

DOI:10.19651/j.cnki.emt.2210695

融合多级注意力机制和信息融合的车辆识别*

李浩¹ 鲍鸿² 詹瑞典¹

(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)

摘要: 不同车型类间外观特征高度相似,同车型类内外观差异大,这对特征提取网络提出了更高的要求。现有的车型识别方案仅依靠车辆外观特征识别,整体识别准确率不高。为此,首先在主干网络设计了多级注意力机制,提高主干网络对车型特征提取和识别能力;其次根据卡口环境下不同车辆位置车辆外观特征的变化提出了车辆位置和外观特征融合结构,从而提取出融合位置信息的复合图像特征,减小类内特征距离,增强主干网络所提取的特征的表达力和稳健性;最后在分析了难例样本注意力热力图基础上,对难例样本注意力区域进行干预,使网络聚焦于车辆细小差异的局部区域。实验结果表明,本文所提出的车型识别方法整体性能比现有方案有显著提升。

关键词: 主干网络;注意力机制;特征融合;热力图

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

Research on vehicle type recognition based on multilevel attention mechanism and information fusion

Li Hao¹ Bao Hong² Zhan Ruidian¹

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

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

Abstract: The appearance characteristics of different models are highly similar and different from those of the same model, which poses a great challenge to the feature extraction network. Existing vehicle type classification schemes only rely on vehicle appearance feature recognition, and the overall recognition accuracy is not high. Therefore, firstly, this paper designs a multi level attention mechanism in backbone network to improve the ability of main network to extract and recognize vehicle features. Secondly, according to the changes of vehicle appearance characteristics at different vehicle locations in the bayonet environment, a feature fusion structure of vehicle location and appearance features is proposed, which extracts the composite image features of the fusion location, reduces the feature distance within the class, and enhances the expressiveness and robustness of the features extracted by the main network. Finally, based on the analysis of the attention heat map of difficult samples, the attention area of difficult samples is intervened to make the network focus on the local area of small differences between vehicles. The experimental results show that the overall performance of the vehicle type recognition method proposed in this paper is significantly improved than the existing scheme.

Keywords: backbone network; attention mechanism; feature fusion; attention heat map

0 引言

随着社会经济的发展,城市的规模和汽车保有量不断的增加。这导致了公共交通日渐拥挤和堵塞,并且严重影响着公共安全。目前我国城市已基本完成视频监控的覆盖,由基于深度学习^[1]的视频分析技术所构成的车型识别系统^[2]为改善公共交通和维护公共安全提供了新的技术手

段。车型识别^[3-5]是智能交通系统的一个重要分支,在交通流量统计、特种车辆检测、交通调度等方面有广泛的应用。然而,由于道路卡口光线的干扰、车型各个类别之间细小的差异、同车型外观差异大,以及车辆与卡口摄像头的角度和距离的不同,造成了同一个车型所呈现的外观特征不同,使得车型识别的性能不高,无法有效的应对交通管控。其次卡口的摄像头一般在正对车辆方向的上方,卡口监控范围

收稿日期:2022-07-15

* 基金项目:国家自然科学基金(62001127)、广东工业大学高等教育研究基金(GXLX20210213)项目资助

内的车辆由于距离摄像头距离的不同,造成了车辆图像特征的变化,目前的方法容易造成不同车型之间的误识别,这使得精确的车辆识别任务具有挑战性。

到目前为止,已经提出了许多基于图像的方法,如基于车辆图像纹理特征和基于深度学习技术自动提取外观特征等多种方法。文献[6]提出了一种基于傅里叶描述子,通过形态学操作得到车辆的轮廓封闭曲线,进而识别车型。并利用积分的方法确定车辆的大小信息,解决了车型识别中的多检和错检的问题。文献[7]提出了一种结合最有效Gabor特征的两阶段车型识别方法。首先利用边缘特征,通过相似k近邻分类器将车辆按大小进行分类,第二阶段基于核稀疏表示的分类器做更具体的车型分类。提高了车型识别的精度核鲁棒性。文献[8]针对车型识别系统中的计算时间长、误识率高的问题,提出了一种联合特征和压缩字典的车型识别算法,将车辆的边缘形状特征和纹理特征串联,构建特征字典进行特征抽取和车型识别,提高了车型识别的准确率和实时性。文献[9]提出了一种基于改进Faster R-CNN的车辆类型识别算法,为了解决外形相似的车辆类型误识别问题,提出了一种空洞卷积金字塔池化结构来增强多尺度信息的获取。文献[10]提出了一种基于YOLO V3的多场景车辆属性识别算法,并针对颜色和车型不同的图像特征属性进行了分级训练,满足了视频场景下的实时性要求。文献[11]针对车辆所处的自然环境复杂多变和车辆精细型号种类繁多等因素,提出一种基于多尺度特征融合的车辆精细型号识别方法,融合了高层和底层的车辆图像特,提高了分类准确率。文献[12]基于 AlexNet 提出了用于车型识别的深度卷积神经网络模型,但 AlexNet 网络提取特征的能力有限,无法应对复杂的交通环境。文献[13]提出了一种将卷积核池化和稀疏栈式编码融合的车辆识别方法,在少量标注数据集上达到了较好的性能,但采用的逐层无监督训练方法不仅增加了训练难度,

