

融合多尺度注意力和双向 LSTM 的行人重识别

闫昊雷¹, 李小春¹, 张仁飞², 张磊², 邱浪波³, 王哲⁴

(1. 空军工程大学信息与导航学院, 西安, 710077; 2. 武警陕西省总队, 西安, 710054;
3. 陕西省信息化工程研究院, 西安, 710061; 4. 陆军装备部, 北京, 100000)

摘要 将深度学习应用于行人重识别领域, 嵌入多尺度注意力融合模块至神经网络中进行多尺度特征提取和表示, 可有效提升注意力机制对深度学习网络的识别性能。提出了一种基于 SE block 的多尺度通道注意力融合模块, 并结合 ResNet50 卷积神经网络提取特征; 然后通过双向 LSTM 网络进一步提取特征序列上下文信息, 在提高模型对图像重要特征的提取能力的同时, 降低对图像冗余特征的关注度; 最后使用级联难采样三元组损失函数和交叉熵损失函数共同训练网络模型, 使样本能够在高维特征空间中实现聚类, 进一步提升模型识别准确性。所提出算法在 Market1501 数据集和 CUHK03 数据集分别进行实验, 并在同等条件下和其他注意力模块算法进行比较。为进一步验证各模块作用, 对算法进行消融实验, 以验证各模块的有效性, 实验结果表明, 所提出方法可有效应用于行人重识别。

关键词 注意力机制; 卷积神经网络; 行人重识别; 深度学习; LSTM

DOI 10.3969/j.issn.2097-1915.2022.05.011

中图分类号 TP391 **文献标志码** A **文章编号** 2097-1915(2022)05-0071-06

A Pedestrian Re-ID with Multi-Scale Attention and Bidirectional LSTM

YAN Haolei¹, LI Xiaochun¹, ZHANG Renfei², ZHANG Lei², QIU Langbo³, WANG Zhe⁴

(1. Information and Navigation School, Air Force Engineering University, Xi'an 710077, China;
2. People's Armed Police of Shaanxi Province, Xi'an 710054, China;
3. Shaanxi Information Engineering Research Institute, Xi'an 710061, China;
4. Army Equipment Department, Beijing 100000, China)

Abstract With the rapid development of the information society, taking video sensors as the front-end for acquiring information is of great significance in effectively finding specific objects through pedestrian re-identification algorithms to protect people's lives and property. This paper applies deep learning to the field of person re-identification, and embeds the multi-scale attention fusion module into the neural network for multi-scale feature extraction and representation, effectively improving the recognition performance of the attention mechanism for deep learning networks. The paper proposes a multi-scale channel attention fusion module based on SE block in combination with the ResNet50 convolutional neural network to extract features, further extract the feature sequence context information through the bidirectional LSTM network, and improve the model's ability to extract important image features. At the same time, the attention to re-

收稿日期: 2022-03-30

基金项目: 陕西省重点发展计划(2021ZDLSF06-09)

作者简介: 闫昊雷(1995—), 男, 陕西西安人, 硕士生, 研究方向为计算机视觉。E-mail: yanhl1995@126.com

通信简介: 李小春(1974—), 女, 河南修武人, 副教授, 研究方向为模式识别研究。E-mail: chunwind@sohu.com

引用格式: 闫昊雷, 李小春, 张仁飞, 等. 融合多尺度注意力和双向 LSTM 的行人重识别[J]. 空军工程大学学报, 2022, 23(5): 71-76. YAN Haolei, LI Xiaochun, ZHANG Renfei, et al. Pedestrian Re-ID with Multi-Scale Attention and Bidirectional LSTM [J]. Journal of Air Force Engineering University, 2022, 23(5): 71-76.

dundant features of images is reduced. Finally, the network model is jointly trained by the cascaded hard-sampling triplet loss function and the cross-entropy loss function, clustering the samples in the high-dimensional feature space, and further improving the model recognition accuracy. Market1501 dataset and CUHK03 dataset are tested by the proposed algorithm respectively, and compared with other attention module algorithms under the same conditions. In order to further verify the function of each module, an ablation experiment is performed by the algorithm to verify the effectiveness of each module. The experimental results show that the proposed method can be effectively applied to person re-identification.

