

用于车辆重识别的基于细节感知的 判别特征学习模型*

邱铭凯, 李熙莹

中山大学智能工程学院 / 广东省智能交通系统重点实验室 / 视频图像智能分析与应用技术公安部
重点实验室, 广东 广州 510006

摘要: 在多摄像头下拍摄的车辆图片集合中, 对属于目标车辆的图片进行匹配为车辆重识别。如何有效的区分具有相似外观的不同车辆的图片, 是车辆重识别的一大挑战。考虑到不同车辆之间的差异集中于车窗等区域的细节, 文章提出了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的提取特征, 不同车辆联合特征之间的欧式距离作为相似度衡量。在算法实验中, 所提出的算法在公开数据集 VehicleID 与 VeRi 上都取得领先于现有车辆重识别算法的结果, 验证了算法的有效性。

关键词: 车辆重识别; 神经网络; 局部特征提取; 判别特征学习

中图分类号: TP183 **文献标志码:** A **文章编号:** 0529-6579 (2021) 04-0111-10

Detail-aware discriminative feature learning model for vehicle re-identification

QIU Mingkai, LI Xiyang

School of Intelligent Systems Engineering, Sun Yat-sen University / Guangdong Provincial Key Laboratory of Intelligent Transportation System / Key Laboratory of Video and Image Intelligent Analysis and Application Technology, Ministry of Public Security, Guangzhou 510006, China

Abstract: Vehicle re-identification (Re-ID) aims to identify a target vehicle from multiple non-overlapping cameras. Vehicle Re-ID is a challenging work because it's hard to distinguish vehicles of the same model with similar appearance. Since the differences between these vehicles are concentrated in some small local regions, a detail-aware discriminative feature learning model is proposed in this paper, based on the assumption that features of network's middle layer is helpful in extracting discriminative feature representation of local regions. In the proposed model, a guided vehicle local feature extraction process is designed, and the final feature representation of vehicle consist of the extracted local feature and the global feature extracted by the backbone network. Extensive experiments over benchmark datasets VehicleID and VeRi have shown that the proposed methods could achieve superior performance than state-of-the-art methods.

Key words: vehicle re-identification; neural network; local feature extraction; discriminative feature learning

* 收稿日期: 2020-03-16

录用日期: 2020-05-07

网络首发日期: 2020-09-15

基金项目: 国家重点研发计划 (2018YFB1601100, 2018YFB1601101); 国家自然科学基金 (U1611461)

作者简介: 邱铭凯 (1996年生), 男; 研究方向: 交通视频图像分析; E-mail: qiunk@mail2.sysu.edu.cn

通信作者: 李熙莹 (1972年生), 女; 研究方向: 交通视频图像分析; E-mail: stslxy@mail.sysu.edu.cn

车辆重识别的一个重要应用就是重构车辆轨迹。通过车辆重识别,可以实现连续时间内同一车辆在不同摄像头下的关联,实现基于道路监控视频的轨迹重构。利用车辆重识别重构的车辆轨迹,结合车辆驾驶员信息,可以实现用户乘客出行规律分析,进一步实现乘客出行预测及出行引导。另外,在公安刑侦方面,车辆重识别可以用于车辆跟踪、被盗车辆检索以及车辆套牌验证等场景。因此,车辆重识别在智能交通系统中具有重要的研究意义。如何在视频大数据中基于计算机视觉技术,自动实现车辆重识别,是智能交通系统发展中一个亟需解决的问题。

车辆重识别在实际研究中的一大挑战就是对外观相似的不同车辆进行区分。在实际场景中,道路上存在多辆车属于同一款式的情况,而在同一摄像头下拍摄到的不同车辆,由于款式和颜色相同,外观相似度高,难以区分。由于相似外观的不同车辆之间的差异主要集中于车窗装饰物等局部区域,提升算法对于局部细节差异的捕获能力成为研究的关键。

现有的车辆重识别研究主要利用深度学习技术,通过设计相应的训练任务及改进网络结构来指导网络提取到具有辨识度的特征。为了加快网络训练,使得网络训练过程更加稳定,Liu等^[1]设计了一个 Couple Cluster Loss 以改进三元组损失;考虑类内差异,Bai等^[2]提出了一个 Group-Sensitive-Triplet Embedding 方法,将同一车辆的不同图片按照属性差异划分为多个组,同一组内的图片被认为是具有相同的属性,然后在损失函数中考虑组间的差异,以及不同车辆之间的差异,达到特征差异化的目的;考虑多视角下车辆的重识别,Chu等^[3]和Zhou等^[4]研究利用对抗生成网络来通过单视角图片生成车辆的多视角融合特征;考虑车辆关键部位特征提取,Khorramshahi等^[5]和Guo等^[6]通过车辆关键部位定位以及关键点检测,利用注意力机制来提取具有辨识度的特征。这些方法的实验结果表明,添加限制性更强的损失函数以及有选择性的引入额外的特征可以提高算法的鲁棒性以及增强算法的识别效果。

卷积神经网络提取的特征可以根据深度被分为三个部分:浅层、中层及深层特征。其中,浅层特征对浅层视觉信息进行表达,包括边缘、角以及圆等;中层特征对图片物体的各个部分进行表达;深层特征得到关于图片整体的语义信息表

