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

Mingkun Li, Chun-Guang Li, Ruopei Guo, and Jun Guo

Contents

- Introduction
- Method
- Result
- Conclusions

Person Re-Identification

Person Re-ID

Person re-identification (Re-ID) aims to match one person with images captured by various cameras.

Supervised method

Traditional supervised methods for person Re-ID demand enormous amount of **manual annotated data** and the performance badly drops when deployed on a new dataset.

Unsupervised method

Unsupervised Person Re-Identification

Method	**Dataset**	**Need label**
Unsupervised domain adaptation (UDA)	Source dataset + Target dataset	Source datadset
Pure unsupervised	Target dataset	None

Unsupervised Person Re-Identification

SPCNet-SI

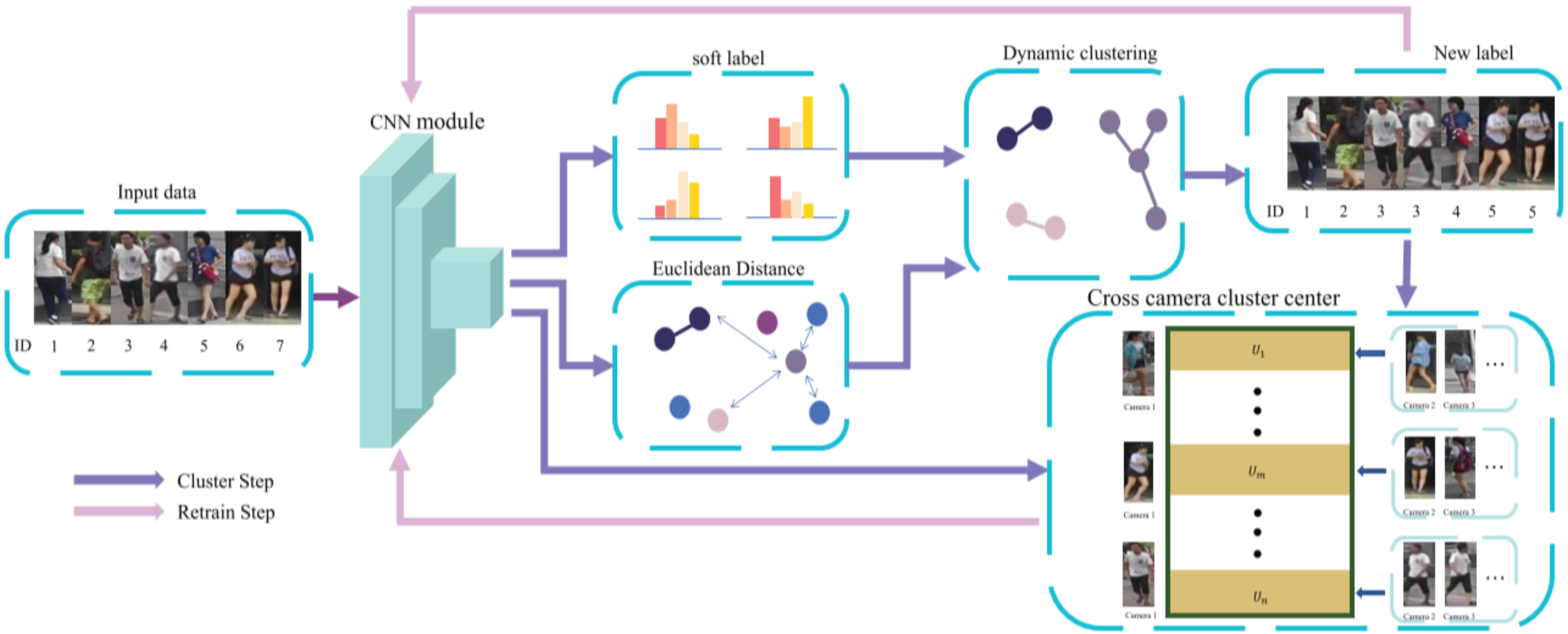
- Pure unsupervised (Target dataset)

+

- Side information (Camera information)

Method

Illustration for the SPCNet-SI.