Key words attention mechanism; convolutional neural network; pedestrian re-ID; deep learning; LSTM

随着我国“智慧城市”“智慧军营”的建设发展,以视频传感器为前端信息获取来源,结合人工智能技术进行后台信息处理的方法受到研究人员的广泛关注。行人重识别是一种应用于数据库中查询特定行人的方法,即对输入行人图像进行解析,提取图像外部特征,根据这些特征,可进行跨区域的行人匹配,寻找特定对象,极大地节省后台信息处理人员的时间和精力,可广泛应用于抓捕逃犯、不明身份人员进入军事涉密区域的跟踪和定位等军事场景和智能安防、无人超市等民用场景,但由于现实场景的复杂多变,不同相机存在质量差异,行人特征受外部环境影响较大,影响检测准确度。因此,行人重识别算法研究现如今仍具有挑战性^[1]。以神经网络为基础的深度学习方法因其强大的识别效果受到研究人员青睐,该方法通过形成深层神经网络对图像提取特征,然后进行图像的分类和检索,避免传统手工算法中依靠人类寻找图像特征带来提取效果不佳的影响。具体来说,行人重识别在深度学习中的研究思路就是将输入图像通过深层神经网络,由网络提取特征加以利用,达到快速捕获其他视频传感设备下的该行人图像^[2]。

注意力机制作为嵌入在神经网络模型中的一种模块,帮助网络识别图像的关键特征,忽略冗余背景特征。融入注意力机制的神经网络模型对图像特征提取效果更好,因此引发研究人员的广泛关注。如 Hu^[3]等人提出的 SENet 将各通道降维后压缩为对特征图的权重,从通道维度寻找与特征图间的关系。Chen^[4]对注意力机制中复杂的高阶统计信息进行建模和利用,从而捕捉行人之间的细微差异。宋晓花^[5]等人将通过含有注意力模块的主干网络后的特征图进行多尺度融合,获取不同尺度下的特征信息。Liu^[6]等人通过将图像划分为不同窗口的形式,降低了运算量,将自注意力^[7]融入目标检测算法中。Zhang^[8]等人通过卷积操作寻找特征图空间维度和通道维度的相互关系,形成注意力权重。当前,对注意力机制研究主要是集中在模块的改进和创新上,而对多尺度的注意力模块融合仍有待研究。由于目

标全局特征不足以充分描述行人图像的细节信息,故可根据先验知识,寻求行人身体各部位之间在图像上的排列规律,即行人图像可看作是从头部到腿部的有规律的排列,沿用循环神经网络可提取出序列上下文信息的思路,尽管行人每个身体部位在不同图像中并不总是位于相同的位置,但所有行人身体部位仍可以按身体排列规则进行建模。长短期记忆网络(LSTM)是门控循环神经网络的一种,其引入自循环方法,以产生梯度长时间持续流动的路径,通过门控自循环的权重,累积的时间尺度可以动态地改变。为进一步寻找行人图像各局部特征间关系,本文在融入多尺度注意力模块的卷积神经网络架构基础上,连接 LSTM^[9]提取行人图像各区域间的序列特征关系。

针对上述问题,本文提出一种基于多尺度注意力融合和双向 LSTM 的行人重识别算法(MAFL-Net),该算法基于卷积神经网络和循环神经网络,使模型通过卷积网络在融入注意力机制有效提取图像特征时,同时关注不同像素区域之间的独特的序列关系。最后使用难采样三元组损失函数^[10]和交叉熵损失函数在 CUHK03^[11]和 Market1501^[12]数据集上对该网络进行训练,检验算法有效性。

1 相关概念

1.1 注意力机制

注意力机制在图像处理和自然语言处理领域已经得到广泛应用,如:图像分类^[13]、图像描述^[14]、物体检测^[15]、行人重识别^[16]。现有的卷积神经网络无法根据输入图像提取特定目标的特征信息,而注意力机制能够作用于图像的中间层特征,帮助模型提取图像所包含的具有辨识度的关键特征。常见的注意力主要有空间注意力和通道注意力。空间注意力旨在提升关键区域的特征表达,将原始图片中的空间信息变换到高维空间中并保留关键信息,为每个区域生成权重掩膜并加权输出;通道注意力能够寻找不同通道对于特征图的相关性,通过神经网络

