



东北师范大学
NORTHEAST NORMAL UNIVERSITY

基于可变形注意力机制的车辆重识别 算法设计与实现



汇报人：唐天彪



指导老师：孙俊喜教授



时间：2024年5月8日



东北师范大学
NORTHEAST NORMAL UNIVERSITY

目 录



研究背景与意义



研究内容与方法



实验结果与分析



系统设计与实现



论文总结与展望



研究
背景



研究
意义

车辆重识别是计算机视觉领域中利用交通大数据信息，提取海量道路车辆图像中处于不同时空、不同摄像头拍摄的相同车辆的技术，能够应用于各种交通安全领域。

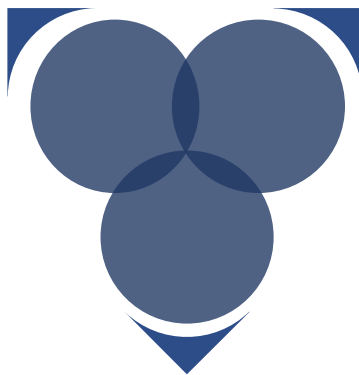
- 推动智慧交通发展建设；
- 保障国民交通安全。



研究内容

● 改进车辆特征表达能力

借助可变形注意力机制优化卷积神经网络的特征学习过程，扩展卷积神经网络的感受野并集中关注图像的显著区域，捕获车辆独特的局部细节，从而增强对车辆个体差异性的表征能力。



● 高效注意力引导的特征筛选

利用可变形注意力动态调整的注意力门控机制对车辆的不同视角、光照条件及遮挡情况下的不变特征进行精准提取，过滤非关键特征，强化对车辆身份鉴别至关重要的特征权重，提高车辆重识别准确度。

● 降低时间复杂度与优化模型性能

相较于传统自注意力机制，可变形注意力机制的一大优势体现在其线性时间复杂度上，预期能够在实际应用中减轻计算负担。在保持或甚至提高识别精度的前提下，将对比分析可变形注意力机制引入车辆重识别系统前后的时间复杂度和运行效率，验证可变形注意力机制在车辆重识别任务中的有效性。



研究方法



聚焦于深度学习架构和可变形注意力机制的融合，提出一种可以集成到 CNN 模型中的关注显著信息的可变形注意力车辆重识别方法DAVReID，由深度可分离模块和变形卷积门控模块组成。深度可分离模块用来加速网络特征提取的效率；变形卷积门控模块用来把网络注意力集中到显著信息上，对重要信息加以强调，不重要信息就过滤掉；最后将原始输入张量和注意力张量进行点乘运算，方便把整个模块整合到 CNN 架构中。



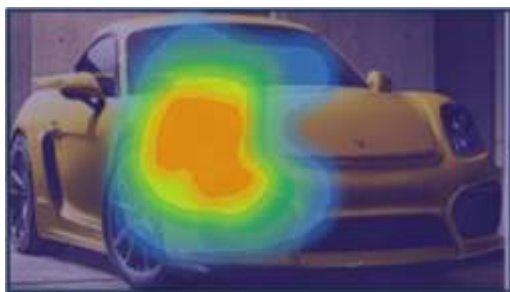
深度可分离模块

深度可分离卷积

充当**瓶颈层**，**减少**输入车辆图像特征映射的通道数量，将**通道数量**从 c 个转换到 $\alpha \times c$ 个，其中 $0 < \alpha < 1$ ，是算法中唯一的**超参数**，其**最佳值**的选取需要平衡网络的准确性和计算效率。

GELU 激活函数

GELU在零点附近具有**非零梯度**，有效地避免了在网络训练时可能发生的**梯度消失**，提高训练过程的稳定性和模型收敛速度，让模型更容易学习和适应车辆重识别任务中的复杂变化。



Step 1



Step 2



Step 3



实例归一化

从输入图像中去除特定于实例的**对比度信息**，最大程度保留特征映射中每个样本的**独特性**，稳定网络内部的信号传播，减轻内部的**协变量偏移**，从而提高训练期间变形卷积门控模块中变形卷积注意力的**鲁棒性**，加快模型的训练**收敛速度**，提高模型的**泛化能力**。

