

中图分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2023)02-0471-12

论文引用格式: Wang Z X, Xu Z M, Xue Y Y, Lang C Y, Li Z and Wei L L. 2023. Global and spatial multi-scale contexts fusion for vehicle re-identification. Journal of Image and Graphics, 28(02): 0471-0482 (王振学, 许喆铭, 雪洋洋, 郎丛妍, 李尊, 魏莉莉. 2023. 融合全局与空间多尺度上下文信息的车辆重识别. 中国图象图形学报, 28(02): 0471-0482) [DOI:10.11834/jig.210849]

融合全局与空间多尺度上下文信息的车辆重识别

王振学, 许喆铭, 雪洋洋, 郎丛妍*, 李尊, 魏莉莉

北京交通大学计算机与信息技术学院, 北京 100044

摘要: **目的** 车辆重识别指判断不同摄像设备拍摄的车辆图像是否属于同一辆车的检索问题。现有车辆重识别算法使用车辆的全局特征或额外的标注信息, 忽略了对多尺度上下文信息的有效抽取。对此, 本文提出了一种融合全局与空间多尺度上下文信息的车辆重识别模型。**方法** 首先, 设计一个全局上下文特征选择模块, 提取车辆的细粒度判别信息, 并且进一步设计了一个多尺度空间上下文特征选择模块, 利用多尺度下采样的方式, 从全局上下文特征选择模块输出的判别特征中获得其对应的多尺度特征。然后, 选择性地集成来自多级特征的空间上下文信息, 生成车辆图像的前景特征响应图, 以此提升模型对于车辆空间位置特征的感知能力。最后, 模型组合了标签平滑的交叉熵损失函数和三元组损失函数, 以提升模型对强判别车辆特征的整体学习能力。**结果** 在 VeRi-776 (vehicle re-identification-776) 数据集上, 与模型 PNVR (part-regularized near-duplicate vehicle re-identification) 相比, 本文模型的 mAP (mean average precision) 和 rank-1 (cumulative matching curve at rank 1) 评价指标分别提升了 2.3% 和 2.0%。在该数据集上的消融实验验证了各模块的有效性。在 Vehicle ID 数据集的大规模测试子集上, 就 rank-1 和 rank-5 (cumulative matching curve at rank 5) 而言, 本文模型的 mAP 比 PNVR 分别提升了 0.8% 和 4.5%。**结论** 本文算法利用全局上下文特征和多尺度空间特征, 提升了拍摄视角变化、遮挡等情况下车辆重识别的准确率, 实验结果充分表明了所提模型的有效性与可行性。

关键词: 车辆重识别; 深度学习; 局部可区分性特征; 特征选择; 多尺度空间特征

Global and spatial multi-scale contexts fusion for vehicle re-identification

Wang Zhenxue, Xu Zheming, Xue Yangyang, Lang Congyan*, Li Zun, Wei Lili

School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China

Abstract: **Objective** Vehicle re-identification issue is concerned of identifying the same vehicle images captured from multiple cameras-based non-overlapping views. Its applications and researches have been developing in computer vision like intelligent transportation system and public traffic security. Current sensor-based methods are focused on hardware detectors utilization as a source of information inputs for vehicle re-identification, but these methods are challenged to get effective information of the vehicle features in related to its color, length, and shape. To obtain feature information about the vehicle, most methods are based on label-manual features in the context of edges, colors and corners. However, such special decorations are challenged to be identified on the aspects of camera view variation, low resolution, and object occlusion of the vehicle-captured images. Thanks to the emerging deep learning technique, vehicle re-identification methods have been

收稿日期: 2021-10-13; 修回日期: 2022-01-10; 预印本日期: 2022-01-17

* 通信作者: 郎丛妍 cylan@bjtu.edu.cn

基金项目: 国家自然科学基金项目 (62072027, 61872032, 62076021); 北京市自然科学基金项目 (4202057, 4202058, 4202060)

Supported by: National Natural Science Foundation of China (62072027, 61872032, 62076021); Beijing Municipal Natural Science Foundation (4202057, 4202058, 4202060)

developing dramatically. Recent vehicle re-identification methods can be segmented into two categories: 1) feature learning and 2) ranges-metric learning. To enhance the re-identification performance, existing methods are restricted by multi-scale contextual information loss and lacking ability of discriminative feature selection because most feature learning and ranges-metric learning approaches are based on vehicle visual features from initial views captured or the additional information of multiple vehicles attributes, spatio-temporal information, vehicle orientation. So, we develop a novel global and spatial multi-scale contexts fusion method for vehicle re-identification (GSMC). **Method** Our method is focused on the global contextual information and the multi-scale spatial information for vehicle re-identification task. Specifically, GSMC is composed of two main modules: 1) a global contextual selection module and 2) a multi-scale spatial contextual selection module. To extract global feature as the original feature, we use residual network as the backbone network. The global contextual selection module can be used to divide the original feature map into several parts along the spatial dimension, and the 1×1 kernel size convolution layer is employed for the dimension-reducing. The softmax layer is used to obtain the weight of each part, which represents the contribution of different parts to the vehicle re-identification task. To extract more discriminative information of vehicles, the feature-optimized is melted into original feature. Additionally, to obtain a more discriminative feature representation, the feature outputs are divided into multiple horizontal local features in this module and these local features are used to replace global feature classification learning. In order to alleviate the feature loss in the boundary area, local features-adjacent have an intersection with a length of 1. What is more, the multi-scale spatial contextual selection module for GSMC is introduced to obtain multi-scale spatial features via different down-sampling, and, to generate the foreground feature response map of the vehicle image, this selected module can be used to optimize those multi-scale features, which can enhance the perception ability of GSMC to the vehicle's spatial location. To enhance the effect of the foreground, an adaptive larger weight can be assigned to the vehicle. To select more robust spatial contextual information, a smaller weight is assigned to the background for alleviating the interference of background information. Finally, as the final feature representation of the vehicle, our approach can fuse the features in the context of the global contextual selection module and the multi-scale spatial contextual selection module. In order to obtain a fine-grained feature space, GSMC is used for the label-smoothed cross-entropy loss and the triplet loss to improve its learning-coordinated ability overall. In the training process, in order to make the model have a faster convergence rate, our model is implemented in the first 5 epochs to keep the model stable in terms of the warm-up learning strategy. **Result** To valid the effectiveness of our approach proposed on vehicle re-identification task, we evaluate our model with some state-of-the-art methods on two public benchmarks of those are VehicleID and vehicle re-identification-776 (VeRi-776) datasets. The quantitative evaluation metrics are related to mean average precision (mAP) and cumulative matching curve (CMC), which can represent the probability that the image of the probe identity appears in the retrieved list. We carry out a series of comparative analysis with other methods, which are additional non-visual information and the multi-view leaning methods. Our analysis is demonstrated that it can surpass PNVR (part-regularized near-duplicate vehicle re-identification) by a large margin significantly. On the VehicleID dataset, we improve the rank-1 by 5.1%, 4.1%, 0.8% and the rank-5 by 4.4%, 5.7%, and 4.5% on three test subsets of different size. Compared to PNVR on the VeRi-776 dataset, GSMC gains 2.3% and 2.0% performance improvements of each in terms of mAP and rank-1. The lower ranks of CMC accuracy illustrates that our method can promote the ranking of rough multi-view captured vehicle images. Furthermore, we use re-ranking strategy as a post processing step over the VeRi-776 dataset and the results have significant improvement in mAP, rank-1 and rank-5 scores. At the same time, to verify the necessity of different modules in the proposed model, we design an ablation experiment to clarify whether a single branch can extract discriminative feature or not and the effectiveness of the feature fusion of the two modules is optimized as well. When different modules are added sequentially, the combination can realize the performance improvement by a large margin on mAP, rank-1 and rank-5. We are able to conclude that our proposed module is effective and can be capable to pull the images of same vehicle identity closer and push the different vehicles far away through the comparative analysis in relevant to the experimental results, the attention heat map visualization and the foreground feature response map. **Conclusion** To resolve the problem of vehicle re-identification, we develop an optimized model in terms of a global contextual selection module and a multi-scale spatial contextual selection module. The proposed model has its potential effectiveness in the extensive experiments in comparison with two popular public datasets mentioned.

