Table 4 The rank-1 (%) accuracy and mAP (%) obtained by the proposed method and the state-of-the-art methods against the different levels of occlusions on the CUHK03 datasets

表 4 本文方法与其他方法在 CUHK03 数据集上不同遮挡比例下的 Rank-1(%)准确率和 mAP(%)

方法	s=0		s=0.3		s=0.6	
	Rank-1	mAP	Rank-1	mAP	Rank-1	mAP
XQDA[3], 2015	44.2	–	36.9	–	32.3	–
NPD[34], 2016	53.7	–	39.5	–	33.8	–
IDE[30], 2016	68.2	62.7	65.1	59.8	46.2	43.6
TriNet[35], 2017	79.1	76.4	68.0	66.9	48.1	49.2
Quad[36], 2017	84.3	82.0	75.3	72.4	58.1	56.5
P2S[37], 2017	54.5	51.3	47.3	45.3	37.2	36.5
GLAD[6], 2017	82.2	–	55.8	–	48.3	–
PAN[38], 2017	85.4	**90.9**	61.0	66.5	53.0	57.6
SVDNet[39], 2017	81.2	84.5	71.2	66.8	63.9	62.1
DPFL[40], 2017	82.0	78.1	–	–	–	–
AACN[5], 2018	89.5	–	–	–	–	–
$RNLSTM_A$[22], 2019	86.6	83.8	77.3	75.8	60.3	60.1
mGD+R	86.5	82.3	80.5	77.5	75.1	71.7
mGD+$RNLSTM_A$	**88.0**	84.2	**83.4**	**80.1**	**77.7**	**74.2**

我们可以得到以下的观察结果.

(1) 相比于传统手工特征和度量学习的方法,例如 NPD[34]和 XQDA[3],其他的行人重识别方法都是基于深度学习相关的方法,这类方法的识别性能有大幅度的提升,这也证明了深度学习在特征表示学习方面的优越性.

(2) 相比于基准的分类识别方法 R,我们的方法 mGD+R,在 3 个数据集的原始图上的识别结果都有不同程度的提高,这证明了我们提出的方法 mGD 有效地增强了数据的多样性,从而提高行人重识别的性能.

(3) 相比于现有的基于深度学习的方法,比如基于局部特征学习的方法(例如 GLAD[6],PAN[38]和 AACN[5]),我们提出的多尺度生成对抗网络结合简单的分类识别网络(mGD+R),就可以达到与现有的方法相近的结果.并且与现有的行人重识别方法,比如与 $RNLSTM_A$[22]结合就可以取得最好的结果.实验结果表明 mGD 对于最终的识别有明显的促进作用.主要原因就是本文的方法不仅可以有效增强训练样本,而且去除遮挡带来的干扰.

(4) 值得注意的是,相比于利用随机遮挡样本进行数据增强的方法 RandEra[18],本文提出的方法利用多尺度 GAN 进行去遮挡预处理,在面对遮挡场景时候的识别性能可以保持比较满意的识别结果.此外,本文提出的方法 mGD+R,在非遮挡场景下(s=0)的识别结果与 $RNLSTM_A$[22]方法相近.但是需要指出的是,mGD+R 利用数据增强的方式,只使用了简单的分类识别网络,而不需要设计特定的网络用于鲁棒特征的学习.同时,在不同比例的遮挡场景下的识别结果远优于 $RNLSTM_A$ 方法.这主要是由于 mGD 利用多尺度生成对抗网络进行去遮挡减少了遮挡物体的干扰,从而提高了遮挡行人重识别的识别效果.通过结合两种方法,即 mGD+$RNLSTM_A$,可以在非遮挡场景下与遮挡场景下均达到最好的识别效果,这主要是结合了去遮挡和鲁棒特征学习的优势.

4 结 论

本文详细介绍了一种新型的遮挡行人重识别方法,并设计了一个简单的多尺度生成对抗网络,其不仅可以用于生成多样的训练数据起到数据增强的作用;并且可以通过恢复遮挡区域的图像信息,进而提高遮挡场景下的行人重识别精度.本文提出的多尺度生成对抗网络可以直接与现有的行人重识别方法结合使用,充分利用去遮挡和鲁棒特征学习的优势.在多个有挑战性的行人重识别数据集上,广泛的实验设计与分析验证了我们提出方法的有效性.目前本文采用人工的方式选取遮挡区域.在未来工作中,可以利用注意力机制对遮挡的区域进行定位,并判断遮挡的严重程度,从而达到对遮挡区域干扰信息的去除以及有用信息的保留.同时,还将研究如何有效处理行人重识别剧烈光照变化带来的影响.

