

一种融合表观与属性信息的车辆重识别方法

谢秀珍^{1,2}, 罗志明^{1*}, 连盛¹, 李绍滋¹

(1. 厦门大学信息学院, 福建 厦门 361005; 2. 龙岩学院数学与信息工程学院, 福建 龙岩 364002)

摘要: 针对基于深度学习的车辆重识别模型缺乏提取车辆局部细节描述的能力, 以及不同环境下同一车辆表观特征变化显著, 增加车辆重识别的难度的问题, 提出一种融合多尺度车辆表观和属性信息重识别模型. 充分利用车辆颜色和车型属性在不同环境下较为稳定且易判断的特性, 将其融合到车辆表观特征中, 获得强化提升的深度特征; 另外使用交叉熵损失函数与 Triplet 损失函数构建多任务的目标优化函数对模型进行参数训练. 该方法在 VeRi-776 车辆数据库上的实验结果表明: 通过融合颜色和车型属性特征可以明显地提高车辆重识别的准确率, 并可以取得优于其他大部分对比方法的性能.

关键词: 车辆重识别; 特征提取; 表观信息; 属性信息; 特征融合

中图分类号: TP 391

文献标志码: A

文章编号: 0438-0479(2021)01-0072-08

随着城市车辆迅猛增加, 对目标车辆的快速查找、跟踪和定位已成为公共交通安全管理的重要内容之一. 目前大部分道路交通监控系统中, 对车辆的识别主要是通过识别特定角度高清摄像头拍摄的前后车牌号来完成的; 但在实际环境中还存在大量摄像头分辨率低、拍摄角度非特定、车牌被遮挡等情况, 导致无法直接从视频中提取车牌号来对车辆进行识别, 因此也需要采用其他方法实现对车辆的识别工作. 车辆重识别也叫作车辆再识别, 其主要目的是构建一个跨摄像头的车辆图像检索模型, 能够对某一监控摄像头下拍摄到的指定目标车辆, 实现快速准确地检索与识别该车辆在其他不同监控摄像头下对应出现的图像^[1]. 在真实的道路交通监控场景中, 车辆图片不可避免受到各种因素的影响: 光照明暗、拍摄角度、障碍物对目标车辆的不同部位的遮盖、摄像设备的分辨率不同等, 都会导致很大的车辆外观差异性. 因此与其他目标检测识别^[2-3]问题相类似, 要实现准确的车辆重识别, 其中的一个重要环节是提取车辆特征来作为重识别的主要依据, 特征的判别力强弱将直接影响重识别的结果.

在深度学习^[4]方法被广泛运用之前, 主要是通过

手动设计提取各种具有较强判断性的视觉特征(整体或部分的特征, 如: 王盼盼等^[5]手动提取车辆图片的 HSV(hue, saturation, value)和局部二值模式(LBP)特征并进行融合, 再进行奇异值分解, 提取特征值; 李熙莹等^[6]运用部件检测算法有针对性地检测和提取区别性较强的车窗和车脸等部位的特征并将其进行融合, 生成新的融合特征, 然后利用图像特征之间的距离进行分类识别.

随着大规模车辆数据集的构建, 例如: VeRi-776 数据集^[7]、VehicleID 数据集^[8]、Toy Car RE-ID^[9]数据集和 VRIC^[10]数据集, 各种基于深度学习的车辆重识别模型也陆续被提出. Liu 等^[1]提出了“PROVID”模型来进行渐进式车辆重识别, 实现在特征方面由粗到细以及监控中由近及远相结合的搜索; Yan 等^[11]提出了一个基于多任务框架的模型, 将车辆图像建模成多粒度的关系, 并提出了广义的二元“相似/不相似”关系的成对排序和基于多粒度的列表排序方法, 渐进地利用多粒度排序约束来缓解精确车辆搜索问题; Liu 等^[12]提出由 4 个分支构成的“RAM”模型, 把整体特征和区域特征联合起来, 提取更详细、更具辨别力的特征; Bai 等^[13]使用敏群三元组嵌入的方法, 显著减

收稿日期: 2020-04-27 录用日期: 2020-10-20

基金项目: 国家自然科学基金(61876159, 61806172, U1705286)

* 通信作者: zhiming.luo@xmu.edu.cn

引文格式: 谢秀珍, 罗志明, 连盛, 等. 一种融合表观与属性信息的车辆重识别方法[J]. 厦门大学学报(自然科学版), 2021, 60(1): 72-79.