Key words: vehicle re-identification; deep learning; local discriminative features; feature selection; multi-scale spatial features

0 引言

车辆重识别是智能交通系统的一部分。其从查询集中任选一幅车辆图像作为查询对象,旨在从图库集中检索出由不同角度摄像机拍摄的与查询对象相同的车辆图像。该任务应用于车辆检索(Liu等,2020)、交通流分析(Jiang等,2018)和车辆跟踪(Xu等,2020)等智能交通系统研究领域中,已成为当前计算机视觉和多媒体分析领域的热门研究课题之一。

早期的车辆重识别方法(Liu等,2016b;Jeng和Chu,2013)主要关注于手工设计的底层特征,诸如颜色和纹理等,并设计一定的计算规则进行车辆重识别任务的研究。此类方法侧重于对车辆图像某方面特征的抽取,对特定任务表现出不错的识别性能。然而这些底层特征容易受到外界环境的干扰,例如光照、遮挡和视角变化等,使这类算法的识别精度受限。随后,一些算法(Liu等,2016c;Liu等,2018a)根据车牌是车辆身份唯一标识这一特性,通过检测图像中车牌的位置,将车牌字符从整幅图像中分割出来,并进一步根据车牌号进行车辆重识别。此类方法在一定程度上缓解了早期方法对手工设计底层特征的依赖。然而,由于车牌像素模糊以及车牌被伪造、遮挡和涉及隐私安全等诸多方面因素,给搜索目标车辆带来极大挑战。

受益于神经网络强大的自主学习能力及对区分力特征的表征能力,深度卷积神经网络(Wang等,2017;Zhou和Shao,2018)在车辆重识别任务上得到了广泛运用,推动了车辆重识别技术的进一步发展。其中,一些算法考虑从车辆的局部位置中提取关键信息。He等人(2019a)利用局部正则化方法,选择性地关注更有区分力的关键区域,例如年检标志粘贴的位置、车内的挂饰等,忽略其他不相关的信息。Liu等人(2018a)将整个图像水平分割成3个部分,并从3个重叠的局部区域中提取相应的区域特征。Chen等人(2020)提出一个专注于不同部位的注意力网络,通过预测不同视角下对象车辆的掩膜,提取对应部分的判别特征。潘海鹏等人(2021)提出一种基于注意力与多尺度融合学习的车辆重识别方法,在多尺度下提取并融合浅层细节信息和深层语义信息。相比于早期对车辆重识别算法的研究,上述方法大幅提升了研究领域的整体识别性能,但在车辆重识别任务中会遇到类内差异性、类间相似性、分辨率低和遮挡等多项挑战,如图1所示,即相机拍摄角度的不同会导致同一辆车在不同视角下形成自身差别,不同车辆在相同视角下因外观相同产生类间相似,或视角被物体遮挡使得拍摄的图像不完整等,这使得现有方法未能充分对多尺度上下文信息进行有效地抽取,不能充分表征与运用丰富的细粒度特征,使这些挑战影响了最终车辆重识别的精度。



图1 车辆重识别任务中的挑战

Fig. 1 The challenges of vehicle re-identification task

((a) intra-class difference; (b) inter-class similarity; (c) low resolution; (d) occlusion)

为了缓解上述问题,本文提出一种融合全局与空间多尺度上下文信息的车辆重识别方法(global and spatial multi-scale contexts fusion for vehicle re-identification, GSMC),利用全局与多尺度空间上下文信息,对判别特征进行有效的选择优化。GSMC

包含全局上下文特征选择和多尺度空间上下文特征选择两个模块。前者首先对车辆判别特征进行划分,然后根据每个部分对最终识别任务不同的贡献程度,动态地学习其对应的权重度量因子,并以此优化车辆判别特征。而后者从多尺度思想出发,对得

到的判别特征通过下采样的方式获得多尺度特征,然后选择性地集成来自多级特征的空间上下文信息,提升模型对于车辆空间位置特征的感知能力。通过两个模块的相互协作,使 GSMC 能够分别从全局与空间局部方面获得强判别力的上下文信息。本文主要贡献如下:1)提出一种融合全局与空间多尺度上下文信息的车辆重识别方法,同时优化全局与空间局部上下文信息,获得更加鲁棒的特征表示。2)设计全局上下文特征选择及多尺度空间上下文特征选择两个模块。同时考虑图像的全局及多尺度空间上下文信息,学习车辆图像的前景特征响应图,以生成强判别力的车辆识别特征。3)在公开数据集 VehicleID 及 VeRi-776 (vehicle re-identification-776)上进行实验,验证了提出的 GSMC 在车辆重识别问题上的性能优于主流算法的性能。

1 相关工作

随着深度学习技术在各领域的应用与发展,基于深度学习的车辆重识别方法备受关注。按照学习方法的不同,车辆重识别方法分为基于表征学习的方法和基于度量学习的方法。

基于表征学习的方法在训练网络时未考虑图像间的相似度,将车辆重识别任务看做分类问题。例如,Shen 等人(2017)将车辆图像和其时空信息构造为一个状态,利用链式马尔可夫模型,生成多条候选的视觉—空间—时间路径,之后使用孪生卷积神经网络计算候选路径中近邻图像之间的相似程度,最后使用长短期记忆网络,以该条候选路径作为先验知识,判断该查询对是否具有相同的 ID (identity),有效规范车辆重识别的结果。Khorramshahi 等人(2019)为了解决车辆视角变化带来的影响,对车辆的关键部位定义并将关键点分为 4 组,使用注意力机制提取车辆具有辨别力的局部特征。Wang 等人(2017)提出一个包含方向不变特征嵌入和时空正则化两个模块的方法,首先选取车身上具有辨识性的 20 个候选区域,得到车辆关键点对应的映射图,再基于车辆的 4 个面,将这些关键点聚合成 4 个候选区域掩码,利用原始的输入图像和生成的候选区域掩码,可以获得一个全局特征和 4 个局部特征。最后,这些特征通过特征聚合模块进行融合,输出一个方向不变特征向量。在时空正则化模块中,利用

