,” *ACM Trans. Multimed. Comput. Commun. Appl.*, 2018, doi: 10.1145/3243316.
17. Q. Yang, H. X. Yu, A. Wu, and W. S. Zheng, “Patch-based discriminative feature learning for unsupervised person re-identification,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2019-June, pp. 3628–3637, 2019, doi: 10.1109/CVPR.2019.00375.2
18. D. Wang and S. Zhang, “Unsupervised person re-identification via multi-label classification,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 10978–10987, 2020, doi: 10.1109/CVPR42600.2020.01099.
19. D. Yi, Z. Lei, S. Liao, and S. Z. Li, “Deep metric learning for person re-identification,” 2014, doi: 10.1109/ICPR.2014.16.
20. H. Chen *et al.*, “Deep Transfer Learning for Person Re-Identification,” 2018, doi: 10.1109/BigMM.2018.8499067.
21. M. Ye, J. Shen, G. Lin, T. Xiang, L. Shao, and S. C. H. Hoi, “Deep Learning for Person Re-identification: A Survey and Outlook,” *IEEE Trans. Pattern Anal. Mach. Intell.*, 2021, doi: 10.1109/TPAMI.2021.3054775.
22. A. Nambiar, A. Bernardino, and J. C. Nascimento, “Gait-based Person Re-identification: A Survey,” vol. 52, no. 2, 2019.
23. Z. Yu *et al.*, “Apparel-invariant Feature Learning for Person Re-identification,” *IEEE Trans. Multimed.*, pp. 1–10, 2021, doi: 10.1109/TMM.2021.3119133.
24. H. X. Yu, A. Wu, and W. S. Zheng, “Unsupervised Person Re-Identification by Deep Asymmetric Metric Embedding,” *IEEE Trans. Pattern Anal. Mach. Intell.*, 2020, doi: 10.1109/TPAMI.2018.2886878.
25. X. Wang, “Intelligent multi-camera video surveillance: A review,” *Pattern Recognit. Lett.*, 2013, doi: 10.1016/j.patrec.2012.07.005.
26. D. Gerónimo, A. M. López, A. D. Sappa, and T. Graf, “Survey of pedestrian detection for advanced driver assistance systems,” *IEEE Trans. Pattern Anal. Mach. Intell.*, 2010, doi: 10.1109/TPAMI.2009.122.
27. P. Dollár, C. Wojek, B. Schiele, and P. Perona, “Pedestrian detection: A benchmark,” 2009, doi: 10.1109/CVPRW.2009.5206631.
28. E. Insafutdinov *et al.*, “ArtTrack: Articulated multi-person tracking in the wild,” 2017, doi:

10.1109/CVPR.2017.142.

1. E. Ristani and C. Tomasi, “Features for Multi-target Multi-camera Tracking and Reidentification,” 2018, doi: 10.1109/CVPR.2018.00632.
2. A. J. Ma, P. C. Yuen, and J. Li, “Domain transfer support vector ranking for person reidentification without target camera label information,” 2013, doi: 10.1109/ICCV.2013.443.
3. T. Xiao, H. Li, W. Ouyang, and X. Wang, “Learning deep feature representations with domain guided dropout for person re-identification,” 2016, doi: 10.1109/CVPR.2016.140.
4. L. Zheng, H. Zhang, S. Sun, M. Chandraker, Y. Yang, and Q. Tian, “Person re-identification in theWild,” 2017, doi: 10.1109/CVPR.2017.357.
5. D. Yi, Z. Lei, and S. Z. Li, “Deep Metric Learning for Practical Person Re-Identification,” *J. Mach. Learn. Res.*, 2014.
6. Y. Zhai, X. Guo, Y. Lu, and H. Li, “In defense of the classification loss for person reidentification,” 2019, doi: 10.1109/CVPRW.2019.00194.
7. Z. Zhong, L. Zheng, D. Cao, and S. Li, “Re-ranking person re-identification with kreciprocal encoding,” 2017, doi: 10.1109/CVPR.2017.389.
8. M. Ye *et al.*, “Person Reidentification via Ranking Aggregation of Similarity Pulling and Dissimilarity Pushing,” *IEEE Trans. Multimed.*, 2016, doi: 10.1109/TMM.2016.2605058.
9. Q. Yang, H. X. Yu, A. Wu, and W. S. Zheng, “Patch-based discriminative feature learning for unsupervised person re-identification,” 2019, doi: 10.1109/CVPR.2019.00375.
10. S. Wu and L. Gao, “Cross-Camera erased feature learning for unsupervised person reidentification,” *Algorithms*, vol. 13, no. 8, 2020, doi: 10.3390/A13080193.
11. M. Li, C.-G. Li, and J. Guo, “Cluster-guided Asymmetric Contrastive Learning for

