Self-paced Bottom-up Clustering Network with Side Information for Person Re-Identification

Mingkun Li, Chun-Guang Li, Ruopei Guo, and Jun Guo
School of Information and Communication Engineering, Beijing University of Posts and Telecommunications
Email: {mingkun.li, lichunguang, guo2016, guojun}@bupt.edu.cn

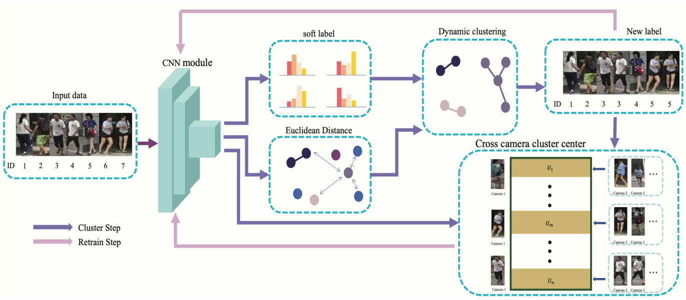
Introduction

In this paper, we propose a Self-Paced bottom-up Clustering Network with Side Information (SPCNet-SI) for person Re-ID. Different from prior methods, we exploit the associated camera information (i.e. from which camera the images comes from) into a bottom-up clustering network to help it learn cross-camera view features. Specifically, we incorporate the camera side information into the repelled loss, and rather than employing a hard assignment in the bottom-up clustering, we employ a soft-label based assignment to find candidate clusters to be merged. Moreover, we propose a dynamic strategy to control the bottom-up merging process to implement an easy-to-hard and slow-to-fast clustering process, which is essentially a self-paced clustering.

Paper Contributions:

- We design a cross-camera repelled loss to exploits the camera side information and encourages to explore the association under different camera views.
- We propose a soft-label based assignment scheme in the bottom-up clustering. Compared to the existing hard assignment, the soft-label based assignment takes into account the ambiguity in data points, making the clusters merge more reliable in the earlier merging stage.
- We present an effective dynamic strategy to regularize the cluster merging process, in which a gradual development scheme is used to control the merging size and an annealing scheme is adopted to adjust the regularization parameter to help select proper clusters to merge.

Our Proposal



A.Cross-camera Repelled Loss

We design a cross-camera repelled loss, which is aimed to explore the latent associations between different cluster centers under different camera views. Specifically, the cross-camera repelled loss is defined as follows:

$$\mathcal{L}(\boldsymbol{x}_i, \tilde{\boldsymbol{u}}_{\omega(i)}, \Theta) = -\ln \frac{\exp(\tilde{\boldsymbol{u}}_i^\top \varphi(\boldsymbol{x}_i, \Theta)/\tau)}{\sum_{\ell=1}^{n} \exp(\tilde{\boldsymbol{u}}_\ell^\top \varphi(\boldsymbol{x}_i, \Theta)/\tau)}, \quad (5)$$

where $\tilde{\boldsymbol{u}}_i$ is the cross-camera average of the nearest neighboring cluster centers to the i-th cluster center which is defined as

$$\tilde{\boldsymbol{u}}_i = \frac{1}{|\mathcal{I}(i)|} \sum_{j \in \mathcal{I}(i)} \boldsymbol{u}_j. \quad (6)$$

where $\boldsymbol{u}_i$ is the i-th cluster center and $\mathcal{I}(i)$ is the set of the indexes for "cross-camera neighbors", which are picked from k different camera views.

B. *Finding Candidate Clusters to be Merged*

To find the proper candidate clusters to merge, we design a soft-label based assignment method. Specifically, rather than using the minimum Euclidean distance, we take into account both the Euclidean distance between two clusters and the inside information of each cluster. Specifically, we define the dissimilarity between a pair of clusters $\mathcal{C}_k$ and $\mathcal{C}_\ell$ where $k, \ell = 1, \cdots, n$ as follows:

$$\Delta(\mathcal{C}_k, \mathcal{C}_\ell) = \delta(\mathcal{C}_k, \mathcal{C}_\ell) + \lambda(\Omega(\mathcal{C}_k) + \Omega(\mathcal{C}_\ell)), \quad (7)$$