深度可分离模块的压缩过程： $X_c = GELU(InstanceNorm(X \cdot W_1))$ ，其中， X 是输入特征映射， W_1 代表深度可分离卷积， X_c 是深度可分离模块的输出特征映射。



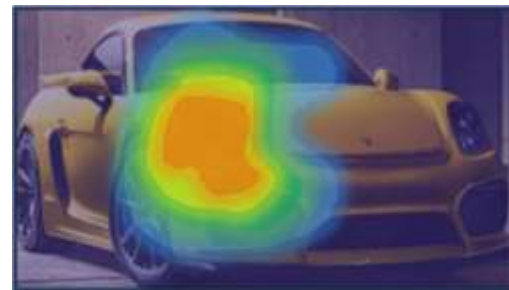
变形卷积门控模块

变形卷积

引入 **偏移量** Δ_p 来构建动态网格，有助于聚焦车辆图像中即使处于卷积核感受野之外的**重要**区域。

Sigmoid 激活函数

门控机制，其输出值介于**0和1之间**，代表**注意力权重**，决定特征被加强或忽略，**通道数**变回 c 个。



Step 1



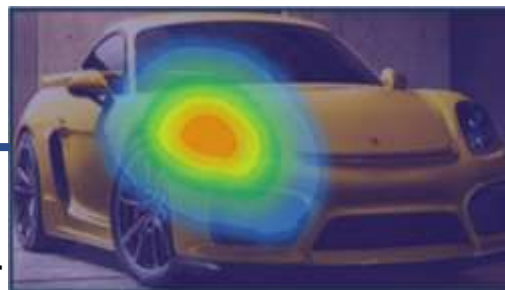
Step 2



Step 3



Step 4



层归一化

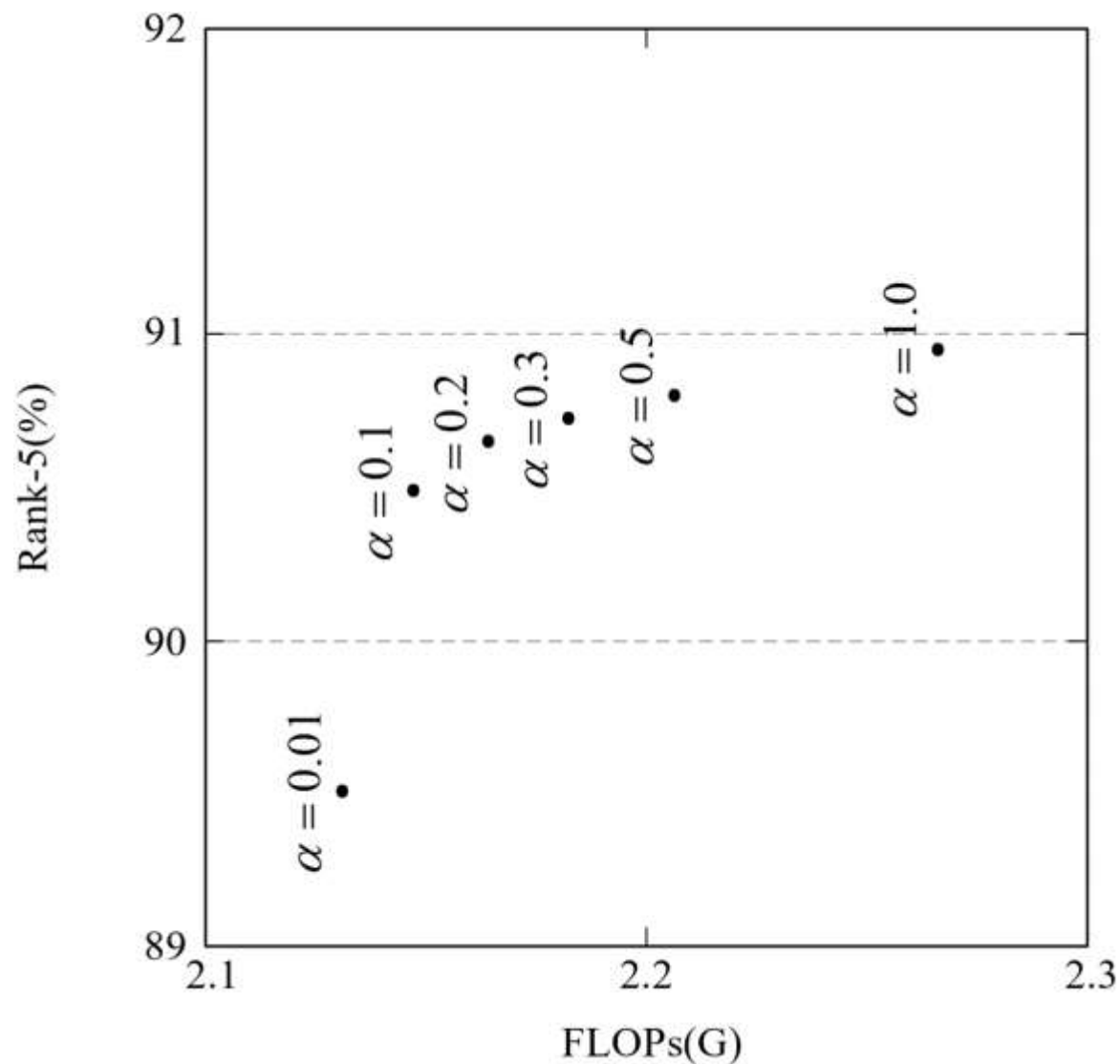
调整变形卷积输出的特征分布，抑制噪声，增强特征表达一致性，利于后续计算的稳定性及模型收敛。

