Unsupervised Person Re-identification by Soft Multilabel Learnin

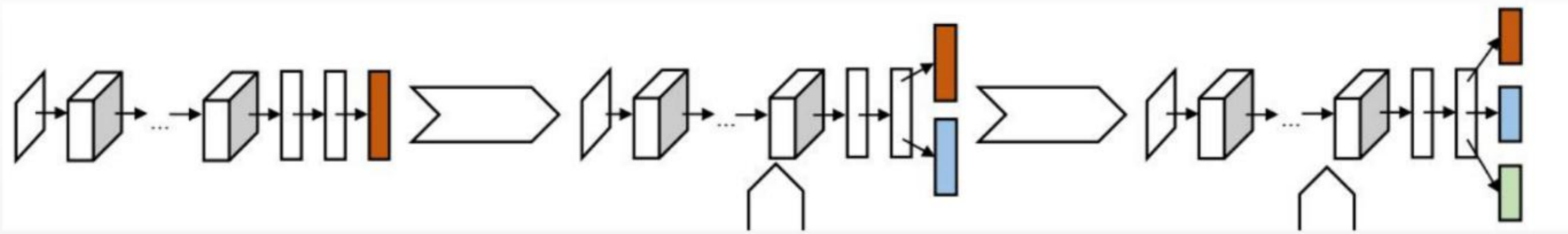
Hong-Xing Yu, Wei-Shi Zheng, Ancong Wu, Xiaowei Guo, Shaogang Gong, and Jian-Huang Lai

https://arxiv.org/pdf/1903.06325.pdf

Intelligent Information Fusion Research Group

引言

learning without forgetting



(1)旧模型在新的数据集上跑，得出这个时候的旧分类的相应的输出，保存下来当作接下来蒸馏训练中的老师

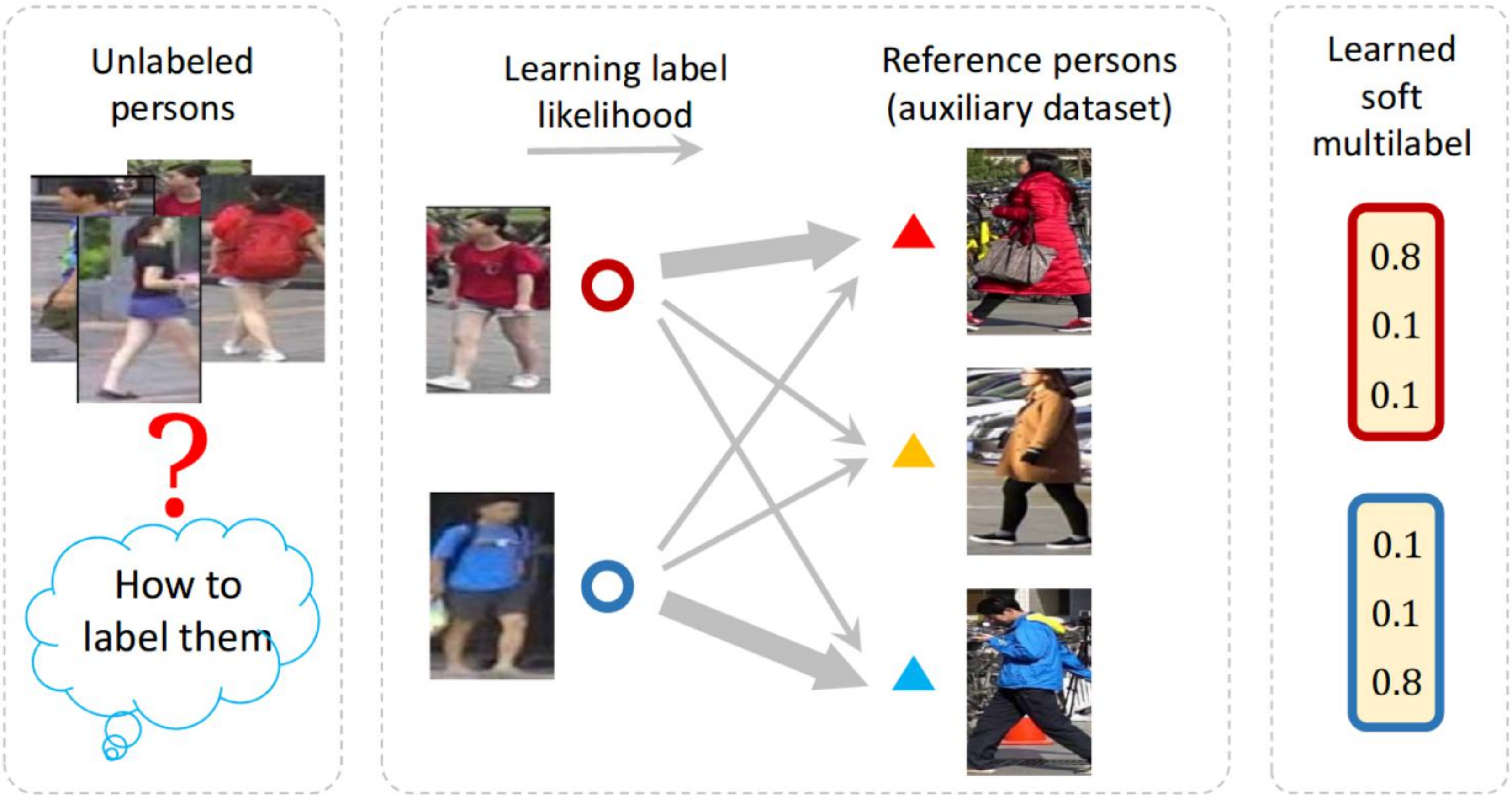
(2)对于新的数据集的新增加的分类采用one-hot“硬监督”的形式进行训练，作为第一个loss

