

基于改进 YOLO 11 模型的棉田地表残膜识别方法研究

孟庆建^{1,2} 翟志强^{1,2} 张连朴^{1,2} 吕继东^{1,2} 王虎挺^{1,2} 张若宇^{1,2}

(1. 石河子大学机械电气工程学院, 石河子 832003; 2. 农业农村部西北农业装备重点实验室, 石河子 832003)

摘要:为实现残膜回收机在自然环境中快速、准确地识别棉田地表残膜目标,本文提出了一种基于 DCA-YOLO 11 轻量化模型的棉田地表残膜识别方法。以 4JMLE-210 型残膜回收机工作后棉田地表残膜为研究对象,在不同时间段采集地表残膜图像 900 幅,通过透视变换、图像裁剪、数据清洗、数据增强等预处理,最终得到 5 215 幅残膜样本图像,按照 4:1 划分为训练集和测试集,实现了对棉田地表残膜的数据集构建。通过在 YOLO 11 模型主干网络中增加深度可分离卷积(DWConv)模块代替通用卷积(Conv)模块,用于减少计算复杂度和参数量;通过在输出检测端末尾加入 CBAM 卷积块注意力机制模块来提高模型的感知能力,减少边缘与背景干扰;通过用 ADown 模块替换主干网络中的 Conv 模块,实现残膜特征图不同层之间的下采样,减少特征图空间维度,保留关键信息来提高残膜目标检测准确性。试验结果表明,在复杂自然环境下,DCA-YOLO 11 模型精确率 P 为 81.9%,召回率 R 为 80.9%,平均精度均值 mAP(重叠率 0.5) 为 86.7%,参数量为 2.20×10^6 ,处理速度为 80 f/s。通过对不同模型进行对比试验,DCA-YOLO 11 模型检测精确率比 YOLO v10、YOLO v9、YOLO v8 分别高 2.9、2.3、3.8 个百分点,召回率比 YOLO v10、YOLO v9、YOLO v8 分别高 2.0、1.0、1.8 个百分点,处理速度比 YOLO v9、YOLO v8 分别提升 12.7%、14.2%,略低于 YOLO v10。DCA-YOLO 11 模型在保证精度的同时,模型最小,参数量最少,证明其轻量化与优越性。模型通过泛化性试验,其在验证数据集上的检测结果, R^2 为 0.72,平均绝对误差和均方根误差分别为 4.92 个和 2.72 个,提出的 DCA-YOLO 11 轻量化模型泛化性较好。该研究可为残膜回收机械在复杂环境下精准、高效捡拾残膜以及残膜回收机回收率车载视觉估测提供理论依据与数据基础。

关键词: 残膜识别; YOLO 11 模型; 目标检测; 模型轻量化

中图分类号: S24

文献标识码: A

文章编号: 1000-1298(2025)05-0017-09

OSID:



Recognition Method of Cotton Field Surface Residual Film Based on Improved YOLO 11

MENG Qingjian^{1,2} ZHAI Zhiqiang^{1,2} ZHANG Lianpu^{1,2} LÜ Jidong^{1,2} WANG Huting^{1,2} ZHANG Ruoyu^{1,2}

(1. College of Mechanical and Electrical Engineering, Shihezi University, Shihezi 832003, China

2. Key Laboratory of Northwest Agricultural Equipment, Ministry of Agriculture and Rural Affairs, Shihezi 832003, China)

Abstract: In response to the issue of estimating the recovery rate of residual film in cotton fields by current residual film recovery machines, a lightweight residual film recognition method named DCA-YOLO 11 was proposed, which enabled rapid and accurate identification of residual film on cotton field surfaces in natural environments. Taking the residual film on cotton field surfaces after the operation of the 4JMLE-210 residual film recovery machine as the research object, totally 900 images of residual film were collected at different time periods. Through preprocessing steps such as perspective transformation, image cropping, data cleaning, and data augmentation, a dataset of 5 215 residual film sample images was constructed, which was divided into training and test sets at a 4:1 ratio. To enhance the model's performance, a depthwise convolution (DWConv) module was added to the backbone network of YOLO 11 to replace a standard convolution (Conv) module, thereby reducing computational complexity and the number of parameters. Additionally, a CBAM attention mechanism module was incorporated at the end of

收稿日期: 2025-03-10 修回日期: 2025-03-27

基金项目: 国家重点研发计划项目(2022YFD2002400)、兵团科技攻关计划项目(2023AB014、2022DB003)、新疆棉花产业技术体系项目(XJARS-03)和自治区科技支疆计划项目(2024E02016)

作者简介: 孟庆建(1984—),男,博士生,主要从事智能检测与装备研究,E-mail: 252720977@qq.com

通信作者: 张若宇(1980—),男,教授,博士生导师,主要从事数字棉花技术与装备研究,E-mail: ry248@163.com

the detection output to improve the model's perception capability and reduce interference from edges and backgrounds. Furthermore, the ADown module was used to replace the Conv module in the backbone network, enabling downsampling between different layers of the residual film feature maps, reducing the spatial dimensions of the feature maps while retaining key information to improve the accuracy of residual film target detection. Experimental results demonstrated that the DCA - YOLO 11 model achieved a precision (P) of 81.9%, a recall (R) of 80.9%, and a mean average precision (mAP) of 86.7% (at an IoU threshold of 0.5) in complex natural environments. The model has about 2.20 million parameters, and an FPS of 80 f/s. Comparative experiments with other models showed that DCA - YOLO 11 outperformed YOLO v10, YOLO v9 and YOLO v8 in precision by 2.9 percentage points, 2.3 percentage points, 3.8 percentage points. In terms of recall, it was improved by 2.0 percentage points, 1.0 percentage points, and 1.8 percentage points compared with that of YOLO v10, YOLO v9, and YOLO v8, respectively. While its processing speed was slightly lower than that of YOLO v10, and it surpassed YOLO v9 and YOLO v8 by 12.7% and 14.2%. DCA - YOLO 11 achieved the smallest model size and the fewest parameters while maintaining high accuracy, demonstrating its lightweight design and superiority. Through generalization test, the model's detection results on the validation dataset showed an R^2 of 0.72, a mean absolute error (MAE) of 4.92 pcs and a root mean square error (RMSE) of 2.72 pcs, indicating good generalization. The research result can provide a theoretical foundation and data support for the precise and efficient collection of residual film by recovery machinery in complex environments, as well as for the visual estimation of the recovery rate of residual film recovery machines.

