

Person Re-Identification based on Data Prior Distribution

Yancheng Wu*, Hongchang Chen, Shaomei Li, Chao Gao, Hongxin Zhi, Yuchao Jiang and Yanchuan Wang

National Digital Switching System Engineering & Technological R&D Center, Zhengzhou 450002 China

*Corresponding author

Abstract—In order to solve the problem of insufficient accuracy of the existing person re-identification methods. We propose a neural network model for identifying pedestrian properties and pedestrian ID. Compared with the existing methods, the model mainly has the following three advantages. First, our network adds extra full connection layer, ensure model migration ability. Second, based on the number of samples in each attribute, the loss function of each attribute has been normalized, avoid number unbalanced among the attributes to effect the identification accuracy. Third, we use the distribution of the attribute data in the prior knowledge, through the number to adjust the weight of each attribute in the loss layer, avoid the number of data sets for each attribute of positive and negative samples uneven impact on recognition. Experimental results show that the algorithm proposed in this paper has high recognition rate, and the rank-1 accuracy rate on DukeMTMC dataset is 72.83%, especially on Market1501 dataset. The rank-1 accuracy rate is up to 86.90%.

Keywords- person re-identification; data prior distribution; weight adjustment; deep learning; neural network

I. INTRODUCTION

In recent years, with the popularization of monitoring equipment in public places, person re-identification technology has been paid more and more attention. Person re-identification is a technique that uses computer vision to determine whether there is a specific person in an image or video sequence. However, due to the problems of light, shelter and person posture, the appearance of the same person in different scenes is quite different, which brings great challenges to the person re-identification research. In order to effectively solve these challenges, a large number of researchers have proposed a number of solutions. Most of them only focus on person ID, which simplifies person re-identification to a multi classification problem, but it does not solve the problem of misperception caused by pedestrian appearance change.

With the appearance of the pedestrian dataset with attribute tags, researchers found that the semantic features of the high level are more stable when pedestrian pose change. Therefore, researchers put forward the attribute based pedestrian recognition method. Because attribute learning method is more suitable for human's search habit and can be applied to zero sample learning, so the current method has become a research

hotspot in this field. Among them, Lin et al. [1] proposed a neural network model combined with pedestrian recognition properties and pedestrian ID, greatly improve the accuracy of person re-identification. First of all, the author mark pedestrian attributes on large-scale person re identification datasets Market1501[2] and DukeMTMC[3], then based on the label image, the author design and implementation the Attribute-Person Recognition (APR) nerve network, extraction and recognition the pedestrian ID and attributes of input images, compare the identification results and picture labeling, comparison results as back propagation basis. Training network, extract the vector represents pedestrians from the network, calculate the distance metric, get the recognition results. The network makes full use of the ID information and attribute information of pedestrians. Compared with the existing methods, the accuracy of person re-identification is improved effectively.

In this paper, based on the APR, we have further improved three aspects. Firstly, the network structure is improved, and a layer of full connection is added to the network. Based on the literature [4], the fully connected layer can improve the network in discrimination ability after fine-tuning, ensure that the source model represents ability to migrate. Then, according to the number of data attribute class imbalance problem, the loss function of each attribute based on the loss of the number of the samples are normalized, to improve the network the unbalanced data processing ability. Finally, according to the number of data on each attribute of positive and negative samples is not balanced, the distribution of the attribute data in the prior knowledge, through the number to adjust the weight of each attribute in the layer of loss. The test results show that our algorithm performs better than the mainstream person re-identification algorithm on the public experimental data set, especially the rank-1 accuracy, which has been greatly improved compared to the APR network.

II. RELATED WORK

In person re-identification field, feature design and distance measure are two key components. Feature representation learning method uses visual features to create a robust and discriminative representation for pedestrians, and then calculates similarity between pedestrians based on traditional similarity measure algorithm (Euclidean distance and so on). Swain et al. [5] first separates pedestrians from the background. It proposes using histogram features to represent targets, and identifies different targets by measuring similarity of

histograms. Farenzena et al. [6] based on the prospect of the pedestrians, divides the foreground into different regions by