变形卷积门控模块的压缩过程： $X_{out} = X \odot \sigma \left(LayerNorm(deform(X_c)) \right)$ ，其中， X 是原始输入特征映射， X_{out} 是CNN下一层的输入特征映射。



确定超参数 α

在VeRi-776数据集上进行消融研究确定 α 的最佳取值，通过选取不同 α 的取值，绘制FLOPs和Rank-5之间的关系图，选择能使FLOPs相对较低但是Rank-5的值较高的 α 取值。右图展示了FLOPs和Rank-5在不同 α 取值下的关系图，可以看出，当 $\alpha > 0.1$ 时，Rank-5的值对其敏感度较低，最终选取 α 为0.2。





相关方法性能比较

将DAVReID在VeRi-776和VehicleID数据集上取得的结果与其它车辆重识别方法进行比较。RGA、ECANet和CBAM是车辆重识别任务中的三个基于注意力的经典方法，可用于两个数据集的对比实验；MRM、VANet和SAN是近几年在VeRi-776数据集上优秀的车辆重识别方法；PRN、PNVR和SAN是近几年在VehicleID数据集上优秀的车辆重识别方法。下表分别展示了DAVReID在两个数据集上和其它方法的比较结果。

方法	mAP	Rank-1	Rank-5
RGA ^[29]	73.42	92.04	95.87
ECANet ^[30]	71.65	91.50	94.34
CBAM ^[31]	73.05	92.31	95.22
MRM ^[32]	68.37	90.24	93.86
VANet ^[33]	66.72	89.50	93.88
SAN ^[34]	72.54	91.47	94.73
DAVReID	75.46	93.37	96.01