References:

[1] Yi D, Lei Z, Liao S, Li SZ. Deep metric learning for person re-identification. In: Proc. of the IEEE Int'l Conf. on Pattern Recognition (ICPR). 2014. 34–39. [doi: 10.1109/ICPR.2014.16]

[2] Song W, Zhao Q, Chen C, Gan Z, Liu F. Survey on pedestrian re-identification research. CAAI Trans. on Intelligent Systems, 2017,12(6):770–780 (in Chinese with English abstract). http://tis.hrbeu.edu.cn/oa/darticle.aspx?type=view&id=201706084 [doi: 10.11992/tis.201706084]

[3] Liao S, Hu Y, Zhu X, Li SZ. Person re-identification by local maximal occurrence representation and metric learning. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). 2015. 2197–2206. [doi: 10.1109/CVPR.2015.7298832]

[4] Zhao H, Tian M, Sun S, Shao J, Yan J, Yi X, Wang X, Tang X. Spindle net: person re-identification with human body region guided feature decomposition and fusion. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). 2017. 1077–1085. [doi: 10.1109/CVPR.2017.103]

[5] Xu J, Zhao R, Zhu F, Wang H, Ouyang W. Attention-aware compositional network for person re-identification. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). 2018. 2119–2128. [doi: 10.1109/CVPR.2018.00226]

[6] Wei L, Zhang S, Yao H, Gao W, Tian Q. Glad: Global-local-alignment descriptor for pedestrian retrieval. In: Proc. of the 25th ACM Int'l Conf. on Multimedia (ACM MM). 2017. 420–428. [doi: 10.1145/3123226.3123279]

[7] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial nets. In: Proc. of the Advances in Neural Information Processing Systems (NIPS). 2014. 2672–2680. [doi: 10.1007/978-1-4842-3679-6_8]

[8] Tang X, Du Y, Liu Y, Li J, Ma Y. Image recognition with conditional deep convolutional generative adversarial networks. Acta Automatica Sinica, 2018,44(5):855–864 (in Chinese with English abstract). http://www.aas.net.cn/CN/Y2018/V44/I5/855 [doi: 10.16383/j.aas.2018.c170470]

[9] Arjovsky M, Chintala S, Bottou L. Wasserstein generative adversarial networks. In: Proc. of the Int'l Conf. on Machine Learning (ICML). 2017. 214–223.

[10] Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V, Courville AC. Improved training of Wasserstein gans. In: Proc. of the Advances in Neural Information Processing Systems (NIPS). 2017. 5767–5777.

[11] Zhu JY, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proc. of the IEEE Int'l Conf. on Computer Vision (ICCV). 2017. 223–2232. [doi: 10.1109/ICCV.2017.244]

[12] Choi Y, Choi M, Kim M, Ha JW, Kim S, Choo J. Stargan: Unified generative adversarial networks for multi-domain image-to-image translation. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). 2018. 8789–8797. [doi: 10.1109/CVPR.2018.00916]

[13] Li D, Chen X, Zhang Z, Huang K. Learning deep context-aware features over body and latent parts for person re-identification. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). 201. 384–393. [doi: 10.1109/CVPR.2017.782]

[14] Ahmed E, Jones M, Marks TK. An improved deep learning architecture for person reidentification. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). 2015. 3908–3916. [doi: 10.1109/CVPR_2015.7299016]

[15] Varior RR, Haloi M, Wang G. Gated siamese convolutional neural network architecture for human re-identification. In: Proc. of the European Conf. on Computer Vision (ECCV). 2016. 791–808. [doi: 10.1007/978-3-319-46484-8_48]

[16] Shi H, Yang Y, Zhu X, Liao S, Lei Z, Zheng W, Li SZ. Embedding deep metric for person re-identification: A study against large variations. In: Proc. of the European Conf. on Computer Vision (ECCV). 2016. 732–748. [doi: 10.1007/978-3-319-46448- 0_44]

[17] Cheng D, Gong Y, Zhou S, Wang J, Zheng N. Person re-identification by multi-channel parts-based cnn with improved triplet loss function. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). 2016. 1335–1344. [doi: 10.1109/CVPR.2016.149]

[18] Zhong Z, Zheng L, Kang G, Li S, Yang Y. Random erasing data augmentation. arXiv Preprint arXiv:1708.04896, 2017.

[19] Huang H, Li D, Zhang Z, Chen X, Huang K. Adversarially occluded samples for person re-identification. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). 2018. 5098–5107. [doi: 10.1109/CVPR.2018.00535]

[20] Zhuo J, Chen Z, Lai J, Wang G. Occluded person re-identification. In: Proc. of the IEEE Int'l Conf. on Multimedia and Expo (ICME). 2018. 1–6. [doi: 10.1109/ICME.2018.8486568]

[21] Su C, Li J, Zhang S, Xing J, Gao W, Tian Q. Pose-driven deep convolutional model for person re-identification. In: Proc. of the IEEE Int'l Conf. on Computer Vision (ICCV). 2017. 3960–3969. [doi: 10.1109/ICCV.2017.427]

[22] Yang W, Yan Y, Chen S. Adaptive deep metric embeddings for person re-identification under occlusions. Neurocomputing, 2019,340:125–132. [doi: 10.1016/j.neucom.2019.02.042]

[23] Hou R, Ma B, Chang H, Gu X, Shan S, Chen X. VRSTC: Occlusion-free video person re-identification. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). 2019. 7183–7192.

