DOI:10.14048/j.issn.1671-2579.2026.03.026 文章编号:1671-2579(2026)03-0238-10

基于双视角特征融合网络的公路收费车型识别研究

肖强[1,2,3],陈心嘉[1*],陈婧[2,3],赵宾[2,3],晏凯锋[2],丁洪伟[1]

(1.云南大学 信息学院;云南 昆明 650011;2.云南公路联网收费管理有限公司,云南 昆明 650011;
3.云南省数字交通重点实验室;云南 昆明 650011)

摘要:为解决收费公路车型分类中的实际问题,该文提出了一种基于双视角特征融合的深度学习网络——ETCLNet。该网络采用双分支架构,分别提取车辆的头部和侧身图像特征,并结合创新性设计的多尺度特征提取模块(IPFE)与基于自适应加权的双线性特征融合机制(AWBF),实现高效、细粒度的车辆特征表达。IPFE模块通过并行多尺度卷积与残差连接设计,有效缓解了梯度消失问题,并增强了对复杂场景的适应能力;AWBF融合机制进一步优化了角度间特征的高阶交互,通过自适应加权和逐元素乘法,实现了多角度特征的高效融合,显著提升了分类性能。试验结果表明:ETCLNet在多项评价指标(如准确率、精确率、召回率、F_1值和S_{AUC})上均优于现有主流模型,解决了现有收费公路稽核系统中摄像机车型识别率低的问题。此外,该设计充分利用现有硬件设备,降低了部署成本。该研究为多视角车辆识别与智慧交通提供了科学高效解决方案,并为深度学习网络设计提供了新思路。

关键词:智慧交通;融合;计算机视觉;双视角;车型识别;ETCLNet;车辆通行费

中图分类号:U495 **文献标志码:**A

Research on Recognition of Highway Toll Vehicles Based on Dual-View Feature Fusion Network

XIAO Qiang[1,2,3], CHEN Xinjia[1*], CHEN Jing[2,3], ZHAO Bin[2,3], YAN Kaifeng[2], DING Hongwei[1]

(1. School of Information Science and Engineering of Yunnan University, Kunming, Yunnan 650011, China;
2. Yunnan Highway Network Toll Management Co., Ltd., Kunming, Yunnan 650011, China;
3. Yunnan Key Laboratory of Digital Communications, Kunming, Yunnan 650011, China)

Abstract: To solve the practical problems in vehicle type classification on toll roads, a deep learning network based on dual-view feature fusion, named ETCLNet, was proposed in this paper. A dual-branch architecture was adopted by the network to extract the features of vehicle head and side body images, respectively, and combined with an innovatively designed multi-scale feature extraction module, namely IPFE and an adaptive weighting-based bilinear feature fusion (AWBF) mechanism, efficient and fine-grained vehicle feature representation was achieved. Through the design of parallel multi-scale convolutions and residual connections, the problem of gradient vanishing was effectively alleviated, and the adaptability to complex scenarios was enhanced by the IPFE module. The high-order interaction of features between angles was further optimized by the AWBF mechanism, and efficient fusion of multi-angle features was achieved, and classification performance was significantly improved through adaptive weighting and element-wise multiplication. Experimental results indicate that ETCLNet outperforms existing mainstream models in multiple evaluation metrics (e.g., accuracy, precision, recall, F_1 value, and S_{AUC}), which solves the problem of the low vehicle type recognition rate of cameras in existing toll road audit systems. In addition, existing hardware devices are fully utilized by this design, which reduces deployment costs. A scientific and efficient solution for multi-view

收稿日期:2025-02-12 **修回日期:**2026-01-26

基金项目:云南省交通运输厅科技创新及示范项目(编号:2022-27(三));云南省数字交通重点实验室项目(编号:202205AG070008)

作者简介:肖强,男,博士研究生,正高级工程师. E-mail:5923384@qq.com

***通信作者:**陈心嘉,女,硕士研究生. E-mail:2240821337@qq.com

vehicle recognition and smart transportation is provided by this paper, and new ideas for deep learning network design are offered.

Keywords: smart transportation; fusion; computer vision; dual-view; vehicle type recognition; ETCLNet; vehicle toll

0　引言

自2020年1月1日中国取消高速公路省界收费站后，中国高速公路进入了“一张网运行、一体化服务”的新阶段，日均通行量由2019年的2 800万辆次增长到3 200万辆次，2023年春节、国庆长假期间高峰通行量超过6 000万辆次。

“一张网”的形成虽然大大地提高了通行效率，但也带来了管理上的困难，一是中国高速公路统一执行《收费公路车辆通行费车型分类》(JT/T 489—2019)标准后，大客车或重型货车长途通行偷逃费可获取高额利润，驱使偷逃费手段层出不穷；二是通行相关流水数据激增，使得稽查追逃取证困难重重[1]。以云南省为例，云南省收费公路日均存在偷逃通行费的流水约500辆次，年涉及金额上千万。因此，搭建一套准确率高的事后稽查系统尤为重要。

中国统一收费车型后，各车型分类中，客车包括载客汽车和乘用车列车，客车分类依据公安机关交通管理部门机动车注册登记的车辆类型和核定载人数；货车包括载货汽车、货车列车和半挂汽车列车，货车列车和半挂汽车列车按牵引车和挂车合并进行车型分类；货车分类依据车辆总轴数、车长和最大允许总质量确定；具体客车和货车收费车型描述见表1、表2。

表1　客车分类明细表

Table 1　Passenger vehicle classification list

类别	核载人数/人	车辆参数说明
1类客车	≤9	车长小于6 000 mm且核定载人数不大于9人的载客汽车(摩托车通行收费公路按1类客车分类)
2类客车	10～19	车长小于6 000 mm且核定载人数为10～19人的载客汽车
3类客车	≤39	车长不小于6 000 mm且核定载人数不大于39人的载客汽车
4类客车	≥40	车长不小于6 000 mm且核定载人数不小于40人的载客汽车

表2　货车分类明细表

Table 2　Cargo vehicle classification list

类别	总轴数(含悬浮轴)	类别	总轴数(含悬浮轴)
1类货车	2	4类货车	4
2类货车	2	5类货车	5
3类货车	3	6类货车	6

注：车辆参数说明：1类货车车长小于6 000 mm且最大允许总质量小于4 500 kg；2类货车车长不小于6 000 mm或最大允许总质量不小于4 500 kg。

