

基于改进 YOLOv7 的麦穗检测算法

陈 森, 徐伟峰, 王洪涛, 雷 耀

(华北电力大学(保定) 计算机系, 河北 保定 071003;

河北省能源电力知识计算重点实验室, 河北 保定 071003)

摘要: 针对麦穗数据集中存在的检测目标密集、遮挡、各地区形态不一致现象引起的漏检、模型泛化能力弱等问题, 提出一种基于改进 YOLOv7 的麦穗检测算法。首先, 在 YOLOv7 网络的骨干特征提取网络引入混合注意力机制加强对位置特征的提取, 缓解检测目标密集导致的漏检问题; 其次, 在骨干特征提取网络引入能结合不同尺寸的可切换空洞卷积 (switchable atrous convolution, SAC), 通过增大感受野实现提取不同尺度的特征信息, 可有效改善因遮挡现象引起的漏检问题; 最后, 在特征融合部分引入增量学习模块 (example vector correction, EVC), 提高模型的鲁棒性和泛化能力。实验结果表明, 改进后的麦穗识别算法在全球小麦麦穗数据集的平均目标检测精度与原 YOLOv7 相比提高了 2.11 个百分点。

关键词: 小麦麦穗检测; 混合注意力; 增量学习; 空洞卷积

中图分类号: TP391 **文献标志码:** A **文章编号:** 1671-5489(2024)04-0886-09

Wheat Ear Detection Algorithm Based on Improved YOLOv7

CHEN Sen, XU Weifeng, WANG Hongtao, LEI Yao

(Department of Computer, North China Electric Power University (Baoding), Baoding 071003, Hebei Province, China;
Key Laboratory of Energy and Power Knowledge Computing of Hebei Province, Baoding 071003, Hebei Province, China)

Abstract: Aiming at the problems of dense detection targets, occlusion, missed detection caused by inconsistent morphology in various regions and weak generalization ability of the model in the wheat ear dataset, we proposed a wheat ear detection algorithm based on improved YOLOv7. Firstly, we introduced a mixed attention mechanism into the backbone feature extraction network of YOLOv7 network to strengthen the extraction of location features and alleviate the missed detection problem caused by dense detection targets. Secondly, switchable atrous convolution (SAC) which could combine different sizes was introduced into the backbone feature extraction network, and the feature information of different scales was extracted by increasing the receptive field, which could effectively improve the missed detection problem caused by occlusion. Finally, an incremental learning module example vector correction (EVC) was introduced into the feature fusion part to improve the robustness and generalization ability of the model. The experimental results show that the average target detection accuracy of the improved wheat ear recognition algorithm in the global wheat ear dataset is 2.11 percentage points higher than that of the original YOLOv7.

Keywords: wheat ear detection; mixed attention; incremental learning; atrous convolution

收稿日期: 2023-07-12.

第一作者简介: 陈 森(1998—), 男, 汉族, 硕士研究生, 从事计算机视觉目标检测的研究, E-mail: chensen@ncepu.edu.cn.

通信作者简介: 徐伟峰(1982—), 男, 汉族, 博士, 讲师, 从事计算机视觉和空管系统的研究, E-mail: weifengxu@163.com.

基金项目: 国家自然科学基金(批准号: 61802124)和中央高校基本科研业务费专项基金(批准号: 2021MS089).

小麦是世界三大粮食作物之一, 其播种面积在谷类作物中位居首位. 使用视觉传感器辅助小麦产量预估已成为该领域目前研究的重要方向之一, 要完成准确预估, 势必涉及最基础的麦穗检测问题^[1]. 本文提出一种结合深度学习^[2]技术基于目标检测的麦穗检测方法^[3], 利用计算机视觉技术完成麦穗的检测任务.