达^[7-8]。因此,利用网络的中层特征将会有助于网络对于车辆局部区域的细节特征的提取,提高网络对于车辆细节差异的辨识能力。

为了提升算法对于局部细节差异的捕获能力,本文提出了一个基于细节感知的判别特征学习网络模型。模型以 InceptionV3^[9] 网络为基础结构,设计了一个指导式的局部特征提取流程。在网络结构中,将基础网络划分为浅层、中层和深层三层,对中层网络设计基于注意力机制的局部特征提取模块;结合局部限制及混合采样策略指导局部特征提取。最后,将网络提取的局部特征及全局特征组合作为车辆的联合特征,计算不同车辆之间特征的欧式距离作为相似度衡量。模型在公开数据集 VehicleID^[1] 及 VeRi^[10] 上都可以取得领先于现有算法的水平。

1 基于细节感知的判别特征学习模型

本文设计的基于细节感知的判别特征学习模型,如图 1 所示。在网络结构设计中,以 InceptionV3 作为基础骨干网络,引入局部特征提取模块,以局部特征及基础骨干网络提取到的全局特征作为车辆的联合提取特征;在模型训练中,设计多任务联合训练,以不同的损失函数指导特征的生成;为了优化模型的训练,在图片输入部分,模型设计了有效的混合采样策略。对于局部特征的生成提取,算法以局部特征提取模块为基础,结合局部限制及混合采样策略,组成了一个完整的指导式局部特征提取流程。

1.1 全局特征提取

判别特征学习模型中的基础骨干网络负责车辆全局特征的提取。在车辆重识别算法中,普遍使用 VGG16^[11], Inception^[9] 以及 ResNet^[12] 等经典网络作为基础骨干网络。其中, InceptionV3 网络的基本结构如表 1 所示,因为 Inception 模块的存在, InceptionV3 可以更好的适用于多尺度特征的提取,所以在本研究中使用 InceptionV3 作为基础骨干网络。

Inception 模块如图 2 所示。Inception 模块设计为先进行多尺寸的卷积、池化并行运算后,将输出并联。一方面能够增加网络的宽度,另一方面增加了网络对于尺度的适应性,从而使得 Inception 网络可以更好地适用于多尺度特征提取。

这里使用 Inception 网络会更有利于对车辆局

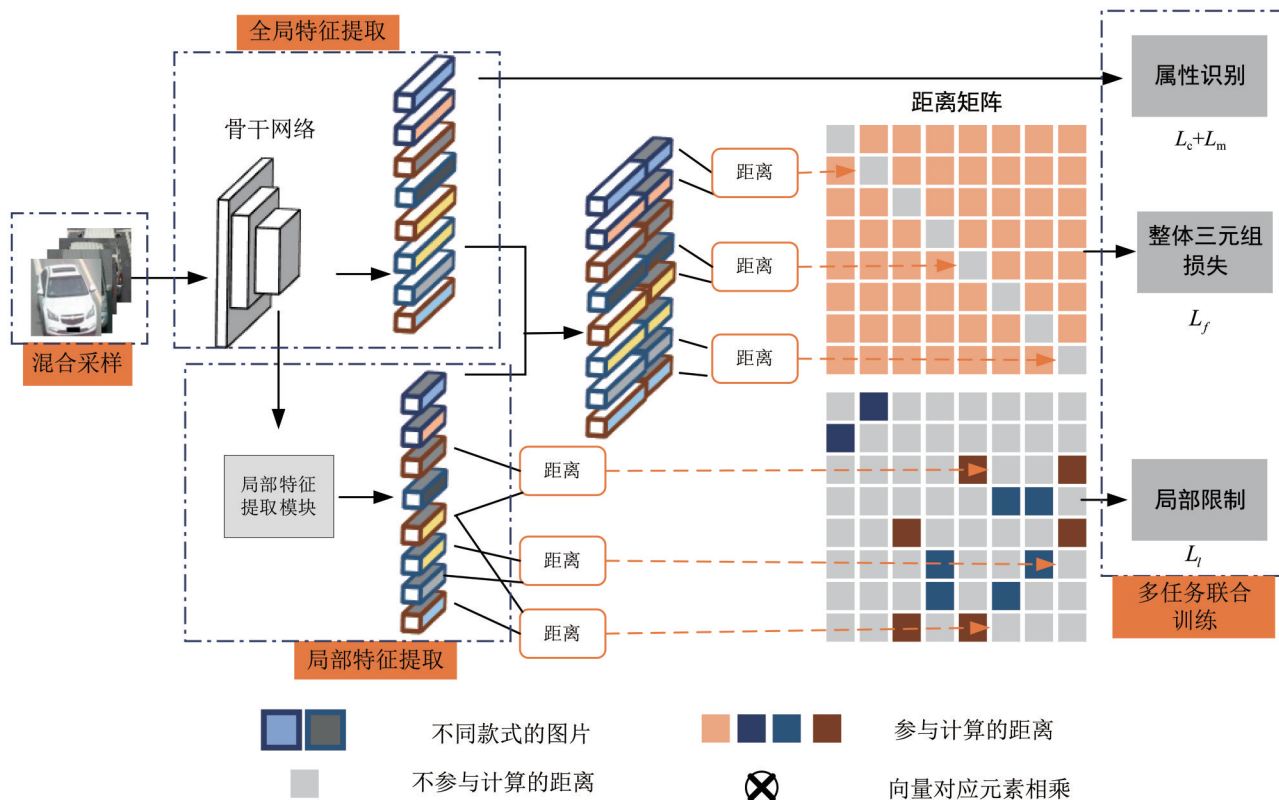


图 1 基于细节感知的判别特征学习模型

Fig. 1 Detail-aware discriminative feature learning model

表 1 InceptionV3 网络结构表

Table 1 Network architecture of InceptionV3

名称	类型	核尺寸	步长	输出尺寸
Conv1	卷积	3×3	2	149×149×32
Conv2	卷积	3×3	1	147×147×32
Conv3	卷积	3×3	1	147×147×64
Pool1	池化	3×3	2	73×73×64
Conv4	卷积	1×1	1	73×73×80
Conv5	卷积	3×3	1	71×71×192
Pool3	池化	3×3	2	35×35×192
Mixed5	Inception	3 × Inception 模块	—	35×35×288
Mixed6	Inception	5 × Inception 模块	—	17×17×768
Mixed7	Inception	2 × Inception 模块	—	8×8×1 280
Pool2	池化	8×8	—	1×1×2 048

