

# 基于改进 YOLOv8 的驾驶路面标识检测与识别方法研究

殷翰文, 嵇小辅

(江南大学 物联网工程学院, 江苏 无锡 214000)

**摘要:** 随着自动驾驶技术的迅速发展, 对路面标识的高效检测与准确识别变得尤为关键; 针对路面标志类繁多、形状复杂、同时检测精度易受到天气、光照等环境因素影响的问题, 提出了一种基于改进 YOLOv8 的驾驶路面标识检测与识别的算法模型; 在骨干网络部分新增多项式核引入模块 (PKI), 使其能够从图像数据中提取更加复杂的特征; 在颈部网络部分引入了一种可变形注意力机制 (DA), 使模型能更灵活地关注图像的相关部分, 适应输入数据的特定空间结构; 将原本的损失函数  $CIoU$  替换为  $CIoU_{inner}$ , 增强模型的泛化能力, 并提高检测精度; 经过实验检测: 改进后的 YOLOv8 网络平均精度均值  $mAP: 50$  和  $mAP: 50 \sim 95$  相较原生网络分别提升 4.4% 和 6.7%, 检测速度也达到 71 Hz, 证明了该算法的有效性。

**关键词:** 自动驾驶; 路面标识检测; YOLOv8; 多项式核引入模块; 可变形注意力;  $CIoU_{inner}$

## Research on Driving Road Sign Detection and Recognition Method Based on Improved YOLOv8

YIN Hanwen, JI Xiaofu

(School of Internet of Things, Jiangnan University, Wuxi 214000, China)

**Abstract:** With the rapid development of autonomous driving technologies, it is particularly critical to efficiently detect and accurately recognize road markings. To address the numerous and complex road markings, as well as the detection accuracy easily affected by environmental factors such as weather and lighting, a driving road marking detection and recognition algorithmic model based on an improved YOLOv8 framework is proposed. In the backbone network, a polynomial kernel introduction (PKI) module is incorporated to extract more complex features from image data. In the neck network, a deformable attention (DA) mechanism is introduced to make the model flexibly focus on the related parts of the image, adapting to the specific spatial structures of input data. Additionally, the original  $CIoU$  loss function is replaced with the  $CIoU_{inner}$  to enhance the model's generalization ability and improve detection accuracy. Experimental results demonstrate that compared with the original network, the improved YOLOv8 network improves the  $mAP: 50$  and  $mAP: 50 \sim 95$  by 4.4% and 6.7%, respectively, with a detection speed of 71 Hz, verifying the effectiveness of the proposed algorithm.

**Keywords:** autonomous driving; road marking detection; YOLOv8; polynomial kernel introduction module; deformable attention;  $CIoU_{inner}$

## 0 引言

驾驶路面标识包括交通信号、道路标线及限速标志等, 它们为车辆行驶提供了重要的行驶指引和交通信息。准确检测和识别驾驶路面标识是自动驾驶系统实现

安全、可靠和高效运行的重要基础, 对于自动驾驶技术的普及和应用具有重要意义<sup>[1]</sup>。

随着深度学习技术的快速发展, 基于深度学习的目标检测算法在图像识别、智能监控及自动驾驶等领域得到广泛应用<sup>[2]</sup>, 成为目标检测领域的主流方法之一。基

收稿日期:2024-11-28; 修回日期:2025-01-02。

作者简介:殷翰文(2000-),男,硕士研究生。

嵇小辅(1979-),男,博士,教授。

引用格式:殷翰文,嵇小辅. 基于改进 YOLOv8 的驾驶路面标识检测与识别方法研究[J]. 计算机测量与控制, 2025, 33(12): 254-261.

于深度学习的目标检测算法主要包括两种类型：双阶段目标检测算法和单阶段目标检测算法。双阶段目标检测算法通过选择性搜索等方法生成候选区域，再对候选区域进行分类和位置精修，从而实现目标检测。典型的双阶段目标检测算法包括 R-CNN 系列算法<sup>[3-4]</sup>、M2Det 算法<sup>[5]</sup>等。单阶段目标检测算法将目标检测任务视为一个端到端的回归问题，通过一个神经网络直接从图像中预测目标的类别和位置。典型的单阶段目标检测算法包括 YOLO 系列算法<sup>[6]</sup>、SSD 算法<sup>[7]</sup>、RetinaNet 算法<sup>[8]</sup>等。与双阶段目标检测算法相比，单阶段目标检测算法具有检测速度快的优势，更适合于自动驾驶之类的实时应用场景。

路面标识种类繁多、形状复杂，同时检测精度易受到天气、光照等环境因素的影响，传统的机器视觉难以实现准确的检测和识别。针对以上问题，刘伟<sup>[9]</sup>提出了改进 DeeplabV3+ 网络的检测方法，减少了网络参数，提高了检测精度，但该方法仅实现了单个路面标识目标的识别，难以直接应用到多类型的路标识别。针对单个路面标识目标的识别，韩建鹏<sup>[10]</sup>提出了改进 LeNet-5 模型，提高了检测精度，但实时性欠佳。高怀序<sup>[11]</sup>给出了 YOLOv5 网络模型方法，提高了检测速度和精度。目前，驾驶路面标识的检测与识别方面取得了一些阶段性成果，但是这些方法不能同时满足检测精度和检测速度的需求，难以直接应用到实时驾驶路面标志的检测与识别。

