Rethinking ReID： Multi-Feature Fusion Person Re-identification Based On Orientation Constraints

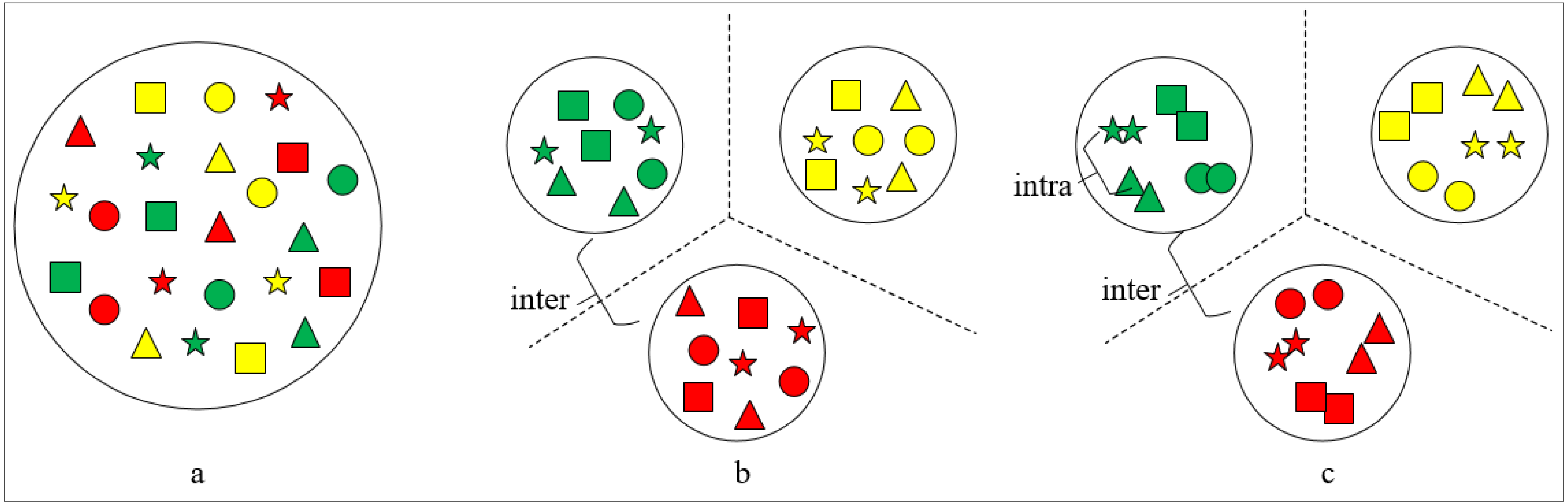
Mingjing Ai, Guozhi Shan, Bo Liu and Tianyang Liu

gzshan1996@buaa.edu.cn

Introduction

One major issue downgrading the ReID model performance lies in that various subjects in the same body orientations look too similar to distinguish by the model, while the same subject viewed in different orientations looks rather different.

