IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 3, 2018, accepted September 14, 2018, date of publication September 28, 2018, date of current version October 25, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2872804

Bike-Person Re-Identification: A Benchmark and a Comprehensive Evaluation

YUAN YUAN[1,2], (Senior Member, IEEE), JIAN'AN ZHANG[1,2], AND QI WANG[1,2,3], (Senior Member, IEEE)

[1]School of Computer Science, Northwestern Polytechnical University, Xi'an 710072, China
[2]Center for Optical Imagery Analysis and Learning, Northwestern Polytechnical University, Xi'an 710072, China
[3]Unmanned System Research Institute, Northwestern Polytechnical University, Xi'an 710072, China

Corresponding author: Qi Wang (crabwq@gmail.com)

This work was supported in part by the National Key Research and Development Program of China under Grant 2017YFB1002202, in part by the State Key Program of National Natural Science Foundation of China under Grant 61632018, in part by the National Natural Science Foundation of China under Grant 61773316, in part by the Natural Science Foundation of Shaanxi Province under Grant 2018KJXX-024, in part by the Fundamental Research Funds for the Central Universities under Grant 3102017AX010, and the in part by the Open Research Fund of Key Laboratory of Spectral Imaging Technology, Chinese Academy of Sciences.

ABSTRACT Existing person re-identification (re-id) datasets only consist of pedestrian images, which are far more behind what the real surveillance system holds. As investigated a real camera in a whole daytime, we find that there are more than 40% persons are riding bikes rather than walking. However, such kind of person re-id (we named bike-person re-id) has not been focused on yet. In this paper, we pay attention to the bike person re-id for the first time and proposed a large new bike-person re-id dataset named BPReid to address such a novel and practical problem. BPReid distinguishes existing person re-id datasets in three aspects. First, it is the first bike-person re-id dataset with largest identities by far. Second, it samples from a subset of real surveillance system which makes it a realistic benchmark. Third, there is a long instance between two cameras which makes it a wide area benchmark. Besides, we also proposed a new pipeline designed for bike person re-id by automatically partitioning a bike person image in two parts (bike and person) for feature extraction. Experiments on the proposed BPReid dataset show the effectiveness of the proposed pipeline. Finally, we also provide a comprehensive evaluation of existing re-id algorithms on this dataset, including feature representation methods as well as metric learning methods.

INDEX TERMS Bike-person re-identification, re-identification, re-identification dataset, splitting method.

I. INTRODUCTION

Person re-identification (re-id) refers to matching the same person across different cameras, which has been made a great progress in the past ten years with the efforts of [1] and [2]. It can be used in various fields in real applications.

A basic application of person re-id is the surveillance. At present, all existing person re-id datasets focus on pedestrians in the camera views. However, we note that in real surveillance scenario especially in campus environment it is not entirely the case that a person walks in the camera view. Apart from pedestrians, it is also a common case that persons ride bikes such as bicycles, motors or eletric-power bikes. To evaluate how the persons riding bikes account for all the persons appearing in camera views, we investigated a real surveillance camera in a whole daytime(from 6:00 am to 7:00 pm) to count the entrance and exit pedestrians and person on bikes (including motor or eletric-power bikes) every 6 minutes 40s (every 10000 frames). Figure 1 shows the statistic of the total persons and persons on bike from morning to the evening. It can be found that there is an average 41 percentage of persons on bikes in all persons and it can reach a maximum of 69 percentage. From this figure, we can conclude that persons on bikes account for a large amount of all persons in camera views and for certain time it will take the majority of all the persons.

The above conclusion motivates us to reconsider the person re-id problem which is basic in surveillance. For an adaption to the real scenarios, it is necessary to take the persons on bikes into consideration and evaluate such a kind of person re-id comprehensively. Next, we will refer to the new type of person re-id as **bike-person re-id**, and the common person re-id as **classical person re-id** for convenience. Ar present, bike person re-id receives little concerns and there are also no common benchmarks as well as methods. This motivates us to begin this work.

TABLE 1. Comparing with existing datasets.

Dataset	Year	#person	#distractors	#BBox	#cameras	Label Method	Image size	Multi-shot
VIPeR [3]	2007	632	0	1264	2	hand	128 × 48	No
GRID [8]	2009	1025	775	1275	8	hand	Varied	No
PRID2011 [5]	2011	934	732	24,541	2	hand	128 × 64	√
CUHK01 [4]	2012	971	0	3,884	2	hand	160 × 60	√
CUHk03 [6]	2014	1467	0	13,164	10	hand/DPM	Varied	√
iLIDS-VID [11]	2014	300	0	42,495	2	hand	Varied	√
Market1501 [7]	2015	1501	2798	32,217	6	hand/DPM	128 × 64	√
PRW [9]	2016	932	0	34,304	6	hand	Varied	√
LSPS [10]	2016	11934	0	34,574	-	hand	varied	no
MARS [13]	2016	1261	0	1,119,003	6	DPM+GMMCP	256 × 128	√
DukeMTMC4ReID [12]	2017	1852	439	46,261	8	Doppia	varied	√
BPReid	**2018**	**4,579**	**10,910**	**200,680**	**6**	**hand**	**varied**	√

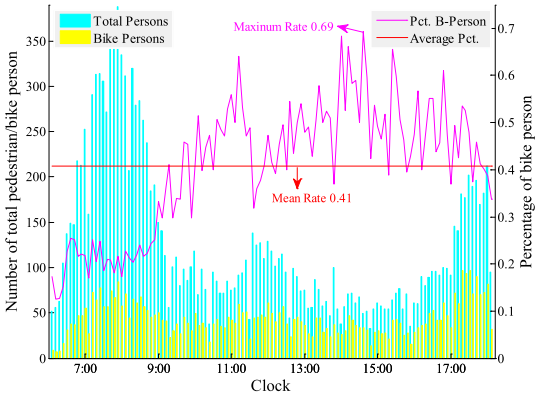


FIGURE 1. A statistic of the persons on bike in/out the doors in the whole daytime. The yellow bars stand for the amount of person on bikes and the cyan bars represent the amount of total pedestrians. A magenta curve show the rate of person on bike to the total persons.

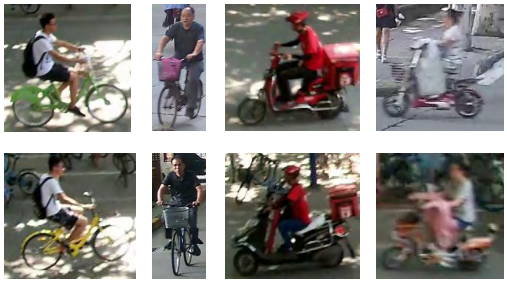


FIGURE 2. Bike persons contain more information than walking person and bikes can assist person re-id.