时空信息对检索结果进一步优化。

基于度量学习的方法旨在通过网络模型学习出图像间的相似度。例如,Liu 等人(2016a)设计了一个双分支的混合差分网络,将输入的车辆图像映射到具有可区分性的特征空间,同时考虑到传统三元组损失的不足,提出了成对簇损失函数,缩小正样本对之间的距离并增大负样本对之间的距离。构建了一个真实监控场景下高质量的车辆重识别数据集 VeRi-776,并使用车辆的外观特征和额外的时空信息作为输入,送入有两个网络输入的孪生网络,度量输入图像之间的相似性。Chu 等人(2019)将车辆重识别分为相似视角和不同视角,通过执行空间内约束与跨空间约束,可以在相似视角图像干扰下检索不同视角的图像。

为了学习到更有区分性的特征映射空间,许多工作联合这两种学习方法一起优化网络,用于获得更具判别力的特征作为最终的特征表示,从而进一步提高算法性能。Zhou 和 Shao(2017)设计了一个 XVGAN (cross-view generative adversarial network),采用对抗学习方案从单幅图像生成多个视角下的车辆图像或特征,可以在一定程度上解决视角变化的挑战,但是可能很难区分外观非常相似的不同车辆。Zhou 和 Shao(2018)使用视图感知注意力模型获取车辆在不同视角下的局部特征,再利用对抗训练方法,设计了一个多视角生成网络,可以从输入图像的单一角度特征推断生成一个包含多角度信息的车辆特征,以实现全方位的外观特征对比。

一些研究重点关注车辆的关键区域,从车辆局部区域获得更具区分性的特征。Chen 等人(2019)设计了一种先分割后聚合的 PRN (partition and reunion network),将特征图分别在空间和通道维度上进行了 4 等分割,并将最后的全连接层输出的子特征向量进行拼接,作为交叉熵损失函数的输入,通过这种先分割再聚合的训练策略,该模型不使用额外的标注信息就可以捕捉到丰富的车辆局部特征,但是缺少对不同尺度大小关键区域的选择。Qian 等人(2019)用水平划分的方式获得局部特征,可以有效抵消水平视角变化引起的不利影响,同时,在车辆属性标签的监督下从特征图中提取全局特征,帮助训练网络。邱铭凯和李熙莹(2021)提出一个基于细节感知的判别特征学习模型,设计了一个指导式的车辆局部特征提取流程,将局部特征与骨干网络提取的

全局特征联合作为车辆特征。Liu 等人(2018a)使用划分的方式进行分块,从一系列局部区域中提取特征,同时学习全局外观和局部区域的特征,但未考虑不同部分对最终的重识别结果的重要程度不同。

不同于上述研究,本文方法不仅为提取的特征动态地分配权重度量因子,而且考虑对多尺度特征融合,更好地实现前景目标与背景的分离,以提升模型性能。

2 GSMC 模型

2.1 模型的整体结构

图2展示了GSMC的整体网络结构,包含全局上下文特征选择和多尺度空间上下文特征选择两个模块。此外,引入基于规则划分模块以获得局部特征(图2右上角)。给定一幅车辆图像,GSMC首先采用预训练好的基础网络提取其初始化的车辆特征。类似于主流算法(Chen等,2020;Khorramshahi等,2019),GSMC利用ResNet-50-ibn-a(Pan等,

2018)作为特征提取的基础网络。不同于原始ResNet-50-ibn-a,GSMC去除了基础网络的全连接层及最后一个池化层。对于ResNet-50-ibn-a最后一个卷积层输出的特征,其编码了丰富的高层语义信息,能够较好地捕捉到车辆的全局与局部信息,因此,在此判别力特征之上,GSMC首先设计一个全局上下文特征选择模块,对判别特征进行规则划分,从整体上学习每个部分的重要性分数,以此增强特征响应大的关键区域,提取具有判别力的细节特征。针对优化后的全局特征,受多尺度思想的启发,GSMC进一步设计一个多尺度空间上下文特征选择模块,对优化后的特征采用多尺度划分,选择性地集成来自多级特征的空间上下文信息,获得可以区分前景与背景的特征响应图,去除噪声以及冗余信息的影响。通过两个模块的相互协作,使GSMC能够分别从全局与空间局部方面获得强判别力的上下文信息,不仅可以挖掘车辆不同部位的细粒度判别信息,也可以从多尺度的空间特征中学习前景特征,获得更加鲁棒的特征表示。

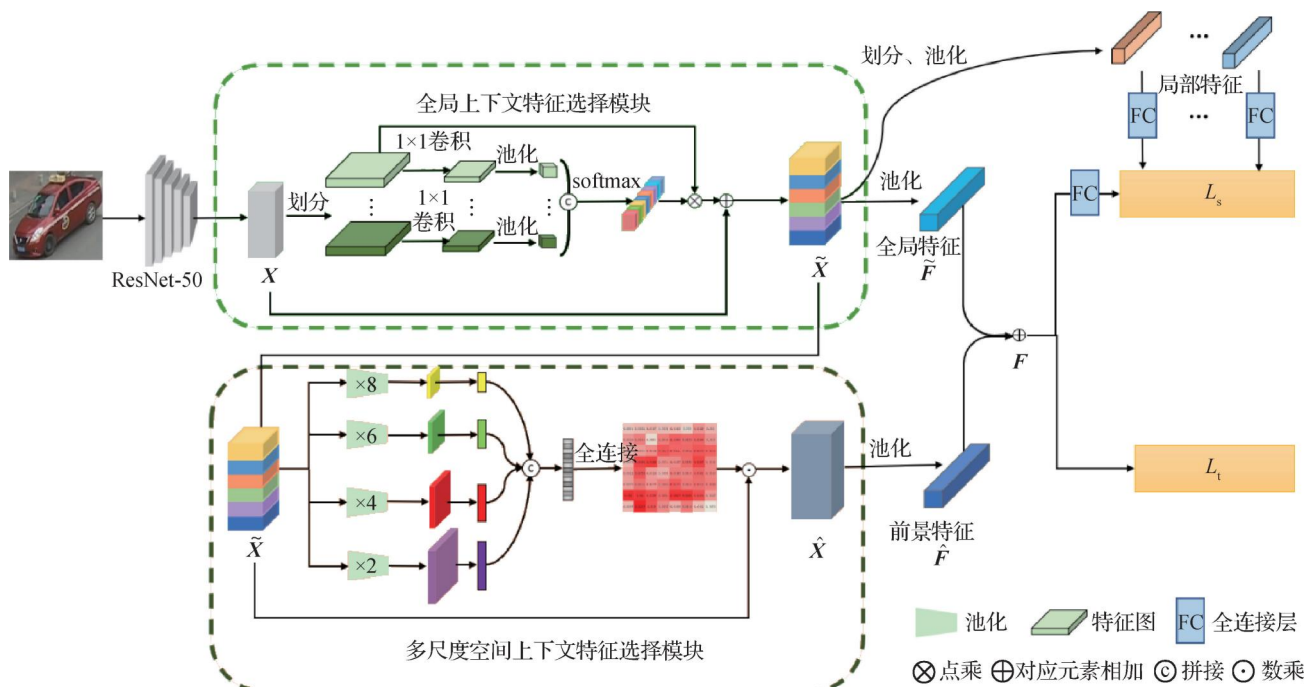


图2 融合全局与多尺度空间上下文信息的车辆重识别模型的整体框架结构示意图

Fig. 2 The overall structure of the global and spatial multi-scale contexts fusion for vehicle re-identification

