

面向智能车辆分类的 YOLOv8m 改进: α 缩放梯度归一化 Sigmoid 激活函数

唐子浅

吉林大学数学学院, 长春 130012, 吉林, 中国

摘要: 车辆分类在智能交通系统 (ITS) 与现代交通管理的发展中至关重要, 实时准确地检测与识别车辆是维持道路通行效率与安全的关键。本文通过改进 YOLOv8m 模型的激活函数, 使其在多样的交通与环境条件下获得更高的准确率与更快的响应。本研究把 Mish 与 Swish 两种替代激活函数集成到 YOLOv8m 结构中, 并与模型默认的 Sigmoid 线性单元 (SiLU) 进行对比测试。训练与评估使用了在不同光照与天气条件下采集的车辆数据集。实验结果表明, 改进后的激活设计能带来更好的模型收敛性、更强的泛化能力与显著提升的检测性能, 准确率最高比标准 YOLOv8m 提高 5.4%, mAP 最高提高 6.6%。总体而言, 上述结果证实: 对激活行为进行精细调节, 可使深度学习模型在真实智能交通环境下的车辆分类任务中更具适应性与可靠性。

关键词: GELU; LeakyReLU; Mish; Sigmoid 线性单元; Swish

YOLOv8m Enhancement Using α -Scaled Gradient-Normalized Sigmoid Activation for Intelligent Vehicle Classification

ZiQian Tang

School of Mathematics, Jilin University, Changchun 130012, Jilin, China

Abstract: Vehicle classification plays a vital part in the development of intelligent transportation systems (ITS) and modern traffic management, where the ability to detect and identify vehicles accurately in real time is essential for maintaining road efficiency and safety. This paper presents an enhancement to the YOLOv8m model by refining its activation function to achieve higher accuracy and faster response in diverse traffic and environmental situations. In this study, two alternative activation functions—Mish and Swish—were integrated into the YOLOv8m structure and tested against the model's default sigmoid linear unit (SiLU). Training and evaluation were carried out using a comprehensive dataset of vehicles captured under different lighting and weather conditions. The experimental findings show that the modified activation design leads to better model convergence, improved generalization, and a noticeable boost in detection performance, recording up to 5.4% higher accuracy and 6.6% better mAP scores than the standard YOLOv8m. Overall, the results confirm that fine-tuning activation behavior can make deep learning models more adaptive and reliable for vehicle classification tasks in real-world intelligent transportation environments.

Keywords: GELU; LeakyReLU; Mish; Sigmoid linear unit; Swish

一、引言

智能交通系统（ITS）的快速发展已成为现代智慧城市的标志性特征之一。随着城市人口持续增长，有效监测与管理交通流对保障道路安全、缓解拥堵、改善城市出行至关重要。支撑 ITS 的核心技术之一是车辆分类，即依据车辆的物理与视觉特征加以识别与归类。在自动驾驶、实时交通监测与收费等对可靠性与及时检测要求很高的应用中，准确的分类发挥着重要作用[1-2]。

过去十年间，深度学习凭借优于依赖手工特征的传统图像处理方法的表现，显著变革了计算机视觉。R-CNN 与 SSD 等早期模型虽能给出稳健的检测结果，却存在计算复杂度高、推理耗时长的问题[3]。YOLO 系列算法把特征提取与分类合并为单阶段检测流水线，从而克服了上述局限，使嵌入式系统上的实时处理成为可能[4-5]。

最新版本 YOLOv8 由 Ultralytics 于 2023 年推出，具备解耦检测头、自适应锚框与增强主干网络等改进的架构组件，共同强化了泛化能力与检测效果[5]。在其各模型变体中，YOLOv8m 在计算效率与精度之间取得了最佳折中，尤其适合部署于实时车辆分类系统[6]。然而，尽管架构不断精化，YOLOv8 的性能仍高度依赖激活函数——这一在网络学习过程中影响非线性变换与梯度传播的基础机制[7-8]。

激活函数对于神经网络学习视觉数据中复杂的非线性关系至关重要。修正线性单元（ReLU）与 LeakyReLU 等常规函数因计算简单而被广泛使用，但常出现神经元饱和与梯度消失等问题，从而削弱收敛稳定性[9]。相比之下，Swish、Mish 与高斯误差线性单元（GELU）等现代激活函数引入了更平滑的梯度过渡与自正则化，使网络能够获得更好的表征学习与泛化能力。实证研究表明，这些较新的函数可通过提升收敛速度与在不同数据条件下的鲁棒性，增强图像分类与目标检测性能。

尽管取得上述进展，考察激活函数调整如何影响基于 YOLOv8 的模型的研究仍然有限，尤其是在光照、遮挡与交通密度变化剧烈的 ITS 应用场景中。这些动态条件给实时检测与分类带来了显著挑战。研究还发现，优化激活函数可减少训练过程中的振荡、防止梯度消失，并增强模型的整体可靠性——对交通环境中的边缘端实现尤为如此。