Key words: residual film identification; YOLO 11 model; target detection; model lightweight

0 引言

新疆棉田地膜多采用聚乙烯材料,在自然条件下降解时间大约 200 年以上^[1],对环境造成“白色污染”。如果地膜不及时回收,将会对土壤造成污染和损害,影响作物生长发育^[2]。近年来,针对残膜回收,我国研发人员先后开发了多种残膜回收机^[3-7],这些残膜回收机基本上无法捡拾尺寸较小的残膜并且残膜回收率也无法实时估测。

为了实现农田残膜的精准捡拾,提高残膜回收机的回收率,江水泉等^[8]提出了利用最小外接矩形法确定残膜目标区域形心来确定残膜位置,残膜形心位置偏差在 1.2 cm 以内,但传统的识别方法依赖于阈值的选择,适应性不好、鲁棒性差。吴雪梅等^[9]基于无人机采集的烟田残膜图像,提出了一种基于颜色特征的残膜识别方法,利用该方法对不同时期的农田残膜进行识别,6 叶期、烟叶收获后、烟杆拔除后和冬季空闲期识别率分别为 96.99%、69.47%、93.55% 和 88.95%,地膜覆盖周期的平均识别率为 87.49%,但利用无人机航拍获取大尺度地膜图像很难直接将识别方法应用于残膜回收机上。张学军等^[10]提出了一种基于 Faster R - CNN 卷积神经网络的农田残膜识别方法,该方法是一种两阶段的目标检测方法,其模型大,参数量多,很难实现轻量化。翟志强等^[11]为了快速准确地监测播前棉田地表残膜污染,提出了一种基于像素块和机器学习的播前棉田地表残膜覆盖率检测方法,结果表

明 20 像素 × 20 像素块下采用人工神经网络算法,残膜覆盖率检测值与真实值的相对误差最小为 0.51%,检测时间最短为 0.29 s。牛玉珩^[12]提出一种基于改进 SegFormer 模型的残膜快速识别分割方法,SegFormer 原始模型平均交并比已达到 83.00%,改进模型较原始模型平均交并比提升 0.42 个百分点,单幅图像检测时间为 51.13 ms。文献[11-12]都是以无人机低空采集图像,很难满足残膜回收机上轻量化、实时检测的要求。

由于田间实际环境中采集的图像含有大量的噪声(光照不均匀、复杂背景等),对目标图像识别、特征提取有非常大的影响,导致目标识别的实时性、准确性难以在实际应用中得到推广。随着人工智能的发展,科研工作者对自然环境下目标对象识别进行了深入研究,起初以机器学习为主^[13-16],后来深度学习成为研究的热点,而卷积神经网络(Convolutional neural network, CNN)是目前最有效的深度学习方法之一^[17-21],能够直接以图像作为输入,具有不受光照、背景复杂影响以及识别准确率高等优点,已在杂草和害虫识别^[22-23]、果实识别^[24-27]、农业遥感图像分类^[28-29]等领域得到了初步的应用。

本文提出一种实现棉田残膜精准识别并且达到实时性要求的识别方法。首先,建立残膜回收机工作后棉田地表残膜图像数据集;然后,针对背景与边缘干扰,提出一种改进的 YOLO 11 模型,并进行轻量化,实现棉田残膜的自动识别。以期为残膜回收

机械在复杂环境下精准、高效捡拾残膜以及残膜回收机回收率评估提供理论依据与数据基础。

1 材料与方法

1.1 研究区域

试验图像分别采集于 2023 年 11 月 3—4 日,采集地点为新疆维吾尔自治区石河子市北泉镇三分场,坐标为 (44.37°N,86.01°E),上一季度种植作物为棉花,地表平整,但残留破碎的地膜,如图 1 所示。

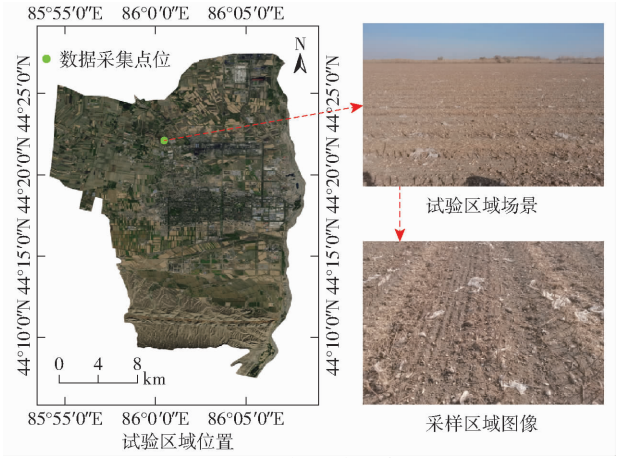


图 1 研究区域示意图

Fig. 1 Sketch map of study area

1.2 数据采集

以 4JMLE-210 型残膜回收机工作后遗留在农田表面的残膜为研究对象,采用深圳鹰眼航拍科技有限公司的 FIREFLY9S 型运动相机(相机镜头距地面高度约 1 600 mm,与地面的拍摄角度为 45°~60°),在自然光、低光照和复杂背景等不同场景下采集图像 900 幅,图像格式为 JPG,分辨率为 4 608 像素×3 456 像素,采集图像示例如图 2 所示。

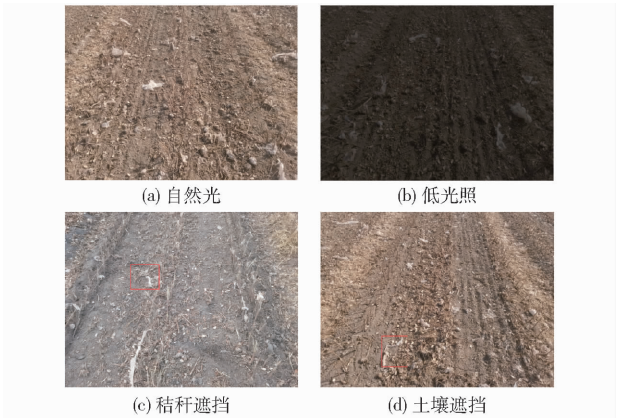


图 2 采集图像示例

Fig. 2 Acquired image examples

1.3 数据预处理与样本数据集构建

1.3.1 透视变换

为适应后期相机安装在残膜回收机上,数据采

集时使相机具有一定拍摄角度,本文采用图像透视变换方法将采集的图像进行校正,它通过一种射影变换,会保持直线的性质,即变换前的直线在变换后依然是直线^[30]。透视变换核心原理是找到两个图像平面上 4 个点之间的映射关系。

透视变换矩阵方程描述为

$$\begin{bmatrix} x' \\ y' \\ w \end{bmatrix} = \begin{bmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (1)$$

其中, (x,y) 是原始图像点的坐标, (x',y') 是变换后图像中点的坐标, h_{ij} 是 3×3 变换矩阵的元素, w 是归一化参数,它通常用于将透视变换后的坐标归一化,使其变为实际坐标值。如图 3 所示,标定框尺寸为 $1.5 \text{ m} \times 1.5 \text{ m}$ 。

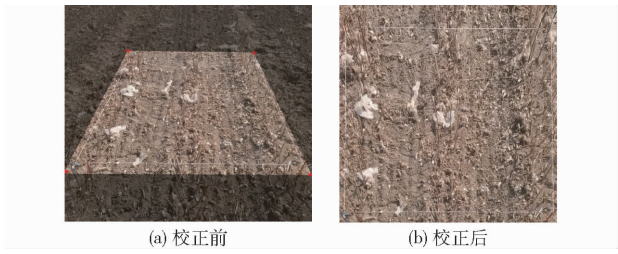


图 3 校正前后图像

Fig. 3 Images before and after correction

1.3.2 图像裁剪与标注

为了保证数据参数准确性以及数据丰富性,借助 Matlab 软件将 1 幅 4 608 像素×3 456 像素的大图像裁剪成 8 幅 1 152 像素×1 728 像素的小图像,如图 4 所示。在模型训练前使用标注工具 Labeling 对数据进行人工标注,由于部分图像中落地棉与残膜极其相似,无法实现人工标注,经过数据清洗,最后构成样本数据集 5 215 幅图像,采用 Pascal VOC 标注格式,其中残膜标注为 residual_film,标注时要把整个残膜的外接矩形作为真实框,标注完成后,生成 .xml 类型的标注文件,将标注完成的数据集按 4:1 的比例划分成训练集和测试集。

