

基于改进 RT-DETR 的极端天气下交通标志检测方法*

秦伦明¹ 张云起¹ 崔昊杨¹ 边后琴¹ 王悉²

(1. 上海电力大学电子与信息工程学院 上海 201306; 2. 北京交通大学自动化与智能学院 北京 100044)

摘要: 针对雨、雾和雪等极端天气下交通标志模糊不清,导致检测精度下降和小目标识别困难等问题,本文提出了一种基于改进 RT-DETR 的交通标志检测算法。首先,采用数据增强方法模拟极端天气环境,以提高模型在这些环境下对交通标志的识别能力。其次,在主干网络中引入 Ortho 注意力机制,利用正交滤波器减少特征冗余,筛选重要通道信息,提高对小目标的检测精度。此外,采用高层筛选特征金字塔网络(HS-FPN)替换原模型中的跨尺度上下文特征混合器(CCFM),通过高层特征筛选并融合低层特征信息,提升模型在极端天气下对低对比度和模糊目标的检测精度。实验结果显示,改进算法在平均检测精度方面达到 87.84%,相比原始 RT-DETR 模型提高了 2.37%,同时参数量减少至 18.22 M,相比原模型降低了 8.4%,对小目标和处于极端天气中的目标识别精度更高,对保障乘客的安全具有实际意义。

关键词: RT-DETR; 正交通道注意力机制; 高层筛选特征金字塔网络; 交通标志识别; 图像增强

中图分类号: TP391.41; TN911 文献标识码: A 国家标准学科分类代码: 510

Improved RT-DETR based method for traffic sign recognition in extreme weather

Qin Lunming¹ Zhang Yunqi¹ Cui Haoyang¹ Bian Houqin¹ Wang Xi²

(1. College of Electronic and Information Engineering, Shanghai University of Electric Power, Shanghai 201306, China; 2. School of Automation and Intelligence, Beijing Jiaotong University, Beijing 100044, China)

Abstract: To address the issues of decreased detection accuracy and difficulty in small object recognition caused by blurred traffic signs in extreme weather conditions such as rain, fog, and snow, this paper proposes a traffic sign detection algorithm based on an improved RT-DETR. First, data augmentation is applied to the TT100K dataset under simulated extreme weather conditions to enhance the model's ability to recognize traffic signs in these environments. Second, the Ortho attention mechanism is introduced into the backbone network, which uses orthogonal filters to reduce feature redundancy and prioritize essential channel information, thereby improving the model's detection accuracy for small objects. Additionally, a high-level screening-feature pyramid network (HS-FPN) replaces the cross-scale contextual feature mixer (CCFM) in the original model. HS-FPN filters and merges low-level feature information using high-level features, enhancing the model's detection accuracy for low-contrast and blurred targets in extreme weather conditions. Experimental results show that the proposed improved algorithm achieves an average detection accuracy of 87.84%, an improvement of 2.37% over the original RT-DETR model, while reducing the parameter count to 18.22 M, an 8.4% reduction compared to the original model. The model demonstrates higher accuracy in recognizing small objects and targets under extreme weather, contributing significantly to passenger safety.

Keywords: RT-DETR; orthogonal attention mechanism; HS-FPN network; traffic sign detection; image enhancement

0 引言

随着人工智能技术的快速发展,自动驾驶作为一个新兴领域受到了广泛关注和深入研究。在自动驾驶系统中,

交通标志的检测与识别是一个至关重要的组成部分,能够为智能车辆提供实时的道路环境信息,辅助其进行安全和合理的决策。然而,极端天气(如雨天、雾天)会严重影响车载传感器成像质量,降低交通标志的对比度和清晰度,进而

影响检测准确性和鲁棒性。因此,在极端天气下对交通标志的准确检测和识别仍是一个巨大的挑战。

传统的交通标志检测方法^[1-2]主要依赖于人工设计的颜色和形状特征,在晴朗的天气条件下通常能取得较好的检测效果,但在复杂天气环境(如雨天或雾天)下,往往表现出较高的误检率和漏检率。为了应对这些问题,近年来深度学习技术被广泛应用于交通标志检测领域。

深度学习模型能够自动学习交通标志的多尺度特征和多样性特征,并具备适应不同光照和天气条件的能力,显现出相较于传统方法更优越的性能。基于深度学习的方法^[3]主要分为一阶段检测算法和两阶段检测算法。两阶段检测算法如 RCNN^[4]、Faster RCNN^[5],首先生成大量候选区域,再对候选区域进行识别,因此检测速度较慢,算法结构复杂,难以满足交通标志检测的实时性需求。相较之下,一阶段检测算法效率更高,检测速度更快,主要有 SSD^[6]算法,CenterNet^[7]算法,RetinaNet^[8]算法和 YOLO^[9-13]系列算法等。然而,现有方法在极端天气条件下的检测性能仍然有限,尤其在雨雾等恶劣天气下,模型的识别准确性仍需进一步提升。文献[14]中提出的深度融合网络结合了传统卷积神经网络和注意力机制,以提高雨天环境下的交通标志检测能力,但对于小目标的检测精度依然较低。文献[15]提出了一种轻量化的 RT-DETR 交通标志检测方法,使用 ShuffNetv2 替换原 Resnet 主干,降低了计算量和参数量,提高了检测速度,但是没有考虑小目标以及处于极端天气下的交通标志存在的漏检和误检现象。由上述改进可以得出,交通标志检测主要面临以下问题。首先,不同天气条件下的检测精度不稳定,特征融合方法的灵活性和适应性不足,其次,小目标检测性能较差。

