

DOI:10.19651/j.cnki.emt.2107881

车辆姿态感知注意力增强的车辆重识别

朱肖磊 吴训成

(上海工程技术大学 机械与汽车工程学院 上海 201620)

摘要: 由于不同道路监控视角下的车辆姿态不断变化,因此车辆重识别仍是智慧交通系统中一项具有挑战性的任务。现有的车辆重识别的方法大多数基于车辆的外观属性,但识别受光照和角度等因素影响导致识别效果较差。因此,设计了一种车辆姿态感知注意力增强网络以提高车辆在光照和角度等因素影响下的重识别效果。首先,将图片输入到卷积姿态网络中生成12个关键点重建车辆姿态信息,然后将输入图像车辆与目标图像车辆进行比较,提取出两辆车公共区域的特征;最后,计算车辆全局特征和局部特征之间的距离,并根据最终结果对识别结果进行排序。本文在Vehicle ID和VeRi776数据集上进行验证,实验结果表明,所提出的网络相较于其他模型top10的检测准确率提高了10%左右。

关键词: 关键点;车辆姿态;注意力机制;车辆重识别

中图分类号: TP391.41; U495 **文献标识码:** A **国家标准学科分类代码:** 580.2010

Enhanced attention of vehicle posture perception for vehicle re-identification

Zhu Xiaolei Wu Xuncheng

(School of Mechanical and Automotive Engineering, Shanghai University of Engineering Sciences, Shanghai 201620, China)

Abstract: Owing to the continuous changes in vehicles under different road monitoring perspectives, vehicle re-identification is still a challenging task in intelligent traffic system. Most of the existing vehicle re-identification methods are based on the appearance attributes of the vehicle, but the recognition is affected by factors such as illumination and angle, which leads to poor recognition results. Therefore, this paper designs a vehicle posture perception attention enhancement network to improve the re-identification effect of vehicles under the influence of factors such as illumination and angle. First, input the image to the convolutional pose machine to generate 12 keypoints to reconstruct the vehicle frame, and then compare the input image vehicle with the target image vehicle to extract the features of the intersecting area between two images. Finally, the global distance and local loss of vehicle features are calculated, and the recognition results are sorted according to the final results. This paper verifies on Vehicle ID and VeRi776 data sets. The experimental results prove that the top10 detection accuracy of the proposed network is increased by about 10% than other models.

Keywords: keypoints; vehicle posture; attention mechanism; vehicle re-identification

0 引言

车辆重识别是一种跨相机车辆跟踪技术,因此,车辆重识别任务在智能交通、刑事侦查和重大事件信息安全等方面具有广泛的应用。但是,在监控视角下车辆的视觉外观在不同的车辆姿态时有很大的不同,相反,从相同的角度看,两辆相似的车外观变化又很小。因此,从多个监控摄像头的外观、姿态、轨迹等方面搜索相同的车辆协助社会公共卫生安全的建立具有一定的重要意义。

为了解决上述两个问题,目前的方法主要集中在元信

息(例如车辆属性,时空信息)来提高特征的代表能力,文献[1]就是采用一种基于属性的车辆重识别方法,它训练了具有各种属性的模型,并在数据集中搜索了具有相似属性的车辆,文献[2]使用三维边界框来纠正车辆的不同图像,然后提取颜色直方图,并将他们连接到一个特征向量,最后,他们训练了SVM分类器来判断两个图像是否具有相同的ID。最近,许多深度学习方法应用于车辆重识别,文献[3]介绍了一种基于孪生网络的度量学习方法,文献[4]采用一种区域感知深度模型进行车辆重识别,文献[5]则考虑了输入邻域之间的差异,文献[6]提出了一种

精细的搜索框架,通过车辆属性和时空信息建模来进行车辆重识别;文献[7]引入了一个深层网络,将摄像机的视野、车辆类型和颜色融合到车辆的特征中,上述方法都侧重于学习车辆的全局表示。但是,在不同的监控视角下,整体外观会发生巨大变化,这导致全局功能的不稳定,并且也带来了第 1 个挑战。相反,局部特征通常提供稳定的判别线索,从而有利于获取更多车辆区别性特征。文献[8]基于车辆的车窗、车顶等局部特征并采用了焦点融合的方法对车辆进行了重识别;文献[9]基于车辆关键点检测生成方向不变特征、文献[10]基于车辆的 3 个均匀分离的区域提取局部特征以获得显著性的车辆局部特征。文献[11]通过 YOLO 检测器检测每辆车的车窗、灯光和品牌,以产生判别特征。以上方法着重于预定义的区域,以学习细微的局部提示。但是,显著的提示(例如排气、贴纸和装饰物)由于其形状小而不宜区分,并且可能出现在车辆的任何部分,增加了检测的复杂性,这导致了第 2 个挑战。上述方法只是从车辆的外观信息来进行重识别,而忽略了车辆的姿态等属性。