(3)训练过程中旧分类的输出按照第一步中的结果以“软监督”的形式进行训练，作为第二个loss

软标签可以来辅助domain记忆，但在行人重识别领域，更多的是domain迁移问题。

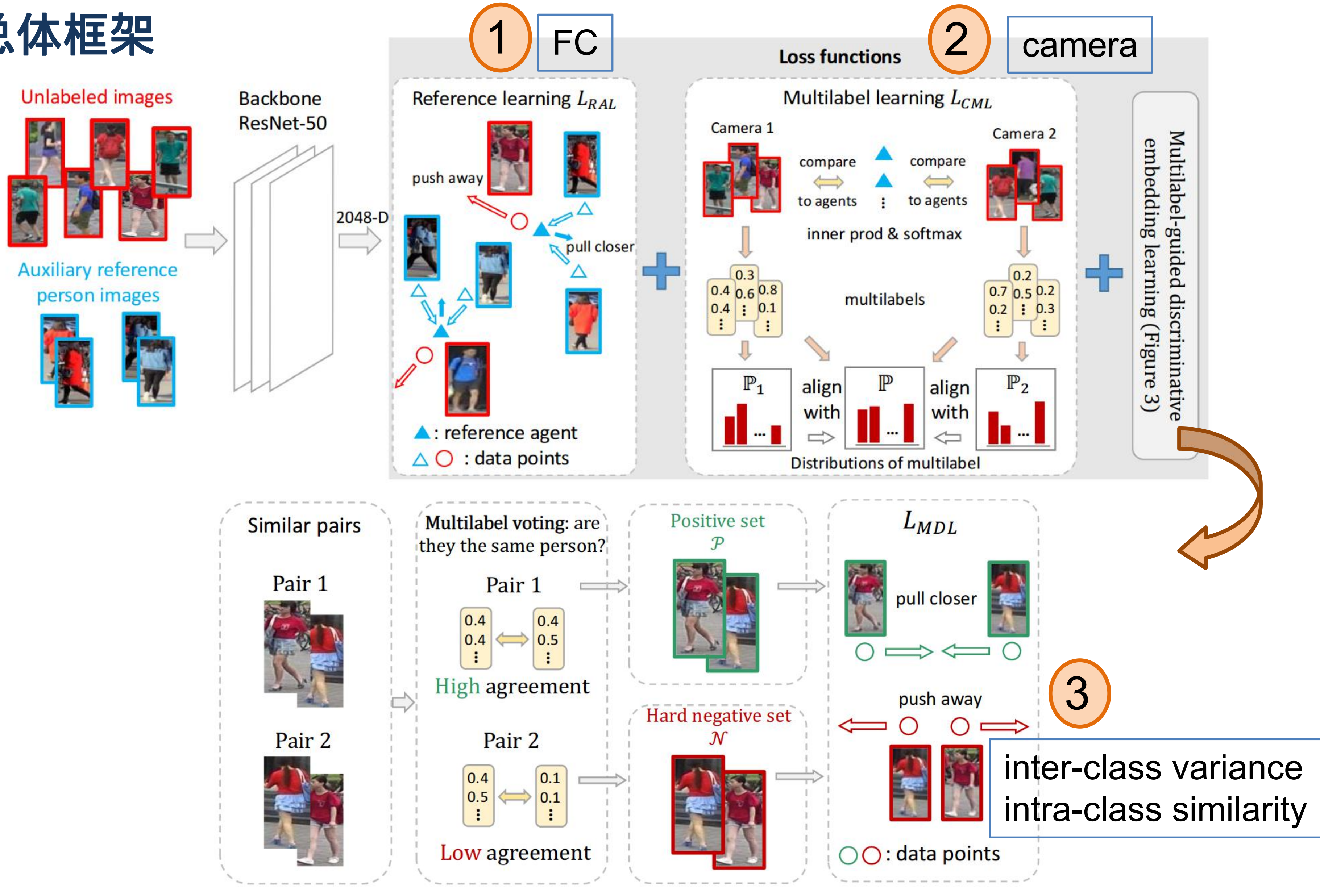
那么在无监督情况下，没有硬监督，只有软标签的时候，怎么利用软标签提高跨domain效果呢？