Apart from the large amount of bike persons, we also find that they contain more useful information than pedestrians, which can assist person re-id. Figure 2 shows these examples. Persons in column 1 and column 2 are riding bicycles. In column 3 and column 4, person are riding motor bikes and eletric-power bikes respectively. As shown, the persons in column 1 wore similar white t-shirts and black short pants with a similar black backpack. It is difficult to identify the person from appearance features while it can easily identify them with the bike color. Similar cases are also in column 2 to column 4 and we can identify the separate person from bicycle basket, motor color and windshield of electric vehicle respectively.

Considering the large amount of bike-persons and the useful information they hold, as the main contribution of this paper, we construct a large new bike-person re-id dataset named "BPReid" for the research of bike person re-id. To our knowledge, this is the earliest work that focuses on such a task. The BPReid datataset contains 4579 matched identities and 10910 unmatched identities collected from 6 cameras, which makes it the large person re-id dataset as far as the amount of identities. All the bounding boxes are hand-drawn from a real monitor video surveillance system. It distinguished from existing datasets in the following aspects: bike-persons, a real surveillance system and long instance between two cameras.

In addition to the proposed dataset, we also proposed a new pipeline for bike person re-id. Based on the investigation of bike person images, we note that a bike person image usually has an explicit structure. Namely, it can be divided into a person part and a bike part. Besides, there often exists noisy regions in a bike person image. All the observations motivate us to develop a new auto-splitting methods for bike person images that can get rid of noisy regions of a bike person image as well as can divide a bike person image into the two parts as mentioned above. The new pipeline first divides a bounding box into two parts based on the auto-splitting method and then extracts features for each divided part. Experiments show the effectiveness of the proposed pipeline.

II. RELATED WORKS

In this section, we will briefly introduce some existing datasets used for person re-id as well as some existing methods for person re-id.

A. DATASETS

There are many existing datasets for person re-id. An overview of the statistic of these datasets is shown in Table 1. Amoung all the datasets, VIPeR [3] is the most popular dataset for it is the first person re-id dataset, but the dataset is small and there is only one image for every identity in each of the two camera views. CUHK01 [4] is proposed with more identities than VIPeR but still with two

camera views. Later, PRID2011 [5] and CUHK03 [6] are constructed with multi-camera views and multi-images for one identity. Recently, with the need of training for deep neural networks, large datasets with multi-camera views and multi-bounding box are proposed one after another, such as Market1501 [7], PRW [9], LSPS [10] and DukeMTMC4ReID [12]. However all these large scale datasets focus only on pedestrians. Different from existing datasets, the proposed BPReid dataset focuses on the bike person for the first time. And it is also a large scale person re-id datset in real surveillance environment with multi-camera views and multi-images for each identity that can be used for training deep neural networks.

B. METHODS

Classical person re-id methods can be classified into two groups: feature-based methods and metric learning methods. Feature-based approaches focus on design discriminative and robust features for pedestrians such as ELF18 [14], [15], SDALF [19], fisher vectors [17], SCNCD [20], Color Names [18], LOMO [32], GOG [16]. Among these features, ELF18 is the most used feature and LOMO and GOG are regarded as the state-of-the art methods. Metric learning methods aim at designing discriminative metrics for matching and most of the methods used for person re-id are on the basis of Mahalanobis distance [21], [22]. A large majority of methods fall into this group. Classical methods include LMNN [23], ITML [24], LDML [25], KISSME [28], PCCA [27], FDA [29], LFDA [30], kLFDA [31], LADF [33], XQDA [32], MLAPG [35] and NFST [34], where XQDA, MLAPG and NFST are regarded as the state-of-the-art.

Recently, deep learning methods have been successfully applied in many computer vision tasks [36] and a lot of deep learning methods are developed and make a great succus in person re-id. Existing deep learning methods for person re-id designed either from the ranking aspects or the classification aspects. The ranking related methods take three images as input where two of them are of the identity and another is from a different identity. A triple loss or improve triple loss is often adopted, such as DRDC [37], ImprTri [38], Quadruple [39], [40]. Classification related works usually adopt classification losses with a carefully designed special structure for person re-id, such as binary classification related methods FPNN [6], [41]–[43]. Multi-classes related methods are also developed such as [44]–[46]. Although the large amount of deep learning methods, there are no neural networks that designed for bike person re-id. It can be developed in the future works. In this work, we will not conduct experiments on deep neural network as we mainly focus on the proposed bike person re-id dataset and give it a benchmark evaluation.

Apart from the above methods, there are also some other kinds of methods, such as [33] and [47] focus on similarity learning, [48]–[50] focus on ranking. We refer readers to [51] and [52] for a comprehensive review.

This paper aims at providing a bike-person re-identification dataset and a basic comprehensive evaluation. Different from general person re-identification methods designed for general pedestrian image, this paper takes the structure of bike-person image into consideration and develop a splitting method used to improve the performance of traditional re-id pipeline. The proposed methods is designed for bike-person re-id and can be easily extended to deep learning methods which we will develop in the future.

III. BIKE-PERSON RE-IDENTIFICATION DATASET

A. DESCRIPTION

To evaluate the bike-person re-id, we first build a new large dataset named BPReid (short for bike person re-id). All frames in the BPReid dataset were captured by 6 static cameras on the Northernwestern Polytechnical University campus at 25 frames per second, which consists of the subset of the real surveillance system. The layout of the 6 cameras is shown in Figure 3 (a).

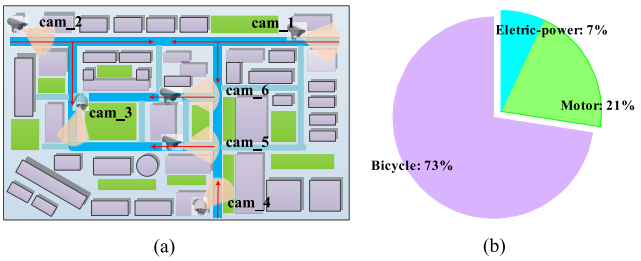


FIGURE 3. (a) Layout of the cameras in BikePerson dataset. (b) The proportions of three kinds of bike persons.

All the 6 cameras are high definition (HD), among which three cameras are 1920 × 1080,two cameras are 2048 × 1536 and one camera is 1280 × 720. We labeled all the 6 cameras in continuous daytime with average 13 hours per camera. In total, this dataset contains 4579 identities captured by 2 cameras and 10910 identities (distractors) captured only by one camera making it the largest person re-id dataset as far as the captured identities. Table 2 shows the detailed information of the BPReid dataset including the detailed configuration of every cameras,labeling time duration and the number of distractors. It also displays the information of every sub-datasets. We can get total 18 sub-datasets and among of these sub-datasets cam45-bike and cam56-bike are two large scale sub-datasets; cam45-motor, cam56-motor, cam35-bike and cam61-bike are middle scale sub-datasets and all the rest are small scale sub-datasets. Besides, we give a statistic comparison with existing datasets in Table 1. It is noted that there is an underlying assumption that persons will be not off the vehicles. That is meant such a case will not be considered in this manuscript. In fact, it is reasonable to make such an assumption because a person often rides until the destination is reached.

