

DOI: 10.19650/j.cnki.cjsi.J2209725

基于局部特征与视点感知的车辆重识别算法

贺晓东,王春艳,孙昊,赵义武
(长春理工大学光电工程学院 长春 130000)

摘要:在车辆重识别任务中,车辆视角的多变性会影响算法的准确性。为了解决视角多变对重识别准确性的影响,本文提出了一种基于局部特征与视点感知的车辆重识别方法。首先,使用语义分割算法将车辆解构为正面、背面、侧面、顶部4个部分,以提高车辆的细粒度表征。通过设计一种车辆视点感知网络,来输出视点的预测概率信息,据此概率信息动态平滑地呈现车辆视点感知效果。利用视点感知效果,为车辆每个局部区域赋予不同的权重,达到缩短类内距离,扩大类间差距,减少视角变化对车辆重识别的影响。利用公开数据集进行实验,其中VeRi776数据集的mAP可达到80.9%。结果表明,本方法可有效提高车辆重识别精度。结合消融实验证明了视点感知的平滑表示对多视角下车辆重识别的有效性。

关键词:深度学习;车辆重识别;语义分割;视点感知

中图分类号: TP391 TH39 **文献标识码:** A **国家标准学科分类代码:** 520.2060

Local-features and viewpoint-aware for vehicle re-identification

He Xiaodong, Wang Chunyan, Sun Hao, Zhao Yiwu

(School of Opto-electronic Engineering, Changchun University of Science and Technology, Changchun 130000, China)

Abstract: The change of vehicle view may affect the accuracy of the re-identification algorithm. To solve the influence of changing viewpoints on the accuracy of re-identification, we propose a vehicle re-identification method based on local features and viewpoint perception. First, a parsing module is trained to parse a vehicle into four different views, front, back, side, and top. In this way, the fine-grained representation of the vehicle is improved. Then, we introduce a vehicle viewpoint-aware network. The output of the network is the predicted probability information of the viewpoint, and the vehicle viewpoint perception effect is dynamically and smoothly represented according to the probability information. Finally, the viewpoint-aware effect is used to assign different weights to each local area of the vehicle to shorten the intra-class distance, expand the inter-class distance, and reduce the impact of viewpoint changes on vehicle re-identification. This method is evaluated on public datasets, including VeRi776 and VehicleID. The accuracy of mAP on VeRi776 dataset has achieved 80.9%. Experimental results show that the proposed method can effectively improve the accuracy of vehicle re-identification. Ablation experiments demonstrate the effectiveness of the viewpoint-aware smooth representation for vehicle re-identification from multiple viewpoints.

Keywords: deep learning; vehicle re-identification; semantic segmentation; viewpoint aware

0 引言

随着计算机视觉和物联网技术的发展,智能交通系统也迅猛发展,车辆已成为智能交通系统中的重要关注对象。车辆检测^[1-2]、跟踪、类型识别等成为热门研究领域。但是随着车辆保有量的剧增,将车辆重识别技术应

用于车辆的监管,对智能交通建设将起到重要作用。

车辆重识别任务是从监控系统中的多个非重叠摄像头中检索目标车辆,可应用于智能停车、可疑车辆跟踪、车辆事件检测、车辆计数等。目前车辆重识别技术的难点是类间差异小和类内差异大。小的类间差异是由于同品牌同型号的不同车辆外观十分相似,大的类内差异是由于同一辆车在不同视角和不同光照条件下外观的多样性^[3-4]。

针对类间差异小的情况,文献[5]提出了一种区域深度感知模型(region-aware deep model, RAM),可同时提取全局特征和局部特征。结合车辆 ID、类型和颜色训练模型,来确定更具判别力的全局特征和局部特征。针对类内差异大的情况,文献[6]提出了一种视点感知注意及多视图推理模型(viewpoint-aware attentive multi-view inference, VAMI)。VAMI 为任意视点的输入图像提取单视角特征,采用视点感知模型来选择不同视点的核心区域,再通过对抗训练生成多个视角特征。VAMI 模型虽然较好的解决了多视角对于车辆重识别精度的影响,但是模型采用了生成对抗网络(generative adversarial networks, GAN),而 GAN 存在训练不稳定、梯度消失、模式崩溃等问题。文献[7]提出了一种基于解构的视点感知嵌入网络模型(parsing-based view-aware embedding network for vehicle re-identification, PVEN)。PVEN 使用解构网络将车辆分为 4 个不同的视图,通过掩码平均池化对齐特征,为共同可见区域的车辆视图赋予不同大小的权重,解决车辆视角多变问题。PVEN 虽然在一定程度上提高了重识别精度,降低了视角变化对重识别结果

