

联合特征胶囊网络在车型精准分类中的应用

周云飞, 余明艳

(广东交通职业技术学院, 广东 清远, 511510)

摘要: 车型精准分类是智能交通领域的研究重点。传统的基于卷积神经网络的车辆分类方法无法表达特征之间的关系, 容易造成误判。为解决上述问题, 引入胶囊网络, 采用不同的卷积核提取不同的特征, 再综合表征它们空间关系的联合特征, 并采用二值交叉熵函数作为损失函数, 加大对错误分类的惩罚。这一方法可用于车型分类, 在训练集和测试集上的分类准确率分别达到 99.8% 和 97.7%, 在某些照明不足的场景中仍可保持较高的准确率。

关键词: 车型精准分类; 卷积神经网络; 胶囊网络; 损失函数; 联合特征

中图分类号: U495; TP391.41

文献标识码: A

Application of the Joint Feature Capsule Network in Precise Classification of Vehicle Types

ZHOU Yunfei, YU Mingyan

(Guangdong Communication Polytechnic, Qingyuan 511510, Guangdong, China)

Abstract: The precise classification of vehicle types is a research focus in the fields of intelligent transportation. Traditional methods based on convolutional neural networks cannot express the mutual relationships among features, and they are prone to cause misjudgments. To solve the above problems, capsule network is introduced, and convolutional kernels with different sizes are used to obtain joint features, expressing the spatial relationships among features. And binary cross-entropy loss function is introduced to increase the penalty for misclassification. The method proposed in this paper can be applied to vehicle classification tasks achieving classification accuracies of 99.8% and 97.7% on the training and test sets, respectively. It maintains high accuracy rate in some scenes with insufficient lighting.

Key words: precise classification of vehicle types; convolutional neural networks; capsule network; loss function; joint features

车型精准分类是智能交通领域的重要研究内容^[1-4], 能提供更准确的车辆信息, 应用于多种场景。比如通过车型分类优化交通流量、分配道路资源, 减少拥堵; 根据车型分类自动收费, 监控不同类型车辆的车速, 提升高速公路通行效率; 在自动驾驶、自动泊车综合调度中提升效率。

传统的车型分类方法主要有支持向量机 (support vector machine, SVM)^[5]、贝叶斯分类法等。为得到更好的分类效果, 一些特征提取算法被引入。局部二值模式^[6] (local binary patterns, LBP) 算子统计图像中像素点及其邻域的灰度比较信息, 提取图像的纹理信息。方向梯度直方图^[7] 算子能够有效地描述车辆局部的梯度

收稿日期: 2025-03-07

作者简介: 周云飞, 博士, 研究方向为计算机视觉技术及应用

基金项目: 2024 年广东省普通高校特色创新类项目“非相干光干涉与高光谱融合技术在薄膜三维形貌及成分测量中的应用研究”(编号: 2024KTSCX379)

纹理特征, 显示不同图像间的特异性, 适合提取车辆轮廓信息。尺度不变特征转换法^[8]采用具有不同标准差的高斯函数对图像进行处理, 在空间中寻找极值点, 并提取出其位置、尺度、旋转不变量, 从而提取出局部特征。这些方法适用于车辆品牌或类型(轿车、货车、越野车等)较粗粒度的识别, 但是对信息的表达能力不够, 对外形相似的车辆难以准确分类。而在交通监控中, 通常只能获得车辆的前部或者侧面部分信息, 可用于分类的外观特征大大减少。再加上光照、分辨率、背景等因素影响, 采集的图片容易丢失关键特征, 造成误判^[9]。

随着深度学习技术在图像分类中广泛应用, 车辆分类技术得到快速发展^[10]。卷积神经网络^[11]能够提取车辆图片的底层特征, VGG^[12]、SSD^[13]、YOLO^[14-15]等模型通过提取高层抽象特征和语义表达不同车型之间的差异, 进一步提升了分类准确率和稳定性^[16]。但卷积神经网络在车型分类中仍存在明显缺陷^[17-18], 影响了分类准确率。主要问题有: ①池化过程中丢失了较多有效信息; ②无法表达特征的空间位置关系和层级关系; ③对背景较暗的车辆图片, 仅能提取车灯区域附近的极少数特征。

