UNSUPERVISED DOMAIN ADAPTIVE PERSON RE-IDENTIFICATION BASED ON ATTRIBUTES

Xiangping Zhu, Pietro Morerio and Vittorio Murino
Pattern Analysis and Computer Vision (PAVIS), Istituto Italiano di Tecnologia (IIT), Genova, Italy

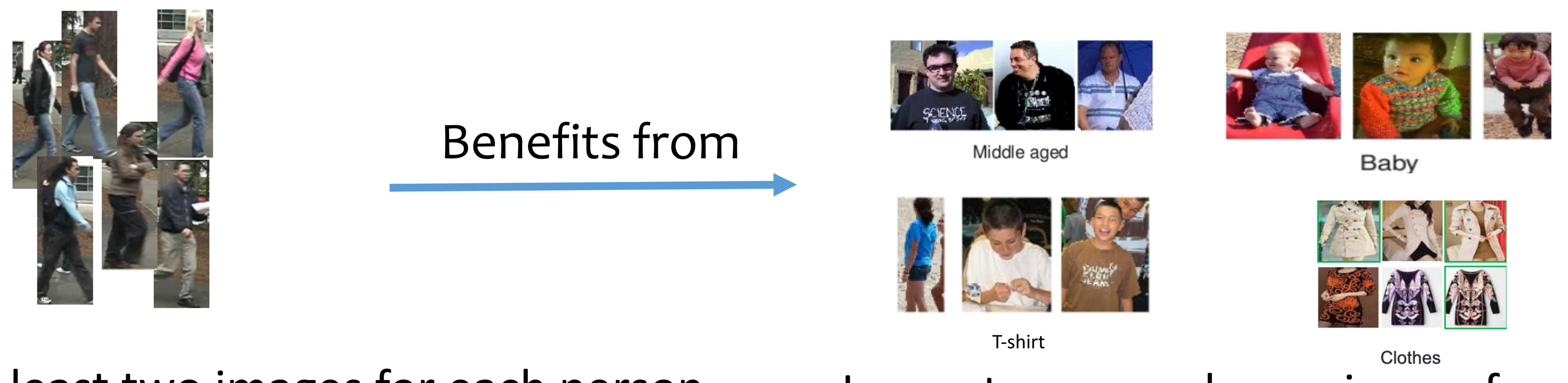
ISTITUTO ITALIANO DI TECNOLOGIA

Introduction

Intuition: pedestrian attributes are coherent across datasets while identity labels are not.

Person **ReID** datasets.

Pedestrian **attribute** recognition datasets [1].



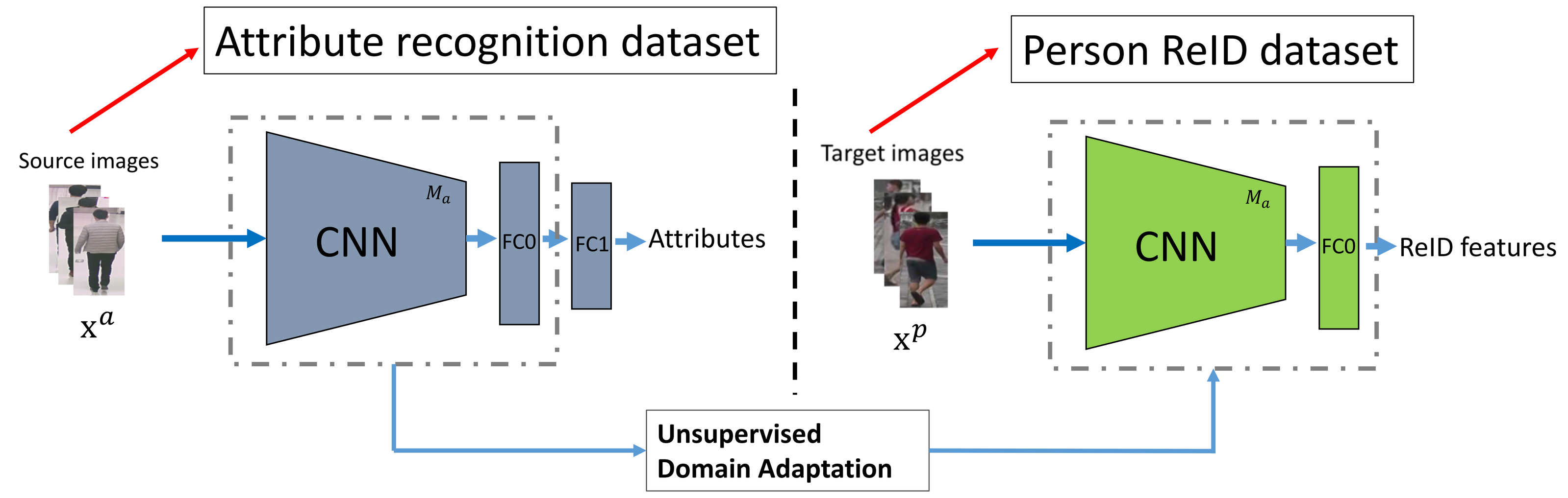
At least two images for each person, usually no person attribute label.