传统的车辆检测器一般采用地感线圈[2]、微波[3]、车辆声纹[4]和压电传感器[5]的方式，对车辆的有无、速度、吨位和车型进行检测。以上车检器的检测方法存在车辆分型不够精确，识别准确率较低、传感器施工工序复杂和维护成本较高等缺点，已无法满足现有收费公路联网收费业务需求。

目前，满足《收费公路车辆通行费车型分类》(JT/T 489—2019)标准，实现收费公路车型进行精确分类的工程技术方案主要有两种：一种是通过激光雷达及其辅助设备对车辆轮廓进行扫描与重建，再从重建后的车辆轮廓数据中，采用计算机视觉的方式提取出用于车型识别的各种特征[6-7]；另一种方案是在路侧架设摄像机，通过拍摄车辆的侧面图片，采用机器学习和计算机视觉技术进行处理，获得车辆的颜色、轴数、车型等特征。

以上两种识别方式各有优缺点，激光雷达检测技术具有检测精度高、抗干扰能力强以及受环境干扰较小等优点，但存在受气候影响大、易受攻击等缺点，按照业务需求，偷逃通行费的证明需有图片作为证明；计算机视觉检测技术具有识别准确性高、获得被测物的信息丰富等优点，但存在需要大量的数据进行训练和测试，数据的采集和标注难度较大及易受光照、姿态、背景等影响的缺点。通过实际工程应用，综合考虑经济性、可维护性及业务特点，本研究采用现有高速公路ETC门架铺设的车牌识别摄像机获取前脸图像，融合侧面摄像机拍摄的图片，运用计算机视觉技术对收费公路车型进行分类识别。该方法提高了检测效率，大幅提升了3类客车和4类客车车型的检测率，减轻了稽核人员的工作强度，具有一定的推广应用价值。

1 研究现状

目前对车辆类型识别及其相关领域的研究主要分为两类:一类是基于传统方法的车型识别;另一类是基于深度学习方法的车型识别。

1.1 基于传统方法的车型识别

基于传统方法的车型识别的主要技术有基于Gabor特征简化的车型识别、基于结构性特征的车型识别等。邓航[8]研究了基于Gabor特征简化的车型识别方法,提出了一种改进的加权稀疏表示的分类算法,试验结果证明平均识别率达到了91.68%;罗亮等[9]为有效识别工程车车型,提出一种改进的HOG特征的自动识别算法,使用RPN自动划分车辆候选区,对每个候选区进行自然度保留的图像增强处理和颜色不变性处理,提取并融合NPE-HOG特征和CIV-HOG特征得到NC-HOG特征,结合一对一支持向量机实现车型自动识别。

基于结构性特征的车型识别技术已在国内外得到广泛研究。其主要优势在于能够有效提取结构信息,从而降低图像的噪声对识别结果的影响,并且可以用于识别具有复杂结构特征的图像。具体而言,研究者通常使用SIFT[10]、SURF[11]、HOG[12]等计算机视觉技术,从图像中提取结构性特征,再使用分类器进行图像分类。例如,采用卷积神经网络提取结构性特征,使用支持向量机作为分类器,结合其他传统方法,如SVM[13]、KNN[14]等。虽然基于结构性特征的车型识别方法能够有效提高车型识别的识别率,但也存在一定问题,例如对于复杂环境下车型识别的适应性和稳定性仍有待提高。因此,基于结构性特征的车型识别算法仍需不断改进和完善。

1.2 基于深度学习方法的车型识别

近年来,基于深度学习的车型识别技术受到越来越多的关注,并取得了大量的研究成果与应用实践,包括基于神经网络的汽车车型识别等。

Yao等[15]提出了一种基于卷积神经网络的车型识别算法。与从车辆的侧视图和前视图中提取车辆外观特征相比,该算法利用车辆后视图像信息进行车型分类的研究相对较少,基于此开发了一种基于车辆后视图像的卷积神经网络车辆类型分类框架;何梓林[16]基于谷歌深度学习框架Tensorflow,研究了深度学习在车型识别领域的应用可能性,得到了可以通过输入图片直接识别汽车类型的卷积神经网络模型,准确率约为92%;为进一步提高智能交通系统对车辆及不同车型识别的泛化性、鲁棒性与实时性,牟亮等[17]提出一种基于梯度压缩的Adam优化算法,进一步提高了YOLO v4算法的训练速度、检测精度以及网络模型的泛化能力。

张念等[18]提出一种基于深度学习的公路货车及其轮轴精细化目标检测方法。采用Retinex理论和加入限制对比度的自适应直方图均衡化等视觉增强方法,预处理所采集图像中的光照不均图像和夜视图像,有效提高了识别精度;封慧杰等[19]针对行驶中车辆的视觉信息易受环境干扰等问题,提出了一种基于ShuffleNet V2的车辆声学特征识别模型S-ShuffleNet,包含对ShuffleNet V2网络深度的缩减以及深度可分离卷积(DSC)中卷积核(Conv)大小的改进。通过提取车辆声信号的对数——梅尔谱图(LMS)特征作为该模型的输入,进行车型识别;徐胜军等[20]针对车辆因姿态、视角等成像差异造成车型难以识别的问题,提出一种基于渐进式多粒度ResNet车型识别网络。以ResNet网络作为主干网络,提出渐进式多粒度局部卷积模块,对不同粒度级别的车辆图像进行局部卷积操作,使网络重构时能够关注到不同粒度级别的车辆局部特征;对多粒度局部特征图利用随机通道丢弃模块进行随机通道丢弃,抑制网络对车辆显著性区域特征的注意力,提高非显著性区域特征的关注度;提出一种渐进式多粒度训练模块,在每个训练步骤中增加分类损失,引导网络提取更具辨别力和多样性的车辆多尺度特征。