Besides the identities we labeled, we also annotated the type of bikes one person rides. We classify the bikes into

TABLE 2. The detailed statistics of the proposed BPReid dataset.

single camera	cam1	cam2	cam3	cam4	cam5	cam6	Total
config.	1920 × 1080	1920 × 1080	1920 × 1080	2048 × 1536	2048 × 1536	1280 × 720	-
#duration	12h 54min	9h 19min	11h 48min	12h 6min	12h 4min	12h 55min	71h 6min
#distractors	4756	2653	561	721	1420	799	10910
camera pairs	cam12	cam23	cma35	cam45	cam56	cam61	Total
#bikes	56	108	215	1309	1428	206	3322
#motors	133	28	34	410	288	47	940
#eletrics	30	5	23	129	115	15	317
#BBox	4380	2820	5440	36934	36586	5356	91586

FIGURE 4. Challenges of the proposed dataset, (a): illumination changes, (b):occlusions, (c):persons of high similarity, (d):different viewpoints, (e):complex background. Each column is corresponding to the same person except group (c) and each row are sampled from the same camera view. In group (c),each column are two persons of high similarity sampled from the same camera view. The red box in the third column of group (b) shows that a bike often carries more than one person and the blue box in the third column of group (d) shows that the scale and the appearance changes.

three types: bicycle, motorbike and eletric-power bike. The first,third,and fourth column of Figure 2 shows the three kinds of bikes respectively. Eletric-power bikes are different from motors that they are smaller than motors and there is often a basket in the front of an eletric-power bike. And they differ from bicycles that they are eletric-power and they are often lower than bicycles. Figure 3 (b) shows the proportion of every kind of bike person. As we can expect, the bicycle person account for 73% to the whole bike person, followed by motor 21% and eletric-power motor 7%.

From the tables and figures, it can be found that three aspects extinguish the proposed BPReid dataset from existing datasets. First, **it is the first bike-person re-id dataset**. To our knowledge, the BPReid dataset is the first dataset evaluated for the novel and practical bike-person re-id problem which has not been focused in previous works. It has 4579 matched identities and 10910 unmatched identities (distractors), which makes it the largest person re-id dataset as for as the number of identities. Apart from the identities we labeled, we also annotate the type of one person rides. This dataset extends the field of traditional person re-id with one realistic new problem.

Second, **the proposed dataset is sampled from real surveillance monitor system with long time annotation**. All the cameras consist of the BPReid dataset are sampled from a real surveillance monitor network, which makes the dataset a real scenario for person re-id. Existing datasets, such as Market1501 and CUHK01, are all sampled from temporal cameras which are not real enough for person re-id. Besides, most of the datasets are sampled in a less time which makes the existing datasets changes a little both in illumination and background. The proposed BPReid dataset is fully annotated from the morning to the evening and it is the most hours hand annotated person re-id dataset up to date with various illumination changes and complex background.

Third, **there is a long distance between every two cameras**. From Figure 3 we can find that the cameras are distributed in the whole campus. Among which, the distance between camera 1 and cam2 is more than 900 meters. Even, the nearest camera pair: cam4 and cam5, the distance between them is also more than 120 meters. The long distance between cameras makes the bike-person a more difficult problem because of the serious illumination changes and complex backgrounds. And it also makes the dataset a more realistic benchmark to evaluate.

B. CAMPARED WITH CLASSICAL PERSON RE-IDENTIFICATION

From what we have introduced in previous sections, it can be found that bike-person re-id is one of the variants of person re-id and it differs from classical person re-id in the aspect that the bike-person re-id focuses on persons on bikes. Hence, it is necessary to discuss the similarities and differences between the BPReid dataset and classical person re-id datasets.

Because the essence of bike-person re-id is an invariant of person re-id, as a result, it has the same challenges as person re-id, such as (1) it suffers from illumination changes. (2) it suffers from occlusions. (3) different persons are often of high similarity. (4) different viewpoints add the it's difficulty. (5)there often exists complex background. Figure 4 shows all these cases. From Figure 4 we can find bike-person re-id is as challenge as the classical person re-id and next we will see that bike-person is often more challenge than classical person re-id.

Apart from the common challenges in person re-id, bike-person re-id also introduces some new challenges.

Bike-person re-id is different from classical person re-id in the following three aspects.

Firstly, the scale of bounding boxes of bike-persons varies much than traditional person re-id. In classical person re-id, the height of the bounding boxes is often larger than the width because no matter how the viewpoint changes a person cannot change himself/herself. However, bike-person re-id is different because a person can change largely under different viewpoints. In Figure 4, the third column of group (d)(shown in blue box) shows an example of such a case, and it can be easily illustrated that the scale of the same person can change largely.

Secondly, the appearance varies more sharply. This is because that a bike can have more appearance changes than a person walks. A bike can have quite different appearance under different viewpoints as well as itself can changes even under the save viewpoint, such as a bike can suddenly change its directions and the appearance of bike can vary sharply accordingly. We also take the person in the third column of group (d) as example, it can be also find that the bike-person's appearance change more sharply than he walks.

Thirdly, there are not only one person in one image. In classical person re-id, a bounding box always contains one one person but in bike-person re-d it is not true. This is because that a bike often carries two or more person rather than only one. In Figure 4, the third column of group (b)(shown in red box) shows an example of such a case where there are three person in one motor bike.

From the above three aspects, we can find that the bike person re-id is more challenge than classical person re-id.

IV. METHODOLOGY

In this section, we will show a new pipeline for bike-person re-id on the basis of investigating the structure of bike person images. Note that this new pipeline is based on traditional person re-id methods by splitting a bike-person image into two parts. Because this paper aims at providing a bike-person re-identification dataset and a comprehensive benchmark evaluation, the proposed method is designed to improve the performance of traditional methods. However, it can be also extended to deep learning methods by taking the structure of bike-person image into consideration as following.

A. SPLITTING METHOD

A bike person image has an explicit structure, namely a person part and a bike part, which can be easily found. Such a strong priori knowledge can be used to develop efficient methods for re-identification. Besides, we also note that a bike person image usually has more noisy information than traditional person re-id images, which can be reduced for better feature extraction. A typical example of bike person image is shown in Fig 5 (a). The regions masked by blue rectangles are noisy regions that cannot assist bike person re-id. The region masked by the yellow rectangle is the bike part while the red rectangle masked part is the person part.