自动获取各通道的权重, 强化重要通道特征并抑制非重要通道特征。

在图像语义理解方面, 选择性注意力和自注意力通过不同方式完成特征提取。选择性注意力模块以模型中间层特征为依据, 通过神经网络学习注意力权重, 并与原特征图融合, 从而达到加强重要特征而抑制非相关成分的目的。而自注意力则从图像自身特点出发, 减少对外部信息的依赖, 尽可能地利用特征内部固有的信息进行注意力的交互。其通过 embedding 变换生成 Key、Query 和 Value, 根据前两项计算权重系数, 然后再针对该权重对 Value 加权求和, 完成对 Query 位置特征的调整。

1.2 LSTM

为进一步获取区域特征间的关系, 本文采用循环神经网络(RNN)获取图像的上下文信息。LSTM 结构如图 1 所示。

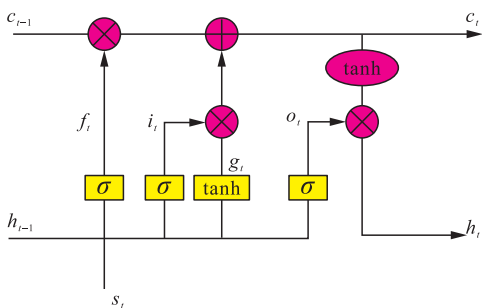


图 1 LSTM 单元结构

图 1 中, 在 2 个输入中, 上面的输入是单元的记
忆状态 c_{t-1} , 下面的输入是上一层的输出 h_{t-1} 。在 2

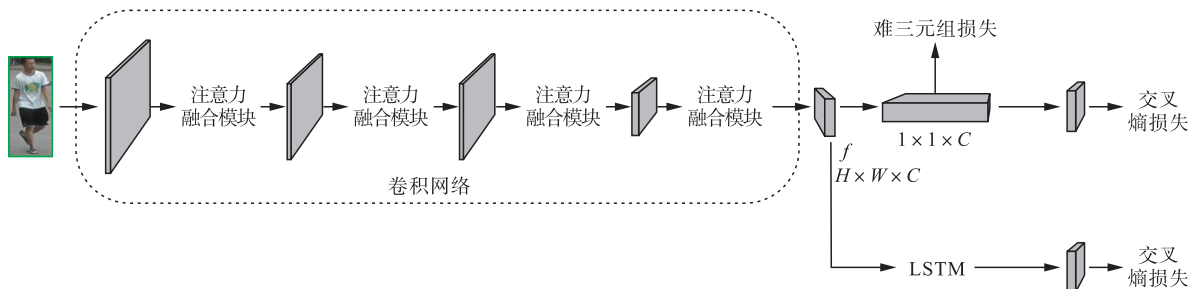


图 2 多尺度注意力融合模块和 LSTM 的改进网络模型(MAFLNet)

2.2 多尺度注意力融合模块

本文通道注意力模块选为 SE block 作为研究对象, 将原始图像通过含有融合多尺度通道注意力模块的 ResNet50 卷积网络, 提取输入图像的全局特征后, 通过难三元组损失函数和交叉熵损失函数进行训练。由于卷积神经网络在处理不同通道维度的特征图时所赋予的各通道权重也有所不同, 故可在注意力融合模块中, 对不同通道下的特征图像进行融合, 即对输入的特征图在通道维度进行切分和组合操作, 使不同通道维度下的注意力有效融合。

个输出中, 上面的输出是下一个单元的记忆状态 c_t , 下面的输出是当前层的输出 h_t 。中间结构的 3 个 Sigmoid 激活函数代表 3 个门控制单元, 由于 Sigmoid 函数的输出是 $0 \sim 1$, 因此可以实现控制阀机制。 f_t 为遗忘门输出, 主要用于对于上一个节点的输出进行选择性的遗忘; i_t 为输入门的输出的值, 由 S_t 和 h_{t-1} 控制。输出门 o_t 用于输出隐藏状态 h_t , tanh 为激活函数。