也阻隔了不同尺度特征的信息交换。文献[14]提出了一种利用半监督的稀疏拉普拉斯滤波器的卷积神经网络,与卷积核参数自学习相比,缺少了一定的自适应性。文献[15]将深度卷积神经网络引入车型识别中,训练分类器实现车型识别,没有考虑车型的不同尺度的特征。

基于以上分析,现如今的基于深度学习的车辆类型识别方法主要集中在多尺度、多属性特征的融合上,模型设计都是在整体上进行优化,没有针对易混淆车辆设计模型,造成了模型对此类样本识别率不高。大部分研究将应用场景设定在了白天光线良好的环境中,在夜间或者白天大光差的环境中无法正确的识别车辆类型。此外,没有应用车辆与卡口摄像头之间相对的位置信息。

针对上述问题,本研究以 ResNetXt^[16] 网络为基础,通过网络注意力机制^[17]精简网络模型参数规模,加快模型推理速度。分析了车型外观差异小的样本,提出了一种融合位置信息的强监督渐进式模型训练方法。将图像增强技术引入到夜间车辆类型识别中,增强算法模型对夜间样本的鲁棒性。

1 车型识别算法

本研究设计的车辆类型识别模型如图1所示,所提出的车辆类型识别方法由图像特征增强、主干网络、热力图注意力转移、位置分支4个部分组成。首先,针对夜间光线较差或者白天大光比环境下造成的车辆图像特征丢失的问题引入了图像特征增强模块,以强化车辆图像细节特征。以优化过后的 ResNetXt 网络作为主干网络完成初次的车辆类型识别,本文针对难例样本在分析热力图的基础上提出了特征图注意力转移机制,通过再训练实现对难例样本的识别,提高车辆类型识别系统的识别性能和稳健性。同时,针对车辆与卡口摄像头由于相对位置不同,造成车辆外观特征变化,从而导致识别错误的问题,本文将位置信息引入,与图像特征融合,共同作用于识别结果。