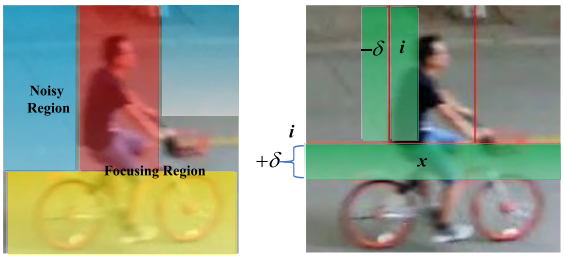


FIGURE 5. Illustration of the structure of bike person image and the method derived from the prior. (a)The motivation of the splitting method. (b) A graphic explanation of notations in Eq. 1.

We named the two regions of the focusing regions as they contain a majority of information we need.

Based on the above observation, it would be natural to come up with that we can get rid of the noisy region of a bike person image for the refine of effective information. And then we can split the focusing region into a person part and bike part for the uses of structure information. Follow this natural idea, we proposed an auto-splitting method for this purpose.

1) AUTO-SPLITTING METHOD

The purpose of a splitting method is to get rid of the noisy region and divide the focusing region into informative parts. Specifically, for bike person images, it is to find a horizontal line that split the lower and upper part and at the same time split the upper part into three regions as showed in Figure 5 (b). We note that a good splitting line should hold that the regions symmetrical on the line would be dissimilar. This property motivates us to propose an anto-splitting method, which can be expressed in the following equation.

$$l_i = \underset{i}{argmax} \sum_{c} \|f(\mathbf{x}_i^{+\delta}) - f(\mathbf{x}_i^{-\delta})\|_2^2, \tag{1}$$

where $\mathbf{x}_i^{+\delta}$ stands for a sub-region of a give image located in the right side of i-th line with the width (or height) δ, while $\mathbf{x}_i^{-\delta}$ is the region on the left side of lien i. $f()$ is the function of feature extraction and we used all the pixels in the region in this paper, And c represents the channels of a given image and l_i is the splitting line of image. The equation shows that the wanted splitting line l_i is the line that maximizes the distance over all channels. Figure 5 (b) shows a graphic explanation of the main notations. The green rectangle masked regions are the regions input x for $f()$ and δ is the region width/height. The wanted output l_i is the line that maximizes the distance of the symmetry regions besides it.

It is noted that a similar equation named chromatic bilateral operator in SDALF [19] has been proposed. The proposed splitting method differs from the chromatic bilateral operator in the following three aspects. First, the chromatic bilateral operator tries to find the line that minimizes the distance between the regions symmetrical on it, which is opposite to the proposed one. Second, the purpose of the two methods is different which the chromatic bilateral operator tries to

find the symmetrical lines. Third, we used a general feature expression rather than the chromatic bilateral operator.

2) HAND-CRAFT SPLITTING METHOD

Apart from the proposed auto-splitting method, we also design a hand-craft splitting method based on priori knowledge. We empirically set the horizontal line the half height of an image and set the first and second vertical line at the $1/5$ and $4/5$ width of the image respectively.

B. PIPELINE FOR BIKE-PERSON RE-ID

Take the proposed splitting method into consideration, we proposed the pipeline for bike person re-id based on traditional feature extraction and metric learning method, which is shown in Figure 6. As for the deep neutral network based method, it can be derived from the pipeline by connecting feature extraction and metric learning step. It consists of three steps: auto-splitting and divided into two parts, feature extraction for each part and metric learning, where the auto-splitting is designed for bike person re-id.

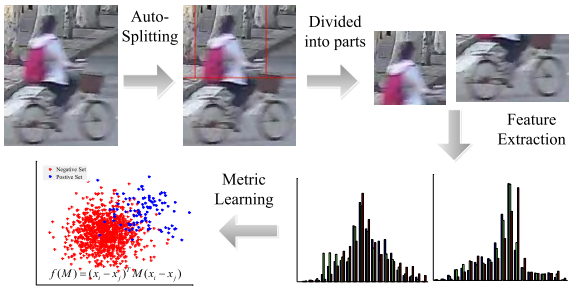


FIGURE 6. The pipeline for the bike person re-id based on traditional feature extraction and metric learning method.

V. EXPERIMENTS

In this section, we will evaluate the proposed pipeline based on the splitting method above. A comprehensive evaluation including three widely used feature representation methods and six metric learning methods is given. Note that, this paper aims at providing a new large bike-person re-id dataset and basic evaluations, so some deep learning based method are not evaluated.

A. EXPERIMENTAL DETAILS

The BPReid dataset consists of 18 sub-datasets and we will evaluate the bike person re-id on all these datasets. For each dataset, the bike persons of the even label are selected for training and the odd label identities are used for testing. For testing set, we randomly select half of the ids as the probe set and the rest consists of the gallery set. For the baseline methods, we resize each image in the BPReid dataset to 128×48 pixels.

We perform multi-shot vs multi-shot experiments one all the sub-datasets to simulate the real scenarios. For multi-probe identities, we max pooling and average pooling is applied respectively. Finally, we report the average performance across all the probe identities as far as CMC rank 1 identification rate and mAP [7].

B. EVALUATION THE PROPOSED PIPELINE

In this section, we will evaluate the proposed pipeline as showed in section IV-B. As can be seen from the pipeline, the splitting methods can influence the performance of bike person re-id. So we first evaluate the auto-splitting method and the hand-craft splitting method. After splitting lines are determined, we can note that the resizing scale of the separate two part is crucial and hence we will test the performance of bike person re-id under different combination of resizing scale. As for feature extraction, it is necessary to evaluate different combinations of features extracted in the two parts. We will conduct various experiments with regards to the above three aspects in the following. It is noted that we test the proposed splitting method and the new pipeline all on a sub-dataset of BPReid, namely the cam12-bike. It is a small size sub-dataset that has 55 bike persons and 1160 bounding boxes, which make it efficient to test the effectiveness of the proposed pipeline.

1) HAND-CRAFT SPLITTING VS AUTO-SPLITTING

We adopted two kinds of methods to split a given image, namely hand-craft splitting method and auto-splitting method as described in section IV. Table 3 shows the performance bike person re-id with the change of these two kind of splitting methods. As expected, the auto-splitting method outperforms the hand-craft splitting methods, which can be easily explained that the hand craft splitting method cannot always remove the noisy regions. This result shows the effective of the proposed the splitting method.