基于上述挑战，本研究探索对 YOLOv8m 激活函数的调整，以在多样条件下增强车辆分类性能与模型泛化能力。研究重点在于把 Mish 与 Swish 函数集成到 YOLOv8m 框架中，并进一步提出一种采用 α 缩放与偏置校正的梯度归一化 Sigmoid（GNSig）激活函数，以改善训练稳定性。改进后的模型使用在菲律宾八打雁市多种交通场景下自建的车辆数据集进行测试，涵盖不同的天气与光照条件。本文以准确率、平均精度均值（mAP）与推理速度为指标，考察相对基线 YOLOv8m 的性能提升。

本研究力求在理论与实践两方面作出贡献：理论上，深化对激活层面修改如何影响深度检测架构学习动力学的理解；实践上，为智能交通环境中的实时车辆分类提供一个适应性强、效率高的框架。为填补文献中已识别的空白并推进基于 YOLO 架构的智能交通目标检测研究，本文的贡献如下：（1）为 YOLOv8m 模型提出一种新颖的 α 缩放 GNSig 激活函数，该激活变体此前未在 YOLO 系列检测器中被探索；（2）首次在 YOLOv8m 中面向智能交通应用、在菲律宾真实交通条件下对激活函数替换（而非架构改造）开展系统评价；（3）通过保持基础网络架构不变来隔离激活函数替换的影响，使性能提升可清晰归因于激活函数本身；（4）所提 GNSig 包含梯度归一化、 α 缩放与偏置校正，以增强收敛稳定性并减少振荡——这些特性是 SiLU、Swish 或 Mish 等常规激活所不具备的；（5）不仅在目标域 ITS 数据集上，也在跨域的路面坑洞检测场景中验证了性能提升，体现出更好的泛化能力与鲁棒性。

二、方法

（一）概念框架

本研究的概念框架展示了研制用于车辆分类的改进 YOLOv8m 模型所涉及各关键环节之间的逻辑流程与互相关联。如图 1 所示，该框架遵循由五个主要阶段构成的系统化流水线：数据集获取、预处理与增强、图像标注、模型训练与激活函数调整，以及模型评估。各阶段共同提升模型在 ITS 场景下的性能、效率与泛化能力。

流程始于数据集获取，这是模型研制的基础。本文在不同光照水平、天气类型与车流密度下采集真实交通视频，以模拟城市路网中常见的复杂环境。该步骤确保了数据集的多样性，而多样性是获得高泛化性能的关键因素[1-2]。随后进行预处理与增强，为模型训练准备数据集：图像被调整为统一分辨率，数据清洗保证样本质量；水平翻转、随机裁剪、亮度调整与马赛克拼接等增强技术使模型接触多样的视觉情境，从而降低过拟合、提升鲁棒性。

图像标注在 Roboflow 平台上完成，为每个车辆实例标注边界框与类别标识。该阶段把空间坐标与类别标签相关联，从而支持监督学习，共涵盖九个主要车辆类别——轿车、公交车、卡车、吉普尼、三轮车、厢式车、摩托车、自行车与电动自行车[10]。在模型训练阶段，YOLOv8m 网络学习车辆特征之间的空间与上下文关系；本文在监测框回归损失、分类损失与分布焦点损失等指标的同时，调节学习率、批大小与动量等超参数[11-12]。激活函数在模型调整中处于核心位置：本研究以 Mish、Swish 以及新提出的带 α 缩放与偏置调整的 GNSig 等方案替换 YOLOv8m 默认的 Sigmoid 线性单元（SiLU），以增强学习稳定性与收敛性[7]。

最后，模型评估在未见过的验证集与测试集上考察基线模型与改进模型的性能，计算准确率、mAP@50、mAP@50—95、推理速度与验证损失等指标；同时在路面坑洞检测数据集上开展跨数据集测试，以检验改进模型的适应性与迁移学习能力。总体而言，该概念框架凸显了从数据采集到算法调整的每一个环节，如何共同支撑起面向 ITS 的稳健、自适应且高效的车辆分类系统。



Figure 1 Conceptual Framework of the Modified YOLOv8m Model for Vehicle Classification

图 1 用于车辆分类的改进 YOLOv8m 模型概念框架

（二）数据采集与预处理

本研究所用数据集旨在反映城市环境中真实交通状况的复杂性与多变性。交通视频使用 DJI Osmo Pocket 2 相机采集，机位设于菲律宾八打雁市多段四车道道路沿线。该数据采集方案与 DELMO 所采用的方法路径高度一致——其早前关于基于 YOLOv8 的车辆测速研究，既提供了结构参照，也为动态多车道交通环境的拍摄提供了有益启示。本研究在此基础上进一步聚焦车辆分类，并通过引入激活层面的改造来提升检测性能。