引言

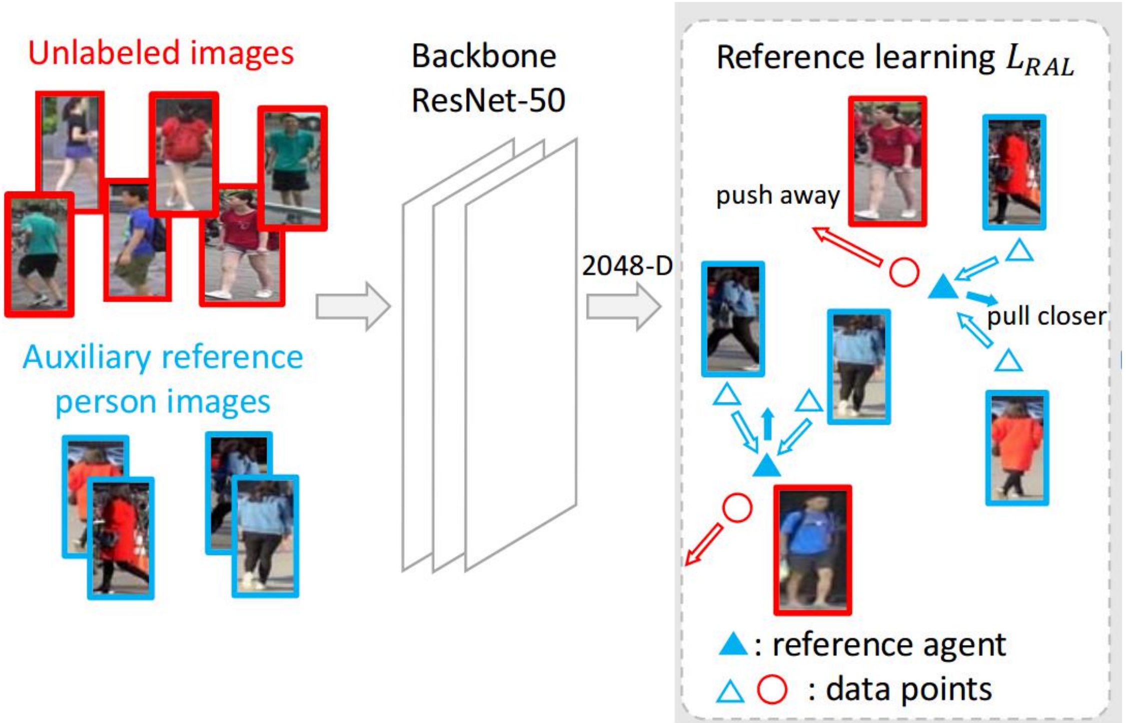


Unlabeled persons
How to label them
Learning label likelihood
Reference persons (auxiliary dataset)
Learned soft multilabel
0.8
0.1
0.1
0.1
0.1
0.8

总体框架


1
FC
2
camera
Loss functions
Unlabeled images
Auxiliary reference person images
Backbone ResNet-50
2048-D
Reference learning L_{RAL}
push away
pull closer
▲: reference agent
△ ○ : data points
Multilabel learning L_{CML}
Camera 1
Camera 2
compare to agents
inner prod & softmax
multilabels
align with
Distributions of multilabel
Multilabel-guided discriminative embedding learning (Figure 3)
Similar pairs
Pair 1
Pair 2
Multilabel voting: are they the same person?
High agreement
Low agreement
Positive set $\mathcal{P}$
Hard negative set $\mathcal{N}$
L_{MDL}
pull closer
push away
○○: data points
3
inter-class variance
intra-class similarity