在目标检测领域, 基于水平框的目标检测算法取得了许多研究成果. 这些算法主要分为两类: 两阶段算法和单阶段算法. 两阶段算法, 如 Faster R-CNN^[4]和 SSD^[5], 具有较高的精度但检测速度较慢; 单阶段算法, 如 YOLO^[6]及无锚框的 YOLOX^[7]等, 单阶段算法具有更快的速度但精度较低. 针对麦穗检测问题, Wang 等^[8]提出了结合角点检测和全卷积层(fully convolutional networks, FCN)^[9]的方法; 张领先等^[10]采用卷积神经网络结合非极大值抑制(non-maximum suppression, NMS)^[11]的方法提升识别效果, 建立了较高准确率的麦穗识别模型. 但这两种方法都只专注于统计麦穗的数量, 并且应用受限于特定地区的小麦检测, 模型泛化性较差. 在单阶段算法中, 陈佳玮等^[12]基于 YOLOv4 算法^[13]构建了轻量级深度学习模型 MobileNetV2-YOLOv4, 检测速度比传统算法更快但检测精度略低. 此外, 张连屹^[14]提出的 YOLOv3-P 模型在 YOLOv3^[15]网络中引入了特征加强模块^[16]和融合注意力^[17-18], 缓解了麦穗遮挡现象引起的漏检问题, 也在全球小麦检测数据集(global wheat head detection, GWHD)上取得了较好的效果, 但模型的检测精度和泛化能力还有待提高. 黄硕等^[19]提出将 CBAM(convolutional block attention module)注意力和 YOLOv5 相结合的网络结构, 但其仅结合注意力机制^[20]导致模型还有进一步改进的空间. 目前, 最新的单阶段算法仍以 YOLO 系列为代表, 尤其是 YOLOv7^[21], 在速度和精度方面性能优异. 为解决麦穗检测任务的特殊性和传统算法的局限性, 本文将基于 YOLOv7 算法框架提出一种新的麦穗图像检测算法.

本文算法主要集中在以下三方面工作: 1) 在 YOLOv7 的骨干特征提取网络的卷积块中引入混合注意力机制 NAM(normalization-based attention module)^[22], 形成新的卷积块 CNAM(convolutional normalization-based attention module), 通过该机制能加强对麦穗目标的特征提取能力, 减轻由于检测目标密集导致的漏检; 2) 在骨干特征提取网络的 MP(MaxPool)模块中, 拟采用能结合不同尺寸的可切换空洞卷积(switchable atrous convolution, SAC)^[23], 使模型能适应不同尺度的麦穗目标, 改善由于遮挡现象导致的漏检; 3) 在特征融合部分的 FPN(feature pyramid network)^[24]处引入增量式学习模块 EVC(example vector correction)^[25], 以提高模型的泛化能力.

1 模型设计

1.1 YOLOv7 模型

YOLO 系列算法是单阶段目标检测算法中最典型的代表之一, 而 YOLOv7^[21]则是目前该系列中最先进的算法. 相比于传统版本, YOLOv7 在准确率和速度上均有显著提升, 其采用了模型重参数化、新的标签分配策略、ELAN 高效网络结构、带辅助头的训练等策略, 因而在检测精度和速度上效果很好, 其网络架构^[26]如图 1 所示.

1.2 融合混合注意力机制

在麦穗图像数据集中, 大部分图片中的麦穗区域只占原图的一小部分, 麦穗在有限范围内分布集中. 但大量背景区域并未提供模型所需的相关信息, 增加了麦穗检测任务的难度. 因此, 本文引入了 NAM 注意力机制^[22]改进麦穗检测网络模型, 抑制其不太显著但颜色与麦穗图片相似的叶片特征, 使其更专注于与麦穗相关的特征, 减少背景特征对模型的影响, 缓解数据集中目标密集导致的漏检问题. NAM 注意力是一种基于归一化的注意力模块, 采用与 CBAM^[27]注意力相同的模块集成方式, 在不增加额外计算量的前提下, 重新设计其通道注意力和空间注意力子模块, 经过实验对比, 将 NAM 添加到 YOLOv7 的 Backbone 中效果最好. NAM 模块结构如图 2 所示.