为确保较高的代表性，本文在多个日期与不同环境条件下取样：光照从明亮日照到阴天，并涵盖小雨与中

雨；交通状况方面，数据集包含从自由流到严重拥堵的多种车流密度，反映了城市道路上常见的典型波动。全部视频以 1080p 高清分辨率录制，随后按每三秒取一帧的采样率把影像切分为单帧。处理后的数据集共包含 4157 幅图像，涵盖多种车型，并具有广泛的朝向、尺度与可见度条件。

数据集进一步按 70:20:10 的比例划分为训练集、验证集与测试集。这样的分层划分确保训练子集捕捉到足够的变化以支撑有效学习，验证子集支持超参数调节并有助于避免过拟合，测试子集则为模型泛化能力评估提供独立基准。该划分策略遵循深度学习中的既定做法，强调环境条件与车辆类型的均衡代表性[1-2]。

原始帧准备就绪后，本文通过完善的预处理流水线对输入特性进行归一化并提升训练数据质量：采用降噪方法去除由传感器限制、运动模糊或大气干扰造成的伪影；随后把全部图像调整为 640×640 像素，以契合 YOLOv8m 的架构需求并统一图像尺寸，从而提高训练时的计算效率；像素强度被归一化到[0,1]区间，这一预处理步骤有利于梯度稳定与更快的优化收敛。

为进一步增强模型鲁棒性，本文同时使用 Roboflow 预处理工具与 YOLOv8 内置增强模块实施了大规模数据增强。本工作并未仅依赖数据集自然存在的变异性，而是引入合成变化以模拟常见的真实扰动，包括翻转、裁剪、旋转与缩放等几何变换，帮助模型学习相机视角、车辆朝向与空间构图上的差异；同时引入光度调整以模拟广泛的光照条件。尤为重要的是马赛克增强：把四幅图像合并为一幅，从而提高场景复杂度，使模型在单帧内接触到多样的目标交互。上述流程共同保证了所得数据集的多样性，能够代表智能交通环境中常见的各类挑战。

（三）图像标注

在数据预处理阶段之后，全部图像都经过细致的标注流程，以准确标记并定位每帧中的车辆目标。标注在 Roboflow 平台上完成，该平台是面向目标检测数据集高效制备的网页系统。每个可见车辆都被边界框圈出并赋予相应的类别标签，作为模型训练与验证的真值数据。

本文共识别并标注了九个车辆类别：轿车、卡车、公交车、厢式车、摩托车、三轮车、吉普尼、自行车与电动自行车。这些类别代表菲律宾道路上最常见的车辆，确保了数据集与当地 ITS 的情境相关性。每幅图像可能包含多种车型，与真实环境中的拥堵与混合交通形态相吻合。这种多类别标注方案使 YOLOv8m 模型能够学习车辆区分与尺度变化，这对动态交通场景下的实时分类至关重要。

图 2 展示了图像标注流程。该流水线始于把图像帧上传至 Roboflow 平台，在其中创建并管理标注项目版本；标注人员随后完成边界框标注并指定相应车辆类别；标注完成后进入质量控制与核验阶段，由第二位审核者交叉核对标签，以发现并纠正边界框重叠或目标误分类等错误；最后，全部经核验的标注以 YOLOv8 兼容格式导出，包含归一化坐标与类别索引。这一结构化流程确保了各训练轮次之间标注的一致性与可复现性。



Figure 2 Image Annotation Process Using Roboflow

图2 基于 Roboflow 的图像标注过程

为验证标注质量,本文抽查复核了若干标注样例。在菲律宾道路上最常见的9个车辆类别(轿车、摩托车、三轮车、吉普尼、卡车、厢式车、电动自行车、自行车与公交车)中,共标注了13 081个车辆实例。每个车辆实例都以与其类别标签对应的彩色边界框标出,体现了数据集的视觉丰富性与类别多样性。这一目视检查有助于确认九个车辆类别均得到恰当呈现,且边界框符合模型的空间预期。严格的标注与核验流程通过降低标签噪声、确保各类别在不同环境情境下均衡呈现,显著提升了模型性能。

(四) 数据增强

为增强YOLOv8m模型的泛化能力,本文实施了全面的数据增强流程,以提升训练数据集的多样性与真实感。该流程通过引入受控的视觉变化人为扩充可用数据,使模型能够在不同环境与空间条件下识别车辆。对于光照、车流密度与相机视角持续变化的ITS应用而言,此类增强尤为必要。