2.2 全局上下文特征选择模块

图2的上半部分展示了全局上下文特征选择模块的整体结构示意图。本文定义ResNet-50-ibn-a最后一个卷积层输出的特征为 $X \in \mathbf{R}^{H \times W \times C}$,其中, H

和 W 分别表示特征图的高度与宽度, C 代表特征图的通道数目。针对总体判别特征 X ,该模块将其按高度方向有序地划分为 N 部分,即 $X = \{x_i\}_{i=1}^N$,且 $x_i \in \mathbf{R}^{H \times W \times C}$ 。对于每部分特征 x_i ,本模块首先采用

一个 1×1 的卷积将其通道数目降为 1;其次,使用全局平均池化操作获得对应部分的特征权重 c_i 。该过程表示为

$$c_i = P(\text{Conv}_{1 \times 1}(\mathbf{x}_i)) \quad (1)$$

式中, P 表示全局平均池化(global average pooling)操作, $\text{Conv}_{1 \times 1}(\cdot)$ 代表 1×1 的卷积函数。此外,为了度量每个部分特征的重要度,本模块连结 N 个部分学习到的特征权重因子,并采用全局归一化操作,通过各部分权重值之间的相互交流,更新对应部分生成的权重值。该更新过程可表示为

$$c_i \leftarrow \zeta(c_i) = \frac{\exp(c_i)}{\sum_{j=1}^N \exp(c_j)} \quad (2)$$

式中, $\zeta(\cdot)$ 代表 softmax 函数, $\exp(\cdot)$ 为指数函数。利用学习到的权重因子,本模块将其与相应的每个部分特征进行点积相乘操作,以增强各部分对应的初始化判别特征。随后,本模块将这些增强后的特征拼接在一起,并采用残差机制,将拼接后的特征与初始化判别特征逐项相加,使网络模型能够关注车辆的关键区域,提升模型对于细粒度判别信息的感知能力,从而获得更好的特征表征 $\tilde{\mathbf{X}}$ 。该整合过程可以表示为

$$\tilde{\mathbf{X}} = (\parallel_{i=1}^N \mathbf{x}_i \cdot c_i) \oplus \mathbf{X} \quad (3)$$

式中, $\parallel$ 表示拼接函数, $\oplus$ 表示残差机制中的对应元素相加操作。

2.3 多尺度空间上下文特征选择模块

在给定由全局上下文特征选择模块生成的全局化判别特征 $\tilde{\mathbf{X}}$ 的基础上,为了可以通过网络学习正确地得到该图像的车辆标签,一种直接的方式是对判别特征进行池化及全连接操作,直接输出车辆类别,此方式需要数据集有清晰完整的车辆图像,然而面对数据集中检索对象存在的遮挡、图像分辨率低等问题时,这种直接分类方式因输出的空间特征受背景影响导致分类性能受限。为了减少这种背景噪声以及冗余信息的影响,GSMD 进一步设计了一个多尺度空间上下文特征选择模块,通过对多尺度特征融合,对前景目标与背景分类,生成前景特征响应图,更好地实现前景目标与背景的分隔,从而提升 GSMD 的整体识别性能。

图 2 的下半部分展示了多尺度空间上下文特征选择模块的整体结构。该模块受多尺度思想启发,采用不同下采样比率 $\{2, 4, 6, 8\}$, 将 $\tilde{\mathbf{X}}$ 分发到不同

的特征尺度空间 $\{\tilde{\mathbf{X}}_j\}_{j=1}^M$, 其中, $\tilde{\mathbf{X}} \in \mathbf{R}^{H \times W \times C}$, M 表示不同尺度特征的个数。对于每一个采样率下车辆的尺度特征 $\tilde{\mathbf{X}}_j \in \mathbf{R}^{H_j \times W_j \times C}$ (H_j 和 W_j 分别代表尺度特征图的高度与宽度), 首先,对其进行维度变换,即 $\tilde{\mathbf{X}}_j \leftarrow \tilde{\mathbf{X}}_j \in \mathbf{R}^{(H_j \times W_j) \times C}$ 。然后,将变换后的不同尺度的空间特征聚合成 $\tilde{\mathbf{F}} = \{\tilde{\mathbf{X}}_j\}_{j=1}^M$, 以获得更为丰富的语义信息。利用此整合特征,模块引入空间增强机制,旨在通过对整合特征的空间上下文信息进行加强的同时,自适应地学习其对应的前景特征响应图 $\mathbf{P}$ 。此过程可以表示为

$$\mathbf{P} = T(\sigma(W(\tilde{\mathbf{F}}))) \quad (4)$$

式中, W 表示全连接层的网络参数,函数 $T(\cdot)$ 表示维度变换操作, $\sigma(\cdot)$ 代表 sigmoid 函数。根据学习到的前景特征响应图,本模块对判别特征 $\tilde{\mathbf{X}}$ 进行空间位置上下文的增强,获得对应前景部分的增强特征 $\hat{\mathbf{X}}$, 即

$$\hat{\mathbf{X}} = \tilde{\mathbf{X}} \odot \mathbf{P} \quad (5)$$

式中, $\odot$ 表示数乘操作。该模块利用多尺度特征之间的信息交互,能够自适应地为车辆分配较大的权重来增强前景的效果,为背景赋予较小的权重来减弱噪声等信息的干扰,以选择更加鲁棒的空间上下文信息。

2.4 模型训练

为了获得细粒度特征空间,得到更有区分力的特征表示,最终特征 $\mathbf{F}$ 可以表示为

$$\mathbf{F} = \tilde{\mathbf{F}} + \hat{\mathbf{F}} \quad (6)$$

式中, $\tilde{\mathbf{F}}$ 是全局上下文特征选择模块中提取的特征 $\mathbf{X}$ 经由全局平均池化、批归一化得到的, $\hat{\mathbf{F}}$ 是由多尺度空间上下文特征选择模块得到的特征 $\hat{\mathbf{X}}$ 转换后得到的。将得到的特征 $\mathbf{F}$ 输入到全连接层,用标签平滑的交叉熵损失函数 L_s 进行优化(Müller 等, 2020),这一过程可以表示为

$$L_s = - \sum_{i=1}^B \tilde{y}_i^t \log(Z_i[t]) + \sum_{j=1, j \neq t}^V \tilde{y}_i^j \log(Z_i[j]) \quad (7)$$

