doi:10.6041/j.issn.1000-1298.2025.05.011

基于改进 YOLO 11n 模型的棉花田间复杂环境障碍物检测方法

韩科立[1,2]　王振坤[1,2]　余永峰[1,2]　刘淑平[2,3]　韩树杰[2,3]　郝付平[2,3]

(1. 中国农业机械化科学研究院集团有限公司, 北京 100083; 2. 现代农装科技股份有限公司, 北京 100083;
3. 农业装备技术全国重点实验室, 北京 100083)

摘要: 针对棉花田间复杂环境障碍物被遮挡致准确检测难、边缘设备算力有限的问题,本文提出一种基于改进 YOLO 11n 模型的田间障碍物检测方法。首先,采用轻量级网络 StarNet 作为主要特征提取网络,并引入 DBA 模块(Dynamic position bias attention block)重构 C2PSA(Convolutional block with parallel spatial attention),以增强多尺度特征之间的交互能力;其次,使用 KAGNConv(Kolmogorov - Arnold generalized network convolution)替换基线模型 C3k2(Cross stage partial with kernel size 2)模块中的瓶颈结构,实现对精细特征提取的同时,给予模型更高灵活性和可解释性;最后,集成分离与增强注意力模块(Separated and enhancement attention module, SEAM)至检测头,增强模型在遮挡场景中的检测能力。试验结果表明,改进模型 YOLO 11n - SKS 与基线模型相比精确率、召回率、mAP_{50}、mAP_{50-95}分别提升 2.3、2.1、1.3、1.4 个百分点,达到 91.7%、88.3%、91.9%、62.3%,模型浮点数运算量仅为 4.4×10^9 FLOPs,模型参数量减少 17.1%。本研究模型在性能和计算复杂度之间实现了较好的平衡,满足棉田收获作业场景中实时检测需求,降低了部署边缘设备算力要求,为采棉机自主安全作业提供技术支撑。

关键词: 采棉机; 障碍物检测; 深度相机; YOLO 11n 模型; 目标识别

中图分类号: TP391.4　　**文献标识码**: A　　**文章编号**: 1000-1298(2025)05-0111-10　　**OSID**:

Obstacle Detection Method for Complex Cotton Field Environments Based on Improved YOLO 11n Model

HAN Keli[1,2]　WANG Zhenkun[1,2]　YU Yongfeng[1,2]　LIU Shuping[2,3]　HAN Shujie[2,3]　HAO Fuping[2,3]

(1. *Chinese Academy of Agricultural Mechanization Sciences Group Co., Ltd., Beijing* 100083, *China*
2. *Modern Agricultural Equipment Co., Ltd., Beijing* 100083, *China*
3. *National Key Laboratory of Agricultural Equipment Technology, Beijing* 100083, *China*)

Abstract: Aiming to address the challenges of accurate obstacle detection in complex cotton field environments due to occlusions and the computational limitations of edge devices, a field obstacle detection method based on improved YOLO 11n model was proposed. Firstly, the lightweight StarNet network was adopted as the primary feature extraction network, and the dynamic position bias attention block module (DBA) was introduced to reconstruct convolutional block with parallel spatial attention (C2PSA) to enhance multi-scale feature interaction. Secondly, Kolmogorov - Arnold generalized network convolution (KAGNConv) was used to replace the bottleneck structure in the cross stage partial with kernel size 2 module (C3k2) of the baseline model, enabling fine-grained feature extraction while improving model flexibility and interpretability. Finally, the separated and enhancement attention module (SEAM) was integrated into the detection head to enhance the model's detection capability in occlusion scenarios. The experimental results showed that, compared with the baseline model, the improved YOLO 11n - SKS achieved increases of 2.3, 2.1, 1.3, and 1.4 percentage points in precision, recall, mAP_{50}, and mAP_{50-95}, reaching 91.7%, 88.3%, 91.9%, and 62.3%, respectively. The model's floating-point operations were reduced to only 4.4×10^9 FLOPs, and the number of model parameters was

收稿日期: 2025-02-03　修回日期: 2025-03-15

基金项目: 国家重点研发计划项目(2022YFD2002402)和中国机械工业集团有限公司重大科技专项(ZDZX2022-1)

作者简介: 韩科立(1983—),男,正高级工程师,博士,主要从事收获机械智能控制系统研究,E-mail: hanky2008@163.com

decreased by 17.1%. This study achieved a favorable balance between performance and computational complexity, meeting the real-time detection requirements of cotton harvesting operations while lowering the computational demands for deployment on edge devices, thereby providing technical support for the autonomous and safe operation of cotton pickers.

Key words: cotton picker; obstacle detection; depth camera; YOLO 11n model; object recognition

0 引言

随着我国农业机械化快速发展,机械化采棉已成为我国当前主要的棉花采收形式,而采棉机是实现机械化采棉的关键核心装备,也是棉花全程机械化和规模化生产不可或缺的重要农机装备。

随着人工智能、大数据、物联网等前沿技术的迅速发展,棉花生产方式从机械化向自动化、智能化迈进。通信技术和智能控制技术等先进科学技术的应用也推动了棉花生产精细农业技术体系的更新迭代[1-2]。而棉花收获装备自主作业技术是实施智慧农业等精细农业的基础,其主要包括棉田环境中障碍物检测识别、路径规划、车辆模型构建与智能控制等关键技术,其中,棉田障碍物检测是其重要一环,也是保障采棉机自主安全作业的技术基础。同时,高效、精准的棉田障碍物检测技术不仅能确保采棉机的安全行驶,还能预防障碍物堵塞机器影响棉花收获进程[3],进一步提高了采棉机作业效率和作业安全。

传统的棉田障碍物检测多依赖于人工巡视,不能很好地适应复杂多变的农田环境,存在检测效率低、实时性差等问题。随着深度学习技术和计算机硬件设备的不断提升,基于图像识别的障碍物检测技术逐渐成为研究的热点。文献[4-7]展示了深度学习在农田复杂环境下对障碍物实现实时检测的潜力,突出了提升模型性能和适应性的重要性,但未充分考虑边缘设备算力的限制。与此同时,众多学者也致力于轻量化目标检测网络相关研究,以兼顾模型的移植部署。文献[8-12]对当前主流单阶段目标检测算法进行了轻量化改进,为研究提供了参考依据和思路。

针对棉花田间复杂环境障碍物被遮挡致准确检测难、边缘设备算力有限的问题,本文提出一种基于改进 YOLO 11n 模型的田间障碍物检测方法。以采棉机为载体通过深度相机获取棉花田间障碍物数据,建立一个包含多类别棉田障碍物数据集;对 YOLO 11n 模型进行改进和训练;通过消融试验和对比试验对改进检测模型进行评估。以期为采棉机自主安全作业提供技术支撑。

