

基于局部特征和焦点融合的车辆重识别算法^{*}李浩^{1,2} 杨超^{1,2} 黄友新^{1,2} 陈嘉哲^{1,3} 詹瑞典^{1,2} 鲍鸿²

(1. 广东工业大学 先进制造学院 揭阳 515200; 2. 广东工业大学 自动化学院 广州 510006;

3. 广东工业大学 实验教学部 广州 510006)

摘要: 由于城市监控中存在大量相似的车辆,造成了车辆重识别匹配率低。车头、车窗、车顶等局部特征是相似车辆细微差异性的所在。根据车辆检测算法卷积特征热力图注意力分布特性,提出了针对车辆局部特征区域检测的MCRF-SSD算法,并与GMM-EM聚类算法相结合,检测性能在公开的数据集上均优于目前主流算法。同时为了增大类间距离、缩小类内距离将Arcface损失函数引入到了特征提取阶段。为了提高车辆重识别匹配性能,在全局特征与局部特征融合阶段提出了一种保留特征图空间分布的焦点融合(FFS)方法,并引入了一个可学习参数,提高了特征融合效率。实验结果表明,所提出的算法在公开的VehicleID和VeRi数据集中性能表现优于目前性能最优的方案。

关键词: 车辆重识别;局部特征;聚类;特征提取;全局特征;特征融合;焦点融合;可学习参数

中图分类号: TP391.41 **文献标识码:** A **国家标准学科分类代码:** 510.1050

Vehicle re-identification based on local feature and focus fusion

Li Hao^{1,2} Yang Chao^{1,2} Huang Youxin^{1,2} Chen Jiazhe^{1,3} Zhan Ruidian^{1,2} Bao Hong²

(1. School of Advanced Manufacturing, Guangdong University of Technology, Jieyang 515200, China;

2. School of Automation, Guangdong University of Technology, Guangzhou 510006, China;

3. Experimental Teaching Department, Guangdong University of Technology, Guangzhou 510006, China)

Abstract: There are many similar vehicles in city monitoring, which brings great challenges to vehicle re-identification. Local features such as front, window and roof are the subtle differences of similar vehicles. According to the attention characteristics of the thermal map of the vehicle detection algorithm, a MCRF-SSD algorithm is proposed to detect the local feature area of the vehicle, and combines it with GMM-EM clustering algorithm. The detection performance is better than the current mainstream algorithm on the open data set. At the same time, in order to increase the inter-instance and reduce the intra-instance, the Arcface loss function is introduced into the feature extraction stage. In order to improve the performance of vehicle re recognition, in the stage of global feature and local feature fusion, a focus fusion structure (FFS) method is proposed, which can preserve the spatial distribution of feature graph, and a learnable parameter is introduced to improve the efficiency of feature fusion. Experimental results show that the performance of the proposed algorithm is better than that of the current best performance scheme in public VehicleID and VeRi datasets.

Keywords: vehicle re-identification; local features; clustering; feature extraction; global features; feature fusion; focus fusion structure; learnable parameter

0 引言

车辆重识别是识别出在城市监控中的不同摄像头所拍摄的同一辆车辆,其在智慧城市的建设和应用中起到了至关重要的作用。车辆重识别可以应用在车辆轨迹、出行规律分析、套牌车、无牌车、嫌疑车辆和被盗抢车辆的检索等。

近几年深度学习在物体检测、特征提取领域取得的巨大成功。现今的车辆重识别的研究主要集中在基于深度卷积神经网络提取出鲁棒性强的特征^[1]。由于现实场景中存在着大量同品牌同车型的相似车辆,车辆某些独特的局部特征是完成车辆重识别的关键,深度卷积神经网络在此方面的研究主要集中在利用注意力网络自动的获取局部特

收稿日期:2021-07-03

^{*} 基金项目:广东省重点领域研发计划项目(2019B010140002)资助

征^[2]。或者利用检测^[3-4]算法、关键点算法^[5]定位局部关键区域。但是如何获取车辆重识别的关键且有效的局部特征以及如何作用于重识别任务仍然是一个难题。所以本研究提出了基于局部特征检测和特征融合的车辆重识别算法,利用基于深度学习的区域检测算法定位车辆局部区域,在全局和局部特征融合阶段出了一种新型、高效率、抗干扰的焦点融合方法。

1 车辆重识别算法

本研究设计的车辆重识别模型如图1所示。整个车辆重识别系统分为局部特征检测、局部分支簇、特征融合与相似度计算3个部分。在局部特征检测部分以优化后的单点多盒(single shot multibox detector, SSD)^[6,7]检测网络为基础,将现实场景中车辆定位出车辆的车标区域、进气格栅区域、车灯区域、车头区域、车窗区域、车顶局部区域。并将上述检测到的高分辨局部特征区域经过局部分支簇提取到能够表达不同车辆之间细微差异的局部特征图,然后通过焦点融合方法计算得到车辆重识别特征向量。最后与模板库中的车辆特征向量计算欧氏距离得到识别结果。