Unsupervised Person Re-Identification,” 2021, [Online]. Available: http://arxiv.org/abs/2106.07846.

[109] T. Wang, S. Gong, X. Zhu, and S. Wang, “Person re-identification by video ranking,” 2014, doi: 10.1007/978-3-319-10593-2\_45.

[110] O. Camps *et al.*, “From the Lab to the Real World: Re-identification in an Airport Camera Network,” *IEEE Trans. Circuits Syst. Video Technol.*, 2017, doi: 10.1109/TCSVT.2016.2556538.

[111] F. Liu and L. Zhang, “View confusion feature learning for person re-identification,” 2019, doi: 10.1109/ICCV.2019.00674.

[112] T. Isobe, D. Li, L. Tian, W. Chen, Y. Shan, and S. Wang, “Towards Discriminative Representation Learning for Unsupervised Person Re-identification,” pp. 8506–8516, 2022, doi: 10.1109/iccv48922.2021.00841.

[113] K. Zeng, M. Ning, Y. Wang, and Y. Guo, “Energy clustering for unsupervised person re-identification,” *Image Vis. Comput.*, vol. 98, pp. 1–8, 2020, doi: 10.1016/j.imavis.2020.103913.

[114] C. C. Loy, T. Xiang, and S. Gong, “Multi-camera activity correlation analysis,” 2009, doi: 10.1109/CVPRW.2009.5206827.

[115] M. Hirzer, C. Beleznai, P. M. Roth, and H. Bischof, “Person re-identification by descriptive and discriminative classification,” 2011, doi: 10.1007/978-3-642-21227-7\_9.

[116] L. Zheng *et al.*, “Mars: A video benchmark for large-scale person re-identification,” 2016, doi: 10.1007/978-3-319-46466-4\_52.

[117] W. Li, R. Zhao, and X. Wang, “Human re-identification with transferred metric learning,” 2013, doi: 10.1007/978-3-642-37331-2\_3.

[118] X. Liu and S. Zhang, “Domain Adaptive Person Re-Identification via Coupling Optimization,” 2020, doi: 10.1145/3394171.3413904.

[119] Z. Hu, C. Zhu, and G. He, “Hard-sample Guided Hybrid Contrast Learning for Unsupervised Person Re-Identification,” 2021, doi: 10.1109/IC-NIDC54101.2021.9660560.

[120] X. Zhang, J. Cao, C. Shen, and M. You, “Self-training with progressive augmentation for unsupervised cross-domain person re-identification,” 2019, doi: 10.1109/ICCV.2019.00831.

[121] S. Zhou, F. Wang, Z. Huang, and J. Wang, “Discriminative feature learning with consistent attention regularization for person re-identification,” 2019, doi: 10.1109/ICCV.2019.00813.

[122] Z. Liu, P. Luo, X. Wang, and X. Tang, “Deep learning face attributes in the wild,” 2015, doi: 10.1109/ICCV.2015.425.

[123] Y. Ge, F. Zhu, D. Chen, R. Zhao, and H. Li, “Self-paced contrastive learning with hybrid memory for domain adaptive object re-ID,” 2020.

