

基于改进 YOLOv8n 的渔光互补池塘作业船导航中线提取方法

孙月平^{1,2}, 孟祥汶¹, 郭佩璇¹, 李自强¹, 刘 勇¹, 赵德安^{1,3}

(1. 江苏大学电气信息工程学院, 镇江 212013; 2. 常州东风农机集团有限公司, 常州 213200;

3. 江苏大学机械工业设施农业测控技术与装备重点实验室, 镇江 212013)

摘要: 在复杂渔光互补池塘环境中, 传统机器视觉算法易受光影变化、池中水草分布和水面障碍物遮挡等因素干扰, 导致视觉导航线检测效果不佳。针对上述问题, 该研究提出了一种基于改进 YOLOv8n 的渔光互补池塘作业船导航中线提取方法。首先从提高检测实时性角度出发, 将 HGNetV2 网络作为主干网络, 采用组归一化方式 (group normalization, GN) 与共享卷积结构, 设计轻量化检测头网络, 减小模型体积; 然后使用 SPPF_LSKA 模块作为特征融合层, 提高模型多尺度特征融合能力; 最后采用 Wise-IoU (weighted interpolation of sequential evidence for intersection over union) 损失函数, 提升边界框回归性能和对中远处小目标的检测精度。利用改进 YOLOv8n 检测框坐标提取两侧水泥立柱定位参照点, 通过最小二乘法拟合两侧水泥立柱行线, 进而使用角平分线提取导航中线。消融试验结果表明, 相对于原始 YOLOv8n 模型, 改进 YOLOv8n 模型的计算量、参数量和模型体积分别下降 36.0%、36.8% 和 32.8%, 平均精度均值 (mean average precision, mAP) 为 97.9%, 查准率为 93.1%, 单张图像检测时间为 6.8 ms, 检测速度提升 42.9%。不同模型对比试验表明, 改进 YOLOv8n 模型在较低计算成本的基础上, 体现出了良好的实时性与精准度检测性能, 具有明显优势。在导航中线定位分析试验中, 提取水泥立柱定位参照点与人工观测标记点平均直线误差在 0~5 和 5~10 m 距离范围内分别为 3.69 cm 和 4.57 cm, 提取导航中线与实际导航中线平均直线误差为 3.26 cm, 准确率为 92%。在导航中线实时性试验中, 导航中线平均提取速度为 22.34 帧/s, 满足渔光互补池塘无人作业船导航要求, 为后续作业船视觉导航系统研究提供参考。

关键词: YOLOv8n; 视觉导航; 无人作业船; 渔光互补池塘; 中线检测

doi: 10.11975/j.issn.1002-6819.202408065

中图分类号: TP391.4; S24

文献标志码: A

文章编号: 1002-6819(2024)-22-0173-10

孙月平, 孟祥汶, 郭佩璇, 等. 基于改进 YOLOv8n 的渔光互补池塘作业船导航中线提取方法[J]. 农业工程学报, 2024, 40(22): 173-182. doi: 10.11975/j.issn.1002-6819.202408065 <http://www.tcsae.org>

SUN Yueping, MENG Xiangwen, GUO Peixuan, et al. Extracting the navigation center line for fishery complementary photovoltaic boat using improved YOLOv8n[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2024, 40(22): 173-182. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.202408065 <http://www.tcsae.org>

0 引言

近年来, 中国光伏产业发展迅速, 为了解决光伏电站用地成本较高的问题, 出现了“渔光互补”独特养殖模式^[1], 将光伏发电设备架设在池塘上方, 实现光伏发电和渔业养殖生产相结合的一种新兴产业模式。传统池塘养殖投饵施药等生产环节主要靠人工作业, 为克服人工养殖劳动强度大和作业效率低等不足, 陆续出现了一些无人投饵施药作业船^[2-4]。由于无人作业船主要依赖北斗/GPS 导航定位, 而渔光互补池塘上方的光伏板阵列会遮挡反射卫星定位信号, 导致作业船导航精度产生较大误差, 严重影响作业船稳定运行。

视觉导航技术因其采集信息丰富、抗干扰能力强等优势被广泛应用于农业机械自动驾驶^[5-7]。对渔光互补池

塘作业环境采集图像进行特征区域提取, 获得导航基准线是无人作业船视觉导航的关键基础。针对导航基准线的研究, 阮承治等^[8]选用 HSV 颜色空间中的 H 分量分割水草图像, 得到水草图像形态特征曲线, 利用最小二乘法拟合出导航线, 平均相对误差为 0.25%, 平均耗时 5 ms。GARCIA-SANTILLAN 等^[9]使用基于 Otsu 分割算法分割玉米幼苗图像得到定位候选点, 通过行扫描法确定幼苗行中心线, 对直线、等距弯曲和不等距弯曲作物行线的检测率分别为 91.8%、86.3% 和 85.6%。JIANG 等^[10]通过改进滑动窗口法获取特征点, 利用 K-means 聚类方法确定消失点的聚类中心, 基于消失点提取作物中心线, 检出率为 90%。

以上研究虽然能获取较为精准的导航基准线, 但面对复杂的作业环境, 如光照条件变化大或障碍物干扰较多等情况, 传统图像分割等技术难以有效应对。随着深度学习算法^[11-12]的出现, 视觉导航的导航线检测技术得到进一步发展。BAH 等^[13]提出一种基于分层方法和无监督图表示的作物行检测方法, 通过 CNN (convolutional neural networks) 和 SLIC (simple linear iterative clustering) 算法检测感兴趣点, 使用霍夫变换法和 RANSAC

收稿日期: 2024-08-08 修订日期: 2024-10-12

基金项目: 国家自然科学基金项目 (62173162); 江苏省农机研发制造推广应用一体化试点专项 (JSYTH14); 江苏省高校优势学科建设项目 (PAPD)

作者简介: 孙月平, 博士, 副教授, 研究方向为农业电气化装置。

Email: sunyujp@ujs.edu.cn

(random sample consensus) 法检测作物行的精确率分别为 85% 和 88%。周学成等^[14]使用轻量化升级后的 DeepLabv3+ 模型分割道路掩码区域, 利用最小二乘法和角平分线法提取导航路径, 分割平均像素准确率为 97.86%, 导航线平均误差占比 2.53%。肖柯等^[15]提出一种基于 Mask R-CNN 和 RANSAC 改进随机采样一致性算法的导航线提取办法, 在不同场景下分割精度为 97%, 导航目标点的偏差率不超过 5.3%。SILVA 等^[16]使用改进 U-Net 模型对马铃薯幼苗冠层进行分割, 提取特征点后确定行中心线, 平均角误差为 2.15°, 平均位移误差为 3.65 cm, 提取速度为 23 帧/s。CAO 等^[17]以残差网络为基础改进 E-Net 语义分割网络模型, 依据分割图像特点改进随机抽样一致性算法得到导航线, 提取精度为 91.2%, 提取速度提升 49%。除了使用基于深度学习的图像分割技术, 目标检测技术也被广泛应用于导航线检测。ZHOU 等^[18]提出一种目标检测和最小二乘法结合的果园场景导航线提取算法, 通过 YOLOv3 检测果树边框坐标得到参考点, 用最小二乘法拟合果树基准线进而确定导航线, 导航线提取准确率为 90%。马驰^[19]使用 Fast R-CNN 算法检测猕猴桃树干得到树干根点的定位基点, 拟合单侧边线最终生成导航线, 不同地貌的平均航向角偏差均在 4.6° 以内。

上述研究主要通过基于深度学习的图像处理技术获取导航基准线, 能解决传统图像处理算法抗干扰能力弱的问题, 但对于无人作业船而言, 提取导航线算法在保证精确度和鲁棒性的同时, 还需要部署在嵌入式平台上满足实时性要求。基于深度学习的 Mask R-CNN 算法虽然拥有较高的检测精度, 但计算成本难以控制, 不利于低成本设备部署; 而 Fast R-CNN 目标检测算法在保证检测精度的条件下, 无法满足实时性的要求, 不适用于惯性较大的无人作业船导航控制。

YOLOv8 是 Ultralytics 公司推出的一个 SOTA 模型, 相对于 YOLOv5, 其更新了骨干网络、Anchor-Free 检测头和损失函数, 模型性能和灵活性得到了提升, 但在低成本的嵌入式设备上, YOLOv8 的中远处小目标检测与实时性等性能仍难以满足无人作业船的导航需求。因此, 本文提出一种基于改进 YOLOv8n 的渔光互补池塘无人作业船导航中线提取方法, 针对提升多尺度融合能力和轻量化升级对 YOLOv8n 进行改进, 并通过模型检测水泥立柱, 获取定位参照点, 使用最小二乘法和角平分线法提取边界行线和导航中线。通过对比原始 YOLOv8n 和改进 YOLOv8n 在低成本嵌入式设备上的实时性表现, 验证本文改进方案的有效性。

1 图像数据采集与数据集构建

选用 RealSense D435i 相机作为图像数据采集设备, 双目测距有效距离为 0~10 m, RGB 帧率为 30 帧/s, RGB 传感器视场角为 69.4°×42.5°(±3°), 如图 1 所示, 在渔光互补池塘中, 相机能够拍摄到两侧水泥立柱, 满足无人作业船视觉导航需求。



a. 自北向南航向
a. From north to south (N-S)

b. 自南向北航向
b. From south to north (S-N)

注: A 为低照度样本; B 为中等照度样本; C 为高照度样本。

Note: A is the low intensity sample, B is the medium intensity sample, C is the high intensity sample.