1

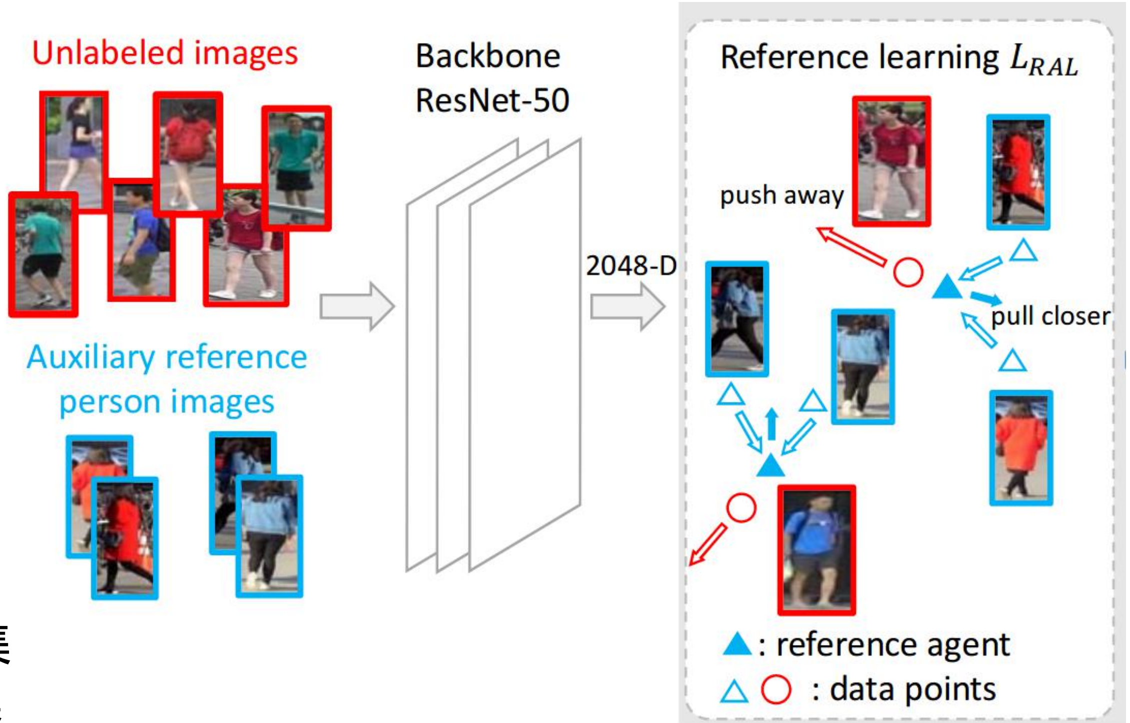


$$L_{AL} = \Sigma_k - \log l(f(z_k), \{a_i\})^{(w_k)} = \Sigma_k - \log \frac{\exp(a_{w_k}^{\mathrm{T}} f(z_k))}{\Sigma_j \exp(a_j^{\mathrm{T}} f(z_k))} \tag{7}$$