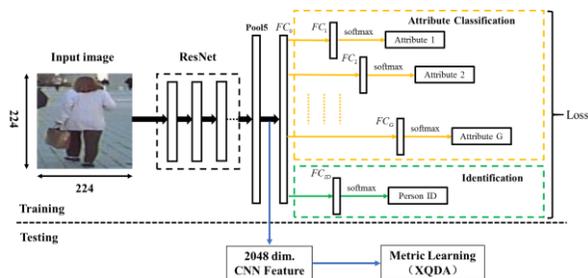


FIGURE I. AN OVERVIEW OF THE NETWORK. DURING TRAINING, IT PREDICTS G ATTRIBUTE LABELS AND AN ID LABEL. THE WEIGHTED SUM OF THE INDIVIDUAL LOSSES IS BACK PROPAGATED. DURING TESTING, WE EXTRACT THE 2048 DIM CNN FEATURE DESCRIPTORS, AND USE XQDA[14] METHOD TO RETRIEVAL.

using the symmetry and asymmetry of the pedestrian area. For each region, it extracts the weight histogram and other characteristics to describe them. Gray et al. [3] divides pedestrian images into 3 parts of the head, upper body and lower body, and then describes the pedestrians with the color histogram of each part.

Considering distance measures, distance metric learning is a measure of similarity between two pedestrians based on the measure learning algorithm. Zhang et al. [8] transformed person re-identification problem into a distance learning problem, and a pedestrian matching model based on probabilistic relative distance is proposed. Du et al. [9] propose a method of learning measure matrix based on statistical inference to measure the similarity of two pedestrian images. Hirzer et al. [10] propose the KISSME algorithm. The algorithm uses the maximum likelihood method to determine whether the two pedestrian vectors are similar.

Attributes are semantic concepts of objects, which are manually defined or directly learned from low level features. Pedestrian attributes have better robustness, which can greatly improve the accuracy of person re-identification. Matsukawa T et al. [11] not only considers the pedestrian attributes, but also combines pedestrian attributes to identify pedestrians. Su et al. [12] uses low rank matrix to describe the similarity between attributes, and obtains a co-occurrence matrix to improve the recognition rate. Li et al. [13] encode the attributes by two values, and the global feature vectors are synthesized at last, and the recognition results are very good.

III. METHODOLOGY

A. Design of Person Re-identification Network Structure Based on Attributes and ID Extract

In this section, we mainly introduce the person re-identification network structure and algorithm flow used to extract pedestrian attributes and ID. In order to extract high robustness pedestrian attribute descriptors, based on prior knowledge of data distribution, we improve the APR network greatly in this paper, and the specific network structure is shown in Figure I.

The network in this paper is mainly composed of two parts. The fully connected layer FC_0 is the boundary. The first half is

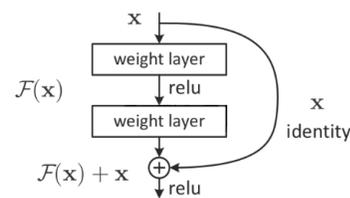


FIGURE II. THE STRUCTURE OF RESIDUAL BLOCK

the residual network: ResNet network [15], and the latter part is the attribute and ID feature classification network.

Firstly, the first half is introduced. The network used in this paper is the Resnet-50 network, that is, a network with 50 layers of depth. The residual block is the core part of the whole network. The specific structure is shown in Figure II. The basic idea is to superimpose an identity map on the shallow network and learn residual function, so that the network will not degenerate and perform better. The residual block has two layers, and the calculation expression is as below:

$$F(x) = W_2 \sigma(W_1 x) \quad (1)$$

$$y = F(x, \{W_i\}) + x \quad (2)$$

where σ is the nonlinear function Relu, W_1 and W_2 are the parameter matrix of two convolution layers.

The latter part of the network is attribute and ID feature classification network, which is mainly used to extract attributes and ID characteristics of person. It consists of fully connected layers, softmax layer and loss layer. Compared with APR network, the biggest improvement of network architecture in this paper is adding FC_0 layers. According to the research of document [16], the main function of the FC_0 layer is to act as a firewall in the process of transferring the ability of model representation. Specifically, our experiment is based on ImageNet's pre-training model to fine-tuning to get the final training results, and ImageNet can be regarded as the source domain. For fine adjustment, if the image in target domain is different from the image in the source domain, such as the person dataset is used in this experiment, compared with ImageNet, the target domain image is not the image of all kinds of objects, but the pedestrian image, the difference is huge. In this case, the result of network fine-tuning without FC is worse than the network containing FC. Therefore, when the source domain is different from the target domain, the addition layer ensures the migration of the model representation capability.