图 1 部分数据集样本

Fig.1 Partial dataset sample

图像数据采集地点位于江苏省宜兴市新建镇 100 MW 渔光互补发电项目基地。将相机安装在无人作业船前端, 距离水面高度为 1.3 m, 遥控船航行完成视频录制, 船速为 0.5 m/s。如图 1 所示, 由于光伏板按照一定倾斜角布置, 因此不同天气和不同航向的光照强度差异较大, 在录制视频同时, 采集并记录同时段光伏板下相同位置的光照强度, 采集频率为 30 s/次。共录制 10 段视频, 每段时长均为 5 min 左右, 共计 3 156 帧, 滤除不含水泥立柱以及较为模糊的样本, 得到 2 895 张图像, 作为数据集, 其中, 将光照强度在 10 000 ~ 28 000 lx 范围内的图像划分为中等照度, 小于 10 000 lx 的图像划分为低照度, 大于 28 000 lx 的图像划分为高照度, 样本具体光照强度分布如表 1 所示。

表 1 样本光照强度分布

Table 1 Distribution of sample light intensity		
光照强度 Light intensity	样本数量 Sample size	水泥立柱数量 Number of concrete column
低照度 Low intensity	876	7 504
中等照度 Medium intensity	1 064	10 161
高照度 High intensity	955	7 602
总计 Total	2 895	25 267

按照 8 : 1 : 1 的比例, 将数据集划分为训练集、验证集和测试集。使用 LabelImg 工具, 分别在两侧选取距离拍摄位较近的 5~8 个清晰水泥立柱标注, 标注位置为水泥立柱主体部位, 标签为 column, 矩形框坐标保存在 xml 文件中。

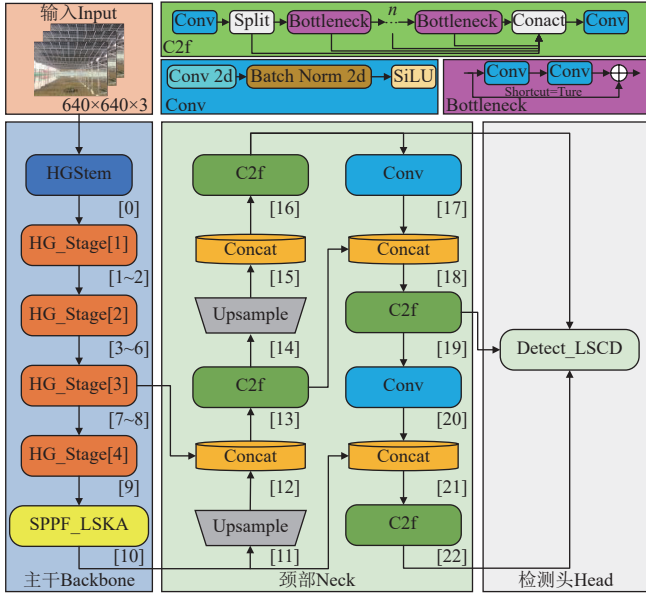
2 渔光互补池塘导航中线检测方法

2.1 改进的 YOLOv8n 轻量化和多尺度检测网络

在渔光互补池塘中, 水泥立柱呈现近大远小的透视效果, 目标尺度不一, 容易出现漏检和误检的情况。在本文低成本的嵌入式设备中, YOLOv8n 模型的检测速度

仅 16 帧/s 左右, 难以满足无人作业船视觉导航需求, 因此, 需要对 YOLOv8n 模型进行改进。

为满足上述应用需求, 本文在维持检测精度的同时, 对模型进行轻量化升级, 提升检测速度。改进后的 YOLOv8n 模型网络结构如图 2 所示, 具体改进如下: 1) 使用 HGNetV2 网络替换 YOLOv8n 主干网络, 大幅降低模型计算参数, 减小模型体积; 2) 特征融合采用 SPPF_LSKA 模块, 提高模型多尺度特征提取能力; 3) 重新设计轻量化检测头, 在保持精度的同时降低计算量; 4) 采用 Wise-IoU 损失函数, 引入更合理的梯度增益分配策略, 提高模型整体检测性能。



注: HGStem 为特征初处理阶段; HG_Stage[1]-HG_Stage[4] 为特征提取阶段; SPPF-LSKA 为融合大可分离核注意力的 SPPF 层; Split 为张量拆分操作; Detect-LSCD 为使用轻量化和共享卷积组成的检测头; C2f 为由 Conv 和 Bottleneck 组成的特征融合模块; Batch Norm 2d 为批归一化处理; n 为 Bottleneck 模块数量; $[i]$ 为模型第 i 层, $i=0\sim22$ 。

Note: HGStem is the feature preprocessing stage; HG_Stage[1]-HG_Stage[4] are the feature extraction stages; SPPF-LSKA is the fusion of a large separable kernel attention SPPF layer; Split is the tensor splitting operation; Detect-LSCD is the detection head composed of lightweight and shared convolutions; C2f is the feature fusion module composed of Conv and Bottleneck; Batch Norm 2d is the batch normalization processing; n is the number of Bottleneck modules; $[i]$ is the i -th layer of the model, $i=0\sim22$.

图 2 改进 YOLOv8n 网络结构

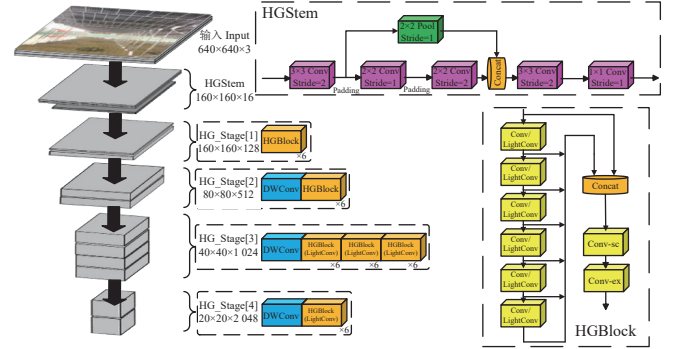
Fig.2 Network structure of improved YOLOv8n

2.1.1 HGNetV2 网络

HGNetV2 网络结构^[20-21] (图 3) 整体由 HGStem、DWConv 和 HGBlock 等模块组成。特征提取主要分为 4 个阶段, HGStem 对输入图像进行初始处理得到特征图, 经由 4 个 HG_Stage 处理, 得到不同尺寸特征图, 适应于检测不同尺度目标。

HGStem 结构在进行一次 3×3 卷积后, 将一部分特征图进行最大值池化, 另一部分进行 2×2 卷积, 降低通道数, 填充尺寸后再次进行 2×2 卷积, 恢复通道数且减半特征图尺寸。两路结果拼接后进行两次卷积, 减半尺寸并缩减通道数, 输出特征图尺寸为原图的 $1/4$, 降低后续过程的计算量。

特征提取的 4 个 HG_Stage 由 DWConv 和 HGBlock 模块构成, 如图 3 所示, HGBlock 模块由大量标准卷积构成。在 HG_Stage[3]、HG_Stage[4] 中, HGBlock 采用多个 LightConv 轻量级卷积方式, 将每一层卷积结果和输入做拼接, 增加目标特征的丰富性和多样性, 然后进行压缩和激励操作, 输出特征信息。在 HGBlock 前端, 采用 DWConv 深度可分离卷积^[22]。相较于普通卷积, DWConv 在特征维度上做分组卷积, 并使用 1×1 卷积聚合通道得到特征图, 可以减少输入通道数, 降低后续卷积运算参数数量, 从而达到精简模型、提高运算速度的目的。



注: Pool 为池化操作; DWConv 为深度可分离卷积操作; LightConv 为轻量化卷积操作; Padding 为填充操作; HGBlock 为特征提取模块; $160\times160\times16$ 、 $160\times160\times128$ 、 $80\times80\times512$ 、 $40\times40\times1024$ 和 $20\times20\times2048$ 分别为该输出特征图维度。

Note: Pool is a pooling operation; DWConv is a deep separable convolution operation; LightConv is a lightweight convolution operation; Padding is a filling operation; HGBlock is the feature extraction module; $160\times160\times16$, $160\times160\times128$, $80\times80\times512$, $40\times40\times1024$ and $20\times20\times2048$ are the dimensions of the output feature map, respectively.