FC

a: 2048 * w_k, w_k 个人
f( ): 2048*1*1

1



z_k: 辅助数据集
x_j: 目标数据集

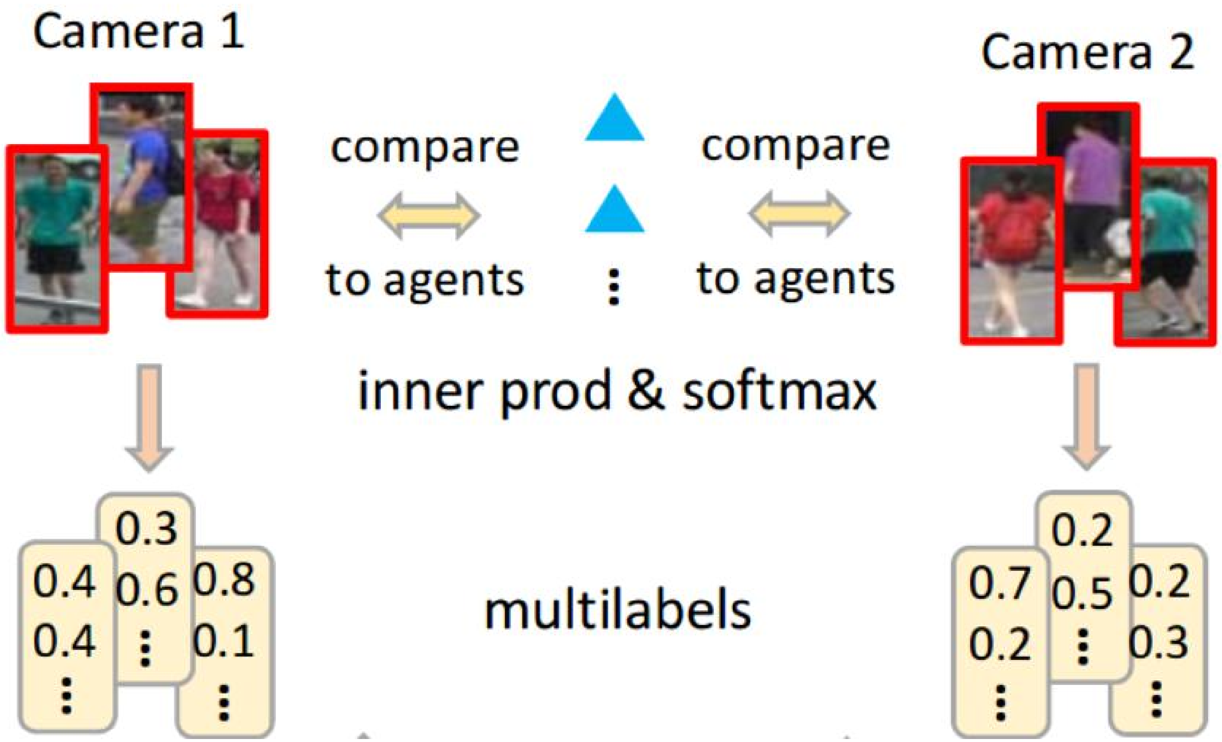
$$L_{RJ} = \Sigma_i \Sigma_{j \in \mathcal{M}_i} \Sigma_{k:w_k=i} [m - \|a_i - f(x_j)\|_2^2]_+ + \|a_i - f(z_k)\|_2^2 \tag{8}$$

$$\mathcal{M}_i = \{j \| \|a_i - f(x_j)\|_2^2 < m\}$$

center loss

$$\boxed{L_{RAL} = L_{AL} + \beta L_{RJ}}$$

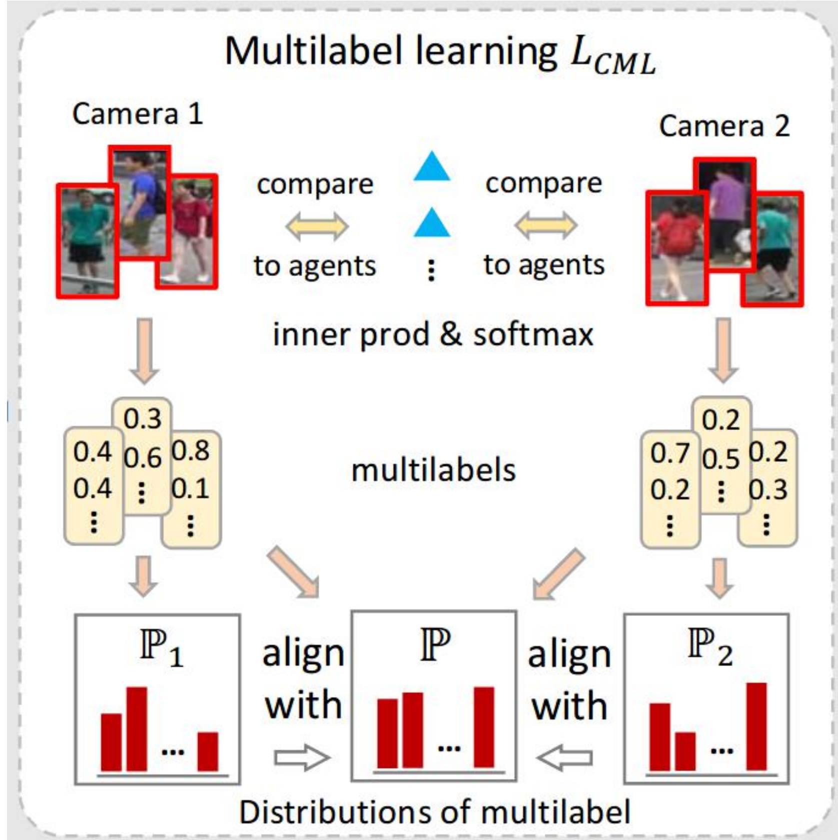
2



$$y^{(k)} = l(f(x), \{a_i\}_{i=1}^{N_p})^{(k)} = \frac{\exp(a_k^{\mathrm{T}} f(x))}{\Sigma_i \exp(a_i^{\mathrm{T}} f(x))} \tag{1}$$