2 多尺度注意力融合和 LSTM 网络

2.1 设计思路

本文采取基于多尺度注意力融合的方法的卷积神经网络获取图像的全局特征, 循环神经网络 LSTM 获取局部特征之间相互关系的方法共同对输入图像进行特征提取。思想框架如图 2 所示, 其中, 长方体表示特征图。具体来说, 首先将输入图像通过基于多尺度通道注意力融合的卷积神经网络 ResNet50, 用以提取全局特征, 然后根据行人各部位构造在图像中具有规律性, 对全局特征序列等份分割后获取相应的局部特征序列, 接着通过双向 LSTM 网络寻找各局部特征之间的相互关系。对通过卷积神经网络获取的全局特征和循环神经网络获取的局部特征使用交叉熵损失函数和难采样三元组损失函数共同进行训练。

多尺度注意力融合模块主要结构如图 3 所示, 为方便阐述, 对网络中间层的批归一化层和 ReLU 激活函数予以省略。主要步骤如下: 假设输入特征图为 $X \in R^{C \times H \times W}$, 其中 C 为图像通道数, H 为图像高度, W 为图像宽度。

首先, 将输入特征图按通道维度等份分割为两部分, 用 X_1, X_2 表示, $X_1 \in R^{\frac{C}{2} \times H \times W}, X_2 \in R^{\frac{C}{2} \times H \times W}$ 。为提升通道特征表示能力, 将 X_1, X_2 融合得到特征图 $X_3 \in R^{\frac{C}{2} \times H \times W}$, 然后对 X_3 使用全局平均池化操作, 将特征图 X_3 空间维度压缩至一维。为下一步获取注意力图, 再利用 1×1 卷积将通道数恢复至 C ,

得到 $X_4 \in R^{C \times 1 \times 1}$, 将 X_4 通过 softmax 函数得到注意力图 $AT \in R^{C \times 1 \times 1}$, 将注意力图 AT 按照通道维度拆分为 AT_1, AT_2 , 其中 $AT_1 \in R^{\frac{C}{2} \times H \times W}$, $AT_2 \in R^{\frac{C}{2} \times H \times W}$, 再将 AT_1, AT_2 与 X_1, X_2 分别相乘, 所得结果组合后得到特征图 X_{out1} , 其中 $X_{out1} \in R^{C \times H \times W}$. 为进一步提取通过融合多尺度注意力模块后特征图