图 3 HGNetV2 网络结构图

Fig.3 Network structure of HGNetV2

2.1.2 SPPF_LSKA 模块

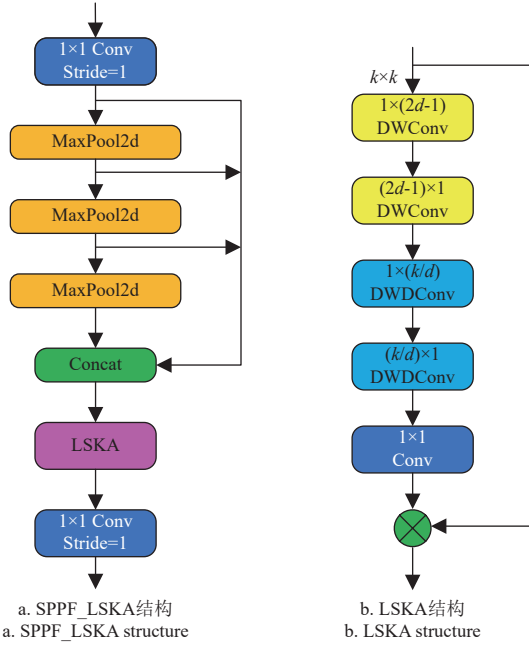
在渔光互补池塘中, 两侧水泥立柱呈现近大远小的特征, 为防止远处小目标漏检和保证不同尺度目标检测精度, 采用 SPPF_LSKA 模块^[23], 其结构图如图 4a 所示。LSKA (large separable kernel attention) 结构利用水平卷积、垂直卷积和深度扩张卷积对不同尺度的信息进行加权和融合, 减少由于尺度变化带来的信息丢失。输入特征图在连续通过 3 个最大池化层后, 输出结果拼接到一起, 并与 LSKA 结构生成的注意力图加权原始特征, 提高 SPPF 层在多个尺度上聚合特征的能力。

LSKA 结构中将深度卷积层的二维卷积核分解, 得到水平和垂直的一维卷积核, 以级联的方式作用于输入特征。如图 4b 所示, 首先分别使用核为 $1\times(2d-1)$ 和 $(2d-1)\times1$ 的卷积, 提取特征图水平和垂直方向特征, 生成初步注意力图; 然后使用空间扩张卷积来进一步提取特征, 覆盖更大的感受野和更广泛的上下文信息; 最后使用普通卷积得到最终注意力图, 并与原始输入特征图作乘法操作, 输出结果。设 LSKA 的输入特征图高、宽、通道数为 $H\times W\times C$, 由图 4b 可计算得出, LSKA 结构的总计算参数量 P_1 和浮点运算量 F_1 分别为

$$P_1 = (2d-1) \times C \times 2 + \left[\frac{k}{d} \right] \times C \times 2 + C \times C \quad (1)$$

$$F_1 = ((2d-1) \times C \times 2 + \left\lceil \frac{k}{d} \right\rceil \times C \times 2 + C \times C) \times H \times W \quad (2)$$

式中 d 为膨胀率, k 为卷积核边长。



注: MaxPool2d 为最大池化层; LSKA 为大可分离核注意力; DWConv 为深度扩张卷积; Stride 为卷积步长; 1×1 、 $k \times k$ 、 $1 \times (k/d)$ 、 $(k/d) \times 1$ 、 $1 \times (2d-1)$ 和 $(2d-1) \times 1$ 分别为卷积核大小; d 为膨胀率; k 为卷积核边长。
Note: MaxPool2d is the maximum pooled layer; LSKA is the large separable kernel attention; DWConv is a deep expanding convolution Stride is the convolution step; 1×1 , $k \times k$, $1 \times (k/d)$, $(k/d) \times 1$, $1 \times (2d-1)$ and $(2d-1) \times 1$ are the convolution kernel sizes; d is the expansion rate; k is the side length of convolution kernel.

图 4 SPPF_LSKA 结构图
Fig.4 SPPF_LSKA structure diagram

若采用普通大核深度卷积和扩展深度卷积, 总计算参数量 P_2 和浮点运算量 F_2 分别为

$$P_2 = (2d-1)^2 \times C + \left\lceil \frac{k}{d} \right\rceil^2 \times C + C \times C \quad (3)$$

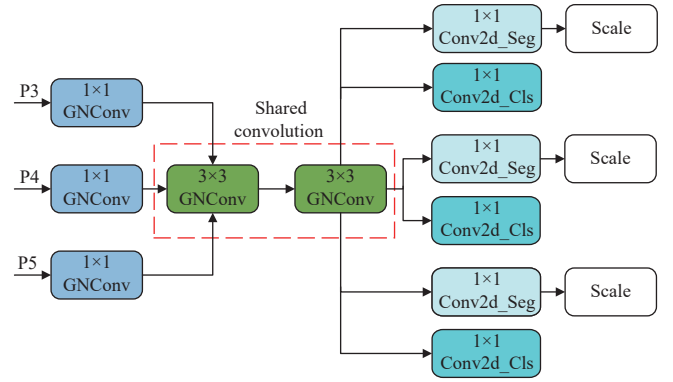
$$F_2 = ((2d-1)^2 \times C + \left\lceil \frac{k}{d} \right\rceil^2 \times C + C \times C) \times H \times W \quad (4)$$

对比式 (1) 和式 (3) 的第 1 项, 前者的卷积方式能节省 $(2d-1)/2$ 倍参数量, 对比第 2 项, 可节省 $(k/d)/2$ 倍参数量, 浮点运算量两者相同。因此在保证特征提取效果的同时, 采用大分离核注意力的模型会有更好的表现。

2.1.3 轻量化检测头设计

为进一步缩减模型大小, 对 YOLOv8n 的检测头做轻量化设计得到 Detect_LSCD, 其结构如图 5 所示。使用组归一化 (group normalization, GN) 层^[24] 替换普通卷积层中的批量归一化 (batch normalization, BN) 层, 通过通道分组, 在组内计算均值和方差进行归一化操作, 优化小批次训练效果。P3、P4、P5 分别为特征融合得到的小尺寸目标、中尺寸目标和大尺寸目标预测层, 分别经历 GNConv^[25] 计算后, 通过共享权重的卷积结构进一步做特征提取, 采用共享卷积结构, 大幅减少参数数量。最后调整通道, 对特征进行尺度缩放, 预测目标框的位

置和类别。



注: GNConv 为采用组归一化的卷积操作; Conv2d_Seg 为回归卷积操作; Conv2d_Cls 为分类卷积操作; Scale 为特征尺度缩放操作; Shared conversation 为共享卷积模块; 1×1 与 3×3 分别为卷积核大小。
Note: GNConv is a convolution operation using group normalization; Conv2d_Seg is a regression convolution operation; Conv2d_Cls is a classification convolution operation; Scale is a feature scale scaling operation; Shared conversation is a shared convolution module; 1×1 and 3×3 are the convolution kernel sizes.