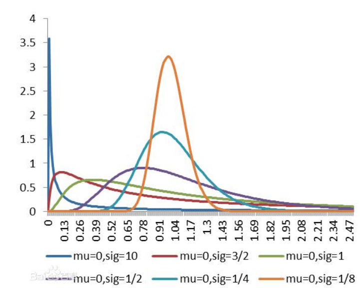
FC

2



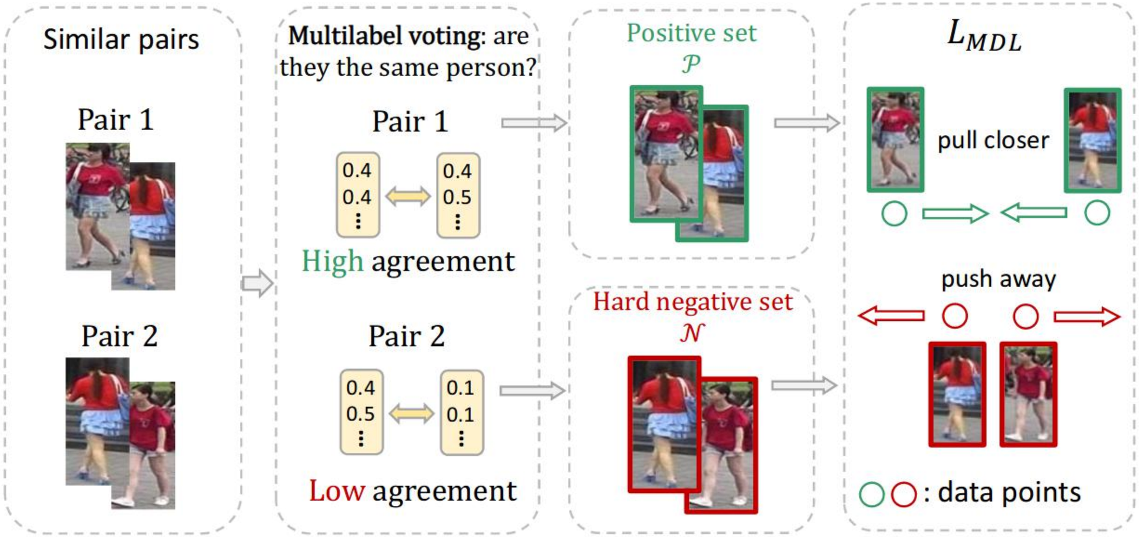
$$L_{CML} = \Sigma_v d(\mathbb{P}_v(y), \mathbb{P}(y))^2$$

数值为对数正态分布



$$L_{CML} = \Sigma_v ||\mu_v - \mu||_2^2 + ||\sigma_v - \sigma||_2^2$$

3

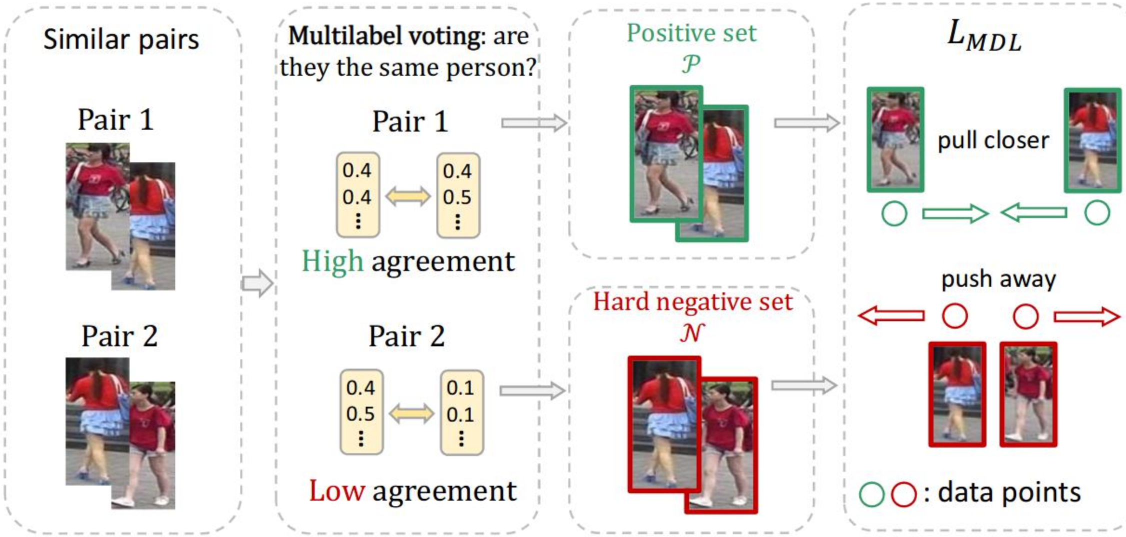


原始feature相似，**软多标签**（类似上下文）也相似，为**正样本：P**
原始feature相似，但是**软多标签**（类似上下文）不相似，为**hard负样本：N**