针对以上问题,本文提出了一种基于 RT-DETR 的极端天气下交通标志检测方法,以提升模型在恶劣天气条件下的检测精度和速度。首先,本文利用 TT100K 数据集并对其进行雨雾数据增强处理,以增加模型的多样化训练样本,提升其在极端天气条件下的泛化能力。其次,在主干网络中引入了改进的 BasicBlock 模块,该模块有效增强了模型在雨雾环境下对模糊目标的特征捕捉能力,进一步提高了模型在恶劣环境中的检测效果。此外,本文在颈部网络中引入了高层筛选特征金字塔网络 (high-level screening-feature pyramid networks, HS-FPN)^[16],以更好地融合多尺度特征信息,提升模型在小目标及复杂背景下的检测能力。

1 模型介绍

1.1 RT-DETR 模型

该网络结构主要由 3 个部分组成,主干网络 (Backbone)、跨尺度上下文特征混合器 (cross-scale contextual feature mixer, CCFM)^[17] 和头部解码器 (Head),其网络结构如图 1 所示。主干网络的主要作用是从输入图

像中提取特征信息。输入图像首先经过多个卷积层和最大池化层,逐步降低图像尺寸,同时提取图像的特征信息。随后,通过 4 个 BasicBlock 模块进一步进行特征提取和空间降维。每个 BasicBlock 模块由多个卷积层,批归一化层 (batch normalization, BN) 和 ReLU 激活函数组成。CCFM 的功能是融合不同尺度的特征,以捕捉目标的多尺度信息和上下文依赖关系。该部分由 RepC3 模块、AIFI 模块和卷积层组成。CCFM 从主干网络获取多尺度特征图,通过卷积层进行特征维度的调整和融合。RepC3 模块对融合后的特征进行增强,而 AIFI 模块自适应地融合不同层次和尺度的特征,提高检测精度。头部解码器负责将混合后的特征图解码为每个检测框的类别置信度和位置坐标,最终完成交通标志的检测和定位。

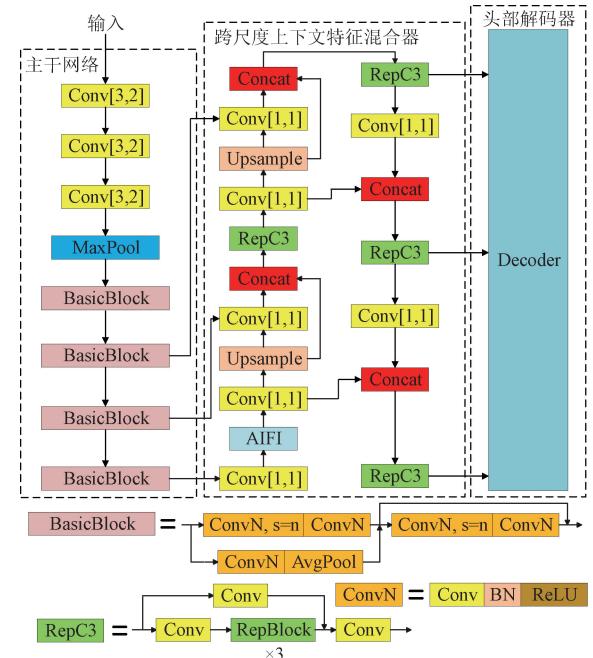


图 1 RT-DETR 网络结构

Fig. 1 RT-DETR network architecture

1.2 BasicBlock_Ortho 模块

由于车辆行驶过程中摄像头捕捉到的交通标志大多数是小目标,且道路环境复杂,模型在检测时容易出现误检和漏检现象。为了提高模型对小目标的识别精度,本文在主干网络的 BasicBlock 模块中融合了正交通道注意力机制 (orthogonal channel attention, Ortho attention)^[18],使得模型在特征提取过程中能够更有效地捕捉目标的关键特征信息。

正交通道注意力机制的核心是通过正交滤波器对输入特征进行压缩,并利用注意力机制对每个通道进行加权,该机制通过对特征的重要性进行自适应调整,提高模型对关键特征的关注度。同时,通过残差连接保持原始信息与增强信息的融合,从而提升特征表示能力。其结构图如图 2 所示。该过程分为以下几个步骤:随机初始化一组滤波器 $F_j \in \mathbb{R}^{H \times W}$,其中 $H \times W$ 表示滤波器的大小,对这一组滤