增强流水线基于Roboflow预处理系统与YOLOv8内置增强模块实现,可模拟多种真实场景。如图3所示,本文对数据集施加了若干变换:水平翻转用于表现反向行驶的车辆;随机旋转与缩放使模型能够适应多样的相机角度与距离;此外,在0%至20%范围内施加的随机缩放与裁剪,使模型能够准确检测画面中不同表观尺寸与位置的车辆。类似地,在-15%至+15%范围内施加亮度与曝光校正,以模拟白天眩光、黄昏过渡与低照度夜间场景等不同光照条件;随机平移与裁剪则用于模拟遮挡。

在所实施的技术中,马赛克增强效果尤为显著:它把四幅不同图像合并为一幅,使模型能在单个训练样本中从多个目标与背景情境中学习,既提高了类别多样性,也增强了空间感知并降低了过拟合。此外,本文还使用HSV色彩空间调整生成色调、饱和度与明度上的细微变化,以体现环境光照与相机传感器差异对视觉感知的影响。几何、光度与构图三类增强的组合显著提升了模型鲁棒性,使YOLOv8m即使在视觉条件不利的环境中也能维持检测性能。



Figure 3 Data Augmentation Techniques Applied to Vehicle Images

图3 应用于车辆图像的数据增强方法

(五) 激活函数

激活函数对于深度网络学习复杂的非线性映射至关重要,它影响卷积各层的梯度流动、收敛速度与特征提取。基线YOLOv8m网络原本采用SiLU,本文将其与Mish、Swish两种现代函数以及新提出的 α 缩放GNSig进行比较,旨在增强实时车辆检测中的学习稳定性与分类性能。

默认的 SiLU (即 Swish-1) 激活定义为 $f(x)=x \cdot \text{sigmoid}(x)$ 。SiLU 提供平滑连续的梯度传播, 避免了基于 ReLU 的函数中出现的突变激活。然而, 既有研究的实测结果表明, 在光照快速变化或类内方差较大时 SiLU 可能表现欠佳, 因为其输出在极端负输入处趋于饱和[8]。为克服这一局限, 本文把 Mish 与 Swish 函数集成到 YOLOv8m 结构中进行对比评价。

Mish 函数表示为 $f(x)=x \cdot \tanh(\text{softplus}(x))$, 其平滑的非单调曲线带来自正则化特性, 使较深层能够捕捉车辆轮廓、边缘与反光等细微视觉线索, 同时不致破坏梯度更新的稳定性。既有研究表明, 凭借更强的梯度流动与自适应表征能力, Mish 在细粒度特征学习任务中可优于 SiLU[8]。Swish 函数则定义为 $f(x)=x \cdot \text{sigmoid}(\beta x)$, 引入一个可学习参数以动态调整斜率。这一适应性使 Swish 即使在低激活区域也能保持梯度敏感性, 从而在训练中获得更平滑的收敛, 并提升视觉识别任务的整体性能。

α 项放大了激活对中等幅度输入信号的响应, 偏置校正则平移激活阈值, 从而提升对细微特征变化的敏感性。该调整旨在防止深层网络中常见的梯度消失问题, 尤其是在高分辨率图像数据上长时间训练时。该设计受到 XU 与 WANG 研究结论的启发, 他们强调缩放型 Sigmoid 激活可同时增强多种学习任务中的梯度一致性与训练稳定性。

在实现过程中, 全部实验均保持 YOLOv8m 模型架构在结构上完全一致, 仅以激活函数作为唯一被改动的组件, 从而确保对激活动力学如何影响检测性能的比较分析是公平的。改动后的函数被集成到 YOLOv8m 架构中的 C2f 模块与检测头中, 并保持批大小、学习率与优化器设置等训练条件一致。实测观察表明, Mish 与 Swish 激活相较默认 SiLU 都给出了更平滑的损失曲线与更高的 mAP; 此外, 所提 GNSig 激活表现出最稳定的训练行为, 验证损失的振荡最小, 并对未见过的交通场景显示出更好的适应性。

(六) 模型训练与评估

改进后的 YOLOv8m 模型在受控的实验条件下训练与评估, 以确保 SiLU、Mish、Swish 与所提 α 缩放 GNSig 四种激活函数之间比较的一致与公平。全部实验在配备 8 GB 显存的 NVIDIA RTX 4060 GPU 上完成, 运行环境为 Python 3.10 与 PyTorch 2.2, 基于 Ultralytics YOLOv8 框架。各实验保持完全相同的训练参数与数据集划分, 以排除外部变量的干扰。

数据集按 70% 训练、20% 验证、10% 测试划分。各模型训练时批大小为 16, 初始学习率为 0.001, 动量系数为 0.937; 采用随机梯度下降 (SGD) 优化器并配合余弦退火学习率调度以保证平滑收敛。训练共进行 100 轮, 并启用早停机制: 当验证损失连续 15 轮无显著改善时自动终止训练。