[124] S. Zhu, X. Gong, Z. Kuang, and J. Du, “Partial person re-identification with two-stream network and reconstruction,” *Neurocomputing*, 2020, doi: 10.1016/j.neucom.2019.04.098.

[125] O. Camps *et al.*, “From the Lab to the Real World: Re-identification in an Airport Camera Network,” *IEEE Trans. Circuits Syst. Video Technol.*, 2017, doi: 10.1109/TCSVT.2016.2556538.

[126] S. Karanam, M. Gou, Z. Wu, A. Rates-Borras, O. Camps, and R. J. Radke, “A Systematic Evaluation and Benchmark for Person Re-Identification: Features, Metrics, and Datasets,” *IEEE Trans. Pattern Anal. Mach. Intell.*, 2019, doi: 10.1109/TPAMI.2018.2807450.

[127] K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick, “Momentum Contrast for Unsupervised Visual Representation Learning,” 2020, doi: 10.1109/CVPR42600.2020.00975.

[128] H. Chen, B. Lagadec, and F. Bremond, “ICE: Inter-instance Contrastive Encoding for Unsupervised Person Re-identification,” 2022, doi: 10.1109/iccv48922.2021.01469.

[129] M. Ye, J. Shen, X. Zhang, P. C. Yuen, and S. F. Chang, “Augmentation Invariant and Instance Spreading Feature for Softmax Embedding,” *IEEE Trans. Pattern Anal. Mach. Intell.*, 2022, doi: 10.1109/TPAMI.2020.3013379.

[130] D. Gray, S. Brennan, and H. Tao, “Evaluating appearance models for recognition, reacquisition, and tracking,” *10th Int. Work. Perform. Eval. Track. Surveill. (PETS),* 2007.

[131] Z. T. D. Salman Khan1, "Dispersion based Clustering for Unsupervised Person Re-identification.," *Cite As: arXiv:1906.01308 [cs.CV],* no. DOI: 10.48550/arXiv.1906.01308, pp. pp. 1-12, 2019.

[132] J. Z. N. W. X. G. De Cheng, "Hybrid Dynamic Contrast and Probability Distillation for Unsupervised Person Re-Id," *IEEE,* vol. 31, no. DOI: 10.1109/TIP.2022.3169693, 2022.