在通道注意力模块中, 使用 Batch Normalization(BN)中的比例因子度量通道的重要性. 这些比例因子是根据每个 batch 的均值 μ 和标准差 σ 计算的, 因此能反映每个通道在当前 batch 中的重要性程度, 用计算公式可表示为

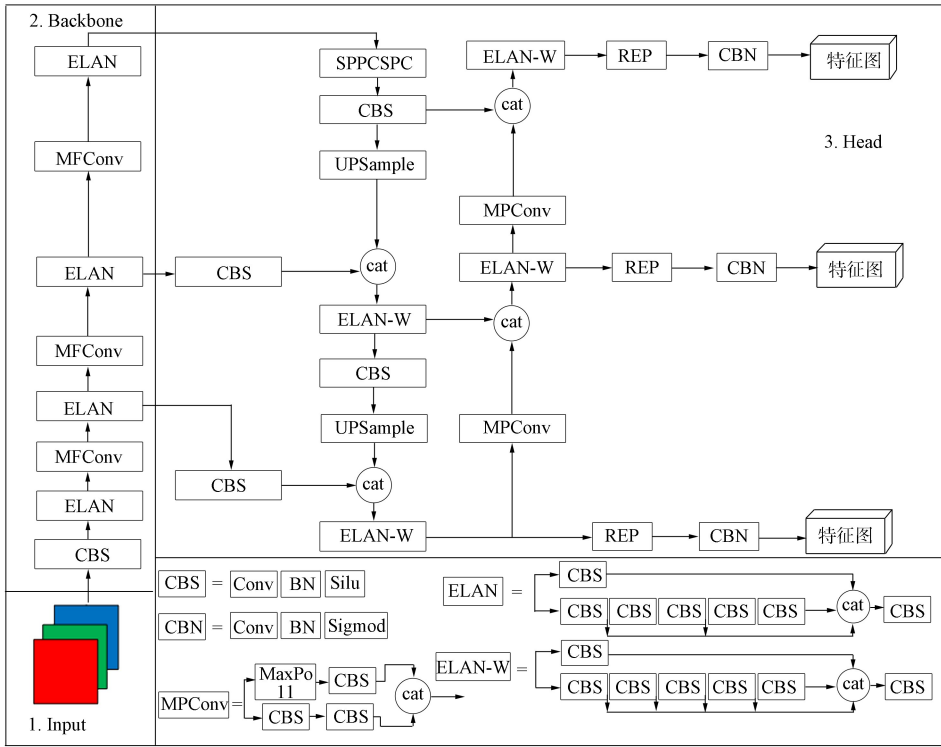


图 1 YOLOv7 网络架构

Fig. 1 Architecture of YOLOv7 network

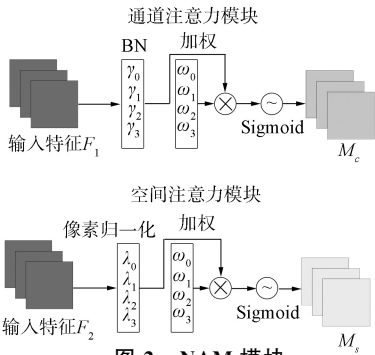


图 2 NAM 模块

Fig. 2 NAM module

$$B_{out} = BN(B_{in}) = \gamma \frac{B_{in} - \mu\beta}{\sqrt{\sigma\beta^2 + \epsilon}} + \beta, \quad (1)$$

其中 ϵ 为无限接近 0 的数值, 目的是避免分母为零. 本文使用这些比例因子加权特征图, 以增强对重要特征的响应. 先将每个通道的特征图乘以可训练的仿射变换参数 γ , 再将它们与可训练位移参数 β 相加得到加权特征图. 通道注意力模块公式为 $M_c = \text{Sigmoid}(W_\gamma(\text{BN}(F_1)))$, 其中 γ 表示每个通道的缩放因子, W_γ 表示权重, F_1 表示输入的特征.

针对空间注意力模块, 将 BN 的比例因子应用于空间维度, 用来衡量像素的重要性, 本文将其命名为像素归一化, 用公式可表示为

$$M_s = \text{Sigmoid}(W_\lambda(\text{BN}_s(F_2))), \quad (3)$$

1.3 改进 MPCConv 模块

针对麦穗数据集中存在麦穗会被其他待检测麦穗遮挡的现象, 为提高麦穗检测模型的抗遮挡能力, 本文在 YOLOv7 骨干特征提取网络的 MPCConv 模块中的 CBS 模块 (Conv 卷积、Batch normalization 归一化、Silu 激活函数) 引入 SAC^[23], 替代原来的普通卷积, SAC 可以自适应地选择不同的膨胀率, 以适应不同大小的检测目标, 其结构如图 3 所示. 由图 3 可见, SAC 包含 3 个核心组件: 两个全局上下文模块和一个 SAC 组件. 这两个上下文模块分别添加在 SAC 组件的前后. 在 SAC 组件中, 用 $y = \text{Conv}(x, w, r)$ 表示采用权重为 w 、空洞率为 r 、输入为 x 的卷积操作, 其中 y 为输出. 从普通卷积到空洞卷积的转换公式为

$$\text{Conv}(x, w, 1) \rightarrow S(x)\text{Conv}(x, w, 1) + (1 - S(x))\text{Conv}(x, w + \Delta w, r), \quad (4)$$