目标关键点定位具有许多重要的应用,例如,面部对

齐^[12-13]、人体姿势估计^[14-15]和目标定位^[16]等。在大多数人脸识别框架中都进行基于关键点的人脸对齐^[17]。关键点的位置非常有用,因为可以通过关键点很好地对齐学习的功能。但是,在现有文献中,对车辆关键点的研究还不够深入。本文提出的方法表明,车辆关键点可以指导学习和对齐,并提高整体车辆重识别的性能。

因此考虑到同一车辆在同一视角下的姿态变化较大,本文提出了一个车辆姿态感知注意力增强框架,首先,将图像输入到 CPM 网络中判别车辆的姿态;其次,对输入图像与目标图像的交叉区域进行过滤,通过注意力网络对公共区域进行特征提取;最后,计算车辆全局特征和局部特征之间的距离,按相似度从高到低排序,筛选出相似度排名前 10 的目标车辆。

1 方 法

本文提出了一种车辆姿态感知注意力增强的车辆重识别方法,包括车辆姿态判别、公共区域特征提取以及全局与局部特征的相似度计算 3 部分。该方法的网络架构如图 1 所示。如下是对每个部分的详细介绍。

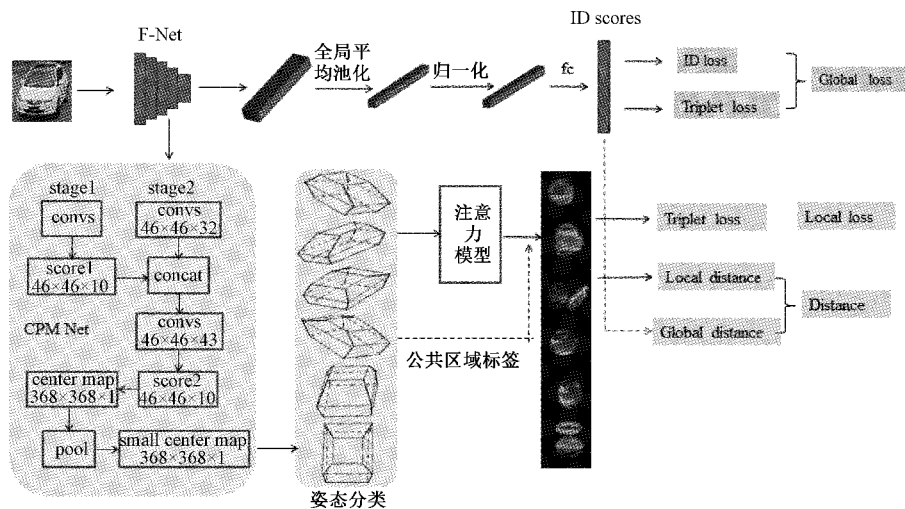


图 1 网络框架

1.1 车辆姿态重建

在本节中,主要介绍车辆框架重建以及姿态判别的方法,利用关键点检测和形状调整的框线来进行车辆姿态的判别。假设每辆车在 3D 中是由 12 个关键点的线框组成,这些关键点对应于车辆上易于识别的位置,比如前灯、反光镜、尾灯、车顶拐角、车轮等,如图 2 所示,圆点即为关键点。

给定输入的车辆图像,使用深度学习网络预测车辆的语义关键点,CNN 输出 12 个带有突出显示的关键点的热图,接着,将输出热图作为输入,并显示热图中最大概率的关键点,关键点将车辆框架划分为 3 个平面,顶层、中间层和底层。第 1 层为前挡风和后挡风玻璃的上方角点;第 2