1 试验材料

1.1 数据采集

本文数据集于 2024 年 10 月 10—30 日在新疆维吾尔自治区生产建设兵团通过实景拍摄获得。图像获取设备为 ZED 2i 型深度相机,以帧率 30 f/s 共采集 7 566 幅原始图像,分辨率为1 920 像素 × 1 080 像素,根据采棉机 5 m 安全制动距离调整相机角度,如图 1 所示。



图 1 数据采集场景

Fig. 1 Data collection scene

棉田收获作业环境中,障碍物分布可按其在棉田位置划分为 2 个区域:棉田间和棉田地头。棉田地头障碍物主要包括地头行人、车辆、农机具等。而棉田间障碍物又可以根据其相对于棉秆高度的不同划分为 3 个层级:棉花冠层以上、棉花冠层与地面间以及地面以下,棉田部分障碍物分布示意图如图 2 所示。棉花冠层上方障碍物主要包括树木、电线杆、行人、杂草堆、编织袋等。棉花冠层与地面间障碍物主要包括喷水桩、编织袋、杂草堆、棉膜和滴灌带等。而棉田间负障碍物则以水坑为主。

此次采集对象为棉田收获作业过程中常见且易影响采棉机正常作业的 4 种障碍物:喷水桩、杂草堆、编织袋和行人。为提高目标检测算法泛化能力,采集不同天气(晴天、阴天)、不同光照角度(顺光、逆光、夜间补光)以及不同遮挡情况下图像数据。为提高数据质量和标注效率,经过人工清洗,剔除重复图像和因曝光过度、灰尘较大导致的不合格图像,筛选出 5 000 幅较高质量图像。

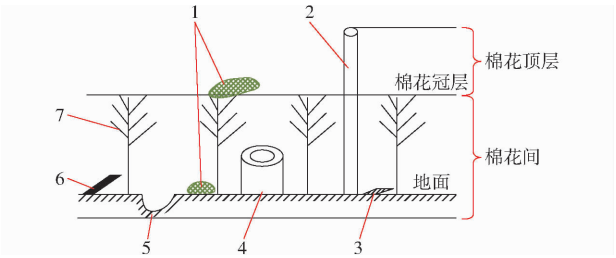

图 2　棉田部分障碍物分布示意图

Fig. 2　Diagram of partial obstacle distribution in cotton fields

1. 杂草堆　2. 电线杆　3. 编织袋　4. 喷水桩　5. 水坑　6. 滴灌带　7. 棉秆

1.2　数据预处理和图像增强

在棉田收获作业过程中，复杂的作业环境使得障碍物检测面临诸多挑战。采棉机在不同行驶状态产生的车体抖动，以及环境中的灰尘、光照变化等因素，导致采集的棉田障碍物数据存在显著差异。由于障碍物分布具有随机性，其受棉花及棉秆遮挡程度各不相同，这也进一步增加了特征提取的难度，从而影响检测模型准确性和泛化能力。

为增强检测模型泛化能力并最大程度地减少过拟合现象，对筛选后数据进行离线数据增强。通过随机亮度方式模拟晴天、阴天或遮挡光线时的光照变化，有效改善了数据集中光照条件分布不均的问题。同时，采用随机裁剪与随机翻转增强模型对多尺度、多方向目标的检测能力。为进一步提高模型对真实环境噪声的适应性，在图像中引入高斯噪声以模拟传感器和环境干扰。通过伪影增强技术模拟棉花及棉秆对障碍物的部分遮挡效应，使模型能够更好地处理特征缺失情况，从而提升在实际作业环境中的检测性能。棉田障碍物图像增强结果如图 3 所示。数据增强后数据集扩充至 8 634 幅。

使用 Labelme 软件对扩充后数据集中的障碍物所在最小外接矩形框进行人工标注，标注障碍物类别和位置信息。对标注完成后棉田障碍物图像按照比例 8∶1∶1进行划分，得到训练集、验证集和测试集图像各 6 907、863、864 幅，各类别障碍物划分具体数量如表 1 所示。

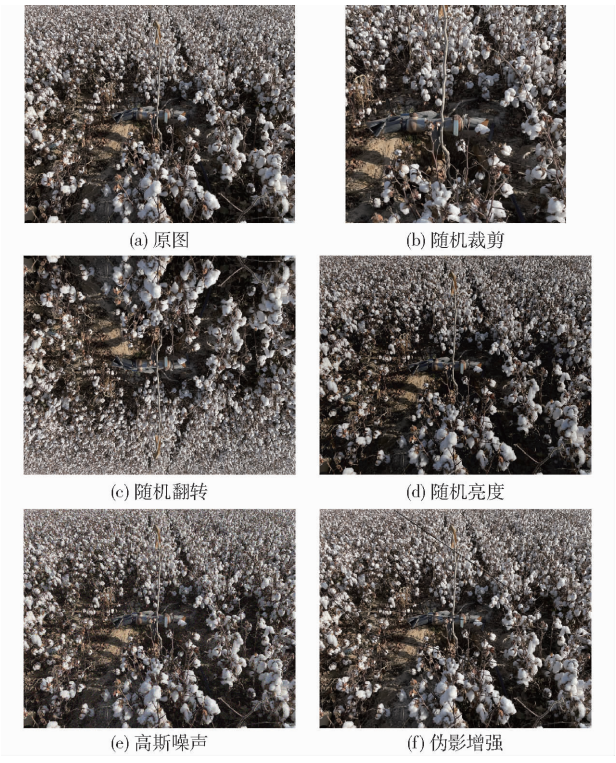

图 3　棉田障碍物图像增强示例

Fig. 3　Augmentation of obstacle images in cotton fields

表 1　各类别障碍物划分数量

Tab. 1　Numbers of obstacles in each category　幅

障碍物类别	训练集	验证集	测试集	总计
杂草堆	2 304	288	289	2 881
喷水桩	1 755	219	219	2 193
编织袋	1 648	206	206	2 060
行人	1 200	150	150	1 500
总计	6 907	863	864	8 634

2　检测方法

2.1　YOLO 11n 目标检测模型结构

YOLO 系列算法是单阶段目标检测算法的代表，使用锚框、分类和回归对目标进行定位，相比于两阶段目标检测算法，如：Fast R－CNN[13]、Faster R－CNN[14]、Mask R－CNN[15]，在网络结构上更加简洁高效，同时具备更快的检测速度。YOLO 11 模型是由 Ultralytics 公司推出的最先进的模型，相较于 YOLO v7[16]、YOLO v9[17]、YOLO v10[18]等版本，在轻量化、检测精度与速度上均有提升。