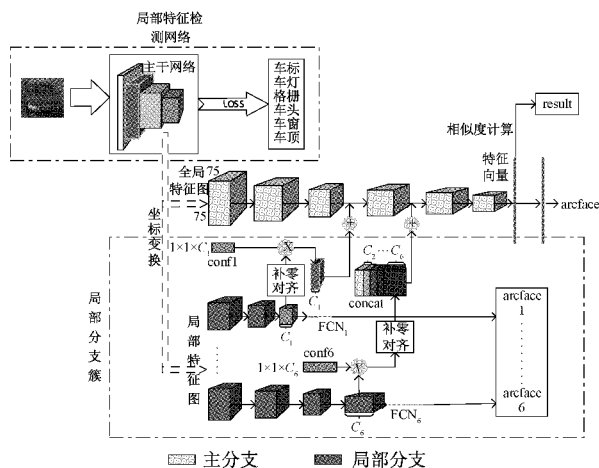


图1 车辆重识别系统

1.1 局部特征

车辆具有强分辨的特性,车辆的某些局部特征体现了车辆独特的特征,这对于车辆重识别中相似车辆的识别具有非常大的影响。车辆的车头具有非常多的高分辨的独特特征,如图2第2行所示不同厂家的车标区域具有很强的区分度,将车标区域加入到车辆重识别中可以极大地减小不同品牌之间的误识别问题。此外不同品牌之间(如图2第3行所示)以及同一品牌(如图3所示)之间不同车型、年款的进气格栅和车灯设计的特征差别较大,将此区域作为局部特征既可以减少不同品牌之间的误识别,同时又能避免同一品牌、同一型号不同年款之间的误识别。上述局部区域使得车辆重识别网络能够有效地识别不同品牌以及同一品牌不同年款的车辆之间的细微的特征差别。但是仅仅采用车标区域、进气格栅和车灯轮廓的局部特征会造成对

于车头很相似的同一个品牌不同车型以及同一个型号不同车辆的误识别。如图4所示对于同一品牌、同一车款的车辆有无天窗、行李架、车窗的年检标、挂件、摆件等是区分此类样本的关键,同时城市中的车辆与摄像头总是存在一定的角度,基于此,车辆的左右车窗边缘弧度特征可以很好地区分同一品牌、车头很相似的不同车型(越野车、轿车)。为此本研究将车顶区域加入到车辆重识别系统中。

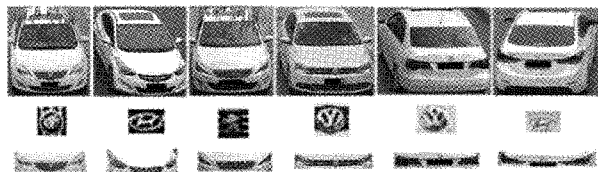


图2 不同品牌车标和车头/车尾区域

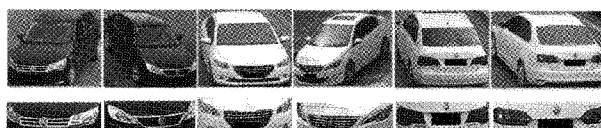


图3 同一品牌车标和车头/车尾区域

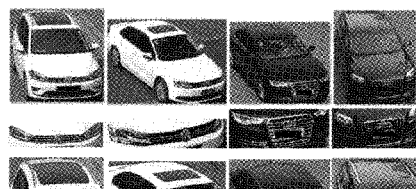


图4 同一品牌不同车型车顶和车头区域

1.2 局部特征检测

为了能够获取车辆的上述(车标、进气格栅、车灯、车头、车窗、车顶)局部区域,并进一步提取对应的局部区域深度特征,本研究利用优化后的单点多盒算法来提取上述局部区域。单点多盒算法根据不同的特征图和感受野大小,在主干网络的不同层级卷积块中设置了不同大小的默认框。底层特征图分辨率更高,感受野更小,被用来检测小目标。但是底层特征语义信息不够丰富,导致了单点多盒算法网络对小目标的检测率很低。高层特征语义丰富,但是分辨率太高,以及感受野太大,不能够用于检测小目标。对小目标的检测既要应用高层特征图丰富的语义信息,又要应用到底层特征图高分辨和低感受野。一般的做法是将特征金字塔结构引入到检测层中,对高层和底层的卷积特征进行融合。考虑到车标、车灯、进气格栅等目标较小,以及车顶、车窗等目标较大,对所有的检测层都引入金字塔,会极大地增加模型的复杂度和计算时间。本研究利用类激活映射^[8]对 conv8-2 卷积层的输出特征做了热力图,如图5所示,经过观察可以发现车头特征丰富,与背景易区分,容易被检测网络捕捉,并且车头区域包含了车标、车灯、进气格栅区域,卷积层的注意力集中在车标、进气格栅和车灯区域。为此只针对车标、车灯、进气格栅对应的检测层引入车

头检测层特征融合构建出多路聚焦、重建的单点多盒检测算法 (multi-channel reconstructed focus based SSD, MCRF-SSD), 以平衡模型复杂度和检测效率。将车头检测层的卷积特征通过上采样的方式与上一层的车标、进气格栅检测层融合。在车标检测层将车头检测层和车灯、进气格栅检测层的特征分双路与车标检测层融合, 从而提高车标的检测率, 网络结构如图 6 所示。