|  |  |
| --- | --- |
| [133] L. W. ,. Z. T. N. Z. H. Haoxuanye Ji, "Meta Pairwise Relationship Distillation for Unsupervised person Re-identification," *computer vision foundation,* no. IEEE Xplore, pp. 1-10, 2021. | |
| [134] X. Z. ,. X. L. X. L. T. H. ,. a. S. W. Zilong Ji, "An Attention-Driven Two-Stage Clustering  Method for Unsupervised Person Re-identiﬁcation.," ECCV, no. European Conference on  computer Vision, pp. 1-17, 2020. | |
| [135] D. L. ,. L. T. ,. W. C. Y. S. S. W. Takashi Isobe, "Towards Discriminative Representation  Learning for Unsupervised Person Re-identiﬁcation.," IEEE Xplore, no. {dongl, lutian,  yishan}@xilinx.com jbj18@mails.tsinghua.edu.cn, pp. 1-11, 2021.  [136] T. L. Y. J. T. W. Y. L. Xue Li, "Camera-aware Style Separation and Contrastive  Learning for Unsupervised Person Re-identification.," Computer Vision and Pattern  Recognition*,* no. DOI: 10.48550/arXiv.2112.10089, pp. 1-6, 2021. | |
| [137] | C. R.-i. i. a. C. network, "Abir Das, Anirban Chakraborty, A. Roy-Chowdhury," *Computer Science,* no. ECCV , 2014. |
| [138] | H. I. a. M. S. Shayan Modiri Assari, "Human Re-identification in Crowd Videos using Personal, Social and Environmental Constraints," *Center for Research in Computer Vision (CRCV),,* no. {smodiri, haroon, shah}@cs.ucf.edu}, pp. 1-17, 2016. |
| [139] | Z. W. Z. W. Y. Z. Y.-Y. C. S. Satoh, "Learning to Reduce Dual-level Discrepancy for Infrared-Visible Person," *National Institute of Informatics,* no. The University of Tokyo, pp. 1-9, 2019. |
| [140] | N. W. L. X. G. Yi Hao, "HSME: Hypersphere Manifold Embedding for Visible Thermal Person Re-Identification.," *in AAAI,* pp. 8385–8392., 2019. |
| [141] | X. L. L. ,. Z. C. Z. ,. Y. S. L. ,. Y. R. Y. W Wang, "Deep reinforcement learning attention selection for person re-identification.," *In BMVC,* 2017. |
| [142] | P. C. Y. Mang Ye, "PurifyNet: A Robust Person Re-Identification Model With Noisy Labels," *IEEE TIFS,* Vols. article num- 8976262, pp. 1-12, 2020. |
| [143] | S. L. L. S. Yanan Wang, "surpassing real-world source training data : Random 3rd characters for generalizable person re-identification," *ACM MM,*  pp. 3422–3430., 2020. |
| [144] | S. P. D. S. R. M. L. Y. W. a. A. K. R.-C. Xueping Wang, "Learing Person Re-identification Models from videos with Weak supervision," Computer Vision and Pattern Recognition, no. arXiv preprint arXiv:2007.10631, pp. 1-11, 2020. |
| [145] | M. A. B. A. S. Siyu Tang, "Multiple People Tracking by Lifted Multicut and Person Re-identification.," in CVPR, pp. 3539–3548., 2017. |
| [146] | S. G. T. X. Weishi Zheng, "Associating Groups of People," BMVC, pp. 1–23, 2009. |
| [147] | S. G. X. Z. a. T. X. Hanxiao Wang(B), "Human-in-the-Loop Person Re-identification.," in ECCV, no. {hanxiao.wang,s.gong,xiatian.zhu,t.xiang}@qmul.ac.uk, pp. 405–422., 2016. |
| [148] | C. D. X. S. X. J. X. G. Z. Y. F. H. R. J. F. Zheng, "Pyramidal person re-identification via multi-loss dynamic training," in CVPR, pp. 8514–8522, 2019. |
| [149] | B. S. a. Y. W. J. Zhou, " Easy Identification from Better Constraints: Multi-Shot Person Re-Identification.," in CVPR, pp. 5373–5381., 2018. |
| [150] | A. O.-W. P. Re-Identification, "Xiang Li, Ancong Wu, Wei-Shi Zheng," in Proceedings of the European Conference on Computer Vision (ECCV), pp. 280-296, 2018. |
| [151] | M.-S. L. w. G. P. W. f. D. M. Learning, "Multi-Similarity Loss with General Pair Weighting for Deep Metric Learning," Shenzhen Malong Artificial Intelligence Research Center, Shenzhen, China, no. in CVPR, p. pp. 5022–5030, 2019. |
| [152] | Z. S. Y. Z. Y. W. Lingxiao He, "Recognizing Partial Biometric Patterns," arXiv preprint arXiv:1810.07399, 2018. |
| [153] | K. Y. Z. Z. A. K. U. D. C. C. L. S. B. Y. W. J. J. X. Z. Z. H. Chengjun Kang, "Growing single crystals of two-dimensional covalent organic frameworks enabled by intermediate tracing study," Nature Communications , no. Article number: 1370 , 2022. |
| [154] | Z. W. W. H. C.-W. L. a. S. S. Z. Huang, ''Learning to Reduce Dual-level Discrepancy for Infrared-Visible Person Re-identification,'' in ACM MM, pp. 1888–1896., 2019. |
| [155] | S. L. L. S. Yanan Wang, "Surpassing Real-World Source Training Data: Random 3D Characters for Generalizable Person Re-Identification," *in ACM MM,* no. Computer Vision and Pattern Recognition, p. pp. 3422–3430., 2020. |
| [156] | A. B. A. R.-C. R. Panda, "Adaptation of person re-identification models for on-boarding new camera(s)," Computer Science, no. Corpus ID: 201249794, December 2019. |
| [157] | A. B. V. M. A. R.-C. R. Panda, "Unsupervised Adaptive Re-identification in Open World Dynamic Camera Networks," in IEEE Conference on Computer Vision and Pattern, 2017. |
| [158] | L. S. Shengcai Liao, "Interpretable and Generalizable Person Re-Identification with Query-Adaptive Convolution and Temporal Lifting.," in European Conference on Computer Vision, 2020. |
| [159] | B. L. J. H. X. G. X.-S. H. Menglin Wang, "Camera-aware Proxies for Unsupervised Person Re-Identification.," no. Computer Vision and Pattern Recognition, pp. 1-9, 2021. |

