Paper ID： 4304

上海市多维度信息处理重点实验室
Shanghai Key Laboratory of Multidimensional Information Processing

PROGRESSIVE MULTI-STAGE FEATURE MIX FOR PERSON RE-IDENTIFICATION

Yan Zhang [1], Binyu He [1], Li Sun [1,2], and Qingli Li [1]

[1]Shanghai Key Laboratory of Multidimensional Information Processing,
East China Normal University, 200241 Shanghai, China
[2]State Key Laboratory of Mathematical Engineering and Advanced Computing, 214125 Wuxi, China

1. Introduction

1.1. Motivation

- CNN suffers from only focusing on small local regions
- Taking batch drop at the intermediate feature level expands the highly-responded areas under no rules

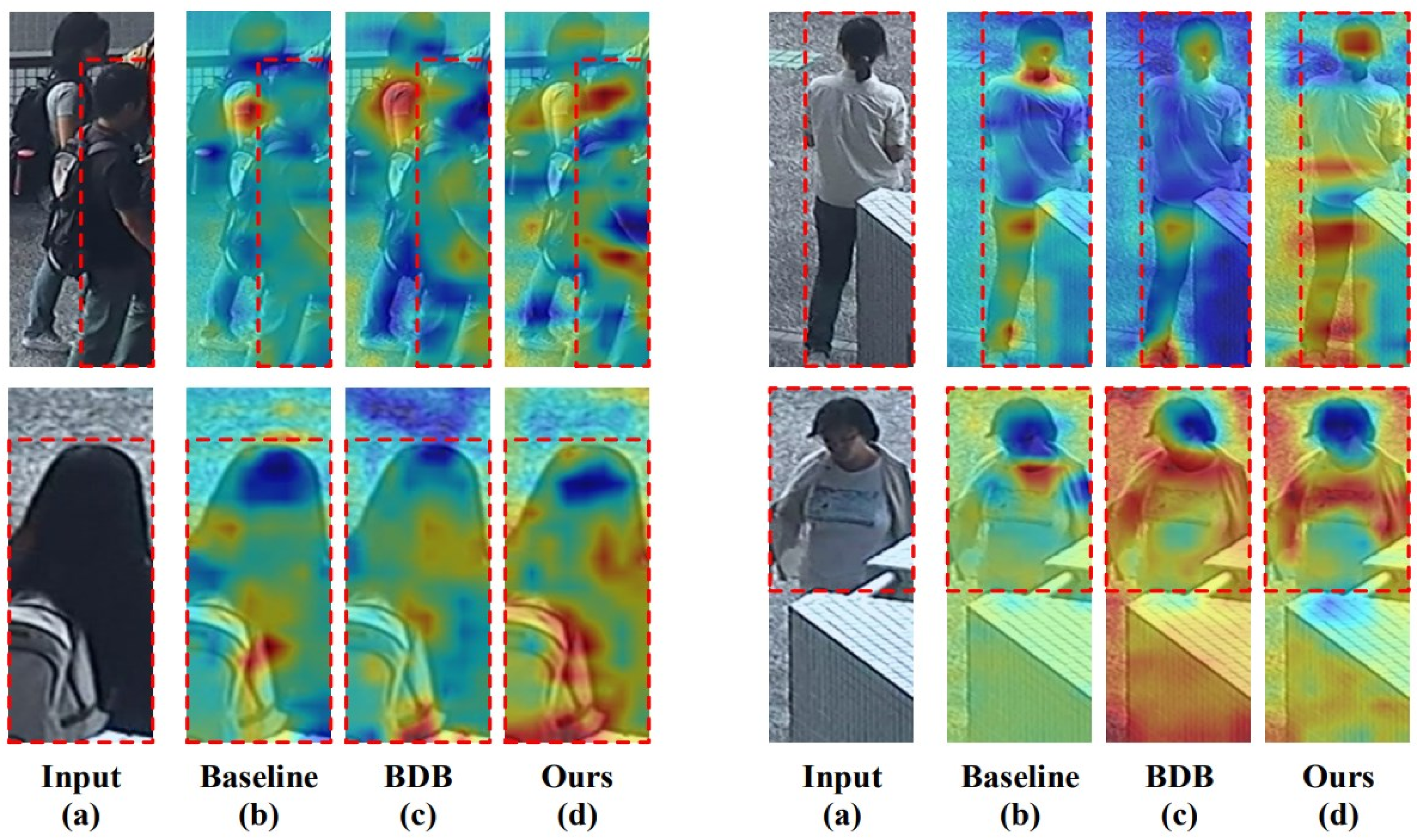

Fig. 1. The visualizations of the gradient based class activation map of baseline (Resnet50+GAP), BDB and our approach. Red boxes indicate the target attention area.