部区域特征的捕捉。对输入图片, 经过 InceptionV3 网络提取, 可以得到大小为 1×2 048 的表示车辆全局特征的向量。

1.2 局部特征提取

1.2.1 局部特征选择 车辆局部特征的提取分为网络划分、中层网络特征提取两步。为了对基础网络结构 InceptionV3 进行划分, 在网络划分阶段, 参考文献 [13], 首先按照图 3 所示的基础算法训练, 训练完成后, 对 InceptionV3 各层网络学习到的特征进行了可视化。

网络每一层的部分特征图的可视化结果如图 4 所示。可以看到, 对于 InceptionV3 网络, 给定输入图片, Conv1 到 Conv5 这 5 层学习到的是颜色、边缘、线段等特征; Mixed5 及 Mixed6 学习到的大部分是车辆的局部区域, 包括车辆栅格、车顶、车窗、车灯等。Mixed7 学习到的则是偏全局性的特征。因此, 将 Mixed5 以及 Mixed6 两个 Inception 模块划分为中层网络。

基于前面的假设, 网络的中层特征有助于网络对于车辆局部区域特征的提取, 所以中间层 Mixed5 以及 Mixed6 被用于提取车辆的局部特征。

1.2.2 基于注意力机制的局部特征提取模块 在实际场景中区分不同车辆时, 车辆的各个区域具有不同的辨识度。高辨识度区域主要集中于车灯、车辆栅格、车标以及车辆挡风玻璃等区域。因此,

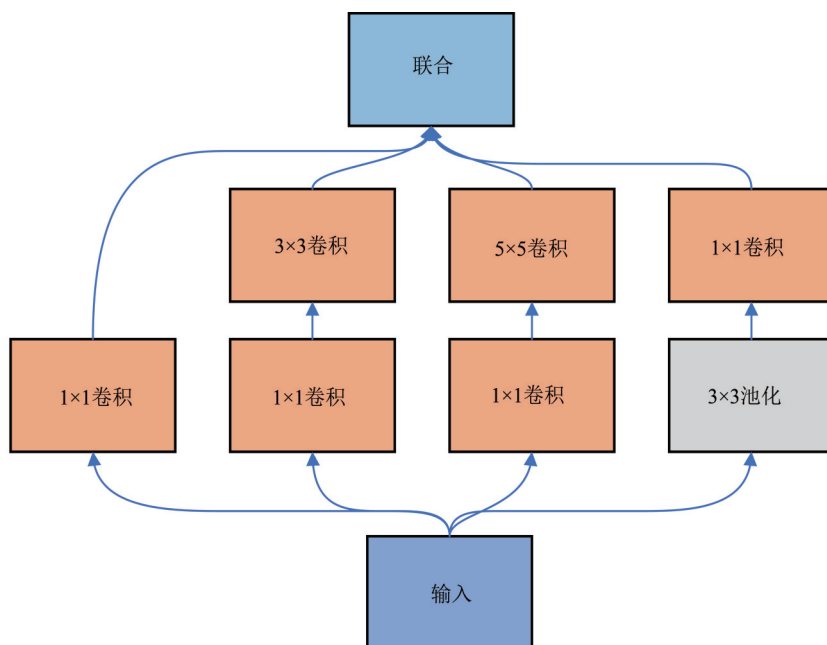


图2 Inception 模块

Fig. 2 Inception Module

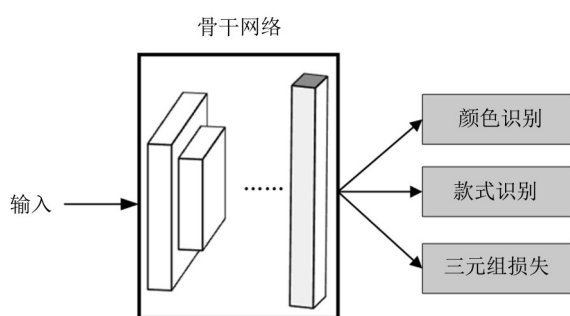


图3 车辆重识别基础算法

Fig. 3 Baseline algorithm for vehicle re-identification

对局部特征进行提取、比对的过程中, 应该考虑不同局部区域特征的辨识度, 增强具有高辨识度的特征, 抑制低辨识度特征的干扰。

基于上面的描述, 在局部特征提取模块的设计中, 我们引入注意力机制, 根据区域的重要性实现网络自学习并自动分配不同的权重, 得到带注意力机制的局部特征提取模块, 模块结构如图5中所示。