其中: r 为一个超参数, 通常情况下, $r = 3$; Δw 为可通过训练得到的权重; 变换函数 $S(\cdot)$ 由一个 5×5 的平均池化层和一个 1×1 卷积层组成.

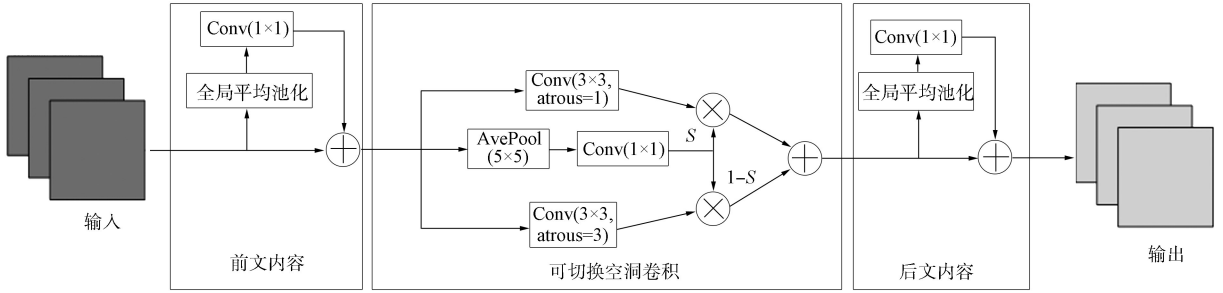


图 3 SAC 模块

Fig. 3 SAC module

YOLOv7 骨干网络中 MPCConv 模块的作用是下采样, 如图 4(A) 所示, 该模块有两个分支: 第一个分支是进行最大池化操作, 将特征图下采样, 然后使用 1×1 卷积进行通道数的变换; 第二个分支是使用 1×1 卷积进行通道数的变换, 然后使用 3×3 卷积核、步长为 2 的卷积块进行下采样. 最终将第一个分支和第二个分支的输出进行拼接, 得到下采样后的结果. 如图 4(B) 所示, 本文将第二个分支的第二个模块中的普通卷积改为 SAC, 使网络更能适应不同尺寸的检测目标, 以提高麦穗检测模型的抗遮挡能力, 改善检测模型的漏检问题.



图 4 改进的 MPCConv 模块

Fig. 4 Improved MPCConv module

1.4 改进特征融合模块

针对麦穗数据集中图片来源广、不同生长时期的图片差异大等因素导致的模型泛化能力弱等问题, 本文在 YOLOv7 的特征融合部分引入 CFPNet 网络的 EVC 模块^[25], 该模块是 CFPNet 网络引入的一种增量式学习模块. EVC 用于解决目标检测中的遗忘问题, 即在原模型基础上, 如何在重新训练整个模型的情况下快速适应新的样本, 其结构如图 5 所示.