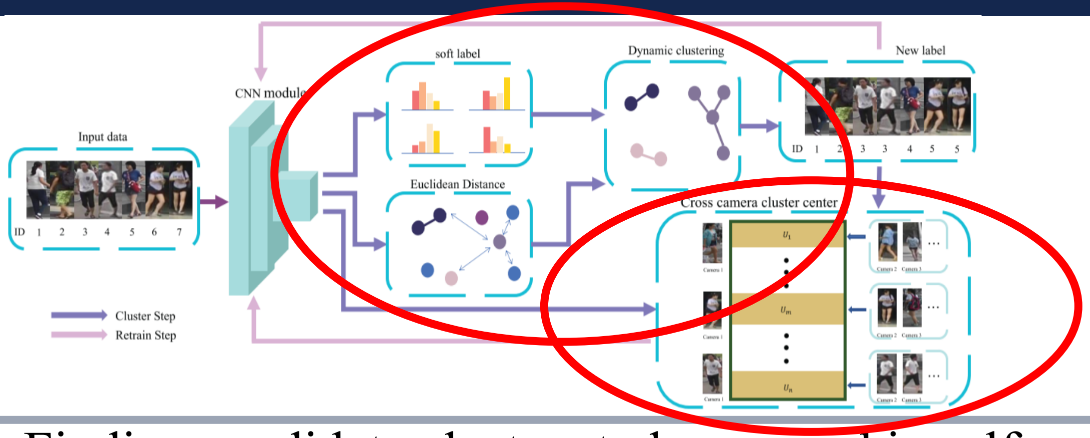


SPCNet-SI

Algorithm 1 SPCNet-SI

Input: Dataset $\mathcal{X} = \{\boldsymbol{x}_1, \ldots \boldsymbol{x}_N\}$, $\lambda^{(0)}$, ν, α, T, initial pseudo labels $\{\omega(i)\}_{i=1}^{N}$, and set $n \leftarrow N$ and $t \leftarrow 0$.
Output: $\{\omega(i)\}_{i=1}^{N}$.

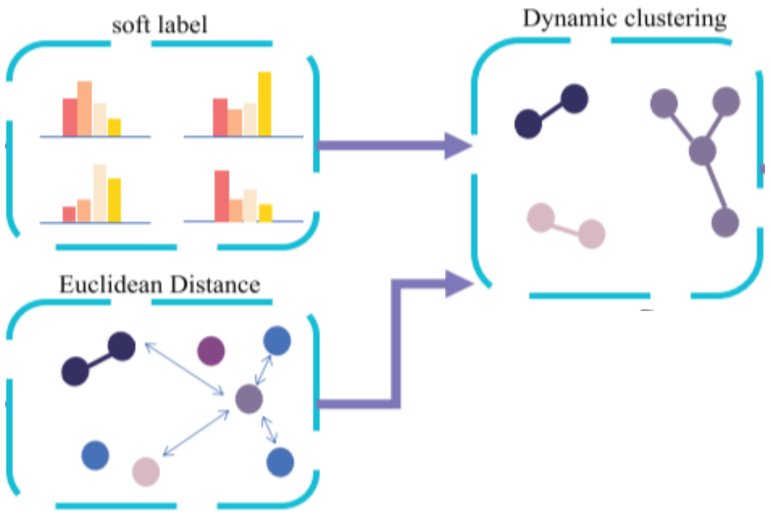
Train the network with $\{\omega(i)\}_{i=1}^{N}$ via the loss in (1);
while $t \leq T$ **do**
  Set $t \leftarrow t + 1$;
  Compute $\{\omega(i)\}_{i=1}^{N}$ via (3);
  Set $s \leftarrow \lfloor \alpha t N \rfloor$ and $\lambda \leftarrow \lambda^{(0)} - \nu t$;
  Find s candidates cluster pairs via (10);
  Set $n \leftarrow n - s$ and update $\tilde{\boldsymbol{u}}_i$ via (5)
  Train network (1) to update Θ via SGD and $\mathcal{U}$ via the loss in (5)
  Evaluate model perofmance P on the validation set
  Once $P_t < P_{t-1}$: train the network with $\{\omega(i)\}_{i=1}^{N}$ via the loss (1) in later iterations;
  if $P_t > P_{t-1}$ **then**
    Best model $= \phi(\cdot, \Theta)$
    $P_{best} = P_t$
  end if
end while



1. Finding candidate clusters to be merged in self-paced way

2. Update the cluster center with side information. And train with cross-camera repelled loss

1. Finding candidate clusters to be merged in self-paced way



Algorithm 1 SPCNet-SI

Input: Dataset $\mathcal{X} = \{\boldsymbol{x}_1, ...\boldsymbol{x}_N\}$, $\lambda^{(0)}$, ν, α, T, initial pseudo labels $\{\omega(i)\}_{i=1}^N$, and set $n \leftarrow N$ and $t \leftarrow 0$.
Output: $\{\omega(i)\}_{i=1}^N$.