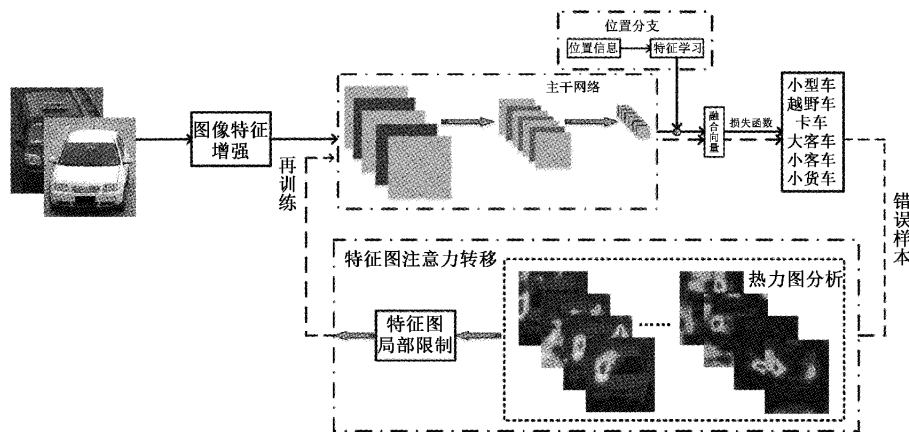


图1 车型识别结构图

1.1 图像增强

车辆类型识别是基于车辆的外观特征实现的,车辆的外观特征直接决定了车辆类型识别的性能。夜间光线太

弱或者白天光线太强形成大光比环境,卡口摄像头所采集的车辆图片过暗、过曝,车辆图片细节、轮廓特征丢失。因此针对此类车辆样本进行图像增强可以尽可能的保留车

辆图像的细节、轮廓等车辆类型识别必要的特征,从而大幅度的提高车辆类型对过曝、过暗样本的识别性能。常见的图像增强方法有直方图均衡化、Retinex^[18]算法等。直方图是图像的一种非常重要的统计特征,常常被用来做相似图像检索^[19]。而直方图均衡化通过重新分配图像的像素值,使一定范围灰度级内的像素均匀分布,这样就造成了原始图像的某些灰度级被抹掉,如图 2 所示造成了图像某些细节的丢失,使得车辆图像与背景难以区分,并且对于灯光等高光范围没有做到很好的压制。这显然不适用于基于车辆细节特征的车辆类型识别。Retinex 图像增强算法是建立在大量科学实验及分析的光学物理特性基础上,如图 2 所示,Retinex 算法可以自适应的在一定的动态范围内实现边缘加强、颜色恒定。相较于直方图均衡化在背景噪声控制、车辆细节还原、高光压制等方面更加优异,更适合对夜间样本的图像增强,所以本文采用 Retinex 图像增强算法。

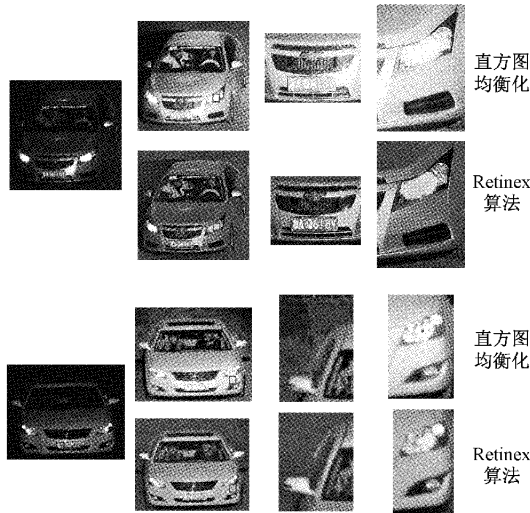


图 2 图像增强效果对比

Retinex 图像增强算法假设一副给定的图像 $S(x, y)$ 由反射图像 $R(x, y)$ 和入射图像 $L(x, y)$, 所以一副给定的图像 $S(x, y)$ 可由式(1)表示。

$$S(x, y) = R(x, y) \cdot L(x, y) \tag{1}$$

Retinex 图像增强算法可由式(2)表示。

$$r(x, y) = \sum_k^K \omega_k \left\{ \log \frac{S(x, y)}{F_k(x, y) * S(x, y)} \right\} \tag{2}$$

$$F_k(x, y) = \lambda_k e^{-\frac{(x^2+y^2)}{c_k^2}} \tag{3}$$

其中, $r(x, y)$ 为输出图像, K 是高斯中心环绕函数的数目, 当 K 取 1 时式(2)就退化为单尺度 Retinex 算法, 相较于单尺度, 多尺度 Retinex 算法能同时保证图像的动态范围和高保真, 所以本文选取 K 为 3, 即同时兼具高、中、低 3 个尺度。其中 ω_k 如(3)所示是高斯中心环绕函数, $*$ 是卷积操作, c_k 是高斯环绕尺度, 与 K 值对应, 经过实

验本文 c_k 取 15、80、120。 ω_k 是 3 个尺度的权重系数, 经过实验, 本文全部取为 1/3。

1.2 注意力模块

深度神经网络在图像识别领域表现出优异的性能, 相较于传统的手工特征, 具有较高性能的同时又兼具良好的抗干扰能力。ResNetXt 网络将分组卷积集成到了 ResNet 网络中, 提高特征提取和图像识别能力, 且模型结构简单, 所以本研究将 ResNetXt 网络作为基础网络。ResNetXt 增加了网络的宽度, 使得模型具有多条支路, 但是多条支路的特征只是进行了简单的相加, 每条支路的权重相同, 在一定程度上限制了模型的特征表达能力。所以本文将通道注意力机制集成到 ResNetXt 网络的各个分支中, 计算公式如式(4)~(6)所示。其中 ck_l 表示第 c 层的第 k 个分支的第 l 通道; H, W 表示特征图的大小; i, j 表示卷积特征的第 i 行 j 列。首先如式(4)所示, 将特征图通过全局平均池化转换成 1×1 , 通道数保持不变。其次通过式(5)对式(4)的结果进行非线性压缩和放大, 在降低复杂度的同时提高模型的泛化能力。其中 $W_1 \in R^{c \times c}, W_2 \in R^{c \times c}$ 。最后通过式(6)得到每个通道的权重。使模型自适应的选取对车型识别重要的特征, 从而提高识别性能。如图 3 所示, 为了使网络更好的抽取表现力强的特征, 本文不仅设计了并行分支的第一级注意力结构, 并且在模型的残差结构位置设计了第二级注意力结构。第一级注意力结构完成对卷积分支输出的特征加权融合, 第二级实现残差特征与主支路特征实现非线性、有选择的融合。