1.2. Contribution

- We design a **Progressive Multi-stage feature Mix (PMM)** to **suppress the most salient features** for the current classifier, and force the head in the later stage to **find other clues**.
- We propose an **attentive Hard-Mix** feature augmentation method, which **synthesizes the harder samples** with mixing the negative pairs.

2. Our Approach

2.1. PMM framework

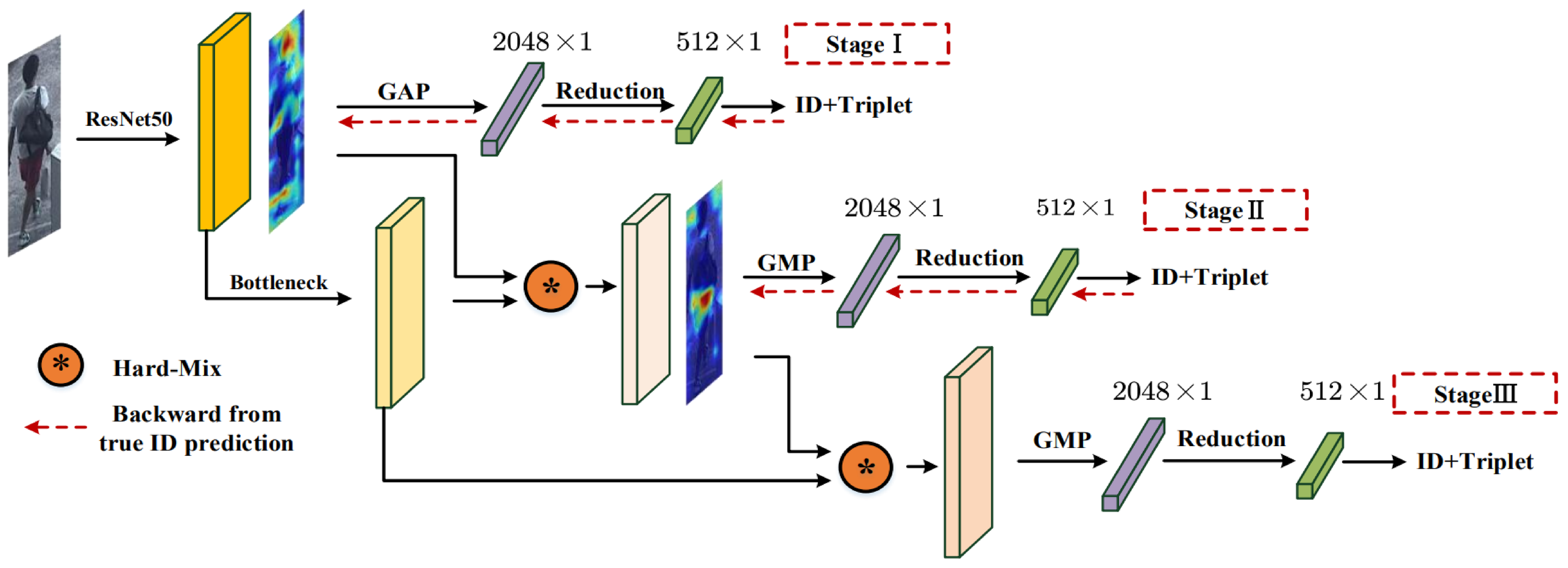

Fig. 2.Three stages are appended after the backbone, and they are supervised by the **same constrains**. Each stage can attain their own Grad-CAM images in the training period, which can be then used to **guide the feature Hard-Mix in the next stage**. In the testing, the green features from different stages are **concatenated together as the final representation**.