图 5 Detect_LSCD 结构图
Fig.5 Detect_LSCD structure diagram

2.1.4 Wise-IoU 损失函数

边框损失函数是目标检测损失函数的重要组成部分, 是模型检测性能的重要影响因素之一。针对高质量训练集的损失函数研究中, 大多致力于强化边框损失的拟合能力, 这不利于提高低质量图像的检测效果。YOLOv8n 采用的 C-IoU 损失函数, 对低分辨率目标检测没有较好的表现。因此本文针对渔光互补池塘远处水泥立柱目标不易识别和识别精度低的问题, 使用 Wise-IoU V3 损失函数^[26], 提升模型对低质量目标的检测能力。在基于注意力边界损失的 Wise-IoU V1 基础上, Wise-IoU V3 采用动态非单调聚焦机制, 通过构造动态梯度增益的计算方法来辅助聚焦机制, 使得模型更集中处理普通质量的锚框, 提高对远处水泥立柱的检测能力。

Wise-IoU V1 根据距离度量, 构造距离注意力 R_{WIoU} , 放大普通质量锚框边框损失 L_{IoU} , Wise-IoU V1 边框损失 L_{WIoUV1} 计算过程为

$$R_{WIoU} = \exp\left(\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2)}\right) \quad (5)$$

$$L_{WIoUV1} = R_{WIoU} L_{IoU} \quad (6)$$

式中 W_g 和 H_g 为真实框和预测框最小外扩框的长和宽, x_{gt} 和 y_{gt} 分别为真实框中心点横纵坐标, x 和 y 分别为预测框中心点横纵坐标。

设梯度增益即聚焦系数为 γ , 由离群度 β 表征, 计算式如下:

$$\beta = \frac{L_{IoU}^*}{L_{IoU}} \in [0, +\infty) \quad (7)$$

$$\gamma = \frac{\beta}{\delta \alpha^{\beta-\delta}} \quad (8)$$

式中 $\overline{L_{IoU}}$ 为滑动平均值, $*$ 为将最小外扩矩形框从图中分离操作, δ 和 α 为超参数。

Wise-IoU V3 边框损失 L_{WIoUV3} 计算方法为

$$L_{WIoUV3} = \gamma L_{WIoUV1} \quad (9)$$

2.2 导航中线提取方法

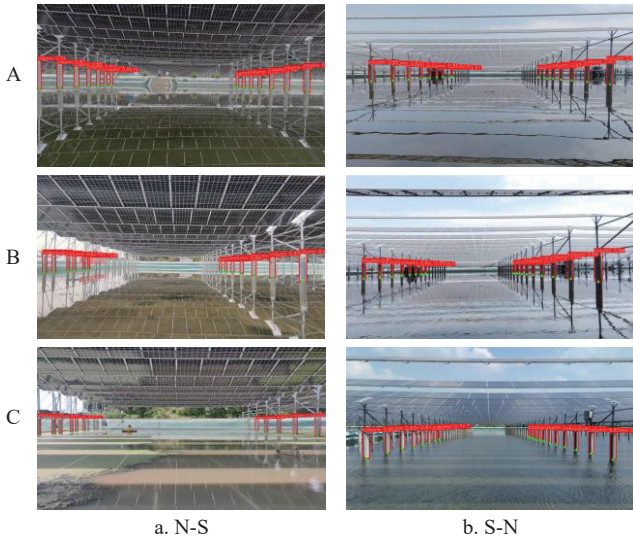
2.2.1 提取定位点

获取作业船导航中线，首要工作是确认定位参照点^[27]。根据安装的位置不同，光伏池塘内设置的水泥立柱大小不一，并且在水平方向上不一定是平行关系。由于水泥立柱为圆柱体，从任意视角观察均可将其近似视为矩形，YOLOv8n 检测框在平面投影上与其贴合度较高，检测框的坐标能较精确反映水泥立柱的定位。若采用水泥立柱检测框几何中心点作为定位参照点，绘制出的导航线位于水面上方，与现实航线存在较大误差。在水泥立柱与水面交界处的位置提取定位参照点，获取的边界线与实际航行边界比较接近，导航线也较为可靠。为减少因图像角度和水泥立柱宽度变化导致的定位偏差，本文选择水泥立柱与水面交界处的中心坐标，即水泥立柱检测框底部中心坐标作为定位参照点。

通过训练改进 YOLOv8n 模型对目标立柱进行检测，可得到预测框坐标，进而计算得出立柱底部中心坐标。设检测框左上角坐标点为 $P_1 = (x_{11}, y_{11})$ ，右下角坐标点为 $P_2 = (x_{22}, y_{22})$ ，单一水泥立柱的底部中心坐标为 $P_C = (x_C, y_C)$ ，则 x_C, y_C 计算如下：

$$\begin{cases} x_C = \frac{x_{11} + x_{22}}{2} \\ y_C = y_{22} \end{cases} \quad (10)$$

图 6 为 6 种不同场景下的水泥立柱定位参照点的提取结果。



注：圆点为根据预测框计算的水泥立柱定位参照点；方框为水泥立柱检测框。
Note: The dots are the reference points for the positioning of the cement column calculated according to the prediction box; the box is the cement column detection box.

图 6 水泥立柱定位参照点提取示意图

Fig.6 Concrete column positioning reference points extraction diagram

为防止误检到相邻池塘的水泥立柱，因此需要对定位参照点进行处理。根据立柱横坐标 x_C 的值，对图像中

所有定位参照点进行升序排序，获取图像长宽尺寸 ($l \times w$)，设置 $l/2$ 作为横坐标预分组条件，得到两组数据，即左右两侧水泥立柱定位参照点。

由于双目相机视角下的两行立柱呈现向中间聚拢的趋势，且定位点离散程度低。经加载视频推理试验可知，由于距离较远，误检测点偶尔出现，其与临近点连线的斜率和正常两点间连线的斜率相差很大，且最大差异点与相邻点连线的斜率是组内两点间斜率均值的 6 倍左右，因此，分别计算组内相邻点坐标的斜率并求其均值 k_a ，设置 $5k_a$ 为阈值。若立柱坐标组中有一点与相邻两点的斜率不在 $(-5k_a, 5k_a)$ 范围内，则判定为隔壁水域的误检测点，并将其滤除，最后得到左右两组精确的水泥立柱定位参照点集，分别为 S_1 和 S_2 。

2.2.2 提取导航中线

目前拟合导航中线的主要算法有 Hough 变换^[28]、最小二乘法^[29]、Bezier 曲线拟合法^[30]以及三次样条插值法^[31]等。本文的研究环境为渔光互补池塘，环境较为结构化，水泥立柱在行线上排列整齐，因此只需要绘制直线来创建航行区域边界，Bezier 曲线拟合法和三次样条插值法多用于曲线拟合，故不采用。Hough 变换需要在参数空间中进行累计操作，计算过程较为复杂，且在样本点数量较少时计算效率并不高，因此本文采用最小二乘法来拟合导航中线。

最小二乘法是用于寻找数据集与某一数学模型之间最佳拟合的一种数学方法，其基本思想是通过最小化观测数据点与拟合模型之间的残差平方和来确定模型的参数。由于渔光互补池塘环境较为结构化，边界线可视为直线，因此，设数学模型为 $y = ax + b$ ，数据点集为 $S = \{(x_i, y_i) | i = 1, 2, 3, \dots, N\}$ ，由定义可确定样本的回归模型为

$$Q = \sum_i^N (y_i - b - ax_i)^2 \quad (11)$$

式中 Q 为残差平方和，当 Q 值达到极小值时的 a 、 b 为目标直线的最优参数。

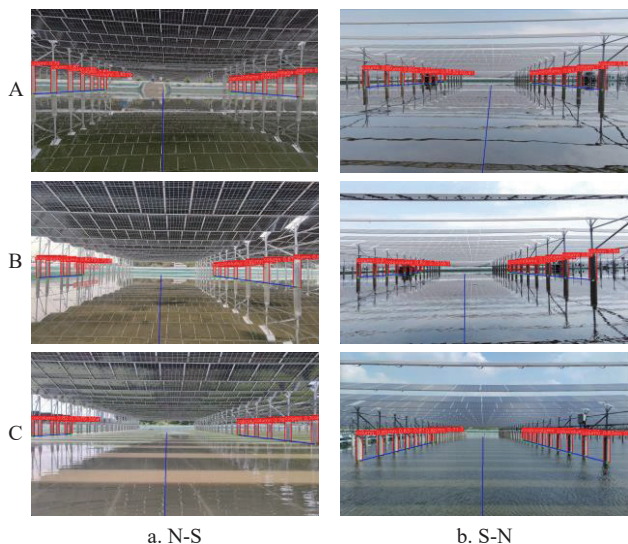
在获得左右两列水泥立柱的定位参照点坐标数组后，代入上式计算，进行两侧立柱行线的拟合，并获取其直线方程，设左右行线及其角平分线的直线方程分别为 $y_1 = k_1x + b_1$ 、 $y_2 = k_2x + b_2$ 和 $y_0 = mx + n$ ，左右行线夹角为 θ ，则由夹角计算式可得：

$$\tan \frac{\theta}{2} = \frac{k_1 - m}{1 + k_1 m} = \frac{m - k_2}{1 + k_2 m} \quad (12)$$