式中, B 与 V 分别表示数据集中的图像总数与类别总数, $\tilde{y}_i^j$ 表示将传统 one-hot 编码的标签向量平滑后,车辆图像 i 对应第 j 类的概率, t 表示该图像的真实标签, $Z_i[t]$ 表示车辆图像 i 的特征 $\mathbf{F}$ 经过全连接层后,对应第 t 类的输出概率。类似地, $Z_i[j]$ 是第 j 类的输出概率。此外,对全局上下文特征选择模块输出的特征 $\tilde{\mathbf{X}}$ 等分为多个水平局部特征 $\{\tilde{\mathbf{X}}_i^l\}$,

$i \in \{1, 2, 3, 4, 5\}$, 利用局部特征而不是整幅图像的全局特征分类学习。为防止边界区域特征丢失, 相邻的局部特征有长度为1的交叉部分, 然后利用标签平滑的交叉熵损失函数 L'_s 训练模型。具体表示为

$$L'_s = \sum_{i=1} L_s^i \quad (8)$$

式中, L_s^i 表示局部特征 $\tilde{X}_i^l$ 对应的损失。

考虑到交叉熵损失函数虽然可以减小类间差异, 但不擅长于使同一类的特征变得更加紧凑, 不足以有效地学习鲁棒的特征表示。因此, 利用三元组损失函数 (Hermans 等, 2017) L_t 在特征空间拉近正样本对的类内距离, 增大负样本对的类间距离, 进一步提高车辆重识别性能。最终的损失函数可以表示为

$$L = L_s + L'_s + L_t \quad (9)$$

通过多个损失函数的联合优化, 可以增强模型的学习能力, 引导网络学习更好的特征嵌入空间, 在训练阶段生成车辆更具判别力的特征表示。

3 实验

3.1 实验设置

实验采用在 ImageNet 数据集 (Deng 等, 2009) 上预训练的 ResNet-50 (Pan 等, 2018) 作为基础网络。在模型训练过程中, 为了使模型具有较快的收敛速度, 在前5个 epoch 内使用预热学习策略保持模型的稳定。训练参数设置如下: epoch 为50, 学习率为0.0001, 且在第15个和第40个 epoch 时衰减, 优化器为 Adam, 动量为0.9, 批处理大小为64, 车辆图像尺寸为 256×256 像素。对于每个批处理, 训练样本来自16辆不同ID(类别)的车辆, 每个ID包括4幅随机选择的图像。

3.2 数据集及评价指标

为了验证提出模型的有效性, 在主流的公开数据集 VehicleID (similar vehicles) (Liu 等, 2016a) 和 VeRi-776 (Liu 等, 2016; Liu 等, 2018a) 上进行实验, 采用 rank-1, rank-5 和 mAP (mean average precision) 评价指标评测模型的性能。rank-1 和 rank-5 分别表示结果列表中排名前1和前5个结果的检索准确率, mAP 表示平均准确率。

VehicleID 数据集由多个视角互不重叠的监控

摄像头在白天拍摄的车辆图像组成, 每辆车的拍摄视角只有前视和后视, 共221 763幅26 267辆车的图像, 平均每辆车8.44幅图像, 其中10 319辆车的90 196幅图像标注了车型等信息。数据集中的车辆共7种颜色, 250种型号。数据集分为训练集与测试集。训练集包含13 134辆车的110 178幅车辆图像。测试集根据车辆图像的数目又分为小规模、中等规模和大规模测试子集, 候选集由每个ID的1幅车辆图像组成, 剩余的图像构成查询集, 每个子集包含的车辆ID数分别为800、1 600和2 400个。由于候选集中只包含每个ID的1幅车辆图像, 因此只使用 rank-1 与 rank-5 作为该数据集的性能评价指标。

VeRi-776 数据集是在 VeRi (Liu 等, 2016c) 数据集上扩充标注后得到的, 数据集中的车辆图像由218个不同位置的摄像头在不同环境中拍摄, 图像具有不同的视角、遮挡和光照等。数据集包括776辆车的近50 000幅图像, 其中, 576辆车的37 778幅图像作为训练集, 200辆车的11 579幅图像作为测试集。数据集中所有图像都标注了车辆ID、车型以及车辆颜色, 共9种车型, 10种颜色。此外, 标注了所有车辆的轨迹、时空关系和各相机之间距离。选择 rank-1、rank-5 和 mAP 作为该数据集的性能评价指标。

3.3 与主流算法的性能比较

为验证本文算法的性能, 在两个公共车辆重识别数据集上与 VAMI (viewpoint-aware attentive multi-view inference) (Zhou 和 Shao, 2018)、AAVER (adaptive attention model for vehicle re-identification) (Khorramshahi 等, 2019)、PNVR (part-regularized near-duplicate vehicle re-identification) (He 等, 2019a)、RAM (region-aware deep model) (Liu 等, 2018a) 和 SPAN + CPDM (semantics-guided part attention network + co-occurrence part-attentive distance metric) (Chen 等, 2020) 等主流算法进行比较。

3.3.1 VehicleID 数据集实验结果

在 VehicleID 数据集上, 选择大规模测试子集进行实验, 因为其含有更多难以区分的车辆图像, 在该测试集上的准确度能更加充分地反映网络模型的鲁棒性和泛化能力。实验时, Baseline 模型同时使用标签平滑的交叉熵损失函数和三元组函数, 以 ResNet-50-ibn-a 为基础网络。表1为本文模型与主流车辆重识别方法在3个测试子集上的对比结果。相比于主流方法和 Baseline, 本文模型在大规模测试

子集上取得了最优的结果, rank-1 和 rank-5 指标分别为 75.0% 和 90.9%。与使用了多视角学习方法的 VAMI 相比,本文方法在大规模测试子集上的评测结果 rank-1 和 rank-5 分别提高了 27.7% 和 20.7%。AAVER 方法可以将全局特征与通过使用自适应关键点和方向的注意力机制获得的细节特征进行融合,但是忽略了不同尺度下的细节信息,致使学习到的局部特征较少。与 AAVER 方法相比,本文提出的融合全局与空间多尺度上下文信息的车辆重识别方法可以得到更多的细粒度判别特征,rank-1 和 rank-5 指标分别提高了 11.5% 和 5.3%。与沿空间维度、通道维度划分的 PRN 相比,本文方法不仅可以得到每个部分的重要程度,还可以避免背景对