为解决上述问题, 胶囊网络^[19]被提出。胶囊网络对卷积神经网络的卷积层、池化层和全连接层进行改进, 通过向量形式的胶囊表示物体的各种特征, 保留了输入目标的姿态、位置、大小、旋转等信息^[20]。而胶囊的向量的分量对应着特定的独立特征, 实现了特征解耦, 并可表示特征间

的位置关系, 在复杂场景的分类任务中表现出更强的可解释性, 提升了模型的表达能力。所以本文采用胶囊网络进行车型分类。

1 相关工作

1.1 胶囊网络

胶囊网络是图像分类领域的前沿技术之一, 早期在手写数字分类任务中表现出优异性。其主要结构如图1所示, 有卷积层(Conv)、基础胶囊层(Primary Caps)、数字胶囊层(Digit Caps)^[20]。网络的输入是一张大小 $28 \text{ p} \times 28 \text{ p}$ 的手写数字灰度图, 首先通过一个卷积层进行灰度特征提取, 得到维度为 $20 \times 20 \times 256$ 的特征图组。然后通过基础胶囊层, 该层由32个胶囊组成。胶囊是胶囊网络的基本单位, 输出一维向量。该向量代表图像的各种属性, 以不同类型的实例化参数表征, 例如速度、形变、色相等。处于底层的基础胶囊层提取图像局部区域的特征, 再将这些特征组合成向量, 于是特征图被封装成特征胶囊。之后, 高层的数字胶囊层将这些特征整合, 动态路由算法对胶囊进行优化处理, 低层胶囊通过将自己的输出乘上权重矩阵(图1中的权重矩阵 W_{ij})来计算预测向量。这种关系通过数字胶囊层输出代表分类结果的10个胶囊, 胶囊向量的欧氏长度即对应该分类的概率。正确类别的胶囊输出向量长度较大, 而错误类别的胶囊输出向量长度较小。通过这种方式, 胶囊网络能够更好地捕捉物体的姿态信息。

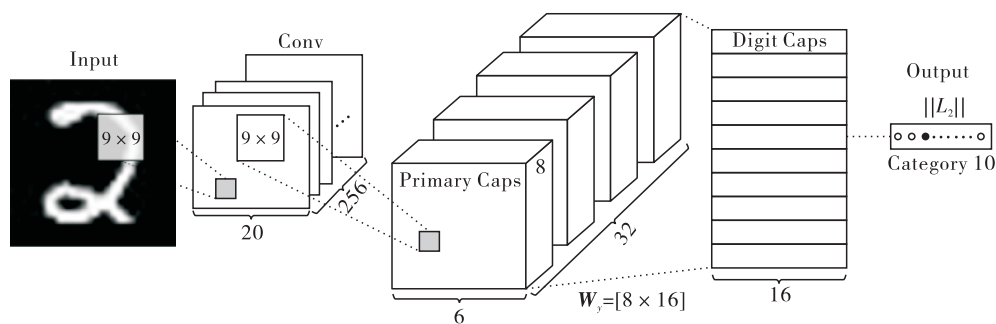


图1 一种胶囊网络的基本结构

1.2 典型问题

假设胶囊网络输出 C 个胶囊, 胶囊网络使用 margin loss (边缘损失) 作为其损失函数, 其表达式如式(1)所示:

$$L_c = T_c \max(0, m^+ - \|V_c\|)^2 + \lambda(1 - T_c) \max(0, \|V_c\| - m^-)^2 \quad (1)$$

