



Hard Samples Rectification for Unsupervised Cross-domain Person Re-ID

Chih-Ting Liu*, Man-Yu Lee*, Tsai-Shien Chen, Shao-Yi Chien

Graduate Institute of Electronics Engineering, National Taiwan University

** denotes equal contributions*

2021 IEEE International Conference on Image Processing, Anchorage, Alaska, USA

Introduction

- What is person re-identification (Re-ID)?
 - Match the images of same person across multiple cameras.
- Challenge for supervised learning methods
 - It is hard to annotate persons across cameras.
 - Unrealistic to deploy in a real-world scenario.
- Unsupervised cross-domain learning is popular





Related Work

- Domain adaptation
 - Clustering-based method_[5,6,7], which are popular and very effective.



Cluster the features of unlabelled data

[5] Hehe Fan, Liang Zheng, Chenggang Yan, and Yi Yang. "Unsupervised person re-identification: Clustering and fine-tuning". In TOMM, 2018

Motivation

- Problems in clustering-based Re-ID methods
 - Hard positive pair \rightarrow Easily be mis-clustered to different groups
 - Hard negative pair → Different people with similar appearance are in the same group



Proposed Method

- Goal : Rectify hard samples in clustering results
 - Inter-Camera Mining (ICM) → Hard positive samples
 - Part-Based Homogeneity (PBH) → Hard negative samples



Clustered-Based Baseline Framework

• Iterative learning of clustering and CNN training



Clustered-Based Baseline Framework



[15] Ergys Ristani and Carlo Tomasi. Features for multi-target multi-camera tracking and re-identification. In CVPR, 2018

Technique 1 : Inter-Camera Mining

- Problem : CNN is only optimized with similar positive pairs within the same camera
- Solution : Shorten the distance of hard positive samples



Inter-Camera Mining

• Obtain *Rank*(*P*) for every image *P*

1. Rank all images in target dataset by feature distance



2. Remove images captured by the same camera as image P



3. Select **top-k** samples to form Rank(P)



Inter-Camera Mining

• Additionally adopt Top-K best buddies pair [18]



Check if P exist in Rank(3)



Inter-Camera Mining

• Additionally adopt Top-K best buddies pair [16]





Technique 2: Part-Based Homogeneity

- Problem: Different people with only subtle difference are possibly grouped into same cluster
- Solution: Partition off hard negative samples



Part-Based Homogeneity

- Define imperfect cluster
 - Calculate Silhouette score [9]

$$Sil(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

b(i) : average distance to i
within closest cluster
a(i) : average distance to i
within same cluster

- mSil(j) : mean Silhouette score of cluster j
- Select cluster that $mSil(j) < \lambda$

[9] Peter J Rousseeuw. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics, 1987

Part-Based Homogeneity

Assign new pseudo labels based on part-based features



Proposed Framework

• Iterative learning of clustering and CNN training



Proposed Framework

• Iterative learning of clustering and CNN training



Experiment Setting

- Dataset
 - Market-1501 [10]
 - DukeMTMC-ReID [11]
- Evaluation measure
 - R1 (Rank-1 accuracy)
 - mAP (mean average precision)



[10] Liang Zheng, Liyue Shen, Lu Tian, Shengjin Wang, Jingdong Wang, and Qi Tian. Scalable person re-identification: A benchmark. In ICCV, 2015.
[11] E. Ristani, F. Solera, R. Zou, R. Cucchiara, and C. Tomasi. Performance measures and a data set for multi-target, multicamera tracking. In European Conference on Computer Vision workshop on Benchmarking Multi-Target Tracking, 2016



• Effectiveness of ICM and PBH



Experimental setting	loss functions & components				$Duke \rightarrow Martket$		Market \rightarrow Duke	
	\mathcal{L}_{CE}	\mathcal{L}_{trip}	\mathcal{L}_{ICM}	PBH	R1	mAP	R1	mAP
Direct Transfer					50.1	20.9	36.2	18.3
Baseline	\checkmark	\checkmark			72.9	46.3	60.2	42.2
Baseline w/ PBH	\checkmark	\checkmark		\checkmark	74.5	47.1	63.5	44.6
Baseline w/ \mathcal{L}_{ICM}	\checkmark	\checkmark	\checkmark		83.8	63.3	73.5	54.4
HSR (Ours)	\checkmark	\checkmark	\checkmark	\checkmark	85.3	65.2	76.1	58.0

• Effectiveness of ICM and PBH



Experimental setting	loss functions & components				$Duke \rightarrow Martket$		Market \rightarrow Duke	
	\mathcal{L}_{CE}	\mathcal{L}_{trip}	\mathcal{L}_{ICM}	PBH	R1	mAP	R1	mAP
Direct Transfer					50.1	20.9	36.2	18.3
Baseline	\checkmark	\checkmark			72.9	46.3	60.2	42.2
Baseline w/ PBH	\checkmark	\checkmark		\checkmark	74.5	47.1	63.5	44.6
Baseline w/ \mathcal{L}_{ICM}	\checkmark	\checkmark	\checkmark		83.8	63.3	73.5	54.4
HSR (Ours)	\checkmark	\checkmark	\checkmark	\checkmark	85.3	65.2	76.1	58.0