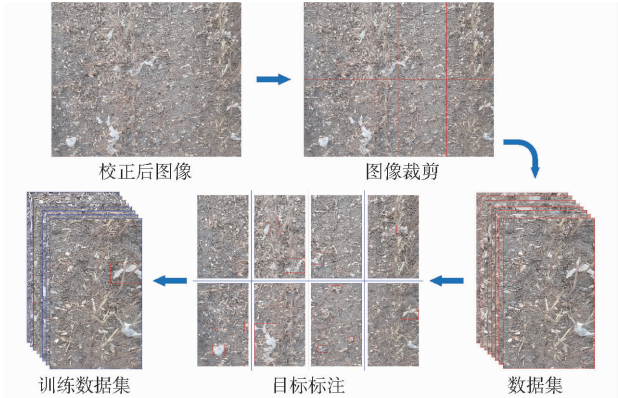


图 4 图像的裁剪与标注

Fig. 4 Annotation and cropping of images

1.4 研究方法

1.4.1 YOLO 11 目标检测网络结构与解析

YOLO 11 模型是 Ultralytics 公司开发的一种集图像分类、目标检测、语义分割等任务于一体的神经网络,属于 YOLO(You only look once)系列的最新迭代版本^[31]。该版本在模型架构上进行了优化和改进,提升了模型的检测性能和速度。相较于以前的版本,YOLO 11 模型在保持实时性的同时,提高了目标检测的准确率。它支持多尺度检测,能够对不同分辨率的目标进行检测,从而更好地适应各种场景和需求。

YOLO 11 模型包括主干网络(Backbone)、颈部网络(Neck)、检测头部网络(Head)等部分,如图 5 所示。主干网络负责输入图像的特征提取,检测头部网络部分延续了 YOLO v8 模型的结构基础,其中 C3k2 模块在代码中通过设置 C3k 参数为 True 或 False 来决定其具体行为,当 C3k 为 False 时,C3k2 模块就是普通的 C2F 模块;当 C3k 为 True 时,会把模块替换成 C3 模块,C3 模块是一种具有特定结构的模块,用于更高效地提取特征。这种可切换的设计使得模型能够根据不同的需求和场景选择合适的特征提取方式,增强了模型适应性和灵活性。

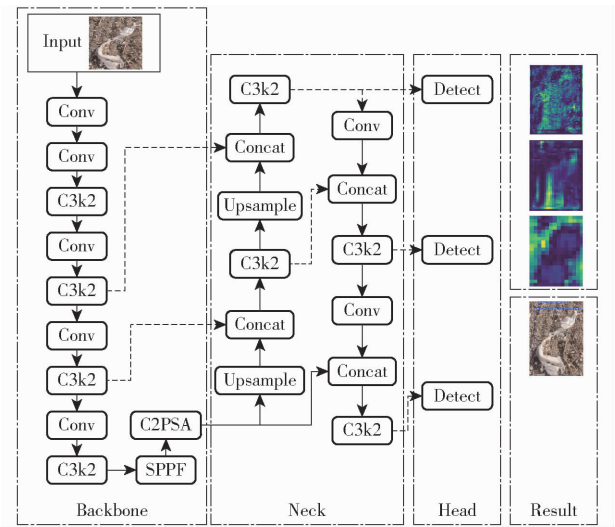


图 5 YOLO 11 模型结构示意图
Fig. 5 YOLO 11 model architecture diagram

主干网络中 C2PSA (Convolutional block with Parallel Spatial Attention) 是个关键模块,是对 C2F 模块的扩展,如图 6 所示。该模块结合了 PSA (Partial self-attention) 块,通过在标准 C2F 模块中引入 PSA 块,提高了模型对重要特征的捕捉能力。PSA 块利用前馈神经网络和多头注意力机制,能够选择性地关注输入特征中的重要信息部分,抑制不重要的信息从而帮助模型更好地聚焦于目标物体的特征,提高检测的准确性。

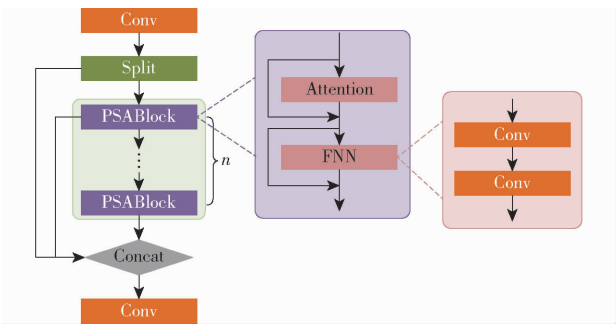


图 6 C2PSA 原理图
Fig. 6 C2PSA schematic

1.4.2 YOLO 11 模型改进

为了解决原始 YOLO 11 网络模型轻量化及在自然环境下残膜识别精度不高的问题,本文提出了改进 YOLO 11 检测模型。首先在主干网络中用了一个深度可分离卷积 (Depthwise convolution, DWConv) 块代替一个 Conv 块,减少模型参数量和运算量,提高模型运行的效率;其次,在输出检测端末尾加入 CBAM (Convolutional block attention module) 卷积块注意力机制模块,通过结合通道注意力和空间注意力 2 个维度的信息,提升模型的特征表达能力以及上下文理解能力;最后,通过融合 ADown (Adaptive downsample) 下采样机制,防止过拟合,减少计算量,增大感受野,提高模型对复杂背景残膜目标的检测精度。

(1) DWConv 模块

DWConv 是一种高效的卷积操作,其卷积核为单通道模式,对输入的每一个通道单独进行卷积,这样就会得到和输入特征图通道数一致的输出特征图。即卷积核个数等于输入特征图通道数,又等于输出特征图个数,如图 7 所示。

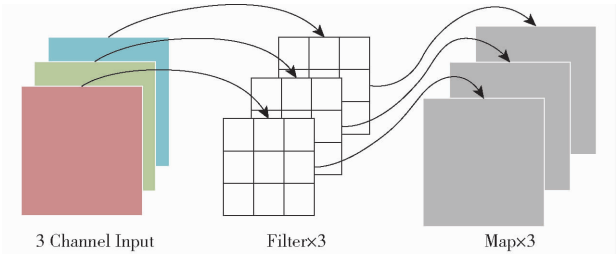


图 7 DWConv 原理图
Fig. 7 DWConv schematic

用数学公式来表达 DWConv,假设卷积核大小为 $D_K \times D_K$,输入通道数为 M ,输出通道数为 N ,输出的特征图尺寸为 $D_F \times D_F$,即:参数量为 $D_K \times M \times N$,计算量为 $D_K \times M \times N \times D_F \times D_F$ 。如表 1 所示,DWConv 参数量只有 2 816。