feature是否相似: $f(x_i)^{\mathrm{T}}f(x_j)$

软多标签（类似上下文）是否相似: $A(y_i, y_j) = y_i \wedge y_j = \Sigma_k \min(y_i^{(k)}, y_j^{(k)}) = 1 - \frac{||y_i - y_j||_1}{2}$

3



原始feature相似，**软多标签**（类似上下文）也相似，为**正样本：P**

原始feature相似，但是**软多标签**（类似上下文）不相似，为**hard负样本：N**

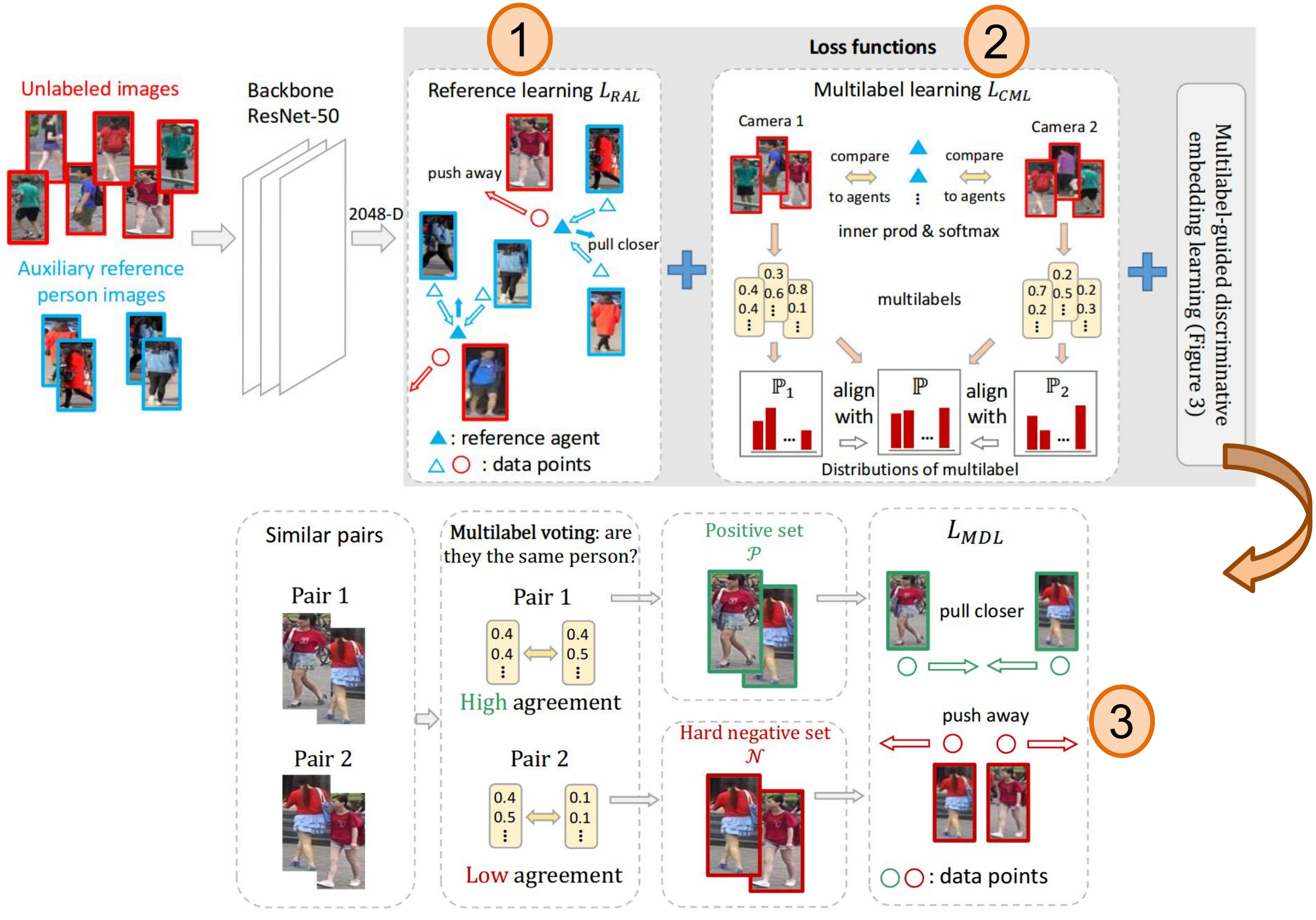
$$L_{MDL} = -\log \frac{\overline{P}}{\overline{P} + \overline{N}},$$