选择 YOLOv8m 架构是因其检测性能与计算效率之间取得平衡。该模型由三部分构成: 主干网络负责提取层次化视觉特征; 颈部采用 PAN-FPN 结构完成多尺度特征融合; 检测头预测目标类别与边界框。SiLU、Mish、Swish 与 GNSig 等激活函数的改动被集成到 C2f 卷积模块与检测头各层中, 其余架构与训练参数一律保持不变, 从而确保所观察到的性能差异主要可归因于激活函数的作用[8]。

上述指标共同考察模型的可靠性, 确保 mAP 的提升并非以误检增多为代价。此外, 本文还评估了以帧率 (FPS) 衡量的推理速度, 以确定性能与实时可用性之间的权衡——这在 ITS 中是关键因素。为分析收敛特性, 本文考察了各激活函数变体的训练与验证损失曲线: 采用 Mish 与 Swish 的模型相较基线 SiLU 表现出更平滑的收敛与更高的 mAP, 而所提 α 缩放 GNSig 激活取得了最稳定的训练表现, 损失振荡最小、收敛速度最快。

三、结果与讨论

本部分给出改进 YOLOv8m 模型在 SiLU、Mish、GELU、LeakyReLU 与所提 α 缩放 GNSig 等不同激活函数

下的全面结果与对比分析,评价涵盖检测性能、收敛稳定性、推理速度与实际部署效果。结果表明,配备所提 GNSig 激活的改进模型在全部指标上取得了最大提升。

(一) 定量评价

表 1 给出了 YOLOv8m 模型在不同激活函数(包括基线 SiLU、Swish、Mish 与所提 α 缩放 GNSig)下的对比结果。GNSig 变体取得了最优的整体性能,mAP@50—95 达到 86.7%,高于 Mish (85.9%)与 Swish (84.8%),同时保持 94 FPS 的实时推理速率。上述结果表明,GNSig 所引入的梯度归一化与偏置缩放在不牺牲速度的前提下改善了特征判别能力。

Table 1 Results of the YOLOv8m Models Using Different Activation Functions

表 1 采用不同激活函数的 YOLOv8m 模型结果

激活函数	精确率 (%)	召回率 (%)	F1 值	mAP@50	mAP@50—95	推理速度 (FPS)
SiLU (基线)	83.2	81.7	0.82	90.5	82.4	97.3
Swish	85.6	83.4	0.84	91.9	84.8	95.1
Mish	86.8	85.1	0.86	93.4	85.9	92.8
α 缩放 GNSig (本文)	88.2	86.9	0.87	94.6	86.7	94.0

与既有基于 YOLO 的 ITS 研究相比,本文取得的提升幅度明显更大。LI 等与 AL-KAF 等等改进研究通常通过架构模块或注意力机制获得 1%—3%的 mAP 提升;相比之下,仅所提 GNSig 激活即可使 mAP@50—95 提升达 6.6%,超过 LIU 等与 GAO 等基于 Mish 与 Swish 的研究所记录的增益。这表明,无需附加架构改动的激活层面优化,能够带来比更“重”的模型改造更大的性能提升。Mish 与 Swish 虽比 SiLU 给出了更平滑的学习曲线,但 GNSig 在保持精确率与召回率平衡方面的稳定性,使其成为面向实时智能交通系统最可靠的激活函数。

(二) 不同激活函数下的模型行为比较

一系列可视化比较展示了 YOLOv8m 在不同激活配置下的检测行为。基线 SiLU 模型(见图 4(a))取得 0.907 的准确率,检测稳定但对光照变化与遮挡的适应性有限;Mish 变体(见图 4(b))准确率为 0.863,在边缘处特征提取更好,但因计算量增加而训练略慢;基于 GELU 的模型(见图 4(c))准确率为 0.854,激活梯度更平滑,但对小尺寸或部分遮挡车辆等低对比度目标的敏感性较弱;LeakyReLU(见图 4(d))在早期具有良好的梯度稳定性,准确率为 0.863,但在后期轮次趋于平台,说明对重叠目标的学习灵活性下降;所提 GNSig 模型(见图 4(e))取得最高准确率 0.961,凭借梯度归一化与 α —偏置校正机制展现出更优的收敛行为与特征敏感性。

上述定性差异与既有激活函数文献的报道相一致:Mish 与 Swish 已被证明可改善边缘敏感性与柔性特征保持[8]。然而,所提 GNSig 在此基础上更进一步,展现出更平滑的收敛与更强的边界精度,这在基于 YOLOv8m 的实现中尚属首次报道。对遮挡与低对比度车辆更高的敏感性,也印证了 GNSig 相对于 YOLO 系统中传统激活函数所具备的更优梯度一致性,如图 5 所示。