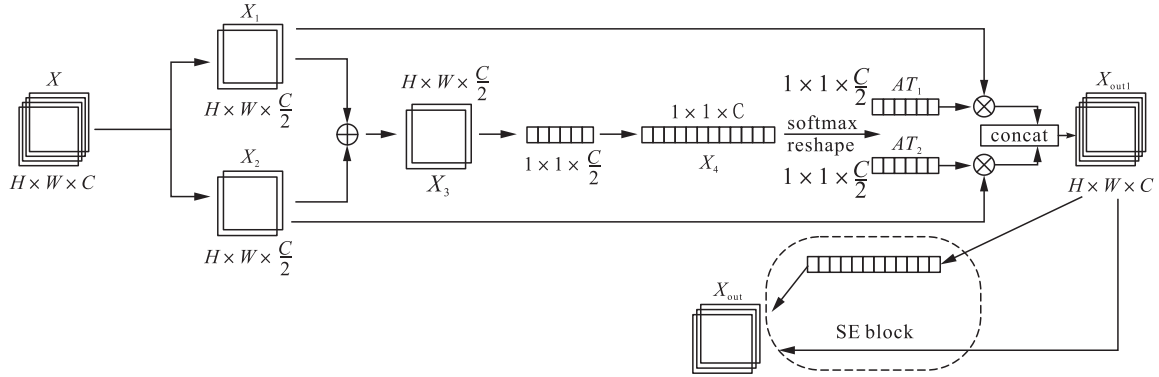


图3 多尺度注意力融合模块

2.3 基于双向 LSTM 网络的序列关系提取方法

为提高模型的预测准确度,进一步使网络理解局部特征间相互关系,使用双向 LSTM 可分别从特征向量的起始点和终点双向移动,用于在二维图像上获取上、下、左、右 4 个方向的信息,为获取更准确的上下文信息,本文算法在设计中将隐藏层设为两层,其具体结构如图 4 所示。将通过多尺度注意力融合网络获取的全局特征向量进行水平分割后,得到该图像各局部特征所对应的特征向量,将所得向量输入至两层双向 LSTM 网络,以获取局部特征之间的相互关系。然后,将所有特征向量连接在一起作为最终的行人局部特征。具体地,由卷积网络所得的输出 $X' \in R^{C \times H \times W}$, 其中 C 为特征图的通道数, H 为图像高度, W 为图像宽度。为获取图像垂直方向的序列关系,首先使用全局平均池化操作将图像宽度压缩至一维,得到 $X_1 \in R^{C \times H \times 1}$, 然后对 X_1 进行等份分割,从而产生 H 个通道数为 C 的特征向量,再将每个特征序列通过两层双向 LSTM 网络,生成 U 个特征向量,从而更好地表示特征序列连接关系。最后,将所有特征向量连接在一起作为最终的行人局部特征关系表示向量,再通过全连接层进行分类,最后使用交叉熵损失函数对模型进行训练。

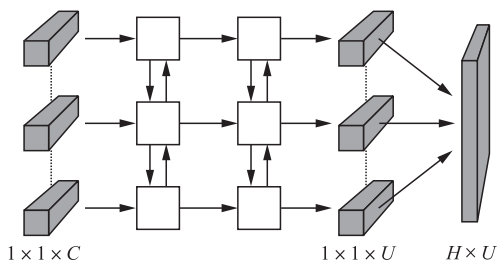


图4 双向 LSTM 结构

的全局特征,将 X_{out1} 通过通道注意力模块 SE block,即得整个多尺度注意力融合模块的输出 X_{out} 。该多尺度注意力融合模块主要设置于 Resnet50 的每个 layer 层。通过卷积网络得到特征向量后,采用交叉熵函数和难三元组损失函数级联的方式共同对模型进行训练。

3 实验结果和分析

本文算法基于 Python 的深度学习框架 Pytorch,使用英伟达 3090 显卡进行模型训练和评估测试。在训练中,为提升模型鲁棒性,对数据集采取 3 种数据增强方法,即随机擦除、随机裁剪和水平翻转。模型训练轮数设置为 400,采用 Adam 优化器,初始学习率设置为 0.000 8,经过每 40 轮训练,将学习率降低为原先的 0.5 倍,难采样三元组损失函数中的超参数设置为 0.3。在训练和测试中,每个 batch 设置为 64,并采用余弦距离计算样本间的距离。

3.1 数据集选取

为检验本文算法的有效性,选用公开数据集 Market1501 和 CUHK03 作为本次实验的数据集。

Market1501:该数据集包含 751 个行人的 12 936 张图像用于训练,750 个行人的 19 732 张图像用于测试。所有图像均由 5 台高分辨率相机和 1 台低分辨率相机拍摄。

CUHK03:该数据集含有 1 467 个不同的人物,由 5 对摄像头采集,其中训练集包含 767 个行人的 7 670 张图像,测试集包含 700 个行人的 7 000 张图像。

3.2 实验评价标准和实施方案

为使各算法具有可比较性,实验中采取相同的实验环境和参数,并采用 mAP 和 Rank-1 作为评价标准。Rank- k 通过计算前 k 个结果中预测正确的个数来表示算法准确性,如 Rank-1 表示预测出的

图像中第 1 张正确的概率。mAP 则反映了算法在整个测试集上的准确率。

为检验模型有效性,文中使用文献[17]的在 ImageNet 网络上预训练模型 ResNet50 作为本次实验的 baseline,将本文算法和当前典型的其他行人重识别算法进行对比,检验算法的有效性和准确率。为检验各环节对算法性能的影响,将网络架构分解为融入 SE block 的 baseline、融入多尺度注意力融合模块网络(MAFNet)和多尺度注意力融合和双向 LSTM 的网络(MAFLNet)进行消融实验,以验证多尺度注意力融合模块和双向 LSTM 网络在行人重识别算法的有效性。