[160] L. Z. Y. Zhedong Zheng, "Unlabeled Samples Generated by GAN Improve the Person Re-identification Baseline in vitro.," in ICCV, no. Computer Vision and Pattern Recognition, pp. 1-10, January 2017.

[161] A. Das, A. Chakraborty, A. K. Roy-Chowdhury, Consistent re-identification in

a camera network, in: European Conference on Computer Vision, Springer , 2014,

pp.330-345.

[162] C. Mao, Y. Li, Z. Zhang, Y. Zhang, X. Li, Pyramid person matching network for person

re-identification, in: Asian Conference on Machine Learning, 2017, pp. 487–497.

[163] C. C. Loy, T. Xiang, and S. Gong, “Multi-camera activity correlation analysis,” 2009,

doi: 10.1109/CVPRW.2009.5206827.

[164] D. Gray, S. Brennan, and H. Tao, “Evaluating appearance models for recognition,

reacquisition, and tracking,” 10th Int. Work. Perform. Eval. Track. Surveill. (PETS),2007

[165]  W. Li, R. Zhao, and X. Wang, “Human reidentification with transferred metric learning,”

2013, doi: 10.1007/978-3-642-37331-2\_3.

[166] S. Zhu, X. Gong, Z. Kuang, and J. Du, “Partial person re-identification with two-stream

network and reconstruction,” Neurocomputing, 2020, doi:

10.1016/j.neucom.2019.04.098.

[167] S. Karanam, M. Gou, Z. Wu, A. Rates-Borras, O. Camps, and R. J. Radke, “A Systematic Evaluation and Benchmark for Person Re-Identification: Features, Metrics, and Datasets,” *IEEE Trans. Pattern Anal. Mach. Intell.*, 2019, doi: 10.1109/TPAMI.2018.2807450.

[168] O. Camps *et al.*, “From the Lab to the Real World: Re-identification in an Airport Camera Network,” *IEEE Trans. Circuits Syst. Video Technol.*, 2017, doi: 10.1109/TCSVT.2016.2556538.

[169] H. Wang, S. Gong, and T. Xiang, “Unsupervised learning of generative topic saliency for person re-identification,” 2014, doi: 10.5244/c.28.48.

[170] S. Liao, Y. Hu, X. Zhu, and S. Z. Li, “Person re-identification by Local Maximal Occurrence representation and metric learning,” 2015, doi: 10.1109/CVPR.2015.7298832.

[171] W. Li, R. Zhao, T. Xiao, and X. Wang, “DeepReID: Deep filter pairing neural network for person re-identification,” 2014, doi: 10.1109/CVPR.2014.27.

[172] R. R. Varior, M. Haloi, and G. Wang, “Gated siamese convolutional neural network architecture for human re-identification,” 2016, doi: 10.1007/978-3-319-46484-8\_48.