对于大小为 $H \times W \times C$ 的中间层特征图, 添加一个卷积核大小为 3×3 、步长为 1×1 的卷积层, 得到一个与原始特征图同样大小的注意力特征图。在这里, 使用 Sigmoid 函数作为激活函数, 从而保证得到的特征图中的数值大小位于 $(0, 1)$ 区间内。得到的注意力特征图与原始特征图相乘后进行全局池化, 即可得到最终的局部特征向量。

1.3 多任务联合训练

输入车辆图片, 模型提取得到两部分特征: 全局特征以及局部特征。针对全局特征、局部特征以及两者组合的联合特征, 分别设计了不同任务及相应的损失函数, 以指导特征的生成。

对于全局特征, 设计了车辆的属性识别任务, 分别对车辆的颜色以及款式进行识别。颜色及款式识别属于多分类问题, 使用 softmax 交叉熵函数作为损失函数:

$$L_c = -\sum_{i=1}^{k_c} y_{c_i} \log(y'_{c_i}), \quad (1)$$

$$L_m = -\sum_{i=1}^{k_m} y_{m_i} \log(y'_{m_i}), \quad (2)$$

其中下标 c 和 m 分别表示颜色及款式。对于颜色识别, k_c 表示车辆颜色种类的数目, 是网络输出的对于车辆属于第 i 种颜色的概率的预测值。 y_{c_i} 为 0 或 1, 表示车辆颜色的真实值, 定义为

$$y_{c_i} = \begin{cases} 1, & \text{输入车辆的实际颜色是第 } i \text{ 种颜色,} \\ 0, & \text{其他,} \end{cases} \quad (3)$$

对于款式识别, k_m , y_{m_i} , y'_{m_i} 的含义与 y_{c_i} , y'_{c_i} 类似。对于车辆的联合特征, 在算法中设计了整体三元组损失函数。

对于车辆的联合特征, 在算法中设计了整体三元组损失函数。从训练集中随机选择一张图片作为 anchor, 记为 x^a ; 再随机选择一张与 x^a 属于同

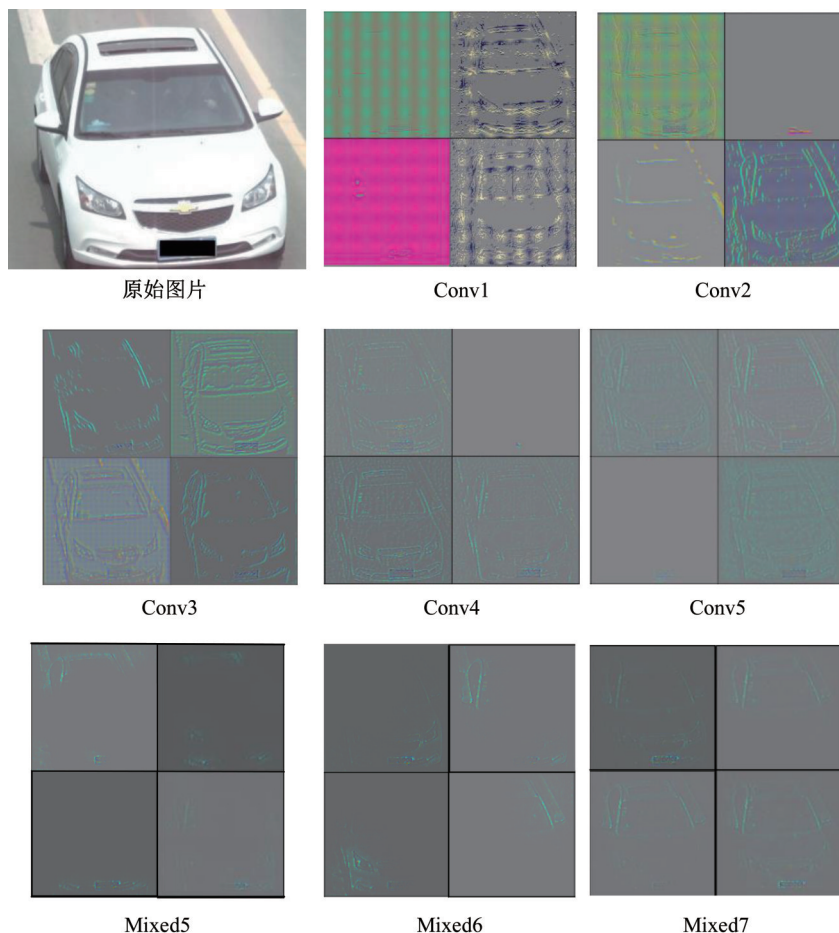


图 4 InceptionV3 网络各层特征可视化结果
Fig. 4 Feature visualization results of InceptionV3

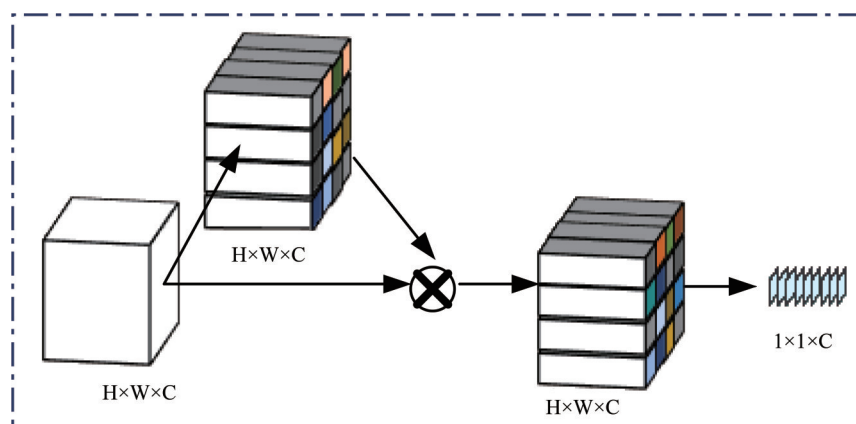


图 5 局部特征提取模块
Fig. 5 Local feature extraction module

一车辆、但是不同的图片作为正样本 (positive), 记为 x^p ; 同时, 选择一张与 x^a 属于不同车辆的图片作为负样本 (negative), 记为 x^n 。这样, 就组成一个训练三元组 $[x^a, x^p, x^n]$, 输入网络后可得到三元组的特征表达 $[G_f(x^a), G_f(x^p), G_f(x^n)]$ 。整体三元组