TABLE 3. The bike person re-id performance with the change of splitting methods.

	r1	r5	r10	r20	mAP
Hand-craft Spl.	28.57	50.00	64.29	76.79	23.81
Auto-Spl.	39.29	64.29	71.43	82.14	**30.63**

2) AUTO-RESIZING VS FIXED RESIZING

The images in BPReid vary sharply from $70 * 50$ pixels to $300 * 350$ pixels, which makes the problem of resizing more important than classical person re-id. We will test two kinds of resizing methods. The first one is auto resizing method based on the distribution of the sizes of the two image parts. It consists of three steps. First, we draw the width-height figure of the two parts of all images and each part is represent a point in the width-height figure. Next, we generate a rectangle region that contains more than 90% of the points. And last we sample five points for every rectangle region. It can show in Figure 7 that triangle cyan points corresponds to lower parts and dot red point corresponds to upper parts. The five points are %25, %50 and %75 of the rectangle size as show in the upper right corner of Figure 7. Note that the

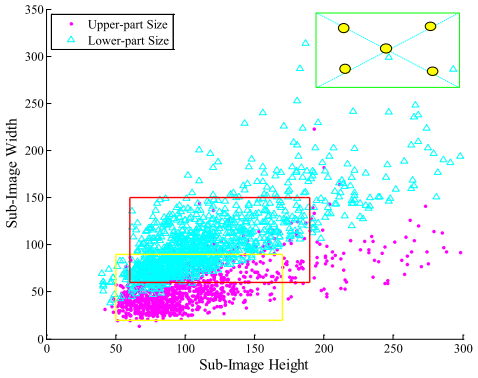


FIGURE 7. Auto-resizing Illustration. Triangle cyan points show the distributions of lower parts' size and dot red points show the upper parts' size. Red rectangle and yellow rectangle are generate in the first step. Green rectangle and five yellow circles shows the generate points.

width and height of the upper and lower parts are generate from the proposed auto-split method.

TABLE 4. The rank 1 identification rate with different size of upper an lower parts. Test with the feature of LOMO + GOG.

Lower \ Upper	110×54	80×38	80 ×72	140×38	140×72
124 × 104	21.24	22.98	21.50	21.39	19.97
92 × 82	21.27	23.04	21.48	21.40	19.96
92 × 128	21.21	**23.06**	21.46	21.39	20.16
158 × 82	21.23	22.90	6.45	21.49	19.95
158 × 128	21.20	22.95	21.49	21.38	19.95

Table 4 shows performance of bike-person re-id with the change of different part sizes. It can be easily found that every column shares similar performance, namely, the performance varies little with the change of lower part of bike person images. This indicates that the size of lower part has little impact on the retrieval accuracy while the size of upper part has much impact on the retrieval accuracy. We argue that this is because the upper part corresponds to torso and head which is more sensitive to resizing. For instance, it is not reasonable the width of upper part is larger than the height because of the prior structure information. While the lower part corresponds to bike which is not sensitive to resizing. This is, resizing operation does little destruction to the structure information of bike.

The other kinds of resizing are designed by ourself taking the existing common resizing scale into consideration. In previous works,a pedestrian image is often resized to 128×64, such as LOMO [32] and GOG [16] methods. Based on this usual rule, we resize the upper part to $64 * 48$ and the lower part to $48 * 128$ for that an upper part is usually the half height of a given image. In the last two rows of Table 5, we show the performance of the fixed size. It can be found such a method performs pool than the auto-resizing methods. This is because the fixed size resizing method does not take the distribution of images sizes into consideration as the auto-resizing methods do.

TABLE 5. The performance of different combination of features with rank 1, 5, 10, 20 and mAP reported. Upper part + Lower part feature.

	r1	r5	r10	f20	mAP
ELF+ELF	17.89	42.86	53.57	60.71	19.84
ELF+LOMO	14.29	35.71	46.43	60.71	17.32
ELF+GOG	39.29	64.29	71.43	82.14	**30.63**
LOMO+ELF	17.86	25.00	35.71	64.29	12.96
LOMO+LOMO	39.29	57.14	64.29	85.71	29.56
LOMO+GOG	28.57	50.00	71.43	85.71	23.13
GOG+ELF	**46.43**	64.29	85.71	89.29	28.03
GOG+LOMO	35.71	60.71	75.00	89.29	24.79
GOG+GOG	32.14	67.86	85.71	89.29	29.94
ELF+GOG+Fixed	34.86	46.43	53.57	57.14	25.19
GOG+ELF+Fixed	28.57	50.00	60.71	75.00	24.96

Different Combination Of Features: The feature of each divided part is important for bike person re-id. We test three kinds of features ELF, LOMO and GOG for each part and we can get 9 combinations in total. Table 5 shows the performance of the 9 combinations on cam12-bike sub-dataset.In this table, we reported both mAP and CMC rank values. As shown, the combination of ELF + GOG and GOG + ELF achieves the best performance as far as mAP and rank 1 identification rate on this dataset. The performance is higher than ELF+ELF and GOG+GOG combinations, which shows the effectiveness of the proposed pipeline. Next, we will adopt a different combination for different sub-datasets for adaption to different cases in a similar way. We argue that ELF feature can capture the low-level feature of an image for it combines color and texture features and GOG feature can capture high level structure information of a image because it considers the distribution of pixel. The combination of ELF and GOG can capture both low-level and high-level feature, which makes the combination achieve a better performance. All the combination are named in the form of the upper part feature plus the lower part feature.

C. COMPARISON OF EXISTING ALGORITHMS A COMPREHENSIVE EVALUATION

In this section, we will evaluate existing algorithms in the proposed BPReid dataset including both feature-based methods and metric learning methods. For feature based methods, the most widely used features ELF [3], LOMO [32] and GOG [16] are adopted. As for metric learning methods, classical FDA [29], KISSME [28], LFDA [30], MFA [31], XQDA [32], MLANG [35] and NFST [34] methods are selected. We evaluated all the combination of the three features and the 6 metric learning methods on the 18 sub-datasets. The l_2-norm metric is used as a baseline. Parameters of the existing methods are set according to the suggestions given in the corresponding works. We used the released codes for all the feature extraction. For metric learning methods, we re-implement the FDA, KISMME, LFDA and LFST according what the corresponding works' setting while used the provided codes for XQDA and MLAPG. The new pipeline method uses auto-splitting Different feature combinations

TABLE 6. Comparison of the existing metric learning methods as well as feature-based methods on the BPReid dataset with mAP reported. In the table, B stands for persons on bicycles in BPReid, M is motor for short and E represent Eltric-power bike. And the best performance in every feature group is bold. Best performance for each sub-dataset is shown in blue.