全局特征造成的影响,在 VehicleID 大规模测试子集上的 rank-1 与 rank-5 分别提高了 3.5% 和 2.5%。SAN 通过水平划分的方式提取全局上下文信息,以学习有效的特征嵌入,但是忽略了不同部分的重要程度。与 SAN 相比,本文方法在大规模测试子集上的 rank-1 与 rank-5 分别提升了 0.6% 和 2.3%。通过实验分析可知,本文模型在不需要额外标注信息的条件下,利用全局上下文特征选择模块和多尺度空间上下文特征选择模块可以明显提升车辆重识别的效果,不仅能获取车辆不同部位的重要度,而且可以区分前景目标与背景,增强特征的鲁棒性。值得注意的是,对于 VehicleID 数据集,候选集中每个 ID 只有 1 幅车辆图像,因此,重排序是不适用的。

表 1 VehicleID 数据集上与主流方法性能对比

Table 1 Performance comparison with state-of-the-art methods on VehicleID dataset

方法	小规模测试子集		中等规模测试子集		大规模测试子集	
	rank-1	rank-5	rank-1	rank-5	rank-1	rank-5
DRDL(Liu 等,2016a)	0.490	0.735	0.428	0.668	0.382	0.616
OIFE(Wang 等,2017)	—	—	—	—	0.670	0.829
VAMI(Zhou 和 Shao,2018)	0.631	0.832	0.528	0.751	0.473	0.702
AAVER(Khorramshahi 等,2019)	0.747	0.938	0.686	0.900	0.635	0.856
RAM(Liu 等,2018a)	0.752	0.915	0.723	0.870	0.677	0.845
GS-TRE(Bai 等,2018)	0.759	0.842	0.748	0.836	0.74	0.827
PNVR(He 等,2019a)	0.784	0.923	0.750	0.883	0.742	0.864
PRN(Chen 等,2019)	0.789	0.948	0.749	0.920	0.715	0.884
SAN(Qian 等,2019)	0.797	0.943	0.784	0.913	0.756	0.883
Baseline	0.682	0.804	0.678	0.777	0.666	0.761
本文	0.835	0.967	0.791	0.940	0.750	0.909

注:加粗字体表示各列最优结果,“—”表示该算法没有实验结果。

3.3.2 VeRi-776 数据集实验结果

表 2 为本文 GSMC 模型与主流算法在 VeRi-776 数据集上的性能比较结果。可以看出,提出模型的性能优于使用额外非视觉特征的方法 FACT + SNN + STR(fusion of attributes and color features + siamese neural network + spatio-temporal relations) (Liu 等,2016c)、OIFE + ST(orientation invariant feature embedding + spatial temporal regularization) (Wang 等,2017)、RAM(Liu 等,2018a)和解决跨视角偏差的方法 VAMI(Zhou 和 Shao,2018)、PNVR(He 等,2019a)和 AAVER(Khorramshahi 等,2019)。与同样

使用全局上下文信息的 RAM(Liu 等,2018a)相比,本文方法在没有引入车辆额外信息的条件下,mAP、rank-1 和 rank-5 指标分别提升了 2.3%、3.0% 和 1.0%。与结合了全局特征和全局上下文特征等的 SAN 相比,本文方法通过提出的多尺度空间上下文特征选择模块,mAP、rank-1 和 rank-5 指标分别提升了 16.1%、7.7% 和 4.1%。与性能第 2 的 PNVR 相比,提出的模型的 mAP 和 rank-1 指标分别实现了 + 3.1% 和 + 2.0% 的明显增益。特别地,当使用 Re-ranking(Zhong 等,2017)作为 VeRi-776 数据集的后处理步骤时,在 k-互反编码(Zhong 等,2017)重排

序算法处理之后, mAP 从 77.6% 增加到 80.8%, 模型的整体识别性能明显提升。实验结果表明, 提出的模型不仅可以对每个局部特征进行学习, 得到每个局部区域的重要性分数, 而且可以整合不同尺度的特征, 对特征进行选择, 避免背景或遮挡对全局特征造成影响, 可以有效抵消水平视角变化带来的不利影响。通过实验性能的对比, 证明了本文模型的有效性。

表2 不同方法在 VeRi-776 数据集上的性能对比
Table 2 Comparison of performance of different methods on VeRi-776 dataset

方法	mAP	rank-1	rank-5
FACT + SNN + STR (Liu 等, 2016c)	0.278	0.614	0.788
OIFE + ST (Wang 等, 2017)	0.514	0.683	0.897
NuFACT (Liu 等, 2018b)	0.485	0.768	0.914
VAMI (Zhou 和 Shao, 2018)	0.613	0.859	0.918
RAM (Liu 等, 2018a)	0.615	0.886	0.940
VANet (Chu 等, 2019)	0.663	0.898	0.960
AAVER (Khorramshahi 等, 2019)	0.611	0.901	0.947
SAN (Qian 等, 2019)	0.725	0.933	0.971
SPAN + CPDM (Chen 等, 2020)	0.689	0.940	0.976
PNVR (He 等, 2019a)	0.743	0.943	0.987
Baseline	0.715	0.933	0.971
本文	0.776	0.963	0.981
Re-ranking	0.808	0.968	0.976

注: 加粗字体表示各列最优结果。

3.4 消融实验

为了证明提出模型各子模块的有效性, 在基于交叉熵损失和三元组损失联合优化的 Baseline 网络的基础上, 通过逐步添加全局上下文特征选择模块和多尺度空间上下文特征选择模块, 在 VeRi-776 数据集上构建消融实验, 得到各模块对算法的提升效果。此外, 为了进一步证明这两个模块的有效性, 增加了使用通道维和空间维注意力机制的消融实验。实验结果表明, 得益于所提模块能够更有效地选择车辆关键特征, 本文方法的 mAP、rank-1 和 rank-5 指标分别提升了 5.6%、2.1% 和 0.6%。不同模块组合得到的实验结果如表 3 所示。其中, 第 1 行是本文的 Baseline 网络框架, GCS (global contextual selection module) 表示全局上下文特征选择模

块, MSCS (multi-scale spatial contextual selection module) 表示多尺度空间上下文特征选择模块, PART (part based partition module) 表示基于规则划分模块, Attention 表示基于通道维和空间维的注意力机制。

表3 不同模块组合得到的实验结果
Table 3 Experimental results obtained by combining different modules

方法	mAP	rank-1	rank-5
Baseline	0.715	0.933	0.971
Baseline + Attention	0.720	0.942	0.975
Baseline + GCS	0.724	0.951	0.979
Baseline + MSCS	0.725	0.945	0.977
Baseline + GCS + MSCS	0.745	0.952	0.979
Baseline + GCS + MSCS + PART	0.776	0.963	0.981

注: 加粗字体表示各列最优结果。

3.4.1 全局上下文特征选择模块

将加入全局上下文特征选择模块的模型与 Baseline 相比, 由表 3 第 3 行与第 1 行可得, mAP、rank-1 和 rank-5 指标分别提高了 0.9%、1.8% 和 0.8%。从表 3 第 4 行与第 5 行可以看出, 基于多尺度空间上下文特征选择模块引入全局上下文特征选择模块后, mAP、rank-1 和 rank-5 指标又分别提高了 2.0%、0.7% 和 0.2%。这些结果都证明了全局上下文选择模块的有效性, 可以学习到细粒度的判别信息。