YOLO 11 模型主要由 4 部分组成：输入层（Input）、特征提取的主干网络（Backbone）、颈部网络（Neck）以及检测头（Head）。首先，输入层接收包含棉田障碍物的图像信息，通过归一化、Mosaic 数据增强和自适应图像缩放等方法进行数据预处理，将图像统一调整成 640 像素 ×640 像素。其次，主干网络负责提取棉田障碍物图像的纹理、语义信息等特征。颈部结构由特征金字塔网络（Feature pyramid network，FPN）和路径聚合网络（Path aggregation network，PAN）架构组成，开展特征融合工作，提高对语义信息和定位信息处理性能。最后，通过 3 个检测头分别进行 8、16、32 倍下采样，以实现棉田障碍物的多尺度检测。

YOLO 11 模型按照模型复杂性和参数数量划分为 5 个版本：YOLO 11n、YOLO 11s、YOLO 11m、

YOLO 11l、YOLO 11x。其中，YOLO 11n 模型在提供高精度检测的同时，提供最低的计算复杂度和最快的推理速度，使其成为在资源受限的边缘设备上进行实时目标检测的理想选择。对于棉田间复杂环境障碍物检测任务来说，实时性和高效性至关重要，因此本文选择 YOLO 11n 模型作为棉田障碍物检测的基线模型，并通过一系列针对性改进举措提升其对棉田障碍物检测的精度与效率。

2.2 棉花田间复杂环境障碍物检测方法

针对棉花田间复杂环境障碍物被遮挡导致准确检测难、边缘设备算力有限的问题，对 YOLO 11n 基线模型进行改进，改进模型命名为 YOLO 11n－SKS，网络结构如图 4 所示。在主干网络层，采用轻量级网络 StarNet 作为主要特征提取网络，并引入 DBA 模块用于重构 C2PSA，以增强多尺度特征之间的交互能力；在中间层部分，使用 KAGNConv 替换基线模型 C3k2 模块中的瓶颈结构，以实现对精细特征提取的同时，给予模型更高灵活性和可解释性；集成分离与增强注意力模块 SEAM 至检测头，进一步增强模型在遮挡场景中的检测鲁棒性，并显著提升整体模型的检测性能。

2.2.1 StarNet 特征提取网络

为了实现在边缘设备上的轻量级部署，选用 StarNet 网络作为 YOLO 11n 模型主要特征提取网络。该网络为 Microsoft 在 2024 年设计的高效轻量级网络，其性能优于 MobileNet[19]、GhostNet[20] 和 ShuffleNetV2[21] 等高效网络。

StarNet[22] 网络采用 4 级分层架构，利用卷积层进行下采样，并通过 Star Block 实现特征提取，其网络结构如图 4 所示。StarNet 网络核心优势在于通