[24] Zhong Z, Zheng L, Zheng Z, Li S, Yang Y. Camera style adaptation for person reidentification. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). 2018. 5157–5166. [doi: 10.1109/CVPR.2018.00541]

[25] Liu J, Ni B, Yan Y, Zhou P, Cheng S, Hu J. Pose transferrable person reidentification. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). 2018. 4099–4108. [doi: 10.1109/CVPR.2018.00431]

[26] Qian X, Fu Y, Xiang T, Wang W, Qiu J, Wu Y, Jiang Y, Xue X. Pose-normalized image generation for person re-identification. In: Proc. of the European Conf. on Computer Vision (ECCV). 2018. 650–667. [doi: 10.1007/978-3-030-01240-3_40]

[27] Deng W, Zheng L, Ye Q, Kang G, Yang Y, Jiao J. Image-image domain adaptation with preserved self-similarity and domain-dissimilarity for person re-identification. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). 2018. 994–1003. [doi: 10.1109/CVPR.2018.00110]

[28] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). 2016. 770–778. [doi: 10.1109/CVPR.2016.90]

[29] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. In: Proc. of the Int'l Conf. on Learning Representations (ICLR). 2015.

[30] Zheng L, Yang Y, Hauptmann AG. Person re-identification: Past, present and future. arXiv Preprint arXiv:1610.02984, 2016.

[31] Ristani E, Solera F, Zou R, Cucchiara R, Tomasi C. Performance measures and a data set for multi-target, multi-camera tracking. In: Proc. of the European Conf. on Computer Vision (ECCV). 2016. 17–35. [doi: 10.1007/978-3-319-48881-3_2]

[32] Li W, Zhao R, Xiao T, Wang X. Deepreid: Deep filter pairing neural network for person re-identification. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). 2014. 152–159. [doi: 10.1109/CVPR_2014.27]

[33] Zheng W, Li X, Xiang T, Liao S, Lai J, Gong S. Partial person re-identification. In: Proc. of the IEEE Int'l Conf. on Computer Vision (ICCV). 2015. 4678–4686. [doi: 10.1109/ICCV.2015.531]

[34] Zhang L, Xiang T, Gong S. Learning a discriminative null space for person reidentification. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). 2016. 1239–1248. [doi: 10.1109/CVPR.2016.139]

[35] Hermans A, Beyer L, Leibe B. In defense of the triplet loss for person reidentification. arXiv Preprint arXiv:1703.07737, 2017.

[36] Chen W, Chen X, Zhang J, Huang K. Beyond triplet loss: A deep quadruplet network for person re-identification. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). 2017. 403–412. [doi: 10.1109/CVPR.2017.145]

[37] Zhou S, Wang J, Wang J, Gong Y, Zheng N. Point to set similarity based deep feature learning for person re-identification. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). 2017. 3741–3750. [doi: 10.1109/CVPR.2017.534]

[38] Zhao L, Li X, Zhuang Y, Wang J. Deeply-learned part-aligned representations for person re-identification. In: Proc. of the IEEE Int'l Conf. on Computer Vision (ICCV). 2017. 3219–3228. [doi: 10.1109/ICCV.2017.349]

[39] Sun Y, Zheng L, Deng W, Wang S. Svdnet for pedestrian retrieval. In: Proc. of the IEEE Int'l Conf. on Computer Vision (ICCV). 2017. 3800–3808. [doi: 10.1109/ICCV.2017.410]

[40] Chen Y, Zhu X, Gong S. Person re-identification by deep learning multi-scale representations. In: Proc. of the IEEE Int'l Conf. on Computer Vision (ICCV). 2017. 2590–2600. [doi: 10.1109/ICCVW.2017.304]

附中文参考文献:

[2] 宋婉茹,赵晴晴,陈昌红,干宗良,刘峰.行人重识别研究综述.智能系统学报,2017,12(6):770–780. http://tis.hrbeu.edu.cn/oa/darticle.aspx?type=view&id=201706084 [doi:10.11992/tis.201706084]

[8] 唐贤伦,杜一铭,刘雨微,李佳歆,马艺玮.基于条件深度卷积生成对抗网络的图像识别方法.自动化学报,2018,44(5):855–864. http://www.aas.net.cn/CN/Y2018/V44/I5/855 [doi: 10.16383/j.aas.2018.c170470]

杨婉香(1993－),女,硕士,主要研究领域为行人重识别.

严严(1984－),男,博士,副教授,CCF 专业会员,主要研究领域为计算机视觉,模式识别.

陈思(1987－),女,博士,副教授,CCF 专业会员,主要研究领域为计算机视觉,模式识别.

张小康(1996－),男,学士,主要研究领域为行人重识别.

王菡子(1973－),男,博士,教授,博士生导师,CCF 专业会员,主要研究领域为计算机视觉,模式识别.