In Figure I, the full link layer FC_{1-G} and FC_{ID} play a major role in classifiers. For every fully connected layer, its parameters are composed of node weight matrix W , offset b and activation function f , which can be expressed as below:

$$y = f(W \cdot x + b) \quad (3)$$

The Softmax layer is mainly calculate probability of the classification results. It can be expressed as below:

$$y_i = \frac{\exp(x_i)}{\sum_{j=1}^n \exp(x_j)} \quad (4)$$

where x_i is the value of the i_{th} node of the softmax layer, y_i is the i_{th} output value, and n is the number of the softmax layer nodes.

The Loss layer uses the cross-entropy loss calculation method, which can be expressed as below:

$$Loss = -\frac{1}{N} \sum_{n=1}^N [p_n \log(\hat{p}_n) + (1 - p_n) \log(1 - \hat{p}_n)] \quad (5)$$

B. Design of Loss Function Based on Data Prior Distribution

The last section introduces the person re-identification network proposed in this paper. This section mainly introduces the improvement of the loss layer calculation in the network. In this paper, we mainly use the prior distribution of data to improve the formula (5) proposed in the previous section, so as to facilitate the description of the problem, we make the following assumptions.

Assuming the training data (Market1501 dataset as an example, Table I) contains N person picture, belonging to M different person, each picture marked G attributes group, e.g., Gender, Age, Hair length, Length of lower-body clothing. For each group, we have $K^{(g)}$ attributes, e.g., child, teenager, adult and old in the Age group, long and short in the Hair length group. The dataset is described in a set way as below:

$$D = \left\{ (x_i, l_i^1, K, l_i^G) \right\}_{i=1}^N \quad (6)$$

where x_i is the i -th image and $l_i^g = (l_{i,1}^g, K, l_{i,K^{(g)}}^g)$ is its attribute label vector for the g -th attribute group. The label $l_{i,k}^g$ takes $l_{i,k}^g \in \{0,1\}$, $l_{i,k}^g = 1$ and $l_{i,k}^g = 0$ respectively represents the presence and the absence of the k -th attribute of group g in x_i .

The loss function in APR network consists of two parts, one is the loss function of attribute recognition, and the other part is the loss function of the person ID recognition. It can be calculated in the following form:

$$L = \lambda L_{ID} + \frac{1}{G} \sum_{i=1}^G L^g \quad (7)$$

where L_{ID} is the loss function of the person ID recognition,

TABLE I. GROUP OF MUTUALLY EXCLUSIVE ATTRIBUTES(MARKET1501)

Group (g)	Attributes	$K^{(g)}$
Gender	male, female	2
Age	child, teenager, adult, old	4
Hair length	long, short	2
Length of lower-body clothing	long, short	2
Type of lower-body clothing	pants, dress	2
Wearing hat	yes, no	2
Carrying bag	yes, no	2
Carrying backpack	yes, no	2
Carrying handbag	yes, no	2
Color of upper-body clothing	black, white, red, purple, yellow, gray, blue, green	8
Color of lower-body clothing	black, white, pink, purple, yellow, gray, blue, green, brown	9

L^g is the loss function of each kind of attribute recognition, λ is a parameter and used to adjust the weights of loss.

The loss function of person ID recognition is in the form of:

$$L_{ID} = -\sum_{m=1}^M \log(p(m))q(m) \quad (8)$$

where the probability value $p(m)$ indicates that the probability of i -th samples belonging to m -th person, it is calculated by the softmax layer which after the FC_{ID} layer. If y is the ground-truth ID label, so that $q(y) = 1$ and $q(m) = 0$ for all $m \neq y$.

The form of loss function of person attribute recognition is as follows:

$$L^g = -\sum_{k=1}^{K^g} \log(p(k))l_{i,k}^g \quad (9)$$

where the probability value $p(k)$ indicates that the probability of i -th samples belonging to k -th attribute in the g -th group, which is calculated by the softmax layer which after FC_{1-G} layer.

On the basis of APR network, this paper mainly improves the loss function of attribute recognition, which is the formula (9). This part of the improvement includes two aspects. One is

the normalization of the loss function based on the number of attribute samples; the other is based on the ratio of the number of positive and negative samples in each attribute, giving different weights to different attributes.