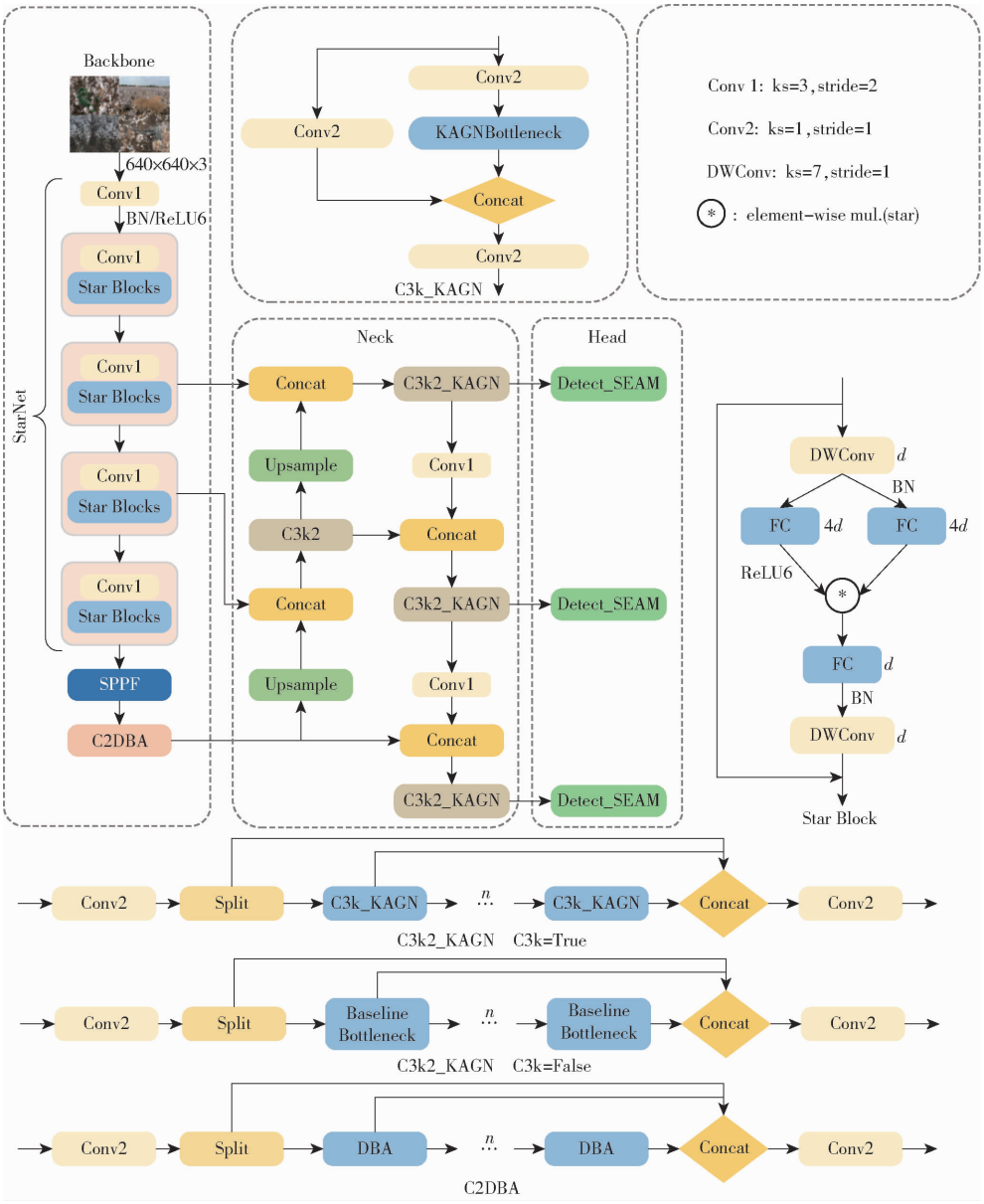

图 4 YOLO 11n－SKS 网络结构图

Fig. 4 Architecture of YOLO 11n－SKS

过简单的元素乘法将输入特征映射至一个隐式的高维特征空间，这种映射不仅有效地扩展了特征空间的维度，还能在保持较低计算复杂度的同时显著增强网络捕捉和建模复杂模式的能力。为实现输出标准化，使用批量归一化（Batch normalization，BN）取代了层归一化（Layer normalization，LN），并将其放置在深度卷积（Depth-wise convolution，DWConv）后。StarConvBN 模块为 StarNet 架构中的一种卷积模块，旨在高效提取和融合图像特征，同时实现计算资源优化。该模块首先通过二维卷积操作（步长为 1、膨胀因子为 1、卷积核大小为 1）生成特征图；随后，将生成的特征图输入到模块中并通过元素乘法运算，即“星型操作”，获得融合后特征。最后，对二维卷积输出特征使用批量归一化处理，以产生完整特征表示，如图 5 所示。



图 5　StarConvBN 模块

Fig. 5　StarConvBN module

Star Block 模块如图 4 所示，首先，通过深度卷积 DWConv 对输入特征进行初级特征提取；其次，使用批量归一化 BN 对输出进行标准化处理，提高处理效率；通过两路扩展因子为 4 的 1×1 卷积进行特征升维，其中一路经过 ReLU6 激活函数进行非线性转换；之后，两路输出经过元素乘法（Star operation）进行特征融合，在全局特征交互的同时指数级扩展特征空间的隐式维度；最后，通过深度卷积 DWConv 进一步提取高阶特征并完成特征降维。

设初始网络层宽度为 d，网络输入特征为 y，第 1 个星型操作的第 i、j 两路通道权重为 $\boldsymbol{w}^{i}_{(1,1)}$、$\boldsymbol{w}^{j}_{(1,2)}$，引入偏置后整体权重为 $\boldsymbol{W}_{(1,1)}$、$\boldsymbol{W}_{(1,2)}$。元素乘法使用“$*$”来表示，则第 l 个星型操作输出 $\boldsymbol{O}_l$ 为

$$\boldsymbol{O}_1 = \sum_{i=1}^{d+1}\sum_{j=1}^{d+1} \boldsymbol{w}^{i}_{(1,1)}\boldsymbol{w}^{j}_{(1,2)} y^{i} y^{j} \quad \left(\boldsymbol{O}_1 \in R^{\left(\frac{d}{\sqrt{2}}\right)^{2^1}}\right) \tag{1}$$

$$\boldsymbol{O}_2 = \boldsymbol{W}^{\mathrm{T}}_{(2,1)}\boldsymbol{O}_1 * \boldsymbol{W}^{\mathrm{T}}_{(2,2)}\boldsymbol{O}_1 \quad \left(\boldsymbol{O}_2 \in R^{\left(\frac{d}{\sqrt{2}}\right)^{2^2}}\right) \tag{2}$$

$$\boldsymbol{O}_3 = \boldsymbol{W}^{\mathrm{T}}_{(3,1)}\boldsymbol{O}_2 * \boldsymbol{W}^{\mathrm{T}}_{(3,2)}\boldsymbol{O}_2 \quad \left(\boldsymbol{O}_3 \in R^{\left(\frac{d}{\sqrt{2}}\right)^{2^3}}\right) \tag{3}$$

$$\vdots \tag{4}$$

$$\boldsymbol{O}_l = \boldsymbol{W}^{\mathrm{T}}_{(l,1)}\boldsymbol{O}_{l-1} * \boldsymbol{W}^{\mathrm{T}}_{(l,2)}\boldsymbol{O}_{l-1} \quad \left(\boldsymbol{O}_l \in R^{\left(\frac{d}{\sqrt{2}}\right)^{2^l}}\right) \tag{5}$$

由上述可知，表征向量空间维数为 $(d/\sqrt{2})^{2^l}$。假定模型星型操作次数为 10，输入通道数为 128 时，网络隐式特征维数约为 $90^{1\,024}$。足以证明，通过多层星型操作，StarNet 网络可以递归增加隐式特征维度到无限维度。

2.2.2　C2DBA 模块

在棉田收获作业场景下，障碍物尺寸和分布存在差异，其中杂草堆尤为突出。此外，受作物遮挡效应的影响，难以实现精确的目标检测。为此，引入 DBA 模块（Dynamic position bias attention block）用于重构 C2PSA，改进后 C2DBA 模块（Convolutional block with dynamic position bias attention block）如图 4 所示。该模块不仅可以增强不同尺度特征之间的交互能力，还能够根据目标物体位置和局部特征动态调整位置偏差，从而赋予感兴趣区域不同的关注权重。这使得网络能够在目标被遮挡的情况下依然有效提取特征。与 C2PSA 相比，该模块避免了对不相关区域的过度计算，减少计算资源的浪费，从而提升检测速度和效率。

DBA 模块由 1 个自注意力机制（Self-Attention）模块和 1 个多层感知机（Multilayer perceptron，MLP）模块组成，并将动态位置偏差（Dynamic position bias，DPB）嵌入其中，以获得位置信息，如图 6 所示。

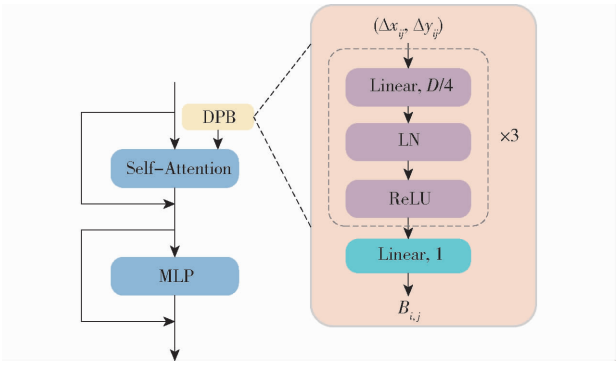

图 6　DBA 模块结构

Fig. 6　Architecture of DBA module

DPB[23] 为基于 MLP 模块动态生成的相对位置偏差，其非线性变换由 3 个全连接层组成，并结合层归一化和 ReLU 激活函数，网络结构如图 6 所示。DPB 输入维度为 2；将中间维度设置为 $D/4$，其中，D 为嵌入维度。表达式为

$$B_{i,j} = \mathrm{DPB}(\Delta x_{ij}, \Delta y_{ij}) \tag{6}$$

式中　DPB( )——动态位置偏差函数

$B_{i,j}$——位置(i,j)处偏置

Δx_{ij}——点$(i,j)$$x$ 方向上相对偏移

Δy_{ij}——点$(i,j)$$y$ 方向上相对偏移

2.2.3　KAGNConv 卷积模块

网络中间层性能表现主要依赖于 C3k2 跨阶段网络融合模块。为提升特征提取能力，采用 KAGNConv 模块替换基线模型 C3k2 模块中的瓶颈结构，并将改进后结构命名为 KAGNBottleneck。其中，C3k2 模块中的瓶颈结构由多个特征提取模块构