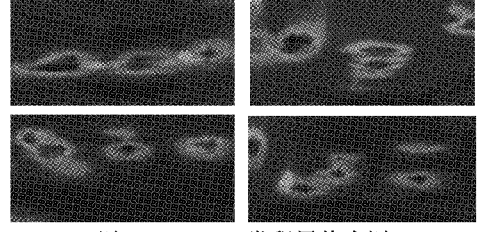


图 5 conv8-2 卷积层热力图

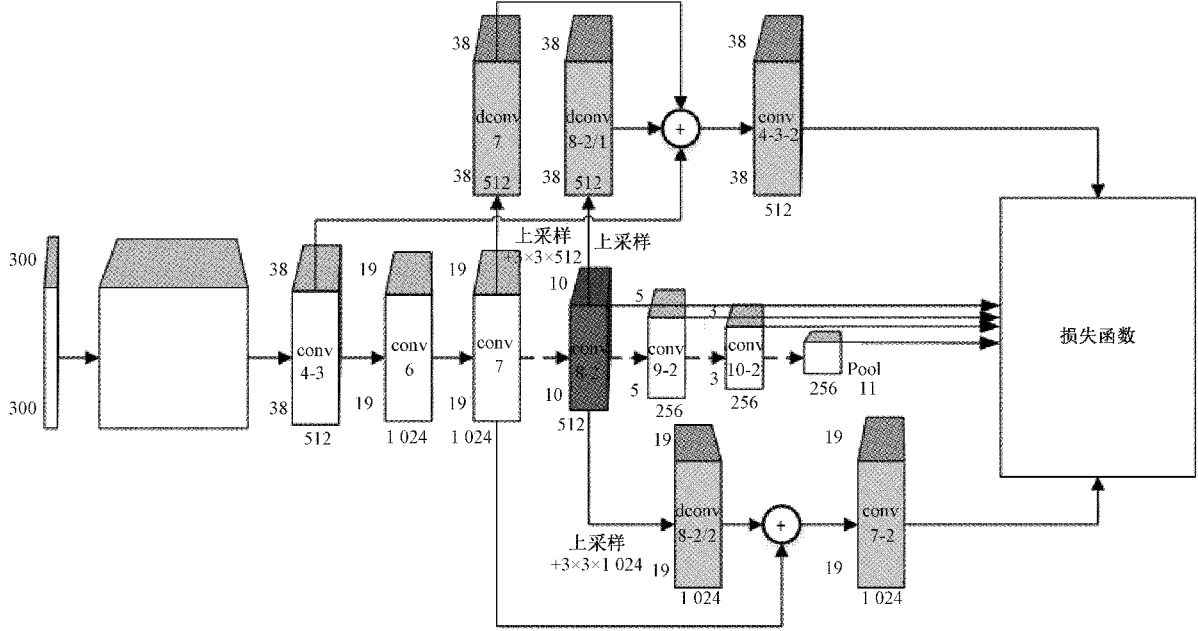


图 6 局部特征检测网络

原始的 SSD 算法默认框是在 VOC 数据集上生成的, 只有获得与训练数据集目标框 (Ground Truth) 综合重叠度 (intersection over union, IoU) 足够高的先验框, 才能使得模型的初始误差较小, 在训练过程中才能完成对待检测区域的位置回归和置信度的预测。研究者的一般做法是使用 K-means 算法对目标框进行聚类分析, 但是本文研究发现, 由于现实场景中的车辆姿态的千变万化, 导致了每个类别目标框之间互有重复, K-means 算法不能够很好地对目标框坐标数据进行聚类。为了使得单点多盒算法更好地适应本文的车辆图像数据, 所以本文采用了性能更好的最大期望估计的高斯混合模型 (Gaussian mixed model & expectation maximization algorithm, GMM-EM)^[9-10] 聚类算法对车辆数据集目标框进行聚类分析。

首先将车辆数据集所标注的 Ground Truth 的宽高值作为二维向量进行聚类训练。假设训练集 Ground Truth 的二维向量为 y , 则其满足如下形式的概率分布模型:

$$P(y | \theta) = \sum_{k=1}^K \alpha_k \phi(y | \theta_k) \quad (1)$$

其中, α_k 是系数, $\alpha_k \geq 0$, $\sum_{k=1}^K \alpha_k = 1$; $\phi(y | \theta_k)$ 是高斯分布密度, $\theta_k = (\mu_k, \sigma_k^2)$, 高斯混合模型由若干个分模型组

成, 分模型 k 的表达式如式 (2) 所示。

$$\phi(y | \theta_k) = \frac{1}{\sqrt{2\pi}\sigma_k} \exp\left(-\frac{(y - \mu_k)^2}{2\sigma_k^2}\right) \quad (2)$$

GMM-EM 聚类算法分为 3 步进行迭代训练, 具体步骤如下。

1) 首先随机的获取初始值, 并开始迭代训练。

2) 根据当前的模型参数, 计算分模型对 y_j 数据的响应。

$$\hat{\gamma}_{jk} = \frac{\alpha_k \phi(y_j | \theta_k)}{\sum_{k=1}^K \alpha_k \phi(y_j | \theta_k)}, j = 1, 2, \dots, N; k = 1, 2, \dots, K \quad (3)$$

3) 计算新一轮迭代的模型参数。

$$\hat{\mu}_k = \frac{\sum_{j=1}^N \hat{\gamma}_{jk} y_j}{\sum_{j=1}^N \hat{\gamma}_{jk}}, k = 1, 2, \dots, K \quad (4)$$

$$\hat{\sigma}_k^2 = \frac{\sum_{j=1}^N \hat{\gamma}_{jk} (y_j - \mu_k)^2}{\sum_{j=1}^N \hat{\gamma}_{jk}}, k = 1, 2, \dots, K \quad (5)$$

$$\hat{\alpha}_k = \frac{\sum_{j=1}^N \hat{\gamma}_{jk}}{N}, k = 1, 2, \dots, K \quad (6)$$