• Effectiveness of ICM and PBH



Experimental setting	loss functions & components			Duke \rightarrow Martket		Market \rightarrow Duke		
	\mathcal{L}_{CE}	\mathcal{L}_{trip}	\mathcal{L}_{ICM}	PBH	R 1	mAP	R1	mAP
Direct Transfer					50.1	20.9	36.2	18.3
Baseline	\checkmark	\checkmark			72.9	46.3	60.2	42.2
Baseline w/ PBH	\checkmark	\checkmark		\checkmark	74.5	47.1	63.5	44.6
Baseline w/ \mathcal{L}_{ICM}	\checkmark	\checkmark	\checkmark		83.8	63.3	73.5	54.4
HSR (Ours)	\checkmark	\checkmark	\checkmark	\checkmark	85.3	65.2	76.1	58.0

Comparison with State-of-the-arts

Methods	Duke -	\rightarrow Market	Marke	$t \rightarrow Duke$	
wiethous	R1	mAP	R1	mAP	
PUL [7]	45.5	20.5	30.0	16.4	-
CAMEL [6]	54.5	26.3	-	-	
SPGAN [3]	58.1	26.9	46.9	26.4	
HHL [4]	62.2	31.4	46.9	27.2	
MAR [19]	67.7	40.0	67.1	48.0	
PAST [8]	78.4	54.6	72.4	54.3	
SSG [9]	80.0	58.3	73.0	53.4	Clustering-based
<i>p</i> MR-SADA [20]	83.0	59.8	74.5	55.8	Ŭ
GDS-H [10]	81.1	61.2	73.1	55.1	
HSR (Ours)	85.3	65.2	76.1	58.1	-

Conclusion

- We propose an inter-camera mining technique to mine potentially hard positive samples
- We propose a part-based homogeneity technique to partition off the imperfect clusters containing hard negative samples
- Conduct extensive experiments and achieve state-of-the-art on two benchmarks among existing unsupervised methods



Reference

[1] Weijian Deng, Liang Zheng, Guoliang Kang, Yi Yang, Qixiang Ye, and Jianbin Jiao. "Image-image domain adaptation with preserved self-similarity and domain-dissimilarity for person re-identification". In CVPR, 2018.

[2] Zhun Zhong, Liang Zheng, Shaozi Li, and Yi Yang. "Generalizing a person retrieval model hetero-and homogeneously". In ECCV, 2018

[3] Hong-Xing Yu, Wei-Shi Zheng, Ancong Wu, Xiaowei Guo, Shaogang Gong, and Jian-Huang Lai. "Unsupervised person re-identification by soft multilabel learning". In CVPR,2019

[4] Chih-wei Wu, Chih Ting Liu, Wei-Chih Tu, Yu Tsao, Yu-Chiang Wang, and Shao-Yi Chien, "Space-time Guided Association Learning for Unsupervised Person Re-Identification". In ICIP, 2020

[5] Hehe Fan, Liang Zheng, Chenggang Yan, and Yi Yang. "Unsupervised person re-identification: Clustering and fine-tuning". In TOMM, 2018

[6] Yang Fu, Yunchao Wei, Guanshuo Wang, Yuqian Zhou, Honghui Shi, and Thomas SHuang. "Self-similarity grouping: A simple unsupervised cross domain adaptation approach for person re-identification". In ICCV, 2019

[7] Xinyu Zhang, Jiewei Cao, Chunhua Shen, and Mingyu You. "Self-training with progressive augmentation for unsupervised cross-domain person reidentification". In ICCV, 2019

[8] Liu, Hongye, et al. "Deep relative distance learning: Tell the difference between similar vehicles." In CVPR. 2016.

[9] Peter J Rousseeuw. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics, 1987

Reference

[10] Liang Zheng, Liyue Shen, Lu Tian, Shengjin Wang, Jingdong Wang, and Qi Tian. Scalable person re-identification: A benchmark. In ICCV, 2015.

[11] E. Ristani, F. Solera, R. Zou, R. Cucchiara, and C. Tomasi. Performance measures and a data set for multi-target, multicamera tracking. In European Conference on Computer Vision workshop on Benchmarking Multi-Target Tracking, 2016

- [12] Wojke, et al. "Simple online and realtime tracking with a deep association metric." ICIP. IEEE, 2017.
- [13] A. Milan, L. Leal-Taixé, I. Reid, S. Roth, and K. Schindler. Mot16: A benchmark for multi-object tracking.arXivpreprintarXiv:1603.00831, 2016
- [14] Liu, Hongye, et al. "Deep relative distance learning: Tell the difference between similar vehicles." In CVPR. 2016
- [15] Ergys Ristani and Carlo Tomasi. Features for multi-target multi-camera tracking and re-identification. In CVPR, 2018
- [16] Tali Dekel, et al. Best-buddies similar-ity for robust template matching, In CVPR, 2015