[173] E. Kodirov, T. Xiang, and S. Gong, “Dictionary Learning with Iterative Laplacian Regularisation for Unsupervised Person Re-identification,” 2015, doi: 10.5244/c.29.44.

[174] M. Farenzena, L. Bazzani, A. Perina, V. Murino, and M. Cristani, “Person re-identification by symmetry-driven accumulation of local features,” 2010, doi: 10.1109/CVPR.2010.5539926.

[175] R. Zhao, W. Ouyang, and X. Wang, “Unsupervised salience learning for person re-identification,” 2013, doi: 10.1109/CVPR.2013.460.

[176] T. Isobe, D. Li, L. Tian, W. Chen, Y. Shan, and S. Wang, “Towards Discriminative Representation Learning for Unsupervised Person Re-identification,” pp. 8506–8516, 2022, doi: 10.1109/iccv48922.2021.00841.

[177] F. Liu and L. Zhang, “View confusion feature learning for person re-identification,” 2019,

doi: 10.1109/ICCV.2019.00674.

[178] K. Zeng, M. Ning, Y. Wang, and Y. Guo, “Energy clustering for unsupervised personre-

identification,” *Image Vis. Comput.*, vol. 98, pp. 1–8, 2020, doi. 10.1016/j.imavis.2020.103913.

[179] T. Wang, S. Gong, X. Zhu, and S. Wang, “Person re-identification by video ranking,” 2014, doi: 10.1007/978-3-319-10593-2\_45.

[180] M. Hirzer, C. Beleznai, P. M. Roth, and H. Bischof, “Person re-identification by descriptive and discriminative classification,” 2011, doi: 10.1007/978-3-642-21227-7\_9.

[181] L. Zheng *et al.*, “Mars: A video benchmark for large-scale person re-identification,” 2016, doi: 10.1007/978-3-319-46466-4\_52.

[182] C. C. Loy, T. Xiang, and S. Gong, “Multi-camera activity correlation analysis,” 2009, doi: 10.1109/CVPRW.2009.5206827.

[183] Z. Hu, C. Zhu, and G. He, “Hard-sample Guided Hybrid Contrast Learning for Unsupervised Person Re-Identification,” 2021, doi: 10.1109/IC-NIDC54101.2021.9660560.

[184] X. Liu and S. Zhang, “Domain Adaptive Person Re-Identification via Coupling Optimization,” 2020, doi: 10.1145/3394171.3413904.

[185] X. Zhang, J. Cao, C. Shen, and M. You, “Self-training with progressive augmentation for unsupervised cross-domain person re-identification,” 2019, doi: 10.1109/ICCV.2019.00831.

[186] S. Zhou, F. Wang, Z. Huang, and J. Wang, “Discriminative feature learning with consistent attention regularization for person re-identification,” 2019, doi: 10.1109/ICCV.2019.00813.

[187] K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick, “Momentum Contrast for Unsupervised Visual Representation Learning,” 2020, doi: 10.1109/CVPR42600.2020.00975.

[188] M. Ye, J. Shen, X. Zhang, P. C. Yuen, and S. F. Chang, “Augmentation Invariant and Instance Spreading Feature for Softmax Embedding,” *IEEE Trans. Pattern Anal. Mach. Intell.*, 2022, doi: 10.1109/TPAMI.2020.3013379.

[189] H. Chen, B. Lagadec, and F. Bremond, “ICE: Inter-instance Contrastive Encoding for Unsupervised Person Re-identification,” 2022, doi: 10.1109/iccv48922.2021.01469.

[190] Y. Ge, F. Zhu, D. Chen, R. Zhao, and H. Li, “Self-paced contrastive learning with hybrid memory for domain adaptive object re-ID,” 2020.

[191] Z. Liu, P. Luo, X. Wang, and X. Tang, “Deep learning face attributes in the wild,” 2015, doi: 10.1109/ICCV.2015.425.

[192] S. Karanam, Y. Li, and R. J. Radke, “Person re-identification with discriminatively trained viewpoint invariant dictionaries,” in ICCV, 2015, pp. 4516–4524.

