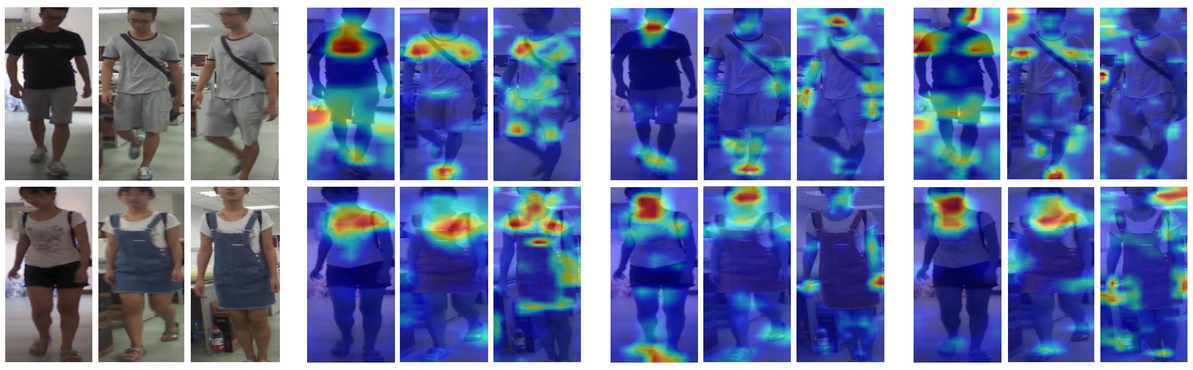
Learning Clothes-irrelevant Cues for Clothes-Changing Person Re-identification

Jingyi Mu, Yong Li , Jun Li and Jian Yang
PCA Lab,
School of Computer Science and Engineering, Nanjing University of Science and Technology.

Introduction

- We propose a Clothes-Relevant information Erasure (CRE) module to erase the clothes relevant information on the original images and put the generated images into the network for training so that the model can learn more biological features which are clothes-irrelevant.
- We further present a Body Shape-Guided Attention (BSGA) module into the network so as to enable the model to learn richer and more discriminative features.

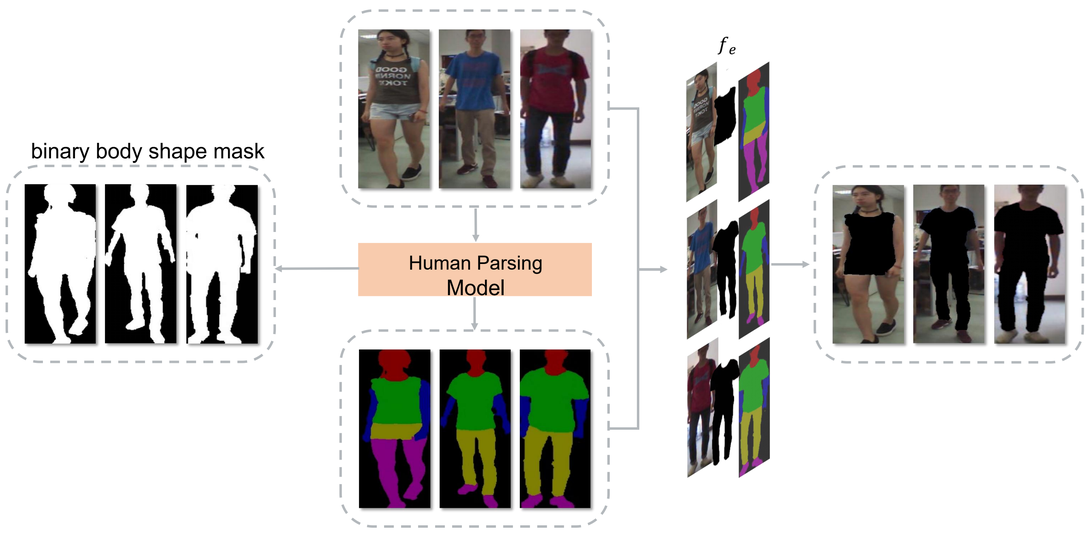
Motivation



(a) (b) (c) (d)

- The visualizations of the activation maps learned by ResNet-50 in (b), our CRE module in (c) and our both CRE and BSGA modules in (d). (a) shows six original images of two persons. We can see that (b) always highlights some clothes-relevant features, (c) highlights head and limbs but occasionally some background noise, while (d) highlights more discriminative features, e.g., face, and body shape.

Method



The structure of the Clothes-Relevant information Erasure (CRE) module.