$$Z_{ckl} = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W v_{ckl}(i, j) \tag{4}$$

$$e_x = W_2 \text{ReLU}(W_1 z) \tag{5}$$

$$\sigma(e_x) = \frac{1}{1 + e^{-e_x}} \tag{6}$$

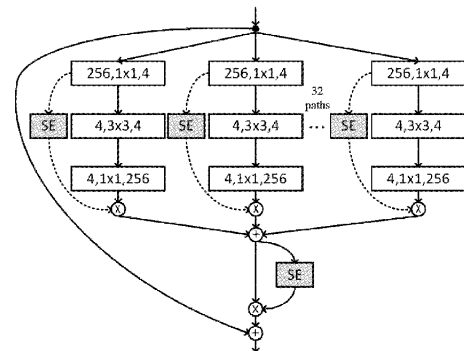


图 3 多级注意力网络结构图

将 ResNetXt 每一个分支都应用注意力机制虽然会提高主干网络特征提取和表征的能力, 但是会增加模型的推理时间。所以, 本文参考 RepVGG 网络的做法, 如式(7)所示将注意力结构在模型推理阶段集成到相应的网络分支中, 简化模型, 提高网络的推理速度。经过上述处理后, 本

文提出的多级注意力网络分支由图4第2列所示。

$$z^t = (w^t \times y^{t-1} + b^t) \times s^{scat} \quad (7)$$

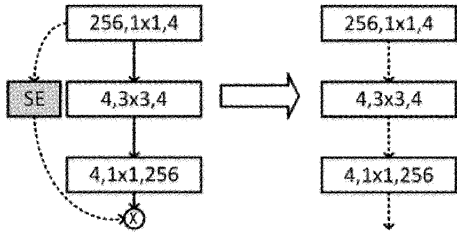


图4 分支参数整合

1.3 细粒度特征选择

如图5所示,1、2、3号样本均为小型车,但是由于和卡口摄像头形成的角度不同,即使同为小型车,外观特征相差很大。1、3号样本分别将车辆的左车窗和右车窗的弧度展现出来,造成了同一类样本外观特征的镜像,而2号样本由于正对卡口摄像头,两侧的车窗弧度均没有很好的展现,容易与5号SUV样本混淆。7、8号样本虽然与卡口摄像头角度相似,但是由于距离的不同,造成了同一辆车车尾特征的不同(距离远无车尾;距离近车尾展现)。4号样本虽然是小型车,但是由于是两厢车,其外观特征整体上与5号SUV样本在卡口摄像头角度下高度相似,但是两厢小型车车顶没有行李架,这成为了区分这两类样本的关

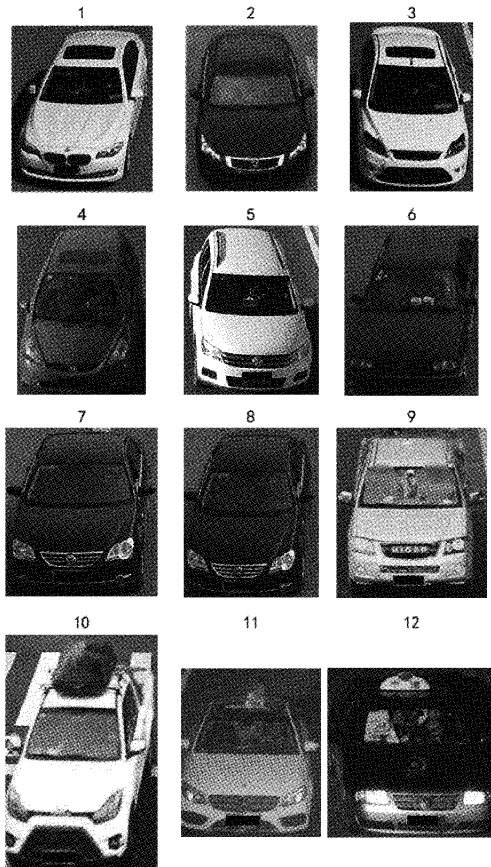


图5 不同角度车辆样例

键特征。6号样本由于其特有的造型,车尾部特征易与9号皮卡的车尾部特征高度相似。10、11、12号样本由于车顶安装有行李箱、天窗站人、特定造型的影响,容易造成识别错误。基于以上分析,造成车辆类型识别错误的原因主要是:1.同一类型车辆由于与卡口摄像头角度的不同,造成车辆外观特征的改变;2.不同类型车辆外观特征的相似,造成识别难度增加。针对问题1本文将车辆的位置信息和车身的大小引入,与车辆图像的卷积特征融合,共同作用于识别结果。针对问题2本文在分析此类错误样本的热力图基础上,加入特征图注意力转移,引导模型识别的更加准确。

