

# 基于改进 YOLOv8s 的雾天目标检测算法<sup>\*</sup>

刘震 杨贤昭 陈洋 曾思航

(武汉科技大学冶金自动化与检测技术教育部工程研究中心 武汉 430081)

**摘要:** 针对现实场景中雾天目标检测困难的问题,提出了一种改进 YOLOv8s 的雾天目标检测方法。设计前端模块 Edge-Dehaze,利用去雾网络和检测网络联合训练并通过 Sobel 算子强化雾天图像边缘信息以改善雾天场景下的检测效果;提出混合注意力特征融合模块 HAFM,通过并行注意力机制和特征图之间的信息交互与融合提升模型对重要特征的关注度;设计轻量化共享注意力卷积检测头 LSACD,通过共享卷积降低检测头参数量,在共享层中引入 SEAM 注意力机制缓解雾天目标检测的遮挡问题。在 RTTS 数据集上的实验结果表明,改进后的 YOLOv8s 网络相对原始 YOLOv8s 网络  $mAP_{50}$  提升了 1.8%, $mAP_{50-95}$  提升了 1.7%,参数量基本持平,从而验证了该算法在雾天目标检测上具有较高的准确性和实用性。

**关键词:** YOLOv8s;计算机视觉;雾天目标检测;注意力机制;特征融合

**中图分类号:** TP391; TN911.73    **文献标识码:** A    **国家标准学科分类代码:** 520.604

## Object detection algorithm for foggy conditions based on improved YOLOv8s

Liu Zhen Yang Xianzhao Chen Yang Zeng Sihang

(Engineering Research Center for Metallurgical Automation and Measurement Technology of Ministry of Education,  
Wuhan University of Science and Technology, Wuhan 430081, China)

**Abstract:** To address the challenges of target detection in foggy conditions in real-world scenarios, this paper proposes an improved foggy target detection method based on YOLOv8s. The design includes a front-end module, Edge-Dehaze, which employs joint training of dehazing and detection networks and uses the Sobel operator to enhance edge information in foggy images, thereby improving detection performance in foggy environments. The proposed hybrid attention feature fusion module (HAFM) utilizes parallel attention mechanisms to enhance information interaction and fusion between feature maps, increasing the model's focus on critical features. Additionally, a lightweight shared attention convolutional detection (LSACD) head is designed, which reduces the parameter count of the detection head through shared convolutions and incorporates the SEAM attention mechanism in the shared layer to alleviate occlusion issues in foggy target detection. Experimental results on the RTTS dataset demonstrate that the improved YOLOv8s network achieves a 1.8% increase in  $mAP_{50}$  and a 1.7% increase in  $mAP_{50-95}$  compared to the original YOLOv8s network, with comparable parameter counts, thereby validating the high accuracy and practicality of the proposed method in foggy target detection.

**Keywords:** YOLOv8s; computer vision; foggy object detection; attention mechanisms; feature fusion

## 0 引言

随着计算机视觉的飞速发展恶劣天气下的图像处理已经成为该领域重要的研究方向<sup>[1]</sup>。雾天场景下的目标检测在自动驾驶、道路监控<sup>[2]</sup>等领域有着广泛的用途。常见的公开数据集和方法大多建立在能见度良好的条件之上,在

雾、霾等恶劣天气条件下拍摄的图像由于气溶胶粒子的吸收和散射作用会产生可视度低<sup>[3]</sup>、失真等退化现象。通用的目标检测方法在雾天环境下无法准确定位和识别出目标,性能会产生显著下降<sup>[4]</sup>,因此雾天环境下的目标检测研究意义重大<sup>[5]</sup>。

近年来,随着深度学习的不断发展,涌现出大量性能优

异的检测模型。目前基于深度学习的目标检测模型主要分为两大类,一类是以 Faster R-CNN<sup>[6]</sup>和 Retinanet<sup>[7]</sup>为代表的二阶段检测算法,先生成候选框,然后对每个框进行分类和位置回归。张骞等<sup>[8]</sup>提出了融合高分辨率网络的雾天目标检测算法,采用高分辨率网络 HRNet 作为特征提取网络,并行连接不同分辨率的子网,减少下采样过程中的信息损失,再通过 CIoU 损失函数和 SoftNMS 机制提升网络检测精度。二阶段算法在检测速度上较慢,且模型较大。另一类是以 YOLO<sup>[9-13]</sup>和 SSD<sup>[14]</sup>为代表的一阶段检测算法,无需生成候选框,直接对图像进行分类和位置回归。在雾天目标检测任务中,检测速度尤为重要,因此有学者直接在 YOLO 系列算法上进行改进研究得到速度快、精度高的模型。苏佳等<sup>[15]</sup>提出了基于维度交互和跨层尺度级联的雾天目标检测算法,使用合成雾图像扩充数据集,通过 TA 和 CLC-PAN 结构增强了模型对特征的捕获能力,结合 SIoU 损失函数提升模型的回归能力。朱磊等<sup>[16]</sup>提出了新型特征增强与融合的雾天目标检测方法 NFF-YOLOX,通过增强网络的特征提取能力有效提高了雾天目标的检测精度。谈爱玲等<sup>[17]</sup>提出了改进的 YOLOv8 模型,通过混合池化的 MixSPPF 和多尺度模块提升了雾天偏振图像中汽车和行人的检测准确率。这类算法并未考虑到雾的干扰对检测模型的影响,因此检测精度并不高。

若将去雾算法考虑到雾天目标检测任务中,解决方法可分为两大类:一是去雾和检测无关联模型,先对图像进行去雾处理,然后将预处理后的图像输入到目标检测模型中进行推理,如 Li 等<sup>[18]</sup>提出轻量级去雾网络 PDR-Net 与目标检测网络 Faster R-CNN 相结合。第一种方法在图像处理过程中会获得较高的图像恢复指标,但是较高的图像恢复指标却不一定对目标检测有利,因此这种方法并不总能提升检测精度。二是去雾和检测相关联模型,去雾与检测同时进行,如 Li 等<sup>[19]</sup>提出基于光学散射模型的 AOD-Net 用来去雾并与目标检测网络 Faster R-CNN 结合形成端到端的训练。Huang 等<sup>[20]</sup>提出了用于雾天目标检测的多任务网络 DSNet,利用两个子网来共同学习可视性增强和目标检测。解宇虹等<sup>[21]</sup>提出了基于先验知识引导的目标检测框架,通过先验知识和模型的联合优化提升雾天场景下的目标检测性能。Liu 等<sup>[22]</sup>针对恶劣天气条件提出可微分的图像处理模块 DIP,使用 YOLOv3 作为检测网络,通过端到端的学习增强雾天目标检测的性能。第 2 种方法虽然可以有效的提升雾天目标检测精度但是模型比较复杂且难以平衡去雾与检测网络权重。针对以上问题,为了进一步提升雾天目标检测的精度,本文提出改进 YOLOv8s 的雾天目标检测算法。主要工作如下:

1)设计前端模块 Edge-Dehaze,采用去雾检测联合优化的策略并显式地提取图像的边缘信息,使模型能更好的应对雾天场景下的目标检测任务。

2)提出了一种混合注意力特征融合模块(hybrid

attention feature fusion module, HAFM),通过注意力机制和增强特征之间的信息交互与融合,使模型具备更强的特征表达能力。

3)设计了轻量化共享注意力卷积检测头(lightweight shared attention convolutional detection, LSACD),通过参数共享的思想降低检测头的参数量,同时在共享层中引入 SEAM 注意力机制增强了对遮挡物体的检测性能。

采用改进后的网络模型在 RTTS<sup>[23]</sup> 数据集上进行测试,并与主流检测模型进行比较,以证明模型的优越性。

## 1 改进 YOLOv8s 网络结构设计

YOLOv8s 是一种轻量化的单阶段检测算法,是目前最先进的检测算法之一,适用于高帧率视频实时检测。其网络结构包括输入端、Backbone 骨干网络(采用 CBS 结构、C2f 结构、SPPF 结构)、Neck 网络(PAN + FPN 结构)、Head 网络(3 个 Detect 检测器,采用解耦头设计,在不同尺度特征图上进行目标检测)。预处理阶段将图像缩放到  $640 \times 640$  大小并进行归一化,训练时采用 Mosaic 数据增强策略。Backbone 提取特征,Neck 融合浅层和深层信息,Head 在不同尺度特征图上预测小、中、大 3 类目标。

由于雾天目标检测有环境复杂、图像分辨率低和严重遮挡等特点,为了使 YOLOv8s 模型更契合雾天目标检测任务,本文对 YOLOv8s 模型进行改进,设计 Edge-Dehaze 前端模块以适应雾天复杂的环境特点、提出 HAFM 特征融合模块以增强特征表示的有效性、设计 LSACD 检测头提升模型对遮挡目标的检测能力。改进后的 YOLOv8s 模型结构如图 1 所示。

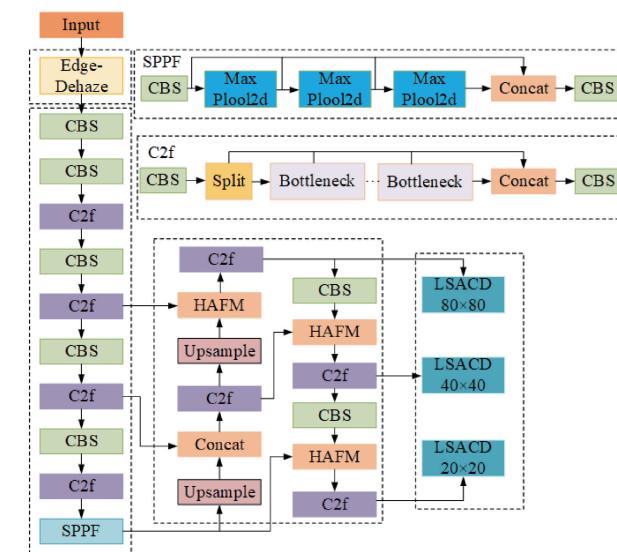


图 1 雾天目标检测模型

Fig. 1 Target detection model in foggy conditions

### 1.1 Edge-Dehaze 模块

直接在雾天图像上训练的模型会难以适应雾天场景中

待检测目标与雾之间的相互作用,导致模型检测精度下降。对雾天场景进行去雾处理,虽然会获得较高的图像恢复指标,但是图像去雾却不一定有益于目标检测,DSNet 的实验也证实了这一点。为了提高真实雾天场景中的检测精度,本文设计前端模块 Edge-Dehaze,考虑到雾天图像整体对比度降低,目标和背景区别不明显,通过增强边缘信息提升目标与背景之间的对比度,并采用联合训练优化策略将去雾与检测网络相结合,使用检测损失函数对整体网络进行训练,去雾网络将不再致力于提升图像恢复指标而是服务于检测任务。有雾图像或无雾图像及其目标检测标签是整个网络的输入,经过前端模块 Edge-Dehaze 进行去雾过程的学习和图像边缘信息的提取,接着经过 YOLOv8s 检测网络进行检测过程的学习,整个网络的最终输出为目标的生成边框和类别预测等信息。

Edge-Dehaze 模块如图 2 所示,分为边缘信息提取和图像去雾两个分支。边缘信息提取采用 Sobel 算子实现,图像去雾采用轻量级去雾网络 AOD-Net 实现。

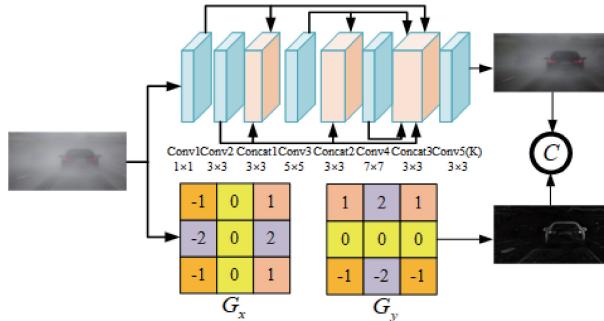


图 2 Edge-Dehaze 模块

Fig. 2 Edge-Dehaze module

在雾霾条件下,虽然图像整体对比度降低,但边缘区域的梯度信息仍然相对明显,通过提取和增强这些边缘信息,可以提高目标与背景之间的对比度,使得物体的轮廓更加清晰。雾霾导致的模糊效应会使得图像中的细节丢失,但边缘信息通常保留了一些关键的结构特征,这些结构特征能够帮助模型区分不同的物体。例如,依据车辆的轮廓等边缘信息依然也可以分辨出不同的车辆。