由式 (12) 可建立如下等式：

$$(k_1 + k_2)m^2 + 2(1 - k_1k_2)m - (k_1 + k_2) = 0 \quad (13)$$

根据图 6 可判断，左右行线所成夹角均为钝角，因此解出式 (13) 后舍弃锐角的斜率即可得角平分线斜率 m ，再将两行线直线方程联立可得交点坐标，进而求得角平分线的直线方程，将其作为导航中线，提取效果如图 7 所示。



注：两侧直线为拟合出的水泥立柱行线；中央直线为拟合出的导航中线。
Note: The straight lines on both sides are the fitted cement column lines; the central line is the fitted navigation center line.

图 7 导航中线提取示意图

Fig.7 Navigation center line extraction diagram

3 结果与分析

3.1 试验环境及参数设置

本文模型的试验主机系统为 Winsdows11 操作系统，CPU 为 Intel 13 th Gen Core i7-13 700，GPU 为 NVIDIA GeForce RTX 4 070 Ti，显卡内存为 16 GB。虚拟环境配置为 Python 3.8、PyTorch 1.13.1 和 CUDA 11.6，训练输入图像大小为 640×640 像素，学习率初始值为 0.01，优化器选择随机梯度下降，工作线程数设置为 8，批次大小设置为 16，训练轮次设置为 300。

3.2 模型评价指标

本文使用查准率 (precision, P)、平均精度均值 (mean average precision, mAP)、模型体积大小、参数

量、计算量和检测速度等指标评价模型检测性能；另外，使用横向偏差、纵向偏差和直线误差距离来评价定位参考点提取效果，使用导航中线中点的横坐标偏差以及偏差角来评价导航中线提取效果，使用模型每秒处理图像帧数来评价导航中线提取速度。

3.3 模型试验检测结果分析

3.3.1 消融试验

为验证本文改进 YOLOv8n 模型的性能提升效果，在使用相同数据集和试验参数的基础上，设计消融试验，结果如表 2 所示。

相较于 YOLOv8n，模型 2 使用轻量化主干网络 HGNetV2，大量使用深度可分离卷积 DWConv 和轻量化卷积 LightConv，使得模型计算量、参数量和模型体积都有明显下降，计算量和参数量分别减少 1.9 G 和 0.65 M，模型体积同时缩减至 4.8 MB，单张图像检测时间减少 39.5%。模型 3 采用 LSCD 检测头，通过共享卷积结构，模型参数进一步得到有效控制，减少过拟合风险的同时，能更好融合来自不同尺度的特征，提高对多尺度特征的敏感度，从而改善检测精度，与模型 2 相比，查准率与平均精度均值分别上涨 4.5、1.5 个百分点，并且单张图像检测时间减少至 7.0 ms，模型轻量化程度进一步提升。模型 4 在模型 3 轻量化的基础上，特征融合采用 SPPF_LSKA 模块，由于引入大可分离核注意力机制，模型查准率和平均精度均值得到一定弥补，分别上涨 0.7、1.9 个百分点，模型计算量、参数量和模型体积没有较大涨幅，单张图像检测时间相较于 YOLOv8n，减少了 42.9%，说明此特征融合模块在应对多尺度目标检测任务中，具有良好的表现。在模型 5 中，损失函数替换为 Wise-IoU，对比模型 4 默认采用的 C-IoU，边界框损失收敛更快，损失精度更少，查准率和平均精度均值得到了进一步提升，分别上涨 1.1、0.1 个百分点，提高了模型边界框回归性能。

表 2 消融试验结果
Table 2 Ablation test results

编号 NO.	模型 Models	计算量 Calculation quantity/ G	参数量 Parameter quantity/M	模型体积 Model size/ MB	查准率 Precision/%	平均精度均值 mAP/%	检测时间 Detection time/ms
1	YOLOv8n	8.9	3.15	6.00	93.4	98.2	11.9
2	YOLOv8n-HGNetV2	7.0	2.50	4.80	86.8	94.4	7.2
3	YOLOv8n-HGNetV2-LSCD	5.4	1.72	3.51	91.3	95.9	7.0
4	YOLOv8n-HGNetV2-SPPF_LSKA-LSCD	5.7	1.99	4.03	92.0	97.8	6.8
5	YOLOv8n-HGNetV2-SPPF_LSKA-LSCD-Wise-IoU	5.7	1.99	4.03	93.1	97.9	6.8

综上，对比原始模型，改进后的 YOLOv8n 模型在查准率和平均精度均值都仅减少 0.3 个百分点的前提下，计算量和参数量分别减少 36.0% 和 36.8%，模型体积减少 32.8%，仅为 4.03 MB，单张图像检测时间为 6.8 ms，检测速度提升 42.9%，完成轻量化升级并保持较高检测精度。

3.3.2 不同模型对比试验

为进一步验证本文改进 YOLOv8n 模型的有效性和优势，在使用相同数据集以及试验参数的基础上，将其与 YOLOv5s、YOLOv6、YOLOv7 以及 YOLOv8s 进行比较，结果如表 3 所示。由表 3 结果分析可知，YOLOv7

计算量和参数量比较大，并且单张图像检测时间最长，不满足轻量化和实时性检测要求。本文改进 YOLOv8n 模型的查准率均优于其余 4 个模型，平均精度均值仅比 YOLOv7 少 0.1 个百分点。YOLOv6 查准率比本文改进 YOLOv8n 模型低 3.2 个百分点，同时模型体积也较大。YOLOv5s 查准率与平均精度均值分别比本文改进 YOLOv8n 模型低 2.0 和 2.8 个百分点，并且计算量、参数量与模型体积分别多出 1.5 G、0.51 M 和 1.00 MB。本文改进 YOLOv8n 模型在查准率和平均精度均值优于 YOLOv8s 的同时，计算量、参数量和模型体积分别减

少 80.1%、82.1% 和 81.2%。由此可知，本文改进 YOLOv8n 模型在众多指标中均能体现出优越性。

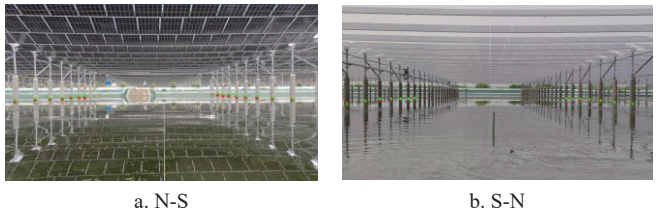
表 3 不同模型检测结果对比
Table 3 Comparison of detection results of different models

模型 Models	计算量 Calculation quantity/ G	参数量 Parameter quantity/M	模型体积 Model size/ MB	查准率 Precision/%	平均精度均值 mAP/%	检测时间 Detection time/ms
YOLOv5s	7.2	2.50	5.03	91.1	95.1	11.1
YOLOv6	11.9	4.24	8.29	89.9	93.5	12.0
YOLOv7	105.1	36.50	74.80	91.0	98.0	14.9
YOLOv8s	28.6	11.10	21.40	92.1	95.6	12.4
YOLOv8n-HGNetV2-SPPF_LSKA-LSCD-Wise-IoU	5.7	1.99	4.03	93.1	97.9	6.8

3.4 导航线提取结果分析

3.4.1 提取定位参照点误差分析

为验证本文算法提取定位参照点的性能，根据不同样本场景，采用人工观测标注的方法与算法提取结果进行对比，对比效果如图 8 所示。横向偏差是指在像素坐标系中，算法提取定位参照点横坐标与人工观测标注定位参照点横坐标差值，纵向偏差是指在像素坐标系中，算法提取定位参照点纵坐标与人工观测标注定位参照点纵坐标差值，直线误差是指在像素坐标系下，算法提取定位参照点与人工观测标注定位参照点的直线距离。



注：绿色点为根据预测框计算的水泥立柱定位参照点；红色点为人工观测标注水泥立柱定位参照点。
Note: The green dots are the reference points for concrete column positioning calculated according to the prediction box; and the red dots are the reference points for the positioning of the concrete column marked by manual observation.