where

$$\delta(\mathcal{C}_k, \mathcal{C}_\ell) = \min_{i \in \mathcal{C}_k, j \in \mathcal{C}_\ell} \|\varphi(\boldsymbol{x}_i, \Theta) - \varphi(\boldsymbol{x}_j, \Theta)\|_2, \quad (8)$$

and

$$\Omega(\mathcal{C}_k) = \sum_{i \in \mathcal{C}_k} (1 - \gamma_{ik}), \quad \Omega(\mathcal{C}_\ell) = \sum_{j \in \mathcal{C}_\ell} (1 - \gamma_{j\ell}), \quad (9)$$

where γ_{ik} and $\gamma_{j\ell}$ are the correspondence scores as defined in (4) to measure the membership of $\varphi(\boldsymbol{x}_i, \Theta)$ to cluster $\mathcal{C}_k$ and $\varphi(\boldsymbol{x}_j, \Theta)$ to cluster $\mathcal{C}_\ell$, respectively.

C. Dynamically Self-Paced Merging

To prevent mis-merging, we propose a dynamically self-paced merging strategy, which is implemented by a gradual development bottom-up merging strategy and assisted with an annealing scheme.

1) Gradual Development Strategy in Bottom-Up Merging:

$$s^{(t)} = \lfloor \alpha t N \rfloor, \quad (11)$$

2) Annealing Scheme for Adjusting the Penalty Parameter in Merging:

At the t-th step, we adjust λ as follows:

$$\lambda^{(t)} = \lambda^{(0)} - \nu t, \quad (13)$$

Experiments

TABLE I
COMPARISON ON MARKET-1501. * MEANS RECURRENCE BY US

method	Reference	Market1501			
		mAP	rank-1	rank-5	rank-10
LOMO [11]	CVPR'15	8.0	27.2	41.6	49.1
Bow [41]	ICCV'15	14.8	35.8	52.4	60.3
PUL [37]	TOMM'18	22.8	51.5	70.1	76.8
DECAMEL [42]	TPAMI'19	32.4	60.2	76.0	
CAMEL[20]	Iccv'17	26.31	54.45	73.10	79.69
PGPPM [43]	CVPR'18	33.9	63.9	81.1	86.4
HHL [32]	ECCV'18	31.4	62.2	78.0	84.0
TJ-AIDL [36]	CVPR'18	26.5	58.2	74.8	-
SPGAN [31]	CVPR'18	26.7	58.1	76.0	82.7
SyRI [23]	ECCV'18	-	65.7	-	-
PTGAN [30]	CVPR'18	15.7	38.6	57.3	-
BUC* [39]	AAAI'19	28.7	60.6	73.5	77.2
SPCNet-SI	**Ours**	**34.9**	**68.7**	**85.2**	**88.9**

TABLE II
COMPARISON ON DUKEMTMC-REID.

method	Reference	DukeMTMC-ReID			
		mAP	rank-1	rank-5	rank-10
LOMO [11]	CVPR'15	4.8	12.3	21.3	26.6
Bow [41]	ICCV'15	8.5	17.1	28.8	34.9
PUL [37]	TOMM'18	22.3	41.1	46.6	63.0
CAMEL[20]	Iccv'17	19.8	40.2	57.5	64.9
PGPPM [43]	CVPR'18	17.9	36.3	54.0	61.6
HHL [32]	ECCV'18	27.2	46.9	61.0	66.7
TJ-AIDL [36]	CVPR'18	23.0	44.3	59.6	-
SPGAN [31]	CVPR'18	26.4	46.9	**62.6**	**68.5**
PTGAN [30]	CVPR'18	13.5	27.4	43.6	-
BUC* [39]	AAAI'19	21.8	40.2	51.3	56.7
SPCNet-SI	**Ours**	**27.7**	**47.8**	61.8	64.1

Conclusions

- Through training with the cross-camera repelled loss, SPCNet-SI gradually incorportates to explore the cross-camera association.
- Owning to the soft-label based assignment and dynamical self-paced mechanisms in regularizing the merging process, SPCNet-SI learns to cluster data points in an easy-to-hard way with a slow-to-fast merging process, leading to more accurate results.
- Experiments on two benchmark datasets Market-1501 and DukeMTMC-ReID demonstrated promising performance.