近年来,研究人员持续探索YOLO算法在交通运输领域的更深层次应用。江屾[21]在YOLO v5的基础上进行了创新性改进,使其不仅能够识别车辆,更具备了在视频流中对特定车辆进行持续跟踪的能力,这为智能交通监控和管理提供了重要的技术支撑。此外,针对复杂环境下的车辆识别需求,王岱[22]以高性能的YOLOX算法为基础,成功实现了在极端天气条件下的车辆类型识别,显著提升了恶劣环境下的交通感知能力;王开明等[23]针对车型识别任务中存在的问题,包括人为组合忽略了其他部位信息以及多网络融合时价值不一致导致的精度下降等,提出了一种基于FIPA(Foundation for Intelligent Physical Agents)模型的自适应协同优化算法,用于车型识别。在特征提取过程中,采用空间到深度卷积层(SPD-Conv)替代YOLO v8中的标准卷积层,以解决细粒度信息丢失和无效检测问题;设计了一种自适应协同优化网络挖掘携带有效信息的汽车

部件，并解决多智能体（Agent）无序竞争问题；引入基于FIPA模型的加权对数极坐标投票模块，整合近距离和远距离位置上细粒度信息，从而解决融合过程中多Agent价值不一致所导致的精度下降问题；董一兵等[24]巧妙地融合了DETR（Detection Transformer）与YOLO的优势，提出了DETR-YOLO算法，并在智能车联网的复杂背景下实现了对目标车辆的精准识别，进一步提升了算法在复杂场景下的鲁棒性和准确性，为未来智能交通系统的发展注入了新的活力；曾永强等[25]设计了一种基于高速路侧毫米波雷达（Millimeter-Wave Radar，MWR）和相机的多模态融合目标检测与跟踪方法，使用基于中心点的雷达-相机融合算法进行目标检测，并使用贪婪算法进行目标关联；该方法在nuScenes数据集中实现了69.1% AMOTA，并优于基准中所有基于视觉的3D跟踪方法；在自建数据集中验证了适用性以及准确性，每张图像的运行时间为35 ms。

由于现有高速公路中通行的车辆类型千变万化，强监督学习的方法难以有效解决真实场景下的车型识别问题。所以，基于深度卷积神经网络的车型识别算法仍需进一步改进和完善。

为解决以上问题，本文提出一个能够提取相似车型细微特征信息并进行有效融合的双视角车辆识别网络，主要贡献如下：

（1）提出一种基于并行分支与金字塔池化的残差特征提取模块。通过深入提取图片多尺度特征，使网络关注车辆图像的细微特征差异。

（2）提出一种基于自适应加权的双线性融合机制。通过加强对高维特征的关注，使网络达成双角度特征的有效交互，突出关键特征。

（3）提出细粒度双角度交互网络。通过多层级多尺度细微特征提取及高维特征有效交互，网络能够分辨相似车型的细微差别，注意关键特征，达到车辆分类高准确率。

本文采集的收费车辆示意图见图1。

2　模型建立

2.1　ETCLNet网络设计

本文提出的基于双视角特征融合的车辆识别网络总体框架如图2所示，该网络为双分支结构，各分支使用提出的基于多尺度特征提取模块的金字塔池化特征，提取机制IPFE模块细粒度特征，通过基于自适应加权的双线性双角度特征融合机制AWBF特征融合模块合并，得到高效稳定的融合特征。最后，将增强后的特征作为输入送入全连接层。通过Softmax分类层计算最终分类结果。

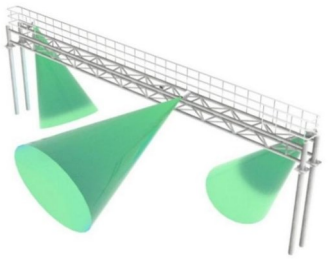

图1　ETC门架车辆图像采集示意图

Figure 1　Schematic diagram of vehicle image acquisition at ETC gantry

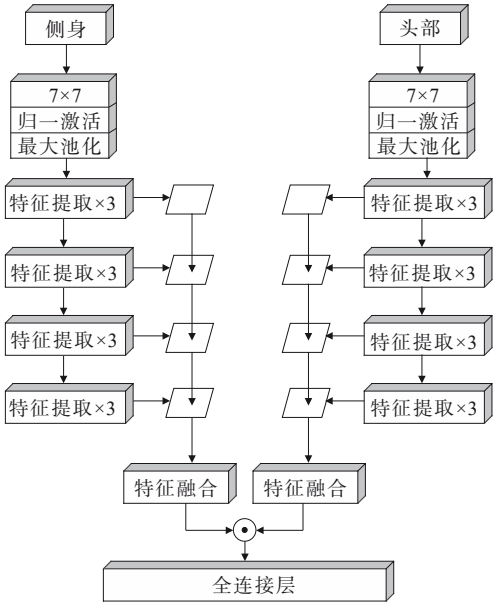

图2　ETCLNet网络结构图

Figure 2　Network architecture diagram of ETCLNet

2.2　IPFE特征提取机制

基于多尺度特征提取模块（Inception）的金字塔池化（Pyramid Pooling Module，PPM）特征提取机制（Inception-Pyramid Feature Extraction，IPFE）模块，采用了与Inception网络相似的并行结构，并在此基础上进行了优化，通过引入残差连接有效缓解梯度消失问题。并为进一步优化特征融合，采用PPM技术，通过对不同尺度的特征图进行池化操作，获得各尺度特征图细粒度特征，最终使用拼接方式生成融合后的输出特征图。这种设计有效提升了网络的稳定性，增强了网络对于细粒度特征的感知能力及对复杂场景的适应能力。

IPFE模块的输入特征图经过4个并行分支处

理，每个分支采用不同的结构设计，旨在实现多尺度特征提取：

第一分支：结合1×1卷积和金字塔池化操作，用于提取全局特征，与其他分支的局部特征提取形成互补。

第二分支：使用两个3×3的卷积核提取较大尺度的车辆特征，用于扩大感受野，并接入金字塔池化，进一步提取该尺度特征图的细粒度特征。

第三分支：结构类似于第二分支，但使用5×5的卷积核后接金字塔池化，进一步增强对更大尺度细微特征的捕获能力。

第四分支：采用两个7×7的卷积核结合金字塔池化，专注于捕获更大感受野范围内的细节特征。

通过这些不同结构的分支设计，IPFE模块可以在不同尺度上提取车辆特征信息，并深入探索各尺度特征图中的细粒度特征，从而提升模型的分类性能。每个分支产生的高维特征图随后在通道维度上进行拼接，实现多尺度特征的整合。