其中: T_c 表示第 c 类是否正确, 是为1, 否为0; $\|V_c\|$ 是胶囊输出向量的长度, 即属于第 c 类的概率; $m^+ = 0.9$, $m^- = 0.1$, 分别表示是假阴性和假阳性惩罚; λ 是比例系数, 调整惩罚结果的权重, 通常设置成常数(如0.5)。可见, Margin Loss 对错误类别的惩罚直接受 λ 影响。如

果 λ 过小, 模型对错误类别的惩罚不足, 导致分类性能下降。

2 胶囊网络设计

2.1 模型结构

本设计采用的网络模型如图2所示, 其主体由两个单元组成。

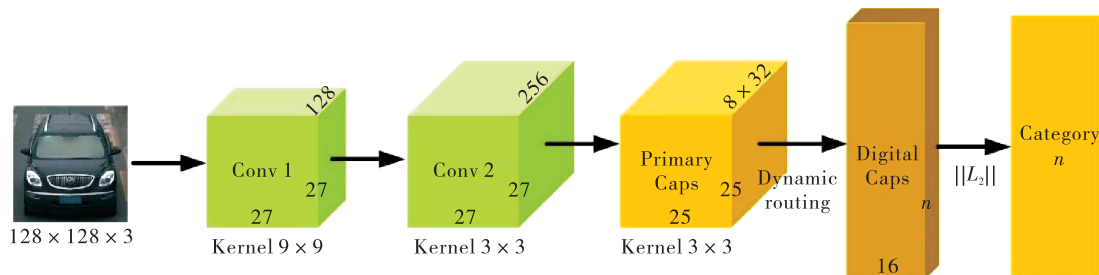


图2 本文采用的基于胶囊网络的车辆精细分类模型结构

为5, 通道数为128, 输出尺寸为 27×27 的特征图。该层提取车辆的宏观特征, 如大小、轮廓等, 步长5保证在准确率不受影响的情况下减少计算量。第二层卷积层的卷积核大小是 3×3 , 步长为1, 通道数为256, 输出为 27×27 的特征图。该层提取车辆的局部特征, 如纹理、结构、局部形状等。两个卷积层分别采用不同大小的卷积核调整感受野大小, 兼顾不同的特征, 可提升车辆分类的准确率。

经两层卷积后的数据被送入胶囊模块。胶囊模块由主胶囊层和数字胶囊层组成。主胶囊层的卷积核大小为 3×3 , 通道数为32, 对输入的特征图进行处理, 得到 $25 \times 25 \times 8$ 的胶囊特征图, 单个胶囊维度为8。数字胶囊层的胶囊个数 n 即为分类数, 胶囊维度为16。主胶囊层与数字胶囊层之间采用全连接, 并引入动态路由算法更新各权值。数字胶囊层的输出是每个胶囊的二范数, 表示车辆属于某类的概率。对每幅图像, 该胶囊网络都分别输出对应 n 种车型的打分值, 分值最高的一类作为分类结果。

2.2 损失函数

二值交叉熵函数通过计算真实标签与预测概率的对数损失, 能够精确地反映模型的预测质量, 适合胶囊网络的分类任务。本文设计的模型采用的损失函数基于二值交叉熵函数, 由式(2)表示:

$$L_{BCE} = -\frac{1}{m} \sum_{i=1}^m [y_i \times \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (2)$$

输入图像为彩色, 具有3个颜色通道。因为数据集的图像尺寸不一, 所以需要经过预处理才能输入模型中进行计算。预处理操作将数据集图片均统一为 $128 \times 128 \times 3$ 的尺寸, 保留车辆主体信息。

预处理后的图像被提取特征, 通过两层卷积实现。第一层卷积层的卷积核大小是 9×9 , 步长

式中, m 是样本数量, i 是样本序号, y_i 和 $\hat{y}_i$ 分别是第 i 个样本被正确和错误分类的概率。与 margin loss 函数相比, 本文所采用的损失函数更简洁, 减小了计算耗时。而且其对错误预测的惩罚力度更大, 当模型预测的概率与真实标签相差较大时, 二值交叉熵函数会输出较大的损失值, 从而在反向传播时产生较大的梯度, 使模型更快地修正错误, 收敛更快更稳定。