边缘信息的提取通过 Sobel 算子实现,Sobel 算子是一种离散微分算子,用于计算图像的梯度,从而实现边缘检测。其卷积核表达形式如图 2 下分支所示。 $G_x$  用于计算水平方向的梯度, $G_y$  用于计算垂直方向的梯度。通过将这两个卷积核分别与图像进行卷积操作,可以得到图像在水平方向和垂直方向的梯度图像。根据式(1)可以计算出图像的边缘强度。

$$G = \sqrt{G_x^2 + G_y^2} \quad (1)$$

AOD-Net 算法是基于简化的大气散射模型设计的端到端去雾算法,用于去雾并恢复清晰的图像。在这个简化模型中,无雾图像  $J(x)$  和其有雾图像  $I(x)$  可表示为式(2):

$$I(x) = J(x)t(x) + A(1 - t(x)) \quad (2)$$

其中, $t(x)$  是透射率, $A$  为大气光值。由于  $t(x)$  和  $A$  的单独估计会累积甚至放大误差,将其统一为一个公式,并直接最小化像素域重建误差,重写式(2):

$$K(x) = \frac{A - 1 + \frac{(I(x) - A)}{t(x)}}{I(x) - 1} \quad (3)$$

$$J(x) = K(x)I(x) - K(x) + 1 \quad (4)$$

AOD-Net 中包含 K-estimation 模块和清晰图像生成模块。K-estimation 模块采用多尺度特征融合网络提取特征,如图 2 上分支所示,该网络中有 5 个卷积层,通过特征融合模块将不同卷积层的输出特征图融合,最后一层卷积层用来估计  $K(x)$ ,清晰图像生成模块根据式(4)恢复清晰无雾图像。最后将 Sobel 算子提取到的边缘信息图像与去雾模块处理后的图像进行融合得到 Edge-Dehaze 模块的输出。

## 1.2 HAFM 混合注意力特征融合模块

雾天目标检测的难点主要在于雾霾对图像质量的严重影响,包括对比度降低、颜色失真和细节模糊,导致网络提取了低质量特征,从而增加了检测的难度。受到雾霾的影响,浅层特征中会包含大量噪声等无关信息,图像的边缘和纹理信息也会受到干扰,深层特征难以准确捕获到目标的高层语义信息<sup>[24]</sup>,降低目标检测的准确性。为了提高模型的特征表达能力,本文提出 HAFM 混合注意力特征融合模块改进 YOLOv8s 的特征金字塔网络,模型结构如图 3 所示。

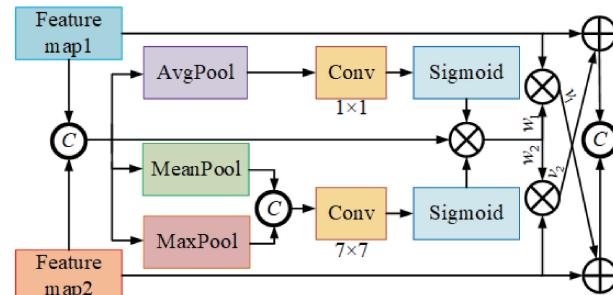


图 3 HAFM 模块

Fig. 3 HAFM module

具体实现步骤如下:

第 1 部分是通过混合注意力机制捕获特征图中的有效信息。将检测网络中提取到的浅层特征和深层特征拼接起来,然后将其输入到混合注意力网络中以生成注意力权重。通道注意力如式(6)所示,空间注意力如式(7)所示。生成的权重通过元素乘法对特征图进行加权,如式(8)所示。

$$F = C(F_1, F_2) \quad (5)$$

$$W_c = \sigma(f^{1 \times 1}(\text{AvgPool}(F))) \quad (6)$$

$$W_s = \sigma(f^{7 \times 7}(C(\text{MeanPool}(F), \text{MaxPool}(F)))) \quad (7)$$

$$F_{\text{weight}} = F \otimes W_c \otimes W_s \quad (8)$$

式中:  $F_1$  和  $F_2$  为输入特征图,  $C$  为特征图的拼接操作,  $\sigma$  为 Sigmoid 激活函数,  $f^{1 \times 1}$  和  $f^{7 \times 7}$  分别代表卷积核大小为  $1 \times 1$  和  $7 \times 7$  的卷积操作,  $AvgPool$ 、 $MeanPool$ 、 $MaxPool$  分别代表全局平均池化、平均池化和最大池化。 $\otimes$  代表元素乘法。

第 2 部分通过特征分割与重新分配增强特征之间的交互性,将  $F_{weight}$  按照通道数分割为  $w_1$  和  $w_2$ , 根据式(9)得到最后的输出。重新分配和融合的特征图能够互相补充和增强,使得最终的输出特征图有更丰富的特征信息。

$$F_{out} = C(F_1 + F_2 \otimes w_2, F_2 + F_1 \otimes w_1) \quad (9)$$

HAFM 混合注意力特征融合模块结合了通道注意力和空间注意力,通过并行处理的方式增强不同特征图之间的信息交互和融合。通道注意力机制能够突出每个通道的重要性,使得重要通道的信息得到加强。空间注意力机制则能够在空间维度上强调关键位置,使得特征图中的重要区域得到突出。通过对输入特征图的通道和空间维度进行注意力加权,提升模型对重要特征的关注度,并抑制无关或冗余特征。该模块强调了特征图之间的信息交互,在特征融合过程中捕捉并利用重要的上下文信息,从而增强特征表示的有效性。

### 1.3 LSACD 检测头

检测头对比如图 4 所示。YOLOv8s 的每个检测层都有独立的卷积模块进行特征提取,如图 4(a)所示,以网络最后 3 个 C2f 层的输出作为输入,采用解耦头设计分为 Cls 分类和 Box 回归两个分支,分别通过 2 个  $3 \times 3$  的卷积和 1 个  $1 \times 1$  的卷积来提取信息,其参数量几乎占据了整个模型的  $1/5$ 。本文设计了 LSACD 检测头,采用参数共享的思想降低了检测头的参数量。在雾天目标检测任务中目标个体之间的遮挡会导致局部重叠和重要特征缺失,导致检测器无法准确识别目标,在轻量化检测头的基础上在共享层中引入 SEAM 注意力模块<sup>[25]</sup>机制可突出图像中的待检测目标区域,弱化背景区域,实现多尺度下的雾天目标检测。