的影响。但是视角的变化具有连续性,车辆视图的常见面积大小并不能完全表示车辆视点平滑的变化。

为了提升在多视角下车辆重识别精度,本文提出了一种使用平滑标签训练的视点感知模型。通过使用视点软标签,平滑过渡不同的离散视点,获得更准确的视点预测信息。然后结合局部特征和全局分支进行多任务联合训练,使模型学习到更具辨识力的特征。最后结合公开数据集的实验证明了该方法的有效性。

1 车辆重识别网络结构

本文的车辆重识别网络结构分为全局分支与局部分支,其中局部分支包括局部特征模块和视点感知模块,如图 1 所示。在训练阶段,全局分支使用全局特征构建车辆 ID Loss 和全局 Triplet Loss^[8]。局部分支中使用局部特征模块生成的局部特征,与视点感知模块输出的视点预测信息构建局部 Triplet Loss。在验证阶段,将全局特征与局部特征的相似性距离按比例相加作为最终车辆间的相似性距离。

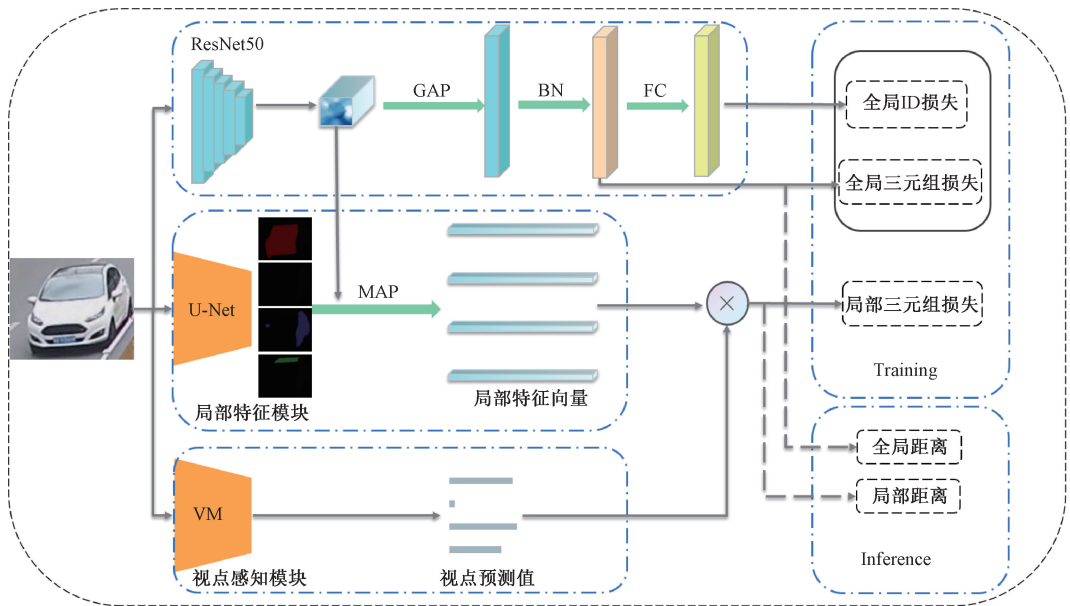


图 1 车辆重识别网络结构图

Fig. 1 Network architecture of vehicle re-identification

1.1 全局分支

全局分支负责提取车辆图像全局特征,将全局特征池化、归一化后输入全局三元组损失函数,全连接层结果输入到车辆 ID 损失函数,最后联合 ID 损失和三元组损失进行全局分支的网络训练。

在车辆重识别算法中,使用 ResNet^[9],Inception^[10]和 VGG16^[11]等经典网络作为基础网络结构。其中深度残

差网络(ResNet50)加入了残差单元,其中包含跨层连接,并与经过卷积操作的结果相加,输出残差函数,使其更容易优化。所以 ResNet50 较于其他网络结构,可以更好的解决梯度问题。并且在残差块中加入了 1×1 卷积,来减少参数,在一定程度上降低计算量。基于此本文选择在 ImageNet^[12]上预训练的 ResNet50 作为特征提取基础架构。