成,每个模块包含 2 个 3×3 卷积层,如图 7 所示。

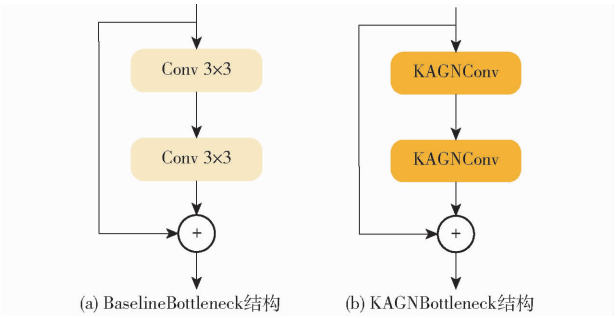

图 7 C3k2 改进前后瓶颈结构对比

Fig. 7 Comparison of two bottleneck structures in C3k2 module

KAN 网络[24](Kolmogorov-Arnold networks)是基于 Kolmogorov-Arnold 表示定理构建的深度学习框架,表示定理公式为

$$f(x)=f(x_1,x_2,\cdots,x_n)=\sum_{q=1}^{2n+1}\Phi_q\left(\sum_{p=1}^{n}\phi_{q,p}(x_p)\right) \tag{7}$$

其中,$\phi_{q,p}:[0,1]\rightarrow R$,$\Phi_q:R\rightarrow R$,该理论核心思想是任何多变量连续函数均可以表示为有限个单变量连续函数组合。

与传统神经网络采用固定激活函数不同,KAN 模型在网络边缘采用可学习的激活函数,使得每个权重参数均可使用一组可学习的单变量函数所替代,这些函数通常以样条函数形式参数化并能够使用更少的参数来模拟复杂函数,从而提供给模型更高灵活性和可解释性。激活函数 $\phi(x)$ 可分解为基函数 $b(x)$ 和样条函数 spline(x) 的线性组合,计算式为

$$\phi(x)=w_b b(x)+w_s\mathrm{spline}(x) \tag{8}$$

其中

$$b(x)=\mathrm{SiLU}(x)=\frac{x}{1+\mathrm{e}^{-x}} \tag{9}$$

$$\mathrm{spline}(x)=\sum_i c_i B_i(x) \tag{10}$$

式中 w_b——基函数可学习权重系数

w_s——样条函数可学习权重系数

c_i——$B_i(x)$ 函数系数

SiLU( )——SiLU 激活函数

$B_i(x)$——B 样条函数

基于 KAN 网络的理论架构,文献[25]提出一种卷积操作——KAN 卷积(Kolmogorov-Arnold network convolution),即将 KAN 模型中可学习的激活函数与传统卷积中的固定激活函数相融合。这种设计既保留了传统卷积操作的局部连接特性,又通过可学习的激活机制实现对复杂特征的高效提取。卷积核为 k 的 KAN 卷积可定义为

$$x_{ij}=\sum_{d=1}^{c}\sum_{a=0}^{k-1}\sum_{b=0}^{k-1}\phi_{a,b,c}(y_{d,i+a,j+b})$$
$$(i\in(1,n-k+1),j\in(1,m-k+1)) \tag{11}$$

其中 $y\in\mathbf{R}^{c\times n\times m}$

式中 c——图像通道数

n——图像高度 m——图像宽度

尽管 KAN 卷积能够有效提升模型特征提取能力,但基函数的引入导致模型参数量大量增加,提高了计算资源消耗,还增加了模型过拟合风险。为在模型复杂度与性能之间取得平衡,选用 KAGN 卷积层,结构框架如图 8 所示。改进 C3k2_KAGN 模块如图 4 所示。

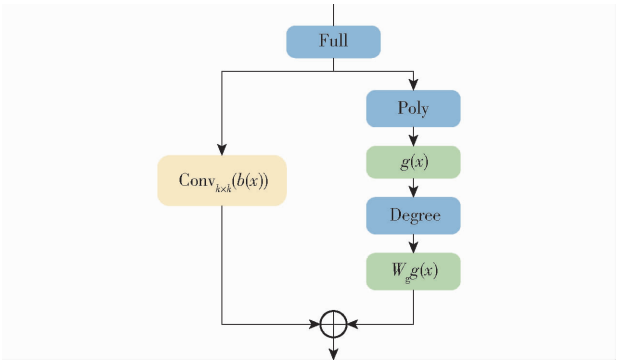

图 8 KAGN 卷积层

Fig. 8 KAGN Conv layer

2.2.4 SEAMHead 检测头

遗落在棉田中的编织袋与灌溉系统的喷水桩由于棉花与棉秆的遮挡,会导致部分原有特征丢失,导致模型难以准确识别和定位目标。因此,将分离与增强注意力模块[26](Separated and enhancement attention module, SEAM)集成至检测头,利用特征图之间的关系在一定程度上弥补遮挡所导致的特征信息丢失,进而提升模型检测准确率。

SEAM 模块可以实现 3 个主要目标:帮助检测被遮挡障碍物、强化图像中障碍物所在区域特征、弱化背景区域干扰。其主要结构由 3 部分组成,网络结构图如图 9 所示。首先,采用带残差连接的深度可分离卷积结构。该卷积对每个通道进行独立的分离卷积操作,能够有效学习各通道的重要性并降低模型参数量,但在此过程中忽略了通道间的交互关系。为弥补这种损失,将深度卷积的输出通过逐点卷积(Pointwise convolution, PWC)进行通道间信息融合。其次,引入 2 个全连接层进一步提取并聚合每个通道的特征信息。这些层通过捕捉空间和通道维度上的长距离关联,建立遮挡与非遮挡目标之间的依赖关系。这个过程有助于传播上下文信息,补偿由于遮挡造成的损失,从而增强网络重建不完整特征的能力,为检测头提供更丰富和完善的决策信