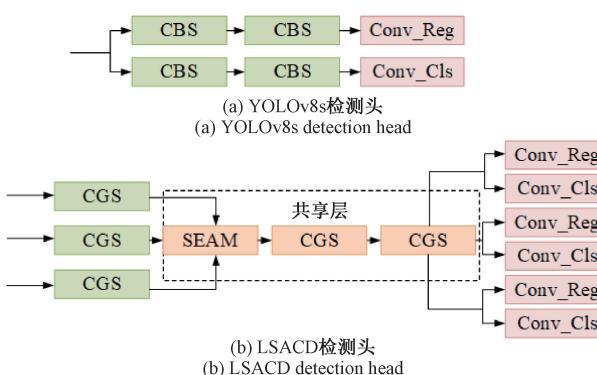


Fig. 4 Comparison of detection head

LSACD 检测头模型如图 4(b)所示,将多层特征图的卷积操作共享,使得同一组卷积核可以处理不同层次的特

征图从而到达减少参数量的目的。每个特征图先通过  $1 \times 1$  的 CGS 卷积调整输出通道数,在共享层中先通过 SEAM 注意力模块再通过两个  $3 \times 3$  的 CGS 卷积提取特征,最后通过  $1 \times 1$  的卷积输出分类和回归结果。CGS 相较于 CBS,将 BN 层替换为了 GN 层<sup>[26]</sup>,GN 层不依赖与批次大小,而是将特征图换分为多个组并在组内进行归一化,在小批量训练时能保持稳定的归一化效果。

SEAM 注意力模块如图 5 所示,其第 1 部分是结合了残差连接的深度可分离卷积。深度可分离卷积的设计目标是通过逐通道操作减少模型参数量,学习不同通道的特征,增强模型的特征提取能力。为了补充通道间的特征联系,通过逐点卷积对不同深度卷积的输出进行组合,然后通过一个两层的全连接网络融合各通道的特征,实现了所有通道之间的连接。第 2 部分是将全连接层的输出结果通过指数函数进行处理,将取值范围从  $[0, 1]$  扩大到  $[1, e]$ ,使得上一步学习到的被遮挡雾天目标的有效特征与无遮挡的有效特征相互联系,最后将 SEAM 模块的输出作为注意力权重与原始特征相乘,使模型能够有效地缓解雾天目标个体间遮挡带来的有效特征丢失的问题。

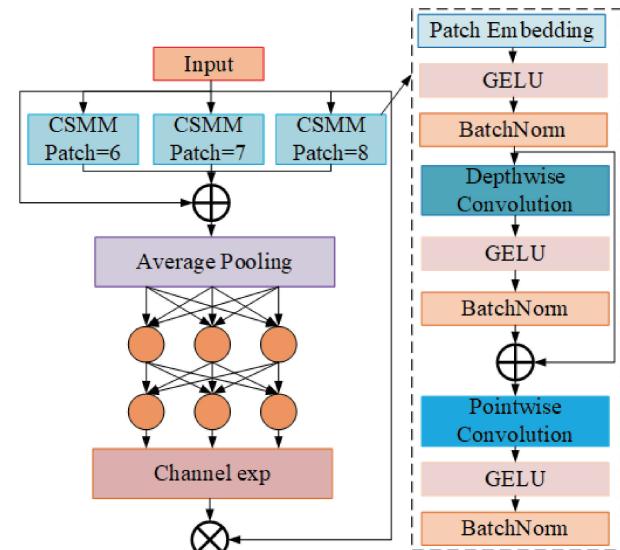


图 5 SEAM 模块

Fig. 5 SEAM module

## 2 分析与讨论

### 2.1 实验平台设置

本实验是在 Pytorch 框架下完成的,开发语言为 Python。模型的训练环境为 Intel Core i7-11700 八核 CPU,运行内存为 16 GB, GPU 为 NVIDIA GeForce RTX3080(10 GB),操作系统为 Win10 系统。Python 版本为 3.8.19,Pytorch 版本为 1.13.1,CUDA 版本为 11.7。模型训练参数设置为:Batch-Size 设为 16,初始学习率为 0.01,权重衰减为 0.0005,使用 SGD 梯度优化器,共迭代

180 个 epoch。

## 2.2 真实世界雾天目标检测数据集

所提出的模型在 RESIDE 数据集的子类数据集 RTTS 上进行评估:RTTS 数据集是雾天条件下的真实世界数据集,有 4 322 幅标注过的自然雾天图像,包含 5 类别即:人、自行车、汽车、公共汽车和摩托车。该数据集主要有两个特点:一是环境和场景复杂,需要适应各种能见度条件下的雾天环境,并且城市和乡村等场景差异较大;二是样本不均衡,类别包含 5 类,各类样本数量严重不均衡,每张图像中包含多类目标,以及各种程度的遮挡。数据集按照 7:1:2 的比例进行训练集、测试集和验证集的划分。

为了衡量模型的性能,采用通用目标检测评价指标:准确率  $P$  (Precision)、平均精度均值  $mAP_{50}$ 、 $mAP_{50-95}$ 。准确率  $P$  指的是模型在所有预测为正类别的样本中,实际为正类别的比例。 $TP$  (True Positives) 表示模型正确地检测到了正样本,  $FP$  (False Positives) 表示模型错误地将负样本预测为正样本,  $FN$  (False Negatives) 表示模型错误地将正样本预测为负样本。计算公式为:

$$P = \frac{TP}{FP + TP} \quad (10)$$

$$R = \frac{TP}{FN + TP} \quad (11)$$

$$AP = \int_0^1 P(R) dR \quad (12)$$

$IoU$  指的是模型检测结果与真实目标边界框之间重叠程度,  $mAP_{50}$  是指将  $IoU$  设为 0.5 时,计算每一类的  $AP$ ,然后所有类别求平均,即  $mAP$ 。 $mAP_{50-95}$  表示在不同  $IoU$  阈值(从 0.50 到 0.95,步长 0.05)上的平均  $mAP$ 。另外,采用参数量来衡量模型的复杂度,利用 Detection Time 评估模型的推理速度,一般 Detection Time 越小模型更符合实际应用需求。