YOLO 是一种基于 PyTorch 框架的单阶段目标检测算法，具有模型复杂度低、检测精确率较高等优点。YOLOv8 作为最新版本，在检测精度、速度、性能方面上都有明显提升，成为实时驾驶场景中路面标识检测和识别的理想选择。本文选择在 YOLOv8 网络模型的基础上进行改进，提出了一种基于改进 YOLOv8 的驾驶路面标识检测与识别方法：引入模块（PKI, poly Kernel inception module）融合骨干网络，使模型能够从图像数据中提取更加复杂的特征；引入了一种可变形注意力机制（DA, deformable attention），使模型能更灵活地关注图像的相关部分，适应输入数据的特定空间结构将原本的损失函数  $CIoU$  替换为  $CIoU_{inner}$ ，通过辅助边界框计算  $IoU$  损失，增强模型的泛化能力，并提高检测精度；实验结果表明：改进后的 YOLOv8 网络在路面标识上的检测性能得到了大幅优化，对复杂路面场景也能进行精准的检测识别，能够满足精测精度和实时性的双重需求。

## 1 YOLOv8 网络结构与原理

### 1.1 YOLOv8 网络模型介绍

YOLOv8 是 YOLO 系列的最新版本，其网络结构主要可以分为输入层、背景特征提取（Backbone）、特

征金字塔网络（FPN）及检测头（Head），其网络结构如图 1 所示。

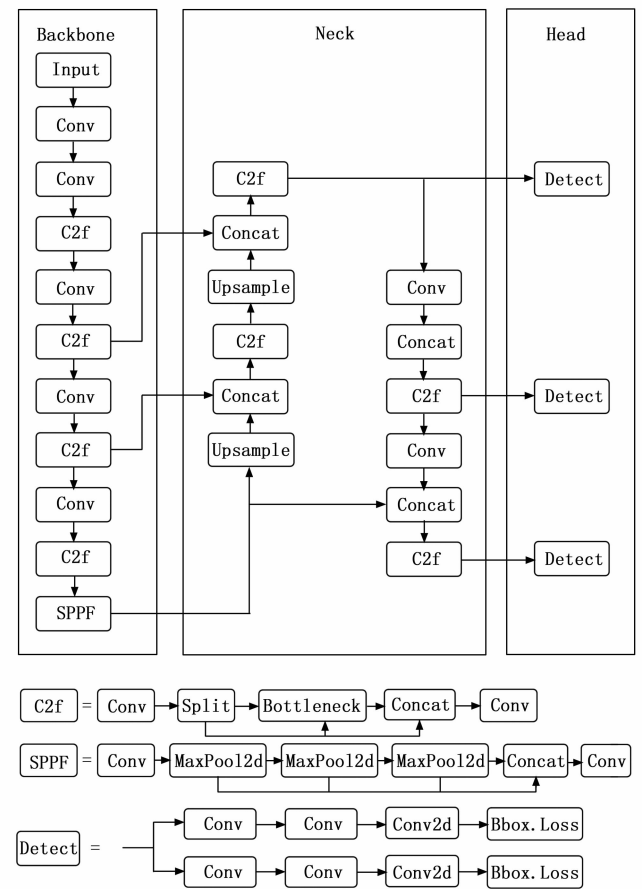


图 1 YOLOv8 网络模型结构图

### 1.2 YOLOv8 网络结构分析

输入层主要负责接收和预处理输入的图像，通常包括图像缩放和归一化等操作，输入图像的尺寸会被调整为固定大小，以适应网络结构的要求；Backbone 负责提取多尺度的图像特征，生成特征金字塔；FPN 是 YOLOv8 中重要的组成部分，它用于融合不同尺度的特征。通过融合来自不同层的特征图，FPN 能够更好地捕捉图像中的多尺度信息，特别是对小目标的检测更加准确<sup>[12]</sup>；YOLOv8 的检测头负责最终的对象检测和分类。

另外，YOLOv8 采用了新的激活函数，如 Mish<sup>[13]</sup>或 Swish<sup>[14]</sup>，相比传统的 ReLU 函数，这些激活函数在提升模型的非线性表示能力方面表现更好，从而提高了整体性能；同时，YOLO-v8 使用了  $IoU$  损失函数来训练模型，这种多任务损失函数帮助模型在不同任务上都能取得良好的表现<sup>[15]</sup>。

## 2 改进措施

### 2.1 改进后的网络模型

原生的 YOLOv8 网络在进行路面标识的检测与识

别中,发现在处理遮挡、光照变化方面的检测表现存在不足:难以充分适应图像的变化;面对复杂场景和多类型目标时也无法捕捉到足够多的细节,限制了其在实际应用中的效果。为了提升 YOLOv8 目标检测算法在复杂场景下对驾驶路面标识的检测与识别上的性能,本文从模型特征提取模块、注意力机制以及损失函数方面进行了改进。

如图 2 所示,本文对 YOLOv8 网络进行了以下几点改进:

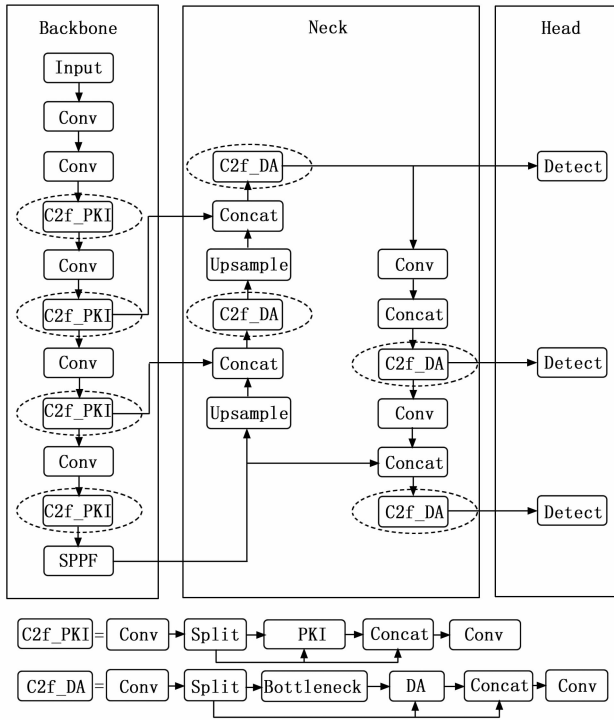


图 2 改进后的 YOLOv8 网络模型结构图

1) 由于路面标识种类繁多、形状复杂,且在路面上分布不均,原生的 YOLOv8 网络在检测中易出现漏检误检,所以在 Backbone 方面新增了多项式核引入模块 PKI 来优化骨干网络。Backbone 网络中的 C2f 模块用于特征提取和融合,如图 2 所示,将 PKI 引入 C2f 模块中,使得 C2f 中原有的 Bottleneck 层被 PKI 替代。PKI 模块采用不同大小的并行深度卷积核,能够有效地捕获不同尺度上的密集纹理特征,从图像数据中提取出更加丰富和详细的特征,使模型能更加精确地进行检测识别。

2) 面对复杂的道路场景以及天气、光照等环境因素的影响,原生的 YOLOv8 网络常常无法准确地识别路面标识,因此在颈部网络引入了 DA 注意力机制,使模型能更灵活地关注图像的相关部分,适应输入数据的特定空间结构。如图 2 所示,将 DA 注意力模块作为 C2f 中 Bottleneck 的增强模块融入其中,替换了 Bottle-

neck 中的部分卷积操作。DA 注意力模块可以动态地调整注意力模型的形状和大小,以更好地适应不同任务和输入数据的特点,从而提高对路面标识的检测和识别能力。

3) 为了应对样本易混淆类别的问题,在损失函数方面,用  $CIoU_{inner}$  替换了原本的  $CIoU$ ,新的损失函数通过生成不同尺度的辅助边界框以计算损失使模型获得了更好的泛化能力,从而获得最佳的训练结果,进一步提升检测精度、速度和稳定性。

通过以上多方面的改进,YOLOv8 算法在检测精度和速度以及泛化能力上均得到了显著提升。

## 2.2 PKI

与一般的目标检测不同,路面标识的识别与检测在单个图像内定位和识别不同大小的对象。为了应对驾驶路面标识的检测与识别中复杂的道路场景所带来的挑战,本文采用了 PKI 模块作为特征提取模块。PKI 模块并行地排列了多个不同大小的无空洞的深度卷积核,并在不同的感受野中提取密集的纹理特征。这些纹理特征沿通道维度自适应地融合,从而收集目标的局部信息<sup>[16]</sup>。PKI 模块通过一个小核的卷积来捕获局部信息,然后利用一系列并行的深度卷积来跨多个尺度捕获上下文信息。

PKI 模块可以用公式表示如下:

$$L_{l-1,n} = \text{Conv}_{k_s \times k_s}(X_{l-1,n}^{(2)}), n = 0, \dots, N_l - 1, \\ Z_{l-1,n}^{(m)} = \text{DWConv}_{k^{(m)} \times k^{(m)}}(L_{l-1,n}), m = 1, \dots, 4 \quad (1)$$

$L_{l-1,n} \in R^{1/2 C_l \times H_l \times W_l}$  是通过  $k_s \times k_s$  卷积提取的局部特征,而  $Z_{l-1,n}^{(m)} \in R^{1/2 C_l \times H_l \times W_l}$  是由第  $m$  个  $k^{(m)} \times k^{(m)}$  深度卷积 (DWConv) 提取的上下文特征。本文中设置  $k_s = 3$ ,  $k^{(m)} = (m+1) \times 2 + 1$ 。对于  $n=0$ , 有  $X_{l-1,n}^{(2)} = X_{l-1}^{(2)}$ 。PKI 模块不使用扩张卷积,从而避免了提取过于稀疏的特征表示。

然后,通过大小为  $1 \times 1$  的卷积来融合局部和上下文特征,表征不同通道之间的相互关系:

$$P_{l-1,n} = \text{Conv}_{1 \times 1}[L_{l-1,n} + \sum_{m=1}^4 Z_{l-1,n}^{(m)}] \quad (2)$$

$P_{l-1,n} \in R^{1/2 C_l \times H_l \times W_l}$  表示输出特征。 $1 \times 1$  卷积作为通道融合机制,用于整合具有不同感受野大小的特征,可以使 PKI 模块在不损害局部纹理特征完整性的情况下捕获广泛的上下文信息。

在目标检测过程中 PKI 模块展现出了较好的适应能力,通过新增 PKI 模块,实现了驾驶路面标识的检测与识别过程中对多类型目标的关注,能够应对复杂道路场景中带来的挑战。