迭代计算步骤 2) 和 3), 直至模型收敛。

1.3 特征融合与识别

1) 卷积特征提取

本文提出的 MCRF-SSD 算法的主干网络有很好的特征提取能力,为了使本文设计的车辆重识别系统更高效,识别效率更高。所以在全局特征、局部特征提取和识别阶段,本文会复用 MCRF-SSD 主干网络的 75×75 尺度特征层以前的所有卷积层,将 MCRF-SSD 网络中 75×75 的所有卷积特征接入到特征融合与识别模块。首先将训练收敛的 MCRF-SSD 网络的参数冻结,然后获取 conv4-3 上一层卷积层的输出,作为特征融合与识别模块的输入。由于 MCRF-SSD 算法网络的输出是相对于先验框的偏移,所以要先将 MCRF-SSD 输出的局部区域的数值还原到输入图像中的真实坐标,然后根据卷积层与输入图像的尺度比截取相应的卷积特征。假设 $d(d^{cx}, d^{cy}, d^w, d^h)$ 是先验框坐标, $b_j(b_j^{cx}, b_j^{cy}, b_j^w, b_j^h)$ 是车辆局部区域边界的坐标, $l_j(l_j^{cx}, l_j^{cy}, l_j^w, l_j^h)$ 是预测的数值,可以由 l_j 求得 b_j 为:

$$b_j^{cx} = d_j^{cx}(\alpha \times l_j^{cx}) + d_j^{cx} \quad (7)$$

$$b_j^{cy} = d_j^{cy}(\beta \times l_j^{cy}) + d_j^{cy} \quad (8)$$

$$b_j^w = d_j^w \exp(\chi \times l_j^w) \quad (9)$$

$$b_j^h = d_j^h \exp(\delta \times l_j^h) \quad (10)$$

其中, $\alpha, \beta, \chi, \delta$ 为调节系数。

根据 $b_j(b_j^{cx}, b_j^{cy}, b_j^w, b_j^h)$ 坐标和卷积组输出尺度可以进一步求出车辆局部区域在卷积组中的位置信息 $b_c(b_c^{cx}, b_c^{cy}, b_c^w, b_c^h)$ 。

$$b_c^{cx} = \frac{b_j^{cx}}{ci/cp} \quad (11)$$

$$b_c^{cy} = \frac{b_j^{cy}}{ci/cp} \quad (12)$$

$$b_c^w = \frac{b_j^w}{ci/cp} \quad (13)$$

$$b_c^h = \frac{b_j^h}{ci/cp} \quad (14)$$

其中, ci 是输入图片的尺寸, cp 是卷积组输出的尺寸。根据计算得到的 $b_c(b_c^{cx}, b_c^{cy}, b_c^w, b_c^h)$ 可以获得对应的卷积组中的局部特征,并输入到特征融合与识别模块。

2) 特征融合与识别模块

VGG^[11] 网络凭借其简单的结构和可靠的性能在工业界占据非常重要的地位,但是其性能无法比肩 ResNet^[12-13] 等极其复杂的网络。但是 VGG 网络凭借着单一结构,计算效率非常高,非常适合实际场景中的应用。所以本文全局分支借鉴了 VGG 网络的最新改进型 RepVGG^[14] 模型,在预测阶段,通过式(15)将模型参数重构,构成单路模型。将多分支模型性能高和单路模型速度快占用内存少的优势结合起来。

$$w'_{i,\dots,i} = \frac{\gamma_i}{\sigma_i} w_{i,\dots,i}, b'_i = -\frac{\mu_i \gamma_i}{\sigma_i} + \beta_i \quad (15)$$

特征融合与识别模块由全局分支和局部分支组成,全局分支负责对车辆信息的全局特征的提取,局部分支负责对车辆的局部特征进行提取。由于主分支需要兼顾低维度的纹理、边缘特征和高维度的抽象特征,因此全局分支采用了高深度、多通道的复杂特征提取网络。局部分支由于只负责车辆的局部特征的提取,其从检测网络获取的特征尺度较小,并且局部分支中的车标、车灯、进气格栅、车窗、车顶、车头等输入特征要注重他们的纹理、形状等浅层的特征参与车辆的重识别。分开来看,车标、车灯、进气格栅等局部特征比较容易识别,并且特征融合与识别模块得到的特征图较小,只需要浅层的网络就能够很好地完成识别。车窗、车顶、车头的局部特征不容易分辨,与背景区分度不大,并且特征融合与识别模块得到的特征图较大,需要采用特征提取较强的识别网络。所以本文会根据局部特征的不同情况有针对性的设计不同配置的 RepVGG 网络来构成局部分支簇。从而兼顾性能和耗时的平衡。全局和局部分支簇网络结构如表 1 所示。