通过加入 DWConv,YOLO 11 模型能够在保持性能的同时降低模型的计算成本,有利于后期在残膜回收机上使用,主要用于减少计算复杂度和参数

量,特别适用于移动端或轻量化网络。

表 1 改进 YOLO 11 内部参数

Tab.1 Internal parameters of improved YOLO 11

序号	模块	参数量	步长	输入尺寸
0	输入图像	/	/	640 × 640 × 3
1	Conv	464	2	320 × 320 × 16
2	Conv	4 672	2	160 × 160 × 32
3	C3k2	6 640	/	160 × 160 × 64
4	Conv	36 992	2	80 × 80 × 64
5	C3k2	26 080	/	80 × 80 × 128
6	ADown	41 216	2	40 × 40 × 128
7	C3k2	87 040	/	40 × 40 × 128
8	DWConv	2 816	2	20 × 20 × 256
9	C3k2	346 112	/	20 × 20 × 256
10	SPPF	164 608	/	20 × 20 × 256
11	C2PSA	249 728	/	20 × 20 × 256
12	Upsample	0	/	40 × 40 × 128
13	Concat	0	/	40 × 40 × 256
14	C3k2	111 296	/	40 × 40 × 128
15	Upsample	0	/	80 × 80 × 128
16	Concat	0	/	80 × 80 × 192
17	C3k2	32 096	/	80 × 80 × 64
18	Conv	36 992	2	40 × 40 × 64
19	Concat	0	/	40 × 40 × 256
20	C3k2	86 720	/	40 × 40 × 128
21	Conv	147 712	2	20 × 20 × 128
22	Concat	0	/	20 × 20 × 384
23	C3k2	378 880	/	20 × 20 × 256
24	CBAM	8 563	/	20 × 20 × 256

注:“/”表示无此项。

(2)CBAM 卷积块注意力机制模块

CBAM 卷积块注意力机制模块是一种用于增强卷积神经网络性能的注意力机制模块,由 WOO 等^[32]提出。模块的主要目标是通过引入通道注意力和空间注意力来提高模型的感知能力,从而在不增加网络复杂性的情况下改善性能。

模块由 2 个关键部分组成:通道注意力模块(Channel attention module)和空间注意力模块(Spatial attention module)。如图 8 所示,通道注意力模块通过全局最大池化和全局平均池化操作生成 2 个特征向量,然后通过全连接层学习每个通道的注意力权重,这些权重用于增强重要的通道并抑制不重要的通道。空间注意力模块通过聚合特征图的特征,生成一个空间注意力图,强调重要的空间区域并抑制不重要的区域。

在复杂的自然环境下拍摄残膜图像,存在大量的干扰因素,增加了检测的难度,原 YOLO 11 模型对复杂背景下残膜经常出现漏检与错检情况。通过在输出检测端末尾加入 CBAM 卷积块注意力机制模块来捕捉残膜图像边缘特征以及通道与空间之间的联系,减少边缘干扰与背景干扰,来提高检测精度。

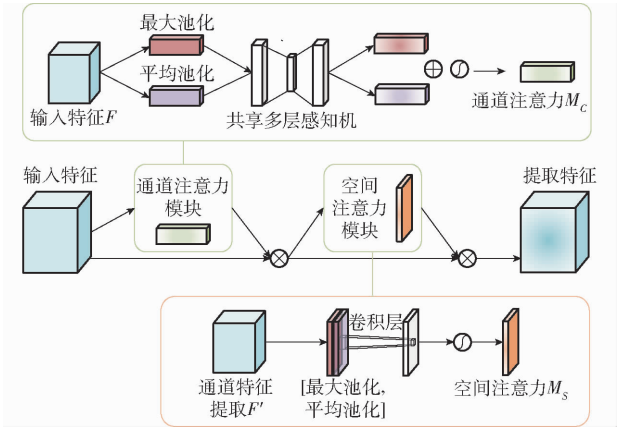


图 8 CBAM 结构
Fig. 8 CBAM structure

(3)ADown 模块

ADown 模块是一种用于目标检测任务下采样操作的卷积块。ADown 模块替换主干网络中的 Conv 模块,用于在残膜特征图的不同层之间进行下采样,减少特征图空间维度,捕捉更高层次上图像的特征,减少计算量,保留关键信息来提高残膜目标检测准确性,改进 YOLO 11 内部参数如表 1 所示。

1.4.3 残膜目标检测识别结果评价指标

为评估改进模型检测效果,采用精确率(Precision)、召回率(Recall)、平均精度均值(mean average precision, mAP)、模型参数量以及处理速度(Frames per second)作为模型评价指标。

2 结果与分析

2.1 试验软件和硬件

本研究模型训练硬件环境为惠普 OM1N 台式计算机,搭载 NVIDIA GeForce RTX4080super 显卡,显存 16 GB;运行环境为 Windows 11,处理器为 Intel(R) Core(TM) i7-14700k,主频为 3 400 MHz,64 位处理器,32 GB 内存;软件环境采用 Pytorch-Cuda 深度学习框架 Pytorch 12.4,编译器为 Pycharm 2022,编程语言为 Python 3.11。

2.2 训练参数设置

考虑实时检测及在移动端上的部署需求,能够更好更轻量化地识别到棉田残膜,采用 YOLO 11 模型为主要框架,采用边界框损失(Box Loss)优化预测残膜边界框的位置和大小,采用分类损失(Cls Loss)用于衡量模型对残膜预测的准确性,设置输入图像尺寸为 640 像素 × 640 像素,以 16 幅图像为一个批处理量,设置总迭代次数为 300 轮,网络权重参数的基础学习率为 0.01,权值衰减率为 0.000 5。

2.3 测试结果与分析

2.3.1 改进 YOLO 11 模型训练过程与结果

在改进 YOLO 11 模型训练过程中记录模型参

数量、Box Loss 损失值、Cls Loss 损失值、精确率、召回率以及 mAP(重叠率 0.5),训练过程中迭代次数与损失值、精确率、召回率、mAP(重叠率 0.5)的关系曲线如图 9 所示。损失值在训练初期下降速度快,当训练到 200 轮时趋于稳定,此时,Box Loss 损失值为 1.365 8,Cls Loss 损失值达到 0.873 5,经过 300 轮训练后,模型完全收敛。最终模型精确率为 81.9%,召回率为 80.9%,mAP 为 86.7%,参数量为 2.20×10^6 。

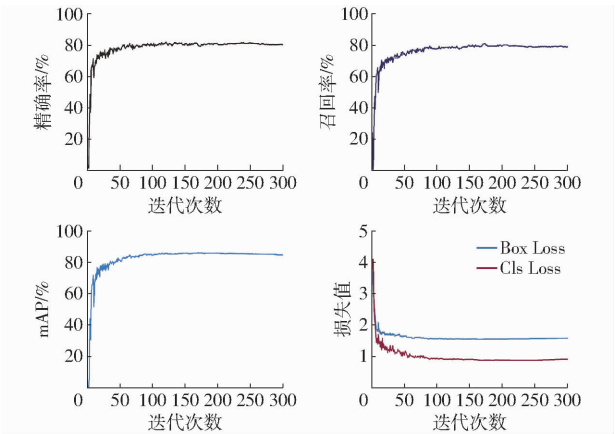


图 9 精确率、召回率、mAP 和损失值变化曲线
Fig.9 Loss function, precision, recall, mAP variation curves of training