## 2.3 实验结果与分析

为了验证 Edge-Dehaze 模块对雾天目标检测的有效性,针对该模块中设计 4 组实验:直接检测、去雾和检测无关联、去雾和检测网络联合训练、带边缘信息增强的去雾和检测网络联合训练。实验结果如表 1 所示,去雾和检测网络联合训练性能均较好,  $mAP_{50-95}$  有明显提升。在联合训练中,AOD-Net 和 YOLOv8s 可以同时学习并优化其特征提取能力,网络参数更新是协调的,这种联合训练策略比单独训练每个网络更有效。引入边缘信息提取模块后,模型的性能也有上升,说明了通过增强图像的边缘信息可以使模型更加有效的应对雾天目标检测。实验结果还表明先对雾天图像进行去雾处理后进行检测的效果并没有直接检测的效果好,在 AOD-Net 处理过程中,优化的是去除雾霾的图像清晰度,而不是优化对目标检测有利的特征,可能会丢失对目标检测网络来说的重要特征信息,验证了图像去雾不一定有益于目标检测的观点。

表 1 Edge-Dehaze 对目标检测的影响结果对比

Table 1 Comparison of the results of the effect of Edge-Dehaze on target detection

| 模型                  | P           | $mAP_{50}$  | $mAP_{50-95}$ % |
|---------------------|-------------|-------------|-----------------|
| YOLOv8s             | 80.0        | 72.9        | 47.9            |
| 去雾检测无关联             | 80.1        | 71.5        | 47.3            |
| 去雾检测相关联             | 80.3        | 73.0        | 48.3            |
| Edge-Dehaze+YOLOv8s | <b>80.4</b> | <b>73.1</b> | <b>48.5</b>     |

为了更直观说明 Edge-Dehaze 模块对雾天目标检测的有效性,对该模块的检测效果进行可视化,实验结果如图 6 所示。绿框代表正确检测,红框代表误检,蓝框代表漏检。图 6(a)直接使用 YOLOv8s 检测时未检测到背景中远景处的两个人,图 6(b)对图像先去雾后检测时不仅未检测到背景区域的两个人还出现了一个误检区域,图 6(c)使用去雾和检测相关联模型时仅漏检了一个目标,图 6(d)使用本文的改进算法完整的检测到所有目标且没有误检。观察图 6(b)可以发现,去雾后的图像尽管在视觉上更符合人眼的习惯,但是图像的整体亮度会降低,导致图像背景区域的模糊像素点更加明显,从而造成模型的误检。

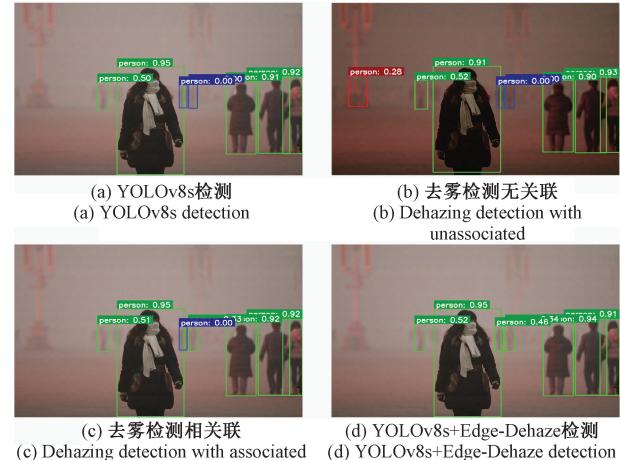


图 6 联合训练模型的检测效果

Fig. 6 Detection performance of the joint training model

在 LSACD 检测头实验中,第 1 部分对检测头进行轻量化,第 2 部分引入 SEAM 以改善遮挡目标的检测性能。其实验结果如表 2 所示。YOLOv8s 为未改进的常规检测头,A 代表使用轻量化检测头,B 代表引入 SEAM 注意力。该模块仅对检测头进行改进,因此使用 Para\_detect 更为直观的表示检测头的参数量。使用共享卷积的思想轻量化检测头后,检测头的参数量降低了大约 80%, $P$  指标提高了 1.5%, $mAP_{50}$  提高了 0.4%, $mAP_{50-95}$  指标几乎没有影响。在共享层中引入 SEAM 注意力后,仅增加少量参数, $P$  指标提高了 0.3%, $mAP_{50}$  提高了 0.6%, $mAP_{50-95}$  提高了 0.5%,说明了在检测头中加入注意力机制的有效性。

表 2 LSACD 检测头对目标检测的影响结果对比  
Table 2 Comparison of the results of the effect of LSACD on target detection

| 模型          | P / %       | mAP <sub>50</sub> / % | mAP <sub>50-95</sub> / % | Para_detect    |
|-------------|-------------|-----------------------|--------------------------|----------------|
| YOLOv8s     | 80.0        | 72.9                  | 47.9                     | 2 117 983      |
| YOLOv8s+A   | 81.5        | 73.3                  | 48.0                     | <b>419 800</b> |
| YOLOv8s+A+B | <b>81.8</b> | <b>73.9</b>           | <b>48.5</b>              | 440 152        |

为了进一步说明模型在遮挡情况下的检测性能,在验证集中选取 2 幅不同遮挡的雾天图像进行实验,使用 YOLOv8s 和 YOLOv8s\_LSACD 模型进行检测。结果如图 7 所示,在图 7(a)中由于绿化带和出租车的轻微遮挡导致没有检测到马路对面的车辆,在图 7(c)中,左右停车位上的车辆相互遮挡,导致未检测到左边最后一辆车。在 YOLOv8s\_LSACD 模型中,正确的检测到了图像中的所有待检测目标,说明模型能有效处理雾天目标检测中的遮挡问题。