对于输入的 256×256 车辆图像,经过基础特征提取网络获得一个 $16 \times 16 \times 2\,048$ 的特征图 F ,将 F 输出到两个分支。对于全局分支,通过全局平均池化获得全局特征 f_g^p ,基于 f_g^p 输出全局 ID 损失 L_{id}^g :

$$L_{id}^g = - \sum_{\forall p} \ln \frac{\exp(\mathbf{W}_y^T \mathbf{f}_g^p)}{\sum_{\forall y} \exp(\mathbf{W}_y^T \mathbf{f}_g^p)} \quad (1)$$

式中: $\mathbf{W}$ 为全连接层的参数矩阵, y 为车辆 ID 标签。使用 L_{id}^g 训练网络可学习到对车辆 ID 进行分类的全局特征。

对于图片对 $I = (p, q)$, 可以根据下式计算它们之间的相似性距离:

$$D(I) = D(p, q) = \|\mathbf{f}_g^p - \mathbf{f}_g^q\|_2 \quad (2)$$

$\mathbf{f}_g^p$ 和 $\mathbf{f}_g^q$ 表示图像对 $I = (p, q)$ 的全局特征, $\|\cdot\|_2$ 表示欧氏距离。

输入 3 张图像 p, p^+, p^- , 其中 p, p^+ 属于同一车辆 ID 的样本, p, p^- 属于不同车辆 ID 的样本。因此, 将 $I_{\text{position}} = (p, p^+)$ 作为正样本, $I_{\text{negative}} = (p, p^-)$ 作为负样本。则全局三元组损失函数如式(3):

$$L_{\text{triplet}}^g(p, p^+, p^-) = \max\{D(I_{\text{position}}) - D(I_{\text{negative}}) + a, 0\} \quad (3)$$

式中: a 表示样本距离的阈值。

联合全局分类损失与全局三元组损失进行网络训练, 可减小车辆类内距离, 扩大类间距离, 可满足车辆重识别任务需求。当全局特征将不同视点的特征融合为一个特征时, 会出现特征无法对齐的情况。而且由于视点变化的多样性, 在出现极端视角变化时, 仅依靠全局分支无法正确进行距离度量, 使重识别的准确度降低。因此, 必须将车辆解构, 提取不同视点的车辆局部特征, 并对车辆进行视点感知来解决此类问题。

1.2 局部特征模块

车辆可看成一个六面刚性物体, 不会产生物理形变。可稳定准确的生成车辆不同部分掩码, 使用掩码可以将不同车辆同一部分对齐, 具有很强的泛化性。由于车辆底部在监控图像中通常是不可见的, 车辆左右两侧不会同时出现, 且具有对称性。所以本文将车辆分为车头、车尾、侧面和顶部作为四个视点基准, 侧面为车辆左侧或者右侧, 前视窗和后视窗分别属于车辆车头和车尾部分。

本文使用 U-Net^[13] 训练 1 个车辆图像语义分割网络, 输出上述 4 个车辆部分的掩码向量。通过最大池化使局部掩码尺寸为 16×16 。如图 2 所示, 每种颜色代表 1 个车辆部分, 对应 1 个车辆局部掩码向量 $\{\mathbf{f}_l^i | i \in 1, 2, 3, 4\}$ 。4 个局部区域可将车辆完整覆盖, 车辆的任何细微特征都可以被捕获。

对 1.1 节中的全局特征图 F 进行掩码平均池化操作 (mask average pooling, MAP), 可得局部特征



图 2 语义分割实例

Fig. 2 Semantic segmentation example

$\{\mathbf{f}_l^i | i \in 1, 2, 3, 4\}$ 。它们分别表示车头、车尾、侧面和顶部的局部特征。 $\mathbf{f}_l^i$ 由式(4) 计算:

$$\mathbf{f}_l^i = \frac{\sum_{j,k=1}^{16} \mathbf{M}_i(j, k) \times \mathbf{F}(j, k)}{\sum_{j,k=1}^{16} \mathbf{M}_i(j, k)} \quad (4)$$

将所得车辆局部特征输入到视点感知模块。

1.3 视点感知模块

视点为图像中车辆的可见区域。例如: 车辆为正前视点时, 车辆只有车头和车顶部是可见的。相当于将车辆可见区域量化为视点值, 在本文中视点值就是视点感知模块输出的视点预测概率。

当前基于视点感知模型的车辆重识别算法大多是对几个固定视点进行感知, 但是车辆的视点并不是离散的而是连续的, 这会弱化模型对视点的感知能力。为此本文根据文献[14]的启发, 提出了一种使用软标签训练的视点感知模型, 对离散视点进行平滑。模型结构如图 3 所示。定义图片 $I_i = (\mathbf{x}_i, v_i)$, 其中 $\mathbf{x}_i$ 表示图像经过 ResNet50 提取到的特征, v_i 表示视点标签, $v_i \in \{\text{front}, \text{side}, \text{back}\}$ 。由于车辆图像是从监控视频中采集的, 都是俯拍。车辆顶部始终可见, 故顶部的区域权重固定为 0.25。