3 实验及结果分析

3.1 数据集

实验使用 Comprehensive Cars 数据集^[21], 该数据集包含互联网场景中和监控场景中的汽车图像。网络场景中的图像画质清晰, 分辨率高, 特征明显, 并且添加了标注信息, 甚至包括车门数量、座椅数量等信息。监控场景中的图像画质较低, 分辨率较低, 特征信息未标注。本文研究公共安全监控中的车辆精准分类问题, 所以从 Comprehensive Cars 数据集监控场景包含的281种车辆分类中选取了10种进行实验。其中训练集896张图像, 测试集386张图像。由于输入图像尺寸不一致, 采用截取的图像大小的方法会丢失部分信息, 所以本文采用以图像尺寸变换的方法, 将输入像素数统一为 $128 \text{ p} \times 128 \text{ p}$, 并保留3个颜色通道的信息。实验在 GPU (型号为 GeForce RTX3060)、显存为12GB的台式计算机上运行。

3.2 模型验证

实验对典型的 VGG16、GoogleNet 图像分类

模型^[22]与本文所采用的基于 CapsNet 的模型进行比较。对比实验采用的 VGG16 模型包含 13 层卷积层和 3 层全连接层,卷积核大小均为 3×3 。GoogleNet 前两层为卷积层,第三、四、五层由若干块组成,包含不同大小的卷积核^[23]。图 3、图 4 和图 5 分别是 VGG16、GoogleNet 和 CapsNet 训练及测试的性能对比。

如图 3 所示, VGG16 在训练过程中的准确率在最开始的 50 次迭代中维持在 24.0% 附近,在 50 ~ 150 次迭代区间逐步上升至 80.0%,之后逐步稳定在 83.2%,达到收敛。测试情况与训练情况类似,但存在若干波动,最终收敛于 81.1%。损失值方面, VGG16 在训练过程中一直存在波动,测试时经过 150 次迭代达到收敛。

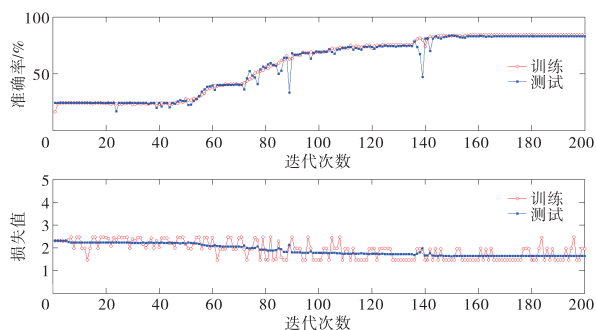


图 3 VGG16 模型训练和测试结果

如图 4 所示, GoogleNet 在训练过程中的准确率在最开始的 5 次迭代中维持在 24.0%,至 100 次迭代逐步上升至 95.5%,至 135 次迭代后稳定在 97.9% 附近,达到收敛。测试情况与训练情况类似,但存在若干波动,最终收敛于 90.1%。在损失值方面, GoogleNet 在训练过程中一直存在波动,测试时经过 145 次迭代后收敛。

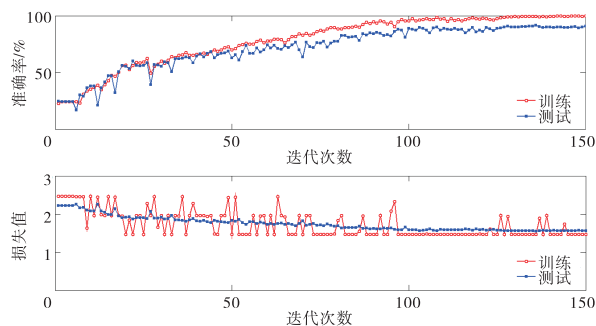


图 4 GoogleNet 模型训练和测试结果

如图 5 所示, CapsNet 在训练过程中的准确率在开始的 4 次迭代中即从 20% 快速地上升至 89.7%,在 8 次迭代后即稳定地维持在 99.8% 附