mAP / Method \ Dataset		cam1-2			cam2-3			cam3-5			cam4-5			cam5-6			cam6-1		
		B	E	M	B	E	M	B	E	M	B	E	M	B	E	M	B	E	M
ELF	l_2-norm	19.6	30.0	16.5	12.5	97.3	38.3	5.74	28.2	21.6	1.2	11.0	4.51	1.32	14.3	5.08	8.58	47.7	15.9
	FDA	12.4	15.1	8.0	6.89	67.7	19.4	3.36	16.5	10.1	0.70	5.21	1.68	0.90	5.67	1.73	3.51	22.8	10.5
	KISSME	26.3	41.1	30.9	26.7	57.7	38.4	10.8	30.2	26.9	2.08	12.6	5.96	1.92	12.5	5.94	15.6	28.2	20.2
	LFDA	26.8	35.4	36.7	31.6	88.9	42.7	27.7	38.5	30.2	8.44	**18.8**	14.7	8.64	18.2	14.2	26.2	46.8	**21.5**
	MLAPG	29.8	**47.3**	33.5	**33.5**	**97.1**	40.8	20.3	35.4	31.6	4.64	14.9	8.86	4.89	**21.0**	9.12	22.1	**52.0**	21.2
	XQDA	**32.3**	43.0	**37.9**	31.0	46.9	**47.5**	**29.7**	**52.1**	**35.8**	**11.2**	18.7	**14.9**	**10.7**	17.9	**14.3**	**28.9**	36.8	21.4
	NFST	28.4	42.2	8.64	10.5	46.9	46.4	3.24	51.6	34.7	0.68	6.12	1.87	0.63	8.10	2.07	3.98	36.8	20.6
LOMO	l_2-norm	7.38	11.9	3.95	3.70	57.8	14.6	2.59	13.1	13.6	0.43	3.38	0.94	0.33	3.21	0.87	2.21	23.4	7.79
	FDA	5.59	10.5	2.97	2.93	60.1	10.5	1.58	13.6	8.57	0.31	2.63	0.80	0.29	2.82	0.77	1.78	22.5	6.38
	KISSME	6.43	11.5	2.63	3.51	58.8	12.3	2.65	15.0	11.7	0.32	3.32	1.07	0.33	3.39	1.10	2.01	19.8	7.94
	LFDA	7.24	**12.9**	3.02	3.73	61.0	12.0	2.39	12.1	12.1	0.52	3.29	1.12	0.35	3.23	1.06	1.96	**24.1**	7.43
	MLAPG	6.94	10.9	3.32	4.00	**61.5**	12.8	3.09	**15.3**	13.2	**0.65**	**3.87**	**1.53**	0.44	3.45	**1.38**	2.19	18.9	7.26
	XQDA	**7.83**	12.1	**4.19**	**4.53**	58.9	**16.0**	**4.20**	15.1	**19.4**	0.40	3.70	1.21	0.36	**4.30**	1.25	**2.53**	23.3	**8.26**
	NFST	6.24	10.822.52		3.12	60.0	12.4	1.61	14.8	11.5	0.28	2.88	0.93	**0.63**	2.96	0.58	1.73	19.6	6.92
GOG	l_2-norm	25.8	33.4	30.6	31.4	92.2	72.1	38.8	**61.1**	70.4	10.8	39.3	21.0	6.40	30.5	18.4	31.3	49.0	33.9
	FDA	5.85	10.5	3.18	3.40	53.9	13.9	9.62	17.8	13.3	0.90	3.15	7.37	0.82	3.51	8.21	5.79	23.8	6.65
	KISSME	**51.8**	43.1	**66.2**	**65.3**	59.4	51.7	**73.2**	57.9	**71.3**	30.2	**57.4**	**45.6**	24.6	**55.0**	**45.2**	**51.8**	29.8	35.3
	LFDA	37.9	44.4	52.6	61.3	81.7	63.7	60.1	55.3	67.3	**35.6**	49.0	40.5	**30.3**	50.6	44.7	48.3	43.4	32.5
	MLAPG	32.5	39.3	40.3	44.9	**93.0**	**72.8**	50.5	56.9	70.4	25.2	39.6	27.2	22.0	41.7	28.6	40.9	**47.8**	**35.8**
	XQDA	43.8	**47.3**	52.8	61.8	85.8	62.3	59.8	58.1	62.4	27.9	50.0	35.3	22.9	52.3	37.1	45.0	44.4	32.7
	NFST	32.9	42.1	42.4	47.8	81.7	56.0	46.5	56.5	57.5	0.68	16.2	1.87	0.63	26.4	1.66	3.51	38.2	27.3
New Pipeline	l_2-norm	31.1	40.2	38.1	31.6	93.6	67.1	28.5	60.2	69.8	12.1	38.2	25.5	8.55	24.0	19.8	35.3	36.0	25.7
	FDA	5.54	11.4	11.6	3.30	68.5	13.5	1.82	14.2	14.3	1.70	3.34	7.09	3.76	7.89	5.48	5.39	24.9	8.77
	KISSME	43.3	32.5	41.6	60.0	58.2	34.0	53.5	52.2	62.7	31.2	52.2	47.1	25.1	48.7	32.8	43.2	35.4	29.4
	LFDA	46.0	43.1	47.5	64.0	86.6	62.9	59.3	54.4	68.9	**35.9**	54.7	45.7	30.8	41.8	25.8	44.0	41.5	29.0
	MLAPG	50.4	47.0	**52.1**	65.7	**95.8**	68.6	**65.3**	58.7	72.5	26.6	55.6	47.9	**46.7**	45.9	-	48.2	48.3	30.6
	XQDA	51.7	44.2	44.6	56.0	95.6	65.9	44.5	56.6	71.2	27.9	47.5	40.9	30.2	40.0	22.0	30.4	**49.4**	25.9
	NFST	**54.1**	**49.5**	50.8	**67.2**	95.1	**73.8**	53.3	**61.3**	**73.7**	0.61	**59.9**	**52.5**	0.40	**58.5**	**29.5**	**53.3**	47.5	**37.0**

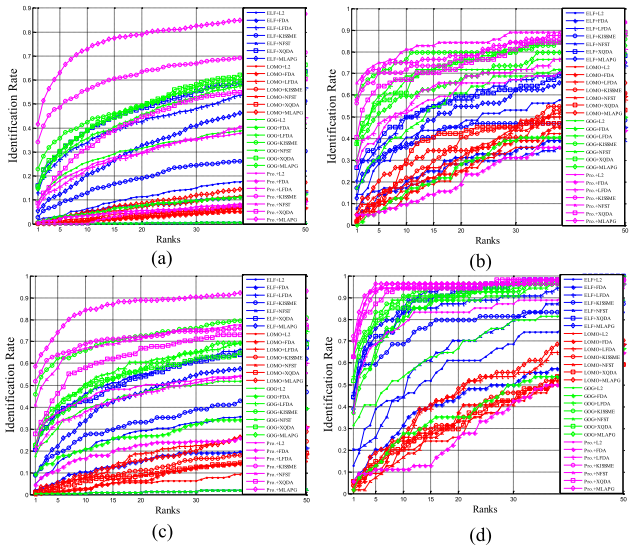


FIGURE 8. CMC curves of four sub-datasets. (a) cam45-bike. (b) cam45-motor. (c) cam45-eletric. (d) cam23-bike. Magenta curves stand for the pipeline methods.

are used for testing and we select the best ones for evaluation,such as for cam12-bike,cam12-motor, cam23-motor and cam23-eletric we apply ELF+GOG, in other sub-datasets, we apply GOG+GOG.