In most cases, only one image for each person, no person ID label and multiple attributes for one person

Contributions:

- A new **unsupervised domain adaptive person re-id framework** only based on pedestrian attributes.
- **A modified unsupervised adversarial adaptation method.**
- Comprehensive experiments.

Methodology



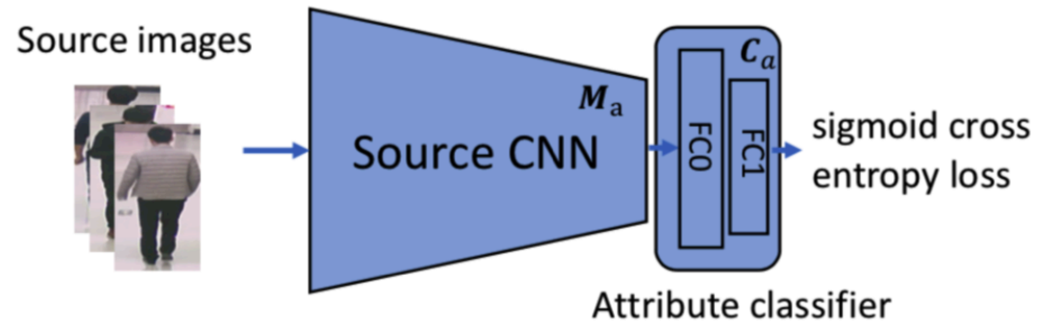
Methodology pipeline:

- ***Step 1***: Train attribute recognition network on pedestrian attribute dataset;
- ***Step 2***: Adapt the trained attribute recognition network to the ReID domain based on unsupervised domain adaptive method;
- ***Step 3***: Use the adapted attribute recognition network for extracting discriminative features for ReID.

Attribute Recognition:

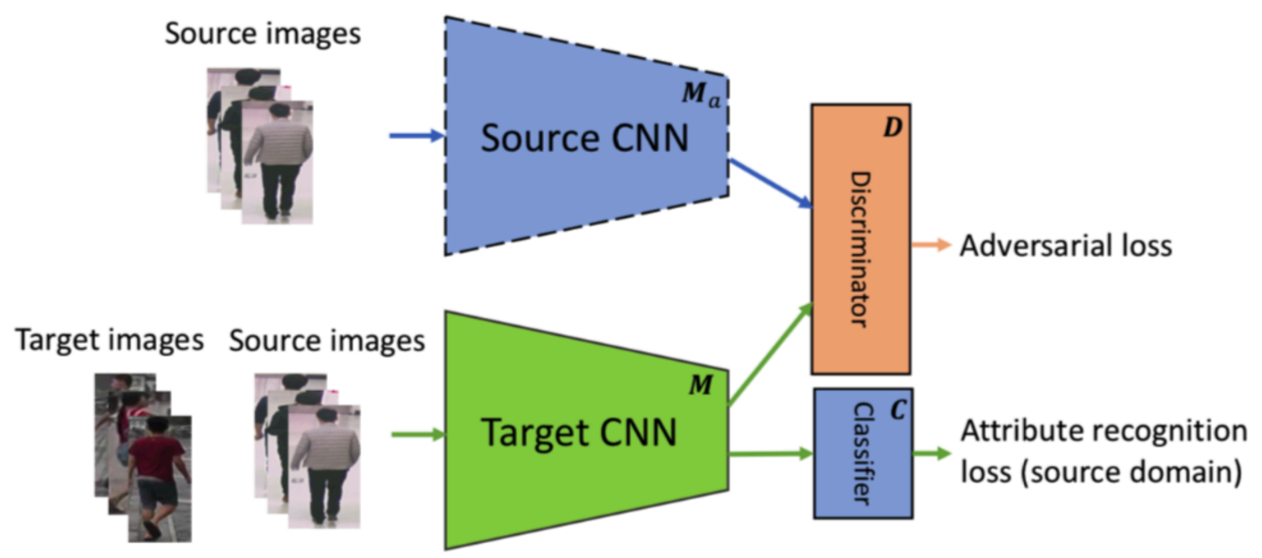
$$L_{attr} = -\mathbb{E}_{\mathbf{x}_i \sim \mathbf{X}^a} \sum_{j=1}^{m} \Big(a_{i,j} \log(\mathbf{C}_a(\mathbf{M}_a(\mathbf{x}_i))) + (1-a_{i,j}) \log(1-\mathbf{C}_a(\mathbf{M}_a(\mathbf{x}_i))) \Big),$$

$a_{i,j}$ denotes the attribute label



Attribute recognition network

Unsupervised Domain Adaptation (UDA):



Compare with ADDA in [2-3], the modified UDA is different in:

- an attribute classifier is added to maintain the attribute recognition performance during domain adaptation;
- both source and target images are fed into the target CNN as in [3].