损失函数为

$$L_f(x^a, x^p, x^n) = \max\left(0, \left\|G_f(x^a) - G_f(x^p)\right\|_2^2 - \left\|G_f(x^a) - G_f(x^n)\right\|_2^2 + \alpha\right), \quad (4)$$

其中 α 是预设值的阈值。

由于相同款式的不同车辆之间, 差异更多的集中于车窗装饰物等细小区域, 所以根据这一现象, 算法设计了局部限制三元组损失以指导局部特征的生成。给定 anchor 图片, 假设图片中车辆的款式为 m , 图片记为 x_m^a , 在训练集中随机选取一张与 x_m^a 属于同一车辆、但不同的图片作为正样本 (positive), 记为 x_m^p 。区别于整体三元组损失函数中, 将不同车辆的图片作为负样本的做法, 在局部限制三元组损失函数中, 对负样本的选取增加了一个限制, 即仅将相同款式不同车辆的图片作为负样本, 记为 x_m^n 。这样组成的局部限制三元组 $[x_m^a, x_m^p, x_m^n]$ 可以保证正样本车辆与负样本车辆之间的差别更集中于局部细节区域, 有利于指导局部特征的生成。将局部限制三元组输入网络, 得到三元组的局部特征表达 $[G_l(x_m^a), G_l(x_m^p), G_l(x_m^n)]$ 。则局部限制三元组损失函数为

$$L_l(x_m^a, x_m^p, x_m^n) = \max\left(0, \left\|G_l(x_m^a) - G_l(x_m^p)\right\|_2^2 - \left\|G_l(x_m^a) - G_l(x_m^n)\right\|_2^2 + \beta\right). \quad (5)$$

1.4 混合采样策略

根据局部限制中对于三元组样本构造的要求, 在网络输入中, 不同于一般算法的直接随机抽取 M 张图片作为输入, 本算法中设计了混合采样策略, 如图 6 所示。

图 6 中, 车辆优先策略属于一种先随机选择车辆, 再随机选择图片的二次随机过程。车辆优先策略以车辆为基本单位, 先对车辆进行随机采样, 选择 K_1 辆车后, 分别对每辆车的所有图片进行随机选取, 选取 P_1 张图片, 则一个批次训练的总图片数目为 $K_1 \times P_1$ 。车辆优先策略的优点是可以保证训练集中每一张图片都可以找到属于相同车辆的其他图片, 从而能够组成有效的三元组, 有效地加速网络训练, 保证网络收敛。

图 6 的款式优先策略在车辆优先的基础上, 加入了对款式随机选择的过程, 属于一种三次随机过程。与车辆优先策略不同, 款式优先策略以款式为基本单位, 先随机选取 C 个款式, 对每一个款式随机选取 K_2 辆车, 对每一辆车再随机选取 P_2 张图片, 则一批次的训练图片总数为 $C \times K_2 \times P_2$ 。款式优先策略的意义在于: 因设计了一个局部限制三元组损失函数, 对于给定的 anchor x_m^a , 只有与 anchor 属于相同款式不同车辆的图片才被确定为负样本。如果仅采用车辆优先策略, 则在同一批次训练中, 存在两辆及以上车辆属于同一款式的概

率极低, 即使用车辆优先策略不利于局部限制三元组损失函数的收敛, 而款式优先策略则可以很好的解决这一问题。

款式优先策略可以满足需求, 但是在本算法中, 使用的是车辆优先策略结合款式优先策略的混合采样策略, 使用混合采样策略的意义会在实验部分进行阐述。

2 实验验证

2.1 实验设置

算法使用在 ImageNet^[14] 数据集上预训练好的模型进行参数的初始化。对于网络的基础设置, 使用 Adam 优化器作为参数优化器, 使用默认的参数设置 ($\epsilon = 10^{-3}, \beta_1 = 0.9, \beta_2 = 0.999$)。初始学习率设置为 0.000 1, 批大小 (batch-size) 设置为 128, 学习率衰减设置为每 5 个 epoch 乘以系数 0.9, 训练 100 个 epoch。对于算法中的各个参数, 在整体三元组损失函数以及局部限制三元组损失函数中, 设置 $\alpha = \beta = 1.0$; 在混合采样策略中, 设置 $C = K_2 = P_1 = P_2 = 4, K_1 = 16$, 即款式优先策略与车辆优先策略采样数目相同, 各采样 64 张图片。

2.2 数据集及评价标准

在测试中, 给定查询集, 对于查询集中每一张图片, 在候选集中进行匹配。为了评价查询集在候选集中的匹配效果, 本研究使用 3 个指标作为评价标准: rank1 准确率, rank5 准确率以及 mAP (mean average precision)。为了对算法的有效性进行验证, 本研究使用车辆重识别公开数据集 VehicleID^[1] 与 VeRi^[10] 作为实验数据集。