此外，IPFE模块在输入和输出之间引入残差连接，不仅强化了信息流的传递，还显著缓解了梯度消失或爆炸问题。模块具体构成如图3所示。

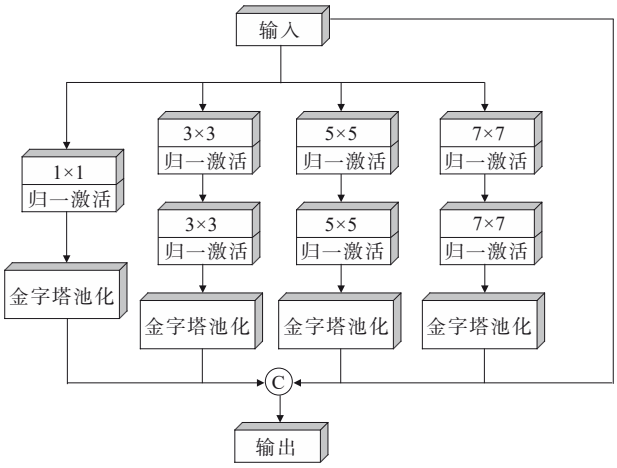

图3 IPFE特征提取机制

Figure 3 IPFE mechanism

2.3 AWBF特征融合机制

AWBF特征融合机制的结构如图4所示。

本文将经过特征提取的头部图像角度特征金字塔表示为$F_h=\{F_{h1},F_{h2},\cdots,F_{hi}\}$，经过特征提取的侧身图像角度特征金字塔表示为$F_s=\{F_{s1},F_{s2},\cdots,F_{si}\}$。模块接收到输入后首先使用Softmax对金字塔中各张特征图进行归一化。以头部角度特征处理为例，初始输入时每张特征图的权重如式(1)所示：

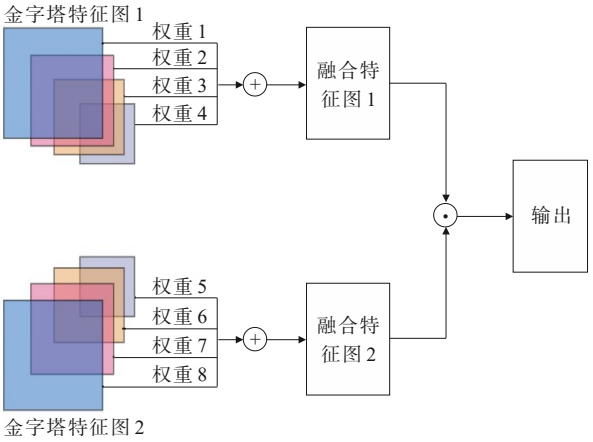

图4 AWBF特征融合机制

Figure 4 AWBF mechanism

$$\boldsymbol{w}_h=[w_{h1},w_{h2},\cdots,w_{hi}]\quad w_i\in R \tag{1}$$

经过Softmax归一化后，得到归一化权重向量如式(2)所示：

$$\boldsymbol{w}'_h=[w'_{h1},w'_{h2},\cdots,w'_{hi}],$$
$$w'_{hi}=\frac{e^{w_{hi}}}{\sum_{i=1}^{N}e^{w_{hi}}}\quad w'_{hi}\in[0,1] \tag{2}$$

为将各尺度特征图融合，使用双线性插值将不同大小特征图上采样至指定的同样尺寸。具体公式如式(3)所示：

$$\hat{F}_{hi}=\mathrm{Interp}\left(F_{hi},H_{\mathrm{target}},W_{\mathrm{target}}\right) \tag{3}$$

式中：H_{target}为指定尺寸的高；W_{target}为指定尺寸的宽。

最终将所有上采样后的特征图按照w'的权重进行加权求和，得到融合后的特征图如式(4)所示：

$$F_{\mathrm{h_fused}}=\sum_{i=1}^{N}w'_{hi}\hat{F}_{hi} \tag{4}$$

经过自适应加权后的头部图像角度和侧身图像角度的总特征图分别表示为$F_{\mathrm{h_fused}}$、$F_{\mathrm{s_fused}}$。双线性融合通过逐元素乘法计算角度间的交互特征，具体融合方程表现如式(5)所示：

$$F_{\mathrm{fused}}=F_{\mathrm{h_fused}}\odot F_{\mathrm{s_fused}} \tag{5}$$

3 模型评价标准

模型的评价指标包括准确率(R_{Acc})、精确率(R_{P})、召回率(R_{R})、F_1值。并采用ROC曲线和S_{AUC}对所得到的网络进行评价。

准确率为模型正确预测的样本数占总样本数的比例，用于衡量总体预测的正确性。在车辆分类中，精确表征了模型的分类正确能力。其公式如式(6)所示：

$$R_{\mathrm{Acc}}=\frac{T_{\mathrm{P}}+T_{\mathrm{N}}}{T_{\mathrm{P}}+T_{\mathrm{N}}+F_{\mathrm{P}}+F_{\mathrm{N}}} \tag{6}$$

式中：T_P(True Positive)为被正确预测为正类的样本数；T_N(True Negative)为被正确预测为负类的样本数；F_P(False Positive)为被错误预测为正类的负类样本数；F_N(False Negative)为被错误预测为负类的正类样本数。

精确率为模型预测为正类的样本中实际为正类的比例，用于衡量模型对正类预测的准确性。在车辆分类中，准确表达了模型对于某一种类别(如1类货车)的分类准确性。其公式如式(7)所示：

$$R_P=\frac{T_P}{T_P+F_P} \tag{7}$$

召回率为实际正类样本中被正确预测为正类样本的比例，用于衡量模型对正类样本的覆盖能力。在车辆分类中，清晰地反映了某一种类型车辆(如1类货车)的漏检情况。其公式如式(8)所示：

$$R_R=\frac{T_P}{T_P+F_N} \tag{8}$$

F_1分数是精确率和召回率的调和平均值，用于综合衡量模型的查准能力和查全能力。在车辆分类中，F_1分数能够平衡精确率和召回率的权重，综合考量了正类的精确率与召回率。其公式如式(9)所示：

$$F_1=2\times\frac{R_P\times R_R}{R_P+R_R} \tag{9}$$

ROC曲线是以假正例率(R_{FPR})为横轴，真正例率(R_{TPR})为纵轴绘制的曲线，用于评估分类模型的区分能力。在车辆分类中，该曲线可准确反映模型在不同分类阈值下的性能变化，当曲线下的面积越大时，模型性能越优。其公式如式(10)所示。

$$R_{TPR}=\frac{T_P}{T_P+F_N},R_{FPR}=\frac{F_P}{F_P+T_N} \tag{10}$$

S_{AUC}表示ROC曲线下的面积，用于定量评估分类模型性能，其值为[0,1]。在车辆分类中，S_{AUC}为衡量模型性能的重要指标。当$S_{AUC}=1$时模型完美区分正负样本；当$S_{AUC}=0.5$时模型与随机猜测无异；当$S_{AUC}<0.5$时模型性能低于随机猜测。