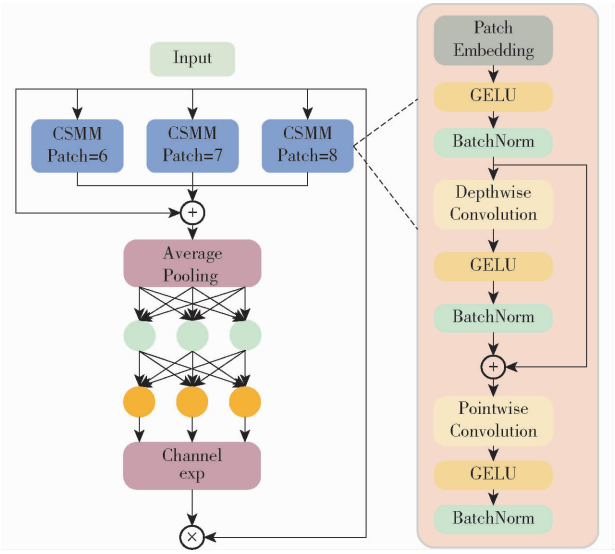

图 9　SEAM 模块结构

Fig. 9　Architecture of SEAM module

息。全连接层的输出范围经过指数函数 $y_i = e^{x_i}$ 从[0,1]扩展到[1,e],通过这种指数归一化建立单调映射关系,增强模型对位置偏差的容忍度。最后,将 SEAM 模块输出作为注意力权重与原始特征进行结合,使得模型在处理遮挡时,能够更准确地聚焦于被部分遮挡的编织袋与喷水桩。

2.3　试验平台与训练方法

试验平台为 Windows 11 系统,采用 PyTorch 深度学习框架,CPU 型号为 Intel Core i7 - 14700;GPU 型号为 NVIDIA GeForce RTX 4070 Ti,16 GB 显存;编程语言为 Python。

为保证试验结果可信度,在整个模型训练过程中将使用相同的训练环境与参数配置。使用随机梯度下降(Stochastic gradient descent, SGD)优化器优化网络,迭代次数设置为 800,批处理大小设置为 32,初始学习率设置为 0.01。

2.4　评价指标

采用精确率(Precision)、召回率(Recall)、平均精度均值(mAP_{50}、mAP_{50-95})衡量模型精度;采用模型参数量(Parameters)、浮点数运算量(FLOPs)衡量模型复杂度;采用帧率(FPS)衡量模型检测效率。

3　试验结果与分析

3.1　消融试验

以 YOLO 11n 模型为基线模型,通过消融试验来验证对主干网络层、中间层以及检测头改进的有效性,试验结果如表 2 所示。

表 2　消融试验结果

Tab. 2　Ablation test results

StarNet	C2DBA	C3k2_KAGN	SEAMHead	精确率/%	召回率/%	mAP_{50}/%	mAP_{50-95}/%	浮点数运算量/FLOPs	参数量
×	×	×	×	89.4	86.2	90.6	60.9	6.3×10^9	2.58×10^6
√	×	×	×	88.9	85.4	90.2	60.6	5.0×10^9	1.94×10^6
√	√	×	×	88.6	87.0	90.8	60.4	4.9×10^9	1.89×10^6
√	√	√	×	90.5	88.6	91.5	61.5	4.9×10^9	2.43×10^6
√	√	√	√	91.7	88.3	91.9	62.3	4.4×10^9	2.14×10^6

注:√表示引入该模块,×表示未引入该模块。

从表 2 得知,在主干网络中,引入 StarNet 轻量级网络后,模型参数量减少 24.8%、浮点数运算量减少 1.3×10^9 FLOPs,但也不可避免地使检测精度和召回率有轻微下降,因为在轻量化模型复杂度的同时,减弱了模型对部分重要特征的提取能力。故同时在主干网络引入 DBA 模块用于重构 C2PSA,使得模型动态调整位置偏差,对感兴趣区域赋予不同关注权重,改进后召回率提升 1.6 个百分点、mAP_{50} 提升 0.6 个百分点,模型复杂度略有下降。在颈部网络中,采用 KAGNConv 模块对基线模型 C3k2 模块中的瓶颈结构改进后,模型参数量增加 5.4×10^5,使计算复杂度增加,但精确率提升 1.9 个百分点、召回率提升 1.6 个百分点、mAP_{50} 提升 0.7 个百分点、mAP_{50-95} 提升 1.1 个百分点且浮点数运算量与主干网络改进后模型保持不变,提高了模型表达与特征提取的能力。将分离与增强注意力模块 SEAM 融合至检测头,增强模型在被遮挡的复杂环境下的目标识别能力,同时优化计算效率,使浮点数运算量下降 10.2%。相比基线模型,最终改进网络模型 YOLO 11n - SKS 精确率、召回率、mAP_{50}、mAP_{50-95} 分别提升 2.3、2.1、1.3、1.4 个百分点,模型浮点数运算量减少 1.9×10^9 FLOPs,模型参数量减少 17.1%。

3.2　不同颈部网络瓶颈结构对比试验

为验证不同颈部网络瓶颈结构优化效果,以在 YOLO 11n 主干网络改进模型基础上进行多组试验。引入 EIEM[27]、Feature Modulation Block[28]、RepViTBlock[29]、KAGNBlock 改进瓶颈结构,进行对比,试验结果如表 3 所示。

表 3 瓶颈结构对比试验结果

Tab. 3 Experimental results comparing bottleneck structures

模型	精确率/%	召回率/%	mAP_{50}/%	mAP_{50-95}/%	浮点数运算量/FLOPs	参数量
YOLO 11n + StarNet + C2DBA	88.6	87.0	90.8	60.4	4.9×10^9	1.89×10^6
YOLO 11n + StarNet + C2DBA + C3k2_EIEM	88.9	86.2	90.4	59.0	5.1×10^9	1.87×10^6
YOLO 11n + StarNet + C2DBA + C3k2_FMB	87.7	84.7	90.0	57.9	5.2×10^9	1.95×10^6
YOLO 11n + StarNet + C2DBA + C3k2_RVB	87.9	86.4	90.0	59.1	4.7×10^9	1.73×10^6
YOLO 11n + StarNet + C2DBA + C3k2_KAGN	90.5	88.6	91.5	61.5	4.9×10^9	2.43×10^6

由表 3 可知，采用 KAGN 模块与其他模块相比，精确率、召回率、平均精度均值 mAP_{50} 与 mAP_{50-95} 均表现出较优效果，分别为 90.5%、88.6%、91.5%、61.5%，且浮点数运算量与改进前模型保持不变。模型参数量虽有上涨，但仍低于基线模型 YOLO 11n 的模型参数量，保持较低的计算开销。综上所述，采用 KAGN 模块替换基线模型 C3k2 模块中的瓶颈结构，使得该检测任务的综合性能显著提升，充分证明选择 KAGN 模块改进的合理性。

3.3 不同检测模型综合对比试验

为验证本文所改进模型的可靠性，选取当前主流且代码开源的轻量级目标检测算法（YOLO v8n、YOLO v9 - t、YOLO v10n）在自建棉田障碍物数据集上进行训练，各模型检测结果如表 4 所示。