模型使用 softmax 函数输出视点预测概率, V 表示视点数量, 按照上述视点定义 $V = 3$, $\mathbf{W}$ 表示权重矩阵 ($\mathbf{W} \in \mathbb{R}^{d \times V}$), 则视点 v 的预测概率如式(5) 所示:

$$S(v) = \frac{\exp(\mathbf{W}_v^T \mathbf{x})}{\sum_{i=1}^V \exp(\mathbf{W}_i^T \mathbf{x})} \quad (5)$$

传统模型会直接使用交叉熵损失进行网络训练, 其计算公式如式(6) 所示:

$$L_v = - \frac{1}{V} \sum_v p_v \ln(S(v)) \quad (6)$$

其中, p_v 表示分类标签, 在传统算法中 p 会按照 one-hot 进行编码, 取值如式(7) 所示。

$$p_v = \begin{cases} 1, & v = V_i \\ 0, & \text{其他} \end{cases} \quad (7)$$

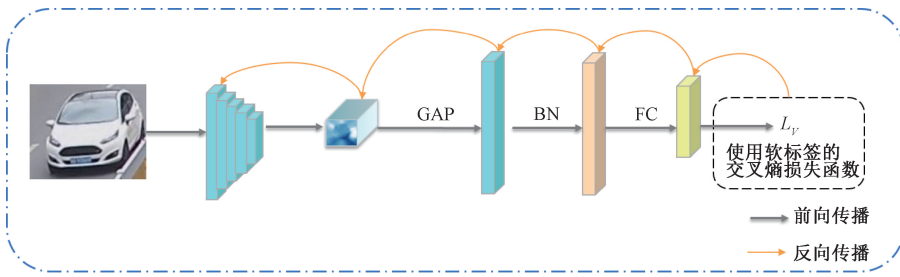


图3 视点感知模型

Fig. 3 Viewpoint-aware model

此时模型输出的视点概率值为离散数据,忽略了在3个固定视角间的连续值。并且模型倾向于最大化一个标签的预测对数似然,这可能会导致过拟合,降低模型的泛化能力。为解决此问题,本文提出了将标签平滑化,使用软标签进行训练。令标签 p_v 为:

$$p_v = \begin{cases} 1 - \varepsilon, & v = V_i \\ \frac{\varepsilon}{V}, & \text{其他} \end{cases} \quad (8)$$

其中, ε 为自定义参数,通过引入参数 ε 来调整概率分布,将 one-hot 硬标签替换为软标签。但由于 ε 是一个固定值,将导致视点感知网络输出的每个样本的预期概率相同,无法得到更准确的视点预测概率。为此,本文提出了平滑参数 $\varepsilon = \alpha(1 - S(v))$, α 为比例缩放系数。当概率输出值 $S(v)$ 大时,可以获得当前视点较大的置信度。相反当 $S(v)$ 小时,可以获得其它视点较大的置信度。将式(8)代入式(6)中,对 L_v 进行网络训练,使视点感知模型最终输出的视点预测概率为视点的连续值。

从重识别数据集中选择部分图像构建视点感知数据集,进行视点感知模型训练。模型训练完成后使用式(5)输出车辆视点预测概率 $\{S(v) | v \in 1, 2, 3\}$ (v 表示车头、车尾、侧面3个视点),加上顶部的固定权重使 $\{S(v) | v \in 1, 2, 3, 4\}$, 然后联合局部特征 $\{f_i^l | i \in 1, 2, 3, 4\}$, 便可获得嵌入视点感知模型的局部三元组损失和车辆相似性度量距离。

图片对 $I = (p, q)$ 的局部相似性距离,在加入视点感知模型的输出之后可由下式计算:

$$D(I) = D(p, q) = \sum_{i=1}^4 D_i(p, q) (S_p(i) S_q(i)) \quad (9)$$

其中, $S_p(i)$ 、 $S_q(i)$ 表示 p 与 q 的视点预测概率, $D_i(p, q)$ 表示车辆4个部分的局部欧式距离。由式(3)可知局部三元组损失值可由式(10)计算:

$$L_{\text{triplet}}^l(p, p^+, p^-) = \max\{D(I_{\text{position}}) - D(I_{\text{negative}}) + a, 0\} \quad (10)$$