4　试验与结果

4.1　试验数据集

本文使用的数据集包含基于高速公路龙门架摄像头抓拍的车辆正面、侧面图片，还未公开发表，由于现有的公开数据集没有收费公路联网收费车型的分类，所以本文采用现有自有的数据集进行训练和测试，该数据集包含客车4类，货车6类，共10类车辆，图片12 276张，涉及白天、黑夜、阴天等不同环境，各类车辆图片数量如表3所示。

表3　数据集中各类别样本数量

Table 3　Number of samples of each category in dataset

车辆类型	原始样本数/张	处理后样本数/张
1类客车	1 752	1 904
2类客车	172	1 910
3类客车	236	1 906
4类客车	332	1 912
1类货车	1 752	1 904
2类货车	1 942	1 942
3类货车	948	1 914
4类货车	1 850	1 842
5类货车	1 528	1 918
6类货车	1 764	1 918
合计	12 276	19 070

原始数据的不同车辆类型数量不均，且存在无法读取的数据。为使试验顺利进行，删除了无法读取的数据，并进行了数据增强。将图片进行左右镜像、上下镜像、顺时针旋转30°、60°、120°、150°、210°、240°、300°、330°操作。得到图片19 070张，各类车辆图片数量见表3。

4.2　试验环境

该试验基于PYTorch 3.9环境，显卡2080 TX平台实现，使用Adam优化器优化模型性能，batchsize为16，学习率设置为0.000 6，训练轮次为150轮。

4.3　对比试验

为验证所提模型的优越性，对本文所提出的ETCLNet与当前主流分类网络进行性能对比。所有模型在4.2节所规定的试验环境下进行训练和测试，以确保试验结果的公平性。

本文使用多种评价指标(准确率、精确率、召回率、F_1分数和S_{AUC})全面评估模型性能。通过对比分析验证特征金字塔双角度网络在分类准确性和鲁棒性上的提升效果，并展示其在细粒度分类任务中的优势。

本文选取了8种具有代表性的深度学习模型作为基准，全面评估所提出ETCLNet算法在车辆分类任务中的性能。基准模型包括AttentionNet[26]、GoogleNet[27]、Inception v3[28]、MobileNet v2[29]和MobileNet v3[30]、SqueezeNet[31]、Mvit[32]、YOLO v9[33]。试验结果见表4。

由表4可知：ETCLNet在所有评价指标中均取得了最高的成绩。分类准确率达到98.08%，相比性

表4 ETCLNet与其他网络模型试验结果对比

Table 4 Comparison of experimental results between ETCLNet and other network models

方法	R_{Acc}/%	R_P/%	R_R/%	F_1/%	S_{AUC}
AttentionNet	70.62	65.92	66.60	65.52	0.82
GoogleNet	92.08	85.61	87.04	85.75	0.93
Inception v3	93.35	93.07	87.38	88.39	0.96
MobileNet v2	85.28	84.75	83.51	83.52	0.91
MobileNet v3	85.42	86.59	81.08	82.85	0.94
SqueezeNet	75.29	63.23	60.59	60.00	0.79
Mvit	85.12	88.70	79.13	81.72	0.97
YOLO v9	88.15	85.96	82.25	86.97	0.92
ETCLNet	98.08	97.84	96.29	96.95	0.99

能次高的Inception v3 93.35%，ETCLNet提升了4.73个百分点，证明了其在车辆分类任务中的显著优势。精确率为97.84%，高于Inception v3的93.07%，说明ETCLNet在降低误差分类方面更具有优势。召回率为96.29%，相较于Inception v3的87.38%提升了8.91个百分点，表明ETCLNet在识别正确类别方面具有更强大的性能。F_1值为96.95%，作为精确率与召回率的综合衡量指标，ETCLNet的F_1值优于其他模型，进一步证明了其分类性能的平衡性。S_{AUC}值达到了0.99，充分表明ETCLNet对不同类别车辆的区分能力显著优于其他模型。

综上所述，ETCLNet不仅在分类任务中展现了全面的优越性能，还验证了其设计在处理复杂多角度信息和细粒度分类问题上的强大能力。

由于本文引用的算法较多，ROC曲线绘制后堆叠不明晰，在文中使用S_{AUC}进行表达，本文引用的各算法中各车型S_{AUC}值详见表5。

表5 ETCLNet与其他网络模型S_{AUC}试验结果

Table 5 Experimental results of S_{AUC} of ETCLNet and other network models

方法	S_{AUC}值									
	1类客车	2类客车	3类客车	4类客车	1类货车	2类货车	3类货车	4类货车	5类货车	6类货车
AttentionNet	0.95	0.94	0.93	0.92	0.82	0.80	0.51	0.64	0.83	0.83
GoogleNet	0.97	0.83	0.92	0.86	0.99	0.96	0.88	0.97	0.95	0.97
MobileNet v3	0.76	0.79	0.96	0.96	1.00	0.99	0.99	1.00	0.99	0.92
Mvit	1.00	1.00	1.00	1.00	0.99	0.97	0.87	0.91	0.98	0.97
Inception v3	0.99	0.99	0.95	0.98	0.99	0.99	0.98	0.83	0.95	0.95
MobileNet v2	0.97	0.94	0.93	0.92	0.99	0.94	0.73	0.87	0.94	0.86
SqueezeNet	0.86	0.50	0.50	0.93	0.98	0.90	0.73	0.85	0.79	0.85
YOLO v9	0.98	0.90	0.90	0.96	0.96	0.95	0.82	0.91	0.91	0.94
ETCLNet	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00