## 2.3 可变形注意力机制

为了应对复杂的道路场景以及天气、光照等环境因素的影响,我们决定引入注意力机制。在近几年的深度

学习研究中, 注意力机制已经成为一种非常重要的工具, 尤其是在目标检测任务中, 表现尤为突出。然而, 标准的自注意力机制在进行路面标识检测时容易出现计算复杂度过高、处理效率低的问题。因此, 引入了可变形注意力机制 (DA), 一种更加灵活且计算效率更高的注意力机制<sup>[17]</sup>。

在传统的自注意力机制中, 权重是通过位置固定的注意力模型进行计算得到的, 序列中的每个位置或图像中的每个空间点都以固定、预先定义的方式与其他位置相互作用<sup>[18]</sup>。这种方式计算复杂度高, 而且在处理大尺寸图像时, 标准的自注意力机制难以在可接受的时间内完成计算。而可变形注意力机制能够通过生成动态偏移来调整采样位置, 动态地调整注意力模型的形状和大小, 减少计算范围, 然后在这些关键位置上应用动态偏移, 以获取更加准确的注意力权重, 以更好地适应不同任务和输入数据的特点。这一创新机制使模型能够处理数据中复杂的关系, 从而在识别图像或序列中的复杂模式时更加灵活和智能。

可变形注意力机制大致流程可以被分成 4 个步骤, 如图 3 所示。

1) 给定特征图  $x \in R^{H \times W \times C}$ , 生成一个点  $p \in R^{H_c \times W_c \times 2}$  的统一网格作为参考;

2) 将特征映射线性投影到  $q = xW_q$ , 然后输入一个轻量级网络  $\theta_{\text{offset}}$ , 生成偏移量  $\Delta\theta_{\text{offset}}(q)$ ;

3) 在变形点的位置采样, 作为键和值, 与查询一同传入多头注意力机制中;

4) 每个头部的特征  $W_o$  连接在一起, 通过投影得到最终输出  $z$ 。

单尺度可变形注意力用公式表达如下:

$$\text{DeformAttn}(z_q, p_q, x) =$$

$$\sum_{m=1}^M W_m \left[ \sum_{k=1}^K A_{mqk} W'_m x(p_q + \Delta p_{mqk}) \right] \quad (3)$$

$A_{mqk}$  直接由输入特征  $z_q$  通过全连接层得到;  $p_q$  代表  $z_q$  的位置, 用来作为参考点;  $\Delta p_{mqk}$  表示第  $m$  个注意

力头中的第  $k$  个采样点相对于参考点的位置偏移;  $K$  是采样的键总数。可以看出, 每个查询都会在每个注意力头部中采样  $K$  个位置, 所以只需要和这些位置的特征进行  $[x(p_q + \Delta p_{mqk})]$  的交互即可。

通过引入可变形注意力, 注意力机制可以自适应地调整, 模型在面对未见过的场景时也能更好地泛化, 减少误检和漏检的情况, 尤其是在处理不规则目标以及复杂背景时表现更加优异, 以实现更加准确和可靠的检测识别, 显著提升了 YOLOv8 网络在驾驶路面标识检测与识别中的性能。

## 2.4 $IoU_{\text{inner}}$ 损失函数

YOLOv8 算法采用了一种综合性的  $IoU$  损失函数来训练模型, 涵盖了分类损失、定位损失和置信度损失等多个方面<sup>[19]</sup>。作为现有主流边界框回归损失函数的基本部分,  $IoU$  定义如下:

$$IoU = \frac{|B \cap B^{\text{gt}}|}{|B \cup B^{\text{gt}}|} \quad (4)$$

YOLOv8 算法的损失函数沿用了 YOLOv5 的  $CIoU$  损失函数, 其可以用公式表示如下:

$$L_{CIoU} = 1 - IoU + \frac{\rho^2(b, b^{\text{gt}})}{c^2} + \alpha \nu \quad (5)$$

其中:  $\nu$  用来衡量纵横比的一致性。然而,  $CIoU$  损失函数仍然集中在通过添加新损失项来加速收敛, 忽略了  $IoU$  损失项本身的局限性, 从理论上讲,  $IoU$  损失可以有效地描述边界框回归的状态, 但在实际的路面标识检测应用中, 面对目标尺度的变化以及天气环境的影响, 它无法根据检测任务进行自适应调整, 且不具备较强的泛化能力。

针对以上问题, 本文使用  $IoU_{\text{inner}}$  替换原本的  $CIoU$  损失函数。 $IoU_{\text{inner}}$  能够通过辅助边框计算  $IoU$  损失, 在不同尺度下优化模型。使用较小尺度的辅助边框计算  $IoU$  损失有助于高  $IoU$  样本的快速回归, 加快收敛过程。相反, 使用较大尺度的辅助边框计算  $IoU$  损失能加速低  $IoU$  样本的回归<sup>[20]</sup>。 $IoU_{\text{inner}}$  引入了尺度因子  $r$ ,