表 1 全局和局部分支簇网络结构

stage	全局	车标	车灯	格栅	车头	车窗	车顶
1	$2 \times 64 \times 75$	$1 \times 32/2$	$1 \times 32/2$	$2 \times 64/2$	$4 \times 128/2$	$2 \times 64/2$	$2 \times 32/2$
2	$2 \times 128 \times 38$	$2 \times 64/2$	$2 \times 64/2$	$4 \times 128/2$	$4 \times 256/2$	$4 \times 128/2$	$2 \times 64/2$
3	$4 \times 256 \times 19$	$2 \times 128/2$	$4 \times 128/2$	$4 \times 128/2$	$8 \times 512/2$	$4 \times 128/2$	$2 \times 64/2$
4	$4 \times 512 \times 10$	—	$8 \times 128/2$	$8 \times 256/2$	$8 \times 512/2$	$8 \times 256/2$	$4 \times 128/2$
5	$8 \times 512 \times 5$	—	—	—	—	—	—
6	$1 \times 1\,024 \times 1$	—	—	—	—	—	—

本文采用全局分支和局部分支特征融合的方法来输出最终的车辆重识别结果。特征融合一般采取的方式是特征向量的串联、特征图的相加、金字塔结构等。特征向量的串联会破坏卷积特征的空间信息,特征图的相加和金字塔结构是对全部的特征图进行相加,没有聚焦到局部重

点区域,计算效率低。同时由于高层卷积核感受野尺度较大,高层的特征图会被其他区域特征干扰。综合以上特征融合的不足本文提出了焦点融合(focus fusion structure, FFS)方法,如图 7 所示是某个局部区域经过局部分支簇得到的一个卷积特征,首先获取局部分支簇的卷积特征,为

了保持局部和全局卷积特征数值尺度一致,通过式(17)对局部卷积特征归一化。然后根据其在原始图像中的坐标,并结合当前卷积组与原图的缩放比例,确定卷积特征上下左右需要补0的维度,从而保证当前局部卷积特征的大小与要融合的全局分支卷积特征一致,并且局部区域保持对齐。假设 x_i 是第 i 个批次局部区域局部分支的输入特征图, x'_{iq} 是第 i 个批次局部区域的第 q 个卷积层输出特征图, k_q 是 x'_{iq} 相对于首层输入 x_0 的下采样尺度。 $x_{iq}^{(c)}$ 是当前批次 c -th 样本第 q 个卷积层节点的输出, x'_{iq} 为 $[x_{iq}^{(1)}, x_{iq}^{(2)}, x_{iq}^{(3)}, \dots, x_{iq}^{(m)}]$ 组成的行向量,长度为批次的样本数量 m , μ 和 σ 是该行的均值和标准差, ϵ 是防止零输入的极小的扰动量, γ 和 β 是该行批次归一化的训练参数。

$$k_q = L \left(\frac{x_0}{x'_{iq}} \right) \quad (16)$$

$$\hat{x}_{iq}^{(c)} = \gamma \times \left(\frac{x_{iq}^{(c)} - \mu(x_i)}{\sqrt{\sigma(x_i)^2 + \epsilon}} \right) + \beta \quad (17)$$

$$x''_{iq} = f(\chi) \{0 \xrightarrow{\cup} \hat{x}_{iq}\} \quad (18)$$

$$X = \phi_{concat} \{f_{q-1}(x''_{iq} \otimes \text{conf})\} \quad (19)$$

$$X_{total} = \phi_{concat} \{f_{q-1}^2(x_{iq}^{main}, X)\} \quad (20)$$

由式(11)~(14)以及卷积特征与输入特征的缩放比例 k_q 可以根据式(18)求得图中补0维度分别为:

$$L_{up}^y = \left(b_c^{cy} - \frac{1}{2} b_c^h \right) / k_q \quad (21)$$

$$L_{low}^y = y_0 / k_q - \left(b_c^{cy} + \frac{1}{2} b_c^h \right) / k_q \quad (22)$$

$$L_r^x = x_0 / k_q - \left(b_c^{cx} + \frac{1}{2} b_c^w \right) / k_q \quad (23)$$

$$L_l^x = \left(b_c^{cx} - \frac{1}{2} b_c^w \right) / k_q \quad (24)$$

为了使局部分支的卷积特征图更加高效的与全局分支融合,本文在局部分支簇每个通道的卷积特征都引入了一个可训练的权重 conf 如式(19)所示,其初值为每个通道卷积特征所属的局部分支分类识别时输出的置信度。

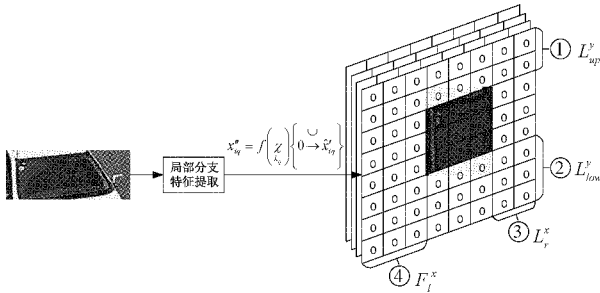


图7 焦点融合计算过程

1.4 损失函数

一般的卷积神经网络的分类损失函数为 Softmax loss,通过该损失函数训练得到的深度特征,会把整个特征空间硬性的按照分类个数划分,确保了训练任务是可分别

的,在以往的分类任务中占据了非常重要的地位。但是对于人脸识别、车辆中识别等类别数特别多、类别之间相似度很大的分类任务,Softmax loss 不能达到很好的分类效果。如式(26)所示, Arcface loss 对特征和权重进行了 L2 正则化,并将角度距离 m 引入进来,使得比 Softmax loss 损失函数分类更加的严格,同时使得类内更加的聚拢,类内间距小,类间间距大。

$$L_{softmax} = -\frac{1}{n} \sum_{i=1}^n \log \frac{e^{w_{y_i}^T f_i}}{\sum_{j=1}^c e^{w_j^T f_i}} = -\frac{1}{n} \sum_{i=1}^n \log \frac{e^{\|w_{y_i}\| \|f_i\| \cos(\theta_{y_i})}}{\sum_{j=1}^c e^{\|w_j\| \|f_i\| \cos(\theta_j)}} \quad (25)$$

$$L_{arcface} = -\frac{1}{n} \log \frac{e^{s \cos(\theta_{y_i}, i-m)}}{e^{s \cos(\theta_{y_i}, i+m)} + \sum_{k \neq y_i} e^{\cos \theta_k, i}} \quad (26)$$