Hard Triplet Loss

where

$$\overline{P} = \frac{1}{|\mathcal{P}|}\Sigma_{(i,j)\in\mathcal{P}} \exp(-||f(z_i) - f(z_j)||_2^2),$$

$$\overline{N} = \frac{1}{|\mathcal{N}|}\Sigma_{(k,l)\in\mathcal{N}} \exp(-||f(z_k) - f(z_l)||_2^2).$$



$$L_{MAR} = L_{MDL} + \lambda_1 L_{CML} + \lambda_2 L_{RAL}$$

3 2 1

总体框架

实验

Table 3. Ablation study. Please refer to the text in Sec. 4.4.

	Methods	Market-1501 rank-1	rank-5	rank-10	mAP
softmax	Pretrained	46.2	64.4	71.3	24.6
	Baseline	44.4	62.5	69.8	21.5
无软多标签在不同camera分布接近	MAR w/o L_{CML}	60.0	75.9	81.9	34.6
	MAR w/o L_{CML}&L_{RAL}	53.9	71.5	77.7	28.2
	MAR w/o L_{RAL}	59.2	76.4	82.3	30.8
	MAR	**67.7**	**81.9**	**87.3**	**40.0**

	Methods	DukeMTMC-reID rank-1	rank-5	rank-10	mAP
	Pretrained	43.1	59.2	65.7	28.8
	Baseline	50.0	66.4	71.7	31.7
	MAR w/o L_{CML}	63.2	77.2	82.5	44.9
无软多标签在不同camera分布接近 和 无源域数据+辅助数据一起训练	MAR w/o L_{CML}&L_{RAL}	60.1	73.0	78.4	40.4
fc + center loss	MAR w/o L_{RAL}	57.9	72.6	77.8	37.1
	MAR	**67.1**	**79.8**	**84.2**	**48.0**

Baseline 包括1，2，但是3最后选正样本和hard负样本的时候，**没用软多标签**。而是原始feature相似，那么用**阈值**选出前多少个为正样本，后多少个为hard负样本

Table 1. Comparison to the state-of-the-art unsupervised results in the Market-1501 dataset. **Red** indicates the best and **Blue** the second best. Measured by %.

Methods	Reference	Market-1501		
		rank-1	rank-5	mAP
LOMO [20]	CVPR'15	27.2	41.6	8.0
BoW [58]	ICCV'15	35.8	52.4	14.8
DIC [16]	BMVC'15	50.2	68.8	22.7
ISR [21]	TPAMI'15	40.3	62.2	14.3
UDML [29]	CVPR'16	34.5	52.6	12.4
CAMEL [52]	ICCV'17	54.5	73.1	26.3
PUL [8]	ToMM'18	45.5	60.7	20.5
TJ-AIDL [48]	CVPR'18	58.2	74.8	26.5
PTGAN [50]	CVPR'18	38.6	57.3	15.7
SPGAN [7]	CVPR'18	51.5	70.1	27.1
HHL [62]	ECCV'18	**62.2**	**78.8**	**31.4**
MAR	This work	**67.7**	**81.9**	**40.0**

Table 2. Comparison to the state-of-the-art unsupervised results in the DukeMTMC-reID dataset. Measured by %.

Methods	Reference	DukeMTMC-reID		
		rank-1	rank-5	mAP
LOMO [20]	CVPR'15	12.3	21.3	4.8
BoW [58]	ICCV'15	17.1	28.8	8.3
UDML [29]	CVPR'16	18.5	31.4	7.3
CAMEL [52]	ICCV'17	40.3	57.6	19.8
PUL [8]	ToMM'18	30.0	43.4	16.4
TJ-AIDL [48]	CVPR'18	44.3	59.6	23.0
PTGAN [50]	CVPR'18	27.4	43.6	13.5
SPGAN [7]	CVPR'18	41.1	56.6	22.3
HHL [62]	ECCV'18	**46.9**	**61.0**	**27.2**
MAR	This work	**67.1**	**79.8**	**48.0**

辅助数据集与目标数据集的相似性，看来也是有影响的。

Visual results

Body shape (median-thin) | Clothes color | w/ backpack

w/ backpack | Trousers color | Clothes color