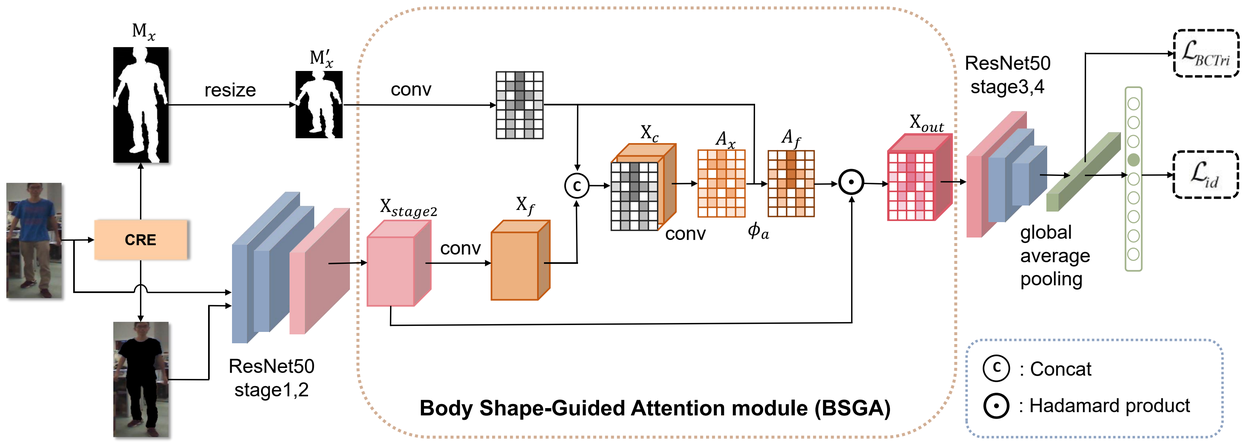
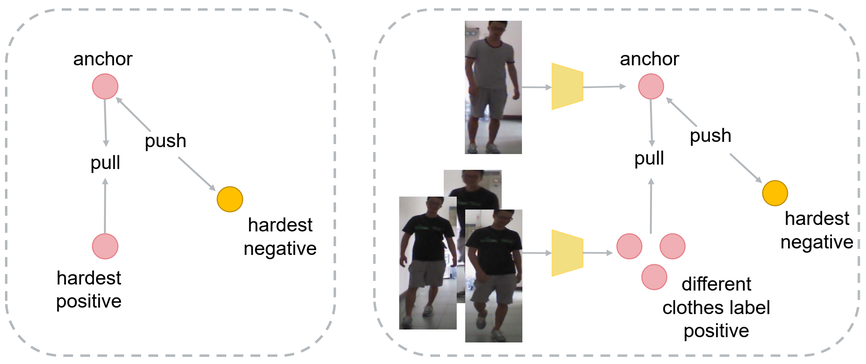


Illustration of our framework.



(a) (b)

Two triplet mining methods. (a) is Batch Hard mining. (b) is the Batch Cross-clothes mining we propose in this paper.

$$\mathcal{L}_{BCTti} = \sum_{n=1}^{N} \max(m + d(x_i, x_j^n) - d(x_i, x_k), 0)$$

Experiments

method	ref	general(all cams)		SC(cam2&cam3)		CC(cam3&cam4)	
		top-1	mAP	top-1	mAP	top-1	mAP
MDLA [36]	ICCV2017	88.9	76.8	93.4	93.9	59.2	60.8
PCB [41]	ECCV2018	87.7	74.6	94.7	94.3	62.0	62.2
Part-aligned [39]	ECCV2018	90.5	79.7	93.9	93.4	69.4	67.3
TransReID [13]	ICCV2021	90.5	80.1	95.1	94.5	70.0	71.8
FSAM [15]	CVPR2021	-	-	94.7	94.8	78.6	78.9
3DSL [2]	CVPR2021	-	-	-	-	79.9	81.2
PS [38]	SPL2021	93.1	84.9	94.7	92.9	82.4	80.3
CAL [11]	CVPR2022	92.9	87.2	**95.1**	**95.3**	81.4	81.7
ours w/o BSGA		94.2	85.7	94.9	93.7	83.5	81.7
ours		**94.4**	**88.2**	94.9	94.4	**84.5**	**84.3**