VehicleID 数据集采集自中国城市道路上监控摄像头于白天拍摄到的视频。总共包含 26 267 辆车共 221 763 张图片, 每张图片根据车牌号都标注有相应的 ID, 其中共 78 957 张图片标注有款式及颜色信息, 共 228 个款式, 7 种颜色。

VehicleID 数据集中提供了三种不同大小的测试集。其中, Test800 测试集包含 800 张查询图片, 6 532 张候选图片; Test1600 测试集包含 1 600 张查询图片, 11 395 张候选图片; Test2400 测试集包含 2 400 张查询图片, 17 638 张候选图片。后续试验会在三个不同测试集上进行, 根据 VehicleID 数据集的设置, 使用 rank1 及 rank5 准确率作为算法在 VehicleID 数据集上的性能评价指标。

VeRi 数据集采集自中国福建省福州市永泰县

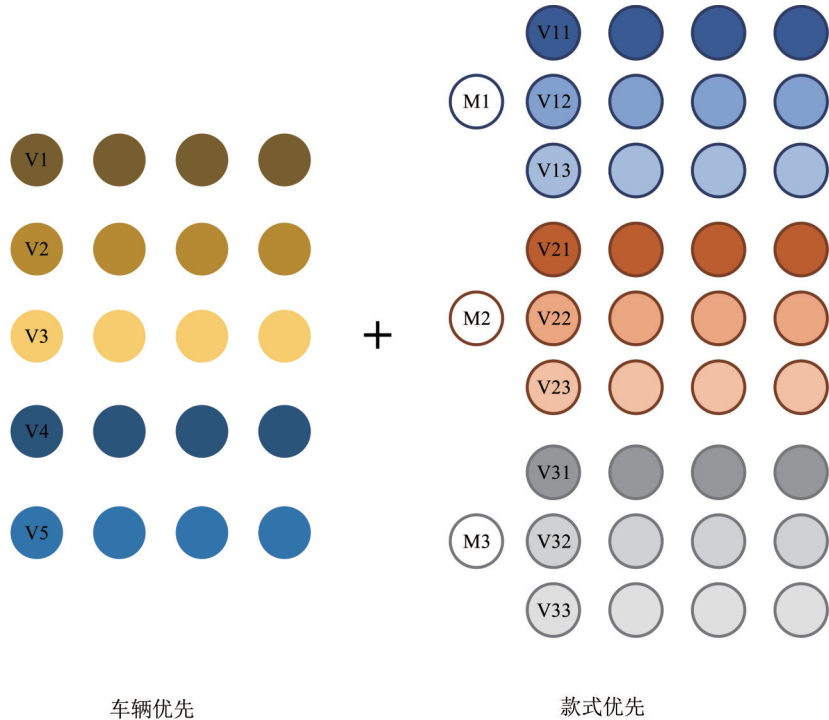


图 6 混合采样策略

Fig. 6 Mixed sampling strategy

一条道路上的道路监控摄像头拍摄到的视频, 总共包含 776 辆车共 49 357 张图片。VeRi 数据集中所有车辆都标注有颜色及车型信息, 共 9 种车型, 10 种颜色。数据集中还提供了每一图片拍摄的位置信息及时间信息。

VeRi 数据集中, 其中 500 辆车共 37 778 张图片作为训练集, 剩余 200 辆车共 11 579 张图片作为测试集。测试集中, 划分 1 678 张图片作为查询集合, 剩余 9 901 张图片作为候选集合。根据 VeRi 数据集的设置, 使用 rank1 准确率, rank5 准确率以及 mAP 作为算法在 VeRi 数据集上的性能评价指标。

2.3 消融分析

本文提出的算法包含了局部特征提取、局部限制、采样策略等各个组件。为了对各个部分选择最有效的设计方式, 以及明确各个部分对于算法效果的影响, 我们对算法进行了消融分析, 从而得到各个部分对于基础算法的提升效果。本部分实验以 VehicleID 数据集的 Test800, Test1600 和 Test2400 测试集进行测试。

首先, 在局部特征提取模块设计中, 考虑中间层以及池化方式对于算法的影响, 我们将 InceptionV3 网络的 Mixed5 以及 Mixed6 两个 Inception 模块划分为中层网络。另外, 利用中层网络特征图进行全局池化提取得到局部特征。在实验中, 我

们针对不同模块分别进行全局均值池化以及全局最大值池化的对比。实验结果如表 2 所示, 其中 Max, Ave 分别表示采用全局最大值池化及全局均值池化。

表 2 不同中间层及池化方式组合的实验结果

Table 2 Results of different combination of middle layers and pooling methods %

模块组	Test800		Test1600		Test2400	
	rank1	rank5	rank1	rank5	rank1	rank5
Baseline	88.1	93.7	85.4	92.5	82.3	91.2
Max-5	89.3	94.3	87.6	93.4	86.5	92.3
Ave-5	88.9	94.5	85.8	91.9	83.0	91.1
Max-6	89.8	94.6	87.5	93.2	86.0	92.3
Ave-6	87.8	93.0	86.2	92.1	83.6	91.5