波器使用 Gram-Schmidt 进行正交化处理, 得到正交滤波器 $K \in \mathbb{R}^{H \times W \times C}$ 。

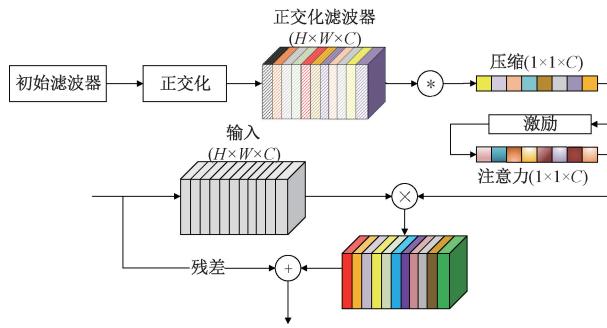


图 2 Ortho attention 网络结构

Fig. 2 Ortho attention network architecture

正交化确保每个滤波器在压缩过程中捕获的特征信息互相独立, 有利于保留更多不同维度的特征信息。随后将大小为 $H \times W \times C$ 的输入特征图通过正交滤波器进行压缩, 得到压缩后的特征向量:

$$\mathbf{F}_{\text{ortho}}(\mathbf{X})_c = \sum_{h=1}^H \sum_{w=1}^W K_{c,h,w} X_{c,h,w} \quad (1)$$

压缩后的特征向量大小为 $1 \times 1 \times C$, 将特征向量 $\mathbf{F}_{\text{ortho}}(\mathbf{X})$ 通过激励函数生成注意力向量 $\mathbf{A}(\mathbf{X})$:

$$\mathbf{A}(\mathbf{X}) = \sigma(\mathbf{W}_2 \delta(\mathbf{W}_1 \mathbf{F}_{\text{ortho}}(\mathbf{X}))) \quad (2)$$

其中, σ 是 Sigmoid 激活函数, δ 是 ReLU 函数, \mathbf{W}_1 和 \mathbf{W}_2 是可学习的权重矩阵。将压缩后的特征向量与原始输入特征 \mathbf{X} 进行逐通道乘法, 得到加权后的输出特征 $\mathbf{X}_{\text{weighted}}$:

$$\mathbf{X}_{\text{weighted}} = \mathbf{A}(\mathbf{X}) \odot \mathbf{X} \quad (3)$$

其中, \odot 表示逐通道乘法。最后, 将加权后的特征与

原始输入特征 \mathbf{X} 进行相加, 得到最终输出 \mathbf{Y} :

$$\mathbf{Y} = \mathbf{X}_{\text{weighted}} + \mathbf{X} \quad (4)$$

为了提高模型对于小目标的检测精度, 本文将 **Ortho** 注意力机制与主干网络中的 **BasicBlock** 模块进行融合, 形成改进后的 **BasicBlock_Ortho** 模块, 该模块结构如图 3 所示。

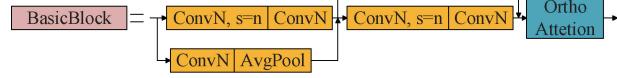


图 3 BasicBlock_Ortho 网络结构

Fig. 3 BasicBlock_Ortho network architecture

将 **Ortho** 注意力机制融入 **BasicBlock** 模块, 使模型能够动态调整各个通道的重要性, 增强对小目标的特征提取能力。

1.3 HS-FPN 模块

HS-FPN 是一种结合特征选择与多尺度特征融合的网络结构, 该网络通过通道注意力模块优化特征选择, 并通过自适应的特征融合模块增强多尺度特征的表达能力, 从而有效提升模型对目标的检测精度, 其结构如图 4 所示。HS-FPN 主要包含以下 3 个模块: 特征选择模块、特征融合模块和通道注意力模块。

在特征选择模块中, S3、S4、S5 代表主干网络中 3 个不同尺度的特征图, 每个尺度的特征图通过通道注意力模块 (channel attention, CA) 进行特征加权处理。特征图先后经过全局最大池化和全局平均池化处理, 随后将两者相加, 再通过 Sigmoid 激活函数计算得到每个通道的权重, 加权后的特征图通过卷积层进行降维处理, 最终得到加权后的 3 个特征图 P3、P4、P5。