图 8 人工观测标注定位参照点和算法提取定位参照点示意图
Fig.8 Schematic diagram of marking of positioning reference points by manual observation and algorithm extraction

根据光照强度的分布时间，拍摄 100 张样本图像作为测试集，共识别出 912 根水泥立柱，进行定位参照点误差分析，误差统计结果如图 9 所示。横向偏差主要分布在 $(-4, 4)$ 像素范围内，纵向偏差主要分布在 $(-3, 3)$ 像素范围内，其中，横向偏差绝对值小于 4 像素的立柱数量为 853 根，占比 93.5%，纵向偏差绝对值小于 3 像素的立柱数量为 874 根，占比 95.8%，像素直线平均误差为 3.79 像素。

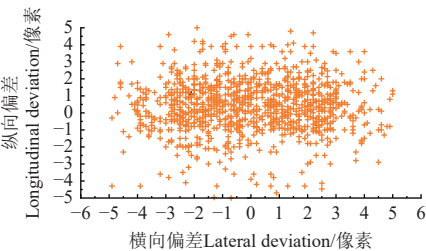


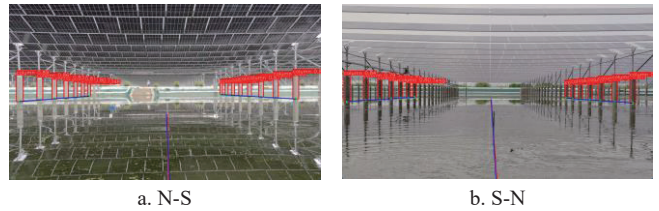
图 9 人工观测标注定位参照点和算法提取定位参照点误差统计
Fig.9 Error statistics of reference points extracted by manual observation and algorithm extraction

通过相机内参计算方法^[32]，将参照点坐标由像素坐标转换至相机坐标系下，计算人工观测标注定位参照点与算法提取定位参照点的欧氏距离，即实际直线误差距离。由于不同立柱定位参照点距离相机位置差异较大，经计算，0~5 m 范围内定位参照点的平均实际直线误差为 3.69 cm，5~10 m 范围内定位参照点的平均实际直线误差为 4.57 cm，结合本文渔光互补池塘作业船视觉导航系统对参照点平均精度的要求为 0~6 cm，说明本文算法提取水泥立柱定位参照点的精度满足导航系统要求。

3.4.2 提取导航中线误差分析

为验证本文算法提取导航中线的性能，引入中线中点偏差以及中线偏差角进行导航中线误差分析。

如图 10 所示，经过测量，在池塘中央区域水面上插杆牵拉直线，建立实际导航中线，作为判断算法提取导航中线的基准线。偏差角 θ' 是指算法提取导航中线与实际导航中线的夹角，已知二者的斜率后， θ' 可通过式 (12) 计算得到。若 $\theta' > 6^\circ$ ，则判定该张图像提取中线不准确，判定为负样本；反之则准确，判定为正样本。在正样本中，计算中线中点偏差 X 。在像素坐标系中，以插杆和拉绳的交点为起点，拉绳和图像底部交点为终点，绘制基准线。为保证中线中点偏差的可靠性，绘制算法提取中线时的起点与终点的纵坐标与基准线一致，分别计算两条中线起始点和终点横坐标的均值，中线中点偏差 X 即为二者均值差值的绝对值。



注：中央红色直线为根据实际导航中线绘制的导航基准线；中央蓝色直线为算法提取导航中线。
Note: The red line in the center is the navigational reference line depicted according to the actual sailing middle line; the central blue line is the navigation center line extracted by the algorithm.

图 10 实际导航中线和算法提取导航中线误差分析示意图
Fig.10 Error analysis diagram of actual navigation center line and algorithm extraction of navigation center line

在 100 张测试集图像中进行测试，结果如图 11 所示，共有 8 张负样本，其中最大偏差角为 11.09° ，最小偏差角 0.48° ，主要偏差角分布在 $(0^\circ, 5^\circ)$ 范围内，导航中线检测准确率达 92%。在所有正样本中，中线中点最大

误差为 7.2 像素, 误差主要分布在 (0, 4.5) 像素范围内, 占比 94%, 平均中线中点误差为 2.51 像素。图像中导航中线中点一般距离相机 4~6 m, 经过相机内参计算后, 中线中点实际平均误差为 3.26 cm, 满足渔光互补池塘无人作业船视觉导航精度需求。

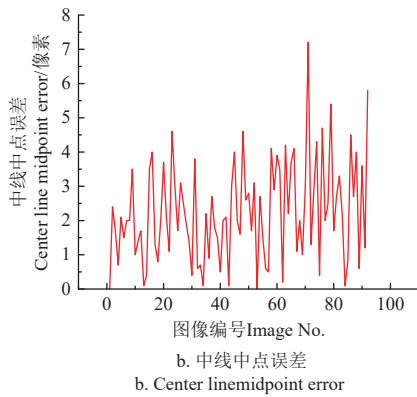
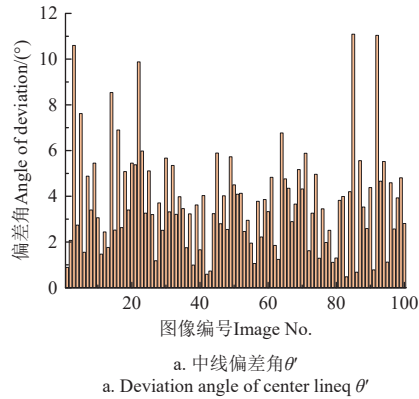


图 11 导航中线误差分析

Fig.11 Analysis of navigation centerline error

3.4.3 提取导航中线速度分析

为验证本文算法提取导航中线实时性, 加载 4 段不同航向下的中照度和高照度条件的视频, 分别计算模型每秒处理图像帧数, 处理过程包括水泥立柱检测及提取导航中线。同时, 为验证本文改进算法在低成本设备上运行性能, 将 pt 模型转换至 engine 模型, 部署至 Jetson Orin Nano 4 GB, pt 格式模型运行环境与 3.1 节环境相同, engine 格式模型运行系统为 Ubuntu 系统, 虚拟环境配置为 Python3.8、PyTorch2.1.0、CUDA11.6 和 TensorRT8.5。在 Windows11 和 Jetson 两种不同试验平台上分别加载 4 类不同航向和光照条件环境视频, 进行导航中线提取, 结果如表 4 所示。

在 Windows11 试验平台上, 相较于 YOLOv8n 模型, 改进 YOLOv8n 模型对视频 1~视频 4 的导航中线提取速度分别提升 39.33%、37.19%、36.25% 和 39.39%, 平均提取速度提升 38.04%, 为 117.01 帧/s。在 Jetson 试验平台上, 相较于 YOLOv8n 模型, 改进 YOLOv8n 模型对视频 1~视频 4 的导航中线提取速度分别提升 33.7%、34.66%、38.19% 和 38.90%, 平均提取速度提升 36.38%, 为 22.34 帧/s, 能满足渔光互补池塘无人作业船视觉导航性能需求。

表 4 4 类环境视频导航中线提取速度

Table 4 Velocities of the navigation center line extracted in four types of environments (帧·s⁻¹)

模型 Models	视频 1 Video 1	视频 2 Video 2	视频 3 Video 3	视频 4 Video 4
YOLOv8n.pt	83.98	85.19	85.59	84.31
改进 YOLOv8n.pt	117.01	116.87	116.62	117.52
YOLOv8n.engine	16.52	16.82	16.21	15.99
改进 YOLOv8n.engine	22.10	22.65	22.40	22.21

注: 视频 1 为中等照度条件下由北向南航向的视频; 视频 2 为中等照度条件下由南向北航向的视频; 视频 3 为高照度条件下由北向南航向的视频; 视频 4 为高照度条件下由南向北航向的视频。改进 YOLOv8n 模型为 YOLOv8n-HGNetV2-SPPF_LSKA-LSCD-Wise-IoU。

Note: Video 1 is the video traveling from north to south under medium indensity condition; video 2 is the video traveling from south to north under medium indensity condition; video 3 is the video traveling from north to south under high indensity condition; video 4 is the video traveling from south to north under high indensity condition. The YOLOv8n model was improved to YOLOv8n-HGNetV2-SPPF_LSKA-LSCD-WISE-IoU.