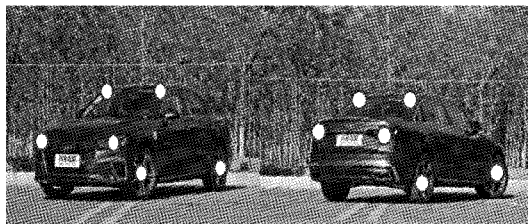


图 2 关键点示意图

层为左前灯、左后灯、右前灯、右后灯;第 3 层为左前轮、左后轮、右前轮、右后轮。根据图像检测出来的关键点构成的车辆框架与形状框线模型进行对比分析,从而判断出车辆所处姿态。车辆可以粗略的看成是具有 8 个表面的模

型,车辆底部在摄像头的下方是不可见的。8个部分(除车辆底部外)进行不同颜色的区域划分,分为顶部、前部、后部和侧面4部分,如图3所示,以便在后续中提取车辆的公共区域的特征。

在监控视角下,车辆的左侧和右侧通常是不能同时出现的,并且在视觉上通常是对称的,基于这些观察,本文将车辆的姿态分为正面、背面、正右侧、背右侧、正左侧和背左侧6个车辆姿态,如图4所示,将判别后的车辆姿态作为第2部分特征提取的已知条件。

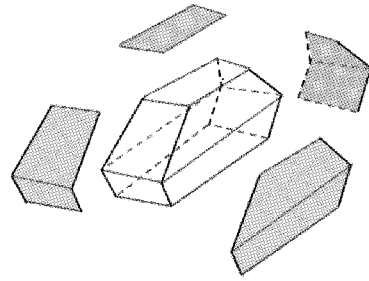


图3 区域颜色划分

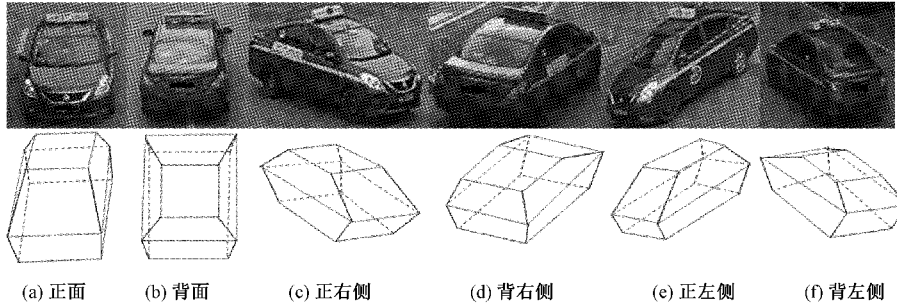


图4 车辆的6个姿态

1.2 公共区域的特征提取

在监控视角下,图像能够覆盖整辆车,因此可以捕获两辆车之间的每一个细微差异;其次,在大多数监控视角下,图像中的车辆框架的3个部分是可见的,因此目标车辆与输入车辆之间至少有两个相同的部分。在上述阶段之后,就可以获得车辆的公共区域的特征,本文采用公共注意力网络,以增强公共区域的特征,这有助于捕获同一车辆的稳定判别信息。

本文提出的该注意力模型能够自动选择显著的区域,并且删除无用的信息,在车辆重识别问题中,本文模型需要关注输入车辆和目标车辆之间车辆重叠的区域,例如,为了从正左侧和正右侧的角度分辨两辆相似的车辆之间的区别,人们通常会关注车辆的共同的外观部分(比如车辆正面和顶部)以区分两辆车是否相同。因此,提出了一种车辆姿态感知的注意力模型。本文提出的注意力网络主要由3部分组成,如图1网络框架所示。首先,通过车辆的外观属性等训练深层的CNN网络,从而筛选出与目标车辆相似的车辆;其次,根据筛选出的相似车辆的图片,对车辆的关键点进行检测,从而重建车辆的线性框架图,进行车辆的姿态判别;最后,将输入车辆姿态与目标车辆姿态进行比较,从而获取公共区域。不同姿态的车辆之间的公共区域如表1所示。根据表格,提取公共区域的特征,为了获取不同区域的特征向量,本文选择了F网络层的第4个连接层,因为它具有高级感知器并保持足够大的空间大小,输入图像公共区域可以表示为 $\{s_1, s_2, s_3, \dots, s_n\}$ 特征向量的集合,其中, n 表示图像中不同区域的数量, s_n 是第 n 个区域的256维特征向量,通过学习由标签指示的弱监督的上下文向量来生成