The performance of these methods are shown in Table 6. From this table, we can find the new pipeline method achieves the best performance in most of the sub-datasets. It shows improved performances for most of the methods and for some method such as NFST it improved more than 20 mAP. But it is also noted that in some dataset, such as cam12-motor, GOG+KISSME achieves the best performance. We argue this is explained by the fact that most of the images in this dataset are shown frontal view and have little noisy information. So an auto-spitting method may remove some useful information which reduces the performance. We also note in some sub-datasets, most of the method can achieve high performance such as cam23-eletric. This is because such datasets have few identities. In addition, we also report CMC curves in Figure 8. We select four representative sub-datasets cam45-bike, cam45-motor, com45-eletric and cam23-bike considering different kinds of bikes and different scale of datasets. From the aspect of identification rate, we can also conclude the new pipeline achieves good performance.

For features, it can be drawn that GOG feature is also the most effective feature for bike person re-id. However, methods based on LOMO perform poorer even than ELF based methods. We argue that this is because bike person images contain more noisy information and LOMO concatenate a noisy information in the maximal process.

From the aspect of metric learning, we can find FDA still reduces the performance than l_2-norm in the multi-query case as usual. However, the LFDA shows high performance as an improved method of FDA. It is also worth mentioning that FDA-related works such as LFDA and XQDA are all of high performance as well as little training time. Compared with FDA-based methods, the MLAPG is time-consuming though it also shows superior performance. Especially, in large scale sub-datasets such as cam45-bike and cam56-bike, it takes more than 6 hours to train. KISSME method also shows competitive performance in some sub-dataset. The NFST algorithm performs best for most of the new pipeline method. But in cam45-bike and cam56-bike, it performs poorly because it requires the feature dimensionality should be larger than the number of training samples, however such a condition cannot be satisfied in these datasets.

VI. CONCLUSION

In this paper, we proposed a new dataset named BPReid which consists of bike persons. It differs from existing datasets in three aspects. First, it is the first bike person re-id dataset. Second, it sampled from a subset of real surveillance network. Third, different camera views have a long distance. In addition, considering the structure of a bike person image has, we develop a new pipeline for bike person re-id. Experiments show the effectiveness of the proposed pipeline. At last, we also test existing person re-id algorithms on this dataset and make a comprehensive evaluation of the dataset.

As can be seen, the proposed pipeline is a naive solution of the bike-person re-id and this work only provides benchmarks dataset as well as methods for this question. In the future work, we will introduce deep learning framework to further improve the performace of this problem.

REFERENCES

[1] L. Sun, Z. Jiang, H. Song, Q. Lu, and A. Men, "Semi-coupled dictionary learning with relaxation label space transformation for video-based person re-identification," *IEEE Access*, vol. 6, pp. 12587–12597, 2018.

[2] T. Ni, Z. Ding, F. Chen, and H. Wang, "Relative distance metric leaning based on clustering centralization and projection vectors learning for person re-identification," *IEEE Access*, vol. 6, pp. 11406–11411, 2018.

[3] D. Gray and H. Tao, "Viewpoint invariant pedestrian recognition with an ensemble of localized features," in *Proc. ECCV*, 2008, pp. 262–275.

[4] R. Zhao, W. Ouyang, and X. Wang, "Human reidentification with transferred metric learning," in *Proc. Asian Conf. Comput. Vis.*, 2012, pp. 31–44.

[5] S. Wang, M. Lewandowski, J. Annesley, and J. Orwell, "Re-identification of pedestrians with variable occlusion and scale," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops*, Nov. 2011, pp. 1876–1882.

[6] W. Li, R. Zhao, T. Xiao, and X. Wang, "DeepReID: Deep filter pairing neural network for person re-identification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 152–159.

[7] L. Zheng, L. Shen, L. Tian, S. Wang, J. Wang, and Q. Tian, "Scalable person re-identification: A benchmark," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 1116–1124.

[8] C. C. Loy, T. Xiang, and S. Gong, "Multi-camera activity correlation analysis," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2009, pp. 1988–1995.

[9] L. Zheng, H. Zhang, S. Sun, M. Chandraker, and Q. Tian, "Person re-identification in the wild," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 3346–3355.

[10] T. Xiao, S. Li, B. Wang, L. Lin, and X. Wang. (2016). "End-to-end deep learning for person search." [Online]. Available: https://arxiv.org/abs/1604.01850

[11] T. Wang, S. Gong, X. Zhu, and S. Wang, "Person re-identification by video ranking," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 688–703.

[12] M. Guo, S. Karanam, W. Liu, O. I. Camps, and R. J. Radke, "DukeMTMC4ReID: A large-scale multi-camera person re-identification dataset," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, Jul. 2017, pp. 1425–1434.

[13] L. Zheng *et al.*, "MARS: A video benchmark for large-scale person re-identification," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 868–884.

[14] Y.-C. Chen, W.-S. Zheng, J.-H. Lai, and P. C. Yuen, "An asymmetric distance model for cross-view feature mapping in person reidentification," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 27, no. 8, pp. 1661–1675, Aug. 2017.

[15] Y.-C. Chen, W. Zheng, and J. Lai, "Mirror representation for modeling view-specific transform in person re-identification," in *Proc. IJCAI*, 2015, pp. 1–7.

[16] T. Matsukawa, T. Okabe, E. Suzuki, and Y. Sato, "Hierarchical Gaussian descriptor for person re-identification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 1363–1372.

[17] B. Ma, Y. Su, and F. Jurie, "Local descriptors encoded by Fisher vectors for person re-identification," in *Proc. Eur. Conf. Comput. Vis.*, 2012, pp. 413–422.

[18] J. van de Weijer, C. Schmid, J. Verbeek, and D. Larlus, "Learning color names for real-world applications," *IEEE Trans. Image Process.*, vol. 18, no. 7, pp. 1512–1523, Jul. 2009.

[19] M. Farenzena, L. Bazzani, A. Perina, V. Murino, and M. Cristani, "Person re-identification by symmetry-driven accumulation of local features," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 2356–2367.

[20] Y. Yang, J. Yang, J. Yan, S. Liao, D. Yi, and S. Z. Li, "Salient color names for person re-identification," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 536–551.

[21] Q. Wang, J. Wan, and Y. Yuan, "Locality constraint distance metric learning for traffic congestion detection," *Pattern Recognit.*, vol. 75, pp. 272–281, Mar. 2018.

[22] Q. Wang, J. Wan, and Y. Yuan, "Deep metric learning for crowdedness regression," *IEEE Trans. Circuits Syst. Video Technol.*, to be published, doi: 10.1109/TCSVT.2017.2703920.