Citation: XIE X Z, LUO Z M, LIANG S, et al. A vehicle re-identification method by fusing the vehicle appearance and attribute information[J]. J Xiamen Univ Nat Sci, 2021, 60(1): 72-79. (in Chinese)



轻了类间相似和类内方差对细粒度识别的影响;Zhou 等^[9]利用车辆角度信息,提出了两种端到端深层架构,学习车辆不同视角角度的转换;Shen 等^[14]结合时空信息,提出了一个两阶段框架,排除具有高相似度的不同车辆间的相互干扰,有效地规范了重识别结果;He 等^[15]通过训练一个额外的车辆部位检测器,可以定位具有显著判别能力的局部区域并提取相应的特征,该方法可以有效地利用车辆的局部特征;Chen 等^[16]将车辆沿着水平、垂直以及特征通道方向划分成多个不同的子部分,然后再将这些子部分的特征进行融合,得到相对包含更多局部信息的车辆特征。

上述方法均有一定的代表性,说明充分利用手动提取或深度提取的特征,对车辆重识别工作中所依赖的车辆表观信息或角度信息等特征进行恰当处理,均能够取得不错的性能。目前大多数方法多从整体层面提取车辆特征,缺乏对局部细节的描述;或者使用额外的检测分支来定位车辆局部区域,增加了人工标注的工作量。此外,不同颜色或者不同车型的车辆肯定具有不同 ID,且颜色与车型在变化的环境因素中亦能较容易被判断和分类;但现有方法更多是将颜色、车型作为属性信息加入到额外的监督损失函数中,并没有很好地将这些属性信息融入到车辆的特征中用于提升车辆重识别的准确率。

基于此,本文提出一种融合“多尺度表观特征”与“车辆属性信息”的车辆重识别算法模型。该模型以预训练的 ResNet-50^[17]作为骨干网络提取车辆的基础特征向量;通过两个分支从中分别计算提取车辆颜色与车型相关的属性特征、局部与整体融合的多尺度表观特征;采用动态自适应加权的方式对这两个分支的特征进行融合,用于获得更具有判别力和鲁棒性的深度特征作为车辆的最终特征表示;并使用一个多任务目标优化函数对整个模型进行参数优化训练。

1 模型框架

本文构建了一个如图 1 所示的车辆重识别网络模型。该模型首先通过一个在 ImageNet 数据库上构建的骨干残差网络-50(ResNet-50)计算提取车辆基础特征向量 z ;接下来分别从基础特征向量中计算提取车辆颜色和车型属性特征 $f_{attribute}$ 以及多尺度车辆表观特征 $f_{appearance}$;然后将车辆属性特征融合到车辆表观特征以辅助提取强化提升的车辆特征向量 f 。本文通过一个多任务的损失函数,在综合考虑车辆分类、车辆颜色分类和车型分类的基础上具体优化网络模型参数。本节后续将分别介绍整个模型的具体计算流程。

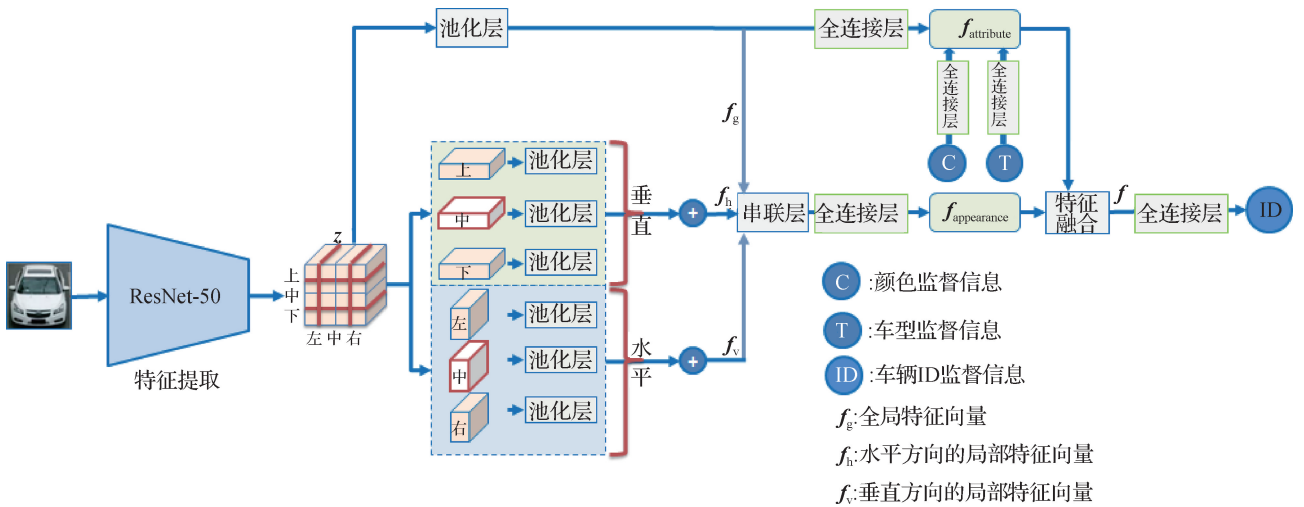


图 1 融合表观和属性信息的车辆重识别网络

Fig. 1 The vehicle re-identification network fused appearance with attribute information