1) *Normalization of loss function based on the number of attribute samples*

The statistics of datasets (e.g., Market1501 dataset in Table II) show that the number of samples is unbalance between attributes, which greatly affects the recognition accuracy of person re-identification. As can be seen from the table, e.g., young in Age group, wearing short sleeved clothes, shorts and other properties of the samples are more, respectively have 569, 712 and 641 samples. Such as handbags, hats, and pink clothes, is very small, only 86, 20 and 2 samples, respectively. In view of the uneven sample data, this paper normalize the loss of each attribute based on the number of samples contained in the loss layer calculation. Finally, the loss layer is rewritten as below:

$$L^g = -\frac{1}{N^g} \sum_{i=1}^N \sum_{k=1}^{K^g} \frac{I_{i,k}^g \log p_{i,k}^g}{N_k^g}, \quad g = 1, K, G \quad (10)$$

where N^g is the number of training samples of the g -th attribute group, N_k^g is the number of training samples of the k -th attributes of the g -th group, probability value $p_{i,k}^g$ is the output of softmax layer which after FC_{1-G} layers, it is the probability of i -th samples belonging to k -th attribute in the g -th group.

For the person ID recognition, the number of samples for each pedestrian is roughly the same, so there are no data imbalances, and the normalization operation is not necessary. The output of the FC_{ID} layer is $z = [z_1, z_2, \dots, z_M] \in R^M$. So the entropy loss of ID classification can be formulated as below:

$$L^{ID} = -\frac{1}{M} \sum_{m=1}^M \log(p_i^m) q(m) \quad (11)$$

where the probability value p_i^m indicates that the probability of i -th samples belonging to m -th person.

For the entire network, the loss can be formulated as below:

$$L = \alpha L^{ID} + (1-\alpha) \frac{1}{G} \sum_{g=1}^G L^g \quad (12)$$

where parameter $0 \leq \alpha \leq 1$ balances the contribution of the two losses and is determined on a validation set of datasets.

2) *Loss function design based on weight adjustment of attribute positive and negative samples*

In the person re-identification datasets, there is not only the imbalance of the number of samples among the attributes, but

also the imbalance of positive and negative data in the attribute. In datasets Market1501 and DukeMTMC, the imbalance of positive and negative sample data in attribute is also very serious. Through statistics, the percentage of positive samples in the data collected from dataset is shown in Table II(e.g.,

TABLE II. THE NUMBER AND RATIO OF ATTRIBUTES TRAINING SAMPLES IN MARKET1501 DATASET

Attribute	Number	Ratio	Attribute	Number	Ratio
upblack	113	0.15	male	431	0.57
upwhite	228	0.30	female	320	0.43
upred	78	0.10	short hair	506	0.67
uppurple	30	0.04	long hair	245	0.33
upyellow	36	0.05	long sleeve	39	0.05
upgray	86	0.11	short sleeve	712	0.95
upblue	46	0.06	long lower body	110	0.15
upgreen	56	0.07	short lower body	641	0.85
handbag no	665	0.89	dress	294	0.39
handbag yes	86	0.11	pants	457	0.61
young	14	0.02	downgray	123	0.16
teenager	569	0.76	downblack	293	0.39
adult	160	0.21	downwhite	58	0.08
old	8	0.01	downpink	29	0.04
bag no	566	0.75	downpurple	2	0.00
bag yes	185	0.25	downyellow	10	0.01
backpack no	552	0.74	downblue	123	0.16
backpack yes	199	0.26	downgreen	14	0.02
hat no	731	0.97	downbrown	69	0.09
hat yes	20	0.03			

Market1501). e.g., the attribute carry bag or not, there is a small number of carry bag, and the majority of them are not carry bag, the ratio is 0.25/0.75, the attribute wear hat or not, there is a small number of wear hat, the ratio is 0.03/0.97. In this case, the effect of the positive sample in the recognition process is too small, and it cannot respond well to the true situation of the pedestrian attributes and affect the effect of the recognition. In order to solve the imbalance between the positive and negative samples, we use the prior distribution of knowledge, positive and negative samples in each attribute proportion, to rewrite the formula (10) as below:

$$L^{g'} = -\frac{1}{N^g} \sum_{i=1}^N \sum_{k=1}^{K^g} \omega_k^g \frac{I_{i,k}^g \log p_{i,k}^g}{N_k^g}, \quad g = 1, K, G \quad (13)$$

$$\omega_k^g = \exp(-p_k^g / \sigma^2) \quad (14)$$

where ω_k^g is the weight of the k -th attribute in the g -th group, p_k^g is the ratio of the positive and negative sample in k -th attribute, σ is the parameter used to adjust the weights.