2.3.2 消融试验

为了验证改进 YOLO 11 网络模型中各模块有效性,在相同训练环境以及超参数设置下开展消融试验。共设置 8 组消融试验,结果如表 2 所示。

表 2 消融试验结果
Tab.2 Ablation experiment

模型	精确率/ %	召回率/ %	平均精度均 值/%	参数量	处理速度/ (f·s ⁻¹)
YOLO 11	79.5	80.0	86.3	2.59×10^6	87
YOLO 11 + ADown	80.7	79.6	86.7	2.48×10^6	75
YOLO 11 + CBAM	81.1	80.2	85.7	2.60×10^6	68
YOLO 11 + DWConv	80.6	79.9	86.3	2.29×10^6	93
YOLO 11 + DWConv + CBAM	80.7	80.1	86.4	2.31×10^6	81
YOLO 11 + ADown + CBAM	80.0	79.4	86.4	2.20×10^6	65
YOLO 11 + DWConv + ADown	80.9	79.6	86.8	2.19×10^6	85
YOLO 11 + DWConv + CBAM + ADown	81.9	80.9	86.7	2.20×10^6	80

YOLO 11 + DWConv 模型比 YOLO 11 模型在更轻量化的情况下,召回率和平均精度均值几乎接近,但模型精确率高 1.1 个百分点。YOLO 11 + CBAM 模型精确率较高,为 81.1%,比 YOLO 11 模型高

1.6 个百分点,说明 CBAM 卷积块注意力机制模块在精度提升上有显著效果,但其平均精度均值最低。YOLO 11 + ADown 模型有较高的平均精度均值,但其召回率较低。YOLO 11 + DWConv + ADown 模型平均精度均值最高,为 86.8%,但是其召回率较低,为 79.6%,意味着模型有残膜漏检情况,会导致模型在识别残膜目标时的表现不佳,从而影响最终的识别性能。YOLO 11 + CBAM 模型、YOLO 11 + DWConv + CBAM 模型和 YOLO 11 + DWConv + CBAM + ADown 模型相较于 YOLO 11 + DWConv 模型均有着更好的识别效果,但是相较于 YOLO 11 + DWConv + CBAM 模型,YOLO 11 + DWConv + CBAM + ADown 模型召回率提升 0.8 个百分点,精确率提升 1.2 个百分点。在图像处理速度方面,YOLO 11 + DWConv + CBAM + ADown 模型处理速度为 80 f/s,改进模型在轻量化以及精度提升的情况下,处理速度没有大幅下滑。模型参数量也略小于原始的 YOLO 11 模型,基本满足在棉田地表复杂背景的自然环境中精准、高效识别残膜的要求。将改进 YOLO 11 模型命名为 DCA - YOLO 11。

2.3.3 不同检测模型对比试验

为了更好地评估 DCA - YOLO 11 模型目标检测性能,与目前主流模型进行对比,在相同环境条件下进行试验,对比算法包括 YOLO v8、YOLO v9 和 YOLO v10,并使用相同的训练集和测试集,以确保结果的可比性。DCA - YOLO 11 模型与其他模型对残膜检测试验结果如表 3 所示。对比 4 种模型在精确率、召回率、平均精度均值、模型参数量和图像处理速度 5 个方面性能指标,DCA - YOLO 11 多个指标均表现出显著优势。

表 3 训练集、测试集试验结果
Tab.3 Test results on training and testing sets

模型	精确率/%	召回率/%	平均精度均值/%	参数量	处理速度/(f·s ⁻¹)
YOLO v8	78.1	79.1	83.9	3.16×10^6	70
YOLO v9	79.6	79.9	85.1	2.71×10^6	71
YOLO v10	79.0	78.9	85.2	2.71×10^6	85
DCA - YOLO 11	81.9	80.9	86.7	2.20×10^6	80

首先,DCA - YOLO 11 模型精确率比 YOLO v10、YOLO v9、YOLO v8 分别高 2.9、2.3、3.8 个百分点,召回率分别比 YOLO v10、YOLO v9 和 YOLO v8 高 2.0、1.0、1.8 个百分点,平均精度均值比 YOLO v10、YOLO v9、YOLO v8 分别高 1.5、1.6、2.8 个百分点,表明了 DCA - YOLO 11 模型在实际应用中能够更精准地捕捉残膜目标,减少残膜漏检率。DCA - YOLO 11 模型参数量最小,为 2.20×10^6 ;

DCA - YOLO 11 模型处理速度为 80 f/s, 相较于 YOLO v8、YOLO v9 分别提升 14.2%、12.7%, YOLO v10 处理速度为 85 f/s, 略高于 DCA - YOLO 11。在模型轻量化和精度提升的情况下, DCA - YOLO 11 仍保持较高的处理速度。

因此, 综合对比可以看出, DCA - YOLO 11 模型在保证精度的同时, 参数量最少, 模型最小, 证明其轻量化与优越性。这使得 DCA - YOLO 11 模型非常适合棉田复杂背景下残膜实时目标轻量化检测的应用。

2.3.4 不同模型检测残膜的可视化效果对比

为更准确地表现 DCA - YOLO 11 模型的实验结果, 基于表 3 的试验数据进行可视化效果分析。如图 10 所示, 蓝色矩形框为模型预测框, 红色和黄色矩形框为人工标注, 黄色框表示地面残膜被棉秆或土块遮挡, 红色框表示残膜无遮挡。在棉田自然环境下, 对于背景较为简单的残膜, 如场景 1, 所有的模型检测残膜都能获得较好的识别效果。但是, 当面对低光照的复杂背景环境时, 如场景 2, YOLO v9、YOLO v10、DCA - YOLO 11 都检测出残膜目标, 但是有一处人工标注的 2 个残膜矩形框识别成 1 个大的矩形框, 仅有 YOLO v8 识别为 2 个残膜矩形框。当环境存在残膜被棉秆遮挡时, 如场景 3 黄色矩形框, 仅有 YOLO v10 和 DCA - YOLO 11 全部检测出目标框, 但是 YOLO v10 存在目标检测框重叠现象, 其他模型都漏检遮挡残膜。当环境存在残膜被土壤或土块遮挡, 如场景 4 黄色矩形框所示, YOLO v8 没有完全检测出目标框, YOLO v9 检测出黄色矩形框, 但是误检 1 处其他目标, YOLO v10 和 DCA - YOLO 11 模型比其他模型表现出明显的残膜识别优势。在场景 4 中, 所有模型都检测出 1 个人工未标注出的残膜目标, 由于不确定是残膜还是落地棉, 该误差由人工标注不确定性引起。通过不同模型检测残膜的可视化效果对比, DCA - YOLO 11 模型在轻量化情况下, 仍能在复杂的检测环境中保持优越的检测性能, 显著减少了残膜漏检和误检现象。