1) 位置信息融合

如图6所示,由于车辆位置与卡口摄像头相对位置的不同,车辆与卡口摄像头之间的角度 δ 不同,不同车型的外观所形成的检测框的宽高比以及大小也不相同。将 δ 角余弦值、 w 、 h 以及 w/h 比值特征利用起来,对车辆类型识别至关重要。

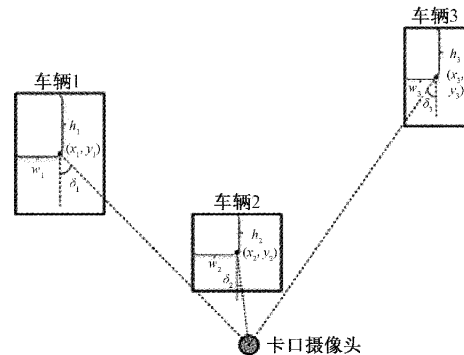


图6 不同车辆位置分布示意图

位置特征由角度 δ 余弦值、 w 、 h 、 w/h 、 x 、 y ,其中目标车辆在卡口摄像头左侧时 δ 为负值,否则为正值。如图7所示,位置分支由三层网络组成,第一层激活函数为Relu,其作用是调整特征分支尺度,以达到和卷积特征更好的融合。第二层激活函数是Sigmoid函数,将特征映射到 $[0,1]$ 区间,自动的学习与卷积特征相加时的权重。

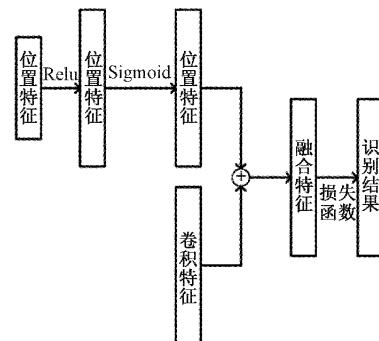


图7 车辆位置分支结构图

2) 特征图注意力转移

由1.3节的分析可知,不同车型的外观特征高度相

似,如两厢小型车和 SUV、皮卡与部分小型车等。造成了车辆类型识别准确率的降低。因此,分析网络注意力,使网络模型注意力聚焦在车辆细小的差异中,对提升车辆类型的识别准确率至关重要。本研究采用类激活映射(class activation mapping,CAM)^[20],分析模型的注意力,从而对网络的注意力进行强监督干预,提升车辆类型的识别准确率。其原理如式(8)所示,其中 A^i 表示进行热力图分析的最后一层特征图的第 i 个通道的特征图; w_i^c 表示分类层中第 c 类对应的第 i 个神经元的权重。因此第 c 类的热力图值 S_c 计算方式如式(9)所示。其中:GAP 表示全局平均池化,计算方式如式(10)所示;特征图的大小为 $c_1 \times c_2$; k 、 j 分别表示特征图的行和列。

$$L_{CAM}^c = \sum_{i=1}^n w_i^c A^i \quad (8)$$

$$S_c = \sum_{i=1}^n w_i^c GAP(A_i) \quad (9)$$

$$GAP(A_i) = \frac{1}{Z} \sum_{k=1}^{c_1} \sum_{j=1}^{c_2} A_{kj}^i \quad (10)$$

CAM 分析了网络模型的注意力集中区域,那么识别错误的样本的注意力所关注的区域显然是错误的。所以,对识别错误样本的注意力所关注的区域进行干预就很有必要。由式(8)可知,经过全局平均池化后每个通道的特征值与对应类别的权重相乘,就得到了一个通道对相应类别的影响。所以权重越大,其所对应的通道特征值对识别结果作用越大。为了对错误的注意力区域进行干预,本文首先对识别性能低的类别的错误样本的权重进行排序,然后根据设定的阈值进行有选择的对权重重置零如式(11)所示,其中 k 为排序后的权重的位置, n 为第 l 层的通道数。本文对排位在前 20%的通道特征值所对应的权重置 0,其他保持不变。将 r^l 与上一层的输出相乘如式(12)所示,对错误的注意力区域的屏蔽,实现注意力区域转移。最后与原有的权重参数 w_c^{l+1} 相乘如式(13)所示,输入到 softmax 函数中完成识别结果的输出。