So, we update the formula (12) as below:

$$L = \alpha L^{ID} + (1-\alpha) \frac{1}{G} \sum_{g=1}^G L^{g'} \quad (15)$$

IV. EXPERIMENT

All experiments in this paper is based on deep learning framework Matconvnet, The experimental platform is a GPU workstation with 64GB memory Intel Core i7 processor and 24GB memory Nvidia TITAN X card.



FIGURE III. EXAMPLE OF DATASET PICTURE (LEFT: MARKET1501, RIGHT: DUKEMTMC)

A. Datasets and Evaluation Protocol

The Market1501 dataset, one of the largest person re-ID datasets, contains 32,668 gallery images and 3,368 query images captured by 6 cameras. Market1501 is split into 751 identities for training and 750 identities for testing. Example of dataset picture was show in Figure III.

The DukeMTMC dataset, it contains 1,812 identities captured by 8 cameras. A number of 1,404 identities appear in more than two cameras, and the rest 408 IDs are distractor images. The training and testing sets both have 702 IDs. So in together, there are 2,228 query images, 16,522 training images and 17,661 gallery images.

Evaluation metrics. For the person re-ID task, we use the Cumulative Matching Characteristic (CMC) curve and the mean average precision (mAP).

B. Implementation Details

We adopt the similar training strategy with APR network. we set the number of epochs to 55. The batch size is set to 64. Learning rate is initialized to 0.001 and changed to 0.0001 in the last 5 epochs. For both networks, the stochastic gradient descent (SGD) is implemented in each mini-batch to update the parameters. The selection process of each parameter is specified in Figure IV.

Parameter α Figure IV(a) shows the performances with varying α . When α takes different values, the recognition accuracy of the network is also changed. For Market1501 dataset, when the attribute loss is not used for the fine-tuning ($\alpha = 1$), the rank-1 rate is 72.36%. When the finetuning is conducted only using the attribute loss ($\alpha = 0$), the rank-1 rate is 76.81%. When $0.1 \leq \alpha \leq 0.9$, the rank-1 rates are better than $\alpha = 0$ or 1. When $\alpha = 0.7$, the rank-1 rate run up to 86.90%. There are similar results on the DukeMTMC dataset. So, we choose $\alpha = 0.7$ in our experiment.

Iteration number Figure IV(b) reports the rank-1 rate per iteration number of the fine-tuning. The performance of person re-identification is evaluated by CNN features every 1,000 iterations. As the number of iterations reaches about 8000 times, the network performance is basically stable.

Attribute selection Figure IV (c)(d) show the recognition accuracy of the network after the different attributes removed from datasets. Some attribute easy to be false detection and missed detection, which brings negative effects to recognition. Therefore, considering the pedestrian attributes tagged from datasets, some attributes with negative effects are excluded. Taking all attributes involved in the training accuracy rate as a benchmark, each time we remove a class of attributes, we get