2 实验

2.1 测试数据集

本文选取的实验数据是 VehicleID 和 VeRi 车辆重识别数据集,VehicleID 是采自中国道路卡口监控所拍摄的图片,数据集中车辆的姿态和所处的环境与实际应用中的姿态和环境一致,采用 VehicleID 数据集所得到的测试指标具有很高的应用参考价值和实际应用意义。VehicleID 一共有 26 267 辆车,总计 221 763 张车辆样本。VehicleID 将测试集分为了 Small、Medium、Large 3 个大小不同的测试数据集,分别有 800、1 600、2 400 个车辆 ID,本文实验将根据 3 个测试机分别展开。VeRi 车辆重识别数据集是福建省道路卡口摄像头所拍摄的真实场景中的图像,数据集包含了 49 357 张总共 776 辆车辆图像。VeRi 数据集将 37 778 张图像总共 500 辆车作为训练集,剩余的 11 579 张图像总共 200 辆车作为测试集。在测试集中 VeRi 数据集将 1 678 张车辆图像划分为查询集,9 901 张车辆图像划分为候选图像集合。

2.2 评估指标

为了验证本文提出的 MCRF-SSD 检测算法的有效性,本文采用平均精度(mean average precision, mAP)作为评估检测网络的指标。其计算方式通过式(29)计算 precision-recall 曲线下的面积得出 AP,然后通过式(30)计算出 AP 的均值 MAP。其中 P 代表 precision , r 代表 recall , TP 表示预测框与目标框重叠度大于 0.5, FP 表示预测框与目标框的重叠度小于 0.5, FN 表示检测网络没有正确的预测出局部区域。 N 代表 precision-recall 曲线插值数量, C 代表了检测类别数,本文中为 6。

$$\text{precision} = \frac{TP}{TP + FP} \quad (27)$$

$$\text{recall} = \frac{TP}{TP + FN} \quad (28)$$

$$AP = \sum_{k=1}^N \max_{\tilde{k} \geq k} P(\tilde{k}) \Delta r(k) \quad (29)$$

$$MAP = \frac{\sum_{i=0}^C AP_i}{C} \quad (30)$$

通过计算测试集与查询集的相似度,对查询集进行相似度排序,本文采用 $Rank-k$ 作为评价指标,其计算方法如式(31)所示。其中 Q 为查询集样本数量, K 值一般取 1、3、5, $g(k)$ 是判断函数,当查询图片相似度排在前 K 张查询图片中为 1,否则为 0。

$$Rank-k = \frac{1}{K} \sum_{k=1}^K g(k), K \leq Q \quad (31)$$

2.3 MCRF-SSD 模型实验分析

MCRF-SSD 模型性能越好,局部区域定位越精确,其检测出的局部区域对车辆重识别的匹配越有利。由表 2 可知,单点多盒算法在应用了 GMM-EM 聚类算法之后在 VehicleID 数据集上 MAP 和 aver-IOU 均有提升,与原始的 SSD 算法相比 MAP 提升了 2.4%, aver-IOU 提升了 10.9%。比应用了 K-means 聚类算法 MAP 和 aver-IOU 分别提高 1.1% 和 8.7%。相比说明了 GMM-EM 聚类算法对定位精度的提升帮助很大,对局部区域特征的提取奠定了基础。其次,本研究提出的 MCRF-SSD 的 MAP 与 SSD 和 YOLO v5 相比分别提升了 10.0% 和 2.8%, aver-IOU 分别提升了 10.9% 和 12.2%,均有很大地提升。证明了本文提出的多路、重建的聚焦结构的有效性。

表 2 MCRF-SSD 测试对比结果

模型	MAP	aver-IOU
SSD	0.870	0.793
SSD & K-means	0.883	0.815
SSD & GMM-EM	0.894	0.902
YOLO v5	0.942	0.780
MCRF-SSD	0.970	0.902

分开来看,局部区域检测的 AP 和平均重叠度如图 8 所示,由图 8 可知所有类别的局部区域的 AP 均达到了 92% 以上,说明本文设计的 MCRF-SSD 具有较好的检测性能,为后续局部分支作用于车辆重识打下了基础。

2.4 消融实验

1) 特征融合实验

车辆重识别的本质是对车辆图像特征的提取,特征提取的好坏直接影响着车辆重识别的匹配率。因此本文通过实验来确定局部分支簇和全局特征之间特征融合方式。本部分实验主要对比在 VehicleID 数据集上不同特征融合方式和上采样对匹配率的影响。由图 9 可知,本文提出的焦点融合方法的在整体上要优于直接将卷积特征串联。验证了本文提出的焦点融合方法的有效性。在卷积通道的组合方式上通道拼接由于加入了可学习参数,使得网络