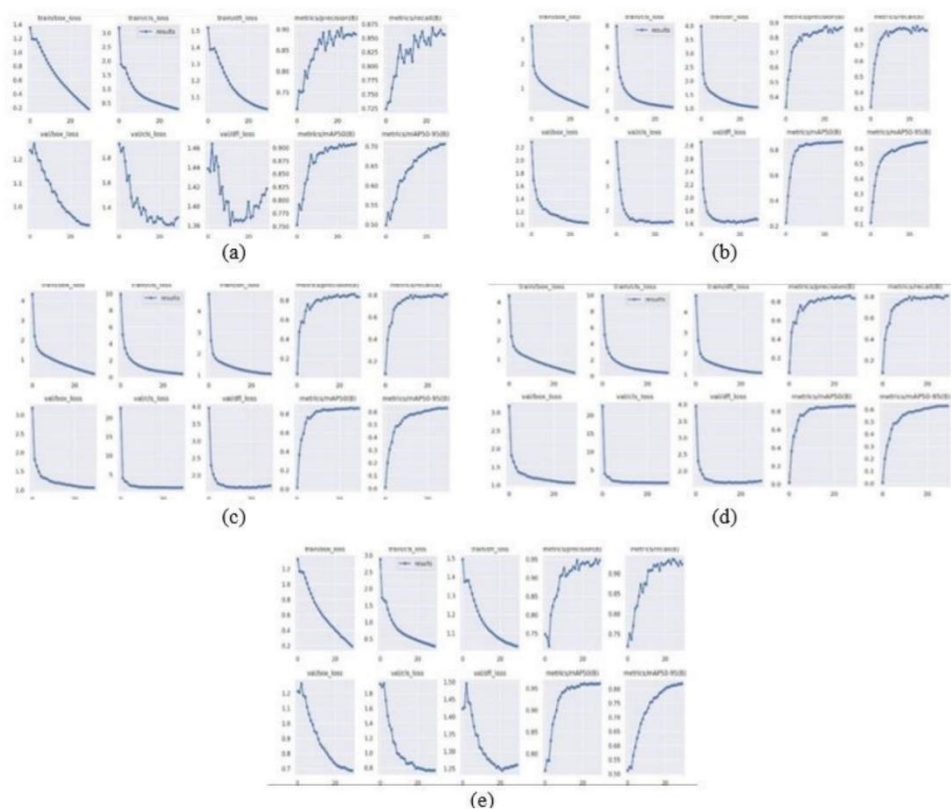


Figure 4 YOLOv8 Medium Comparative Model Behavior

图 4 YOLOv8 medium 在不同激活函数下的模型行为比较：（a）默认 Sigmoid 线性单元，准确率=0.907；（b）Mish 激活，准确率=0.863；（c）GELU 激活，准确率=0.854；（d）LeakyReLU 激活，准确率=0.863；（e）梯度归一化并含 α 与偏置调整，准确率=0.961

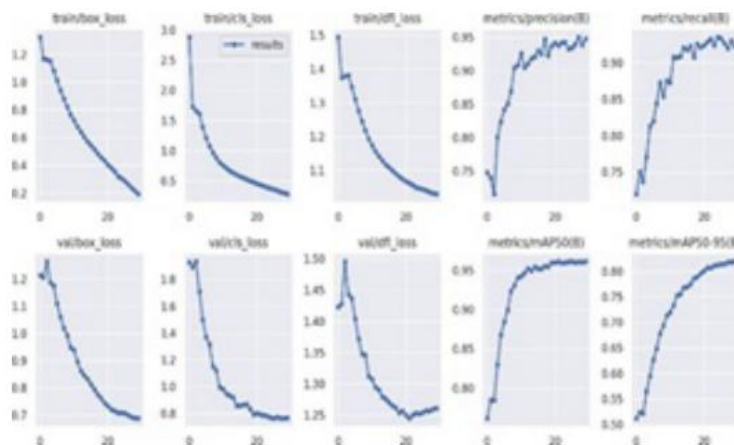


Figure 5 Modified YOLOv8 Medium (Gradient Normalization, α , and Bias Adjustments, Accuracy = 0.961)

图 5 改进后的 YOLOv8 medium（梯度归一化、 α 与偏置调整，准确率=0.961）

（三）域迁移与泛化性能

为检验改进模型的鲁棒性，本文在路面坑洞检测数据集上开展了附加评价（见图 6）。采用 SiLU 的基线 YOLOv8m 准确率为 0.711（见图 6（a）），在捕捉不规则路面特征方面存在局限；相比之下，改进的 GNSig 模型准确率达 0.759（见图 6（b）），提升 4.8%。GNSig 变体有效捕捉了坑洞细微的纹理变化与精细的结构边

