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

Cheng Guan

July 26, 2018

Abstract

With the development of the Re-Identification (ReID), although the performance of person Re-Identification has been significantly boosted, many challenging issues in real scenarios have not been fully investigated, e.g., the complex scenes and lighting variations, viewpoint and pose changes, and the large number of identities in a camera network. To facilitate the research towards conquering those issues, this paper contributes a new dataset called MSMT17 with many important features, e.g., 1) the raw videos are taken by an 15-camera network deployed in both indoor and outdoor scenes, 2) the videos cover a long period of time and present complex lighting variations, and 3) it contains currently the largest number of annotated identities, e.g., 4,101 identities and 126,441 bounding boxes. The authors also observe that, domain gap commonly exists between datasets, which essentially causes severe performance drop when training and testing on different datasets.

1. Introduction

Person Re-Identification (ReID) targets to match and return images of a probe person from a large-scale gallery set collected by camera networks. Because of its important applications in security and surveillance, person ReID has been drawing lots of attention from both academia and industry. Thanks to the development of deep learning and the availability of many datasets, person ReID performance has been significantly boosted.

Although the performance on current person ReID datasets is pleasing, there still remain several open issues hindering the applications of person ReID. The currently largest *DukeMTMC-reID* [1] contains less than 2,000 identities and presents simple lighting conditions. Those limitations simplify the person ReID task and help to achieve high accuracy. Another challenge they observe is that, there exists domain gap between different person ReID datasets, *i.e.*, training and testing on different person ReID datasets results in severe performance drop. As shown in Fig.1 the domain gap could be caused by many reasons like different lighting conditions, resolutions, human race, seasons, backgrounds, *etc*.

Figure 1. Illustration of the domain gap between $CUHK03$ and $PRID$. It is obvious that, $CUHK03$ and $PRID$ present different styles, *e.g.*, distinct lightings, resolutions, human race, seasons, backgrounds, etc., resulting in low accuracy when training on $CUHK03$ and testing on $PRID$.

To address the second challenge, they propose to bridge the domain gap by transferring persons in dataset *A* to another dataset *B*. The transferred persons from *A* are desired to keep their identities, meanwhile present similar styles, *e.g.*, backgrounds, lightings, *e.g.*, with persons in *A*. they model this transfer procedure with a Person Transfer Generative Adversarial Network (PTGAN), which is inspired by the Cycle-GAN [2].

Their contributions can be summarized into three aspects. 1) *A* new challenging large-scale *MSMT17* dataset is collected and will be released. Compared with existing datasets, *MSMT17* defines more realistic and challenging person ReID tasks. 2) They propose person transfer to take advantages of existing labeled data from different datasets. It has potential to relieve the expensive data annotations on new datasets and make it easy to train person ReID systems in real scenarios. An effective PTGAN model is presented for person transfer. 3) This paper analyzes several issues hindering the applications of person ReID. The proposed *MSMT17* and algorithms have potential to facilitate the future research on person ReID.

2. Related Work

This work is closely related with descriptor learning in person ReID and image-to-image translation by GAN. they briefly summarize those two categories of works in this section.

References

[1] 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] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *ICCV*, 2017. 1