由表 2 可以看出, 对于两个模块组, 使用全局最大值池化进行特征提取在三个测试集上的表现都要优于使用全局均值池化。在使用全局最大值池化时, 使用 Mixed6 模块组在 Test800 上的 rank1 准确率和 rank5 准确率要优于 Mixed5 模块组, 但是在 Test1600 和 Test2400 两个较大的测试集上的 rank1 和 rank5 准确率则都要劣于 Mixed5 模块组。考虑到 Mixed6 模块组提取到的局部特征维度为

768, 远远大于 Mixed5 模块组提取到的 288 维特征向量, 结合表格 3 的数据, 在最终算法中将局部特征提取模块设计为利用全局最大值池化方式, 对 Mixed5 模块组提取局部特征向量。

与 Baseline 相比, 加入局部特征向量, 在不同的测试集上的 rank1 和 rank5 的准确率都有所提升。在 Test800 测试集上, Max-5 的 rank1 和 rank5 准确率分别提高 1.2% 和 0.6%, 而在 Test1600 测试集上, Max-5 的 rank1 和 rank5 准确率分别提高 2.2% 和 0.9%; 在 Test2400 测试集上, Max-5 的 rank1 和 rank5 准确率分别提高 4.2% 和 1.1%。可以看到, 随着测试集的数量增大, 加入局部特征向量对于准确率的提升越大。这是因为测试集的数量越大, 则意味着存在外观相似的车辆数量越多, 算法更大概率会因混淆而导致误判。而引入局部特征向量, 则能够增强算法对于局部区域特征的辨识能力, 从而能够更好的对易混车辆进行区分。所以, 测试集数量越大, 引入局部特征算法相比于原始算法的优势则越大。

其次, 考虑引入注意力机制以及局部限制对于算法的影响。实验结果如表 3 所示, 其中 att 表示注意力模块, LC 表示局部限制, 最后的 Max-5+att+LC 则为最终的算法。

表 3 组件消融分析结果

模块组	Test800		Test1600		Test2400	
	rank1	rank5	rank1	rank5	rank1	rank5
	Baseline	88.1	93.7	85.4	92.5	82.3
Max-5	89.3	94.3	87.6	93.4	86.5	92.3
Max-5+att	89.9	94.6	87.6	92.9	86.2	93.1
Max-5+LC	90.4	94.7	88.4	93.5	87.3	92.9
Max-5+att+LC	91.1	94.7	88.9	93.5	88.5	93.4

由表 3 可以看出, 相比于原始的加入局部特征向量, 单独引入注意力机制, 在 Test800 测试集的 rank1 准确率提升 0.9%; 而在 Test1600 和 Test2400 两个测试集上的 rank1 和 rank5 准确率则近似没有变化。另外, 单独加入局部限制, 在三个测试集上的准确率都要优于单独引入注意力机制。相比于原始的仅加入局部特征向量, 在 Test800 测试集的 rank1 准确率提升 1.1%, 在 Test1600 和 Test2400 上的 rank1 准确率都提升了 0.8%。而, 同时引入注意力机制及局部限制, 相比于原始的加入局部特征向量, 在 Test800 测试集上的 rank1 和 rank5 准

准确率分别提高 1.8% 和 0.4%, 在 Test1600 上的 rank1 准确率提高 1.3%, 在 Test2400 上的 rank1 和 rank5 准确率则分别提升 2% 和 1.1%。可以看到, 引入注意力机制和局部限制对算法在不同大小测试集上的 rank1 准确率影响最大, 而对于 rank5 准确率的影响则较小。

在实验中, 我们使用了款式优先结合车辆优先的混合采样策略。对混合采样策略和款式优先策略进行比较, 实验结果如表 4 所示。其中 Mix 表示混合采样策略, MF 表示款式优先策略。

表 4 不同采样策略的结果

采样策略	Test800		Test1600		Test2400	
	rank1	rank5	rank1	rank5	rank1	rank5
	Max-5+att+LC+Mix	91.1	94.7	88.9	93.1	88.5
Max-5+att+LC+MF	88.9	93.9	86.2	91.9	85.3	91.2

可以看到, 仅使用款式优先策略, 相比于使用混合采样策略, 在 Test800 测试集的 rank1 和 rank5 准确率分别下降了 2.2% 和 0.8%, 在 Test1600 上的 rank1 和 rank5 准确率分别下降了 2.7% 和 1.2%, 在 Test2400 上的 rank1 和 rank5 准确率分别下降了 3.2% 和 2.2%。出现准确率下降的原因为: 如果单独采用款式优先策略, 则对于每一车辆, 在样本集中都存在与其相同款式但是不同 ID 的车辆图片。由于在整体三元组损失的计算中, 对于每一个 anchor, 都是选择与其距离最近的负样本, 即外观最相似的不同车辆的图片。则由于款式相同的车辆之间外观最接近, 对于每一个 anchor, 选择的都是与其相同款式但是不同 ID 的车辆图片进行训练, 这也就意味着网络在训练的过程中没办法学习到不同款式车辆之间的差异。而采用款式优先结合车辆优先的混合策略, 则一方面可以保证在样本集中存在可以满足局部限制三元组的样本对, 另一方面可以保证网络可以学习到相同款式以及不同款式车辆之间的差异。

2.4 结果对比

将本研究的算法与现有的其他车辆重识别最新算法在 VehicleID 及 VeRi 数据集上进行对比。我们选择的对比算法包括 DRDL^[1]、PROVID^[10]、VAMI^[3]、JFSDL^[15]、ABLN^[4]、DHMVI^[16]、TAMR^[6]、AAVER^[5]、EALN^[17]。算法对比结果如