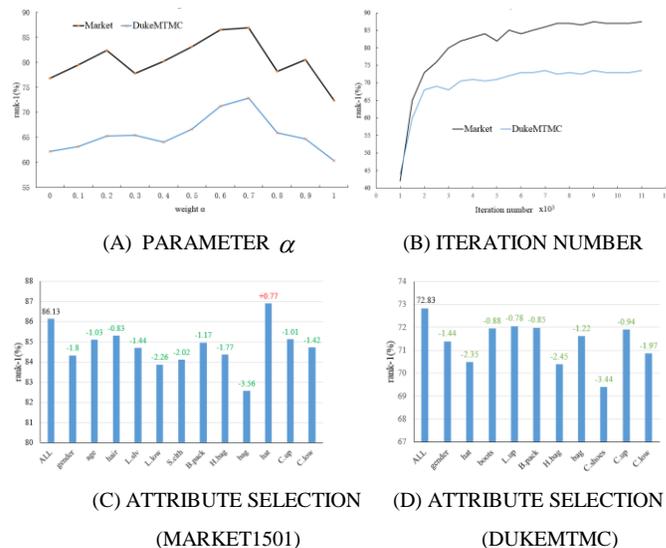


FIGURE IV. SELECTION OF NETWORK PARAMETERS

the recognition accuracy, which is compared with the benchmark, and abscissa is the attribute removed. It can be found that in Market1501 dataset, the accuracy of recognition is increased without using Hat attribute, so this attribute is not used in this experiment. In the DukeMTMC dataset, the recognition effect of pedestrians will be reduced after reducing any attributes, so all attributes should be used.

C. Evaluation of Attribute Recognition

We test attribute recognition on the galleries of the Market1501 and DukeMTMC datasets in Table III and Table IV, respectively. We compare the model learned by APR and ours. We can find two observations point.

First, on both the Market-1501 and DukeMTMC datasets, the overall attribute recognition accuracy is improved by the proposed our network to some extent. The improvement is 0.99% and 2.03% on Market-1501 and DukeMTMC, respectively. So overall speaking, the use of data prior distribution helps in learning a more discriminative attribute model.

Second, we observe that the strong data imbalance attribute, e.g., the attribute Carrying bag or not, the recognition accuracy rate were increased by 7.29% and 4.23%. For some balance data, e.g., the attribute Gender, the recognition accuracy rate were increased by 0.28% and 0.12% only. This indicates that the weight adjustment strategy based on prior knowledge of data presented in this paper has obvious effect on improving pedestrian attributes, especially those with strong data imbalance.

We show two examples of attribute prediction in Figure V. Our system makes correct predictions of all the attributes for the person on the left. For the person on the right, incorrect recognition is observed on the attributes Type of lower-body clothing and Carrying handback.

TABLE III. ACCURACY RATE OF EACH ATTRIBUTE RECOGNITION (MARKET1501)

	gender	age	hair	L.slv	L.low	S.cloth	B.pack	H.bag	bag	C.up	C.low	mean
APR	86.45	87.08	83.65	93.66	93.32	91.46	82.79	88.98	75.07	73.4	69.91	85.33
Ours	86.73	88.14	84.12	93.5	94.54	91.86	85.99	90.67	82.36	77.83	73.82	86.32

TABLE IV. ACCURACY RATE OF EACH ATTRIBUTE RECOGNITION (DUKEMTMC)

	gender	hat	boots	L.up	B.pack	H.bag	bag	C.shoes	C.up	C.low	mean
APR	82.61	86.94	86.15	88.04	77.28	93.75	82.51	90.19	72.29	41.48	80.12
Ours	82.73	89.02	87.17	89.33	81.33	95.81	86.74	93.12	73.04	43.21	82.15



FIGURE V. TWO EXAMPLE OF ATTRIBUTE RECOGNITION

carries out a person re-identification experiment. The experimental results are shown in Table V and Table VI. The "-" indicates that the experiment result is not found.

1) Market1501 dataset

We made a comparative experiment for three improvement points proposed in this paper. In Table V, we got the experimental results separately in the following three aspects: "Ours- FC_0 ", "Ours-normalization" and "Ours-ratio". It can be found that compared with APR network, the rank-1 have been improved, respectively increased 1.08%, 0.63%, 1.38%. The mAP accuracy also has 0.44%, 0.15%, 0.56% promotion, which shows that the three improvements have a great effect on improving the results of person re-identification.

Comparison with the state-of-the-art methods. In this paper, DADM[26], MBC[27] and other methods are compared. It can be found that the rank-1 of these methods are generally not very high. After using the deep learning method, the accuracy rate has been greatly improved, and the APR network has achieved the rank-1 of 84.29% and the mAP of 64.67%. Our paper further improves the accuracy of recognition, achieves 86.90% rank-1 and 65.87% mAP, and the 5,10,20 matching rate has also been promoted correspondingly.

2) DukeMTMC dataset

Similarly, we made a comparative experiment for three improvement points proposed in this paper. In Table VI, we can find that compared with APR network, the rank-1 have been improved, respectively increased 0.87%, 0.26%, 1.13%. The mAP accuracy also has 0.48%, 0.25%, 0.79% promotion,

D. Evaluation of Person Re-ID

Based on Market1501 and DukeMTMC datasets, this paper

which shows that the three improvements have a great effect on improving the results of person re-identification.