2.3.5 模型泛化性验证

根据预测模型, 随机选取 2024 年 11 月 21 日在库尔勒市尉利县拍摄的 40 幅分辨率为 4 608 像素 × 3 456 像素的图像, 进行泛化性验证测试。首先对原图进行人工标注, 残膜碎片总数为 701 个, 利用平均绝对误差 (MAE)、均方根误差 (RMSE) 作为评价指标^[33], 对比图中残膜碎片个数真实值和预测值。MAE 越小, 表示残膜预测值与真实值误差越小; RMSE 越小, 表示残膜预测值与真实值分散程度越

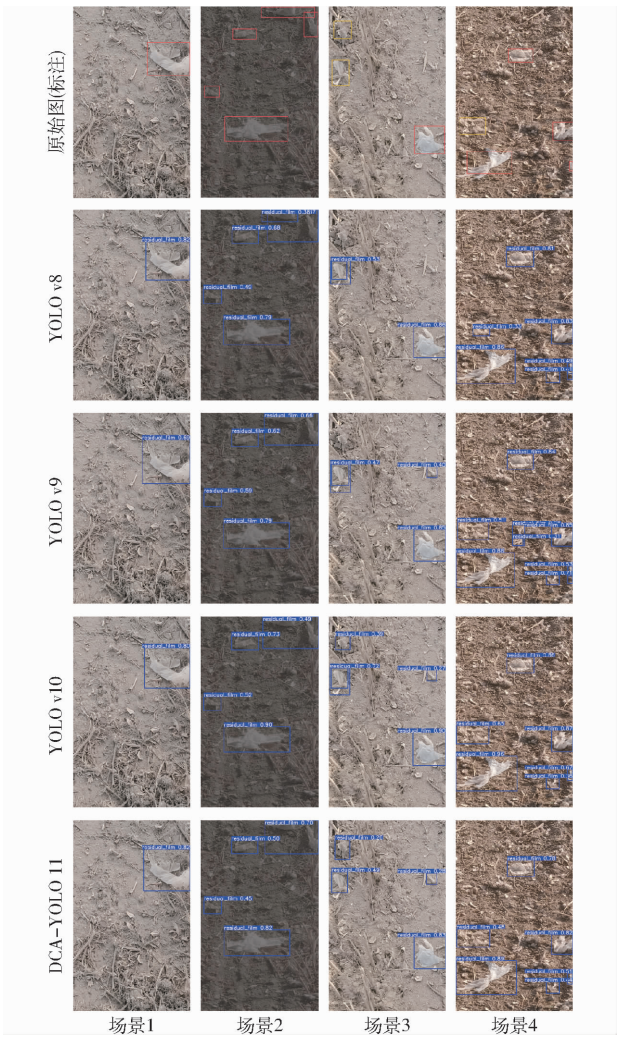


图 10 不同模型识别效果

Fig. 10 Recognition effect diagram of different models

小; R^2 表示趋势线拟合程度, 其值越靠近 1, 拟合程度越高。

提出的模型能够精准检测到残膜碎片, 其在验证数据集上检测结果, R^2 为 0.72, 平均绝对误差和均方根误差分别为 4.92 个和 2.72 个。如图 11 所示, 线性拟合线可以有效反映残膜预测值与真实值关系, 在棉田地表复杂的自然环境下, 拟合程度较好。总体而言, 提出的轻量化模型泛化性较好, 可以准确地对残膜碎片数进行检测并计数。

3 结论

(1) 提出了一种改进 YOLO 11 模型的棉田地表残膜识别方法, 通过在主干网络中将普通卷积模块替换为 DWConv 模块, 能够在保持性能同时降低模型大小, 减少计算复杂度和参数量; 通过在输出检测端末尾加入 CBAM 卷积块注意力机制模块来增强对残膜图像边缘特征的提取能力以及通道与空间之间的联系, 减少边缘干扰与背景干扰, 来提高检测精度; 通过融合 ADown 下采样机制, 实现残膜特征图

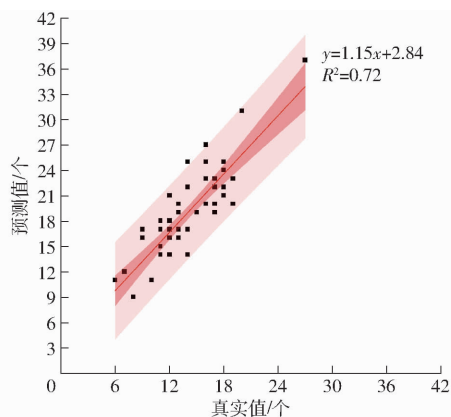


图 11 残膜碎片数真实值和预测值拟合结果

Fig. 11 Results of true and predicted values of residual film

不同层之间的下采样,减少特征图空间维度,捕捉更高层次上图像的特征,减少计算量,保留关键信息来提高残膜目标检测准确性;最终模型目标识别精确率为 81.9%,召回率为 80.9%,平均精度均值(重

叠率 0.5)为 86.7%,参数量为 2.20×10^6 ,图像处理速度为 80 f/s。

(2)改进 YOLO 11 模型与 YOLO v8、YOLO v9、YOLO v10 模型进行对比试验,检测的精确率比 YOLO v10、YOLO v9、YOLO v8 分别高 2.9、2.3、3.8 个百分点,召回率分别比 YOLO v10、YOLO v9 和 YOLO v8 高 2.0、1.0、1.8 个百分点,平均精度均值比 YOLO v10、YOLO v9、YOLO v8 分别高 1.5、1.6、2.8 个百分点,处理速度比 YOLO v9、YOLO v8 分别提升 12.7%、14.2%。DCA-YOLO 11 模型在保证精度的同时,参数量最少,模型最小,非常适合棉田地复杂背景下轻量化残膜实时目标检测的应用。

(3)最后经过模型泛化性验证测试,其在验证数据集上的检测结果, R^2 为 0.72,平均绝对误差和均方根误差分别为 4.92 个和 2.72 个,本文提出的 DCA-YOLO 11 轻量化模型泛化性较好。