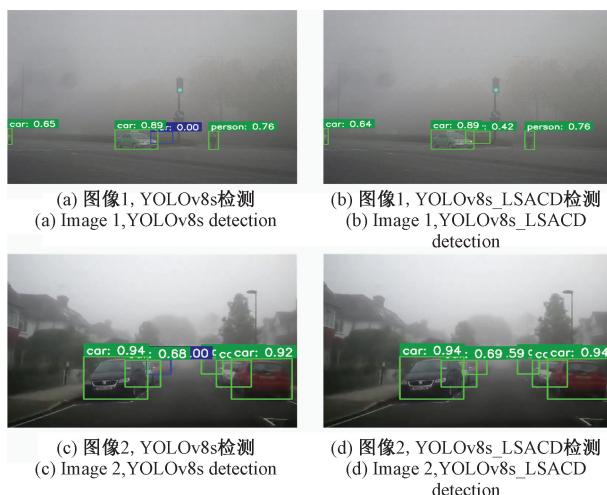


图 7 不同遮挡下的检测效果

Fig. 7 Detection performance under different occlusions

HAFM 特征融合模块通过注意力机制和加强浅层特征与深层特征的交互性,提升模型对重要特征的关注度。使用热力图的形式说明该模块的有效性,Grad-CAM<sup>[27]</sup>利

用网络反向传播的梯度计算出特征图每一个通道的权重得到热力图,可将模型对雾天待检测目标特征的学习程度以热力图的方式显示。分别对 YOLOv8s 模型和 YOLOv8s\_HAFM 模型进行热力图可视化,实验结果如图 8 所示,图 8(a)中部分车辆并未完全被热力区域覆盖且模型未关注到图像右侧和远景处的车辆,说明模型对这些目标的特征提取不够充分,存在漏检的风险。图 8(b)中待检测目标所在的区域大多被热力区域覆盖且热力值较高,说明 HAFM 模块充分利用了不同层次的特征信息,提高了特征表示的丰富性和准确性,使得特征提取更加精确,减少了漏检的可能性,从而提升了模型在雾天条件下目标检测的可靠性。

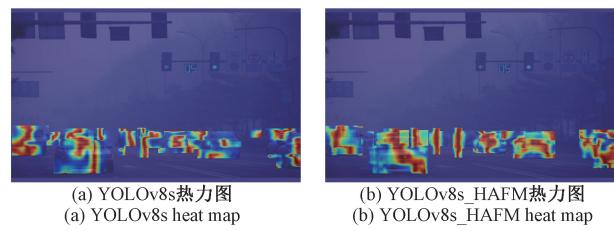


图 8 热力图效果对比  
Fig. 8 Comparison of heat map effect

为了验证改进算法的有效性,在 RTTS 数据集上进行了消融实验,评估 Edge-Dehaze 前端模块、HAFM 特征融合模块、LSACD 检测头模块对雾天目标检测模型性能的影响,实验结果如表 3 所示。BaseLine 为 YOLOv8s 基线,单独使用 Edge-Dehaze, P 提升 0.4%, mAP<sub>50</sub> 提升 0.2%, mAP<sub>50-95</sub> 提升 0.6%, 推理时间仅增加 0.1 ms。单独使用 HAFM, P 提升 0.6%, mAP<sub>50</sub> 提升 0.7%, mAP<sub>50-95</sub> 提升 0.8%, 推理时间不变。单独使用 LSACD, P 提升 1.8%, mAP<sub>50</sub> 提升 1.0%, mAP<sub>50-95</sub> 提升 0.5%, 推理时间不变。验证了本文提出的算法在真实雾天场景下的目标检测任务中的有效性。观察表中的推理时间,在单独应用 Edge-Dehaze 时推理时间会明显增加,说明本文所提模型主要延迟反映在 Edge-Dehaze 模块上。在组合使用模块时,LSACD 模块也会导致推理时间的少量增加,相比之下,Edge-Dehaze 模块对检测时间的影响更为显著。最终本文所提模型较基线, P 提升 2.9%, mAP<sub>50</sub> 提升 1.8%, mAP<sub>50-95</sub> 提升 1.7%, 推理时间增加了 0.2 ms。

表 3 在 RTTS 数据集上测试每个模块的实验结果

Table 3 Experimental results for testing each module on the RTTS dataset

| 基线 | Edge-Dehaze | HAFM | LSACD | P / %       | mAP <sub>50</sub> / % | mAP <sub>50-95</sub> / % | Detection Time / ms |
|----|-------------|------|-------|-------------|-----------------------|--------------------------|---------------------|
| ✓  |             |      |       | 80.0        | 72.9                  | 47.9                     | <b>1.8</b>          |
| ✓  | ✓           |      |       | 80.4        | 73.1                  | 48.5                     | 1.9                 |
| ✓  |             | ✓    |       | 80.6        | 73.6                  | 48.7                     | <b>1.8</b>          |
| ✓  |             |      | ✓     | 81.8        | 73.9                  | 48.5                     | <b>1.8</b>          |
| ✓  | ✓           | ✓    |       | 82.8        | 74.0                  | 48.9                     | 1.9                 |
| ✓  | ✓           |      | ✓     | 82.0        | 73.2                  | 49.1                     | 2.0                 |
| ✓  |             | ✓    | ✓     | 81.9        | 74.2                  | 49.0                     | 1.9                 |
| ✓  | ✓           | ✓    | ✓     | <b>82.9</b> | <b>74.7</b>           | <b>49.6</b>              | 2.0                 |

为了验证改进算法的先进性,对改进模型与主流模型在 RTTS 数据集上进行了对比试验,具体实验结果如表 4 所示。两阶段目标检测算法 Faster\_rcnn 和 Retinanet 在该数据集上表现较差,精度较低的同时参数量也较大。在单阶段算法中,相较于 SSD 算法,改进后的 YOLOv8s 算法在  $mAP_{50}$  和  $mAP_{50-95}$  上分别高出 2.3% 和 4.4%,并且参数量也更少。与 YOLO 系列算法比较,在检测精度方面,改进后的 YOLOv8s 模型  $mAP_{50}$  和  $mAP_{50-95}$  分别比 YOLOv3 高 2.7%、9.2%,比 YOLOv4s 高 2.3%、4.9%,比 YOLOv5s 高 4.4%、6.3%,比 YOLOv5s-Fog<sup>[28]</sup> 高 1.1%、3.9%,比 YOLOv6s 高 4.3%、5.6%,比 YOLOv7-tiny 高 3.1%、7.0%,比 YOLOv8s 高 1.8%、1.7%,比 YOLOv10s 高 2.5%、2.1%。在参数量方面,改进后的 YOLOv8s 模型比 YOLOv3 和 YOLOv6s 小,与 YOLOv8s 模型相近。