VC-Clothes

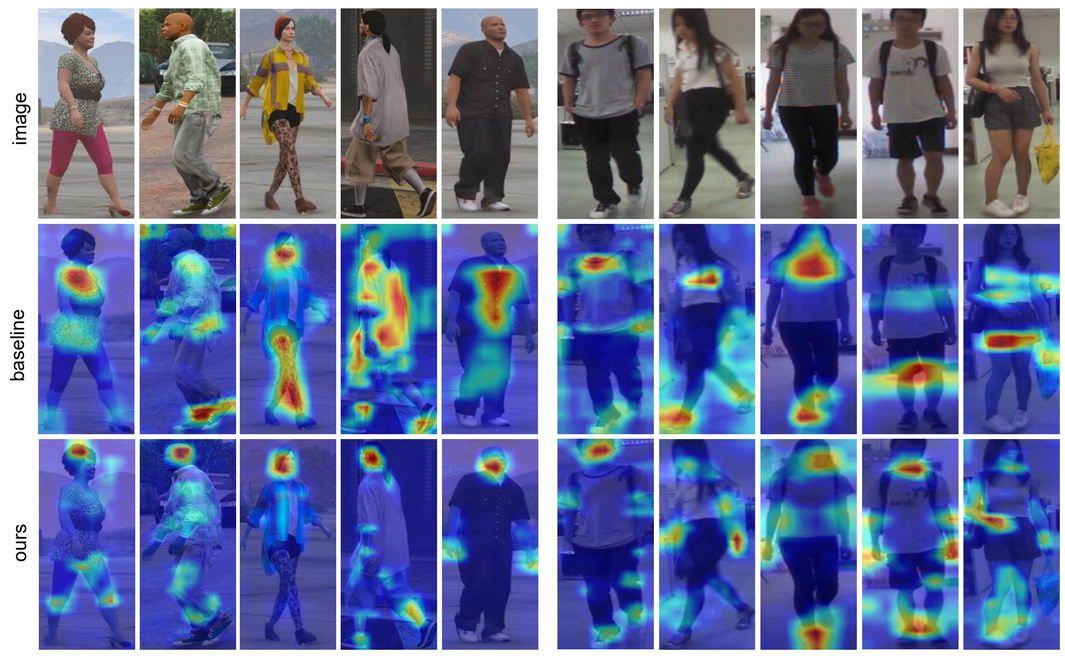
method	ref	SC		CC	
		top-1	mAP	top-1	mAP
HACNN [27]	CVPR2018	82.5	-	21.8	-
PCB [41]	ECCV2018	99.8	97.0	41.8	38.7
IANet [16]	CVPR2019	99.4	98.3	46.3	45.9
TransReID [13]	ICCV2021	97.3	95.9	47.1	49.3
SPT+ASE [54]	TPAMI2019	64.2	-	34.4	-
GI-ReID [21]	CVPR2022	80.0	-	33.3	-
RCSANet [19]	ICCV2021	100	97.2	50.2	48.6
3DSL [2]	CVPR2021	-	-	51.3	-
FSAM [15]	CVPR2021	98.8	-	54.5	-
PS [38]	SPL2021	99.2	96.6	61.1	58.3
CAL [11]	CVPR2022	**100**	**99.8**	55.2	55.8
ours w/o BSGA		98.8	96.0	59.7	56.8
ours		99.6	97.3	**61.8**	**58.7**

PRCC

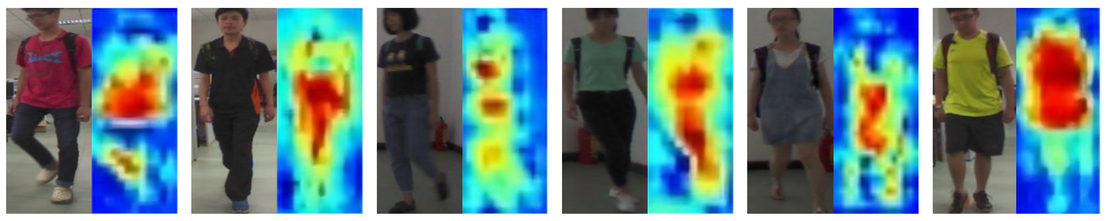
method	ref	top-1	top-5	top-10
se-resnext [18]	CVPR2018	16.7	25.5	31.2
senet [18]	CVPR2018	18.2	25.2	30.6
PCB [41]	ECCV2018	16.9	25.6	30.6
MGN [44]	ACM MM2018	18.8	28.8	33.0
TransReID [13]	ICCV2021	21.2	32.4	37.6
PS [38]	SPL2021	18.8	**35.2**	**45.5**
LSD [53]	IVC2021	16.4	27.9	34.8
baseline		20.6	29.1	33.3
baseline+BSGA		22.1	30.9	36.7
baseline+CRE		**23.0**	33.6	40.0

NKUP

Visualization



The visualization of heat maps on PRCC (right) and VC-Clothes (left).



Visualization of the original images and learned attention maps.

Conclusion

- We erase the clothes-relevant information from the original images so that the model can adaptively explore clothes-irrelevant cues.
- We further utilize the body shape mask to guide the learning of the attention map, which makes the model focus on richer and more discriminative features.
- Thorough experiments on three clothes-changing re-ID benchmarks demonstrate the performance advantages of our method.

Acknowladgements

- This work was supported by the National Natural Science Foundation of China under Grant Nos. 62102180, U1713208, 62072242, the Natural Science Foundation of Jiangsu Province under Grant Nos. BK20210329. Note that the PCA Lab is associated with Key Lab of Intelligent Perception and Systems for High-Dimensional Information of Ministry of Education, and Jiangsu Key Lab of Image and Video Understanding for Social Security, Nanjing University of Science and Technology.