Comparison with the state-of-the-art methods. In this paper, Bow[2], LOMO[28] and other methods are compared. It can be found that the rank-1 of these methods are generally not very high. After using the GAN[3] (Generative Adversarial

TABLE V. COMPARISON WITH STATE OF THE ART ON MARKET1501

Methods	rank-1	rank-5	rank-10	rank-20	mAP
DADM[16]	39.4	-	-	-	19.6
MBC[17]	45.56	67	76	82	26.11
SML[18]	45.16	68.12	76	84	-
DLDA[19]	48.15	-	-	-	29.94
SL[20]	51.9	-	-	-	26.35
DNS[21]	55.43	-	-	-	29.87
LSTM[22]	61.6	-	-	-	35.3
S-CNN[23]	65.88	-	-	-	39.55
2Stream[24]	79.51	90.91	94.09	96.23	59.87
GAN[25]	79.33	-	-	-	55.95
Pose[26]	78.06	90.76	94.41	96.52	56.23
Deep[27]	83.7	-	-	-	65.5
APR	84.29	93.2	95.19	97	64.67
Ours- FC_0	85.37	94.05	96.13	97.31	65.11
Ours-normalization	84.92	93.75	95.92	97.46	64.82
Ours-ratio	85.67	94.69	96.75	97.94	65.23
Ours	86.9	95.37	97.03	98.17	65.87

Networks) method, the accuracy rate has run up to 67.68%, and the APR network has achieved the rank-1 of 70.69% and the mAP of 51.88%. Our paper further improves the accuracy of recognition, achieves 72.83% rank-1 and 53.42% mAP.

Finally, a group of re-ID results sample on the Market-1501 dataset is shown in Figure VI. It can be found that although there are some misunderstandings, the overall recognition effect has reached a high degree.

V. CONCLUSIONS

With the development of deep learning technology and the appearance of attribute datasets, person re-identification based on attributes has effectively improved the recognition accuracy in recent years. Based on previous studies, a new deep neural network for person re-identification is designed based on prior knowledge of data. Firstly, the increase in the network structure

of the fully connected layer, improve the migration ability of the network. Secondly, the number of attribute between classes is not balanced, the number of samples was normalized, reducing the imbalance between the number of samples. Finally, use prior knowledge to adjust the imbalance of positive and negative sample data. The experimental results verify the effectiveness of the proposed method.

TABLE VI. COMPARISON WITH STATE OF THE ART ON DUKEMTMC

Methods	rank-1	mAP
BoW+kissme[2]	25.13	12.17
LOMO+XQDA[28]	30.75	17.04
GAN[3]	67.68	47.13
APR	70.69	51.88
Ours- FC_0	71.56	52.36
Ours-normalization	70.92	52.03
Ours-ratio	71.82	52.67
Ours	72.83	53.42



FIGURE VI. A GROUP OF RE-ID RESULTS SAMPLE ON THE MARKET-1501 DATASET

ACKNOWLEDGMENT

This research was supported by The National Natural Science Fund of China (61601513) and China National Digital Switching System Engineering & Technological R&D Center (NDSC).

REFERENCES

- [1] Lin, Y., Zheng, L., Zheng, Z., Wu, Y., & Yang, Y. (2017). Improving person re-identification by attribute and identity learning.
- [2] Zheng, L., Shen, L., Tian, L., Wang, S., Wang, J., & Tian, Q. (2016). Scalable Person Re-identification: A Benchmark. IEEE International Conference on Computer Vision (pp.1116-1124). IEEE.
- [3] Zheng, Z., Zheng, L., & Yang, Y. (2017). Unlabeled samples generated by gan improve the person re-identification baseline in vitro.
- [4] C.-L. Zhang, J.-H. Luo, X.-S. Wei, and J. Wu. (2017). In Defense of Fully Connected Layers in Visual Representation Transfer. Pacific-Rim Conference on Multimedia. 2017
- [5] Swain, M. J., & Ballard, D. H. (1990). Indexing via color histograms. International Conference on Computer Vision, 1990. Proceedings (pp.390-393). IEEE.
- [6] Farenzena, M., Bazzani, L., Perina, A., Murino, V., & Cristani, M. (2010). Person re-identification by symmetry-driven accumulation of local features. Computer Vision and Pattern Recognition (Vol.23, pp.2360-2367). IEEE.
- [7] Gray, D., Brennan, S., & Tao, H. (2007). Evaluating appearance models for recognition, reacquisition, and tracking.
- [8] Zheng, W. S., Gong, S., & Xiang, T. (2011). Person re-identification by probabilistic relative distance comparison. Computer Vision and Pattern Recognition (Vol.42, pp.649-656). IEEE.