注意力图,该标签指示的是输入图像和目标图像之间的共同区域。通过式(1),步骤 t 的上下文向量可以参与输入图像的某些区域。

$$c^t = \text{Attention}(c^{t-1}, \{s_n\}_{n=1}^N, v) \quad (1)$$

其中, c^{t-1} 是步骤 $t-1$ 的上下文向量, v 表示目标图像与输入图像公共区域的特征。采用软注意力机制,将所有输入特征向量的加权平均值用于计算上下文向量。注意权重 $\{\alpha_n^t\}_{n=1}^N$ 是通过两层非线性转换和softmax函数来计算的:

$$h_n^t = \tanh(W_c^t(c^{t-1} \odot v) + b_c^t) \odot \tanh(W_u^t s_n + b_u^t) \quad (2)$$

$$\alpha_n^t = \text{softmax}(W_h^t h_n^t + b_h^t) \quad (3)$$

$$c^t = \sum_{n=1}^N \alpha_n^t s_n \quad (4)$$

其中, W_c^t, W_u^t, W_h^t 和偏置项是可学习的参数, h_n^t 是隐藏状态, \odot 表示逐元素相乘。上下文向量 c^0 通过如下方式初始化:

$$c^0 = \frac{1}{N} \sum_{n=1}^N s_n \quad (5)$$

本文的注意力模型主要是由输入图像和目标图像之间的共同区域的标签来弱监督的,两辆车之间的公共区域如表1所示,括号内按顺序依次代表的是正面、背面、侧面和顶面。其中,侧面表示车辆的左侧和右侧。当两个图像有公共区域时,则设置为1,否则设置为0。通过交叉熵优化了注意力损失,如果目标图像是正右侧,输入图像为正面,那么公共区域为正面和顶面,则标签记为(1,0,0,1),一旦对注意力模型进行训练,就会输出一个注意力图,该注意力图只会输入车辆与目标车辆的公共区域出现高响应。

表 1 车辆公共区域标签

输入图像	查询图像					
	正面	背面	正右侧	正左侧	背左侧	背右侧
正面	(1,0,0,1)	(0,0,0,1)	(1,0,0,1)	(1,0,0,1)	(0,0,0,1)	(0,0,0,1)
背面	(0,0,0,1)	(0,1,0,1)	(0,0,0,1)	(0,0,0,1)	(0,1,0,1)	((0,1,0,1)
正右侧	(1,0,0,1)	(0,0,0,1)	(1,0,0,1)	(1,0,0,1)	(0,0,0,1)	(0,0,1,1)
正左侧	(1,0,0,1)	(0,0,0,1)	(1,0,0,1)	(1,0,1,1)	(0,0,1,1)	(0,0,0,1)
背左侧	(0,0,0,1)	(0,1,0,1)	(0,0,0,1)	(0,0,1,1)	(0,1,1,1)	(0,1,0,1)
背右侧	(0,0,0,1)	(0,1,0,1)	(0,0,1,1)	(0,0,0,1)	(0,1,0,1)	(0,1,1,1)

1.3 车辆特征相似度计算

在上述阶段之后,本文获得了注意力模型提取出来的局部特征,本节将结合全局特征对图像进行相似度计算。对于给定的输入图像 p 与目标图像 q 以及他们对应输出的注意力特征图 M_p, M_q , 计算其相似度得分 v_p^i 和 v_q^i , 这表示每个对应区域相似度的大小, 相似度得分定义为 $v_i = \sum_{j,k=1}^{16} M_i(j,k)$, 本文引入公共注意力机制来计算公共区域的得分 $score_i^{p,q}$, 则:

$$score_i^{p,q} = \left(\sum_{i=1}^N v_i^p v_i^q \right)^{-1} v_i^p v_i^q \quad (6)$$