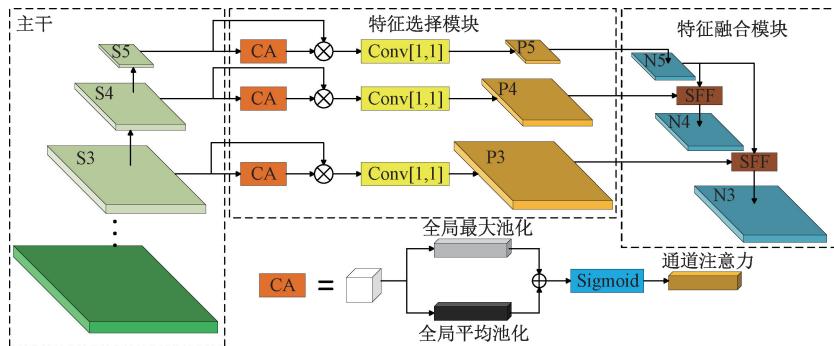


图 4 HS-FPN 网络结构

Fig. 4 HS-FPN network architecture

在特征融合模块中, 高分辨率的特征图包含丰富的细节信息, 但语义信息较少。相反, 低分辨率的特征图具备丰富的语义信息, 但目标定位较为粗糙。为解决这一问题, HS-FPN 网络中的 SFF 模块选择性融合高分辨率和低分辨率特征图, 使模型自适应地融合不同层级的特征信息, 从而提高目标识别精度。SFF 的网络结构如图 5 所示。

首先, 将包含丰富语义信息的高层特征 f_{high} 通过转置卷积和双线性插值处理, 使其与包含丰富细节信息的低层特征 f_{low} 在维度上匹配, 得到特征图 f_{att} 。其次, 使用 CA 模块对高层特征进行加权, 并与低层特征融合, 形成多尺度输出特征 f_{out} 。通过 HS-FPN 结合多尺度特征, 实现多尺度特征的有效融合, 从而提高模型对极端天气下低对比度及模糊目标的检测精度。

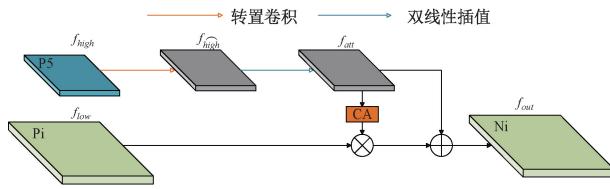


图 5 SFF 网络结构

Fig. 5 SFF network architecture

1.4 BH-RT-DETR 网络结构

本文针对交通标志识别中的问题,提出了一种改进算法 BH-RT-DETR (BasicBlock_Ortho and HS-FPN based RT-DETR)。改进后的网络结构如图 6 所示。该网络由主干网络,颈部网络和头部解码器 3 部分组成。

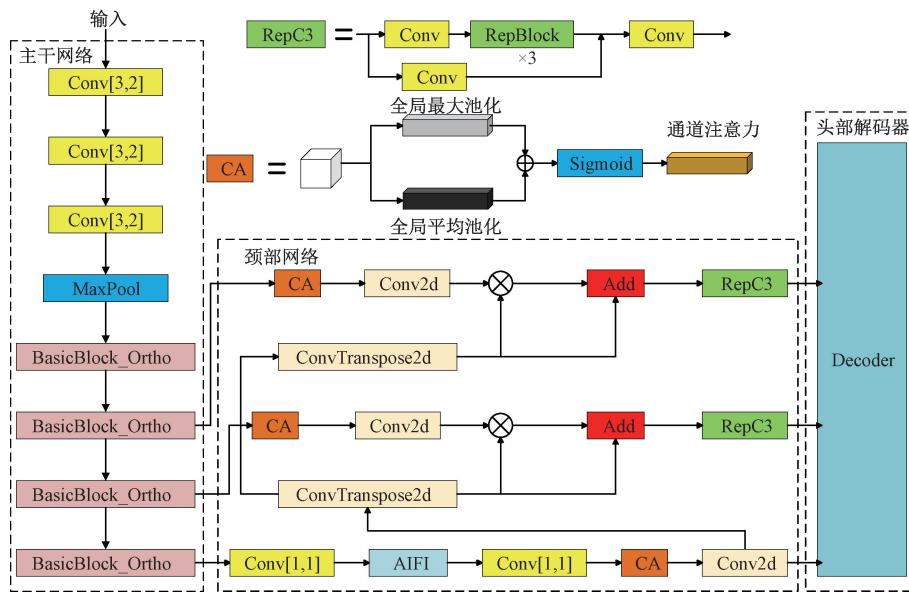


图 6 BH-RT-DETR 网络结构

Fig. 6 BH-RT-DETR network architecture

2 实验结果与分析

2.1 实验环境及数据集

为验证本文算法的优越性,本文进行了基于同一交通标志数据集的多种算法对比实验和消融实验,实验软硬件环境配置如表 1 所示。模型训练过程中,输入图像分辨率为 320 pixel \times 320 pixel,初始学习率设置为 0.01,训练周期设置为 150,批处理大小(batch)为 16,采用 AdamW 优化器。实验数据集选用 TT 100K 数据集,该数据集由清华大学与腾讯联合实验室共同整理并公布,共标注了 128 类交通标志,涵盖城市道路、乡村道路以及高速公路等多种场景。然而,该数据集存在标签分布不均的问题,部分类别标签数量不足 50 个,导致网络在数量较少的类别上难以收敛,进而影响模型的整体训练效果。因此,本文选取了标签数量超过 300 个的 15 类交通标志。图 7 展示了用