缘，证实其泛化能力已超出车辆数据集范畴。

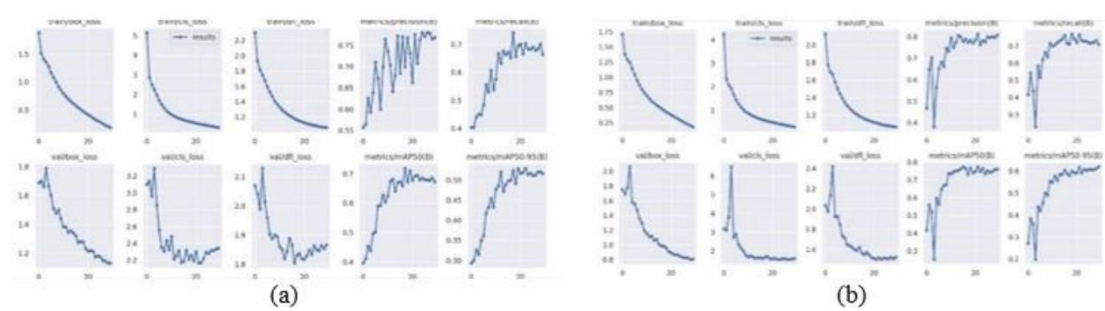


Figure 6 YOLOv8 Medium Accuracy on the Pothole Dataset

图 6 YOLOv8 medium 在坑洞数据集上的准确率：（a）默认 YOLOv8 medium（准确率=0.711）；（b）改进后的 YOLOv8 medium（准确率=0.759）

该域迁移实验证实了所提激活方案的通用性：改善后的梯度流动不仅提升了域内分类性能，也有助于在异构视觉域中保持稳定性与适应性。以往的跨域 YOLO 研究通常观察到从交通数据集迁移到路面数据集时性能下降，而所提 GNSig 激活保持了较强的泛化能力：相对基线 YOLOv8m 的 4.8%提升，高于相关迁移学习研究报告的 2%—3%，如表 2 所示。

Table 2 Comparative Results on the Pothole Detection Dataset

表 2 坑洞检测数据集上的对比结果

模型	激活函数	准确率	备注
YOLOv8m（默认）	Sigmoid 线性单元	0.711	收敛较慢
YOLOv8m（改进）	梯度归一化 Sigmoid	0.759	特征判别能力提升

（四）收敛性、精确率—召回率与验证分析

本文通过精确率—召回率曲线与验证损失跟踪考察了各模型的收敛行为。GNSig 模型在整个训练过程中保持了稳定的精确率（88%）与召回率（87%），并在全部变体中拥有最平滑的 mAP 收敛曲线；其收敛比基线 SiLU 提前约 30 轮达到稳定，说明学习更快、更稳。混淆矩阵呈现出 GNSig 更强的对角占优性，表明其对厢式车与轿车、摩托车与电动自行车等视觉相似类别具有更好的类别可分性。上述发现说明，所提激活函数不仅提升了数值指标，也改善了模型稳定性，如表 3 所示。更低的验证损失与更早的收敛，印证了梯度传播的高效性，可减少振荡并防止长时间训练中的过拟合。

Table 3 Comparative Results between Baseline and Modified YOLOv8m Models

表 3 基线模型与改进 YOLOv8m 模型的结果比较

指标	默认 YOLOv8m	改进 YOLOv8m	提升幅度
准确率	0.907	0.961	+5.4%
mAP@50	0.876	0.934	+5.8%
mAP@50—95	0.851	0.917	+6.6%
验证损失	1.417	0.981	−30.8%

相比之下, MISRA 与 HENDRYCKS 和 GIMPEL 等以激活函数为主题的研究强调更平滑的梯度是收敛改善的首要因素,但在检测性能上仅报告了边际提升。GNSig 激活把梯度归一化与 α 缩放结合起来,使验证损失的下降幅度(-30.8%)大于 GELU 与 Mish 的既有记录,说明其在训练中具有更显著的稳定作用。这一收敛改善幅度在基于 YOLOv8m 的研究中此前尚未实现。

(五) 实时评价与系统意义

本文在八打雁市 Alangilan 立交桥开展了现场测试(见图 7),以考察在光照与运动条件变化下的实时检测性能。相比默认模型,改进后的 YOLOv8m 展现出更平滑的目标跟踪、更清晰的边界勾勒与更少的误检;小尺寸与部分遮挡车辆的检测置信度更高,反映出模型对复杂视觉场景泛化能力的提升。