由表5可知：在经典深度学习分类算法中，对于不同类别的车型分类性能存在较大差异，Inception v3的分类效果较好，S_{AUC}值普遍较高，该模型在试验中对多类别的车型区分能力较强。SqueezeNet的分类效果最差，存在显著的S_{AUC}低值（2类客车、3类客车均为0.50），模型分类的能力鲁棒性不强。

本文所提出的ETCLNet对于各类别的分类性能都十分均衡且分类结果较好，S_{AUC}值高且各类别识别能力均衡，模型的鲁棒性强。

注意力热图是一种对神经网络注意力权重的可视化表示，用颜色强度或亮度来表示不同位置的重要性。在热图中，颜色越亮或越浓，表示模型对该部分的关注度越高；颜色越暗或越浅，表示关注度越低。所有试验网络的注意力热图如图5所示。

由图5可知：本文所提出的ETCLNet在注意力分布上展现了更好的均衡性，能够在头部视角下同时关注车标、车灯和车窗等关键特征，而在侧身视角下能够覆盖车顶、车窗和车轮等区域。设计中的多层特征金字塔（FPN）与自适应加权机制（Adaptive Weighting）有效提升了全局感知能力，使得网络能够提取多尺度、多区域的全局特征。

该网络在两种视角下均保持稳定且高效的特征提取能力，相较于其他受限于设计复杂性和特征覆盖范围的轻量化模型（如SqueezeNet、MobileNet v2和MobileNet v3），在特征表达的丰富性和分布的均衡性上均有显著提升。

4.4 消融试验

为了量化每个模块对整体模型性能的贡献，设计消融试验逐步去除或替换关键模块，使用准确率、F_1分数和S_{AUC}等指标对消融试验结果进行量化分析，揭示每个模块在细粒度分类任务中的作用。通过对比试验结果，证明模块设计的合理性与必要性，为进

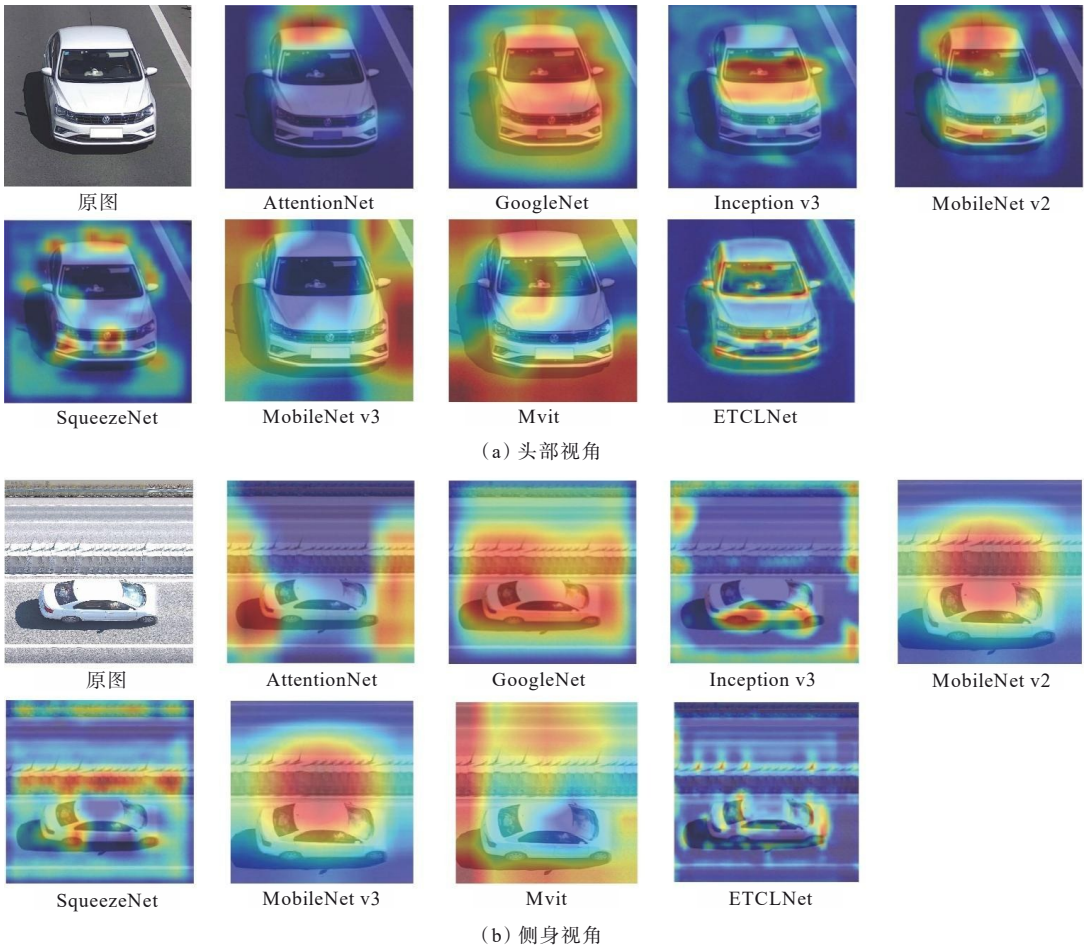

图5　本文所使用试验网络的注意力热图结果

Figure 5　Attention heatmap results of experimental networks used in this paper

一步优化模型提供参考依据。

为验证特征融合方法的有效性，本文将AWBF与常见的特征融合方法进行了对比。如表6所示，Add表示逐元素求和融合，Max表示最大值融合，Concat表示通道维度拼接，而Multiply表示逐元素相乘。

表6　不同特征融合方式消融试验结果

Table 6　Ablation test results of different feature fusion methods

方法	R_{Acc}/%	R_P/%	F_1/%
Add	93.56	94.31	91.54
Concat	89.12	87.56	85.62
Max	90.56	89.46	89.23
Multiply	91.32	89.45	88.41
AWBF	98.08	97.84	96.29