表 4 检测模型对比试验结果

Tab. 4 Experimental results of detection models comparison

模型	精确率/%	召回率/%	mAP_{50}/%	浮点数运算量/FLOPs	参数量	帧率/(f·s^{-1})
YOLO v8n	87.6	88.0	90.6	8.1×10^9	3.01×10^6	98
YOLO v9 - t	88.4	87.6	90.1	1.1×10^{10}	2.62×10^6	78
YOLO v10n	87.9	84.6	89.3	8.2×10^9	2.69×10^6	90
YOLO 11n	89.4	86.2	90.6	6.3×10^9	2.58×10^6	102
YOLO 11n - SKS	91.7	88.3	91.9	4.4×10^9	2.14×10^6	85

改进模型 YOLO 11n - SKS 中检测帧率低于除 YOLO v9 - t 外的其他目标检测模型，使得检测速率下降，但其拥有较高的精确率、召回率和平均精度均值 mAP_{50}，如图 10 所示。该模型所需参数量仅为 2.14×10^6，且相较于 YOLO v8n、YOLO v9 - t、YOLO v10n、YOLO 11n 模型，浮点数运算量分别降低 45.6%、60.0%、46.3%、30.1%。YOLO 11n - SKS 中检测速度略有下降，但仍达 85 f/s，能满足采棉机作业过程中的实时障碍物检测需求。综上所述，在保持高精度检测前提下，YOLO 11n - SKS 模型降低了计算成本，减少了模型参数量，更适合部署在有限资源的边缘设备中。

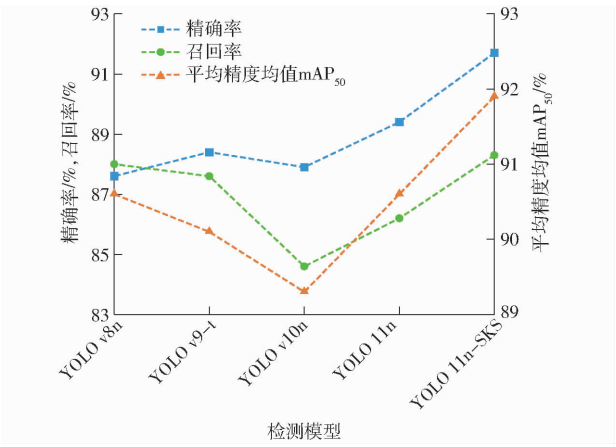

图 10 不同检测模型性能

Fig. 10 Performance of different detection models

为更直观对比改进模型与基线模型检测效果，对棉花田间复杂环境障碍物检测结果进行展示，如图 11 所示。改进模型目标置信度在多类别障碍物下均有提升，能够较好地处理复杂棉田环境下被遮挡的障碍物。综合表明，YOLO 11n - SKS 模型具备更强棉田障碍物特征提取与分析能力，可为采棉机环境感知系统提供可靠的技术支撑。

在未来研究中，扩展棉田障碍物检测的类别范围，以提升模型的泛化能力和实用性；结合其他传感器与相机进行多模态信息融合提升棉株交错复杂环境下的障碍物检测精度，特别是针对相对密集区域的目标识别难题；基于多层次环境感知信息，进一步深化避障策略研究，提升采棉机自主作业能力。

4 结论

（1）针对棉田收获作业复杂环境，提出了一种根据相对于棉秆高度不同划分棉田间障碍物层级的方法，为棉田障碍物检测提供一种思路。提出了一种基于 YOLO 11n 的棉花田间复杂环境下障碍物检

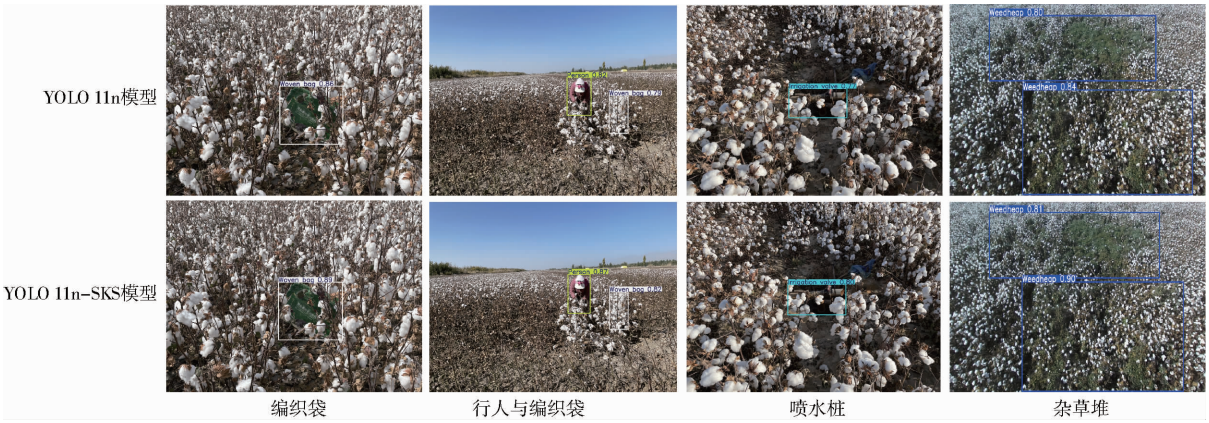

图 11　基线与改进模型检测结果对比

Fig. 11　Comparisons of detection results between baseline and improved models

测模型 YOLO 11n－SKS，在主干网络层，采用轻量级网络 StarNet 作为主要特征提取网络，并引入 DBA 模块用于重构 C2PSA；采用 KAGNConv 替换基线模型 C3k2 模块中的瓶颈结构完成颈部网络改进；集成分离与增强注意力模块 SEAM 至检测头。

（2）改进后 YOLO 11n－SKS 模型与基线模型相比精确率、召回率、mAP_{50}、mAP_{50-95} 分别提升 2.3、2.1、1.3、1.4 个百分点，模型浮点数运算量仅为 4.4×10^9 FLOPs，模型参数量减少 17.1%。YOLO 11n－SKS 对棉田障碍物检测效果更加显著且检测速度满足实时性需求，为采棉机自主安全作业提供技术支撑。

参　考　文　献

[1] 何创新，巩蕾，苗中华，等. 基于平行轨迹导航的采棉机自动对行控制方法[J]. 农业机械学报，2024，55(6)：34－41.
HE Chuangxin，GONG Lei，MIAO Zhonghua，et al. Automatic row-follow control method for cotton picker based on parallel trajectory navigation[J]. Transactions of the Chinese Society for Agricultural Machinery，2024，55(6)：34－41.（in Chinese）