表 4 先进目标检测模型在 RTTS 数据集上的性能对比

Table 4 Performance comparison of advanced target detection models on RTTS datasets

| 模型               | $mAP_{50}$ / % | $mAP_{50-95}$ / % | Para/M     |
|------------------|----------------|-------------------|------------|
| Faster_rcnn_r101 | 60.6           | 32.9              | 60.1       |
| Retinanet_r101   | 67.6           | 41.9              | 55.2       |
| SSD              | 72.4           | 45.2              | 25.0       |
| YOLOv3           | 72.0           | 40.4              | 61.6       |
| YOLOv4s          | 72.4           | 44.7              | 9.1        |
| YOLOv5s          | 70.3           | 43.3              | 7.2        |
| YOLOv5s-Fog      | 73.6           | 45.7              | 26.9       |
| YOLOv6s          | 70.4           | 44.0              | 17.2       |
| YOLOv7-tiny      | 71.6           | 42.6              | <b>6.0</b> |
| YOLOv8s          | 72.9           | 47.9              | 11.1       |
| YOLOv10s         | 72.2           | 47.5              | 8.0        |
| Ours             | <b>74.7</b>    | <b>49.6</b>       | 11.1       |

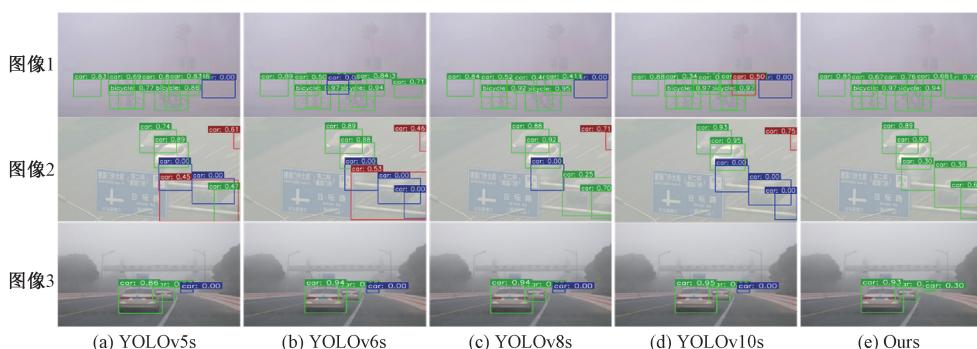


图 10 检测效果对比图

Fig. 10 Detection performance comparison chart

为了更直观的体现出本文所提模型在 YOLO 系列算法中的优势,训练过程中的  $mAP_{50-95}$  对比如图 9 所示,在 RTTS 数据集上,YOLOv8s 模型和改进的 YOLOv8s 模型均表现较好,在整个曲线上都优于其他模型。在曲线后半段,改进的 YOLOv8s 模型相较于 YOLOv8s 也有明显优势。结果表明本文提出的改进算法相比于其他主流检测模型在检测精度上有较大提升,  $mAP_{50-95}$  分别达到最优的 49.6%。

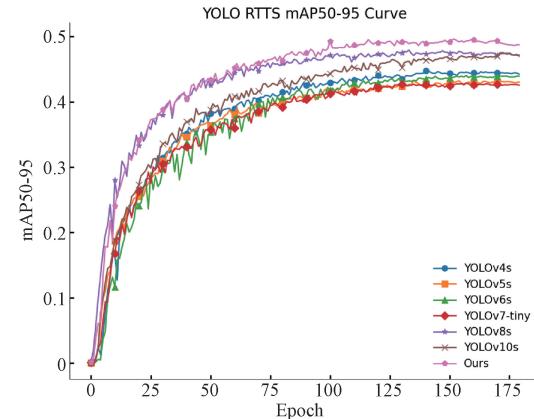


图 9 YOLO 算法对比图

Fig. 9 Comparison of YOLO algorithms

图 10 展示了 YOLOv4s、YOLOv5s、YOLOv6s、YOLOv8s、YOLOv10s 和改进后的 YOLOv8s 模型在检测效果上的对比。选取 4 种不同应用场景下的图像进行测试:在图像 1 中,即使存在雾霾严重干扰,本文所提模型仍然能完整检测到图像中目标;在图像 2 中,部分车辆模糊且存在路牌的遮挡,本文所提模型完整检测了全部目标且未对图像边缘区域产生误检;在图像 3 中,本文所提模型针对远景处的小目标车辆也有较好的检测效果。上述实验表明在复杂的实际应用场景中,改进后的 YOLOv8s 模型表现出更高的准确性和鲁棒性,具有更高的实际应用价值。

### 3 结论

针对雾天目标检测易受雾气干扰检测精度低的问题,本研究提出了一种改进 YOLOv8s 的轻量化雾天目标检测算法。设计前端模块 Edge-Dehaze,增强去雾与检测任务关联性并强化边缘特征以增强模型在雾天场景中的检测效果;提出 HAFM 混合注意力特征融合模块,通过注意力机制和加强浅层特征与深层特征的交互性,提升模型对重要特征的关注度;设计 LSACD 检测头,通过参数共享的思想降低检测头复杂度,在共享层中引入 SEAM 注意力机制缓解雾天目标检测中的遮挡问题。本文所提方法可有效缓解雾天目标检测效果差的难题,mAP<sub>50</sub> 和 mAP<sub>50-95</sub> 分别达到 74.7% 和 49.6%,模型参数量与 YOLOv8s 相当。在今后的工作中,将聚焦于雨雪等恶劣天气对目标检测的影响问题,并构建适应复杂环境的目标检测模型。