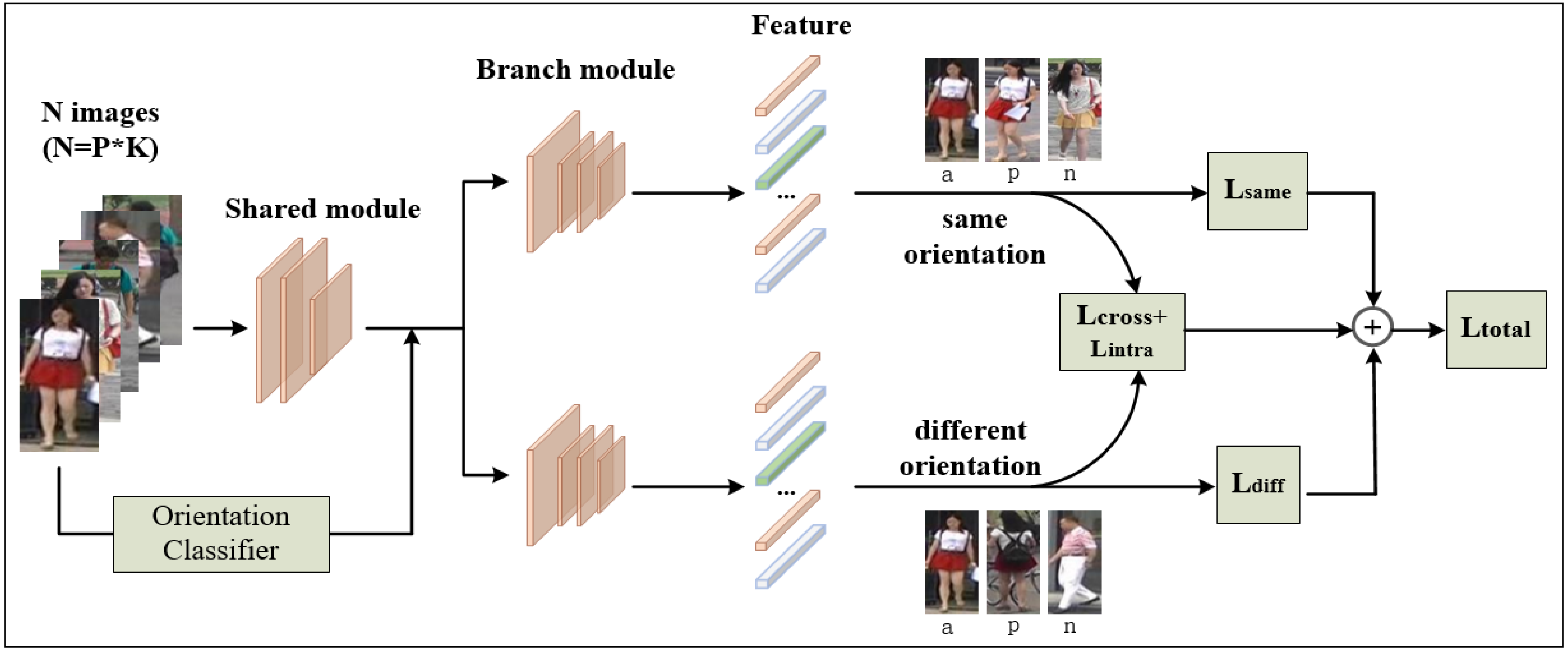
Introduction



* Different colors represent different people, and shapes represent orientations.

As shown in Fig(c), our network can not only distinguish different people, but also form perspective-level clusters, so that the model can preferentially identify the person with same orientation.

Method



The pipeline of our proposed network model (OCMFPR).

Method

A. Orientation-based sample sampling strategy

We consider the orientation of every pedestrian and propose a strategy for sampling difficult samples based on orientation. We also randomly select P persons for each training batch. But in the selection of K images of each person, it is not simply random, but to **ensure that there are both samples with the same orientation and samples with different orientations in the K images.**

$$L_{triHard} = \frac{1}{P*K} \sum_{a \in batch} [\max d(a,p) - \min d(a,n) + \alpha]_{+}$$

B. Global and Local Feature Fusion

We chose the **AlignedReID method** because the method obtains body parts through horizontal division, and at the same time, the **dynamic programming method** based on the shortest path realizes automatic alignment.

In the same orientation branches, negative sample pairs in the same direction may be very similar. Local features can capture some detailed features, so as **to better distinguish positive and negative samples**.

Method

C. Training strategy

(1) Same-orientation branch

$$L_{triSame} = \frac{1}{P*K}\sum_{a\in batch}[\max d(a,ps) - \min d(a,ns) + \alpha]_+$$

$$L_{ceSame} = -\sum_{i=1}^{N}\log\frac{e^{f_i(label\_i)}}{\sum_{k=1}^{M*T} e^{f_k(label\_k)}}$$

$$L_{center} = \frac{1}{2}\sum_{i=1}^{N}\| f_i - C_{label\_i}\|_2^2$$

$$\boxed{L_{same} = L_{triSame} + L_{ceSame} + \lambda L_{center}}$$

(2) Different-orientation branch

$$L_{triDiff} = \frac{1}{P*K}\sum_{a\in batch}[\max d(a,pd) - \min d(a,nd) + \beta]_+$$