图 3 是 4 组加入全局上下文特征选择模块前后的特征注意力热图。每组左列是原始车辆图像, 中间列是加入该模块前的注意力热图, 右列是经过该模块后输出的注意力热图, 颜色越深代表权重越大, 对识别车辆身份越重要。通过该模块使网络可以自动关注重要的区域, 提高车辆重识别的性能。在不使用额外标注信息的情况下, 该模块能够有效捕捉丰富的局部特征, 对全局特征进行细节上的补充。

3.4.2 多尺度空间上下文特征选择模块

在实验的训练阶段, 使用带有 ID 注释的车辆图像进行训练, 通过损失函数进行监督, 获取能够区分前景与背景的分类器。特征响应值越接近 1, 表示前景特征的概率越大; 越接近 0, 表示背景的概率越大。在测试阶段, 多尺度空间上下文特征选择



图3 进入全局上下文特征选择模块前后的特征注意力热图

Fig. 3 Feature attention heatmaps before and after entering the global contextual selection module

((a) vehicle ID-1; (b) vehicle ID-2; (c) vehicle ID-3; (d) vehicle ID-4)

模块可直接预测输入车辆图像的前景特征响应图,而无需通过人工标注。通过比较表3第4行与Baseline的实验结果,使用了多尺度空间上下文特征选择模块的网络在 mAP、rank-1 和 rank-5 上分别提高了 1.0%、1.2% 和 0.6%。将表3第5行与第3行的结果对比,可以看出,该模块在 mAP 指标上增加了 2.1%。这些都表明该模块对特征进行了过滤选择,为车辆特征分配了较大的权重来增强前景的效果,为背景赋予较小的权重来减小背景的影响,获得车辆更加鲁棒的全局特征描述。图4显示了4组

使用多尺度空间上下文特征选择模块生成车辆图像的前景特征响应图。可以看到,使用该模块可以准确地检测到车的部分。

除了比较不同模块对车辆重识别实验结果的影响外,使用规则划分的方法验证实验效果,该方法可以对每一个局部特征而不是利用整幅图像的全局特征进行分类学习。比较表3第6行与第5行的实验结果,该方法的 mAP、rank-1 和 rank-5 指标又分别提高了 3.1%、1.1% 和 0.2%,相比较于 Baseline,在各评价指标上有了很大的性能提升。

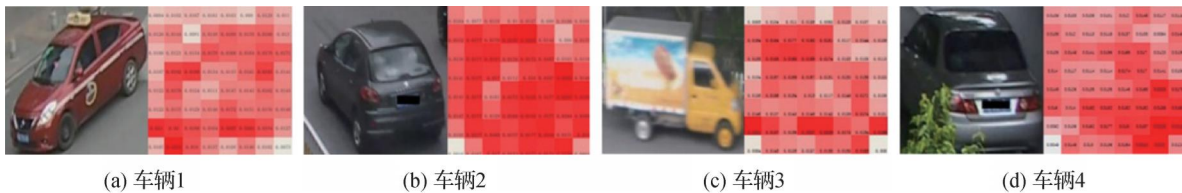


图4 由多尺度空间上下文特征选择模块得到的车辆前景特征响应图

Fig. 4 The vehicle foreground feature response maps obtained by the multi-scale spatial contextual selection module

((a) vehicle ID-1; (b) vehicle ID-2; (c) vehicle ID-3; (d) vehicle ID-4)

4 结 论

本文探讨了车辆识别任务面临的挑战,提出了一个新颖有效的车辆重识别网络。全局上下文特征选择模块动态地学习不同部分对应的权重度量因子,有效选择对车辆识别贡献大的关键区域,对判别特征进行优化。多尺度空间上下文特征选择模块自适应地为车辆分配较大的权值来增强前景对重识别准确率的影响,为背景赋予较小的权重去除冗余信息,旨在将前景目标与背景进行分离,选择出车辆特征,提升模型对于车辆空间位置特征的感知能力,获得车辆更加鲁棒的全局特征描述。在两个公开数据集上与主流车辆重识别算法相比,通过消融实验和对比实验,本文方法可以明显提升实验效果,证明了

本文方法的有效性。

本文算法利用车辆的视觉信息提取车辆的细粒度特征,未来希望进一步将局部区域之间的空间几何结构关系、车辆的属性信息(如颜色、车型等)融合到模型中,进一步提升算法各方面的性能。

参考文献(References)

- Bai Y, Lou Y H, Gao F, Wang S Q, Wu Y W and Duan L Y. 2018. Group-sensitive triplet embedding for vehicle reidentification. IEEE Transactions on Multimedia, 20(9): 2385-2399 [DOI: 10.1109/TMM.2018.2796240]
- Chen H, Lagadec B and Bremond F. 2019. Partition and reunion: a two-branch neural network for vehicle re-identification//Proceedings of 2019 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Long Beach, USA: IEEE: 184-192