本文将 RT-DETR 的 BasicBlock 模块与 Ortho 注意力机制相结合,得到 BasicBlock_Ortho 模块,使得网络在特征提取过程中能够更加精确地捕捉目标特征信息。同时,将原网络中的 CCFM 替换为 HS-FPN 特征融合网络,通过 HS-FPN 网络的自适应特征选择和融合机制,充分利用多尺度特征信息,进一步提升网络对不同尺度目标的检测能力。改进后的 RT-DETR 模型中,BasicBlock_Ortho 模块通过引入 Ortho 注意力机制,有效增强了特征通道之间的区分能力,使得网络能够更加关注关键目标特征,同时抑制不重要的背景干扰。HS-FPN 模块通过选择性特征融合模块和通道注意力机制,对不同层级的特征图进行自适应选择和融合,从而提高网络在检测小目标及复杂背景目标时的精度。

表 1 训练环境与训练参数设置

Table 1 Training environment and training parameter settings

项目	配置
操作系统	Windows11
CPU	Intel(R) Core(TM) i9-13980H
GPU	NVIDIA GeForce RTX 4060
内存	32 G
Python 版本	3.9.7
CUDA、CUDNN 版本	CUDA 11.6, CUDNN 8.7.0

此外,本文使用 imgaug 库对 9 050 张图片进行数据增强,例如使用 Rain 函数添加雨滴、GaussianBlur 函数增加

不同程度的模糊效果模拟雾霾,以及通过 Snow 函数引入雪花噪声模拟雪天,增强后的图像如图 8 所示。通过数据

增强,最终得到了 13 239 张图片,并按 8:1:1 的比例划分为训练集,验证集和测试集进行模型训练。



图 7 选取的交通标志

Fig. 7 Selected traffic signs



图 8 增强前后对比

Fig. 8 Comparison before and after augmentation

2.2 评价指标

为了验证改进后的 BH-RT-DETR 网络模型的性能,本文采用参数量和平均精准度(*mAP*)^[19]对算法进行评估。精确率(*P*)和召回率(*R*)的计算方法为:

$$P = \frac{TP}{TP + FP} \quad (5)$$

$$R = \frac{TP}{TP + FN} \quad (6)$$

其中, *TP* 表示检测结果为正的正样本, *FP* 表示检测结果为正的负样本, *FN* 表示检测结果为负的正样本。通过不同的 *P* 和 *R* 可以画出 PR 曲线, PR 曲线下的面积表示一类目标的平均检测精度 *AP*, *AP* 为所有检测类别 *mAP* 的均值,其计算方法如下:

$$mAP = \frac{\sum_{i=1}^N AP_i}{N} \quad (7)$$

式中: *AP_i* 表示第 *i* 类目标的平均检测精度, *N* 为数据集中所有类别的数量。

2.3 消融实验

为了验证本文提出的各项改进对交通标志检测算法的提升效果,本文以 RT-DETR 作为基准模型,依次加入各个改进模块,并在相同的数据集上进行消融实验,结果如表 2 所示。

表 2 消融实验结果

Table 2 Ablation study results

模型	HS-FPN	BasicBlock_Ortho	<i>mAP</i> /%	参数量 /M	FLOPs /G
RT-DETR			85.47	19.89	57.0
A	✓		86.68	18.13	53.4
B		✓	86.33	19.98	57.0
C	✓	✓	87.84	18.22	53.4

表 2 中, A 表示在 RT-DETR 模型中引入 HS-FPN 特征融合网络,以增强小目标检测能力。实验结果表明,加入 HS-FPN 特征融合网络后,模型对小目标的识别精度显著提高, *mAP* 从 85.47% 提升至 86.68%,同时参数量减少至 18.13 M,说明 HS-FPN 特征融合网络在特征提取方面表现出更高的效率。B 表示将模型中的 BasicBlock 模块替换为 BasicBlock_Ortho 模块,通过引入正交化卷积,该模块不仅降低了模型的参数量,还提高了模型的目标识别精度,相比原模型 *mAP* 提高了 0.86%。C 表示在结合 HS-FPN 特征融合网络的基础上加入 BasicBlock_Ortho 模块,实验结果显示,该改进组合的 *mAP* 达到了 87.84%,相较于基础模型提升了 2.37%,同时,参数量和计算量分别减少至 18.22 M 和 53.4 G,各项实验指标均优于 RT-DETR,进一步验证了改进算法的有效性。