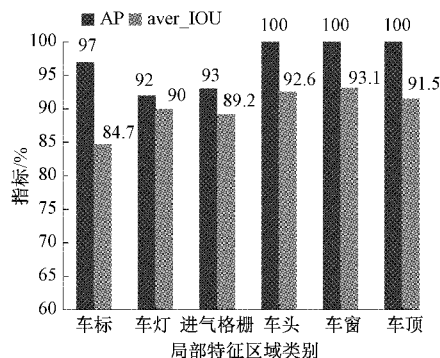


图 8 局部区域检测结果

自动的学习通道之间的组合关系,要优于直接将卷积特征相加。

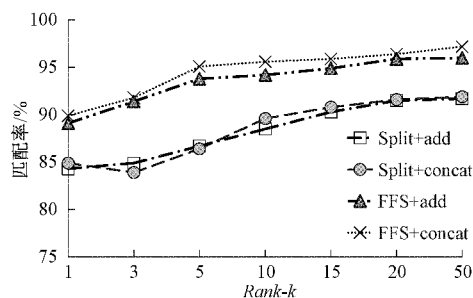


图 9 不同特征融合方式的匹配率

2) 相似车辆分析

为了验证本文提出的重识别方法对于图 3、4 出现的针对同一品牌车头很相似的样本识别性能的提升,从 VehicleID 测试集中分品牌挑选出了 10 个主品牌,平均每个主品牌 10 个 ID,总计 100 个 ID 的相似样本。由于本部分实验只是测试同一个品牌之间的匹配率,所以只测试了 Rank1、Rank3、Rank5 的指标,测试结果如表 3 所示,相比仅采用主分支的重识别方案,将局部分支簇加入进来, Rank1、Rank3、Rank5 均有了大幅度的提升,并且 Rank1 的平均匹配率达到了 0.854,说明了将局部分支簇引入后,大部分同一辆车样本的相似度维持在了较高的水平。验证了本文提出的局部分支簇对于同一品牌相似样本在匹配率上的提升。

表 3 相似车辆测试结果

方法	Rank1	Rank3	Rank5
仅主分支	0.601	0.670	0.765
主分支+局部分支簇	0.854	0.892	0.917

3) 损失函数分析

卷积特征提取的好坏一方面在于特征提取网络结构的设计,另一方面在于特征提取网络损失函数的设计。本文在主分支和局部分支簇的损失函数均采用了 Arcface loss。本部分实验将分别实验不同的损失函数,测试他们

对匹配率的影响。由表 4 可知,即使采用了最普通的 Softmax 损失函数,本文提出的车辆重识别方法在 3 个测试集上 $Rank1$ 、 $Rank5$ 均达到了较高的水平。由表 4 第 2 行可知 Center loss 相较于 Softmax loss 在匹配率上有一定的提升,但经过测试发现 Arcface loss 性能表现最优。

2.5 总体分析

本文选取了 VAMI^[15]、QD-DLF^[16]、TAMR^[17]、EALN^[18]、RAM^[19]、Part-Regularized^[20]、XVGAN^[21]、DDDFL^[22]、文献[23]在 VehicleID 和 VeRi 测试集上进行对比测试,测试结果如表 5、6 所示。本文所提出的算法在 $Rank1$ 、 $Rank5$ 的测试指标上均优于现存的算法,在 Samll 测试集上 $Rank1$ 至少提高了 14.0%, $Rank5$ 至少提高了

表 4 不同损失函数对匹配率的影响

损失函数	Small		Medium		Large	
	$Rank1$	$Rank5$	$Rank1$	$Rank5$	$Rank1$	$Rank5$
Softmax	0.893	0.949	0.889	0.931	0.872	0.916
Center	0.902	0.952	0.891	0.933	0.880	0.918
Arcface	0.924	0.974	0.892	0.944	0.889	0.942

5.0%。在 Medium 测试集上 $Rank1$ 、 $Rank5$ 至少提高了 14.2%和 5.5%,在 Large 测试集上 $Rank1$ 、 $Rank5$ 至少提高了 14.7%和 10.8%。在 VeRi 测试集上 $Rank1$ 、 $Rank5$ 至少提高了 1.5%和 0.7%,说明了本文提出的车辆重识别算法的可行性和性能提升。

表 5 VehicleID 匹配率对比

方法	Small		Medium		Large	
	$Rank1$	$Rank5$	$Rank1$	$Rank5$	$Rank1$	$Rank5$
VAMI	0.631	0.832	0.529	0.751	0.473	0.703
TAMR	0.660	0.797	0.629	0.768	0.597	0.739
QD-DLF	0.723	0.924	0.707	0.889	0.641	0.834
EALN	0.751	0.881	0.718	0.839	0.693	0.814
RAM	0.752	—	0.723	—	0.677	—
Part-Regularized	0.784	—	0.750	—	0.742	—
文献[23]	0.770	0.850	0.718	0.808	0.663	0.784
本文	0.924	0.974	0.892	0.944	0.889	0.942

表 6 VeRi 匹配率对比

方法	$Rank1$	$Rank5$
XVGAN	0.602	0.770
VAMI	0.770	0.908
EALN	0.844	0.941
QD-DLF	0.885	0.945
DDDFL	0.887	0.957
文献[23]	0.881	—
本文	0.902	0.964