[2] 张漫，季宇寒，李世超，等. 农业机械导航技术研究进展[J]. 农业机械学报，2020，51(4)：1－18.
ZHANG Man，JI Yuhan，LI Shichao，et al. Research progress in agricultural machinery navigation technology[J]. Transactions of the Chinese Society for Agricultural Machinery，2020，51(4)：1－18.（in Chinese）

[3] 王皓，韩科立，韩树杰，等. 基于粒子群算法和 SDAE 的采棉头故障诊断研究[J]. 农业机械学报，2023，54(增刊 2)：164－172.
WANG Hao，HAN Keli，HAN Shujie，et al. Fault diagnosis method and experiment of cotton picking head based on particle swarm optimization algorithm and SDAE[J]. Transactions of the Chinese Society for Agricultural Machinery，2023，54(Supp. 2)：164－172.（in Chinese）

[4] 杨昊霖，王其欢，李华彪，等. 基于改进 YOLOv5 的田间复杂环境障碍物检测[J]. 中国农机化学报，2024，45(6)：216－222，256.
YANG Haolin，WANG Qihuan，LI Huabiao，et al. Obstacle detection in complex farmland environment based on improved YOLOv5[J]. Journal of Chinese Agricultural Mechanization，2024，45(6)：216－222，256.（in Chinese）

[5] 蔡舒平，潘文浩，刘慧，等. 基于 D2－YOLO 去模糊识别网络的果园障碍物检测[J]. 农业机械学报，2023，54(2)：284－292.
CAI Shuping，PAN Wenhao，LIU Hui，et al. Orchard obstacle detection based on D2－YOLO deblurring recognition network[J]. Transactions of the Chinese Society for Agricultural Machinery，2023，54(2)：284－292.（in Chinese）

[6] QIU Z J，ZHAO N，ZHOU L，et al. Vision-based moving obstacle detection and tracking in paddy field using improved YOLOv3 and Deep SORT[J]. Sensors，2020，20(15)：4082.

[7] SU F，ZHAO Y P，SHI Y X，et al. Tree trunk and obstacle detection in apple orchard based on improved YOLOv5s model[J]. Agronomy，2020，12(10)：2427.

[8] 王卫星，刘泽乾，高鹏，等. 基于改进 YOLO v4 的荔枝病虫害检测模型[J]. 农业机械学报，2023，54(5)：227－235.
WANG Weixing，LIU Zeqian，GAO Peng，et al. Detection of litchi diseases and insect pests based on improved YOLO v4 model[J]. Transactions of the Chinese Society for Agricultural Machinery，2023，54(5)：227－235.（in Chinese）

[9] 乔琛，韩梦瑶，高苇，等. 基于 Faster－NAM－YOLO 的黄瓜霜霉病菌孢子检测[J]. 农业机械学报，2023，54(12)：288－299.
QIAO Chen，HAN Mengyao，GAO Wei，et al. Quantitative detection of cucumber downy mildew spores at multi-scale based on Faster－NAM－YOLO[J]. Transactions of the Chinese Society for Agricultural Machinery，2023，54(12)：288－299.（in Chinese）

[10] 冀汶莉，刘洲，邢海花. 基于 YOLO v5 的农田杂草识别轻量化方法研究[J]. 农业机械学报，2024，55(1)：212－222，293.
JI Wenli，LIU Zhou，XING Haihua. Research on lightweight method for weed recognition in farmland based on YOLO v5[J]. Transactions of the Chinese Society for Agricultural Machinery，2024，55(1)：212－222，293.（in Chinese）

[11] FAN X, SUN T, CHAI X, et al. YOLO - WDNet: a lightweight and accurate model for weeds detection in cotton field[J]. Computers and Electronics in Agriculture, 2024, 225: 109317.
[12] ZENG T, LI S, SONG Q, et al. Lightweight tomato real-time detection method based on improved YOLO and mobile deployment[J]. Computers and Electronics in Agriculture, 2023, 205: 107625.
[13] GIRSHICK R. Fast R - CNN[C]// Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440 - 1448.
[14] REN S, HE K, 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, 2016, 39(6): 1137 - 1149.
[15] HE K M, GKIOXARI G, DOLLAR P, et al. Mask R - CNN[C]// Proceedings of the IEEE International Conference on Computer Vision, 2017: 2961 - 2969.
[16] 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]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464 - 7475.
[17] WANG C Y, YEH I H, LIAO H Y M. YOLOv9: learning what you want to learn using programmable gradient information[C]// European Conference on Computer Vision, 2024: 1 - 21.
[18] WANG A, CHEN H, LIU L, et al. YOLOv10: real-time end-to-end object detection[J]. arXiv Preprint, arXiv: 2405.14458, 2024.
[19] HOWARD A, SANDLER M, CHU G, et al. Searching for MobileNetv3[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 1314 - 1324.
[20] HAN K, WANG Y, TIAN Q, et al. GhostNet: more features from cheap operations[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 1580 - 1589.
[21] MA N, ZHANG X, ZHENG H T, et al. ShuffleNet v2: practical guidelines for efficient CNN architecture design[C]// Proceedings of the European Conference on Computer Vision (ECCV), 2018: 116 - 131.
[22] MA X, DAI X, BAI Y, et al. Rewrite the stars[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 5694 - 5703.
[23] WANG W, YAO L, CHEN L, et al. CrossFormer: a versatile vision transformer hinging on cross-scale attention[J]. arXiv Preprint, arXiv: 2108.00154, 2021.
[24] LIU Ziming, WANG Yixuan, SACHIN V, et al. Kan: Kolmogorov-Arnold networks[J]. arXiv Preprint, arXiv: 2404.19756, 2024.
[25] DROKIN I. Kolmogorov-Arnold convolutions: design principles and empirical studies[J]. arXiv Preprint, arXiv: 2407.01092, 2024.
[26] YU Z, HUANG H, CHEN W, et al. YOLO - Facev2: a scale and occlusion aware face detector[J]. Pattern Recognition, 2024, 155: 110714.
[27] ZHANG J, DING A, LI G, et al. A pyramid attention network with edge information injection for remote-sensing object detection[J]. IEEE Geoscience and Remote Sensing Letters, 2023, 20: 1 - 5.
[28] ZHENG M, SUN L, DONG J, et al. SMFANet: a lightweight self-modulation feature aggregation network for efficient image super-resolution[C]// European Conference on Computer Vision, 2024: 359 - 375.
[29] WANG A, CHEN H, LIN Z, et al. Repvit: revisiting mobile CNN from vit perspective[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 15909 - 15920.