2.4 数据增强前后对比实验

为了验证数据集在加雨雾和雪处理前后的检测效果,本文分别使用加雨雾和雪前后的数据集对 RT-DETR 模型进行训练,实验结果如表 3 所示。由表 3 可知,经过雨雾和雪的增强处理后,RT-DETR 模型在大多数类别上的检测性能显著提升。对于 *pn* 和 *pne* 类别,增强后的精度分别达到 95.2% 和 96.7%,召回率达到 90.8% 和 92.5%,表明模型在复杂气象条件下的识别能力得到了显著增强。此外, *il60* 和 *pl100* 类别的改进较为明显,增强后的精度分别达到 95.9% 和 88.1%,召回率提升至 98.9% 和 90.4%。同样, *p11* 和 *p26* 类别在数据增强后也表现出较大幅度的提升,精度分别提高到 90.5% 和 89.0%。这些结果表明,通过对数据集进行加雨雾和雪等气象条件的增强处理,模型在复杂环境下的检测精度得到了提高。

与最新的图像增强方法相比,本文使用的 imgaug 库中的增强方法具有较高的效率和定制化优势。通过

模拟雨滴、雾霾、雪花噪声等极端天气条件,能够增强图像的真实感和环境适应性。相较于需要大量计算资源和训练过程的深度学习增强方法,imgaug 提供的操作简单高效、易于实现,且不依赖复杂的模型训练,能够快速生成多样化的训练数据,特别适用于资源有限的实际应用场景。

表 3 数据集加雨雪前后对比实验结果

Table 3 Comparison of dataset before and after adding

rain and snow effects

%

类别	加雨雾、雪前		加雨雾、雪后	
	P	R	P	R
pn	96.4	87.3	95.2	90.8
pne	96.8	84.6	96.7	92.5
p11	84.5	68.3	90.5	79.6
p26	86.3	63.6	89.0	81.0
i2	91.7	85.5	86.1	76.0
i4	93.7	82.4	92.0	86.5
i5	89.8	87.7	92.7	89.9
il60	92.3	97.9	95.9	98.9
pl5	75.1	62.5	85.6	71.5
pl30	78.9	55.2	89.3	81.7
pl40	80.5	67.2	86.6	80.2
pl50	87.1	62.6	86.7	84.7
pl60	74.0	67.1	83.0	73.6
pl80	77.3	74.0	86.4	83.4
pl100	88.1	79.3	88.1	90.4

2.5 特征融合网络对比实验

为了验证 HS-FPN 的有效性,本文将特征融合网络分别替换为 Slimneck^[20] 和 Bifpn^[21] 分别进行训练,对比实验结果如表 4 所示。

表 4 特征融合网络对比实验

Table 4 Feature fusion network comparison experiment

特征融合网络	mAP/%	参数量/M	FLOPs/G
RT-DETR-Bifpn	85.51	20.32	64.3
RT-DETR-Slimneck	85.91	19.32	53.3
RT-DETR-HS-FPN	86.68	18.13	53.4

由表 4 可知,添加了 HS-FPN 特征融合网络的模型表现最佳,其 mAP 达到了 86.68%,优于 Slimneck 的 85.91% 和 Bifpn 的 85.51%。同时,HS-FPN 的参数量最少,仅为 18.13 M,而 Slimneck 和 Bifpn 的参数量分别为 19.32 和 20.32 M。在计算复杂度方面,HS-FPN 与 Slimneck 的 FLOPs 接近,分别为 53.4 和 53.3 G,均优于 Bifpn 的 64.3 G。综合来看,HS-FPN 在保持较低的计算

资源消耗和参数量的同时,取得了较好的识别精度。

2.6 注意力机制对比实验

为了验证 Ortho 注意力机制的性能,在主干网络的 BasicBlock 模块的同一位置融合 ELA^[22], CA^[23], SE^[24] 等不同的注意力机制进行对比实验,实验结果如表 5 所示。

表 5 注意力机制对比实验

Table 5 Attention mechanism comparison experiment

模块	mAP/%	参数量/M	FLOPs/G
BasicBlock_SE	84.59	19.98	57.0
BasicBlock_CA	85.21	20.18	58.4
BasicBlock_ELA	85.64	24.98	58.9
BasicBlock_Ortho	86.33	19.98	57.0

由表 5 可知,添加了 Ortho 注意力机制的模型在性能上优于其他几种添加不同注意力机制的模型。具体来说,添加 Ortho 注意力机制的模型 mAP 为 86.33%,相比其他注意力机制的模型,平均提升了 1.18%。在参数量方面,添加了 Ortho 注意力机制的模型与添加 SE 注意力机制的模型参数量相同,均为 19.98 M,而添加 CA 注意力机制的模型参数量稍高,为 20.18 M,添加 ELA 注意力机制的模型参数量最高,为 24.98 M。在计算复杂度(floating point operations, FLOPs)上,添加了 Ortho 注意力机制和 SE 注意力机制的模型均为 57.0 G,而添加 CA 和 ELA 注意力机制的模型 FLOPs 分别为 58.4 和 58.9 G。综合来看,添加了 Ortho 注意力机制的模型在保持较低参数量和计算复杂度的同时,检测精度更高。