[23] K. Q. Weinberger and L. K. Saul, "Distance metric learning for large margin nearest neighbor classification," *J. Mach. Learn. Res.*, vol. 10, no. 2, pp. 207–244, Feb. 2009.

[24] J. V. Davis, B. Kulis, P. Jain, S. Sra, and I. S. Dhillon, "Information-theoretic metric learning," in *Proc. Int. Conf. Mach. Learn.*, 2007, pp. 209–216.

[25] M. Guillaumin, J. Verbeek, and C. Schmid, "Is that you? Metric learning approaches for face identification," in *Proc. IEEE Int. Conf. Comput. Vis.*, Sep. 2009, pp. 498–505.

[26] W.-S. Zheng, S. Gong, and T. Xiang, "Person re-identification by probabilistic relative distance comparison," in *Proc. IEEE Int. Conf. Comput. Vis.*, Jun. 2011, pp. 649–656.

[27] A. Mignon and F. Jurie, "PCCA: A new approach for distance learning from sparse pairwise constraints," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 2666–2672.

[28] M. Kostinger, M. Hirzer, P. Wohlhart, P. M. Roth, and H. Bischof, "Large scale metric learning from equivalence constraints," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 2288–2295.

[29] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Ann. Eugenics*, vol. 7, no. 2, pp. 179–188, 1936.

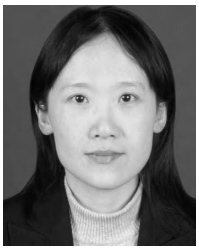
[30] S. Pedagadi, J. Orwell, S. Velastin, and B. Boghossian, "Local Fisher discriminant analysis for pedestrian re-identification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 3318–3325.

[31] F. Xiong, M. Guo, O. Camps, and M. Sznaier, "Person re-identification using kernel-based metric learning methods," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 1–16.

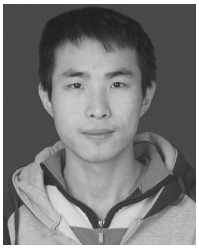
[32] S. Liao, Y. Hu, X. Zhu, and S. Z. Li, "Person re-identification by local maximal occurrence representation and metric learning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 2197–2206.

[33] Z. Li, S. Chang, F. Liang, T. Huang, L. Cao, and J. R. Smith, "Learning locally-adaptive decision functions for person verification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 3610–3617.

[34] L. Zhang, T. Xiang, and S. Gong, "Learning a discriminative null space for person re-identification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 1239–1248.
[35] S. Liao and S. Z. Li, "Efficient PSD constrained asymmetric metric learning for person re-identification," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 3685–3693.
[36] Q. Wang, J. Gao, and Y. Yuan, "A joint convolutional neural networks and context transfer for street scenes labeling," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 5, pp. 1457–1470, May 2018.
[37] S. Ding, L. Lin, G. Wang, and H. Chao, "Deep feature learning with relative distance comparison for person re-identification," *Pattern Recognit.*, vol. 48, no. 10, pp. 2993–3003, Oct. 2015.
[38] D. Cheng, Y. Gong, S. Zhou, J. Wang, and N. Zheng, "Person re-identification by multi-channel parts-based CNN with improved triplet loss function," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 1335–1344.
[39] W. Chen, X. Chen, J. Zhang, and K. Huang, "Beyond triplet loss: A deep quadruplet network for person re-identification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 1–10.
[40] A. Hermans, L. Beyer, and B. Leibe, "In defense of the triplet loss for person re-identification," in *Proc. IEEE Int. Conf. Comput. Vis.*, Nov. 2017, pp. 1–10.
[41] E. Ahmed, M. Jones, and T. K. Marks, "An improved deep learning architecture for person re-identification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 3908–3916.
[42] L. Wu, C. Shen, and A. Hengel. (2016). "PersonNet: Person re-identification with deep convolutional neural networks." [Online]. Available: https://arxiv.org/abs/1601.07255
[43] R. R. Varior, M. Haloi, and G. Wang, "Gated siamese convolutional neural network architecture for human re-identification," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 791–808.
[44] H. Zhao *et al.*, "Spindle net: Person re-identification with human body region guided feature decomposition and fusion," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 1077–1085.
[45] D. Li, X. Chen, Z. Zhang, and K. Huang, "Learning deep context-aware features over body and latent parts for person re-identification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 384–393.
[46] C. Su, J. Li, S. Zhang, J. Xing, W. Gao, and Q. Tian, "Pose-driven deep convolutional model for person re-identification," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 3980–3989.
[47] D. Chen, Z. Yuan, B. Chen, and N. Zheng, "Similarity learning with spatial constraints for person re-identification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 1268–1277.
[48] C. C. Loy, C. Liu, and S. Gong, "Person re-identification by manifold ranking," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2014, pp. 3567–3571.
[49] S. Paisitkriangkrai, C. Shen, and A. van den Hengel, "Learning to rank in person re-identification with metric ensembles," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 1846–1855.
[50] Z. Zhong, L. Zheng, D. Gao, and S. Li, "Re-ranking person re-identification with k-reciprocal encoding," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 3652–3661.
[51] S. Gong, M. Cristani, S. Yan, and C. C. Loy, Eds. *Person Re-Identification*. Berlin, Germany: Springer, 2014.
[52] L. Zheng, Y. Yang, and A. G. Hauptmann. (2016). "Person re-identification: Past, present and future." [Online]. Available: https://arxiv.org/abs/1610.02984



YUAN YUAN (M'05–SM'09) is currently a Full Professor with the School of Computer Science and also with the Center for OPTical IMagery Analysis and Learning, Northwestern Polytechnical University, Xi'an, China. She has authored or co-authored over 150 papers, including about 100 in reputable journals, such as the IEEE TRANSACTIONS and *Pattern Recognition*, and also conference papers in CVPR, BMVC, ICIP, and ICASSP. Her current research interests include visual information processing and image/video content analysis.



JIAN'AN ZHANG received the B.E. degree in information and computing science from the Ocean University of China, Qingdao, China, in 2015. He is currently pursuing the Ph.D. degree with the Center for Optical Imagery Analysis and Learning, Northwestern Polytechnical University, Xi'an. His research interests include computer vision and pattern recognition.

QI WANG (M'15–SM'15) received the B.E. degree in automation and the Ph.D. degree in pattern recognition and intelligent systems from the University of Science and Technology of China, Hefei, China, in 2005 and 2010, respectively. He is currently a Professor with the School of Computer Science, with the Unmanned System Research Institute, and also with the Center for Optical Imagery Analysis and Learning, Northwestern Polytechnical University, Xi'an, China. His research interests include computer vision and pattern recognition.