近,达到收敛。测试情况与训练情况类似,最终收敛于 97.7%。损失值方面,训练和测试过程中均快速稳定地收敛。

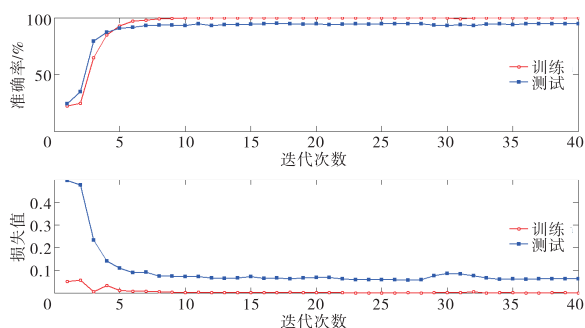


图 5 CapsNet 模型训练和测试结果

由上述比对分析可知, CapsNet 的性能具有一定优势。VGG16 和 GoogleNet 在训练过程中存在明显的波动,这是由于样本在特征方面表现得比较精细。而这两种网络在精细特征提取方面存在一定不足,导致模型在精细样本训练的时候,容易过拟合,损失值存在波动。而测试样本相较于训练样本数量较少,在上述过拟合的情况下,表现相对稳定。

此外,本文还与传统方法进行比较。采用的传统方法先通过 LBP 算法对车辆图像进行特征提取,再用 SVM 进行分类。为验证上述算法的性能,选取部分实验图像进行对比,如图 6 所示。所有图像均来自交通监控场景,每张图像上方标注了其在数据集中的编号,左下角标注了车辆正确分类号。每组图片下方是四种算法的分类结果,红色表示分类错误,绿色表示分类正确。



图 6 四种车辆分类算法的部分结果对比

在光照条件较为理想的情况下,图像对比度较高,车辆特征明显,所有算法都有较好的性能,如图6(a1)、(c1)。但是在条纹干扰的情况下,算法受不同程度的影响,如图6(c3),车辆引擎盖上的树枝倒影对LBP算法的低层特征提取造成影响,使其分类错误。图6(a4)、(b4)、(c4)是夜间采集的图像,车辆全局特征不明显,局部特征主要依靠车灯周围对比度较高的区域,其他算法都出现了错误,而CapsNet能够利用联合特征所表达的局部特征之间的空间关系,所以分类结果正确。在光照强度不足,图像对比度较低时,其他算法效果也明显不如CapsNet。

3.3 对比分析

实验的比较结果如表1所示,其中训练准确率(测试准确率)是在训练集(测试集)上对所有分类结果统计的平均准确率,耗时是对单张图像进行实验的平均耗时。

表1 车辆分类的四种算法性能比较

算法	训练准确率/%	测试准确率/%	耗时/ms
LBP + SVM	78.7	70.6	21.45
VGG16	83.2	81.1	8.39
GoogleNet	97.9	90.1	8.08
CapsNet	99.8	97.7	8.30

LBP + SVM 算法采用的是传统基于图像处理的特征提取算法,需要耗费较多的时间来提取图像统计学特征,耗时明显高于基于人工神经网络的算法。另外,其受图像整体灰度分布以及背景纹理的影响较大,提取的底层特征较为粗糙,容易导致将同一类样本分为多类,且准确率不会随着训练样本数的增加而显著提升。VGG16 采用 CNN 提取图像特征,能够减少图像背景条纹带来的影响,模型最后通过3层全连接层进行抽象特征分类,所以准确率较LBP + SVM 算法有一定提升。GoogleNet 引入 inception 块,在增加网络深度和宽度的同时减少了参数,提升了算法性能,故准确率明显提升。VGG16、GoogleNet 等以卷积神经网络为基本结构的模型虽然节省了统计学图像特征提取的耗时,比基于统计模型算法的速度更快,但是池化操作丢失了图像的结构信息,无法表示整体与局部特征间的空间位置关系。CapsNet 网络用动态路由算法取代卷积神经网络中的池化操作,保存了图像的结构信息,不仅拥有前述所有方法特征提取的能力,而且对局部特征之间的空间关系进行了分析。这种特征间的关系也被作为一种联合特征用于分类任务,提