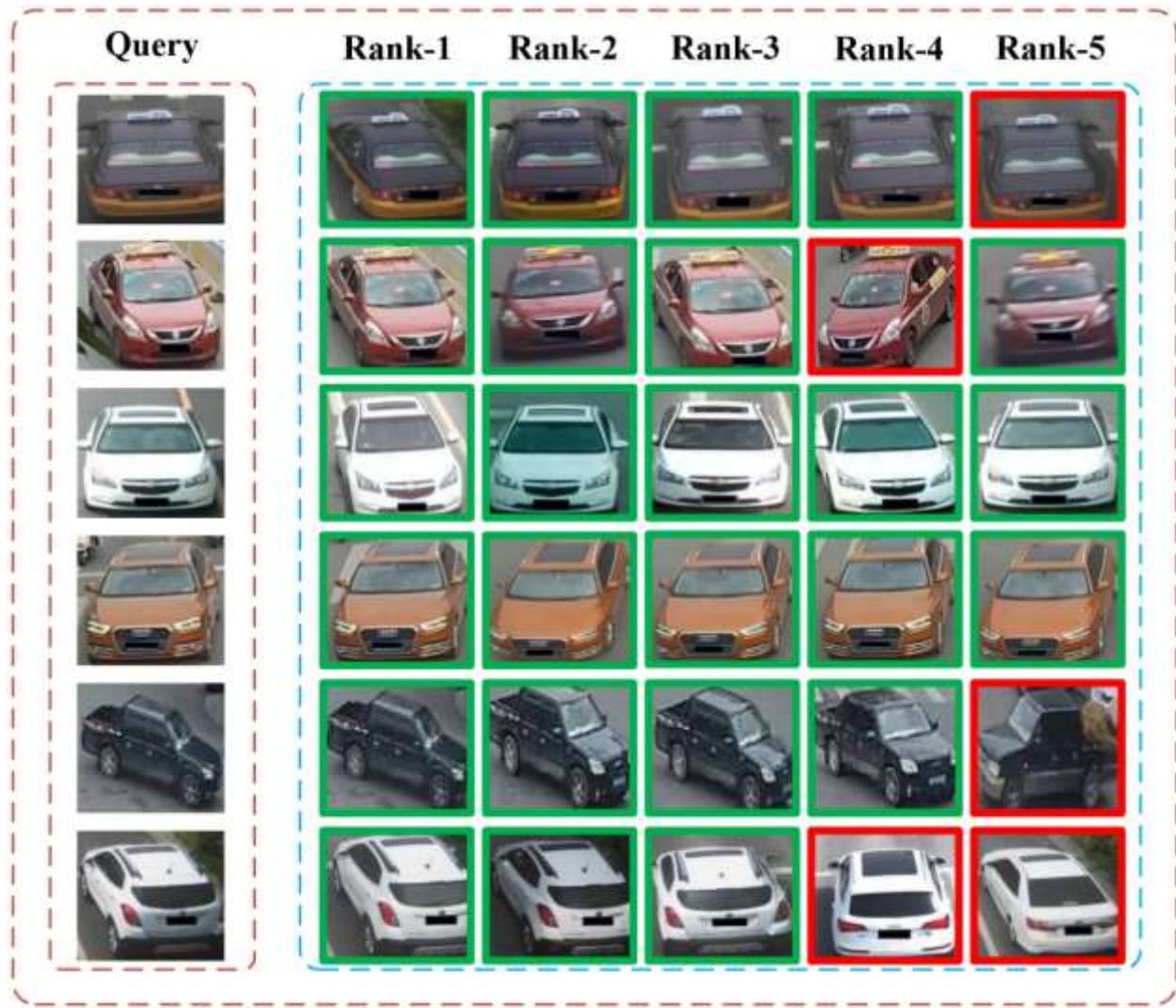
方法	Small		Medium		Large	
	Rank-1	Rank-5	Rank-1	Rank-5	Rank-1	Rank-5
RGA ^[29]	78.57	91.43	74.75	89.24	66.53	85.41
ECANet ^[30]	77.08	85.64	70.54	81.40	65.32	78.44
CBAM ^[31]	72.76	83.66	69.89	80.42	66.39	77.54
PRN ^[35]	77.55	93.21	73.04	91.35	70.82	87.93
PNVR ^[12]	77.43	91.39	74.83	87.37	73.35	85.20
SAN ^[34]	78.63	93.52	77.29	90.53	74.28	86.32
DAVReID	79.37	94.11	73.84	91.59	67.23	84.81