- Chen T S, Liu C T, Wu C W and Chien S Y. 2020. Orientation-aware vehicle re-identification with semantics-guided part attention network//Proceedings of the 16th European Conference on Computer Vision. Glasgow, UK; Springer: 330-346 [DOI: 10.1007/978-3-030-58536-5_20]
- Chu R H, Sun Y F, Li Y D, Liu Z, Zhang C and Wei Y C. 2019. Vehicle re-identification with viewpoint-aware metric learning//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South); IEEE: 8281-8290 [DOI: 10.1109/ICCV.2019.00837]
- Deng J, Dong W, Socher R, Li L J, Li K and Li F F. 2009. ImageNet: a large-scale hierarchical image database//Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA; IEEE: 248-255 [DOI: 10.1109/CVPR.2009.5206848]
- He B, Li J, Zhao Y F and Tian Y H. 2019a. Part-regularized near-duplicate vehicle re-identification//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA; IEEE: 3992-4000 [DOI: 10.1109/CVPR.2019.00412]
- He L X, Wang Y G, Liu W, Zhao H, Sun Z and Feng J S. 2019b. Foreground-aware pyramid reconstruction for alignment-free occluded person re-identification//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South); IEEE: 8449-8458 [DOI: 10.1109/ICCV.2019.00854]
- Hermans A, Beyer L and Leibe B. 2017. In defense of the triplet loss for person re-identification [EB/OL]. [2021-06-03]. <https://arxiv.org/pdf/1703.07737.pdf>
- Jeng S T C and Chu L Y. 2013. Vehicle reidentification with the inductive loop signature technology. *Journal of the Eastern Asia Society for Transportation Studies*, 10: 1896-1915 [DOI: 10.11175/easts.10.1896]
- Jiang N, Xu Y, Zhou Z and Wu W. 2018. Multi-attribute driven vehicle re-identification with spatial-temporal re-ranking//Proceedings of the 25th IEEE International Conference on Image Processing (ICIP). Athens, Greece; IEEE: 858-862 [DOI: 10.1109/icip.2018.8451776]
- Khorramshahi P, Kumar A, Peri N, Rambhatla S S, Chen J C and Chellappa R. 2019. A dual-path model with adaptive attention for vehicle re-identification//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South); IEEE: 6131-6140 [DOI: 10.1109/ICCV.2019.00623]
- Li Y Q, Li Y H, Yan H F and Liu J Y. 2017. Deep joint discriminative learning for vehicle re-identification and retrieval//Proceedings of 2017 IEEE International Conference on Image Processing (ICIP). Beijing, China; IEEE: 395-399 [DOI: DOI: 10.1109/ICIP.2017.8296310]
- Lin W P, Li Y D, Yang X L, Peng P X and Xing J L. 2019. Multi-view learning for vehicle re-identification//Proceedings of 2019 IEEE International Conference on Multimedia and Expo (ICME). Shanghai, China; IEEE: 832-837 [DOI: 10.1109/icme.2019.00148]
- Liu H Y, Tian Y H, Yang Y W, Pang L and Huang T J. 2016a. Deep relative distance learning: tell the difference between similar vehicles//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA; IEEE: 2167-2175 [DOI: 10.1109/CVPR.2016.238]
- Liu X B, Zhang S L, Huang Q M and Gao W. 2018a. RAM: a region-aware deep model for vehicle re-identification//Proceedings of 2018 IEEE International Conference on Multimedia and Expo (ICME). San Diego, USA; IEEE: 1-6 [DOI: 10.1109/ICME.2018.8486589]
- Liu X C, Liu W, Ma H D and Fu H Y. 2016b. Large-scale vehicle reidentification in urban surveillance videos//Proceedings of 2016 IEEE International Conference on Multimedia and Expo (ICME). Seattle, USA; IEEE: 1-6 [DOI: 10.1109/ICME.2016.7553002]
- Liu X C, Liu W, Mei T and Ma H D. 2016c. A deep learning-based approach to progressive vehicle re-identification for urban surveillance//Proceedings of the 14th European Conference on Computer Vision. Amsterdam, the Netherlands; Springer: 869-884 [DOI: 10.1007/978-3-319-46475-6_53]
- Liu X C, Liu W, Mei T and Ma H D. 2018b. PROVID: progressive and multimodal vehicle reidentification for large-scale urban surveillance. *IEEE Transactions on Multimedia*, 20(3): 645-658 [DOI: 10.1109/TMM.2017.2751966]
- Liu X C, Liu W, Zheng J K, Yan C G and Mei T. 2020. Beyond the parts: learning multi-view cross-part correlation for vehicle re-identification//Proceedings of the 28th ACM International Conference on Multimedia. Seattle, USA; ACM: 907-915 [DOI: 10.1145/3394171.3413578]
- Müller R, Kornblith S and Hinton G. 2020. When does label smoothing help? [EB/OL]. [2021-06-03]. <https://arxiv.org/pdf/1906.02629.pdf>
- Pan H P, Wang Y T and Ma M. 2021. Vehicle re-identification methods based on attention mechanism and multi-scale fusion learning. *Journal of Zhejiang Sci-Tech University (Natural Sciences Edition)*, 45(5): 657-665 (潘海鹏, 王云涛, 马森. 2021. 基于注意力机制与多尺度融合学习的车辆重识别方法. *浙江理工大学学报(自然科学版)*, 45(5): 657-665) [DOI: 10.3969/j.issn.1673-3851(n).2021.05.011]
- Pan X G, Luo P, Shi J P and Tang X O. 2018. Two at once: enhancing learning and generalization capacities via IBN-net//Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich, Germany; Springer: 484-500 [DOI: 10.1007/978-3-030-01225-0_29]
- Qian J J, Jiang W, Luo H and Yu H Y. 2019. Stripe-based and attribute-aware network: a two-branch deep model for vehicle re-identification [EB/OL]. [2021-06-03]. <https://arxiv.org/pdf/1910.05549.pdf>

- Qiu M K and Li X Y. 2021. Detail-aware discriminative feature learning model for vehicle re-identification. *Acta Scientiarum Naturalium Universitatis Sunyatseni*, 60 (4): 111-120 (邱铭凯, 李熙莹. 2021. 用于车辆重识别的基于细节感知的判别特征学习模型. *中山大学学报(自然科学版)*, 60 (4): 111-120) [DOI: 10.13471/j.cnki.acta.snus.2020.03.16.2020B023]
- Shen Y T, Xiao T, Li H S, Yi S and Wang X G. 2017. Learning deep neural networks for vehicle Re-ID with visual-spatio-temporal path proposals//*Proceedings of 2017 IEEE International Conference on Computer Vision*. Venice, Italy: IEEE: 1918-1927 [DOI: 10.1109/ICCV.2017.210]
- Wang Z D, Tang L M, Liu X H, Yao Z L, Yi S, Shao J, Yan J J, Wang S J, Li H S and Wang X G. 2017. Orientation invariant feature embedding and spatial temporal regularization for vehicle re-identification//*Proceedings of 2017 IEEE International Conference on Computer Vision*. Venice, Italy: IEEE: 379-387 [DOI: 10.1109/ICCV.2017.49]
- Xu Z M, Wei L L, Lang C Y, Feng S H, Wang T and Bors A G. 2020. HSS-GCN: a hierarchical spatial structural graph convolutional network for vehicle re-identification//*Pattern Recognition. ICPR International Workshops and Challenges*. Milan, Italy: Springer: 356-364 [DOI: 10.1007/978-3-030-68821-9_32]
- Zhong Z, Zheng L, Cao D L and Li S Z. 2017. Re-ranking person reidentification with k-reciprocal encoding//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, USA: IEEE: 3652-3661 [DOI: 10.1109/CVPR.2017.389]
- Zhou Y and Shao L. 2017. Cross-view GAN based vehicle generation for re-identification//*Proceedings of 2017 British Machine Vision Conference (BMVC)*. [s. l.]: BMVA Press: 186.1-186.12 [DOI: 10.5244/c.31.186]
- Zhou Y and Shao L. 2018. Viewpoint-aware attentive multi-view inference for vehicle re-identification//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Salt Lake City, USA: IEEE: 6489-6498 [DOI: 10.1109/CVPR.2018.00679]

作者简介

王振学,男,硕士研究生,主要研究方向为车辆重识别、行人重识别。E-mail: 18140082@bjtu.edu.cn

郎丛妍,通信作者,女,教授,主要研究方向为计算机视觉、多媒体图像处理与分析。E-mail: cylang@bjtu.edu.cn

许喆铭,女,博士研究生,主要研究方向为车辆重识别和多视图学习。E-mail: 21112016@bjtu.edu.cn

雪洋洋,男,硕士研究生,主要研究方向为计算机视觉、车辆重识别。E-mail: 19120418@bjtu.edu.cn

李尊,女,博士研究生,主要研究方向为计算机视觉、多媒体图像处理与分析。E-mail: lznus2018@gmail.com

魏莉莉,女,博士研究生,主要研究方向为计算机视觉、多媒体图像处理与分析。E-mail: 20112014@bjtu.edu.cn