1.1 基础特征提取

目前 ImageNet 数据库上预训练的深度学习模型被广泛作为各种计算机视觉任务的特征提取骨干网络。预训练模型指的是已经训练好的能执行大量数据任务的深度学习框架卷积神经网络(CNN)。预训练结束时,会得到结果较好的一组权重值供他人研究共

享。随着深度学习网络层数的增加,CNN 表达力增强,有利于提高分类准确率;但层数并不是可以无限增加的,到达瓶颈值后网络收敛缓慢,分类准确率不升反降。这种情况下,采用 ResNet 能解决这个问题:通过增加残差结构单元,让网络在层数增加的同时不会出现退化现象。本文综合考虑准确率和计算复杂度,采

用了目前其中之一的 ResNet-50^[17]作为骨干特征提取器,用于提取车辆图片的基础特征向量 z . 实际计算过程中,为了提取局部特征,本文中删除了 ResNet-50 网络中的全局池化层及后续的全连接(FC)层,并把最后一个残差模块的步长设置为 1,因此对于一个 256×256 大小的输入图片,初始特征向量的维度为 $z \in \mathbf{R}^{16 \times 16 \times 2048}$.

1.2 属性特征提取

车辆的颜色与车型是从整体上描述车辆的相关属性信息,因此在本文中从共享的全局特征向量 z 中计算提取出车辆的颜色车型属性. 首先通过一个全局的池化层将骨干网络计算提取的初始基础特征向量 z 转换成一维的特征向量,然后再通过一个 FC 层映射成最终的属性特征向量 $f_{\text{attribute}}$:

$$f_{\text{attribute}} = \delta(W_1 P(z)), \quad (1)$$

其中, $W_1 \in \mathbf{R}^{512 \times 2048}$ 为属性特征对应 FC 层的参数矩阵, P 是全局平均池化操作, δ 是 ReLU 激活函数.

1.3 表观特征提取

在车辆重识别过程中,车辆的全局表观特征是从整体上识别车辆,在特定环境下具有较强的判断力,是必不可少的重识别特征;但在环境影响下,同一个车辆的整体表观特征差异较大,由于全局特征缺乏对局部细节的细节描述,在一定程度上降低了表达力. 而局部特征是对车辆某个部分的具体描述,例如车辆前部的车标、车尾的形状、两侧的外观等,对车辆的特征表达比较细腻具体,区分度明显,能补充增加整体特征的表达力;尤其是在处理相同颜色、相同车型整体属性的车辆时,局部特征更是发挥了关键的辅助“点睛”作用. 基于此,本文在提取车辆表观特征时,除了提取车辆的全局表观特征外,同时分别沿着水平方向和垂直方向将车辆划分成左、中、右和上、中、下等不同的区域提取对应的局部特征,这里各区域的划分不是平均分配,而是有所重叠. 每个区域的划分量占整体的 1/2,其中水平方向的中部和垂直方向的中部各被使用 2 次(如图 1 所示的虚线框部分). 最后将各局部特征与整体特征进行融合,得到最终的多尺度车辆表观特征.

1) 全局特征提取. 全局特征从整体上表示车辆的表观信息,与属性特征提取相类似,本文同样通过一个全局的池化层将初始特征向量 z 转换成一维的全局特征向量 f_g :

$$f_g = P(z). \quad (2)$$

2) 水平方向的局部特征提取. 车辆前部的车标、

车窗等局部区域能够提供较为具体的车辆信息,这些局部信息对车辆的重识别非常有帮助,但精确地提取这些信息需要大量的人工标注,工作量大. 通过观察,本文中发现一些比较固定的事实可利用,如车辆的车窗通常位于整个图片的上半部分,车标和车灯位于图片的下半部分等.

因此,本文在提取水平方向局部特征时,首先沿着水平方向把 z 重叠的部分分成左、中、右 3 个局部特征,即 $h_1 = z[0:7;:::]$ 、 $h_2 = z[4:11;:::]$ 、 $h_3 = z[8:15;:::]$; 然后通过池化操作分别把 h_1 、 h_2 和 h_3 转换成一维特征向量;最后使用求和操作把这 3 个局部特征融合成表示整个车辆水平方向的局部特征 f_h :