参 考 文 献

- [1] 梁荣庆,陈学庚,张炳成,等. 新疆棉田残膜回收方式及资源化再利用现状问题与对策[J]. 农业工程学报,2019,35(16): 1-13.
LIANG Rongqing, CHEN Xuegeng, ZHANG Bingcheng, et al. Problems and countermeasures of recycling methods and resource reuse of residual film in cotton fields of Xinjiang[J]. Transactions of the CSAE, 2019, 35(16): 1-13. (in Chinese)
- [2] 赵岩,陈学庚,温浩军,等. 农田残膜污染治理技术研究现状与展望[J]. 农业机械学报,2017,48(6): 1-14.
ZHAO Yan, CHEN Xuegeng, WEN Haojun, et al. Research status and prospect of control technology for residual plastic film pollution in farmland[J]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(6): 1-14. (in Chinese)
- [3] 蒋永新,曲衡,刘旋峰,等. 残膜回收机铲齿式起膜装置设计与试验[J]. 农业机械学报,2025,56(1): 282-289.
JIANG Yongxin, QU Heng, LIU Xuanfeng, et al. Design and experiment of spade tooth film lifting device for residual film recycling machine[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(1): 282-289. (in Chinese)
- [4] WANG P, CHEN X, WEN H. Research and experiment on the removal mechanism of light impurities of the residual mulch film recovery machine[J]. Agriculture, 2022, 12(6): 775.
- [5] 寇嘉乐,王吉奎,王磊,等. 夹指链式棉田残膜回收秸秆粉碎还田一体机设计与试验[J]. 农业工程学报,2023,39(23): 14-22.
KOU Jiale, WANG Jikui, WANG Lei, et al. Integrating clamping finger-chain residual film recycling and straw crushing return to cotton field[J]. Transactions of the CSAE, 2023, 39(23): 14-22. (in Chinese)
- [6] LIU X, ZHAO W, ZHANG H, et al. Design and experiment of an upper-side-discharge straw-returning and bundle self-unloading integrated corn residual film recycling machine[J]. International Journal of Agricultural & Biological Engineering, 2023, 16(5): 79-95.
- [7] CAO S, XIE J, WANG H, et al. Design and operating parameters optimization of the hook-and-tooth chain rail type residual film picking device[J]. Agriculture, 2022, 12(10): 1717.
- [8] 江水泉,张海东,华英杰. 基于计算机视觉的农田残膜定位研究[J]. 中国农机化学报,2016,37(11): 150-154.
JIANG Shuiquan, ZHANG Haidong, HUA Yingjie. Research on location of residual plastic film based on computervision[J]. Journal of Chinese Agricultural Mechanization, 2016, 37(11): 150-154. (in Chinese)
- [9] 吴雪梅,梁长江,张大斌,等. 基于无人机遥感影像的收获期后残膜识别方法[J]. 农业机械学报,2020,51(8): 189-195.
WU Xuemei, LIANG Changjiang, ZHANG Dabin, et al. Identification method of plastic film residue based on UAV remote sensing images[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(8): 189-195. (in Chinese)
- [10] 张学军,黄爽,靳伟,等. 基于改进 Faster R-CNN 的农田残膜识别方法[J]. 湖南大学学报(自然科学版),2021,48(8): 161-168.
ZHANG Xuejun, HUANG Shuang, JIN Wei, et al. Identification method of agricultural film residue based on improved Faster R-CNN[J]. Journal of Hunan University(Natural Sciences), 2021, 48(8): 161-168. (in Chinese)
- [11] 翟志强,陈学庚,邱发松,等. 基于像素块和机器学习的播前棉田地残膜覆盖率检测[J]. 农业工程学报,2022,38(6): 140-147.
ZHAI Zhiqiang, CHEN Xuegeng, QIU Fasong, et al. Detecting surface residual film coverage rate in pre-sowing cotton fields using pixel block and machine learning[J]. Transactions of the CSAE, 2022, 38(6): 140-147. (in Chinese)
- [12] 牛玉珩. 机载残膜回收智能监测设备研究设计[D]. 乌鲁木齐:新疆农业大学,2023.

NIU Yuheng. Research and design of intelligent monitoring equipment recycling monitoring equipment [D]. Urumqi: Xinjiang Agricultural University, 2023. (in Chinese)

[13] 刘云玲, 张天雨, 姜明, 等. 基于机器视觉的葡萄品质无损检测方法研究进展 [J]. 农业机械学报, 2022, 53 (增刊 1) : 299 – 308.

LIU Yunling, ZHANG Tianyu, JIANG Ming, et al. Review on non-destructive detection methods of grape quality based on machine vision [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53 (Supp. 1) : 299 – 308. (in Chinese)

[14] 丁永军, 杨文涛, 赵一龙. 基于一致性半监督学习的苹果叶片病斑分割模型研究 [J]. 农业机械学报, 2024, 55 (12) : 314 – 321.

DING Yongjun, YANG Wentao, ZHAO Yilong. Apple leaf spot segmentation model based on consistency semi-supervised learning [J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55 (12) : 314 – 321. (in Chinese)

[15] 白恩龙, 张周强, 郭忠超, 等. 基于机器视觉的棉花颜色检测方法 [J]. 纺织学报, 2024, 45 (3) : 36 – 43.

BAI Enlong, ZHANG Zhouqiang, GUO Zhongchao, et al. Cotton color detection on method based on machine vision [J]. Journal of Textile Research, 2024, 45 (3) : 36 – 43. (in Chinese)

[16] 夏彬, 史书伟, 张若宇, 等. 棉花轧工质量机器视觉检测系统设计与试验 [J]. 农业机械学报, 2023, 54 (11) : 189 – 197.

XIA Bin, SHI Shuwei, ZHANG Ruoyu, et al. Design and test of machine vision inspection system for cotton preparation [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54 (11) : 189 – 197. (in Chinese)

[17] 李国厚, 李泽旭, 金松林, 等. 混合卷积神经网络用于高光谱小麦品种鉴别 [J]. 光谱学与光谱分析, 2024, 44 (3) : 807 – 813.

LI Guohou, LI Zexu, JIN Songlin, et al. Mix convolutional neural networks for hyperspectral wheat variety discrimination [J]. Spectroscopy and Spectral Analysis, 2024, 44 (3) : 807 – 813. (in Chinese)

[18] 樊湘鹏, 许燕, 周建平, 等. 基于改进卷积神经网络的在体青皮核桃检测方法 [J]. 农业机械学报, 2021, 52 (9) : 149 – 155, 114.

FAN Xiangpeng, XU Yan, ZHOU Jianping, et al. Green walnut detection method based on improved convolutional neural network [J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52 (9) : 149 – 155, 114. (in Chinese)

[19] 张凡, 王文秀, 王春山, 等. 高光谱图像结合卷积神经网络的马铃薯干腐病潜育期识别 [J]. 光谱学与光谱分析, 2024, 44 (2) : 480 – 489.

ZHANG Fan, WANG Wenxiu, WANG Chunshan, et al. Study on hyperspectral detection of potato dry rot in gley stage based on convolutional neural network [J]. Spectroscopy and Spectral Analysis, 2024, 44 (2) : 480 – 489. (in Chinese)

[20] 李齐明, 唐世彬, 翁芳文, 等. 波形识别与过程解释的微小地震监测基于改进 LeNet5 卷积神经网络 [J]. 中南大学学报, 2023, 30 (3) : 904 – 918.

LI Qiming, TANG Shibin, WENG Fangwen, et al. Waveform recognition and process interpretation of microseismic monitoring based on an improved LeNet5 convolutional neural network [J]. Journal of Central South University, 2023, 30 (3) : 904 – 918.

[21] 李丛悦, 胡义华, 江家伟, 等. 基于 Gramian 角场和卷积神经网络的船用柴油发动机故障诊断 [J]. 浙江大学学报-理学版, 2024, 25 (6) : 470 – 482.

LI Congyue, HU Yihua, JIANG Jiawei, et al. Fault diagnosis of a marine power-generation diesel engine based on the Gramian angular field and a convolutional neural network [J]. Journal of Zhejiang University-Science A (Applied Physics & Engineering), 2024, 25 (6) : 470 – 482.

[22] 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. Lightweight method for identifying farmland weeds based on YOLO v5 [J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55 (1) : 212 – 222, 293. (in Chinese)

[23] 唐延龄, 韩巧玲, 赵玥, 等. 基于多尺度序列特征融合的林分害虫检测算法 [J]. 林业科学, 2025, 61 (2) : 21 – 31.