最终得到整体模型训练时使用的损失函数如式(11)所示。

$$L = L_{\text{id}}^g + L_{\text{triplet}}^g + L_{\text{triplet}}^l \quad (11)$$

2 实验结果与分析

2.1 数据集与评价标准

实验使用公开数据集 VeRi776^[15] 和 VehicleID^[16] 对模型进行可靠性评估。VeRi776 是广泛使用的车辆重识别数据集,通过20个监控摄像头从不同视角收集了776辆车的50 000张图像。训练集包括576辆车的37 778张图像,测试集包括200辆车的11 579张图像。VehicleID 主要从前后两个视角收集车辆图像,共有26 267辆车的221 763张图像,其中训练集包含13 134辆车的110 178张图像。共有3种不同尺寸的测试集:Test800、Test1600和Test2400。

基于上述数据集,本文使用平均精度值(mean average precision, mAP)和累计匹配曲线(CMC@k)作为模型评价标准。

mAP 表示所有待检索图像的准确率的平均数,由式(12)计算:

$$mAP = \frac{\sum_{q=1}^Q AP(q)}{Q} \quad (12)$$

Q 表示查询图片的数量,其中 $AP(q)$ 可由式(13)计算:

$$AP = \frac{\sum_{k=1}^n p(k) g(k)}{N} \quad (13)$$

其中, n 为测试集图像数量, N 为当前检索目标图像数量。 $p(k)$ 表示检索结果前 k 个的准确率,当检索结果第 k 个匹配正确时 $g(k) = 1$, 否则 $g(k) = 0$ 。 CMC@k 表示了待检索图像在前 k 个结果中正确的概率。

2.2 实验过程

车辆语义分割模型使用 U-Net 网络,训练使用的数据集在 VeRi776 中进行标注。为了提高分割模型对不同视角的适应能力,尽量选择车辆的多个视角。共选择3 000张车辆图像,其中2 500张为训练集,500张为验证集。训练40个 epoch 后在验证集上达到81%的 IoU 分数。

视点感知模型使用在 ImageNet 上预训练的 ResNet50 作为基础网络进行特征提取,数据集在 VeRi776 中标注。选择前、侧、后 3 个固定视角作为基准,赋予不同的置信度。共训练 2 000 张图像,将提取到的特征经分类层输出后使用交叉熵损失函数计算 Loss,共训练 50 个 epoch。式(8)中的比例缩放系数 α 设置为 0.1。在验证阶段,特征经分类层输出后,使用 softmax 函数计算视点预测概率。

在整车重识别网络训练中,车辆图像先经过随机反转和随机擦除等预处理,尺寸调整为 256×256 ,并进行批量归一化操作后输入网络中。在训练阶段,只有 L_{id}^s 使用全连接层的输出作为输入进行 Loss 的计算,其余损失函数的输入使用图像特征批量归一化的值。批量大小设置为 64,采用 warm-up^[17]策略调整学

习率。共训练 150 个 epoch,使用 Adam 优化器进行优化, TripleLoss 中的距离阈值为 0.3。在验证阶段,去除全连接层获取图像特征后使用式(2)计算欧氏距离作为全局相似性距离,使用式(9)计算局部相似性距离。在检索图像时总的相似性距离时,采用式(14)计算。

$$D(I) = \lambda_1 D_{global}(I) + \lambda_2 D_{local}(I)$$

(14)

2.3 实验结果与分析

从数据集中随机选取 5 个不同车辆,使用本文方法进行车辆重识别,根据相似度距离排名如图 4 所示。第 1 列为查询车辆,2~10 列为排名结果。其中实线边框车辆表示查询结果正确,属于同一辆车,虚线边框表示结果错误。图 5 表示本文方法与 RAM^[5]和 PRN^[18]两种方法针对同一辆车的重识别结果对比图。



图 4 车辆重识别结果
Fig. 4 Results of vehicle re-identification



图 5 车辆重识别结果对比
Fig. 5 Comparison of vehicle re-identification results

为了验证本文提出的 PVM 车辆重识别方法的有效性,分别在 VeRi776 和 VehicleID 数据集上进行了对比实验。使用 mAP, CMC@1 和 CMC@5 作为评价指标。PVM 与其他近年来优秀的车辆重识别方法的对比表现

如表 1 和 2 所示。
其中 PVEN 对于视角的处理,采用了共同可见注意力机制。将车辆解构后的对应区域的掩码面积作为不同区域相似性度量的权重。这种方法虽然在一定程度上解