$$f_h = P(h_1) + P(h_2) + P(h_3). \quad (3)$$

3) 垂直方向的局部特征提取. 在提取垂直方向的局部特征时,本文采用与提取水平方向的局部特征相类似的方法. 首先沿着垂直方向把 z 有重叠的分成上、中、下 3 个局部的特征,即 $v_1 = z[:,0:7;:]$ 、 $v_2 = z[:,4:11;:]$ 、 $v_3 = z[:,8:15;:]$; 接下来也采用池化与求和操作得到融合后的垂直方向的局部特征 f_v :

$$f_v = P(v_1) + P(v_2) + P(v_3). \quad (4)$$

在计算得到 f_g 、 f_h 和 f_v 之后,将这 3 种特征进行融合,得到总的表观特征 $f_{\text{appearance}}$. 本文中主要通过一个串联层和一个 FC 层计算得到车辆最终的 $f_{\text{appearance}}$. 首先使用串联层得到一个高维的特征向量,再通过 FC 层将高维的特征向量进行降维,同时保持与属性特征的维度一致. 具体的计算式为:

$$f_{\text{appearance}} = \delta(W_2 [f_g, f_h, f_v]), \quad (5)$$

其中 $W_2 \in \mathbf{R}^{512 \times 6144}$ 为表观特征对应 FC 层的参数矩阵.

1.4 特征融合

在完成属性特征与多尺度表观特征的提取工作之后,通过注意力的加权求和将属性特征与表观特征进行融合. 首先,通过一个注意力机制,利用多尺度表观特征计算属性特征的加权权重,使用该权重对属性特征进行加权;再与多尺度表观特征进行融合得到车辆最终的特征向量 f :

$$f = f_{\text{appearance}} + \varphi_1(W_3 f_{\text{appearance}}) f_{\text{attribute}}. \quad (6)$$

其中: $W_3 \in \mathbf{R}^{1 \times 512}$ 为计算属性特征注意力权重的 FC 层参数; φ_1 是 Sigmoid 激活函数,将权重值转换到 $[0, 1]$ 之间.

1.5 优化目标函数

针对车辆的颜色与车型属性识别,本文在属性特征 $f_{\text{attribute}}$ 的后面增加两个并行 FC 层,用于计算将提

取的属性特征分类到不同的颜色和车型的概率:

$$p_{\text{color}} = \varphi_2(\mathbf{W}_{\text{color}} \mathbf{f}_{\text{attribute}}), \quad (7)$$

$$p_{\text{type}} = \varphi_2(\mathbf{W}_{\text{type}} \mathbf{f}_{\text{attribute}}), \quad (8)$$

其中, $\mathbf{W}_{\text{color}} \in \mathbf{R}^{C \times 512}$ 是颜色分类 FC 层的参数矩阵, $\mathbf{W}_{\text{type}} \in \mathbf{R}^{T \times 512}$ 是车型分类 FC 的参数矩阵, φ_2 是 Softmax 激活函数, C 和 T 分别是整个数据库中车辆颜色与车型的种数。

针对车辆 ID 标签分类部分, 本文在特征向量 \mathbf{f} 后面增加一个 FC 层, 用于计算将不同的车辆图片分类到对应的车辆 ID 类别的概率。在训练过程中, 将同一辆车的所有不同摄像头拍摄得到的车辆图片的 ID 分类标签设为一致, 作为一个相同的类别, 该类别的概率为:

$$p_{\text{ID}} = \varphi_2(\mathbf{W}_{\text{ID}} \mathbf{f}), \quad (9)$$

其中, $\mathbf{W}_{\text{ID}} \in \mathbf{R}^{N \times 512}$, 是车辆分类 FC 的参数矩阵, N 是训练集中所有不同车辆的 ID 数。

得到 p_{color} 、 p_{type} 和 p_{ID} 后, 使用交叉熵损失函数计算得出各自的分类损失 L_{color} 、 L_{type} 和 L_{ID} , 计算如下:

$$L = -\mathbf{q}^T \log(\mathbf{p}), \quad (10)$$

其中, \mathbf{q} 是训练数据的真实的独热(one-hot)标签, \mathbf{p} 是模型的输出概率。

最后再使用了一个多任务的优化目标函数用于优化整个网络模型的参数。该优化目标函数 L 由 3 个部分组成, 分别是车辆的 ID 分类损失 L_{ID} 、颜色分类损失 L_{color} 和车型的分类损失 L_{type} :

$$L = \lambda_1 L_{\text{ID}} + \lambda_2 L_{\text{color}} + \lambda_3 L_{\text{type}}, \quad (11)$$