- [9] Du, Y. N., & Ai, H. Z. (2014). A statistical inference approach for person re-identification. Journal of Electronics & Information Technology.
- [10] Hirzer, M. (2012). Large scale metric learning from equivalence constraints. IEEE Conference on Computer Vision and Pattern Recognition (pp.2288-2295). IEEE Computer Society.
- [11] Matsukawa, T., & Suzuki, E. (2017). Person re-identification using CNN features learned from combination of attributes. International Conference on Pattern Recognition (pp.2428-2433). IEEE.
- [12] Su, C., Yang, F., Zhang, S., Tian, Q., Davis, L. S., & Gao, W. (2015). Multi-Task Learning with Low Rank Attribute Embedding for Person Re-Identification. IEEE International Conference on Computer Vision (pp.3739-3747). IEEE.
- [13] Li, D., Chen, X., & Huang, K. (2016). Multi-attribute learning for pedestrian attribute recognition in surveillance scenarios. Pattern Recognition. IEEE.
- [14] Liao, S., Hu, Y., Zhu, X., & Li, S. Z. (2015). Person re-identification by Local Maximal Occurrence representation and metric learning. Computer Vision and Pattern Recognition (Vol.8, pp.2197-2206). IEEE.
- [15] Visin, F., Kastner, K., Cho, K., Matteucci, M., Courville, A., & Bengio, Y. (2015). Renet: a recurrent neural network based alternative to convolutional networks. Computer Science, 25(7), 2983-2996.
- [16] Su, C., Zhang, S., Xing, J., Gao, W., & Tian, Q. (2016). Deep attributes driven multi-camera person re-identification. 475-491.
- [17] Ustinova, E., Ganin, Y., & Lempitsky, V. (2017). Multiregion bilinear convolutional neural networks for person re-identification. Computer Science, 48(10), 2993-3003.
- [18] Jose, C., & Fleuret, F. (2016). Scalable metric learning via weighted approximate rank component analysis. 875-890.
- [19] Wu, L., Shen, C., & Hengel, A. V. D. (2017). Deep linear discriminant analysis on fisher networks: a hybrid architecture for person re-identification. Pattern Recognition, 65, 238-250.
- [20] Chen, D., Yuan, Z., Chen, B., & Zheng, N. (2016). Similarity Learning with Spatial Constraints for Person Re-identification. Computer Vision and Pattern Recognition (pp.1268-1277). IEEE.
- [21] Zhang, L., Xiang, T., & Gong, S. (2016). Learning a Discriminative Null Space for Person Re-identification. Computer Vision and Pattern Recognition (pp.1239-1248). IEEE.
- [22] Varior, R. R., Shuai, B., Lu, J., Xu, D., & Wang, G. (2016). A siamese long short-term memory architecture for human re-identification. 135-153.
- [23] Varior, R. R., Haloi, M., & Wang, G. (2016). Gated Siamese Convolutional Neural Network Architecture for Human Re-identification. European Conference on Computer Vision (pp.791-808). Springer, Cham.
- [24] Zheng, Z., Zheng, L., & Yang, Y. (2016). A discriminatively learned cnn embedding for person re-identification.
- [25] Zheng, Z., Zheng, L., & Yang, Y. (2017). Unlabeled samples generated by gan improve the person re-identification baseline in vitro.
- [26] Zheng, L., Huang, Y., Lu, H., & Yang, Y. (2017). Pose invariant embedding for deep person re-identification.
- [27] Geng, M., Wang, Y., Xiang, T., & Tian, Y. (2016). Deep transfer learning for person re-identification.
- [28] Liao, S., Hu, Y., Zhu, X., & Li, S. Z. (2015). Person re-identification by Local Maximal Occurrence representation and metric learning. Computer Vision and Pattern Recognition (Vol.8, pp.2197-2206). IEEE.