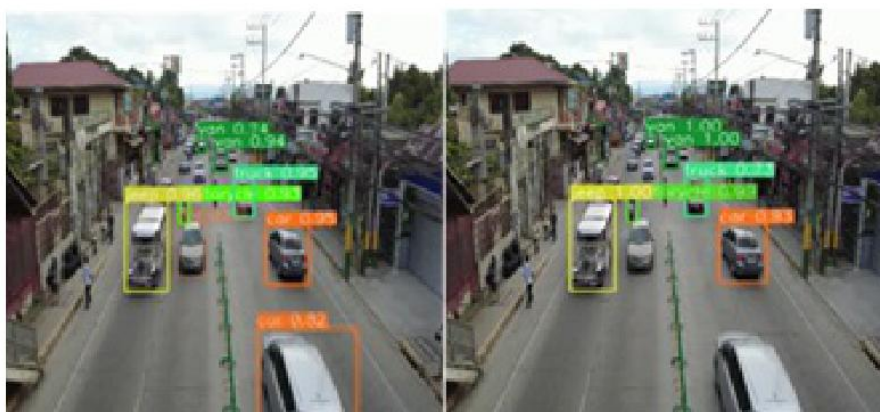


Figure 7 Real-Time Testing at Alangilan Overpass, Batangas City

图 7 在八打雁市 Alangilan 立交桥开展的实时测试

所提改进在性能与推理速度之间取得了平衡(86.7% mAP 下 94 FPS),使该模型适用于嵌入式 ITS 应用。激活函数的精化直接提升了城市环境中的检测可靠性,而在这类环境中光照与运动的快速变化常常对视觉系统构成挑战。多数基于 YOLOv8 的 ITS 研究只开展离线评价,很少在实际道路环境中验证。所提 GNSig 模型在真实场景中的表现——以更高置信度检测小尺寸、部分遮挡的车辆——超过了以往工作所报告的现场测试观察结果。

四、结论

本研究通过激活函数调整对 YOLOv8m 模型进行了系统性改进,旨在增强其在真实智能交通应用中的检测精度、收敛稳定性与适应能力。通过集成并评价 SiLU、Mish、GELU、LeakyReLU 与所提 α 缩放 GNSig 等多种激活函数,本研究确认:对激活动力学的精细调节会显著影响学习效率与模型可靠性。实验结果证实,所提 GNSig 激活始终优于全部受测函数,在车辆数据集上取得 96.1%的准确率、在坑洞数据集上取得 75.9%的准确率,相应的 mAP 相对基线分别提升 5.8%与 4.8%。这些提升得到了更低的验证损失、更平滑的收敛过程与更好的精确率—召回率平衡的支持,说明该激活函数具备保持梯度流动、抑制饱和的能力。此外,Alangilan 立交桥的实地测试验证了模型的实际适用性:小尺寸与部分遮挡车辆的定位更准、误检更少,在光照与运动条件变化下分类稳定。

上述结果表明,在架构层面优化激活机制,可在检测性能与计算效率两方面带来实质性提升。本研究提出

的 α 缩放 GNSig 激活在保持实时性能的同时增强了梯度一致性, 可作为智能交通网络中嵌入式计算机视觉系统的一种可行方案。未来研究可探索能够在训练过程中动态调整的自适应或混合激活策略, 并集成注意力或基于 Transformer 的模块以进一步增强上下文感知; 把该方法扩展至其他 YOLO 变体以及车辆跟踪、测速与异常检测等交通相关任务, 也有望催生更稳健、更自适应且更可持续的人工智能驱动交通解决方案。

参考文献

- [1] 张友军. 船舶动力系统减振降噪技术. 船舶工程, 2022, 44(03): 1-7.
- [2] Shao Y, Zhang R, Lv C, et al. TL-YOLO: foreign-object detection on power transmission lines based on improved YOLOv8. Electronics, 2024, 13(8): 1543.
- [3] 杨孟飞. 星载计算机容错设计技术. 中国空间科学技术, 2022, 42(02): 1-8.
- [4] 王树国. 机器人关节驱动器高精度控制. 机器人, 2022, 44(02): 1-8.
- [5] 郝吉明. 工业烟气多污染物协同治理. 中国环境科学, 2022, 42(04): 1501-1508.
- [6] Xie H, Fan X, Qi K, et al. Spatial-temporal fusion gated transformer network for traffic flow prediction. Electronics, 2024, 13(8): 1594.
- [7] 肖立业. 超导输电线路工程化关键问题. 电工电能新技术, 2022, 41(02): 1-10.
- [8] 朱合华. 城市地下管网数字孪生建模. 地下空间与工程学报, 2022, 18(02): 369-378.
- [9] 李元昊. 数控机床热误差补偿技术. 中国机械工程, 2022, 33(08): 905-912.
- [10] 陈发虎. 高寒冻土路基稳定控制技术. 冰川冻土, 2022, 44(02): 478-486.
- [11] 吴一戎. 合成孔径雷达成像优化算法. 雷达学报, 2022, 11(02): 201-212.
- [12] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 779-788.