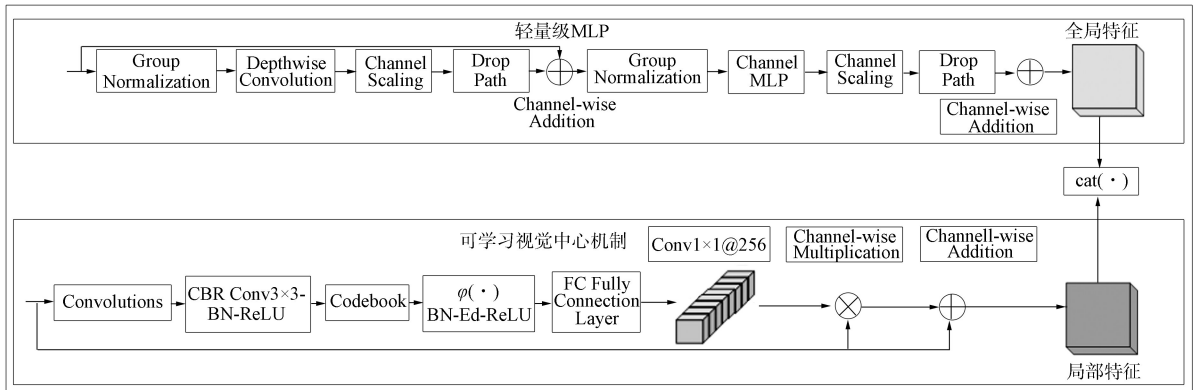


图 5 EVC 模块

Fig. 5 EVC module

EVC 模块由两个并行连接的块构成, 其中上方分支使用轻量级多层感知机 (MLP) 捕获顶级特征的全局信息, 同时, 为保留局部信息, 下方分支使用可学习的视觉中心机制 (LVC) 聚合层内局部区域特征. 将这两个模块的结果特征图沿通道维度连接在一起, 作为 EVC 的输出, 用于下游识别, 用公式可表示为

$$X = \text{cat}(\text{MLP}(X_{in}); \text{LVC}(X_{in})). \tag{5}$$

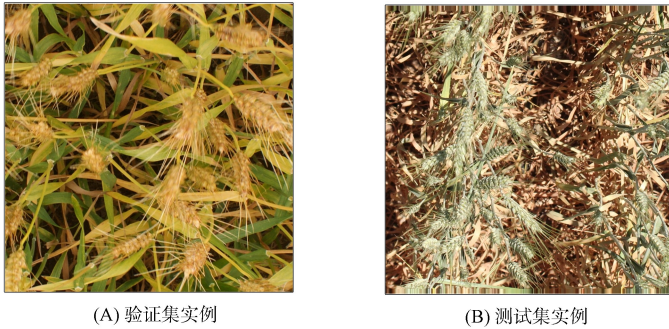
2.1 数据集及处理

由于全球小麦公开数据集 GWHD 为 .csv 格式, 因此本文先将其转换为 YOLO 格式, 经筛选后数据集共有 6 515 张图片, 尺寸大小为 $1\ 024 \times 1\ 024$, 其中训练集为 3 657 张图片, 验证集和测试集分别为 1 476 张和 1 382 张图片. 训练集、验证集和测试集实例分别如图 7 和图 8 所示.



图 7 训练集示例

Fig. 7 Example of training set



(A) 验证集实例

(B) 测试集实例

图 8 验证集(A)和测试集(B)实例

Fig. 8 Examples of validation set (A) and test set (B)

2.2 实验参数设置

本文实验中, 由于设备性能所限, 将输入图像的大小缩放至 $1\ 024 \times 1\ 024$ 像素, 并使用 SGD 作为优化器, 设学习率为 1×10^{-2} , 动量为 0.9, 权重衰减为 5×10^{-4} . 为调整学习率, 使用余弦退火算法. 在训练过程中, 每个 batch 中包含 6 张图片, 在添加 EVC 模块后, 将 batch 尺寸调整为 2. 对模型进行 300 个 epoch 的训练, 并交替进行 10 个训练 epoch 和 1 个测试 epoch.

2.3 实验评价指标

在小麦麦穗检测实验中, 本文使用 $mAP@0.5$ 和 $mAP@(0.5 \sim 0.95)$ 的平均值作为评价指标. 其中, $mAP@0.5$ 表示在交并比(IoU)阈值大于 0.5 的情况下的平均准确率均值. 准确率(Precision)、召回率(Recall)、平均准确率 AP 和平均准确率均值 mAP 的计算公式分别为

$$\text{Precision} = \frac{TP}{TP + FP}, \quad (8)$$

$$\text{Recall} = \frac{TP}{TP + FN}, \quad (9)$$

$$\text{AP} = \int_0^1 \rho(r) dr, \quad (10)$$

$$\text{mAP} = \frac{1}{c} \sum_{i=1}^c \text{AP}_i, \quad (11)$$

其中: TP 表示正确的预测; FP 表示错误的预测; FN 表示预测中漏掉的检测目标; Precision 表示预测为 TP 的比例; Recall 表示被正确预测的目标数与数据集中的所有目标数的比例; 每个类别的平均准确率 AP 通过在准确率-召回率曲线下计算曲线下面积获得; 平均准确率 mAP 是所有类别的 AP 的平

均值, 通常使用 AP@0.5 作为衡量指标.