通过不同特征融合模块的对比，在网络其他结构相同的情况下，本文所提出的BRFF特征融合模块能够将高维特征与低维特征有效融合，为正确分类提供了准确有效的信息。

为了验证所提出的IPFE特征提取模块在车辆分类中的有效性，本文开展了消融试验。设计了一组基于AWBF融合模块和GoogleNet的Inception特征提取模块的双分支模型，作为对比试验的基线(Baseline)模型。其试验结果详见表7。

表7　不同特征提取方式消融试验结果

Table 7　Ablation test results of different feature extraction methods

方法	R_{Acc}/%	R_P/%	F_1/%
Baseline	93.73	92.21	89.72
IPFE	98.08	97.84	96.29

经试验证明，基线方法(Baseline)在车辆分类任务中取得了93.73%的分类准确率，精确率达到92.21%，F_1分数为89.72%。这一结果表明，基线方法能够在一定程度上完成车辆分类任务，说明其特征提取机制具有一定的效果。然而，在精确率和F_1分数方面表现相对不足，尤其是F_1分数低于90%，这表明基线方法在处理数据类别不平衡或特征复杂的情况下存在一定的分类不足，难以全面捕获细粒度特征。

通过引入IPFE模块，分类性能得到了显著提升。在分类准确率方面，IPFE模块的准确率从基线方法的93.73%提升至98.08%。这一结果充分说明，IPFE模块能够有效提取多尺度特征，在复杂场景中具有更强的泛化能力。精确率从92.21%提升至97.84%，表明IPFE模块能够更加精准地捕捉目标特征，有效减少错误分类的发生。F_1分数从89.72%提升至96.29%，表明IPFE模块在提升精确率的同时，也显著改善了对不同类别样本的召回能力，使得分类结果更加平衡。

5 结论

（1）本文提出了一种基于双视角融合的深度学习网络——ETCLNet，用于解决收费公路收费车型分类的实际问题。通过引入创新性的网络设计和特征融合机制，ETCLNet展现出卓越的分类性能，解决了现有侧面抓拍摄像机只有侧面图片的车型识别率过低的问题，减少了收费公路投资方的经济损失。

（2）提出了一种新颖的双分支网络架构ETCLNet，该网络创新性地利用车辆的头部和侧身角度进行并行特征提取，充分挖掘不同视角信息的互补性。每个分支均集成了创新设计的基于多尺度特征提取的模块（IPFE），该模块通过并行多尺度卷积和残差连接，有效提取车辆的细粒度特征，缓解梯度消失，并增强网络泛化能力，为后续特征融合提供高质量的特征表示。

（3）提出基于残差的双线性特征融合机制。本文提出的AWBF融合机制，通过双线性融合捕捉头部和侧身角度特征之间的高阶交互，并利用自适应权重和残差连接，增强融合特征的表达能力和鲁棒性。

（4）下一步研究可从扩充车型数据样本（恶劣天气下）、多模态感知融合（与红外图像、毫米波雷达数据或激光雷达点云等）方面入手，进一步优化分类模型的精准度，通过克服极端条件下的数据瓶颈并有效融合多模态信息，将进一步推动AI技术在智能交通领域的深入应用，为构建更安全、更高效的交通系统奠定坚实基础。

参考文献：

References：

[1] 匡永强，胡先全. 基于多数据融合的高速公路收费稽核系统设计[J]. 中国交通信息化, 2022(增刊1): 254-256，278.
KUANG Yongqiang, HU Xianquan. Design of expressway toll audit system based on multi-data fusion[J]. China ITS Journal, 2022(sup 1): 254-256,278.

[2] 史元超. 基于新型传感器的交通状态获取技术的研究[D]. 北京: 北京交通大学, 2009.
SHI Yuanchao. Research on traffic state acquisition technology based on new sensor[D]. Beijing: Beijing Jiaotong University, 2009.

[3] 彭堉斌. 基于FMCW毫米波雷达的车辆轨迹跟踪和车型识别系统[D]. 哈尔滨: 哈尔滨工业大学, 2020.
PENG Yubin. Vehicle trajectory tracking and vehicle identification system based on FMCW millimeter wave radar[D]. Harbin: Harbin Institute of Technology, 2020.

[4] 王君逸. 基于多特征融合的交通车辆声信号识别方法研究[D]. 南昌: 华东交通大学, 2022.
WANG Junyi. Research on traffic vehicle acoustic signal recognition method based on multi-feature fusion[D]. Nanchang: East China Jiaotong University, 2022.

[5] 柳长源, 何先平, 毕晓君. 融合注意力机制的高效率网络车型识别[J]. 浙江大学学报(工学版), 2022, 56(4): 775-782.
LIU Changyuan, HE Xianping, BI Xiaojun. Efficient network vehicle recognition combined with attention mechanism[J]. Journal of Zhejiang University (Engineering Science), 2022, 56(4): 775-782.

[6] 曾夏明, 何家雄, 曹林根, 等. 基于激光雷达车检器的车型识别研究[J]. 中国工程机械学报, 2021, 19(4): 324-330.
ZENG Xiaming, HE Jiaxiong, CAO Lingen, et al. Research on vehicle recognition based on laser radar vehicle inspector[J]. Chinese Journal of Construction Machinery, 2021, 19(4): 324-330.

[7] 郑健龙, 陈梦洁, 刘超超. 公路基础设施智能建造发展与展望综述[J]. 中外公路, 2025, 45(2): 1-20.
ZHENG Jianlong, CHEN Mengjie, LIU Chaochao. Review of intelligent construction development of highway infrastructures and its prospects[J]. Journal of China & Foreign Highway, 2025, 45(2): 1-20.

[8] 邓航. 基于Gabor特征简化的车型识别[D]. 西安: 西安电子科技大学, 2017.
DENG Hang. Vehicle recognition based on Gabor feature simplification[D]. Xi'an: Xidian University, 2017.

[9] 罗亮, 吕俊杰, 李涛, 等. 基于改进的NC-HOG特征的工程车车型自动识别算法[J]. 计算机工程与设计, 2021, 42(11): 3164-3173.
LUO Liang, LYU Junjie, LI Tao, et al. Automatic recognition of engineering vehicle type based on improved NC-HOG feature[J]. Computer Engineering and Design, 2021, 42(11): 3164-3173.