$$L_{ceDiff} = -\sum_{i=1}^{N}\log\frac{e^{f_i(id\_i)}}{\sum_{k=1}^{M} e^{f_k(id\_k)}}$$

$$\boxed{L_{diff} = L_{triDiff} + L_{ceDiff} + \lambda L_{center}}$$

Method

C. Training strategy

(3) Cross-constrained training

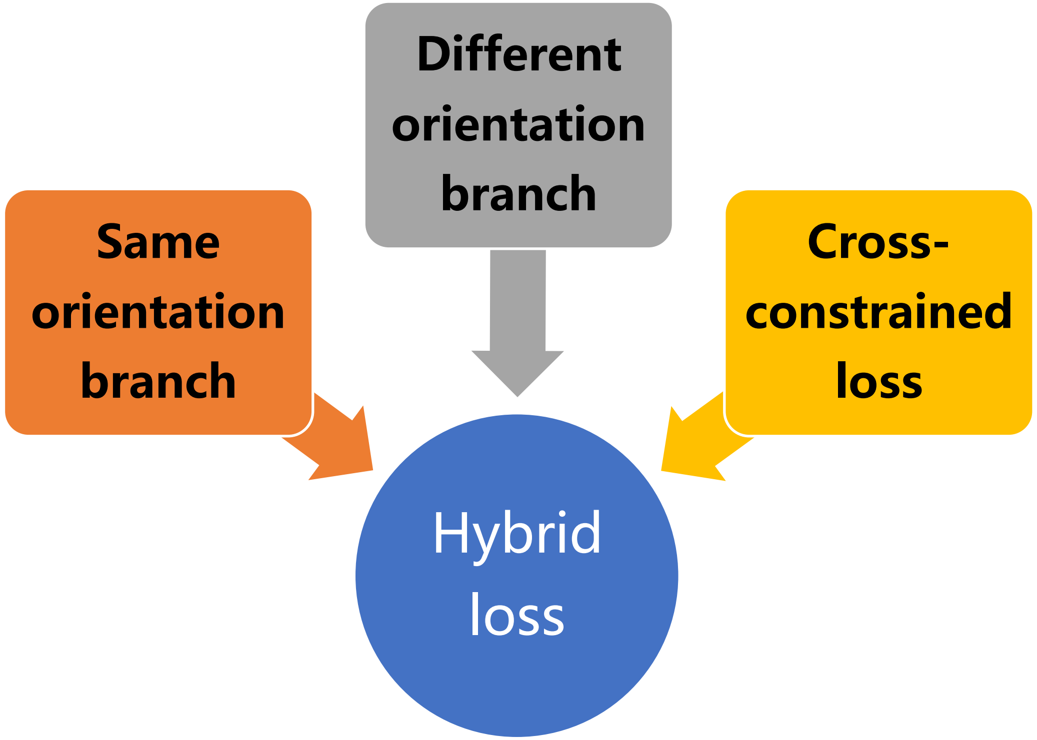
Cross constraint:

$$L_{cross} = \frac{1}{P*K}\sum_{a \in batch}[\max d(a, pd) - \min d(a, ns) + \theta]_+$$

Intra class constraint:

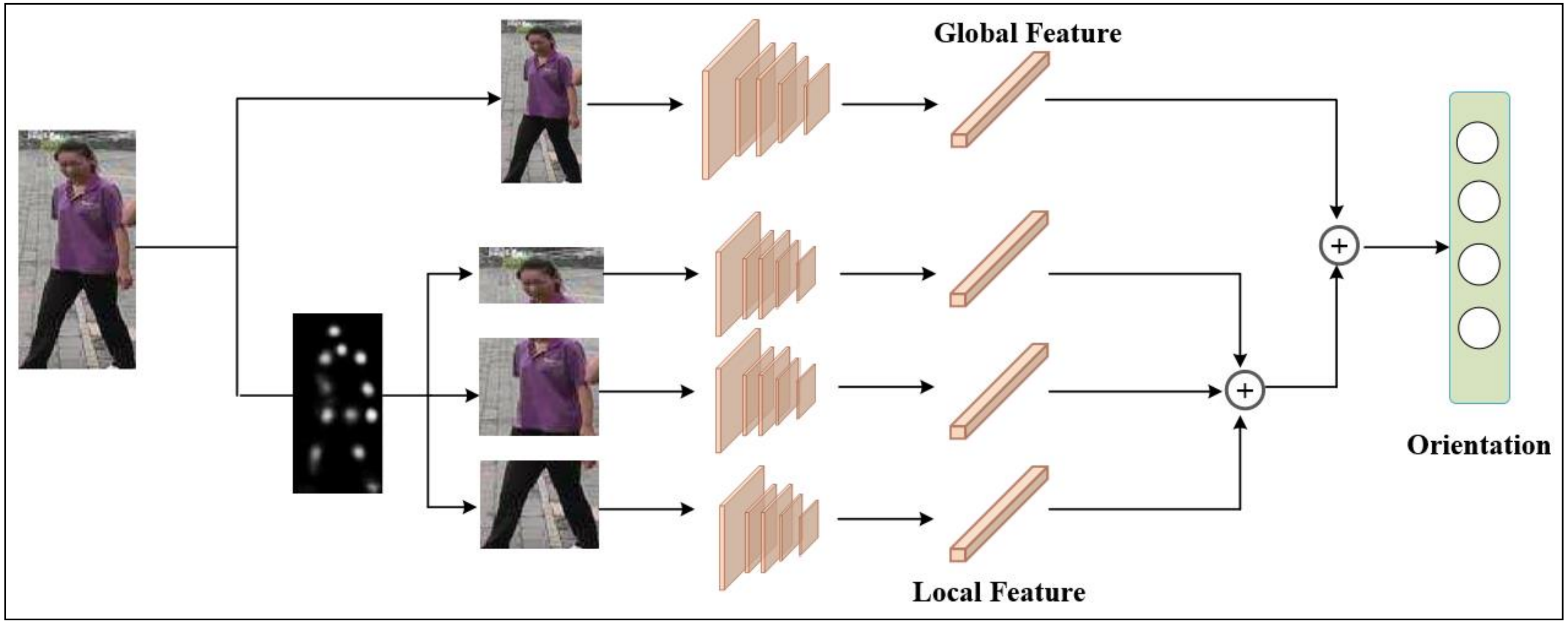
$$L_{intra} = \frac{1}{P*K}\sum_{a \in batch}[\max d(a, ps) - \min d(a, pd) + \delta]_+$$

$$\boxed{L_{Total} = L_{same} + L_{diff} + L_{cross} + \mu L_{intra}}$$



Method

D. The Design of Orientation Classifier



Orientation classifier based on global and three body parts.

Method

D. The Design of Orientation Classifier

1). **Classify based on the relative positions of the pose joints. (OpenPose)**

Use the left and right shoulder points to form a vector (from left to right).Then calculate the clockwise angle between this vector and the vertical direction. According to this angle range, we can determine the orientation of each person.

2). **Orientation classification based on global features**. (use ResNet50 as the backbone)

3). **Global and local features fusion classification.**

Methods	The Results of classification	
	Method introduction	*Accuracy (%)*
1	Based on pose joints	82.07
2	Global Features	87.33
3	**Global and Local Features**	**89.03**

* RAP dataset

* The red dot represents the left shoulder, and the green dot represents the right shoulder.

Experiments Data comparison

Methods	Market-1501		
	Rank-1(%)	*Rank-5(%)*	*mAP(%)*
PCB	92.3	97.2	77.4
AlignedReID	91.8	97.1	79.3
PIE	87.33	95.56	69.25
GLAD	89.9	-	73.9
Spindle	76.9	91.5	-
HA-CNN	91.2	-	75.7
TriHard	86.67	93.38	81.07
HPM	94.2	97.5	82.7
PGR	93.87	97.74	77.21
OSCNN	83.9	-	73.5
OCMFPR(ours)	**94.71**	**98.06**	**84.11**
OCMFPR(ours+RR)	**94.87**	**98.30**	**92.71**