$$r^l = \begin{cases} 1 & \frac{k}{n} > 0.2 \\ 0 & \frac{k}{n} < 0.2 \end{cases} \quad (11)$$

$$\tilde{y}^l = r^l \times y^l \quad (12)$$

$$z_c^{l+1} = w_c^{l+1} \times \tilde{y}^l \quad (13)$$

2 实 验

2.1 数据集

本探究提出的车辆类型识别方法采用 BIT-vehicle 数据集做对比实验,验证本文所提方法的有效性和性能提升。BIT-vehicle 车辆数据集由 9 850 张卡口摄像头所拍摄的车辆图像构成,共由小型车、越野车、卡车、大客车、小客车、小货车 6 种车辆类型。首先根据预先标注的车辆位置坐标将车辆区域分割出来,并将车辆的位置信息进行记

录。随机抽取 20%的样本作为测试集,剩余的作为训练集。样本分布如表 1 所示。

表 1 BIT-vehicle 数据分布

类别	训练集	测试集	总数
小型车	4 737	1 184	5 921
越野车	1 112	278	1 390
卡车	698	174	872
大客车	446	111	557
小客车	706	177	883
小货车	381	95	476

2.2 评价指标

本文采用精确率(precision,P)、召回率(recall,R)、准确率(accuracy,Acc)作为评估算法性能的指标。其计算公式分别如式(14)~(16)所示。

$$P = \frac{TP}{TP + FP} \quad (14)$$

$$R = \frac{TP}{TP + FN} \quad (15)$$

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (16)$$

其中,TP 是正例判断为正例的样本数;TN 是负例判断为负例的样本数;FP 是负例判断为正例的样本数;FN 是正例判断为负例的样本数。

2.3 实验结果分析

1)整体性能对比

由表 2 可知,本文所提出的车辆类型识别优于目前所有的车辆类型识别方案,准确率达到 94.55%。比目前准确率最高的方案提高了 4.94%,在 BIT-vehicle 数据集上验证了本文方法的有效性和可行性。

表 2 不同方案性能对比

方法	准确率/%
文献[12]	79.23
文献[13]	80.15
SLFL ^[14]	88.11
文献[15]	89.61
本文方法	94.55

2)主干网络实验分析

由表 3 可知,在仅使用 ResNetXt 网络的情况下,车型识别性能只能达到较低的水平。平均召回率、精确率、准确率仅达到了 84.58%、79.26%、85.04%。由表 4 具体来看,主要是小型车和 SUV 以及卡车与小货车之间等不同车型之间外观特征的高度相似,车型之间存在大量的样本互相识别错误。如表 5 所示,在增加一级注意力后,平均召回率、精确率、准确率分别提高了 0.79%、1.62%、

2.03%。提高了网络对车辆外观细粒度特征的提取能力,主要表现在小型车、SUV召回率和精确率大幅度提高。

表3 不同注意力配置性能对比

网络配置	平均召回率/ %	平均精确率/ %	平均准确率/ %
ResNetXt	84.58	79.26	85.04
一级	85.37	80.88	87.07
二级	88.83	84.43	89.85

表4 ResNetXt网络识别结果样本分布

	小型车	越野车	卡车	大客车	小客车	小货车	召回率 /%
小型车	1 025	127	5	3	8	16	86.57
越野车	49	212	2	1	3	11	76.26
卡车	2	0	147	1	3	21	84.48
大客车	0	0	0	107	0	4	96.40
小客车	0	1	1	4	152	19	85.88
小货车	0	1	12	2	6	74	77.89
P/%	95.26	62.17	88.02	90.68	88.37	51.03	

表5 一级注意力识别结果样本分布

	小型车	越野车	卡车	大客车	小客车	小货车	召回率 /%
小型车	1 054	102	1	2	9	16	89.02
越野车	37	229	2	0	3	7	82.37
卡车	2	0	147	1	3	21	84.48
大客车	0	0	2	104	1	4	93.69
小客车	1	2	1	5	150	18	84.75
小货车	0	1	12	2	6	74	77.89
P/%	96.34	68.56	89.09	91.23	87.21	52.86	