消融研究

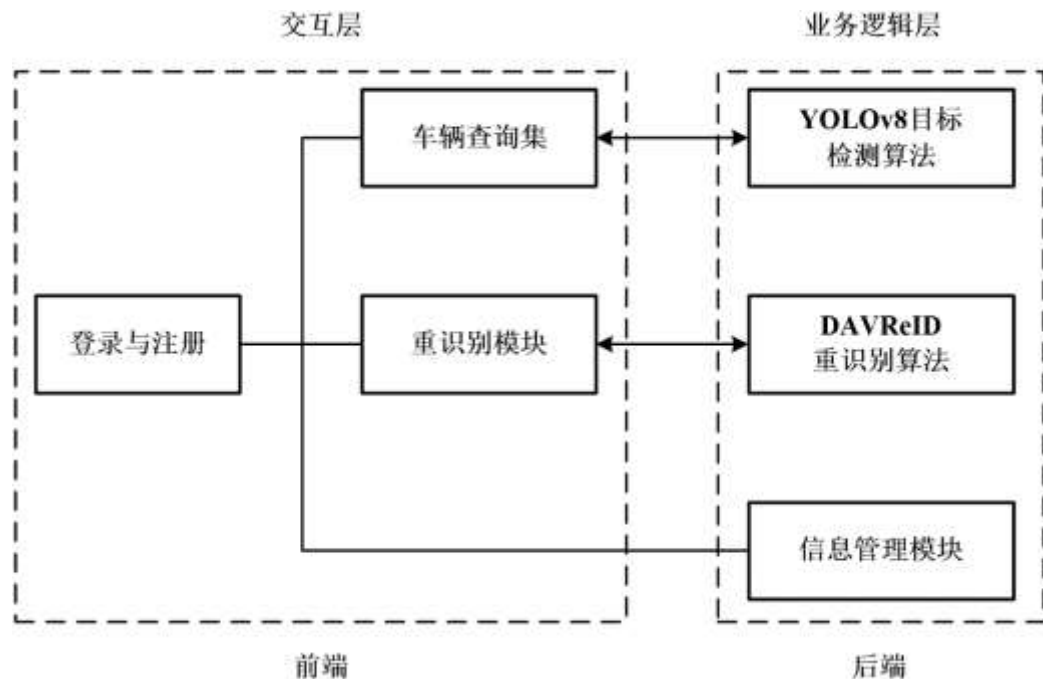
研究DAVReID的不同模块选择对车辆重识别性能的影响。具体来说，主要研究了仅含ResNet-50网络架构、ResNet-50+深度可分离模块、ResNet-50+变形卷积门控模块和ResNet-50+DAVReID四种不同的框架在VeRi-776数据集上对车辆重识别性能的影响，为了方便直观比较出性能的优劣，仅关注mAP和Rank-1指标。下表可以看出，加入DAVReID两个模块的性能表现最佳。

ResNet-50	深度可分离模块	变形卷积门控模块	mAP	Rank-1
✓			73.64	88.34
✓	✓		74.39	90.45
✓		✓	74.31	91.32
✓	✓	✓	75.46	93.37