表5-6所示。

表5 算法在 VehicleID 数据集上的对比结果

算法	%					
	Test800		Test1600		Test2400	
	rank1	rank5	rank1	rank5	rank1	rank5
DRDL	49.0	73.5	42.8	67.0	38.2	61.6
ABLN	52.6	80.5	—	—	—	—
DHMMVI	48.6	81.0	—	—	—	—
JFSDL	54.8	85.3	48.3	78.8	41.3	70.6
VAMI	63.1	83.3	52.9	75.1	47.3	70.3
TAMR	66.0	79.7	62.9	76.8	59.7	73.9
AAVER	74.7	93.8	68.6	90.0	63.5	85.6
EALN	75.1	88.1	71.8	83.9	69.3	81.4
DDFL (Ours)	91.1	94.7	88.9	93.1	88.5	93.4

从表5-6可以看出, 在与其他算法的对比中, 本研究所提出的算法模型在两个数据集中都可以取得最好的效果。在 VehicleID 数据集中, 相比于效果排在第二位的算法, 本研究所提出的算法在 Test800 测试集的 rank1 和 rank5 的准确率、分别提高 16% 和 6.6%, 在 Test1600 和 Test2400 两个测试集上的 rank1 准确率分别提高 17.1% 和 19.2%。在 VeRi 数据集中, 相比于效果排在第二位的算法, 本研究所提出的算法的 rank5 准确率和 mAP 指标分别提高了 1.6% 和 7.1%。

参考文献:

- [1] LIU H Y, TIAN Y H, WANG Y W, et al. Deep relative distance learning: Tell the difference between similar vehicles [C] // In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016:2167-2175.
- [2] BAI Y, LOU Y H, GAO F, et al. Group-sensitive triplet embedding for vehicle reidentification [J]. IEEE Transactions on Multimedia, 2018, 20(9):2385-2399.
- [3] CHU R H, SUN Y F, LI Y D, et al. Vehicle re-identification with viewpoint-aware metric learning [C] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019:8281-8290.
- [4] ZHOU Y, SHAO L. Vehicle re-identification by adversarial bi-directional LSTM network [C] // In 2018 IEEE

表6 算法在 VeRi 数据集上的对比结果

Table 6 Comparisons of our method with state-of-the-arts on VeRi dataset

算法	rank1	rank5	mAP
ABLN	60.5	77.3	24.9
JFSDL	82.9	91.6	53.5
VAMI	77.0	90.9	50.1
DHMMVI	60.8	78.6	25.1
PROVID	81.6	95.1	53.4
AAVER	88.7	94.1	58.5
EALN	84.4	94.1	57.4
DDFL (Ours)	88.7	95.7	64.5

3 结论

本文设计了一个基于细节感知的判别特征学习模型。考虑深度网络可以根据表述的特征划分为浅层、中层和深层三类, 其中中层特征对图片的局部区域特征进行表达, 因此网络的中层特征会有利增强网络对于局部区域细节的辨识力。基于这一假设, 模型以 InceptionV3 网络为基础结构, 设计了一个基于注意力机制的局部特征提取模块, 结合局部限制及混合采样策略, 指导局部特征提取生成。以提取的局部特征及基础网络生成的全局特征结合作为车辆的联合特征, 通过计算车辆特征之间的绝对距离作为相似度度量。在公开数据集 VehicleID 及 VeRi 上的实验证明, 所提的算法可以取得优于现有的车辆重识别算法的效果。

- Winter Conference on Applications of Computer Vision (WACV), 2018:653-662.
- [5] KHORRAMSHAHI P, KUMAR A, et al. A dual-path model with adaptive attention for vehicle re-identification [C] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019:6131-6140.
- [6] GUO H Y, ZHU K, TANG M, et al. Two-level attention network with multi-grain ranking loss for vehicle re-identification [J]. IEEE Transactions on Image Processing, 2019, 28(9): 4328-4338.
- [7] LIN G S, MILAN A, SHEN C H, et al. ReID refinenet: multi-path refinement networks for high-resolution semantic segmentation [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),

- 2017:5168–5177.
- [8] YOSINSKI J, CLUNE J, NGUYEN A, et al. Understanding neural networks through deep visualization [J/OL]. [2015–06–22]. <https://arxiv.org/abs/1506.06579v1>.
- [9] SZEGEDY C, LIU W, JIA Y Q, et al. Going deeper with convolutions[C]// 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015: 1–9.
- [10] LIU X C, LIU W, MEI T, et al. PROVID: Progressive and multimodal vehicle reidentification for large-scale urban surveillance[J]. IEEE Transactions on Multimedia, 2018, 20(3):645–658.
- [11] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J/OL]. [2015–04–10]. <https://arxiv.org/abs/1409.1556>.
- [12] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016:770–778.
- [13] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks [J/OL]. [2013–11–28]. <https://arxiv.org/abs/1311.2901>.
- [14] DENG J, DONG W, SOCHER R, et al. Imagenet: a large-scale hierarchical image database [C]// IEEE Computer Vision & Pattern Recognition, 2009: 248–255.
- [15] ZHU J Q, ZENG H Q, DU Y Z, et al. Joint feature and similarity deep learning for vehicle re-identification [J]. IEEE Access, 2018(6): 43724–43731.
- [16] 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.
- [17] LOU Y H, BAI Y, LIU J, et al. Embedding adversarial learning for vehicle re-identification [J]. IEEE Transactions on Image Processing, 2019, 28(8): 3794–3807.

(责任编辑 王海蓉)