Code: https://github.com/crazydemo/Progressive-Multi-stage-Feature-Mix-for-Person-Re-Identification

2.2. Attentive Hard Mix

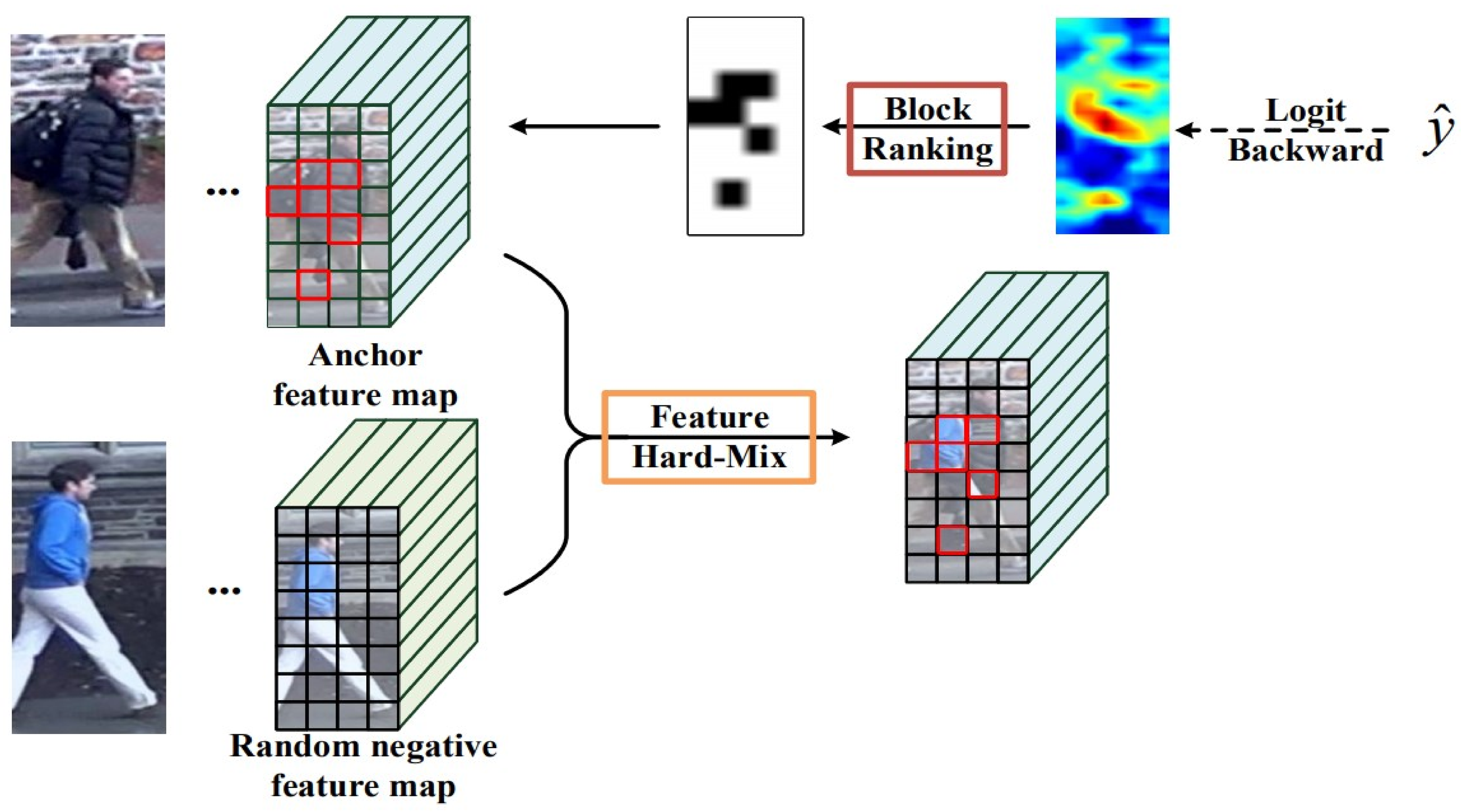

Fig. 3. There are two steps, i.e. **the block ranking step** and the **feature Hard-Mix** step. Firstly, taking a Grad-CAM image, corresponding to the expected class, the block ranking turns it into a binary mask (0 for the highlighted regions in Grad-CAM and 1 for the rest). Then, according to the binary mask, the **highlighted regions of an anchor** feature map will be **replaced by** the corresponding region features from **a random negative** feature map in the feature Hard-Mix step.

3. Experiments

Table 1. Comparison results with the state-of-the-art methods on the classic reID datasets.

Methods	CUHK03		DukeMTMC		Market-1501	
	Rank-1	mAP	Rank-1	mAP	Rank-1	mAP
AOS	54.6	56.1	79.2	62.1	91.3	78.3
PCB+RPP	62.8	56.7	83.3	69.2	93.8	81.6
CAMA	66.6	64.2	85.8	72.9	**94.7**	84.5
BDB*	73.5	69.8	87.1	74.5	94.0	84.9
Ours	**76.3**	**73.0**	**88.0**	**74.6**	94.1	**85.2**

Table 2. Comparison results between different feature augmentation methods.