4 结 论

本文提出一种基于改进 YOLOv8n 的渔光互补池塘环境下无人作业船导航中线提取方法。从提高模型多尺度特征融合能力和轻量化程度的角度出发, 通过使用 HGNetV2 轻量化网络、SPPF_LSKA 模块、LSCD 检测头和 Wise-IoU 损失函数改进 YOLOv8n 算法, 提高水泥立柱定位参考点的提取精度, 采用最小二乘法和角平分线法提取作业船导航中线。主要结论如下:

1) 本文改进的 YOLOv8n 模型相较于原始模型, 计算量和参数量分别减少 36.0%、36.8%, 模型体积缩小至 4.03 MB, 单张图像检测时间为 6.8 ms, 检测速度提升了 42.9%, 查准率和平均精度均值分别为 93.1% 和 97.9%。与 YOLOv5s、YOLOv6、YOLOv7 和 YOLOv8s 模型对比, 本文改进 YOLOv8n 模型的查准率和平均检测精度均优于其他模型, 并且在保持较高检测精度的前提下, 大幅降低计算量、参数量和模型体积, 保持了最快的速度, 实现了轻量化和检测精度的最佳平衡, 更适宜在无人作业船低成本设备上部署运行。

2) 为验证基于改进 YOLOv8n 的导航中线提取方法的有效性, 本文使用 100 张代表不同光强区间的样本图像, 对 6 种场景下提取的水泥立柱定位参照点和导航中线分别进行误差分析。模型提取水泥立柱定位参照点横向误差绝对值小于 4 像素的占比 93.5%, 纵向误差绝对值小于 3 像素的占比 95.8%, 直线平均误差为 3.79 像素, 平均实际直线误差在 0~5 m 和 5~10 m 距离内分别为 3.69 和 4.57 cm。提取导航中线的准确率为 92%, 中线中点误差在 7.2 像素以内, 误差小于 4.5 像素的占比 94%, 平均误差为 2.51 像素, 实际平均误差为 3.26 cm, 满足无人作业船视觉导航精度要求。

3) 为验证基于改进 YOLOv8n 的导航中线提取方法的实时性, 本文分别加载 4 类环境视频, 在 Windows11 试验平台上导航中线平均提取速度为 117.01 帧/s, 在 Jetson 试验平台上导航中线平均提取速度为 22.34 帧/s, 满足渔光互补池塘无人作业船视觉导航性能要求。

[参 考 文 献]

- [1] 张万顺, 邓浪浪, 彭虹, 等. 柳堰集水库渔光互补工程对水环境影响的数值模拟研究[J]. *水资源保护*, 2023, 39(5): 1-8. ZHANG Wanshun, DENG Langlang, PENG Hong, et al. Numerical simulation study on the influence of fish-light complementary project on water environment in Liuyan Reservoir[J]. *Water Resources Protection*, 2023, 39(5): 1-8. (in Chinese with English abstract)
- [2] 赵德安, 罗吉, 孙月平, 等. 河蟹养殖自动作业船导航控制系统设计与测试[J]. *农业工程学报*, 2016, 32(11): 181-188. ZHAO De'an, LUO Ji, SUN Yueping, et al. Design and experiment of navigation control system of automatic operation boat in river crab culture[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2016, 32(11): 181-188. (in Chinese with English abstract)
- [3] 洪剑青, 赵德安, 孙月平, 等. 水产养殖自动导航无人明轮船航向的多模自适应控制[J]. *农业工程学报*, 2017, 33(1): 95-101. HONG Jianqing, ZHAO De'an, SUN Yueping, et al. Multi model adaptive control of paddlewheel vehicle's course in aquaculture[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2017, 33(1): 95-101. (in Chinese with English abstract)
- [4] 胡庆松, 曹佳瑞, 郑波, 等. 明轮驱动虾塘自主导航投饵船设计与可靠性试验[J]. *农业机械学报*, 2019, 50(11): 121-128. HU Qingsong, CAO Jiarui, ZHENG Bo, et al. Design and reliability test of paddle wheel driven shrimp pond autonomous navigation feeding boat[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2019, 50(11): 121-128. (in Chinese with English abstract)
- [5] 兰玉彬, 闫瑜, 王宝聚, 等. 智能施药机器人关键技术研究现状及发展趋势[J]. *农业工程学报*, 2022, 38(20): 30-40. LAN Yubin, YAN Yu, WANG Baoju, et al. Current status and future development of the key technologies for intelligent pesticide spraying robots[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2022, 38(20): 30-40. (in Chinese with English abstract)
- [6] ZHANG S, LIU Y, XIONG K, et al. A review of vision-based crop row detection method: Focusing on field ground autonomous navigation operations[J]. *Computers and Electronics in Agriculture*, 2024, 222: 109086.
- [7] Ahmadi A, Halstead M, McCool C. Towards autonomous visual navigation in arable fields[C]// Proceeding of the IEEE Conference on Intelligent Robots and Systems. Kyoto, Japan: IEEE/RSJ, 2022: 6585-6592.
- [8] 阮承治, 林根深, 陈旭, 等. 基于图像处理的蟹塘水草图像导航线拟合方法[J]. *中国农机化学报*, 2023, 44(7): 147-153, 235. RUAN Chengzhi, LIN Genshen, CHEN Xu, et al. Navigation line fitting method of crab pond aquatic plant image based on image processing[J]. *Journal of Chinese Agricultural Mechanization*, 2023, 44(7): 147-153, 235. (in Chinese with English abstract)
- [9] GARCIA-SANTILLAN I, GUERRERO J M, MONTALVO M, et al. Curved and straight crop row detection by accumulation of green pixels from images in maize fields[J]. *Precision Agriculture*, 2018, 19(1): 19-41.
- [10] JIANG G Q, WANG X J, WANG Z H, et al. Wheat rows detection at the early growth stage based on Hough transform and vanishing point[J]. *Computers and Electronics in Agriculture*, 2016, 123: 211-223.
- [11] ANTUN V, RENNA F, POON C, et al. On instabilities of deep learning in image reconstruction and the potential costs of AI[J]. *Proceedings of the National Academy of Sciences*, 2020, 117(48): 30088-30095.
- [12] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]// European Conference on Computer Vision, German: Springer, 2016: 21-37.
- [13] BAH M D, HAFIANE A, CANALS R. Hierarchical graph representation for unsupervised crop row detection in images[J]. *Expert Systems with Applications*, 2023, 216: 119478.
- [14] 周学成, 肖明玮, 梁英凯, 等. 基于改进 DeepLabv3+的火龙果园视觉导航路径识别方法[J]. *农业机械学报*, 2023, 54(9): 35-43. ZHOU Xuecheng, XIAO Mingwei, LIANG Yingkai, et al. Navigation path recognition between dragon orchard using improved deepLabv3+Network[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023, 54(9): 35-43. (in Chinese with English abstract)
- [15] 肖珂, 夏伟光, 梁聪哲. 复杂背景下果园视觉导航路径提取算法[J]. *农业机械学报*, 2023, 54(6): 197-204, 252. XIAO Ke, XIA Weiguang, LIANG Congzhe. Visual navigation path extraction algorithm in orchard under complex background[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023, 54(6): 197-204, 252. (in Chinese with English abstract)
- [16] SILVA D R, CIENLNIAK G, WANG G, et al. Deep learning-based crop row detection for infield navigation of agri-robots[J]. *Journal of Field Robotics*, 2023, 41(7): 1-23.
- [17] CAO M Y, TANG F F, PENG J, et al. Improved real-time semantic segmentation network model for crop vision navigation line detection[J]. *Frontiers in Plant Science*, 2022, 1: 898131.
- [18] ZHOU J J, GENG S Y, QIU Q, et al. A deep-learning extraction method for orchard visual navigation lines[J]. *Agriculture*, 2022, 12(10): 1650.
- [19] 马驰. 猕猴桃树干智能检测与行间导航线生成方法研究[D]. 杨凌: 西北农林科技大学, 2021. MA Chi. Research on Intelligent Detection of Kiwi Trunk and Generation of Navigation Line Between Rows[D]. Yangling: Northwest Agriculture & Forestry University, 2021. (in Chinese with English abstract)
- [20] ZHAO X F, ZHANG W W, ZHANG H, et al. ITD-YOLOv8: An infrared target detection model based on YOLOv8 for unmanned aerial vehicles[J]. *Multidisciplinary Digital Publishing Institute*, 2024, 8(4): 161-178.
- [21] 滕文想, 王成, 费树辉. 基于 HGTC-YOLOv8n 模型的煤研识别算法研究[J]. *工矿自动化*, 2024, 50(5): 52-59. TENG Wenxiang, WANG Cheng, FEI Shuhui. Research on coal waste identification algorithm based on HGTC-YOLOv8n model[J]. *Industrial and mining automation*, 2024, 50(5): 52-59. (in Chinese with English abstract)
- [22] 王泽宇, 徐慧英, 朱信忠, 等. 基于改进 YOLOv8 的轻量化鱼苗检测算法: FD-YOLO[J/OL]. *计算机工程*, (2024-06-13) [2024-07-13]. <https://doi.org/10.19678/j.issn.1000-3428.0069278>. WANG Zeyu, XU Huiying, ZHU Xinzong, et al. Lightweight fry detection algorithm based on improved YOLOv8: FD-YOLO[J/OL]. *Computer Engineering*, (2024-06-06)[2024-07-13]. <https://doi.org/10.19678/j.issn.1000-3428.0069278>. (in Chinese with English abstract)
- [23] LAU K W, PO L, REHMAN Y A U. Large separable kernel attention: Rethinking the large kernel attention design in CNN[J]. *Expert Systems with Applications*, 2024, 236: 121352.
- [24] TIAN Z, SHEN C H, CHEN H, et al. FCOS: Fully convolutional One-Stage object detection[C]// Proceeding of the IEEE Conference on Computer Vision, Seoul, Korea (South): IEEE/CVR, 2019: 9626-9635.
- [25] WU R K, LIANG P C, HUANG X, et al. MHorUNet: High-order spatial interaction UNet for skin lesion segmentation[J].