2.7 不同算法对比实验

为了验证改进的检测算法性能,将该模型与其他目标检测算法进行对比实验,实验均在相同的环境下使用相同的数据集,并在相同测试集上进行比较,结果如表 6 所示。由表可知,BH-RT-DETR 的 mAP 达到了 87.84%,远高于其他对比算法。该性能提升主要原因为在模型的主干网络中融合了 Ortho 注意力机制并且将 CCFM 特征融合网络替换为 HS-FPN 特征融合网络。两项改进增强模型对特征空间的全局上下文感知能力,有效提升了目标检测的精度,同时优化了多尺度特征融合过程,使得模型在处理小目标和复杂背景时具有更高的检测精度。然而,这些改进也导致 BH-RT-DETR 的参数量和计算复杂度有所增加。总体而言,实验结果表明,BH-RT-DETR 算法在交通标志检测任务中相比其他算法,在检测精度方面有明显提升,尽管其计算复杂度略有增加,但在实际应用中能够提供更精确的检测结果。

为了进一步验证本文算法在交通标志的检测中的性能优势,本文针对一些包含小目标和复杂背景的交通标志图像进行测试,对比实验结果如图 9 所示。

表 6 不同算法对比实验

Table 6 Comparison experiment of different algorithms

网络模型	mAP / %	参数量 / M	FLOPs / G
RetinaNet	68.31	36.62	168.4
CenterNet	56.16	32.68	109.7
YOLOXs	40.84	8.93	26.8
YOLOv5s	49.64	7.14	15.9
改进 YOLOv5s ^[25]	82.41	7.28	16.2
YOLOv7	59.41	37.36	105.4
YOLOv8s	50.18	11.13	28.5
YOLOv8m	54.24	25.85	78.7
YOLOv9	62.57	7.12	26.3
RT-DETR	85.47	19.89	57.0
BH-RT-DETR	87.84	18.22	53.4

该测试选取了 3 组交通标志图像进行对比。在第 1 组

图片中, 交通标志尺寸较小, 尤其是远距离的目标。YOLOv5s 和 YOLOv7 在小目标检测上出现了误检和漏检现象, 而 YOLOv8m 和 RT-DETR 一定程度上改善了这一问题, 但其置信度仍不及 BH-RT-DETR。第 2 组图片中, 部分目标被遮挡, 背景较为复杂, 干扰信息较多。YOLOv5s, YOLOv7 和 YOLOv8m 均出现漏检现象, 而 RT-DETR 模型出现多检问题。相比之下, BH-RT-DETR 能够准确识别目标。在第 3 组图片中, 包含多个交通标志, YOLOv5s, YOLOv7 和 YOLOv8m 均存在不同程度的漏检, 虽然 RT-DETR 能够准确识别所有交通标志, 但其置信度低于 BH-RT-DETR。BH-RT-DETR 的高检测精度和置信度得益于其采用的 Ortho 注意力机制和 HS-FPN 特征融合网络。这些改进显著提升了模型的全局特征感知能力和多尺度特征融合效果, 使得模型在复杂场景下仍能稳定且精准地完成检测任务。



图 9 BH-RT-DETR 与其他算法在不同场景下的交通标志检测结果

Fig. 9 BH-RT-DETR and other algorithms in traffic sign detection across different scenarios

3 结论

本文提出了一种基于 RT-DETR 的改进交通标志检测算法, 针对极端天气下的交通标志检测问题, 通过引入 Ortho 注意力机制和 HS-FPN 网络, 有效提升了模型在雨雾等恶劣环境中的检测精度, 尤其是在小目标和复杂背景下的目标检测能力。然而, 改进后的模型仍存在一些不足

之处: 首先, 尽管在复杂场景中的检测性能显著提升, 但模型的计算复杂度有所增加, 可能在实时性要求较高的自动驾驶场景中带来额外的计算负担。其次, 模型对某些类别的交通标志检测仍存在误检和漏检的现象, 特别是在极端光照变化条件下的表现有限。为解决这些问题, 后续工作可进一步优化模型结构, 减少计算成本, 并结合更强大的数据增强技术和自适应特征提取机制, 提高模型在极端天

气和复杂光照条件下的泛化能力。

参考文献

[1] ZHANG X. Traffic sign detection based on YOLOv3[C]. 2023 IEEE 3rd International Conference on Power, Electronics and Computer Applications (ICPE-CA), Shenyang, China. IEEE, 2023: 1044-1048.

[2] DHYANI S, KUMAR V. Multi-class traffic sign recognition system using one-stage detector YOLOv5s[C]. International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies (ViTECoN), Vellore, India. IEEE, 2023: 1-5.