Unsupervised domain adaptation loss:

$$\min_{\mathbf{M}} \max_{\mathbf{D}} L_{adv} = \mathbb{E}_{\mathbf{x}_i \sim \mathbf{X}^a \cup \mathbf{X}^p} \|\mathbf{D}(\mathbf{M}(\mathbf{x}_i)) - 1\|^2 + \mathbb{E}_{\mathbf{x}_i \sim \mathbf{X}^a} \|\mathbf{D}(\mathbf{M}_a(\mathbf{x}_i))\|^2,$$

Final loss of the modified unsupervised domain adaptation:

$$\min_{\{\mathbf{M},\mathbf{C}\}} \max_{\mathbf{D}} L_{adv} + \alpha L_{attr},$$

α is the hyper-parameter, L_{attr} is the same as in attribution recognition with $\mathbf{C}, \mathbf{M}$ in place of $\mathbf{C}_a, \mathbf{M}_a$

Experimental Results

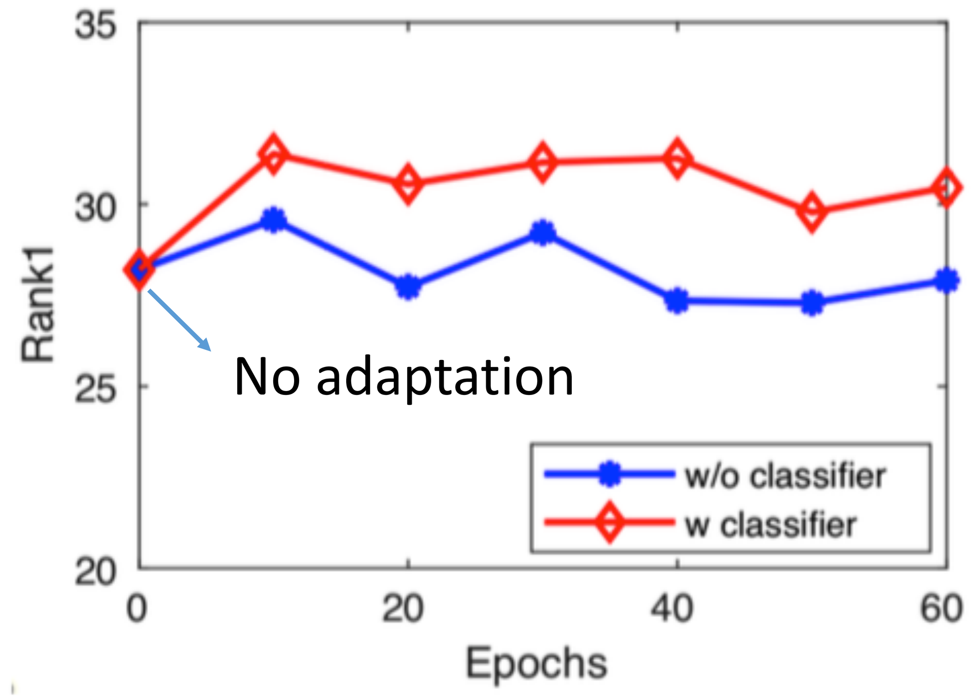
Three datasets are considered:

- RAP is an attribute recognition dataset [4];
- Market1501 and DukeMTMC-reID are two ReID datasets, but they are also labeled with pedestrian attributes.

Effectiveness of UDA:

Table 1: Comparing experimental results before and after adaptation

	Market-1501		DukeMTMC-reID	
Metric	Rank1	mAP	Rank1	mAP
w/o adaptation	28.2%	8.7%	15.6%	4.9%
w adaptation	**32.1%**	**10.6%**	**18.7%**	**6.5%**



Effectiveness of the classifier in UDA

Comparisons with SOTA results:

Table 2: Performance comparisons with existing attribute based unsupervised domain-adaptive person ReID methods[5-6]:

Source → Target	Market-1501 → DukeMTMC-reID			
Metric	Rank1	Rank5	Rank10	mAP
TJ-AIDL	24.3%	38.3%	45.7%	10.0%
MMFA	15.8%	26.0%	48.2%	5.7%
Ours	**28.6%**	**44.2%**	**51.7%**	**13.1%**
Source → Target	DukeMTMC-reID → Market-1501			
Metric	Rank1	Rank5	Rank10	mAP
TJ-AIDL	38.0%	59.2%	67.6%	13.6%
MMFA	35.5%	55.3%	64.0%	12.7%
Ours	**43.0%**	**63.3%**	**70.6%**	**17.1%**

References:

[1] Deng Y, Luo P, Loy CC, Tang X. Pedestrian attribute recognition at far distance. ACM MM, 2014.
[2] Tzeng E, Hoffman J, Saenko K, Darrell T. Adversarial discriminative domain adaptation. CVPR, 2017.
[3] Volpi R, Morerio P, Savarese S, Murino V. Adversarial feature augmentation for unsupervised domain adaptation. CVPR, 2018.
[4] Li D, Zhang Z, Chen X, Ling H, Huang K. A richly annotated dataset for pedestrian attribute recognition, IEEE TIP, 2018
[5] Wang J, Zhu X, Gong S, Li W. Transferable joint attribute-identity deep learning for unsupervised person re-identification. CVPR, 2018.
[6] Lin S, Li H, Li CT, Kot AC. Multi-task mid-level feature alignment network for unsupervised cross-dataset person re-identification, BMVC, 2018