如表6所示,在融合了二级注意力机制后召回率、精确率、准确率相较于表5只增加一级注意力机制分别提高了3.46%、3.55%、2.78%,网络模型对于细粒度特征提取能力进一步提升,验证了在主路分支和残差分支特征融合阶段引入注意力机制的有效性。相较于在卷积分支加入注意力机制,性能提升幅度更大。由表6可知,存在大量的小货车被识别成了卡车、小客车,造成了小货车召回率的降低以及其他车型精确率的降低。以及卡车被错误的识别成了小货车,造成小货车精确率只达到了59.52%。因此本文对于此类难例样本的识别引入位置信息融合加热力图分析。

3) 细粒度特征选择实验分析

如表7第1列所示,在引入位置信息后平均召回率、平均精确率、平均准确率相较于表6分别提升了3.04%、4.74%、2.72%。位置信息的引入使得模型更加专注与卡

表6 二级注意力识别结果样本分布

	小型车	越野车	卡车	大客车	小客车	小货车	召回率 /%
小型车	1 076	83	0	0	9	16	90.88
越野车	28	243	0	0	0	7	87.41
卡车	0	0	154	1	3	16	88.51
大客车	0	0	0	110	1	0	99.10
小客车	2	2	0	5	156	12	88.14
小货车	1	0	10	2	7	75	78.95
P/%	97.20	74.09	93.90	93.22	88.64	59.52	

口摄像头两侧车窗弧度的细小差异和卡车后车厢大小的差异,同时位置信息融合模块消除了对于卡口摄像头左侧和右侧车辆车窗所呈现出的弧度差异的镜像,使得同一个车型类内特征距离更小。由表8所示使得小型车召回率提高了2.28%,越野车召回率提高了3.60%。使得SUV的精确率提高了4%,卡车的召回率提高了4.59%。由于卡车被识别成小货车样本的减少,小货车的精确率达到了80.81%,减少了车型之间的混淆。

表7 不同细粒度网络配置性能对比

网络配置	平均 召回率/%	平均 精确率/%	平均 准确率/%
+位置信息融合	91.87	89.17	92.57
+注意力转移	93.93	91.67	94.55

表8 位置信息融合识别结果样本分布

	小型车	越野车	卡车	大客车	小客车	小货车	召回率 /%
小型车	1 103	69	0	0	9	3	93.16
越野车	23	253	0	0	0	2	91.01
卡车	0	0	162	1	3	8	93.10
大客车	0	0	0	109	2	0	98.20
小客车	2	2	0	5	162	6	91.53
小货车	1	0	8	2	4	80	84.21
P/%	97.70	78.09	95.29	93.16	90.00	80.81	

如图8第3列所示在融合了注意力转移后,与图8第2列未融合注意力转移相比网络模型对车辆与卡口摄像头角度很小的样本注意力更加集中于车顶尾部、车窗边缘等细小的差异,同时联合了同车型类内左右两侧不同角度的样本的车辆外观特征。使得小型车多召回了21个样本,召回率提升了1.77%;SUV注意力集中在车顶行李架等SUV特有的特征区域上面,使得SUV多召回了11个样本,召回率提升了3.95%;小货车则将注意力集中在车厢区域上,使得样本多召回了5个样本,召回率提升了5.26%;如表9所示,在总体性能上平均召回率、平均精确

率、平均准确率分别达到了 93.93%、91.67%、94.55%。验证了注意力转移方法的有效性。



图 8 注意力转移效果样例

表 9 注意力转移性能对比

	小型车	越野车	卡车	大客车	小客车	小货车	召回率 / %
小型车	1 124	48	0	0	9	3	94.93
越野车	12	264	0	0	0	2	94.96
卡车	2	1	163	0	1	7	93.68
大客车	0	0	1	108	2	0	97.30
小客车	2	2	0	6	165	2	93.22
小货车	0	0	6	0	4	85	89.47
精确率	98.60	83.81	95.88	94.74	91.16	85.86	

3 结 论

车型识别由于类内特征距离大、类间特征小，卡口环境下背景复杂，车辆数量多，造成了车型识别任务挑战性强。本文从车型识别的难点出发首先改进了主干网络，引入多级的注意力机制。首次将车辆位置信息融合进车辆类型识别任务，解决了车辆类型在不同卡口摄像头角度下识别困难的问题。在分析网络注意力热力图的基础上，强制干预网络模型的注意力区域，改善了对外观高度相似的且与卡口摄像头角度很小的样本识别性能。综合性能优于目前最优方案。后续的研究内容将考虑更多车辆特征，如品牌、颜色提高车辆类型的识别性能。进一步优化网络模型，降低复杂度和训练难度。