表 1 不同方法在 VeRi776 数据集上的表现

Table 1 The performance of methods on VeRi776 dataset

方法	mAP	CMC@ 1	CMC@ 5
GoogleNet[19]	0. 170	0. 498	0. 712
FACT[15]	0. 185	0. 510	0. 735
PROVID[20]	0. 278	0. 614	0. 788
OIFE[21]	0. 480	0. 894	–
RAM[5]	0. 615	0. 886	0. 940
VANet[22]	0. 663	0. 878	0. 960
PRN[18]	0. 743	0. 943	0. 989
PVEN[7]	0. 794	0. 956	0. 984
VAT[23]	0. 804	0. 958	–
PVM(本文)	0. 809	0. 961	0. 986

表 2 不同方法在 VehicleID 数据集上的表现

Table 2 The performance of methods on VehicleID dataset

方法	Small		Medium		Large	
	CMC@ 1	CMC@ 5	CMC@ 1	CMC@ 5	CMC@ 1	CMC@ 5
OIFE[21]	–	–	–	–	0. 670	0. 829
VAMI[6]	0. 631	0. 833	0. 529	0. 751	0. 473	0. 703
RAM[5]	0. 752	0. 915	0. 723	0. 870	0. 677	0. 845
AAVER[24]	0. 747	0. 938	0. 686	0. 900	0. 635	0. 856
PRN[18]	0. 784	0. 923	0. 750	0. 883	0. 742	0. 864
VAT[23]	0. 845	–	0. 805	–	0. 782	–
PVEN[7]	0. 847	0. 970	0. 806	0. 945	0. 778	0. 920
PVM(本文)	0. 849	0. 976	0. 811	0. 943	0. 780	0. 904

决了视点变化的问题。但是,这样训练的网络最终很难学习到每个视角的所有可判别特征。而 PRN 为了获得细微的差异,仅仅使用了车窗、车灯和车牌的局部特征进行相似性度量。此方法确实可以提取到重要区域的细微特征,但忽略了其他区域的特征信息,效果并不好。包括 RAM 提出的区域感知深度模型,也只是针对局部特征,没有解决视角问题。本文构建的视点感知模型(VM)可以输出视点预测概率,将其作为局部区域相似性度量的权重,解决了视点变化问题。结合局部特征模型(PM),可以获取细微的并覆盖车辆完整图像的特征。

由图 5 可知,本文方法不仅可以检索到与查询图像属于同一视角的车辆图像,而且可以检索到不同视角下的同一车辆。而仅仅使用局部特征的 RAM 和 PRN 两种方法虽然对于相同视角下的车辆检索精度较高,但是不能检索到车辆的其他视角。证明了本文方法对于多视角下车辆重识别的有效性。

由表 1 和 2 结果可知,PVM 在 VeRi 数据集上的 mAP 达到了 80. 9%,CMC@ 1 达到了 96. 1%,高于其他方法。CMC@ 5 为 98. 6% 也处于较高水平。在 VehicleID 数据集上的 CMC@ 1 和 CMC@ 5 整体比较高,其中在 Small 测试子集中 CMC@ 1 和 CMC@ 5 是最高的。表中的实验结果证明了本文方法的有效性。

为了进一步证明 PVM 对实际车辆图像的效果,以不同视角对车辆图像进行采集。由于数据量较少,无法用来训练模型,所以本文使用在 VeRi776 数据集上训练的模型进行测试。测试结果如图 6 所示。

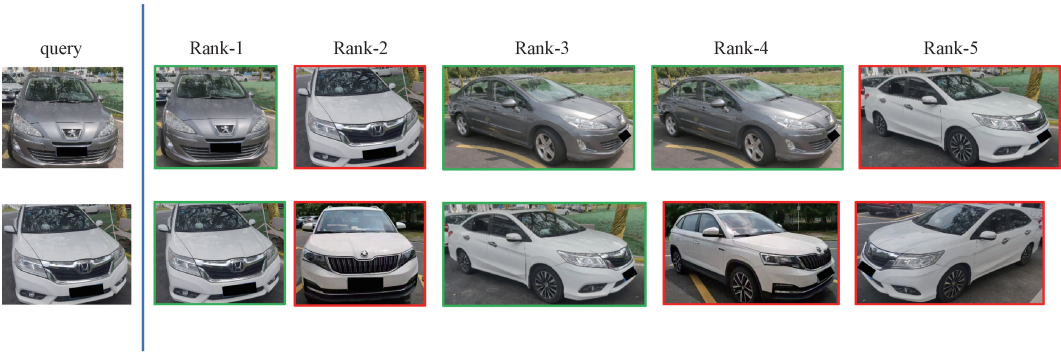


图 6 车辆重识别结果

Fig. 6 The Result of Vehicle re-identification

图 6 表明对于实际采集的车辆图像,PVM 的重识别结果比较准确,同时解决了多视角下的车辆重识别问题。在实测数据集上的 mAP 为 0. 432,相较于在 VeRi776 数据集上的测试结果仍有较大差距。这是由于数据集的图像风格变化,背景差异大造成的。