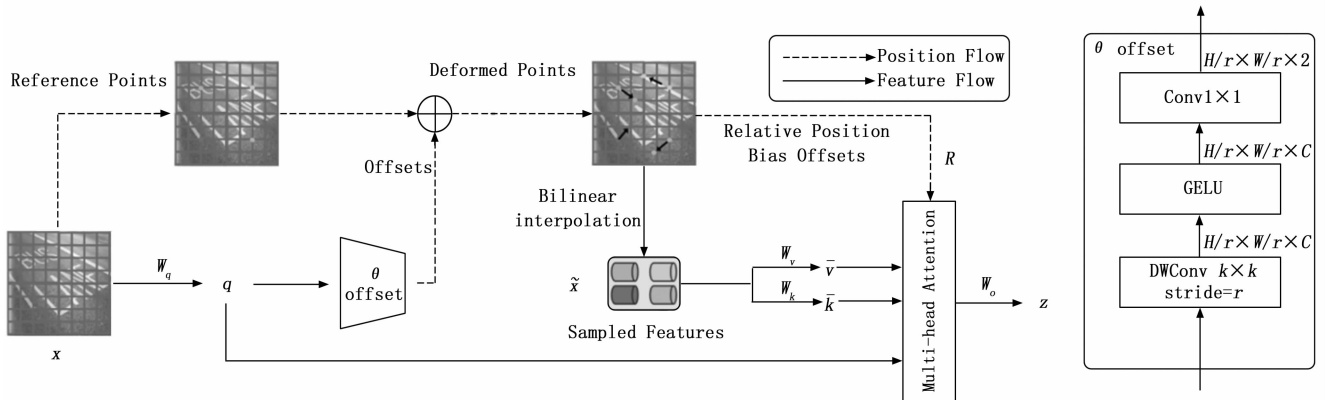


图 3 DA 注意力机制作用流程

根据不同的数据集和检测器控制辅助边框的大小,对于小目标数据集,选择较小的尺度因子,对于多尺度目标数据集,则根据目标类别动态调整。不同检测器对辅助框的需求各异,  $IoU_{inner}$  通过尺度因子适配不同检测器的设计,从而更快、更有效地实现回归。

图 4 的左右图分别为较小尺度和较大尺度下的  $IoU_{inner}$  计算方式。变量  $r$  对应尺度因子,通常在  $[0.5, 1.5]$  的范围内,当  $r > 1$  时,辅助框变大,更倾向于涵盖更大的目标区域;而当  $r < 1$  时,辅助框变小,更倾向于捕捉更精确的目标区域。

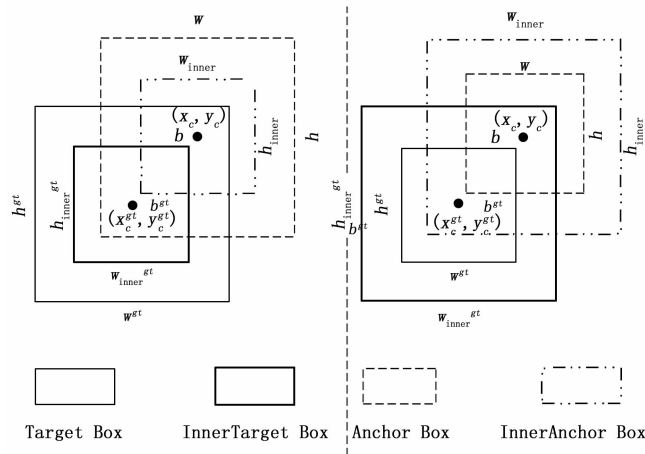


图 4  $IoU_{inner}$  示意图

$IoU_{inner}$  损失函数的计算方式如下所示:

$$b_l^{gt} = x_c^{gt} - \frac{w^{gt} \times r}{2}, b_r^{gt} = x_c^{gt} + \frac{w^{gt} \times r}{2} \quad (6)$$

$$b_t^{gt} = y_c^{gt} - \frac{h^{gt} \times r}{2}, b_b^{gt} = y_c^{gt} + \frac{h^{gt} \times r}{2} \quad (7)$$

$$b_l = x_c - \frac{w \times r}{2}, b_r = x_c^{gt} + \frac{w \times r}{2} \quad (8)$$

$$b_t = y_c - \frac{h \times r}{2}, b_b = y_c^{gt} + \frac{h \times r}{2} \quad (9)$$

$$inter = [\min(b_r^{gt}, b_r) - \max(b_l^{gt}, b_l)] \times [\min(b_b^{gt}, b_b) - \max(b_t^{gt}, b_t)] \quad (10)$$

$$union = (w^{gt} \times h^{gt}) \times r^2 + (w \times h) \times r^2 - inter \quad (11)$$

$$IoU_{inner} = \frac{inter}{union} \quad (12)$$

$IoU_{inner}$  损失继承了  $IoU$  损失的一些特征,同时具有自己的特点。 $IoU_{inner}$  损失的范围与  $IoU$  损失相同,为  $[0, 1]$ 。由于辅助边界框与实际边界框之间仅存在尺度差异,损失函数的计算方法相同,  $IoU_{inner}$  偏差曲线与  $IoU$  偏差曲线相似。

与  $IoU$  损失相比,当比例小于 1 且辅助边界框尺寸小于实际边界框时,回归的有效范围小于  $IoU$  损失,但梯度的绝对值大于从  $IoU$  损失获得的梯度,可以提

高  $IoU$  样本的回归速度,进而加速收敛。相反,当比例大于 1 时,较大尺度的辅助边界框扩展了回归的有效范围,对低  $IoU$  样本的回归具有增强效果。