升了准确率。

4 结论

本文研究设计了一种多尺度联合特征的胶囊网络,并将其用于车型精准分类任务。通过在 Comprehensive Cars 数据集上的对比实验,证明本文设计的模型相比于 LBP + SVM、VGG16、GoogleNet,在分类准确率方面有一定优势。原因在于,LBP + SVM 算法主要提取的是图像统计学特征,在复杂场景中容易受到干扰。VGG16、GoogleNet 等卷积神经网络使用池化操作丢失了图像的部分信息,导致在复杂场景中分类准确率降低。胶囊网络用动态路由算法取代卷积神经网络中的池化操作,兼顾了整体与局部之间的联合特征,作为一种新的特征用于分类任务,提升了准确率。另一方面,本文设计的胶囊网络模型采用基于二值交叉熵的损失函数,加大了对错误分类的惩罚力度,在复杂场景的车型分类中有较好的效果。

参考文献:

- [1]李鹏,杨泽伟,王磊,等.基于视频车型识别的车型库平台[J].中国交通信息化,2024(3):104-107.
- [2]陆文新.南平高速公路收费站车型识别系统应用分析[J].黑龙江交通科技,2024,47(12):155-158,162.
- [3]任文娟,陈建华.基于多属性融合的车辆重识别研究[J].自动化与仪器仪表,2021(12):72-75.
- [4]崔安东.基于深度学习的车辆多特征识别方法研究[D].西安:西安电子科技大学,2021.
- [5]张依林,王学颖.基于CNN-SVM的车辆检测与类型分类[J].电子技术与软件工程,2022(7):190-193.
- [6]潘慧,杨锦润.基于双重细化局部二值模式的纹理图像分类[J].价值工程,2024,43(19):135-138.
- [7]吴逸,吴静静.改进HOG特征的支持向量机工程车辆识别方法[J].计算机与数字工程,2023,51(10):2405-2412.
- [8]董小灵.局部特征的局部敏感哈希专利二值化图像检索[J].电视技术,2022,46(5):54-60,66.
- [9]SHOKRAVI H, SHOKRAVI H, BAKHARY N, et al. A review on vehicle classification and potential use of smart vehicle-assisted techniques [J]. Sensors (Basel), 2020, 20(11):3274.
- [10]YU S, WU Y, LI W, et al. A model for fine-grained vehicle classification based on deep learning [J]. Neurocomputing, 2017, 257: 97-103.
- [11]解丹,陈立潮,曹玲玲,等.基于卷积神经网络的车辆分类与检测技术研究[J].软件工程,2023,26(4):10-13.

(下转第128页)

其次,智慧党建可实现党组织对党员的精细化管理,借助数据分析和人工智能技术,通过对党员的数据收集与分析,及时了解党员的思想动态、理论学习、党建与业务的融合情况等,据此,可针对性地开展主题党日、专题党课、党建共建等活动,从而提高党建策划的针对性、党员管理的精准性。最后,智慧党建系统能够提升党务工作效率,利用其信息传播的即时性特点,提高组织之间、党员之间、组织和党员之间交流的时效性;利用其强大的系统工具,灵活设置学习专栏或模块,利于党的创新理论学习,成果展示,或主题教育的安排等,打造泛在的学习环境;利用其丰富的交互功能,快捷的安排会议、交流学习活动,不仅能实现线上+线下模式的无缝衔接,还能提高集中学习和自主学习的融合度,为学习交流提质增效。