2.4 YOLOv7 网络模型与改进后网络模型实验结果对比

本文以递增的方式考察每个模块对整个网络模型的作用. 表 1 列出了本文算法在数据集 GWHD 上的消融实验结果. 由表 1 可见, 与原始的 YOLOv7 相比, 本文设计的检测模型在数据集 GWHD 上的 mAP@(0.5~0.95) 提升了 2.11 个百分点.

表 1 本文算法在数据集 GWHD 上的消融实验结果

Table 1 Results of ablation experiment of proposed algorithm on GWHD dataset

方法	输入尺寸/像素	mAP@0.5/%	mAP@(0.5~0.95)/%
YOLOv7	1 024×1 024	93.89	49.89
YOLOv7+NAM	1 024×1 024	94.07	50.92
YOLOv7+SAC+NAM	1 024×1 024	94.83	51.47
YOLOv7+EVC	1 024×1 024	94.21	51.17
本文	1 024×1 024	95.57	52.00

针对数据集实际情况, 原 YOLOv7 与改进后的 YOLOv7 网络模型检测效果对比如图 9 和图 10 所示. 图 9 为目标密集图片检测结果对比. 由图 9 可见, 图 9(A) 左下角存在由于检测目标密集导致的漏检情况, 而图 9(B) 改进后的网络模型则成功检测出目标. 图 10 为遮挡严重图片检测结果对比. 由图 10 可见, 图 10(A) 中心的上方和下方都存在由于遮挡导致的漏检, 而图 10(B) 改进后的网络模型则成功检测出遮挡麦穗. 且改进后的网络检测效果整体置信度得分都高于原 YOLOv7 的检测效果.

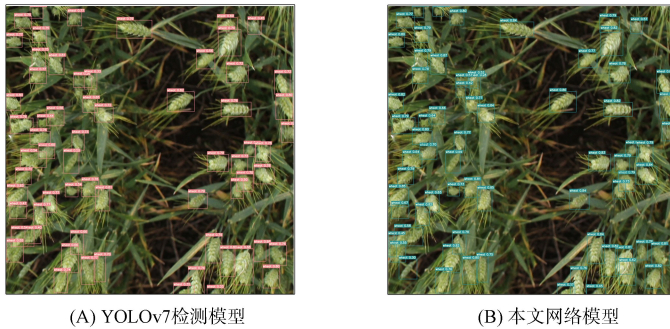


图 9 目标密集图片检测结果对比

Fig. 9 Comparison of detection results of target dense pictures

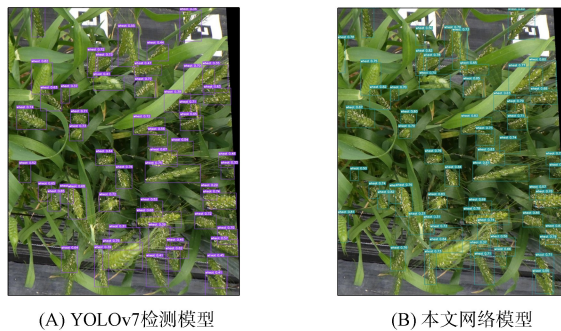


图 10 遮挡图片检测结果对比

Fig. 10 Comparison of masked image detection results

2.5 改进后网络模型与其他网络模型的实验结果对比

下面对改进后的 YOLOv7 网络模型与其他经典网络模型进行对比实验, 以验证其有效性. 在保持配置环境和初始训练参数一致的前提下, 对其进行比较, 结果列于表 2. 由表 2 可见, 当对输入图片做相同的设置后, 改进后的 YOLOv7 网络模型的检测效果最好, 其 mAP 值也有显著提升, 表明改进后的网络模型更适合用于麦穗检测的场景.