本文将  $IoU_{inner}$  损失应用于现有的基于  $CIoU$  的边界框回归损失函数,可以用公式表示如下:

$$L_{CIoU_{inner}} = L_{CIoU} + IoU - IoU_{inner} \quad (13)$$

$IoU_{inner}$  在促进训练过程收敛和提高检测精度方面有显著优势。在驾驶路面标识的检测与识别中采用  $IoU_{inner}$  损失函数可以显著提升检测系统的性能和可靠性。

### 3 实验结果与分析

#### 3.1 数据集介绍

如图 5 所示,本文使用的数据集为 Kaggle 官网的公开数据集 Road,包含 2 776 张驾驶路面标识的图片,数据集被分为训练集 (Train)、验证集 (Valid)、测试集 (Test) 共 3 部分,其中训练集 2 167 张,验证集 417 张,测试集 192 张,图片包括公交车道、禁停线、白色纵向实线、白色横向实线、斑马线、人行横道预告标识、减速标识、左转道线、右转道线、直行道线、直行或左转道线、直行或右转道线、自行车道线共计 13 种路面标识类型,基本涵盖日常行驶道路中遇到的各种场景,包括高速公路、城市街道和农村路段,能够很好地还原真实世界。



图 5 驾驶路面标识数据集图片实例

#### 3.2 实验环境搭建以及参数配置

1) 实验环境搭建: 本文硬件配置如表 1 所示。

表 1 硬件配置

系统	Windows10(64 位)
CPU	Intel core i5-12 400 f
RAM	32 GB
GPU	NVIDIA RTX 4060Ti 8 G
深度学习框架	PyTorch
GPU 加速库	CUDA 11.3

文中所有实验中模型的训练均在此台计算机上完成。

2) 实验参数配置: 实验所使用的训练参数如表 2 所示。

表 2 训练参数

输入图片尺寸	640×640
训练轮数	300
批量大小	24
初始学习率	0.01
尺度因子	$r=0.7$
优化器	SGD

3.3 模型评估指标

本文采用的评估指标包括精确率 (P, precision)、召回率 (R, recall) 和平均精度均值 (mAP, mean average precision)。精确率表示在所有被检测为正样本的目标中, 实际为正样本的比例; 召回率表示在所有实际为正样本的目标中, 被正确检测出的比例精确率和召回率分别可以用公式表示为:

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

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

TP 表示正确检测出的正样本数量, FP 表示错误检测出的正样本数量, FN 表示未能检测出的正样本数量。

为了综合评估模型在不同阈值下的表现, 通常采用平均精度均值 (mAP) 这一指标。mAP 是在不同阈值下的计算的平均精度 (AP) 值的均值, 其中 AP 是精确率-召回率曲线下的面积, 可以通过对不同召回率点的精确率进行插值计算得到:

$$AP = \int_0^1 P(R) dR$$
 (16)

mAP 则是对所有类别的 AP 值求平均值, 衡量了模型在多类别检测任务中的总体性能, 通常有两种不同的计算方式: mAP : 50 和 mAP : 50~95。mAP : 50 为在 IoU 阈值为 0.5 时计算的平均精度均值, mAP : 50~95 则为在 IoU 阈值从 0.5~0.95, 以 0.05 为间隔, 共计 10 个阈值下的计算的平均精度均值。P、R、mAP 指标共同构成了目标检测模型性能评估的重要标准, 能够有效衡量模型的检测准确性和鲁棒性。

3.4 模型训练结果展示

图 6 为改进后的 YOLOv8 模型对驾驶路面标识的检测训练结果的展示。

在实验中, 我们对基于改进 YOLOv8 的驾驶路面标识检测模型进行了充分的训练和评估, 结果显示出



图 6 部分驾驶路面标识训练结果展示

此模型在不同天气、尺度距离、光照环境以及多目标条件下都有出色的表现, 且大部分路面标识能达到 0.9 以上的识别率, 能够充分说明改进后的模型的高效与可靠性。

3.5 消融实验

为了验证所提出的改进 YOLOv8 模型在驾驶路面标识检测与识别中的有效性, 我们在所选数据集上设计并进行了消融实验。本实验旨在评估各个改进模块对整体模型性能的贡献, 从而确定添加的每个改进点在提升模型检测和识别能力方面的重要性。在原始 YOLOv8 模型的基础上, 逐一增加各个改进模块并观察模型性能的变化, 添加了改进表格内则用“√”表示, 未使用则表格内为空。通过对比分析, 揭示各个模块在提高检测平均精度均值上的作用。

根据所添加的改进策略做了相应的消融实验, 表 3 为消融实验对比结果。

表 3 消融实验

改进策略	第 1 组	第 2 组	第 3 组	第 4 组	第 5 组
DA		√			√
PKI			√		√
CIoU <sub>inner</sub>				√	√
mAP : 50/%	87.7	89.1	90.5	89.9	91.6
mAP : 50~95/%	66.1	67.9	69.2	68.3	70.5

第 1 组为原生 YOLOv8 算法模型的训练结果, 作为基准结果, 展示了原生 YOLOv8 算法在未进行任何改进策略时的性能。原生模型具备一定的检测能力, 但为了在更广泛和复杂的驾驶环境中实现更高的检测准确性和鲁棒性, 需要引入进一步的改进策略来优化模型性能。

第 2 组引入了可变形注意力机制, mAP : 50 和 mAP : 50~95 分别由 87.7%和 66.1%提升至 89.1%和 67.9%, 可以看出加入的 DA 模块通过动态调整注意力权重, 使模型能够更好地聚焦于目标区域的关键特征,