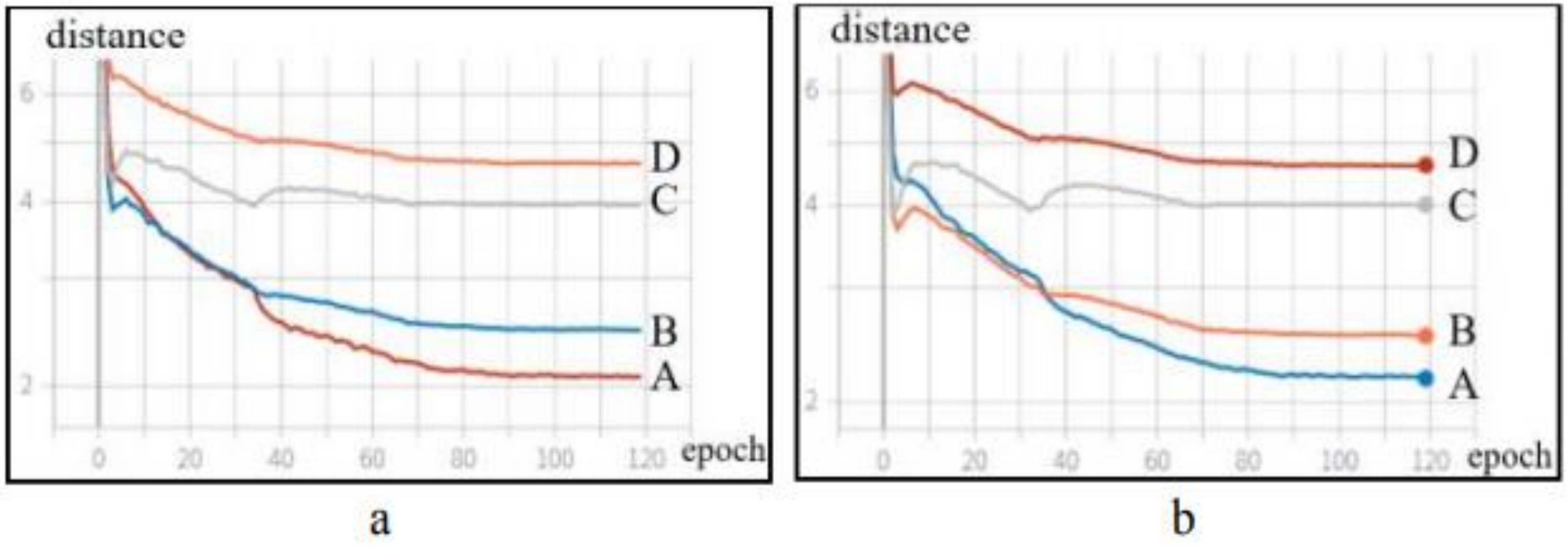
Methods	DukeMTMC-ReID		
	Rank-1(%)	*Rank-5(%)*	*mAP(%)*
PCB	81.7	89.7	66.1
AlignedReID	81.2	-	67.4
PIE	80.84	88.30	64.09
HA-CNN	80.5	-	63.8
HPM	86.6	-	74.3
PGR	83.63	91.66	65.98
SVDNet	76.7	-	56.8
OCMFPR(ours)	**87.31**	**93.54**	**73.20**
OCMFPR(ours+RR)	**90.63**	**94.25**	**87.67**

Experiments Ablation Study

Methods	Ablation Study		
	Rank-1(%)	***Rank-5(%)***	***mAP(%)***
BaseLine	91.5	97.02	81.71
Sampling by orientation	92.2	97.15	81.76
Same orientation branch	89.19	95.26	71.64
Different orientation branch	90.11	95.37	74.52
OCMFPR(4)	93.54	97.54	82.74
OCMFPR(4+Local)	94.28	97.61	83.66
OCMFPR(3+Local)	**94.71**	**98.06**	**84.11**

* Market1501 Dataset

Experiments Training effect



* a is the result in Market1501 Dataset, b is the result in DukeMTMC-ReID dataset.

In the picture above, A represents the positive sample pair (is the same person) with the same body orientation, B represents the positive sample pair with different body orientations, C represents the negative sample pair with same orientation and D means the negative sample pair with different orientations.

Experiments Visual display

Thanks

gzshan1996@buaa.edu.cn