4 结语

智慧党建是传统党建工作和数字技术、智慧化策略融合发展的结果,是高校党建工作发展的必然趋势,是智慧教育发展的必然结果。高校的党建工作和业务发展是相辅相成的关系,党建工作是高校工作的重要组成部分,应把党建工作的智慧化发展同高校智慧教育实践有机融合,以高质量党建引领教育事业高质量发展。泛在学习理论为如何建设智慧党建系统、智慧化开展党建工作提供了重要的参考,有助于加强师生党员的思想

政治教育、党建知识学习,提高党务组织和管理工作的精准化、个体化水平,进一步促进党建工作的发

参考文献:

- [1]刘伟.新时代高校智慧党建的现实问题与对策分析[J].河南教育(高等教育),2023(11):19-21.
- [2]窦毅.新时代高校智慧党建高质量发展路径研究[J].辽宁经济管理干部学院学报,2024(4):82-84.
- [3]武晶晶.数字技术赋能高校智慧党建工作体系建设[J].中国高等教育,2023(12):46-49.
- [4]李鲲鹏.全媒体时代高校“智慧党建”体系构建的三维审视[J].新闻研究导刊,2024(16):180-185.
- [5]肖君,姜冰倩,许贞,等.泛在学习理念下无缝融合学习空间创设及应用[J].现代远程教育研究,2015(6):96.
- [6]孙雪.2006—2019年我国泛在学习研究综述:以中国知网文献刊登情况为例[J].中国现代教育装备,2020(14):34-37.
- [7]祝智庭,管珏琪.教育变革中的技术力量[J].中国电化教育,2014(1):1-9.
- [8]梁颖华,王佳发,傅艳晓.高校“智慧党建”系统建设的实践与思考[J].北京教育(德育),2022(10):24-27.
- [9]黄先开.数字技术在高校基层党建创新的应用分析[J].人民论坛,2021(32):82-84.
- [10]黄艳,成黎明.“互联网+”背景下构建高校党建新常态论析[J].思想理论教育导刊,2021(5):127-31.
- [11]齐勇.高校“智慧党建”信息平台建构实践研究[J].北京教育(高教版),2019(7):112-115.
- [12]ZHANG D, LV J, CHENG Z. An approach focusing on the convolutional layer characteristics of the VGG network for vehicle tracking [J]. IEEE access, 2020, 8: 112827-112839.
- [13]李昭璐.基于迁移学习的交通视频智能解译算法研究[D].西安:长安大学,2022.
- [14]李杨,武海燕,王凡.基于YOLO V5的改装车辆识别技术研究[J].科学技术创新,2024(15):83-86.
- [15]赵立波.基于改进YOLO的车辆及行人识别算法研究[D].南京:南京邮电大学,2023.
- [16]邵萍.复杂背景下细粒度车型识别技术研究[D].大连:大连交通大学,2022.
- [17]陈暄,吴吉义.基于优化卷积神经网络的车辆特征识别算法研究[J].电信科学,2023,39(10):101-111.
- [18]左隆.基于注意力机制与多尺度特征融合的车辆重识别方法研究[D].南昌:南昌大学,2024.
- [19]HINTON G E, KRIZHEVSKY A, WANG S D. Transforming auto-Encoders [C]// International conference on artificial neural networks. Berlin: Springer, 2011:44-51.
- [20]SABOURS, FROSSTN, HINTON G E. Dynamic routing between capsules [C]// Proceeding of the 31st conference on neural information systems. Cambridge, MA:MIT Press, 2017:1-11.
- [21]YANG L, LUO P, LOY C C, et al. A large-scale car dataset for fine-grained categorization and verification [C]// 2015 IEEE conference on computer vision and pattern recognition (CVPR), 2015.
- [22]SEO Y, SHIN K S. Image classification for vehicle type dataset using state-of-the-art convolutional neural network architecture [C]// Proceedings of the 2018 artificial intelligence and cloud computing conference, 2018.
- [23]SZEGEDY C, WEI L, YANGQING J, et al. Going deeper with convolutions [C]// 2015 IEEE conference on computer vision and pattern recognition (CVPR), 2015: 1-9.

(上接第48页)