其中, $score_i^{p,q}$ 测量常见可见区域的一致性。

然后,两辆车之间的局部区域特征距离 \hat{D} 计算为:

$$\hat{D}^{p,q} = \sum_{i=1}^N score_i^{p,q} D(f_i^p, f_i^q) \quad (7)$$

其中, D 表示欧几里得距离。

如果车辆某些区域不是公共区域,则相应区域的分数将相对较小,因此,只有得分高的特征图才有助于最终距离,本文通过构建全局特征的 ID 损失和三元组损失以及局部特征的三元组损失来优化网络,根据上述局部特征的距离计算三元损失,则:

$$L_{triplet}^l = \max(\hat{D}^{ap} - \hat{D}^{an} + \gamma, 0) \quad (8)$$

最后,该注意力机制网络的目的是最大程度的减少后续损失,则:

$$L = L_{id}^g + L_{triplet}^g + L_{triplet}^l \quad (9)$$

2 实 验

2.1 数据集

本文选用 Vehicle ID 和 VeRi776 数据集来验证所提出方法的性能。

Vehicle ID: 是由监控相机收集的,该数据集包含 26 267 种车辆以及 221 763 张车辆图像,该数据集的训练集由 13 182 辆车的 100 182 张图像组成,测试集由 2 400 辆车的 19 777 张图像组成,Vehicle ID 中的图像是从车辆前方视角和车辆后方视角中捕获的,它还提供了 250 种车辆模型的注释。

VeRi776:是从实际监控场景中 20 个不重叠的交通监

控摄像头中收集的,它包含 776 辆车的 51 035 个边界框,该数据集划分了 576 辆车用于训练,200 辆车用于测试,该数据集中的车辆包含 3 个视点,即正视图、后视图和侧视图, VeRi-776 提供如下信息的注释:ID、型号、车辆颜色、摄像机之间的关系以及轨迹信息。

2.2 评价指标

受文献[18]行人重识别的研究启发,本文采用 top@1, top@5, top@10 的平均精度(mAP)作为本实验的评估指标。

top@k:在测试集中查询给定的一个目标图像, top@k 表示查询结果中前 k 个图像是正确结果的概率,假设车辆的显著性特征有 N 处,那么输入图像的特征可以表示为 $f_{Inquire} = [q_1, q_2, \dots, q_n]$ 。

测试集包含 M 个图像,则所有图像的特征可以表示为:

$$f_{Test} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ a_{M1} & a_{M2} & \dots & a_{MN} \end{bmatrix} \quad (10)$$

本文使用欧几里得距离来计算图像特征之间的相似度,欧几里得距离可以表示为:

$$S_M = \sqrt{\sum_{k=1}^N (q_k - a_{MK})^2} \quad (11)$$

其中, q_k 表示查询图像的特征向量, a_{MK} 是测试集中第 M 个图像的特征向量, S_M 越小,表示图像间的相似度越高,因此,首先根据 S_M 从低到高的顺序排序,然后计算 top@k, $k=1, 5, 10$ 的精度。

平均精度(mAP)度量标准评估车辆重识别的整体模型性能,给定目标图像,平均精度为:

$$AP = \frac{\sum_{k=1}^M top@k \times gt(k)}{N_{gt}} \quad (12)$$

其中, N_{gt} 是真实值的数量, $gt(k)$ 是指示第 k 个回归图像是否正确的标记函数。因此,对于所有目标图像, mAP 的计算如下:

$$mAP = \frac{\sum_{q=1}^Q AP(q)}{Q} \quad (13)$$

其中, Q 表示输入图片的数量。

2.3 实验分析

1) 关键点回归分析

在本节中,将根据回归的精度以及标签与车辆姿态之间的关系对关键点回归器进行全面的研究。为了训练和评估关键点回归器,在整个 VeRi776 数据集图像上手动标注 12 个关键点位置,如图 3 所示。在测试阶段,提取测试图像的响应图,并预测具有最大响应的关键点位置。如果回归的关键点位置与真实位置之间的距离小于阈值 r_0 , 则认为该点为正确的预测的关键点。否则为错误的预测关键点,在评估阶段中不可见的关键点将被忽略。本文采用两个损失函数来训练关键点回归模型,如表 2 所示,由实验结果可知,在 $r_0=3$ pixel 内正确预测出关键点的概率是 89.2%。如图 5 所示为车辆关键点的回归结果。

表 2 关键点回归模型的评估结果

模型	$r_0=5$	$r_0=3$
L2 损失	91.56	87.91
交叉熵损失	92.15	89.22

表 3 不同模型在 VeRi776 数据集上的准确度

准确度	top@5	top@10	top@20	top@30	top@40	top@50
外观属性	0.790	0.836	0.878	0.894	0.907	0.915
多视点生成模型	0.811	0.857	0.903	0.924	0.938	0.945
多视图模型	0.861	0.904	0.942	0.958	0.968	0.976
本文模型	0.918	0.945	0.966	0.977	0.982	0.986