3 结 论

本文设计了一种基于局部特征检测和特征融合的车辆重识别模型。针对车辆局部区域对车辆重识别的重要性和车辆局部区域分布的特性,设计出了多路聚焦、重建的单点多盒检测算法。在公开数据集上验证了检测模型在车辆局部区域检测任务上性能的提升。在局部特征提取中针对局部特征的不同特性,有针对性的设计了局部特征簇。在特征融合阶段考虑了特征的空间分布和卷积核感受野的干扰,提出了高效率的焦点特征融合方法,结合可学习参数自动的完成全局特征和局部特征的融合。在 VehicleID 和 VeRi 数据集上实验表明,提出的方法性能均

优于现存算法。但本文所提出的方法仍具有一定的局限性,在下一步的研究中结合注意力网络和车辆的多种属性研究,分析更多的局部特征和属性,如车辆颜色、车型、车辆大小等特征,进一步提高算法的 $Rank1$ 性能。

参考文献

- [1] 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.
- [2] 潘海鹏,王云涛,马森. 基于注意力机制与多尺度融合学习的车辆重识别方法[J]. 浙江理工大学学报(自然科学版),2021,45(5):657-665.
- [3] 刘鸣瑄,刘惠义. 基于特征融合 SSD 的远距离车辆检测方法[J]. 国外电子测量技术,2020,39(2):28-32.
- [4] 李熙莹,周智豪,邱铭凯. 基于部件融合特征的车辆重识别算法[J]. 计算机工程,2019,45(6):12-20.
- [5] 窦鑫泽,盛浩,吕凯,等. 基于高置信局部特征的车辆重识别优化算法[J]. 北京航空航天大学学报,2020,46(9):1650-1659.
- [6] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. ECCV 2016: Computer Vision-ECCV, 2016,9905:21-37.

- [7] 李晖晖,周康鹏,韩太初. 基于 CReLU 和 FPN 改进的 SSD 舰船目标检测[J]. 仪器仪表学报, 2020, 41(4): 183-190.
- [8] ZHOU B, KHOSLA A, LAPEDRIZA A, et al. Learning deep features for discriminative localization [J]. IEEE Conference on Computer Vision and Pattern Recognition, 2016:2921-2929.
- [9] LIU Y, YUAN X, LIANG Y C, et al. Machine learning based iterative detection and multi-interference cancellation for cognitive IoT [J]. IEEE Communications Letters, 2020, 24(9):1995-1999.
- [10] 谭博,郭静波,常广,等. 基于正弦高斯混合模型的磁目标通用快速检测[J]. 仪器仪表学报, 2019, 40(2): 1-10.
- [11] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]. International Conference on Learning Representations, San Diego: IEEE, 2015: 1-14.
- [12] 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.
- [13] 王飞,张莹,邵豪,等. 多尺度残差网络模型的研究及其应用[J]. 电子测量与仪器学报, 2019, 33(4):19-28.
- [14] DING X, ZHANG X, MA N, et al. RepVGG: Making VGG-style ConvNets great again[J]. ArXiv, 2021, ArXiv:2101.03697.
- [15] YI Z, LING S. Viewpoint-aware attentive multi-view inference for vehicle re-identification[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018:6489-6498.
- [16] ZHU J, ZENG H, HUANG J, et al. Vehicle reidentification using quadruple directional deep learning features[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(1):410-420.
- [17] 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 A Publication of the IEEE Signal Processing Society, 2019, 28(9):4328-4338.
- [18] LOU Y, BAI Y, LIU J, et al. Embedding adversarial learning for vehicle re-identification [J]. IEEE Transactions on Image Processing, 2019, 28(8): 3794-3807.
- [19] 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 (ICME), IEEE, 2018, DOI:10.1109/ICME.2018.848659.
- [20] HE B, LI J, ZHAO Y, et al. Part-regularized near-duplicate vehicle re-identification [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019:3992-4000.
- [21] ZHOU Y, SHAO L. Cross-view gan based vehicle generation for re-identification [C]. British Machine Vision Conference, 2017, DOI:10.5244/C.31.186.
- [22] 邱铭凯,李熙莹. 用于车辆重识别的基于细节感知的判别特征学习模型[J]. 中山大学学报(自然科学版), 2021, 60(4):111-120.
- [23] WANG Q, MIN W, HE D, et al. Discriminative fine-grained network for vehicle re-identification using two-stage re-ranking [J]. Science China Information Sciences, 2020, 63(11):1-12.

作者简介

李浩, 硕士, 助教, 主要研究方向为人工智能、图像处理。

E-mail: 2854584059@qq.com

杨超, 博士研究生, 副教授, 主要研究方向为图像处理/电气工程。

E-mail: jxdxyc@163.com

黄友新, 硕士, 高级工程师, 主要研究方向为人工智能、模式识别。

E-mail: 83220110@qq.com

陈嘉哲, 硕士, 助教, 主要研究方向为谐振变换器拓扑、控制及其应用。

E-mail: 1636368187@163.com

詹瑞典, 硕士, 工程师, 主要研究方向为信号/图像处理、集成电路设计。

E-mail: zhanruidian@chipeye.cn

鲍鸿(通信作者), 博士, 教授, 主要研究方向为图像处理、人工智能。

E-mail: bhong@gdut.edu.cn