[10] BISHARA A, MAZE E H, MAZE M. Considerations for the implementation of machine learning into acute care settings[J]. British Medical Bulletin, 2022, 141(1): 15-32.

[11] ZHAO P N, LAI L F. Minimax rate optimal adaptive nearest neighbor classification and regression[J]. IEEE Transactions on Information Theory, 2021, 67(5): 3155-3182.

[12] 宋嘉冀. 基于视频的车辆信息提取与检索技术研究[D]. 南京: 东南大学, 2017.
SONG Jiaji. Research on vehicle information extraction and retrieval technology based on video[D]. Nanjing: Southeast University, 2017.

[13] 程淑红, 高许, 周斌. 基于多特征提取和SVM参数优化的车型识别[J]. 计量学报, 2018, 39(3): 348-352.
CHENG Shuhong, GAO Xu, ZHOU Bin. Vehicle recognition based on multi-feature extraction and SVM parameter optimization[J]. Acta Metrologica Sinica, 2018, 39(3): 348-352.

[14] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.

[15] YAO D H, ZHU W X, CHEN Y J, et al. Chinese license plate character recognition based on convolution neural network[C]//Jinan: 2017 Chinese Automation Congress (CAC). 2017: 1547-1552.

[16] 何梓林. 基于Tensorflow的车型识别系统研究[J]. 农业装备与车辆工程, 2022, 60(8): 135-138.
HE Zilin. Research on vehicle recognition system based on Tensorflow[J]. Agricultural Equipment & Vehicle Engineering, 2022, 60(8): 135-138.

[17] 牟亮, 赵红, 李燕, 等. 基于梯度压缩的YOLO v4算法车型识别[J]. 工程科学学报, 2022, 44(5): 940-950.
MU Liang, ZHAO Hong, LI Yan, et al. Vehicle recognition based on gradient compression and YOLO v4 algorithm[J]. Chinese Journal of Engineering, 2022, 44 (5): 940-950.

[18] 张念, 张亮. 基于深度学习的公路货车车型识别[J]. 交通运输工程学报, 2023, 23(1): 267-279.
ZHANG Nian, ZHANG Liang. Type recognition of highway trucks based on deep learning[J]. Journal of Traffic and Transportation Engineering, 2023, 23(1): 267-279.

[19] 封慧杰, 赵红东, 于快快, 等. 基于轻量级卷积神经网络的车辆声学识别[J]. 传感器与微系统, 2024, 43(7): 136-140.
FENG Huijie, ZHAO Hongdong, YU Kuaikuai, et al. Vehicle acoustic recognition based on lightweight CNN [J]. Transducer and Microsystem Technologies, 2024, 43 (7): 136-140.

[20] 徐胜军, 荆扬, 李海涛, 等. 渐进式多粒度ResNet车型识别网络[J]. 光电工程, 2023, 50(7): 36-51.
XU Shengjun, JING Yang, LI Haitao, et al. Progressive multi-granularity ResNet vehicle recognition network[J]. Opto-Electronic Engineering, 2023, 50(7): 36-51.

[21] 江屾. 基于改进YOLO v5的车辆检测及跟踪方法研究[D]. 重庆: 重庆交通大学, 2023.
JIANG Shen. Research on vehicle detection and tracking method based on improved YOLO v5[D]. Chongqing: Chongqing Jiaotong University, 2023.

[22] 王岱. 基于改进YOLOX的恶劣天气车道线与车辆检测识别[D]. 北京: 中国地质大学(北京), 2023.
WANG Dai. Detection and identification of vehicle lanes and vehicles in bad weather based on improved YOLOX [D]. Beijing: China University of Geosciences, 2023.

[23] 王开明, 罗哲皓, 孙文琪, 等. 基于FIPA模型的自适应协同优化的车型识别算法[J]. 计算机应用, 2024, 44(增刊2): 267-273.
WANG Kaiming, LUO Zhehao, SUN Wenqi, et al. Vehicle recognition algorithm based on FIPA model and adaptive collaborative optimization[J]. Journal of Computer Applications, 2024, 44(sup 2): 267-273.

[24] 董一兵, 曾辉, 李建科, 等. DPRT-YOLO: 智能网联汽车复杂驾驶环境实时目标检测器[J]. 计算机工程与应用, 2025, 61(14): 148-162.
DONG Yibing, ZENG Hui, LI Jianke, et al. DPRT-YOLO: Real-time object detector for intelligent and connected vehicles in complex driving environments[J]. Computer Engineering and Applications, 2025, 61(14): 148-162.

[25] 曾永强, 张永华, 赵辉, 等. 高速路侧雷达和相机融合的目标检测与跟踪[J]. 中外公路, 2025, 45(6): 261-267.
ZENG Yongqiang, ZHANG Yonghua, ZHAO Hui, et al. Object detection and tracking based on fusion of radar and camera on highway side[J]. Journal of China & Foreign Highway, 2025, 45(6): 261-267.

[26] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[PP/OL]. arXiv.(2011-06-03)[2025-02-12]. https://arxiv.org/abs/1706.03762.

[27] SZEGEDY C, LIU W, JIA Y Q, et al. Going deeper with convolutions[C]//Boston: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015: 1-9.

[28] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision [C]//Las Vegas: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 2818-2826.

[29] SANDLER M, HOWARD A, ZHU M L, et al. MobileNet v2: Inverted residuals and linear bottlenecks[C]//Salt Lake City：2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition，2018: 4510-4520.

[30] KOONCE B. Convolutional Neural networks with swift for tensorflow: Image recognition and dataset categorization[M]. Berkeley, CA: Apress, 2021.

[31] IANDOLA F N, HAN S, MOSKEWICZ M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size[PP/OL]. arXiv.(2016-02-07)[2025-02-12].https://arxiv.org/abs/1602.07360.

[32] DAI Y Q, ZHENG T, XUE C B, et al. MViT-PCD: A lightweight ViT-based network for Martian surface topographic change detection[J]. IEEE Geoscience and Remote Sensing Letters, 2023, 20: 3000405.

[33] WANG C Y, YEH I H, LIAO H M. YOLO v9: Learning what you want to learn using programmable gradient information[PP/OL]. arXiv. (2024-02-13) [2025-02-12]. https://arxiv.org/abs/2402.13616.