并减少了漏检和误检的情况出现，从而提升了检测精度。

第 3 组加入了 PKI 模块的实验结果，PKI 模块通过多种卷积核的组合，捕捉不同尺度和不同特征模式，从而提升模型的特征提取能力和检测精度。从结果来看， $mAP:50$  和  $mAP:50\sim95$  分别由 87.7% 和 66.1% 提升至 90.5% 和 69.2%，效果显著。

第 4 组实验把原生模型的损失函数  $CIoU$  替换为  $CIoU_{inner}$ ， $CIoU_{inner}$  通过辅助边框计算  $IoU$  损失，进一步优化目标边界框的定位损失函数，在不同尺度下优化模型，能够更准确地调整和优化目标检测结果， $mAP50$  和  $mAP50\sim95$  也分别提升至 89.9% 和 68.3%，充分证明了在模型优化过程中，选择更加合适的损失函数对提升模型性能的重要性。

第 5 组实验结果表明，组合 DA、PKI 和  $CIoU_{inner}$  三种改进策略能够在各自优势的基础上，进一步提升 YOLOv8 模型的整体性能。相比于原生模型，组合 3 种改进策略的模型在  $mAP:50$  和  $mAP:50\sim95$  分别由 87.7% 和 66.1% 提升至 91.6% 和 70.5%，均有显著提升，证明了改进后的模型的高效可靠，能够实现优秀的检测和识别性能。

3.6 损失函数收敛曲线

为了评估 YOLOv8 网络的改进效果，我们也重点关注了改进前后模型的收敛曲线。图 7 和图 8 分别为模型改进前后的损失函数收敛曲线。

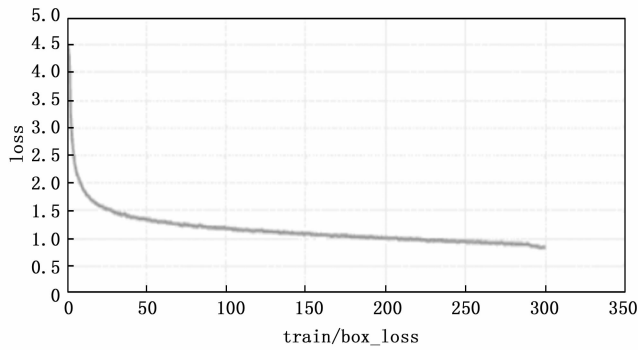


图 7 改进前模型的收敛曲线

对比图 7、8 可以看出：改进前损失下降速度较为缓慢，并且在达到一定次数的迭代后，损失函数的下降趋势趋于平缓，难以进一步显著降低。这表明改进前的 YOLOv8 在收敛性上存在一定的局限性，可能需要更多的迭代次数才能达到相对理想的效果。经过优化和调整后的 YOLOv8 网络，损失函数值在初始阶段迅速下降，表现出更快的收敛速度。随着训练的进行，损失函数持续显著降低，且在相同的迭代次数下，改进后的 YOLOv8 达到了更低的损失值。这表明改进后的网络在

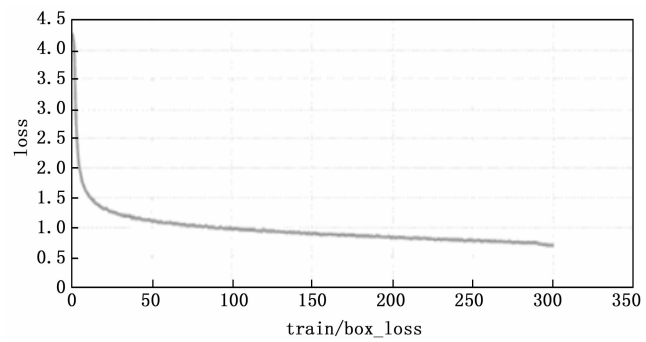


图 8 改进后模型的收敛曲线

收敛性上有了显著提升，能够更高效地进行训练，从而在较少的迭代次数内达到更优的性能。

通过对比改进前后 YOLOv8 网络的收敛曲线，可以清晰地看到改进后的网络在训练效率和最终性能上的优势。改进后的 YOLOv8 不仅加快了收敛速度，减少了训练时间，同时在损失函数收敛值上也达到了更低的水平，表明模型在训练过程中的稳定性和准确性均得到了提升。这些改进能够使模型在实际的路面标识的检测与识别应用中更高效地完成目标检测任务，证明改进后的模型具有更强的实用价值。

3.7 对比实验

为了全面评估改进后的 YOLOv8 模型在驾驶路面标识检测与识别任务中的性能，本文将其与现有的多个主流目标检测模型以及其他改进的 YOLO 模型在本文所选数据集上的训练结果进行了对比，包括 SSD、Faster R-CNN、YOLOv5、YOLOv7、YOLOv8-EMA 和 YOLOv8-CBAM。这些模型代表了当前目标检测领域的不同方法和技术路线，具有广泛的应用和较高的影响力。通过对比这些模型的检测精度和实时性，能够更全面地验证改进策略的有效性和优势。对比模型的参数均为默认值，对比试验结果如表 4 所示。