[193] S. Bak, S. Zaidenberg, B. Boulay, and F. Bremond, “Improving person re-identification by viewpoint cues,” in AVSS, 2014, pp. 175–180.

[194 ] X. Li, W.-S. Zheng, X. Wang, T. Xiang, and S. Gong, “Multi-scale learning for low-resolution person re-identification,” in ICCV, 2015, pp. 3765–3773.

[195] Y. Wang, L. Wang, Y. You, X. Zou, V. Chen, S. Li, G. Huang, B. Hariharan, and K. Q. Weinberger, “Resource aware person reidentification across multiple resolutions,” in CVPR, 2018, pp. 8042–8051. [5] Y. Huang, Z.-J. Zha, X. Fu, and W. Zhang, “Illumination-invariant person re-identification,” in ACM MM, 2019, pp. 365–373.

[196] Y.-J. Cho and K.-J. Yoon, “Improving person re-identification via pose-aware multi-shot matching,” in CVPR, 2016, pp. 1354–1362.

[197] H. Zhao, M. Tian, S. Sun, and et al, “Spindle net: Person reidentification with human body region guided feature decomposition and fusion,” in CVPR, 2017, pp. 1077–1085.

[198] M. S. Sarfraz, A. Schumann, A. Eberle, and R. Stiefelhagen, “A pose-sensitive embedding for person re-identification with expanded cross neighborhood re-ranking,” in CVPR, 2018, pp. 420–429.

[199] H. Huang, D. Li, Z. Zhang, X. Chen, and K. Huang, “Adversarially occluded samples for person re-identification,” in CVPR, 2018, pp. 5098–5107.

[200] R. Hou, B. Ma, H. Chang, X. Gu, S. Shan, and X. Chen, “Vrstc: Occlusion-free video person re-identification,” in CVPR, 2019, pp. 7183–7192.

[201] A. Wu, W.-s. Zheng, H.-X. Yu, S. Gong, and J. Lai, “Rgb-infrared cross-modality person re-identification,” in ICCV, 2017, pp. 5380– 5389.

[202] C. Song, Y. Huang, W. Ouyang, and L. Wang, “Mask-guided contrastive attention model for person re-identification,” in CVPR, 2018, pp. 1179–1188.

[203] A. Das, R. Panda, and A. K. Roy-Chowdhury, “Continuous adaptation of multi-camera person identification models through sparse non-redundant representative selection,” CVIU, vol. 156, pp. 66–78, 2017.

[204] N. Martinel, A. Das, C. Micheloni, and A. K. Roy-Chowdhury, “Temporal model adaptation for person re-identification,” in ECCV, 2016, pp. 858–877.

[205] J. Garcia, N. Martinel, A. Gardel, I. Bravo, G. L. Foresti, and C. Micheloni, “Discriminant context information analysis for post-ranking person re-identification,” IEEE Transactions on Image Processing, vol. 26, no. 4, pp. 1650–1665, 2017.

[206] W.-S. Zheng, S. Gong, and T. Xiang, “Towards open-world person re-identification by one-shot group-based verification,” IEEE TPAMI, vol. 38, no. 3, pp. 591–606, 2015.

[207] A. Das, R. Panda, and A. Roy-Chowdhury, “Active image pair selection for continuous person re-identification,” in ICIP, 2015, pp. 4263–4267.

[208] J. Song, Y. Yang, Y.-Z. Song, T. Xiang, and T. M. Hospedales, “Generalizable person re-identification by domain-invariant mapping network,” in CVPR, 2019, pp. 719–728.

[209] A. Das, A. Chakraborty, and A. K. Roy-Chowdhury, “Consistent re-identification in a camera network,” in ECCV, 2014, pp. 330– 345.

[210] Q. Yang, A. Wu, and W. Zheng, “Person re-identification by contour sketch under moderate clothing change.” IEEE TPAMI, 2019.