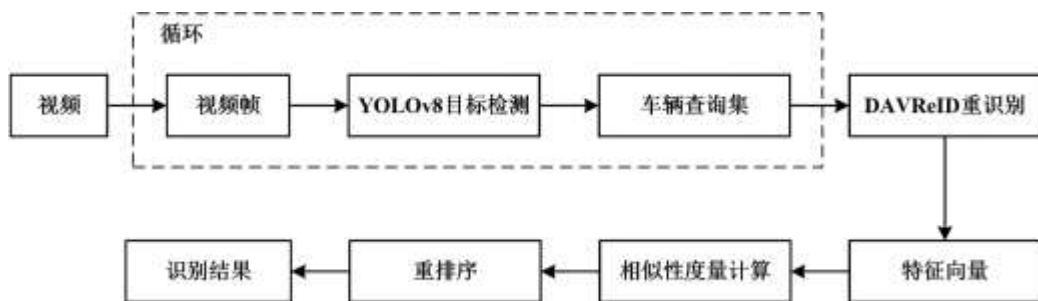




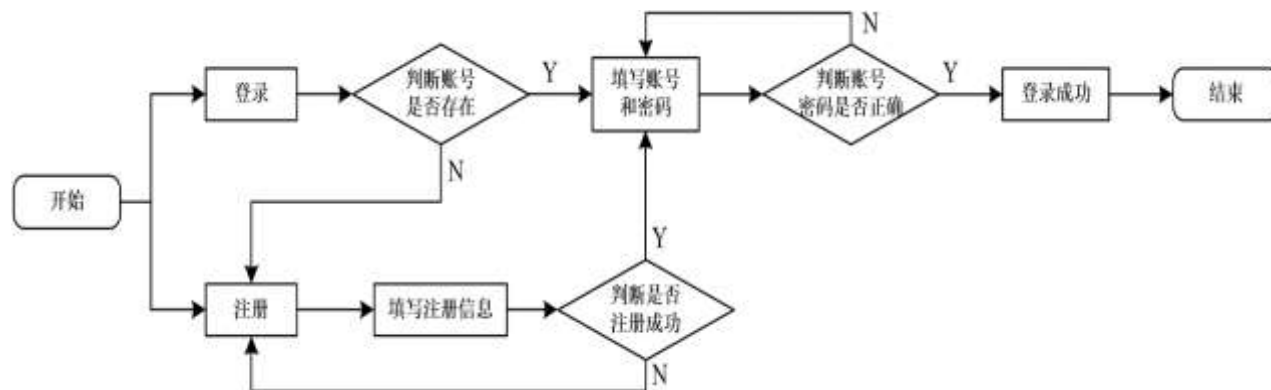
系统框架



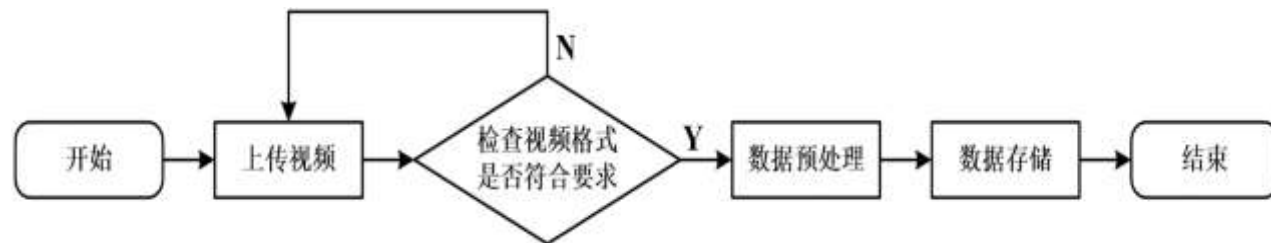
● 车辆重识别模块



● 用户管理模块



● 数据采集模块





系统功能

● 登录页面



● 注册页面



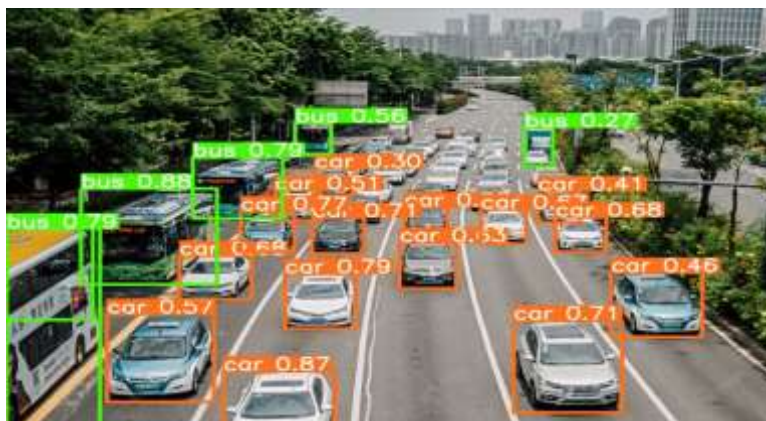
● 系统主页面



● 车辆查询集页面



● YOLOv8定位检测页面



● 车辆重识别页面





系统测试

为了确保该系统的各项功能均能有效运作，针对系统展开一系列**功能性验证测试**。首先设定一系列具有代表性的**测试场景**及对应的**预期输出结果**，随后在实际系统环境中执行这些测试用例。通过比较系统**实际运行**得到的结果与预先设定的**预期结果**是否一致，来评估系统的稳定性和功能性表现。下表是测试用例以及实际运行结果。

编号	测试用例名称	预期运行结果
1	账号或密码输入为空，登录	提示：登录账号或密码不能为空
2	输入不存在的账号或密码，登录	提示：账号或密码错误
3	输入正确的账号和密码，登录	跳转到系统主页面
4	点击注册按钮	跳转到注册页面
5	输入已存在的账号，注册	提示：该账号已存在
6	输入不存在的账号和密码，注册	提示：用户注册成功
7	账号或密码输入为空，注册	提示：注册账号或密码不能为空
8	点击“车辆查询集”按钮	跳转到车辆查询集页面
9	点击“上传数据”按钮	弹出选择文件窗口
10	点击系统主页面目标查询车辆	跳转到该车辆重识别页面