其中, λ_1 、 λ_2 和 λ_3 为 L_{ID} 、 L_{color} 和 L_{type} 相应的权重。

另外, 在车辆重识别模型的训练过程中, 三元组 Triplet 损失函数^[18]也经常被用于减小相同 ID 车辆图片的欧式距离, 同时增加不同 ID 车辆图片之间的欧式距离。锚点车辆图片 a 、相同 ID 的其他车辆图片 p 、其他 ID 的车辆图片 n 构成三元组 (a, p, n) , 相应的 Triplet 损失函数为:

$$L_{\text{triplet}} = \sum_{\substack{a, p, n \\ y_a = y_p \neq y_n}} \max(m + D_{a, p} - D_{a, n}, 0). \quad (12)$$

其中: D 为两个样本间的欧式距离; m 是正负样本之间的欧式距离间隔, 在本文中取值为 0.3; y_a 、 y_p 、 y_n 分别是车辆图片 a 、 p 、 n 对应的 ID。

为进一步增加特征的判别力, 在优化目标函数 (11) 的基础上, 本文将 Triplet 损失函数引入模型的训练, 得到最终的优化目标函数:

$$L = \lambda_1 L_{\text{ID}} + \lambda_2 L_{\text{color}} + \lambda_3 L_{\text{type}} + \lambda_4 L_{\text{triplet}}, \quad (13)$$

其中 λ_4 是 L_{triplet} 对应的权重。

2 实验与结果分析

为验证本文所提算法模型的有效性, 在 VeRi-776 数据集上进行了训练与测试, 并与其他几个目前性能较优的算法进行对比。

2.1 VeRi-776 数据集

VeRi-776 数据集是由北京邮电大学构建的用于车辆重识别的数据集, 该数据集有城市监控场景下 20 个摄像头拍摄的 776 辆汽车的 50 000 多张图片, 将其其中 576 辆汽车的 37 778 张图片用于构建训练集, 剩下的 200 辆汽车的 11 579 张和 1 678 张图片分别用于构建测试集和查询集。数据库中每一辆车的多张图片均由 2~18 个摄像头在不同角度、光照、分辨率和遮挡情况下拍摄得到, 同时也标注了每一辆车的颜色与车型信息。车辆的颜色有黑、灰、白、红、绿、橙、黄、金、棕和蓝 10 种颜色; 车型有轿车、运动型多用途车 (SUV)、两厢车、多用途车 (MPV)、厢式车、皮卡车、公共汽车、卡车和房车 9 种车型。

2.2 评价指标

为评价模型的准确率, 本文采用平均精度均值 (mean average precision, mAP) 和 Rank- k 作为车辆重识别效果的评价指标。

1) mAP. 车辆重识别是一种图像检索任务, mAP 是图像检索任务中常用的评测指标, 通过综合考虑召回率和准确率来评价算法全局性能。对查询集中的某一张图片, 假设在测试集中与之相关的正确图片数为 k , 该图片检索的平均精度 (average precision, AP) p_{av} 定义如下:

$$p_{\text{av}} = \frac{1}{k} \sum_{i=1}^k \frac{i}{R_i}, \quad (14)$$

其中, R_i 指按照相似度从高到低排序后包含前 i 个正确检索结果所需的最少检索图片数。在计算得到查询集中每一张图片的 AP 之后, 对所有查询图片的 AP 求平均得到 mAP。

2) Rank- k . 车辆重识别任务就是要在车辆图片数据集中寻找与被查询车辆最相似的车辆。对于每一张被查询图像, 计算它与数据库中所有图像的相似度。Rank- k 即表示根据相似度进行排序后的结果中, 与被查询车辆可能属于同一 ID 的前 k 张图像。

2.3 实验参数设置

输入到模型的图片尺寸均缩放像素为 256×256 , 批大小为 64, 使用随机梯度下降法 Nesterov 训练参数,

新添加层的初始学习率为 0.1,其他预训练层的初始学习率为 0.01. 网络模型共训练 40 轮,在第 20 轮时把学习率乘以 0.1. 训练的过程中,本文也采用将图片随机翻转和随机擦除的数据增强方法. 式(11)和(13)中的 $\lambda_1, \lambda_2, \lambda_3$ 和 λ_4 的取值分别为 1, 0.5, 0.5 和 0.5. 在 Triplet 损失训练时,采用困难样本挖掘策略.

在进行车辆重识别检索测试时,本文采用 f 的相似度作为判别标准,相似度计算采用的度量距离为欧式距离.

2.4 实验结果分析

本文提出的算法模型在提取特征时融合了属性特征与表观特征,其中多尺度表观特征融合了根据水平方向和垂直方向划分的多个局部特征和全局特征的多尺度特征,因此在第一部分实验中,主要分析网络模型中不同特征模块组合对车辆重识别准确率的影响. 只包含单一全局表观特征,不包含属性特征、水平方向和垂直方向局部特征的基准模型记为 Baseline;在 Baseline 模型的基础上增加局部特征的模型记为 Baseline + Part,增加属性特征的模型记为

Baseline + Attribute;增加局部特征与属性特征的模型记为 Baseline + Part & Attribute.

