Person Transfer GAN to Bridge Domain Gap for Person Re-Identification

Cheng Guan

July 28, 2018

1. Related Work

1.1. Descriptor Learning in Person ReID

Deep learning based descriptors have shown substantial advantages over hand-crafted features on most of person ReID datasets. Some works [33][41][28] learn deep descriptors from the whole images with classification models, where each person ID is treated as a category. Some other works combine verification models with classification models to learn descriptors. Hermans *et al.* [5] show that triplet loss effectively improves the performance of person ReID. Similarly, Chen *et al.* [1] proposed the quadruplet network to learn representations.

The above works learn global descriptors and ignore the detailed cues which might be important for distinguishing persons. To explicitly utilize local cues, Cheng *et al.* proposed a multi-channel part-base network to learn a discriminative descriptor features could be complementary with deep features. They divide the global image into five-length regions. For each region, a histogram descriptor is extracted and concatenate with the global performance, they ignore the misalignment issue caused by fixed body part division.

1.2. Image-to-Image Translation by GAN

Since GAN proposed by Goodfellow *et al* [3], many variants of GAN have been proposed to tackle different tasks, *e.g.*, natural style transfer, super-resolution, sketch-to-image generation, image-to-image translation, *etc*. Among them, image-to-image translation has attracted lots of attention. PTGAN proposed by authors is similar to Cycle-GAN, it also performs image-to-image translation. Differently, extra constraints on person identity are applied to ensure transferred images can be used for model training. Zheng *et al.* [10] adopt GAN to generate new samples for data augmentation in person ReID. Their work differs from ours in both motivation and methodology.

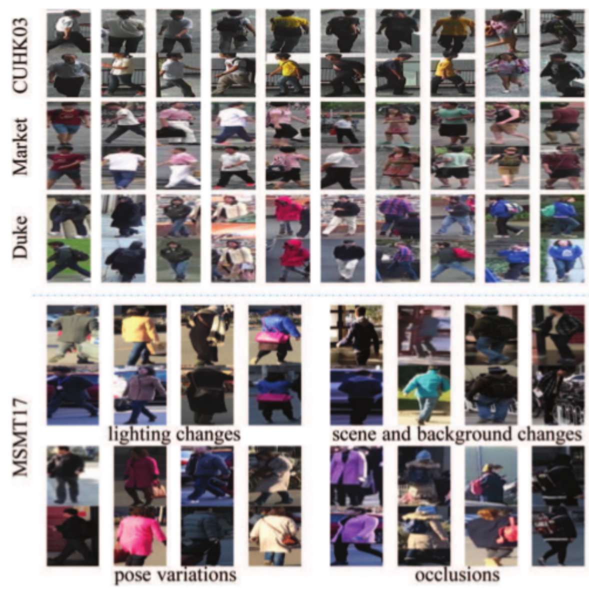

Figure 1. Comparison of person images.Each column shows two sample images of the same identity.

2. MSMT17 Dataset

2.1. Overview of Previous Datasets

Current person ReID datasets have significantly pushed forward the research on person ReID. As shown in Table.1, *DukeMTMC-reID* [10], *CUHK03* [8], and *Market1501* [9] involve larger numbers of cameras and identities than *VIPeR* [4] and *PRID* [6]. The amount of training data makes it possible to develop deep models and show their discriminative power in person ReID. Although current algorithms have achieved high accuracy on those datasets, person ReID is far from being solved and widely applied in real scenarios. Therefore, it is necessary to analyze the limitations of existing datasets.

2.2. Description to MSMT17

Targeting to address above mentioned limitations, the authors collect a new Multi-Scene Multi-Time person ReID

Table 1. Comparison between *MST17* and other person ReID datasets.

Dataset	*MSMT17*	*Duke* [10]	*Market* [9]	*CUHK03* [8]	*CUHK01* [7]	*VIPeR* [4]	*PRID* [6]	*CAVIAR* [2]
BBoxes	**126,441**	36,411	32,668	28,192	3,884	1,264	1,264	1,134
Identities	**4,101**	1,812	1501	1,467	971	632	632	934
Cameras	**15**	8	6	2	10	2	2	2
Detector	**Faster RCNN**	hand	DPM	DPM,hand	hand	hand	hand	hand
Scene	**outdoor,indoor**	outdoor	outdoor	indoor	indoor	outdoor	outdoor	outdoor

dataset (*MSMT17*) by simulating the real scenarios as much as pos- sible. They utilize an 15-camera network deployed in cam- pus. This camera network contains 12 outdoor cameras and 3 indoor cameras. They select 4 days with different weather conditions in a month for video collection. For each day, 3 hours of videos taken in the morning, noon, and after- noon, respectively, are selected for pedestrian detection and annotation. Their final raw video set contains 180 hours of videos, 12 outdoor cameras, 3 indoor cameras, and 12 time slots. Sample images from *MSMT17* are shown and compared in Fig.1. Compared with existing datasets.

References

[1] W. Chen, X. Chen, J. Zhang, and K. Huang. Beyond triplet loss: A deep quadruplet network for person re-identification. In *CVPR*, 2017. 1

[2] D. S. Cheng, M. Cristani, M. Stoppa, L. Bazzani, and V. Murino. Custom pictorial structures for re-identification. In *BMVC*, 2011. 2

[3] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In *NIPS*, 2014. 1

[4] D. Gray and H. Tao. Viewpoint invariant pedestrian recognition with an ensemble of localized features. In *ECCV*, 2008. 1, 2

[5] A. Hermans, L. Beyer, and B. Leibe. In defense of the triplet loss for person re-identification. *arXiv preprint arXiv:1703.07737*, 2017. 1

[6] M. Hirzer, C. Beleznai, P. M. Roth, and H. Bischof. Person re-identification by descriptive and discriminative classification. In *SCIA*, 2011. 1, 2

[7] W. Li, R. Zhao, and X. Wang. Human reidentification with transferred metric learning. In *ACCV*, 2012. 2

[8] W. Li, R. Zhao, T. Xiao, and X. Wang. Deepreid: Deep filter pairing neural network for person re-identification. In *CVPR*, 2014. 1, 2

[9] L. Zheng, L. Shen, L. Tian, S. Wang, J. Wang, and Q. Tian. Scalable person re-identification: A benchmark. In *ICCV*, 2015. 1, 2

[10] Z. Zheng, L. Zheng, and Y. Yang. Unlabeled samples generated by GAN improve the person re-identification baseline in vitro. *arXiv preprint arXiv:1701.07717*, 2017. 1, 2