表 2 不同网络模型实验结果对比

Table 2 Comparison of experimental results of different network models

网络模型	输入尺寸/像素	mAP@0.5/%	mAP@(0.5~0.95)/%
SSD	1 024×1 024	90.49	47.79
YOLOv5	1 024×1 024	89.07	46.40
YOLOv6	1 024×1 024	88.19	43.55
YOLOv6+SAC	1 024×1 024	89.21	46.71
本文	1 024×1 024	95.57	52.00

综上所述, 针对麦穗数据集检测中存在的问题, 本文引入了一系列改进措施. 首先, 引入混合注意力机制加强对位置特征的提取, 缓解了检测目标密集导致的漏检问题; 其次, 引入能结合不同尺寸的空洞卷积 SAC, 实现提取不同尺度的特征信息, 进一步改善了由于遮挡导致的漏检问题; 最后, 引入增量学习模块 EVC 提高模型的鲁棒性和泛化能力. 实验结果表明, 该算法在全球小麦麦穗数据集 GWHD 上的平均目标检测精度比原 YOLOv7 模型提高了 2.11 个百分点.

参 考 文 献

- [1] MA J C, LI Y X, DU K M, et al. Segmenting Ears of Winter Wheat at Flowering Stage Using Digital Images and Deep Learning [J]. Computers and Electronics in Agriculture, 2020, 168: 105159-1-105159-16.
- [2] 于增源. 基于深度学习的麦穗检测与精准分割方法研究 [D]. 南京: 南京农业大学, 2019. (YU Z Y. Research on Wheat Ears Detection and Accurate Segmentation Method Based on Deep Learning [D]. Nanjing: Nanjing Agricultural University, 2019.)
- [3] 李柯泉, 陈燕, 刘佳晨, 等. 基于深度学习的目标检测算法综述 [J]. 计算机工程, 2022, 48(7): 1-12. (LI K Q, CHEN Y, LIU J C, et al. A Review of Deep Learning Based Target Detection Algorithms [J]. Computer Engineering, 2022, 48(7): 1-12.)
- [4] REN S 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, 2016, 39(6): 1137-1149.
- [5] LIU W, ANGELOV D, ETHAN D, et al. SSD: Single Shot MultiBox Detector [C]//European Conference on Computer Vision. Berlin: Springer, 2016: 21-37.
- [6] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2016: 779-788.
- [7] 李美霖, 芮杰, 金飞, 等. 基于改进 YOLOX 的遥感影像目标检测算法 [J]. 吉林大学学报(地球科学版), 2023, 53(4): 1313-1322. (LI M L, RUI J, JIN F, et al. Target Detection Algorithm for Remote Sensing Images Based on Improved YOLOX [J]. Journal of Jilin University (Earth Science Edition), 2023, 53(4): 1313-1322.)
- [8] WANG D Y, FU Y Y, YANG G J, et al. Combined Use of FCN and Harris Corner Detection for Counting Wheat Ears in Field Conditions [J]. IEEE Access, 2019, 7: 178930-178941.
- [9] SHELHAMER E, LONG J, DARRELL T. Fully Convolutional Networks for Semantic Segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39(4): 640-651.
- [10] 张领先, 陈运强, 李云霞, 等. 基于卷积神经网络的冬小麦麦穗检测计数系统 [J]. 农业机械学报, 2019, 50(3): 144-150. (ZHANG L X, CHEN Y Q, LI Y X, et al. Convolutional Neural Network-Based Wheat Spike Detection Counting System for Winter Wheat [J]. Journal of Agricultural Machinery, 2019, 50(3): 144-150.)
- [11] NEUBECK A, VAN GOOL L. Efficient Non-maximum Suppression [C]//18th International Conference on Pattern Recognition (ICPR'06). Piscataway, NJ: IEEE, 2006: 850-855.
- [12] 陈佳玮, 李庆, 谭巧行, 等. 结合轻量级麦穗检测模型和离线 Android 软件开发的田间小麦测产 [J]. 农业工程学报, 2021, 37(19): 156-164. (CHEN J W, LI Q, TAN Q X, et al. Combining Lightweight Wheat Spike Detection Model and Offline Android Software Development for Field Wheat Yield Measurement [J]. Journal of Agricultural Engineering, 2021, 37(19): 156-164.)
- [13] 张宝朋, 康谦泽, 李佳萌, 等. 轻量化的 YOLOv4 目标检测算法 [J]. 计算机工程, 2022, 48(8): 206-214.