3.3 实验结果分析

本文算法和其他当前基于注意力模块的行人重识别算法实验结果如表 1 所示,通过实验结果对比发现,在 CUHK03 和 Market1501 数据集上,baseline 的特征提取效果一般,mAP 分别只有 69% 和 83.7%,Rank-1 分别为 73.8% 和 94.2%,而本文所提出的 MAFLNet 算法 mAP 分别达到 77.2% 和 87.1%,Rank-1 达到 79.4% 和 95.4%,在 mAP 值平均提高 5.8%,Rank-1 平均提高 3.4%,和其他算法进行对比,准确率也有较大提升。

表 1 本文算法和其他算法对比

数据集	CUHK03		Market1501	
	Rank-1	mAP	Rank-1	mAP
baseline	73.8	69	94.2	83.7
OSNet ^[18]	72.3	67.8	94.8	84.9
MHN-6(PCB) ^[4]	77.2	72.4	95.1	85
文献[19]	78.9	75.8	95.9	88.7
文献[5]	—	—	92.7	80.4
AMFC ^[20]	75.6	76.7	94.6	92.5
MAFLNet	79.4	77.2	95.4	87.1

为消除各环节间的相互影响,按照实验方案分为 3 个部分进行消融实验。实验结果如表 2 所示。通过结果发现,使用 SE block 后,mAP 在 2 个数据集上分别提高至 71.2% 和 85.5%,Rank-1 提升至 73.8% 和 94.2%,表明该通道注意力模块可有效获取相应通道权重,提取通道重要信息;在该网络后添加双向 LSTM 网络,mAP 在 2 个数据集上分别提高至 72.5% 和 86%,Rank-1 提升至 75.8% 和 94.5%。通过融合注意力模块的网络后,双向 LSTM 网络根据行人图像特点,进一步学习输出特征序列的上下文信息,使模型预测准确率进一步提升。采用 MAFNet 对输入图像进行特征提取,mAP 在 2 个数据集上分别提高至 75.6% 和 86.4%,

Rank-1 提升至 78.4% 和 94.8%,表明相较于基于 SE block 的单一尺度注意力网络,基于 SE block 的多尺度注意力模块可学习特征图在多个尺度下的通道信息权重并进行融合,在保证不低于单一尺度下的特征提取效果情况下,进一步挖掘不同尺度下的通道信息,从而关注重要特征而忽略冗余特征;为进一步验证和提升模型提取效果,在该网络后加入双向 LSTM 结构,形成本文算法 MAFLNet。实验结果表明,mAP 在 2 个数据集上分别提高至 77.2% 和 87.1%,Rank-1 提升至 79.4% 和 95.4%,说明各分支结构均可提升模型预测精准度,多尺度注意力模块可有效引导模型关注关键特征,双向 LSTM 结构能有效获取图像的局部特征间的关系。

表 2 各环节网络实验结果

数据集	CUHK03		Market1501	
	Rank-1	mAP	Rank-1	mAP
baseline	73.8	69.0	94.2	83.7
SENet	74.3	71.2	94.3	85.5
SENet+LSTM	75.8	72.5	94.5	86.0
MAFNet	78.4	75.6	94.8	86.4
MAFLNet	79.4	77.2	95.4	87.1

为进一步直观感受各网络特征提取效果,对各个环节训练出的网络进行可视化处理,如图 5 所示。通过热点图可看出,本文算法 MAFLNet 更能关注行人的关键性特征,忽略冗余的背景特征,因此预测效果较好。

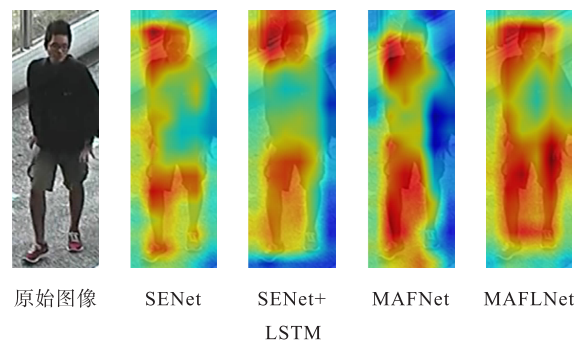


图 5 网络热点图

4 结语

本文提出一种基于多尺度融合和双向 LSTM 的行人重识别算法,旨在将多尺度下注意力模块进行融合,提取与单一尺度不同的且更为关键的特征,后采用双向 LSTM 网络寻找图像上下文信息,并通过交叉熵损失函数和难采样三元组损失函数进行训练,实验结果表明本文算法的有效性,提高了行人重