2. 4 消融实验

为了进一步验证本文所提出方法中各模块的有效性,在 VeRi776 数据集上进行了消融实验。评价指标同样使用 mAP, CMC@ 1 和 CMC@ 5,结果如表 3 所示。

表 3 模型有效性验证结果

Table 3 The model validity verification

方法	mAP	CMC@ 1	CMC@ 5
baseline	0. 726	0. 925	0. 963
Global	0. 775	0. 958	0. 983
Local(PM+VM)	0. 767	0. 958	0. 982
PVM	0. 809	0. 961	0. 986

其中 baseline 作为基线,使用 ResNet50 提取全局特征,仅使用 ID 损失函数 L_{id} 进行训练。Global 表示全局分支,在 baseline 的基础上添加全局三元组损失函数 $L_{triplet}^g$ 训练。Local 表示局部分支,是本文提出的视点感知模块结合局部特征模块,在 baseline 的基础上添加局部三元组损失函数 $L_{triplet}^l$ 进行训练。由表 3 得 baseline 的 mAP 为 72.6%,添加 $L_{triplet}^g$ 后 mAP 增加了 4.9%,添加局部分支后 mAP 增加了 4.1%。可见全局分支与局部分支对于重识别效果都有提升。而将局部分支与全局分支构成 PVM 时,mAP 提升了 8.3%,证明了 PM 和 VM 的有效性。

在验证阶段车辆的检索结果,按照相似性距离排名,通过式(14)计算。为了研究 λ_1 和 λ_2 取值对检索结果的影响,在 VeRi776 数据集上进行了实验。实验结果如表 4 所示。

表 4 参数 λ_1 和 λ_2 对检索结果的影响

Table 4 The effect of λ_1 and λ_2 for query

λ_1	λ_2	mAP	CMC@ 1	CMC@ 5
1	0	0. 784	0. 952	0. 973
1	0. 3	0. 795	0. 959	0. 983
1	0. 5	0. 809	0. 961	0. 986
1	0. 7	0. 798	0. 960	0. 986
1	1	0. 768	0. 957	0. 972
0	1	0. 726	0. 915	0. 964

结果表明局部距离权重增大对检索结果有改进,当 λ_2 达到 0.5 时效果最好,所以本文取 $\lambda_1 = 1, \lambda_2 = 0.5$ 计算验证阶段中的相似性距离。

3 结 论

本文提出了一种基于局部特征与视点感知的车辆重识别模型,二者结合不仅可以稳定获取相同车辆的判别信息,而且可利用视点感知能力降低视点变化对车辆重识别的影响。不仅可以缩短类内距离,还可增大类间距离。且使用软标签训练的视点感知模型可以将离散视点

平滑化,获得连续的视点感知能力。使在局部区域相似性度量中的各区域权重更加准确。在两个数据集上的实验结果证明了本方法的有效性,结果优于当前大多数方法。

由于视点感知模型属于独立训练,会对重识别模型的性能有一定影响,后续将会研究如何将视点感知模块嵌入,构成一个完整的端对端网络。对于图像风格变化导致重识别精度下降的问题,未来应当研究如何将迁移学习应用到车辆重识别算法中,使算法可以落地到实际应用。

参考文献

[1] 郑少武,李巍华,胡坚耀. 基于激光点云与图像信息融合的交通环境车辆检测[J]. 仪器仪表学报,2019, 40(12):143-151.
ZHEN SH W, LI W H, HU J Y. Vehicle detection in traffic environment based on fusion of laser point cloud and image information [J]. Chinese Journal of Scientific Instrument, 2019, 40(12): 143-151.

[2] 梁继然,陈壮,董国军,等. 结合注意力机制和密集连接网络的车辆检测方法[J]. 电子测量与仪器学报, 2022,36(3):210-216.
LIANG J R, CHEN ZH, DONG G J, et al. A vehicle detection method combining attention mechanisms and densely connected networks [J]. Journal of Electronic Measurement and Instrumentation, 2022, 36 (3): 210-216.