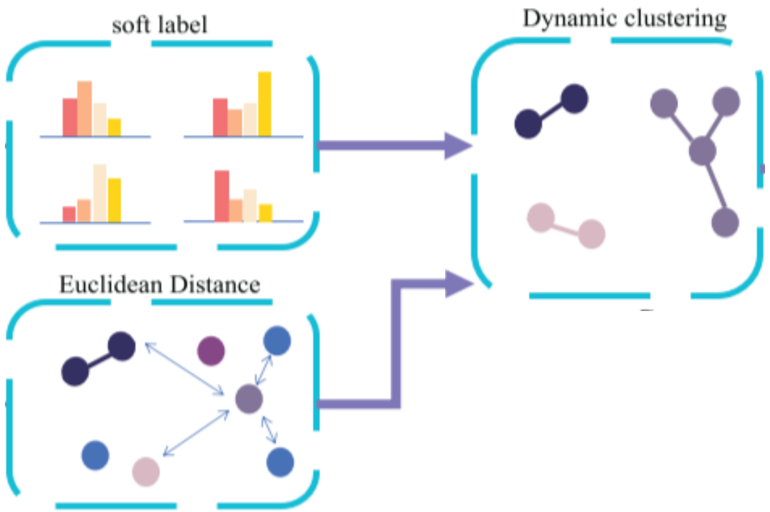
```
Train the network with {ω(i)}_{i=1}^N via the loss in (1);
while t ≤ T do
    Set t ← t + 1;
    Compute {ω(i)}_{i=1}^N via (3);
    Set s ← ⌊αtN⌋ and λ ← λ^(0) − νt;
    Find s candidates cluster pairs via (10);
    Set n ← n − s and update ũ_i via (5)
    Train network (1) to update Θ via SGD and U via the loss in (5)
    Evaluate model perofmance P on the validation set
    Once P_t < P_{t−1} : train the network with {ω(i)}_{i=1}^N via the loss (1) in later iterations;
    if P_t > P_{t−1} then
        Best model = φ(·, Θ)
        P_best = P_t
    end if
end while
```

Find candidate clusters:

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

$$\begin{cases} \delta(\mathcal{C}_k, \mathcal{C}_\ell) = \min\limits_{i \in \mathcal{C}_k, j \in \mathcal{C}_\ell} \|\varphi(\boldsymbol{x}_i, \Theta) - \varphi(\boldsymbol{x}_j, \Theta)\|_2 \\ \Omega(\mathcal{C}_k) = \sum\limits_{i \in \mathcal{C}_k} (1 - \gamma_{ik}), \quad \Omega(\mathcal{C}_\ell) = \sum\limits_{j \in \mathcal{C}_\ell} (1 - \gamma_{j\ell}) \\ \gamma_{ij} = \dfrac{\exp(\boldsymbol{u}_j^\top \varphi(\boldsymbol{x}_i, \Theta)/\tau)}{\sum_{\ell=1}^n \exp(\boldsymbol{u}_\ell^\top \varphi(\boldsymbol{x}_i, \Theta)/\tau)} \end{cases}$$

1. Finding candidate clusters to be merged in self-paced way



Algorithm 1 SPCNet-SI

Input: Dataset $\mathcal{X} = \{\boldsymbol{x}_1, ...\boldsymbol{x}_N\}$, $\lambda^{(0)}$, ν, α, T, initial pseudo labels $\{\omega(i)\}_{i=1}^{N}$, and set $n \leftarrow N$ and $t \leftarrow 0$.

Output: $\{\omega(i)\}_{i=1}^{N}$.