参考文献

- [1] LEE J. Integration of digital twin and deep learning in cyber-physical systems: Towards smart manufacturing[J]. 2020, 38(8):901-910.
- [2] CHAN R, LIM M Y, PARTHIBAN R. A neural network approach for traffic prediction and routing with missing data imputation for intelligent transportation system [J]. Expert Systems with Applications, 2021, DOI:10.1016/j.eswa.2021.114573.
- [3] STAROLETOV S M, LAPTEV M A, NEKRASOV D V. Development and testing of algorithms for vehicle type recognition and car tracking with photo and video traffic enforcement cameras [J]. Pattern Recognition and Image Analysis, 2021, 31(2):323-333.
- [4] WANG J, ZHENG H, HUANG Y, et al. Vehicle type recognition in surveillance Images from labeled Web-Nature data using deep transfer learning [J]. IEEE Transactions on Intelligent Transportation Systems, 2018, 19(9):2913-2922.
- [5] DING X, WANG J, DONG C, et al. Vehicle type recognition from surveillance data based on deep active learning [J]. IEEE Transactions on Vehicular Technology, 2020, 69(3):2477-2486.
- [6] 朱金荣, 张广杰, 夏长权. 基于图像处理的车辆类型识别的研究[J]. 现代电子技术, 2017, 40(4):16-18, 23.
- [7] SUN W, ZHANG X R, HE X Z, et al. A two-stage vehicle type recognition method combining the most effective gabor features[J]. Cmc-Computers Materials & Continua, 2018, 65(3):2489-2510.
- [8] 孙伟, 施顺顺, 张小瑞. 基于联合特征和压缩字典学习的车辆类型识别算法[J]. 汽车工程, 2018, 40(3):259-264, 276.
- [9] 宁俊, 王年, 朱明. 基于改进 Faster R-CNN 的车辆类型识别算法[J]. 安徽大学学报(自然科学版), 2021, 45(3):26-33.
- [10] 王林, 柴江云. 深度神经网络在多场景车辆属性识别中的研究[J]. 计算机工程与应用, 2021, 57(9):162-167.
- [11] 刘廷建, 顾乃杰, 张孝慈, 等. 基于多尺度特征融合 CNN 模型的车辆精小型号识别[J]. 计算机工程与应用, 2018, 54(18):154-160.
- [12] 石磊, 王亚敏, 曹仰杰, 等. 基于深度卷积神经网络的车型识别[J]. 计算机科学, 2018, 45(5):280-284.
- [13] 代乾龙, 孙伟. 基于改进稀疏栈式编码的车型识别[J]. 计算机工程与应用, 2020, 56(1):136-141.
- [14] DONG Z, WU Y W, PEI M T, et al. Vehicle type classification using a semisupervised convolutional neural network[J]. IEEE Trans Actions on Intelligent

- Transportation Systems, 2015, 16(4):2247-2256.
- [15] 袁公萍, 汤一平, 韩旺明, 等. 基于深度卷积神经网络的车牌识别方法[J]. 浙江大学学报(工学版), 2018, 52(4):694-702.
- [16] XIE S, GIRSHICK R, P DOLLAR, et al. Aggregated residual transformations for deep neural networks[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 5987-5995, DOI: 10.1109/CVPR.2017.634.
- [17] 徐健, 陆珍, 刘秀平, 等. 注意力机制优化 RetinaNet 的密集工件检测方法研究[J]. 电子测量与仪器学报, 2022, 36(1):227-235.
- [18] ZHANG S, WANG T, DONG J Y, et al. Underwater image enhancement via extended multi-scale retinex[J]. Neurocomputing, 2017, 245:1-9, DOI: 10.1016/j.neucom.2017.03.029.
- [19] 杨迪, 陈宁. 跨域情形下的服装图像检索综述[J]. 国外电子测量技术, 2021, 40(11):24-34.
- [20] 潘云龙, 王森, 张印辉, 等. 结合特征复用注意力与精细化分层残差的细微裂纹密集连续检测[J]. 仪器仪表学报, 2021, 42(2):285-296.

作者简介

李浩, 助教, 硕士, 主要研究方向为图像处理、人工智能。

E-mail: lhao@gdut.edu.cn

鲍鸿, 教授, 博士, 主要研究方向为图像处理、智能信息处理。

詹瑞典, 工程师, 硕士, 主要研究方向为信号/图像处理、集成电路设计。