各模型的准确率如表 1 所示. 单独增加局部特征和属性特征时,相比于 Baseline 模型, mAP 分别提高了 0.8 和 1.1 个百分点, Rank-1 提高了 1 和 1.2 个百分点;在同时增加这两个特征时, mAP 可达到 72.9%, Rank-1 达到 95.1%. 该结果表明本文构建的局部特征和属性特征对车辆重识别的准确率的提升有显著作用,且这两种特征互补.

除了对比增加不同特征模块对车辆重识别准确率的影响外,本文也计算了 Baseline + Part & Attribute 模型对于车辆颜色和车型分类的准确率. 在 VeRi-776 数据集的测试集与查询集上的颜色识别准确率分别为 94.7% 和 95.1%, 车型识别准确率分别为 93.3% 和 93.9%. 另外,通过对测试集上混淆矩阵分析发现:在颜色识别过程中,易出现将白色分类为灰色,棕色分类为黑色的错误;在车型识别过程中主要存在将 SUV 分类为轿车,两厢车分类为轿车的错误.

表 1 不同特征组合得到的 mAP 和 Rank- k 准确率
Tab. 1 The mAP and Rank- k accuracies of different feature combinations %

模型	mAP	Rank-1	Rank-5	Rank-10	Rank-20
Baseline	70.9	93.1	97.2	98.3	99.2
Baseline+Part	71.7	94.1	97.2	98.4	99.0
Baseline+Attribute	72.0	94.3	97.3	98.5	99.1
Baseline+Part & Attribute	72.9	95.1	97.6	98.9	99.2

接下来,本文比较了所构建的车辆重识别算法模型与近年来性能较优的算法模型的准确率,按照取得的 mAP 排序,结果如表 2 所示. 可以看出,本文提出的模型取得比其他大部分模型更高的 mAP. 其中, DenseNet121^[19]、VAMI^[20]、PROVID^[1]、VGG + CTS^[21] 和 VSTP^[14] 取得的 mAP 分别为: 45.1%, 50.1%, 53.4%, 58.3% 和 58.8%, 远低于本文采用的 ResNet-50 骨干模型的准确率 78.2%. 其中 DenseNet121^[19] 和 VGG + CTS^[21] 采用的 DenseNet121 和 VGG 模型,其提取的基础特征判别力低,从而导致最终的准确率偏低;VAMI^[20] 首先为每张输入图像提取单视图特征,再使用单个角度的特征生成多角度的特征,最终得到全局多视图特征,但生成多角度的特征与真实的多角度特征还是存在一定的差异,因此效果一般;PROVID^[1] 和 VSTP^[14] 结合时间或空间信

息,虽然能从时序上对检索结果进行一定的重排序,但是所使用的基础网络判别力以及时序信息的计算会增加相应的计算复杂度. SSL^[22] 通过(GAN)生成大量的虚拟样本用于数据扩充,并使用半监督的方法进行训练,可以在一定程度提高识别的准确率,但使用 GAN 会增加模型的训练时间和计算复杂度. RAM^[12] 提出了 4 个分支构成的深度模型,也是通过把整体特征和区域特征进行联合起来,用于提取更详细、更具辨别力的特征,但该模型提取的特征维度更高,易造成对训练数据的过拟合. QD-DLF^[23] 设计四向深度学习网络将基本特征映射压缩为水平、垂直、对角线和反对角方向特征图,最后将这些特征归一化用于获取多维度的表观特征,但该方法并没有考虑车辆的属性等相关信息. MTCRO^[24] 使用多任务卷积神经网络和新型排序优化的方法,可以在一定程度上提高特征提

取后的检索精度. BS^[25] 中使用了联合 Triplet 损失函数与交叉熵分类损失函数训练网络模型, 可以取得比单独使用交叉熵分类损失函数更优的性能, 但整体性能并未见优势. JDRN^[26] 提出了一种多域联合学习框架, 通过利用多个数据集的车辆图片训练一个更加鲁棒的模型, 该方法虽然可以提高准确率, 但对数据集要求更高, 需要利用多个不同的数据集来训练模型. 最后, 我们注意到: 由于 MRL+Softmax^[27] 在增加不同车辆之间基础上通过度量学习, 使相同车辆在不同视角下的距离相对增大, 使用 Softmax 函数训练后在 Veri-776 数据库上的 mAP 和 Rank-5 可以达到 78.5% 和 99.0%, 超过了本文的 78.2% 和 97.0%, 但 MRL+Softmax 通过聚类估计车辆视角时会引入新的数据噪声, 导致 Rank-1 指标上要低于本文的 95.8%, 其次 MRL+Softmax 引入了视角约束也会增加模型的训练复杂度.