表 4 对比实验

模型	$mAP:50/\%$	$mAP:50\sim95/\%$	FPS/Hz
SSD	75.1	53.2	33
Faster R-CNN	76.7	55.4	26
YOLOv5	81.9	60.4	39
YOLOv7	83.2	65.3	47
YOLOv8	87.7	66.1	52
YOLOv8-EMA	86.2	65.9	51
YOLOv8-CBAM	86.9	66.8	58
本文模型	91.6	70.5	71

由此可见，改进后的 YOLOv8 模型在驾驶路面标识的检测识别中检测精度显著高于其他模型。同时，我们也把改进后模型的检测速度（FPS）与其他检测模型进行了对比，原生的 YOLOv8 模型检测速度为

52 Hz, 改进后的模型的检测速度达到了 71 Hz, 表现最为优异, 这表明, 通过引入 DA、PKI 和  $CIoU_{inner}$  等改进策略, 改进后的 YOLOv8 能够在复杂驾驶场景中更准确地识别和定位路面标识物, 提供更可靠的检测性能, 在现有的多个主流目标检测模型具有很大优势。

#### 4 结束语

针对驾驶路面标识的检测与识别目前面临的诸多问题, 本文提出了一种基于改进 YOLOv8 的检测模型, 在 Backbone 部分新增了 PKI 模块, 使模型能够从图像数据中提取更加复杂的特征, 并引入 DA 注意力机制, 让模型更加灵活地关注图像的相关部分, 适应输入数据的特定空间结构, 最后将原本的损失函数  $CIoU$  替换为  $CIoU_{inner}$ , 通过辅助边界框计算  $IoU$  损失, 增强模型的泛化能力, 提高检测精度。

经过实验, 改进后模型的平均精度均值  $mAP: 50$  和  $mAP: 50 \sim 95$  相较于原生模型分别提升了 4.4% 和 6.7%, 检测速度也由原本的 52 Hz 提升至 71 Hz, 充分证明改进后的 YOLOv8 模型在复杂路况和多样化标识场景下表现出色, 显示出较高的检测精度和速度。尽管改进后的模型在大多数情况下表现优异, 但在极端天气条件下, 如暴雨、大雾等环境中的检测效果仍有待验证。我们在未来的研究中, 需要进一步优化模型的鲁棒性, 以应对各种极端环境的挑战。其次, 模型的计算资源消耗仍然是一个需要关注的问题。未来将尝试在保证精度的前提下进一步减少计算资源的消耗, 提高模型在资源受限设备上的应用能力。

#### 参考文献:

- [1] GRIGORESCU S, TRASNEA B, COCIAS T, et al. A survey of deep learning techniques for autonomous driving [J]. Journal of Field Robotics, 2020, 37 (3): 362 - 386.
- [2] LI Y, LI X, LI T. Vision-based pedestrian detection for automated driving: a benchmark and a new model [C] // International Conference on Computer Vision (ICCV), 2019: 1899 - 1908.
- [3] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C] // Computer Vision and Pattern Recognition (CVPR), 2014: 580 - 587.
- [4] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [C] // Neural Information Processing Systems, 2015: 91 - 99.
- [5] ZHAO Q, SHENG L, WANG Y, et al. M2Det: a single-shot object detector based on multi-level feature pyramid network [C] // Artificial Intelligence (AAAI), 2019: 9259 - 9266.
- [6] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C] // Conference on Computer Vision (ICCV), 2017: 2980 - 2988.
- [7] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [C] // ECCV, Lecture Notes in Computer Science, 2016: 21 - 37.
- [8] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C] // Computer Vision and Pattern Recognition, 2016: 779 - 788.
- [9] 刘伟. 基于深度神经网络的道路标识标线识别系统研究 [D]. 重庆: 重庆交通大学, 2023.
- [10] 韩建鹏. 基于深度学习的交通标志检测与识别优化研究 [D]. 兰州: 兰州交通大学, 2022.
- [11] 高怀序. 基于深度网络的道路标识检测算法研究 [D]. 辽宁: 辽宁科技大学, 2023.
- [12] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C] // Computer Vision and Pattern Recognition (CVPR), 2017: 2117 - 2125.
- [13] MISRA D. Mish: a self regularized non-monotonic neural activation function [J]. Arxiv Preprint, 2019, Arxiv: 1908.08681.
- [14] RAMACHANDRAN P, ZOPH B, LE Q V. Searching for activation functions [J]. Arxiv Preprint, 2017, Arxiv: 1710.05941.
- [15] ZHENG Z, WANG P, LIU W, et al. Distance-IoU-loss: faster and better learning for bounding box regression [C] // Artificial Intelligence, 2020, 34 (7): 12993 - 13000.
- [16] CAI X, LAI Q, WANG Y, et al. Poly Kernel inception network for remote sensing detection [C] // Computer Vision and Pattern Recognition (CVPR), 2024: 27706 - 27716.
- [17] XIA Z, PAN X, SONG S, et al. Vision transformer with deformable attention [C] // Computer Vision and Pattern Recognition (CVPR), 2022: 4794 - 4803.
- [18] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C] // European Conference on Computer Vision (ECCV), 2018: 3 - 19.
- [19] MILLETARI F, NAVAB N, AHMADI S. V-Net: fully convolutional neural networks for volumetric medical image segmentation [C] // IEEE, 2016: 565 - 571.
- [20] ZHANG H, XU C, ZHANG S.  $IoU_{inner}$ : more effective intersection over union loss with auxiliary bounding box [J]. Arxiv preprint, 2023, Arxiv: 2311.02877.