[3] 刘凯,李浥东,林伟鹏. 车辆再识别技术综述[J]. 智能科学与技术学报, 2020, 2(1): 10-25.
LIU K, LI X D, LIN W P. Review of vehicle re-identification technology [J]. Journal of Intelligent Science and Technology, 2020, 2(1): 10-25.

[4] WANG H, HOU J, CHEN N. A survey of vehicle re-identification based on deep learning[J]. IEEE Access, 2019, 7: 172443-172469.

[5] LIU X, ZHANG S, HUANG Q, et al. Ram: A region-aware deep model for vehicle re-identification[C]. 2018 IEEE International Conference on Multimedia and Expo (ICME), IEEE, 2018: 1-6.

[6] ZHOU Y, SHAO L. Aware attentive multi-view inference for vehicle re-identification[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 6489-6498.

[7] MENG D, LI L, LIU X, et al. Parsing-based view-aware embedding network for vehicle re-identification [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 7103-7112.

[8] SCHROFF F, KALENICHENKO D, PHILBIN J.

- Facenet: A unified embedding for face recognition and clustering[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 815-823.
- [9] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [10] SXEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 1-9.
- [11] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. ArXiv Preprint, 2014, ArXiv: 1409. 1556.
- [12] RUSSAKOVSKY O, DENG J, SU H, et al. Imagenet large scale visual recognition challenge[J]. International Journal of Computer Vision, 2015, 115(3): 211-252.
- [13] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]. International Conference on Medical Image Computing and Computer-Assisted Intervention, 2015: 234-241.
- [14] ZHIHUI Z, XINYANG J, FENG Z, et al. Viewpoint-aware loss with angular regularization for person re-identification[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 13114-13121.
- [15] LIU X, LIU W, MA H, et al. Large-scale vehicle re-identification in urban surveillance videos [C]. 2016 IEEE International Conference on Multimedia and Expo (ICME), IEEE, 2016: 1-6.
- [16] LIU H, TIAN Y, YANG Y, et al. Deep relative distance learning: Tell the difference between similar vehicles[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 2167-2175.
- [17] XIONG R, YANG Y, HE D, et al. On layer normalization in the transformer architecture [C]. International Conference on Machine Learning, PMLR, 2020: 10524-10533.
- [18] HE B, LI J, ZHAO Y, et al. Part-regularized near-duplicate vehicle re-identification[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 3997-4005.
- [19] YANG L, LUO P, CHANGELOY C, et al. A large-scale car dataset for fine-grained categorization and verification[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 3973-3981.
- [20] LIU X, LIU W, MEI T, et al. Provid: Progressive and multimodal vehicle reidentification for large-scale urban surveillance [J]. IEEE Transactions on Multimedia, 2017, 20(3): 645-658.
- [21] WANG Z, TANG L, LIU X, et al. Orientation invariant feature embedding and spatial temporal regularization for vehicle re-identification [C]. Proceedings of the IEEE International Conference on Computer Vision, 2017: 379-387.
- [22] CHU R, SUN Y, LI Y, et al. Vehicle re-identification with viewpoint-aware metric learning[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 8282-8291.
- [23] YU Z, PEI J, ZHU M, et al. Multi-attribute adaptive aggregation transformer for vehicle re-identification [J]. Information Processing & Management, 2022, 59(2): 102868.
- [24] KHORRAMSHAHI P, KUMAR A, PERI N, et al. A dual-path model with adaptive attention for vehicle re-identification [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 6132-6141.

作者简介



贺晓东, 2019 年于长春理工大学获得学士学位, 现为长春理工大学硕士研究生, 主要研究方向为深度学习与图像处理。

E-mail: 15568804518@163.com

He Xiaodong received his B.Sc. degree from Changchun University of Science and Technology in 2019. He is currently a master student at Changchun University of Science and Technology. His main research interests include deep learning and image processing.



王春艳(通信作者), 分别在 1995 年和 2002 年于长春光学精密机械学院获得学士学位和硕士学位, 2005 年于长春理工大学获得博士学位。现为长春理工大学教授, 主要研究方向为光学仪器设计及测试。

E-mail: 245044961@qq.com

Wang Chunyan (Corresponding author) received her B.Sc. and M.Sc. degrees both from Changchun Institute of Optics and Fine Mechanics in 1995 and 2002, and received her Ph.D. degree from Changchun University of Science and Technology in 2005. She is currently a professor at Changchun University of Science and Technology. Her main research interests include optical instrument design and testing.