2) 注意模型的作用

由于输入图像的每个车辆的姿态都不相同,对于车辆重识别任务相对困难,因此,本文考虑关注输入车辆与目标车辆之间的公共区域,从而探索注意力机制在基线中的意义。如表 4 所示,如果不使用注意力机制,mAP 会降低 10%左右,由于注意力模型可以通过深度 k 步建立,因此,对于变量 k 进行进一步评估以选出最佳性能的 k 值,表 4 显示了当 $k=2$ 时达到了最高的 mAP,因此,本文模型的 k 取值为 2。如果不使用注意力模型,则结果不令人满意。

表 4 注意力模型评价(k 是注意步骤)

基准	mAP	top@1	top@5	top@10
无注意机制	40.36	70.01	82.16	91.85
$k=1$	43.87	73.59	88.63	94.31
$k=2$	52.39	77.43	91.15	98.37
$k=3$	46.51	75.22	88.29	94.61
$k=4$	41.72	71.94	83.44	92.87

3) 全局和局部损失权重的选择

本节通过实验验证了车辆姿态感知增强注意力模型对车辆重识别性能的影响,如表 5 所示为全局损失和局部

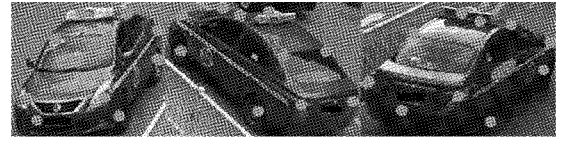


图 5 关键点回归结果

本文还对关键点位置和方向分类之间的关系进行了研究,由于 VeRi776 数据集没有提供方向信息,因此为 VeRi776 数据集手动标注了方向信息,即正面、背面、正左侧、正右侧、背左侧和背右侧,利用带有注释的图像进行训练,通过使用关键点响应图作为输入来训练车辆方向分类器,经过训练的分类器在测试图像上的准确度达到 93.2%,这表明关键点响应图包含足够的信息来推断车辆的方向,本文还验证了根据此方法判别车辆姿态的合理性。如表 3 所示为不同模型在 VeRi776 数据集上的验证效果,从表 3 可以看出,本文模型总体上呈现上升趋势,当查询车辆数目越多时,模型趋于稳定。从表格的纵向来看,在 top@10 处本文模型相对于其他模型有明显地提升,提升的幅度在平均 10%左右。总的来说,车辆姿态感知模型优于车辆外观属性等方法。

损失之间不同权重对结果的影响,从表 5 中可以发现,本文所提出的车辆姿态注意力模型在 mAP、top@1、top@5、top@10 上都得到改善。当 $\lambda_1=1, \lambda_2=1$ 时 top@1 的准确度最高,当 $\lambda_1=1, \lambda_2=0.5$ 时 top@1 准确度低了 0.5%,但其他结果都达到最佳,因此本文选择了 $\lambda_1=1, \lambda_2=0.5$ 的模型。

表 5 在 VeRi776 上全局距离和局部距离权重的选择

λ_1	λ_2	mAP	top@1	top@5	top@10
1	0	0.787	0.955	0.982	0.990
1	0.3	0.794	0.956	0.984	0.992
1	0.5	0.795	0.956	0.984	0.992
1	0.7	0.794	0.959	0.984	0.992
1	1	0.793	0.961	0.984	0.991
0	1	0.713	0.922	0.967	0.992

最后,本文验证了车辆姿态感知增强注意力模型中每一部分在 Vehicle ID 数据集上的准确度。如表 6 所示,从表 6 中可以看出,本文模型在整个 top 排序中准确度均高于其他模型,其中在 top@10 处出现幅度较大地提升。

无注意力机制模型在 top@20 前表现较优于其他模型,但在 top@20 后由于搜索范围变大,模型影响因素变多,鲁棒性反而降低;无度量计算和无姿态分类模块的模型虽