[3] 闫世洋, 罗素云. 基于 SC-YOLOv8 的交通标志检测算法研究[J]. 电子测量技术, 2024, 47(15): 117-124.

YAN SH Y, LUO S Y. Research on traffic sign detection algorithm based on SC-YOLOv8 [J]. Electronic Measurement Technology, 2024, 47(15): 117-124.

[4] LI X, GENG S. Improved traffic sign detection algorithm for YOLOv5s[C]. International Conference on Computer Engineering and Application (ICCEA). IEEE, 2023: 696-699.

[5] REN SH Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.

[6] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. European Conference on Computer Vision (ECCV), Amsterdam, Netherlands. Springer, 2016: 21-37.

[7] DUAN K, BAI S, XIE L, et al. CenterNet: Keypoint triplets for object detection [C]. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea(South). IEEE, 2019: 6568-6577.

[8] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]. IEEE International Conference on Computer Vision, 2017: 2980-2988.

[9] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016: 779-788.

[10] REDMON J, FARHADI A. YOLOv3: An incremental improvement[J]. ArXiv preprint arXiv: 1804.02767, 2018.

[11] QIAO S, CHEN L C, YUILLE A. Detectors: Detecting objects with recursive feature pyramid and switchable atrous convolution [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 10213-10224.

[12] WANG C Y, BOCHKOVSKIY A, LIAO H Y. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2023: 7464-7475.

[13] GE ZH, LIU S, WANG F, et al. YOLOx: Exceeding YOLO series in 2021 [J]. ArXiv preprint arXiv: 2107.08430, 2021.

[14] LIU M, XU P, YANG S, et al. Traffic sign recognition under adverse weather using a deep fusion network with attention mechanisms[C]. 2024 IEEE/CVF International Conference on Computer Vision (ICCV), Paris, France. IEEE, 2024: 2043-2050.

[15] WANG Z X, LEI X M, ZHOU SH S. Traffic sign detection based on lightweight RT-DETR[C]. 2024 IEEE 2nd International Conference on Sensors, Electronics and Computer Engineering (ICSECE). IEEE, 2024: 1593-1597.

[16] CHEN Y, ZHANG C, CHEN B, et al. Accurate leukocyte detection based on deformable-DETR and multi-level feature fusion for aiding diagnosis of blood diseases[J]. ArXiv preprint arXiv: 170.107917, 2024.

[17] ZHAO Y, CHEN B, LIN J, et al. DETRs beat YOLOs on real-time object detection [J]. ArXiv preprint arXiv: 2304.08069, 2023.

[18] SALMAN H, PARKS C, SWAN M, et al. OrthoNets: Orthogonal channel attention networks[C]. 2023 IEEE International Conference on Big Data (BigData). IEEE, 2023: 829-837.

[19] 易磊, 黄哲玮, 易雅雯. 改进 YOLOv8 的输电线路异物检测方法[J]. 电子测量技术, 2024, 47(15): 125-134.

YI L, HUANG ZH W, YI Y W. Improved YOLOv8 method for foreign object detection on transmission lines[J]. Electronic Measurement Technology, 2024, 47(15): 125-134.

[20] OGUNTOLA I, BIRODKAR V, SHETH A. SlimNets: An exploration of deep model compression and acceleration[J]. ArXiv preprint arXiv: 1808.00496, 2018.

[21] TAN M, PANG R, LE Q V. EfficientDet: Scalable and efficient object detection [J]. ArXiv preprint arXiv: 1911.09070, 2020.

[22] XU W, WAN Y. ELA: Efficient local attention for deep convolutional neural networks [J]. ArXiv preprint arXiv: 2403.01123, 2024.

[23] HOU Q, ZHOU D, FENG J. Coordinate attention for

efficient mobile network design [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2021:13708-13717.

[24] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2018:7132-7141.

[25] GUO J B, YU L, YU CH Q. Application of improved YOLOv5s algorithm in traffic sign detection and recognition [J]. Journal of National University of Defense Technology, 2024, 46(6): 123-130.

作者简介

秦伦明,博士,讲师,硕士生导师,主要研究方向为目标

检测、语义分割、多源图像视频信息处理、电力设备在线监测。

E-mail:lunming.qin@shiep.edu.cn

张云起,硕士,主要研究方向为目标检测。

E-mail:yunqi202407@163.com

崔昊杨,博士,教授,博士生导师,主要研究方向为电力设备状态检测、红外探测与检测。

E-mail:cuihy@shiep.edu.cn

边后琴(通信作者),博士,副教授,主要研究方向为智能控制、计算机视觉和模式识别。

E-mail:bianhouqin@163.com

王悉,博士,教授,硕士生导师,主要研究方向为列车智能驾驶、模型预测控制。

E-mail:xwang@bjtu.edu.cn