TANG Yanling, HAN Qiaoling, ZHAO Yue, et al. A forest pest detection algorithm based on multi-scale sequence feature fusion [J]. Scientia Silvae Sinicae, 2025, 61 (2) : 21 – 31. (in Chinese)

[24] 张宏建, 孙智霖, 亓新春, 等. 基于改进 YOLO v8 的苹果树树干精准识别方法 [J]. 农业机械学报, 2024, 55 (增刊 1) : 246 – 255, 262.

ZHANG Hongjian, SUN Zhilin, QI Xinchun, et al. Accurate apple tree trunk recognition method based on improved YOLO v8 [J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55 (Supp. 1) : 246 – 255, 262. (in Chinese)

[25] 熊俊涛, 韩咏林, 王潇, 等. 基于 YOLO v5-Lite 的自然环境木瓜成熟度检测方法 [J]. 农业机械学报, 2023, 54 (6) : 243 – 252.

XIONG Juntao, HAN Yonglin, WANG Xiao, et al. Method of maturity detection for papaya fruits in natural environment based on convolutional neural networks YOLO v5-Lite [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54 (6) : 243 – 252. (in Chinese)

[26] 郑文轩, 杨瑛. 基于频域数据增强与轻量化 YOLO v7 模型的成熟期香梨目标检测方法 [J]. 农业机械学报, 2024, 55 (5) : 244 – 253.

ZHENG Wenxuan, YANG Ying. Mature stage pear detection method based on frequency domain data augmentation and lightweight YOLO v7 model [J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55 (5) : 244 – 253. (in Chinese)

[27] 林家豪, 张元泽, 梁千月, 等. 基于视触觉与深度学习的猕猴桃无损硬度检测方法 [J]. 农业机械学报, 2023, 54 (10) : 390 – 398.

LIN Jiahao, ZHANG Yuanze, LIANG Qian Yue, et al. Non-destructive firmness testing of kiwifruit based on vision-based tactile sensor and fusion learning [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54 (10) : 390 – 398. (in Chinese)

CAE, 2021, 23(4): 1–9. (in Chinese)

[18] 农业食品标准化研究所. ISO 智慧农业标准化战略咨询组中国专家工作组获批成立[EB/OL]. (2022–08–31)[2025–03–23]. https://www.cnis.ac.cn/bydt/zhxw/202208/t20220831_53739.html.

[19] 姚艳敏,白玉琪. 农业大数据标准体系框架研究[J]. 农业大数据学报,2019,1(4):76–85.
YAO Yanmin, BAI Yuqi. A framework for agricultural big data standards[J]. Journal of Agricultural Big Data, 2019, 1(4): 76–85. (in Chinese)

[20] 中国标准化研究院. 标准体系构建原则和要求:GB/T 13016—2018[S]. 北京:中国标准出版社,2018.

[21] 中国标准化研究院. 综合标准化工作指南:GB/T 12366—2009[S]. 北京:中国标准出版社,2009.

[22] 刘志强. 2024 年全国棉花产量同比增长 9.7% [N]. 人民日报,2024–12–27(02).

[23] 陈澈,孟庆刚. 基于切克兰德方法论的中国健康养老产业分析[J]. 中国老年学杂志,2018,38(21):5371–5375.
CHEN Che, MENG Qinggang. Analysis of china's healthy aging industry based on clandestine methodology [J]. Chinese Journal of Gerontology, 2018, 38(21): 5371–5375. (in Chinese)

[24] 杨建梅. 切克兰德软系统方法论[J]. 系统辩证学学报,1994(3):86–91.
YANG Jianmei. On the soft systems methodology of checkland[J]. Chinese Journal of Systems Science, 1994(3): 86–91. (in Chinese)

[25] 刘光盛,王红梅,胡月明,等. 中国土地利用工程标准体系框架构建[J]. 农业工程学报,2015,31(13):257–264.
LIU Guangsheng, WANG Hongmei, HU Yueming, et al. Infrastructure of standard system for land use engineering in China [J]. Transactions of the CSAE, 2015, 31(13): 257–264. (in Chinese)

[26] 王同林,邵志勇,聂智星,等. 基于德尔菲法及层次分析法构建鲜食番茄品质评价体系[J/OL]. 浙江农业学报,1–14.
<http://kns.cnki.net/kcms/detail/33.1151.S.20250212.1410.002.html>.
WANG Tonglin, SHAO Zhiyong, NIE Zhixing, et al. Construction of a quality evaluation system for fresh-eating tomatoes based on Delphi method and analytic hierarchy process [J/OL]. Acta Agriculturae Zhejiangensis, 1–14. <http://kns.cnki.net/kcms/detail/33.1151.S.20250212.1410.002.html>. (in Chinese)

[27] 董君. 层次分析法权重计算方法分析及其应用研究[J]. 科技资讯,2015,13(29):218,220.
DONG Jun. Analysis of hierarchical analysis weighting method and its application research [J]. Science & Technology Information, 2015,13(29): 218,220. (in Chinese)

[28] 李锦清. 基于模糊层次分析法的知识产权质押融资风险评价研究[D]. 成都:电子科技大学,2021.
LI Jinqing. Research on risk evaluation of intellectual property pledge financing based on fuzzy analytic hierarchy process [D]. Chengdu: University of Electronic Science and Technology of China, 2021. (in Chinese)

(上接第 25 页)

[28] YANG J,ZHAI Z,LI Y,et al. Design and research of residual film pollution monitoring system based on UAV[J]. Computers and Electronics in Agriculture,2024,217:108608.

[29] 岳继博,冷梦蝶,田庆久,等. 叶片多理化参数的高光谱遥感与深度学习估算[J]. 光谱学与光谱分析,2024,44(10): 2873–2883.
YUE Jibo, LENG Mengdie, TIAN Qingjiu, et al. Estimation of leaf physical and chemical parameters based on hyperspectral remote sensing and deep learning technologies [J]. Spectroscopy and Spectral Analysis, 2024, 44(10): 2873–2883. (in Chinese)

[30] 蒋凯伟,刘彪,刘国豪,等. 基于卷积神经网络的车道线逆透视变换算法[J]. 北京交通大学学报,2023,47(2):106–113.
JIANG Kaiwei, LIU Biao, LIU Guohao, et al. An inverse perspective transformation algorithm for lane lines based on CNN[J]. Journal of Beijing Jiaotong University, 2023, 47(2): 106–113. (in Chinese)

[31] JJOCHER G, NISHIMURA K, MINEEVA T, et al. YOLO11[EB/OL]. 2024, <https://github.com/ultralytics/ultralytics>.

[32] WOO S, PARK J, LEE J Y, et al. CBAM:convolutional block attention module[C]// Computer Vision-ECCV 2018. Cham: Springer International Publishing, 2018:3–19.

[33] 鲍文霞,苏彪彪,胡根生,等. 基于 FE–P2Pnet 的无人机小麦图像麦穗计数方法[J]. 农业机械学报,2024,55(4):155–164,289.
BAO Wenxia, SU Biaobiao, HU Gensheng, et al. Method for counting wheat ears in UAV images based on FE–P2Pnet[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(4): 155–164, 289 (in Chinese)