然整体的准确度不高,但一直呈现上升趋势。总的来说,本文所提出的模型的每一部分都起到了提升模型匹配度的效果。

表 6 在 Vehicle ID 数据集上验证模型每一部分的性能

准确度	top@5	top@10	top@20	top@30	top@40	top@50
无车辆姿态分类	0.794	0.839	0.885	0.913	0.933	0.949
无局部损失	0.828	0.874	0.918	0.953	0.965	0.973
无注意力机制	0.849	0.891	0.932	0.939	0.947	0.962
本文模型	0.867	0.946	0.965	0.975	0.981	0.982

3 结 论

本文提出了一种车辆姿态感知增强注意力模型,抓住车辆姿态的属性,基于车辆关键点重构车辆的姿态,通过 CPM 网络划分车辆的姿态,通过注意力机制改进了特征提取的方式,进一步提取出车辆重识别中显著性的特征,最终计算全局特征与局部特征的距离。本文所提出的模型有助于获得同一车辆的稳定判别信息。并且在数据集上的实验数据表明,本文所提出的模型相较于其他模型在 top10 准确度上提高了 10% 左右。该技术可广泛引用语智能公共安全防护、智能交通环境感知领域,对于追捕违法车辆、治安防控、刑事侦破等社会重大公共安全卫生具有重要意义。

参考文献

- [1] FERIS R S, SIDDIQUIE B, PETERSON J, et al. Large-scale vehicle detection, indexing, and search in urban surveillance videos [J]. IEEE Transactions on Multimedia, 2011, 14(1): 28-42.
- [2] ZAPLETAL D, HEROUT A. Vehicle re-identification for automatic video traffic surveillance [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2016: 25-31.
- [3] YI D, LEI Z, LIAO S, et al. Deep metric learning for person re-identification [C]. 2014 22nd International Conference on Pattern Recognition, IEEE, 2014: 34-39.
- [4] LIU X, ZHANG S, HUANG Q, et al. Ram: A region-aware deep model for vehicle re-identification [C]. The 2018 IEEE International Conference on Multimedia and Expo (ICME), Piscataway: IEEE Press, 2018: 1-6.
- [5] AHMED E, JONES M, MARKS T K. An improved deep learning architecture for person re-identification [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 3908-3916.
- [6] LIU X, LIU W, MEI T, et al. Provid: Progressive

and multimodal vehicle re-identification for large-scale urban surveillance [J]. IEEE Transactions on Multimedia, 2017, 20(3): 645-658.

- [7] WANG H, PENG J, CHEN D, et al. Attribute-guided feature learning network for vehicle re-identification [J]. IEEE MultiMedia, 2020, 27(4): 112-121.
- [8] 李浩, 杨超, 黄友新, 等. 基于局部特征和焦点融合的车辆重识别算法 [J]. 电子测量技术, 2021, 44(18): 167-174.
- [9] 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.
- [10] LIU X, ZHANG S, HUANG Q, et al. RAM: A region-aware deep model for vehicle re-identification [J]. IEEE, 2018: 1-6.
- [11] 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.
- [12] REN S, CAO X, WEI Y, et al. Face alignment at 3000 fps via regressing local binary features [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 1685-1692.
- [13] 唐杰, 陈仁文, 余小庆. 基于卷积神经网络的车载疲劳驾驶检测系统的实现 [J]. 国外电子测量技术, 2018, 37(1): 116-121.
- [14] NEWELL A, YANG K, DENG J. Stacked hourglass networks for human pose estimation [C]. European Conference on Computer Vision, Springer, Cham, 2016: 483-499.
- [15] WEI S E, RAMAKRISHNA V, KANADE T, et al. Convolutional pose machines [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 4724-4732.
- [16] 夏浩宇, 索双富, 王洋, 等. 基于 Keypoint RCNN 改进

模型的物体抓取检测算法[J]. 仪器仪表学报, 2021, 42(4): 236-246.

- [17] 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.
- [18] QIAN X, FU Y, XIANG T, et al. Pose-normalized image generation for person re-identification [C].

Proceedings of the European Conference on Computer Vision(ECCV), 2018: 650-667.

作者简介

朱肖磊, 硕士研究生, 主要研究方向为计算机视觉、图像处理等。

E-mail: zhuxl96@163.com

吴训成(通信作者), 博士, 教授, 主要研究方向为智能网联汽车技术等。

E-mail: 18861120776@163.com