Methods	CUHK03		DukeMTMC		Market-1501	
	Rank-1	mAP	Rank-1	mAP	Rank-1	mAP
CutOut	73.5	69.8	87.1	74.5	94.1	84.7
CutMix	74.3	69.9	85.8	72.0	92.8	82.4
A-CutMix	65.6	62.8	83.8	68.7	91.6	78.1
A-HardMix	**76.3**	**73.0**	**88.0**	**74.6**	**94.1**	**85.2**

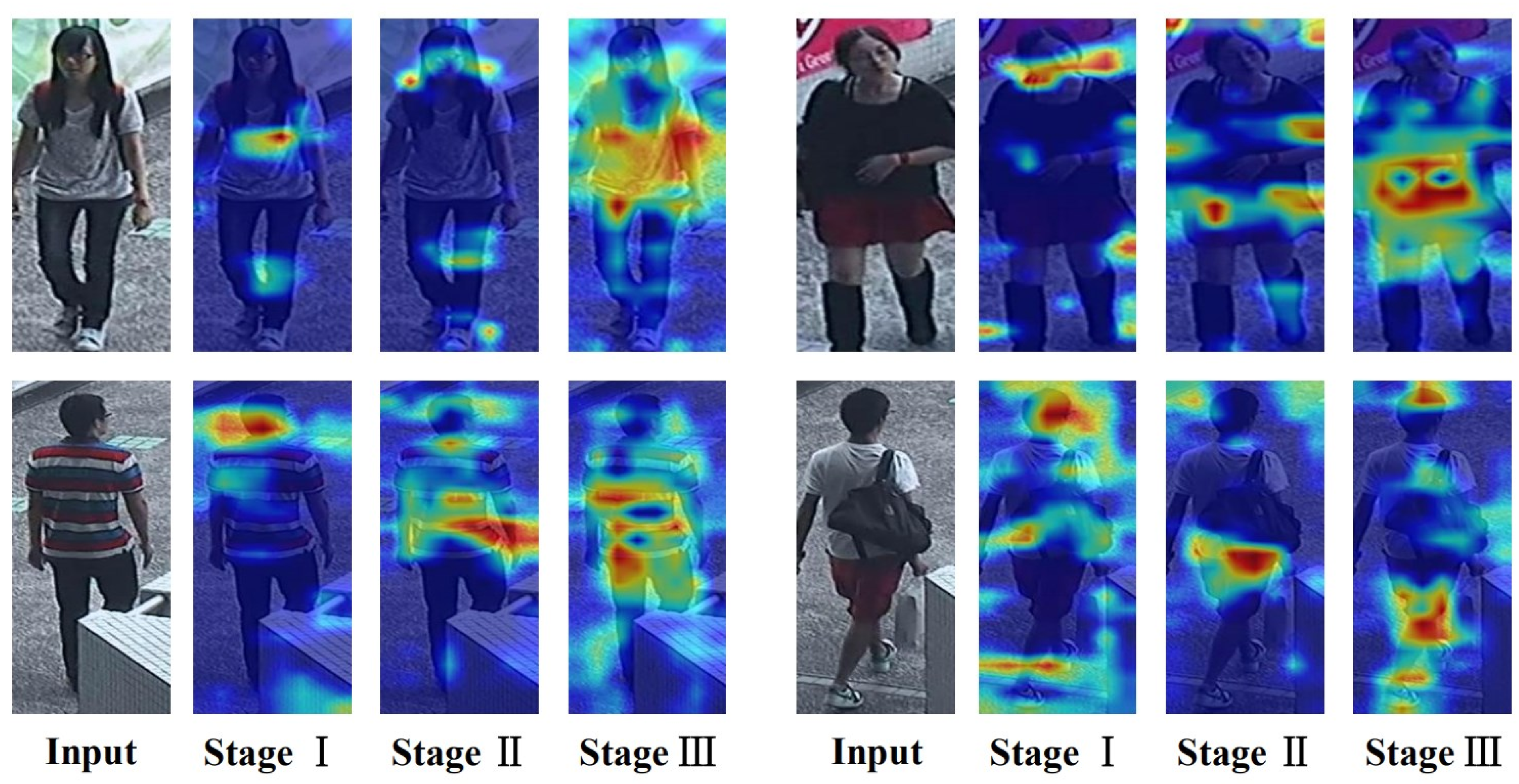

Fig. 4. The Grad-CAM image g from each stage during test (gradients are back propagated from the predicted class). The visualization is conducted on CUHK03-Detected.