识别算法的准确度。由于本文算法在同一数据集下进行训练和测试,模型泛化能力较弱。因此,有关模型的迁移学习仍有进一步研究空间。

参考文献

- [1] 朱宽堂,张建勋,谭暑秋. 基于全局特征和多种局部特征的行人重识别[J/OL]. 微电子学与计算机: 1-8. [2021-11-12]. <https://doi.org/10.19304/J.ISSN1000-7180.2021.0837>.
- [2] SONG W R, ZHAO Q Q, CHEN C H, et al. Survey on Pedestrian Re-identification Research[J]. CAAI Transactions on Intelligent Systems, 2017, 12(6): 770-780.
- [3] HU J, SHEN L, SUN G. Squeeze-and-Excitation Networks[J]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [4] CHEN B, DENG W, HU J. Mixed High-Order Attention Network for Person Re-Identification[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea; IEEE, 2019: 371-381.
- [5] 宋晓茹,杨佳,高嵩,等. 基于注意力机制与多尺度特征融合的行人重识别方法[J]. 科学技术与工程, 2022, 22(4): 1526-1533.
- [6] LIU Z, LIN Y, CAO Y, et al. Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows[EB/OL]. (2021-08-17)[2022-03-30]. <https://doi.org/10.48550/arXiv.2013.14030>.
- [7] VASWANI A, SHAZEER N, Parmar N, et al. Attention Is All You Need[EB/OL]. (2017-12-06)[2022-03-30]. <http://arXiv.org/abs/706.03762>. 2013.14030.
- [8] ZHANG Z, LAN C, ZENG W, et al. Relation-Aware Global Attention for Person Re-Identification[C]//IEEE/cvf Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2020: 3186-3195.
- [9] HOCHREITER S, SCHMIDHUBER J. Long Short-Term Memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [10] HERMANS A, BEYER L, LEIBE B. In Defense of the Triplet Loss for Person Re-Identification[EB/OL]. (2017-11-21)[2022-03-30]. <http://arXiv.org/abs/1703.0773>.
- [11] LI W, ZHAO R, XIAO T, et al. DeePreId: Deep Filter Pairing Neural Network for Person Re-Identification[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA: IEEE, 2014: 152-159.
- [12] ZHENG L, SHEN L, TIAN L, et al. Scalable Person Re-Identification: A Benchmark[C]//2015 IEEE International Conference on Computer Vision. Santiago, Chile: IEEE, 2015: 1116-1124.
- [13] WOO S Y, PARK JH, LEE JY, et al. CBAM: Convolutional Block Attention Module[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018: 3-19.
- [14] XU K, BA J, KIROS R, et al. Show, Attend and Tell: Neural Image Caption Generation with Visual Attention[EB/OL]. (2015-02-10)[2022-03-30]. <http://arXiv.org/abs/1502.03044>.
- [15] BA J, MNIH V, KAVUKCUOGLU K. Multiple Object Recognition with Visual Attention[EB/OL]. (2014-12-24)[2022-03-30]. <http://arXiv.org/abs/1412.7755>.
- [16] ZHAO L, LI X, ZHUANG Y, et al. Deeply-Learned Part-Aligned Representations for Person Re-Identification[C]//2017 IEEE International Conference on Computer Vision (ICCV). Venice: IEEE, 2017: 3239-3248.
- [17] HE K, GIRSHICK R, DOLLAR P. Rethinking ImageNet Pre-Training[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South): IEEE, 2019: 4917-4926.
- [18] ZHOU K, YANG Y, CAVALLARO A, et al. Omni-Scale Feature Learning for Person Re-Identification[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South): IEEE, 2019: 3701-3711.
- [19] 董虎胜,钟珊,杨元峰,等. 基于多粒度区域相关深度特征学习的行人重识别[J]. 计算机科学, 2021, 48(12): 269-277.
- [20] 张正一,丁建伟,魏慧雯,等. 基于注意力机制的多级特征级联行人重识别[J]. 激光与光电子学进展, 2021, 58(22): 374-383.

(编辑:徐楠楠)