另外, Re-ranking^[28] 被作为一种后处理的方法, 被用于提高重识别排序的准确率, 在本文中也将 Re-ranking 用于提高车辆重识别的准确率. 从表 2 中可以看到, SSL^[22] 与 JDRN^[26] 以及本文所提方法在使用 Re-ranking 进行后处理之后, mAP 均有提高, 其中, 本文方法从 75.0% 提高到 78.2%, 优于 SSL^[22] 与 JDRN^[26] 分别使用 Re-ranking 后的准确率 69.9% 和 73.1%.

最后, 随机选取 7 个不同车辆, 使用本文方法结合 Triplet 损失进行训练, 重识别结果按照相似度从

表 2 不同车辆重识别方法的对比

Tab. 2 The comparison of different

Re-identification methods		%		
方法	mAP	Rank-1	Rank-5	
DenseNet121 ^[19]	45.1	80.3	91.1	
VAMI ^[20]	50.1	77.0	90.8	
PROVID ^[1]	53.4	81.6	95.1	
VGG+CTS ^[21]	58.3	83.5	90.0	
VSTP ^[14]	58.8	86.4	92.9	
SSL ^[22]	61.1	88.6	93.6	
RAM ^[12]	61.5	88.6	94.0	
QD-DLF ^[23]	61.8	88.5	94.5	
MTCRO ^[24]	62.6	88.0	94.6	
BS ^[25]	67.6	90.2	96.4	
JDRN ^[26]	69.1			
SSL+Re-ranking ^[22]	69.9	89.7	95.4	
JDRN+Re-ranking ^[26]	73.1			
MRL+Softmax ^[27]	78.5	94.3	99.0	
本文方法	72.9	95.1	97.6	
本文方法+Triplet	75	95.4	98.3	
本文方法+Triplet+re-ranking	78.2	95.8	97	

注: 表中缩写词的中文注释见附录 (<http://jxmu.edu.cn/upload/20210110>) 表 S1.

高到低进行可视化排序, 如图 2 所示. 可以看出, 本文构建的方法可以准确地重识别不同角度、不同光照条



图 2 本文所构建模型相似度排名前 10 的检索结果

Fig. 2 The top-10 retrieval result of our method

件拍摄的另一车辆,同时也发现通过红色框标注出来的识别错误主要集中在相同车型或相同颜色的车辆之间的区分错误。

3 结 论

本文尝试将车辆的颜色、车型属性信息融合到多尺度车辆外观信息来获得既具判别力又具鲁棒性的车辆特征。在改良的 ResNet-50 网络计算基础特征向量的基础上,用两个分支分别计算车辆属性特征和多尺度外观特征,再进行融合得到最终特征向量,通过多任务优化目标函数对网络参数进行优化。在 VeRi-776 测试数据集上的实验结果证明了本文方法的识别效果远远优于目前的大多数性能较优的方法。然而,本文方法也不可避免地存在不足之处:在颜色识别过程中易将白色分类为灰色,棕色分类为黑色;在车型识别过程中的错误主要有将 SUV 分类为轿车,将两厢车分类为轿车等。后续的研究工作中将尝试解决这些问题。

参考文献:

- [1] LIU X, LIU W, MEI T, et al. A deep learning-based approach to progressive vehicle re-identification for urban surveillance [C] // European Conference on Computer Vision. Amsterdam: Springer, 2016: 869-884.
- [2] 张涛,费树岷,李晓东,等. 基于 GA-RBF 神经网络及边界不变特征的车辆识别[J]. 智能系统学报, 2009, 4(3): 278-272.
- [3] 姚会,苏松志,王丽,等. 基于改进的混合高斯模型的运动目标检测方法[J]. 厦门大学学报(自然科学版), 2008, 47(4): 505-510.
- [4] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436.
- [5] 王盼盼,李玉惠. 基于特征融合和 L-M 算法的车辆重识别方法[J]. 电子科技, 2018, 31(4): 12-15.
- [6] 李熙莹,周智豪,邱铭凯,等. 基于部件融合特征的车辆重识别算法[J]. 计算机工程, 2019(6): 12-20.
- [7] LIU X, LIU W, MA H, et al. Large-scale vehicle re-identification in urban surveillance videos [C] // International Conference on Multimedia and Expo. Seattle: IEEE, 2016: 1-6.
- [8] LIU H, TIAN Y, WANG Y, et al. Deep relative distance learning: tell the difference between similar vehicles[C] // IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 2167-2175.
- [9] ZHOU Y, LIU L, SHAO L. Vehicle re-identification by

deep hidden multi-view inference[J]. IEEE Transactions on Image Processing, 2018, 27(7): 3275-3287.