- (ZHANG B P, KANG Q Z, LI J M, et al. Lightweight YOLOv4 Target Detection Algorithm [J]. Computer Engineering, 2022, 48(8): 206-214.)
- [14] 张连屹. 基于改进 YOLOv3 的麦穗检测算法研究 [D]. 长春: 吉林大学, 2021. (ZHANG L Y. Research on Wheat Ears Detection Algorithm Based on Improved YOLOv3 [D]. Changchun: Jilin University, 2021.)
- [15] REDMON J, FARHADI A. YOLOv3: An Incremental Improvement [EB/OL]. (2018-04-08)[2023-01-20]. <https://arxiv.org/abs/1804.02767>.
- [16] 陈乔松, 周丽, 毛彦媚, 等. 基于浅层空间特征融合与自适应通道筛选的目标检测方法 [J]. 江苏大学学报(自然科学版), 2022, 43(1): 67-74. (CHEN Q S, ZHOU L, MAO Y M, et al. A Target Detection Method Based on Shallow Spatial Feature Fusion and Adaptive Channel Screening [J]. Journal of Jiangsu University (Natural Science Edition), 2022, 43(1): 67-74.)
- [17] 欧阳继红, 王梓明, 刘思光. 改进多尺度特征的 YOLO_v4 目标检测方法 [J]. 吉林大学学报(理学版), 2022, 60(6): 1349-1355. (OUYANG J H, WANG Z M, LIU S G. Improved YOLO_v4 Target Detection Method with Multi-scale Features [J]. Journal of Jilin University (Science Edition), 2022, 60(6): 1349-1355.)
- [18] 逢晨曦, 李文辉. 基于注意力改进的自适应空间特征融合目标检测算法 [J]. 吉林大学学报(理学版), 2023, 61(3): 557-566. (PANG C X, LI W H. Adaptive Spatial Feature Fusion Object Detection Algorithm Based on Attention Improvement [J]. Journal of Jilin University (Science Edition), 2023, 61(3): 557-566.)
- [19] 黄硕, 周亚男, 王起帆, 等. 改进 YOLOv5 测量田间小麦单位面积穗数 [J]. 农业工程学报, 2022, 38(16): 235-242. (HUANG S, ZHOU Y N, WANG Q F, et al. Improvement of YOLOv5 to Measure the Number of Spikes per Unit Area of Wheat in the Field [J]. Journal of Agricultural Engineering, 2022, 38(16): 235-242.)
- [20] 高陈强, 董亚盼, 谌放, 等. 一种基于注意力机制的红外小目标检测方法 [J]. 重庆邮电大学学报(自然科学版), 2023, 35(2): 219-226. (GAO C Q, DONG Y P, CHEN F, et al. An Infrared Small Target Detection Method Based on Attention Mechanism [J]. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), 2023, 35(2): 219-226.)
- [21] 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 [EB/OL]. (2022-06-06)[2023-01-30]. <https://arxiv.org/abs/2207.02696>.
- [22] LIU Y C, SHAO Z R, TENG Y Y, et al. NAM: Normalization-Based Attention Module [EB/OL]. (2021-11-24)[2023-02-01]. <https://arxiv.org/abs/2111.12419>.
- [23] QIAO S Y, CHEN L C, YUILLE A. DetectoRS: Detecting Objects with Recursive Feature Pyramid and Switchable Atrous Convolution [EB/OL]. (2020-06-03)[2023-01-15]. <https://arxiv.org/abs/2006.02334>.
- [24] LIU S, QI L, QIN H F, et al. Path Aggregation Network for Instance Segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2018: 8759-8768.
- [25] QUAN Y, ZHANG D, ZHANG L Y, et al. Centralized Feature Pyramid for Object Detection [J]. IEEE Transactions on Images Processing, 2023, 32: 4341-4354.
- [26] 戚玲珑, 高建瓴. 基于改进 YOLOv7 的小目标检测 [J]. 计算机工程, 2023, 49(1): 41-48. (QI L L, GAO J L. Small Target Detection Based on Improved YOLOv7 [J]. Computer Engineering, 2023, 49(1): 41-48.)
- [27] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional Block Attention Module [C]//Proceedings of the European Conference on Computer Vision (ECCV). New York: ACM, 2018: 3-19.

(责任编辑: 韩 啸)