- Biomedical Signal Processing and Control, 2024, 88: 105517.
- [26] TONG Z, CHEN Y, XU Z, et al. Wise-IoU: Bounding box regression loss with dynamic focusing mechanism [EB/OL]. (2023-04-08)[2024-07-13] <https://arxiv.org/abs/2301.10051>.
- [27] 彭书博, 陈兵旗, 李景彬, 等. 基于改进 YOLOv7 的果园行间导航线检测[J]. 农业工程学报, 2023, 39(16): 131-138. PENG Shubo, CHEN Bingqi, LI Jingbin, et al. Detection of the navigation line between lines in orchard using improved YOLOv7[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2023, 39(16): 131-138. (in Chinese with English abstract)
- [28] 李霞, 苏筠皓, 岳振超, 等. 基于中值点 Hough 变换玉米行检测的导航线提取方法[J]. 农业工程学报, 2022, 38(5): 167-174. LI Xia, SU Yunhao, YUE Zhenchao, et al. Extracting navigation line to detect the maize seedling line using median-point Hough transform[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2022, 38(5): 167-174. (in Chinese with English abstract)
- [29] 刘星星, 张超, 张浩, 等. 最小二乘法与 SVM 组合的林果行间自主导航方法[J]. 农业工程学报, 2021, 37(9): 157-164. Liu Xingxing, Zhang Chao, Zhang Hao, et al. Inter-row automatic navigation method by combining least square and SVM in forestry[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(9): 157-164. (in Chinese with English abstract)
- [30] 奚小波, 史扬杰, 单翔, 等. 基于 Bezier 曲线优化的农机自动驾驶避障控制方法[J]. 农业工程学报, 2019, 35(19): 82-88. Xi Xiaobo, Shi Yangjie, Shan Xiang, et al. Obstacle avoidance path control method for agricultural machinery automatic driving based on optimized Bezier[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2019, 35(19): 82-88. (in Chinese with English abstract)
- [31] 赵汝健, 杨威, 巫震宇, 等. 基于 LSTM 和海图约束的舰船航迹预测方法[J/OL]. 北京航空航天大学学报, (2024-01-23) [2024-07-13]. <https://doi.org/10.13700/j.bh.100-595.2023.0516.0516>. ZHAO Rujian, YANG Wei, WU Zhenyu, et al. Ship track prediction method based on LSTM and nautical chart constraints[J/OL]. Journal of Beihang University, (2024-01-23)[2024-07-13]. <https://doi.org/10.13700/j.bh.100-595.2023.0516.0516>. (in Chinese with English abstract)
- [32] 孙月平, 孙杰, 袁必康, 等. 基于改进 YOLOv5s 的轻量化蟹塘障碍物检测与定位方法[J]. 农业工程学报, 2023, 39(23): 152-163. SUN Yueping, SUN Jie, YUAN Bikang, et al. Lightweight crab pond obstacle detection and location method based on improved YOLOv5s[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2023, 39(23): 152-163. (in Chinese with English abstract)

Extracting the navigation center line for fishery complementary photovoltaic boat using improved YOLOv8n

SUN Yueping^{1,2}, MENG Xiangwen¹, GUO Peixuan¹, LI Ziqiang¹, LIU Yong¹, ZHAO De'an^{1,3}

(1. College of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, China; 2. Changzhou Dongfeng Agricultural Machinery Group Co., LTD., Changzhou 213200, China; 3. Key Laboratory of Agricultural Measurement and Control Technology and Equipment for Machinery Industry, Jiangsu University, Zhenjiang 212013, China)

Abstract: Photovoltaic panel arrays can often shade over the fishing and light complementary ponds. BeiDou/GPS positioning signals are then affected to significantly reduce the autonomous navigation accuracy of unmanned workboats. Additionally, traditional machine vision can easily cause the suboptimal visual navigation of line detection, due to the low robustness. The resulting images are also confined to the variations in the light and shadow, the distribution of aquatic plants, and surface obstacles. In this study, an improved YOLOv8n model was proposed to extract the navigation center line in the fish and light complementary ponds. Firstly, the HGNetV2 network was used as the backbone network, in order to improve the real-time detection. Group normalization (GN) and shared convolution structure were also used to design a lightweight detection head network, in order to reduce the size of the model. Then the SPPF_LSKA module was used as the feature fusion layer to improve the multi-scale feature fusion of the model. Finally, the Wise-IoU (weighted interpolation of sequential evidence for intersection over union) loss function was used to improve the bounding box regression performance and the detection accuracy of remote small targets. The detection frame coordinates of improved YOLOv8n were used to extract the reference points for the positioning of the cement columns on both sides. The lines of the cement columns on both sides were fitted by the least square method. The middle line of the navigation was then extracted by the angle bisection line. Ablation test results showed that the calculation amount, parameter number and model volume of the improved YOLOv8n model decreased by 36.0%, 36.8% and 32.8%, respectively, compared with the original, where the mean average precision (mAP) was 97.9%. The detection speed increased by 42.9%, where the accuracy was 93.1%, and the detection time of a single image was 6.8 ms. Comparison test showed that the improved YOLOv8n model exhibited the smallest size and the highest degree of lightweight, while maintaining a high level of detection accuracy, compared with the YOLOv5s, YOLOv6, YOLOv7, and YOLOv8. Excellent performance was also achieved in detecting the concrete columns. In the positioning test of the navigation center line, the average linear errors between the reference point of the extraction cement column and the manual observation mark were 3.69 and 4.57 cm in the range of 0-5 m and 5-10 m, respectively. The average linear error between the extracted and actual navigation center line was 3.26 cm, with an accuracy of 92%. The real-time test showed that the average extraction speed of the navigation midline was improved by 38.04% (117.01 frame/s) on the Windows11 test platform, compared with the original. Furthermore, the average extraction speed of navigation midline increased by 36.38% (22.34 frame/s) on the Jetson test platform. Consequently, the improved model can fully meet the navigation requirements of unmanned fishing and light complementary pond boat. The findings can provide a strong reference for the subsequent research on the visual navigation system of operation boats.

Keywords: YOLOv8n; visual navigation; unmanned operation boat; fish and light complementary pond; center line detection