### 参考文献

- [1] 朱开,陈慈发. 基于 YOLOv5 的雾霾天气下交通标志识别[J]. 电子测量技术,2023,46(8):31-37.
- ZHU K, CHEN C F. Traffic sign recognition under fog weather based on YOLOv5 [J]. Electronic Measurement Technology,2023,46(8):31-37.
- [2] 熊李艳,涂所成,黄晓辉,等. 基于 MobileVit 轻量化网络的车辆检测方法[J]. 计算机应用研究,2022,39(8):2545-2549.
- XIONG L Y, TU S CH, HUANG X H, et al. Vehicle detection method based on MobileVit lightweight network [J]. Application Research of Computers, 2022,39(8):2545-2549.
- [3] 彭熙舜,陆安江,龙纪安,等. 结合图像分层与暗通道的雾霾图像增强[J]. 电子测量技术,2022,45(2):123-128.
- PENG X SH, LU AN J, LONG J AN, et al. Haze image enhancement combining image layering and dark channel [J]. Electronic Measurement Technology, 2022,45(2):123-128.
- [4] 王子昭,景明利,史金钢,等. 一种改进 CBAM 机制和细节恢复的单幅图像去雾算法[J]. 电子测量技术,2023,46(2):161-168.
- WANG Z ZH, JING M L, SHI J G, et al. A single image defogging algorithm based on improved CBAM mechanism and detail recovery [J]. Electronic Measurement Technology,2023,46(2):161-168.
- [5] 李任斯,石蕴玉,刘翔,等. 基于 Double-Head 的雾天图像目标检测[J]. 液晶与显示, 2023, 38 (12): 1717-1727.
- LI R S, SHI Y Y, LIU X, et al. Object detection in foggy image based on Double-Head [J]. Chinese Journal of Liquid Crystals and Displays,2023,38(12): 1717-1727.
- [6] 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.
- [7] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2):318-327.
- [8] 张骞,陈紫强,孙宗威,等. 融合高分辨率网络的雾天目标检测算法[J]. 计算机工程与科学,2023,45(11):1970-1981.
- ZHANG Q, CHEN Z Q, SUN Z W, et al. A fog target detection algorithm fusing high-resolution network[J]. Computer Engineering and Science,2023, 45(11):1970-1981.
- [9] REDMON J, FARHADI A. YOLOv3: An incremental improvement [EB/OL]. (2018-04-08) [2024-06-25]. <https://doi.org/10.48550/arXiv.1804.02767>.
- [10] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection[J]. ArXiv preprint arXiv. 2020. 10934,2020.
- [11] LI CH Y, LI L L, JIANG H L, et al. YOLOv6: A single-stage object detection framework for industrial applications [EB/OL]. (2020-09-07) [2024-06-25]. <https://doi.org/10.48550/arXiv.2209.02976>.
- [12] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. 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, 2023:7464-7475.
- [13] WANG AO, CHEN H, LIU L H, et al. YOLOv10: Real-time end-to-end object detection[EB/OL]. (2024-05-23) [2024-06-25]. <https://doi.org/10.48550/arXiv.2405.14458>.
- [14] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multiBox detector[C]. Computer Vision-ECCV 2016,2016.
- [15] 苏佳,梁奔,冯康康,等. 基于维度交互和跨层尺度级联的雾天目标检测算法[J]. 微电子学与计算机,2024,41(1):53-62.
- SU J, LIANG B, FENG K K, et al. Object detection algorithm in foggy weather based on dimensional interaction and cross-layer scale cascade [J]. Microelectronics & Computer,2024,41(1):53-62.

- [16] 朱磊, 赵涵, 王伟丽. 基于新型特征增强与融合的雾天目标检测方法 [J]. 西安工程大学学报, 2023, 37(6):106-113.  
ZHU L, ZHAO H, WANG W L. An object detection method in foggy weather based on novel feature enhancement and fusion [J]. Journal of Xi'an Polytechnic University, 2023, 37(6):106-113.
- [17] 谈爱玲, 李晓航, 赵勇, 等. 基于偏振成像和 YOLOv8 的雾天道路目标检测 [J]. 计量学报, 2024 (11): 1626-1634.  
TAN AI L, LI X H, ZHAO Y, et al. Road object detection in foggy weather based on polarization imaging and YOLOv8 [J]. Acta Metrologica Sinica, 2024(11):1626-1634.
- [18] LI CH Y, GUO CH L, GUO J CH, et al. PDR-Net: Perception-inspired single image dehazing network with refinement [J]. IEEE Transactions on Multimedia, 2020, 22(3):704-716.
- [19] LI B Y, PENG X L, WANG ZH Y, et al. AOD-Net: All-in-one dehazing network [C]. IEEE International Conference on Computer Vision, 2017:4770-4778.
- [20] HUANG S C, LE T H, JAW D W. DSNet: Joint semantic learning for object detection in inclement weather conditions [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43 (8): 2623-2633.
- [21] 解宇虹, 谢源, 陈亮, 等. 真实有雾场景下的目标检测 [J]. 计算机辅助设计与图形学学报, 2021, 33(5):733-745.  
XIE Y H, XIE Y, CHEN L, et al. Object detection in real misty scenes [J]. Journal of Computer Aided Design and Graphics, 2021, 33(5):733-745.
- [22] LIU W Y, REN G F, YU R SH, et al. Image-adaptive YOLO for object detection in adverse weather conditions [C]. AAAI Conference on Artificial Intelligence, 2022, 36(2):1792-1800.
- [23] LI B Y, REN W Q, FU D P, et al. Benchmarking single image dehazing and beyond [J]. IEEE Transactions Image Processing, 2019, 28(1):492-505.
- [24] TANG L F, ZHANG H, XU H, et al. Rethinking the necessity of image fusion in high-level vision tasks: A practical infrared and visible image fusion network based on progressive semantic injection and scene fidelity [J]. Information Fusion, 2023, 99:101870.
- [25] YU Z P, HUANG H B, CHEN W J, et al. YOLO-Facev2: A scale and occlusion aware face detector [J]. Pattern Recognition, 2024, 155:110714.
- [26] TIAN ZH, SHEN CH H, CHEN H, et al. FCOS: Fully convolutional one-stage object detection [C]. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019.
- [27] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization [C]. IEEE Conference on Computer Vision, 2017:618-626.
- [28] MENG X L, LIU Y, FAN L L, et al. YOLOv5s-Fog: An improved model based on YOLOv5s for object detection in foggy weather scenarios [J]. Sensors, 2023, 23(11):5321.

### 作者简介

刘震, 硕士研究生, 主要研究方向为深度学习、目标检测。  
E-mail: 3355768368@qq.com

杨贤昭(通信作者), 副教授, 博士, 主要研究方向为信号处理、电磁场与电磁波。

E-mail: yangxianzhao@wust.edu.cn

陈洋, 教授, 博士生导师, 主要研究方向为机器人建模、规划与控制。

E-mail: chenyag@wust.edu.cn

曾思航, 硕士研究生, 主要研究方向为深度学习、目标检测。

E-mail: 624462167@qq.com