编号	测试用例名称	预期运行结果	实际结果
1	账号或密码输入为空，登录	提示：登录账号或密码不能为空	一致
2	输入不存在的账号或密码，登录	提示：账号或密码错误	一致
3	输入正确的账号和密码，登录	跳转到系统主页面	一致
4	点击注册按钮	跳转到注册页面	一致
5	输入已存在的账号，注册	提示：该账号已存在	一致
6	输入不存在的账号和密码，注册	提示：用户注册成功	一致
7	账号或密码输入为空，注册	提示：注册账号或密码不能为空	一致
8	点击“车辆查询集”按钮	跳转到车辆查询集页面	一致
9	点击“上传数据”按钮	弹出选择文件窗口	一致
10	点击系统主页面目标查询车辆	跳转到该车辆重识别页面	一致



论文总结

针对车辆重识别任务中存在的类内差异性、类间相似性等难题，本文提出了基于可变形注意力机制的车辆重识别算法DAVReID，并设计实现了一个相应的车辆重识别系统。DAVReID算法融合深度学习架构与可变形注意力机制，以提升车辆特征表达、实现高效特征筛选以及优化模型计算效率为目标，具体体现在以下几个技术要点：

(1) 改进车辆特征表达能力：采用可变形注意力机制，有效地扩展了CNN的感受野，使网络能够聚焦于车辆图像中的显著区域，如车灯、车身装饰、车牌等局部细节，从而增强了对车辆个体差异的特征提取与表达能力；

(2) 高效注意力引导的特征筛选：在车辆特征映射阶段，通过可变形注意力机制的动态调整能力，对不同视角、光照条件及遮挡情况下的车辆图像实施精准特征提取，过滤非关键特征，强化对车辆身份识别至关重要的特征权重，提高了跨摄像头场景下的车辆重识别准确度；

(3) 降低时间复杂度与优化模型性能：在确保识别精度不变甚至有所提升的前提下，对比分析了引入可变形注意力机制前后系统的时间复杂度和运行效率。得益于可变形注意力相对于传统自注意力机制的线性时间复杂度优势，DAVReID算法在实际应用中显著降低了计算负担，提高了模型运行效率，有利于实现实时车辆重识别；

(4) 遵循软件工程原则的车辆重识别系统：该系统以DAVReID算法为核心，涵盖了用户管理、数据采集和车辆重识别等功能模块。系统设计充分考虑了车辆重识别应用场景的实际需求，确保了系统的可用性和可维护性。通过对系统进行全面的测试，验证了其基础功能的正常实现，证实了DAVReID算法在智慧交通监控、自动驾驶、跨摄像头追踪等领域的实用价值。



论文展望

尽管DAVReID算法及其配套系统在车辆重识别任务上取得了积极成果，但仍存在若干值得进一步探索与优化的方向：

(1) **适应性与泛化能力提升**：随着新型车辆、个性化装饰、临时遮挡物等不断涌现，模型需要具备更强的适应性以应对更为复杂的现实场景。未来可通过引入自适应学习机制，使模型在持续更新的环境中自我调整，提升识别新类别和异常情况的能力。

(2) **多模态融合与时空信息利用**：结合车辆图像以外的信息，如车辆运动轨迹、交通流量、环境声音等多模态数据，以及拍摄时间、地点等时空信息，构建多源信息融合的车辆重识别模型，有望进一步提升识别精度与鲁棒性。

(3) **轻量化与边缘计算适应**：随着物联网的发展，边缘计算设备在智能交通系统中的应用日益增多。研究模型的轻量化技术和高效推理算法，以适应资源受限的边缘计算环境，是确保车辆重识别技术广泛落地的关键。

(4) **隐私保护与合规性考量**：在大规模车辆重识别系统的设计与实施中，应充分考虑数据隐私保护与法律法规要求，探索安全多方计算、差分隐私等技术，确保在高效识别的同时，严格保护用户隐私，满足监管合规要求。



恳请各位老师批评指正

2024年5月8日



汇报人：唐天彪