Train the network with $\{\omega(i)\}_{i=1}^{N}$ via the loss in (1);
while $t \leq T$ **do**
    Set $t \leftarrow t + 1$;
    Compute $\{\omega(i)\}_{i=1}^{N}$ via (3);
    Set $s \leftarrow \lfloor \alpha t N \rfloor$ and $\lambda \leftarrow \lambda^{(0)} - \nu t$;
    Find s candidates cluster pairs via (10);
    Set $n \leftarrow n - s$ and update $\tilde{\boldsymbol{u}}_i$ via (5)
    Train network (1) to update Θ via SGD and $\mathcal{U}$ via the loss in (5)
    Evaluate model perofmance P on the validation set
    Once $P_t < P_{t-1}$: train the network with $\{\omega(i)\}_{i=1}^{N}$ via the loss (1) in later iterations;
    if $P_t > P_{t-1}$ **then**
        Best model $= \phi(\cdot, \Theta)$
        $P_{best} = P_t$
    end if
end while

Dynamic clustering:

a) Gradual Development Strategy:

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

b) Annealing Scheme:

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

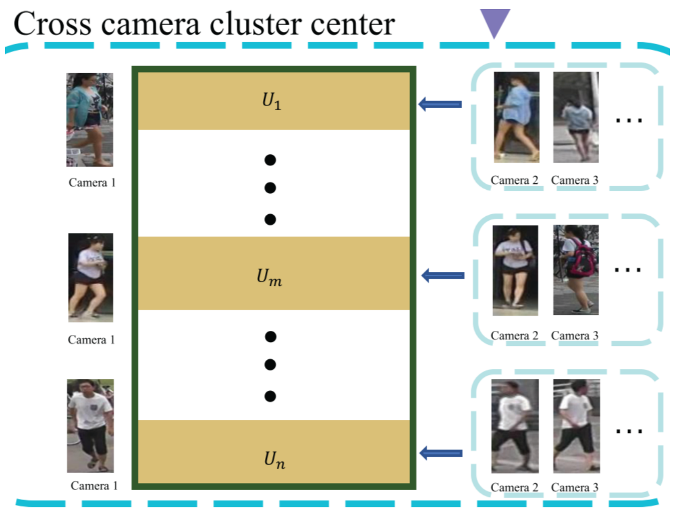
2. Update the cluster center with side information and train with cross-camera repelled loss

Algorithm 1 SPCNet-SI

Input: Dataset $\mathcal{X} = \{\boldsymbol{x}_1, \ldots \boldsymbol{x}_N\}$, $\lambda^{(0)}$, ν, α, T, initial pseudo labels $\{\omega(i)\}_{i=1}^{N}$, and set $n \leftarrow N$ and $t \leftarrow 0$.
Output: $\{\omega(i)\}_{i=1}^{N}$.

```
Train the network with {ω(i)}_{i=1}^N via the loss in (1);
while t ≤ T do
    Set t ← t + 1;
    Compute {ω(i)}_{i=1}^N via (3);
    Set s ← ⌊αtN⌋ and λ ← λ^(0) − νt;
    Find s candidates cluster pairs via (10);
    Set n ← n − s and update ũ_i via (5)
    Train network (1) to update Θ via SGD and U via the loss in (5)
     Evaluate model perofmance P on the validation set
     Once P_t < P_{t−1} : train the network with {ω(i)}_{i=1}^N via the loss (1) in later iterations;
    if P_t > P_{t−1} then
        Best model = φ(·, Θ)
        P_best = P_t
    end if
end while
```

1.Update the cluster center with side information：



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

2.cross-camera repelled loss：

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

Result

Ablation Study

method	Market-1501		DukeMTMC-ReID	
	mAP	rank-1	mAP	rank-1
Baseline*	30.1	64.1	22.7	43.0
Baseline	33.4	67.2	25.4	45.6
$s + \lambda^{(t)}$	34.9	67.7	26.6	45.8
$s^{(t)} + \lambda$	**35.1**	68.2	25.7	47.1
$s^{(t)} + \lambda^{(t)}$	34.9	**68.7**	**27.7**	**47.8**

Compare to the SOTA

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**

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

1.We designed a cross-camera repelled loss to exploits the camera side information and encourages to explore the association under different camera views.

2. We proposed a soft-label based assignment scheme in the bottom-up clustering.

3. We presented an effective dynamic strategy to regularize the cluster merging process to help select proper clusters to merge.

THANKS