- [10] KANACI A, ZHU X, GONG S, et al. Vehicle re-identification in context [C] // German Conference on Pattern Recognition. Stuttgart: Springer, 2018: 377-390.
- [11] YAN K, TIAN Y, WANG Y, et al. Exploiting multi-grain ranking constraints for precisely searching visually-similar vehicles[C] // IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 562-570.
- [12] LIU X, ZHANG S, HUANG Q, et al. RAM: a region-aware deep model for vehicle re-identification[C] // IEEE International Conference on Multimedia and Expo. San Diego: IEEE, 2018: 1-6.
- [13] BAI Y, LOU Y, GAO F, et al. Group-sensitive triplet embedding for vehicle re-identification [J]. IEEE Transactions on Multimedia, 2018, 20(9): 2385-2399.
- [14] SHEN Y, XIAO T, LI H, et al. Learning deep neural networks for vehicle re-ID with visual-spatio-temporal path proposals[C] // IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 1900-1909.
- [15] HE B, LI J, ZHAO Y, et al. Part-regularized near-duplicate vehicle re-identification [C] // IEEE Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 3997-4005.
- [16] CHEN H, LAGADEC B, BREMOND F. Partition and reunion: at wo-branch neural network for vehicle re-identification[C] // IEEE Conference on Computer Vision and Pattern Recognition Workshops. Long Beach: IEEE, 2019: 184-192.
- [17] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C] // IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 770-778.
- [18] SCHROFF F, KALENICHENKO D, PHILBIN J. FaceNet: a unified embedding for face recognition and clustering [C] // IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015: 815-823.
- [19] HUANG G, LIU Z, MAATEN L V D, et al. Densely connected convolutional networks[C] // IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 2261-2269.
- [20] ZHOU Y, SHAO L. Viewpoint-aware attentive multi-view inference for vehicle re-identification [C] // 2018 IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 6489-6498.
- [21] ZHANG Y, LIU D, ZHA Z J. Improving triplet-wise training of convolutional neural network for vehicle re-identification [C] // IEEE International Conference on Multimedia and

- Expo. Hong Kong: IEEE, 2017: 1386-1391.
- [22] WU F, YAN S, SMITH J S, et al., Vehicle re-identification in still images: application of semi-supervised learning and re-ranking[J]. Signal Processing: Image Communication, 2019(76): 261-271.
- [23] ZHU J, ZENG H, HUANG J, et al. Vehicle re-identification using quadruple directional deep learning features[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(1): 410-420.
- [24] XU D, LANG C, FENG S, et al. A framework with a multi-task CNN model joint with a re-ranking method for vehicle re-identification [C] // International Conference on Internet Multimedia Computing and Service. Nanjing: ACM, 2018: 1-7.
- [25] KUMAR R, WEILL E, AGHDASI F, et al. Vehicle re-identification: an efficient baseline using triplet embedding [C] // International Joint Conference on Neural Networks. Budapest: IEEE, 2019: 1-9.
- [26] LIU C T, LEE M Y, WU C W, et al. Supervised joint domain learning for vehicle re-identification [C] // IEEE Conference on Computer Vision and Pattern Recognition Workshops. Long Beach: IEEE, 2019: 45-52.
- [27] LIN W, LI Y, YANG X, et al. Multi-view learning for vehicle re-identification [C] // IEEE International Conference on Multimedia and Expo (ICME). Shanghai: IEEE, 2019: 832-837.
- [28] ZHONG Z, ZHENG L, CAO D, LI S. Re-ranking person re-identification with k -reciprocal encoding [C] // IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 3652-3661.

A vehicle re-identification method by fusing the vehicle appearance and attribute information

XIE Xiuzhen^{1,2}, LUO Zhiming^{1*}, LIANG Sheng¹, LI Shaozi¹

(1. School of Informatics, Xiamen University, Xiamen 361005, China; 2. College of Mathematics and Information Engineering, Longyan University, Longyan 364002, China)

Abstract: Although global features extracted from the deep learning-based model have secured strong discrimination abilities, they endure lack of descriptions about local details. Besides, the appearance of the same car in different environments changes significantly, thus increasing the difficulty of vehicle re-identification. Therefore, in this study, we propose a vehicle re-identification model which integrates multi-scale vehicle appearance and attribute information. Due to the robustness of the attribute information in different environments, we fuse attribute features into the appearance features to obtain an enhanced feature for accurate vehicle re-identification. Finally, a multi-task loss function containing the cross-entropy loss and the triplet loss is adopted to train model parameters. We conduct extensive experiment on the large-scale VeRi-776 dataset, and experimental results demonstrate the effectiveness of our proposed method, which can significantly improve the accuracy and outperform most compared methods.

Keywords: vehicle re-identification; feature extraction; appearance information; attribute information; feature fusion