

MNTH-YOLOv8: 一种用于食品包装中蚊虫高效检测的深度学习方法

doi:10.3969/j.issn.1674-7100.2024.03.012

王晓红 张 微

上海理工大学

出版印刷与艺术设计学院

上海 200093

摘 要: 食品安全一直是社会关注的焦点,而在食品包装印刷生产过程中,蚊虫的夹杂会对食品安全构成威胁。针对食品包装质检过程中蚊虫检测仍是人工筛查的现状,以及蚊虫目标尺寸小、所处背景复杂的特点,提出了一种基于深度学习的全自动 MNTH-YOLOv8 检测方法。该方法是在 YOLOv8 强大的目标检测功能基础上,结合通道特征部分卷积模块、SimAM 注意力机制和改进的特征融合模块,并以 CIoU 与归一化 Wasserstein 距离作为定位回归损失函数的优化模型。对真实数据集的检测结果表明, MNTH-YOLOv8 表现出显著优势,不仅有效提高了小目标蚊虫的检测精度,还在保持检测速度的前提下减少了参数量。MNTH-YOLOv8 在食品包装中蚊虫的实时检测应用上拥有广阔前景。

关键词: 食品包装安全; 蚊虫检测; YOLOv8; 小目标检测; SimAM 注意力机制; 特征融合; 归一化 Wasserstein 距离

中图分类号: TB484.1; TP317.4

文献标志码: A

文章编号: 1674-7100(2024)03-0091-08

引文格式: 王晓红, 张 微. MNTH-YOLOv8: 一种用于食品包装中蚊虫高效检测的深度学习方法 [J]. 包装学报, 2024, 16(3): 91-98.

1 研究背景

食品安全一直是社会关注的重点问题。在食品包装生产过程中,蚊虫夹杂在生产纸张当中,不仅难以发现,而且对食品安全构成了威胁。因此,迫切需要一种高效和精确检测食品包装中蚊虫的技术。不同于其他的包装检测技术^[1],蚊虫检测技术需要考虑食品包装复杂多变的背景以及蚊虫目标微小的特点。

目前,食品包装纸上较主流的蚊虫检测方法通常包括两个步骤。首先,采用 AVT (automatic vision technology) 质检系统对包装纸表面进行各种瑕疵检测。

随后,通过人工筛查的方式对所检测到的图像进行进一步的分析,以识别并发现其中存在的蚊虫。这样的做法不仅效率低,而且也无法保证其准确率,常常有漏检、错检的现象。而随着计算机视觉技术的进步,机器学习和深度学习已应用于害虫识别领域。

传统机器学习是用特征描述子提取待检测目标的纹理、颜色、形状等特征,再使用支持向量机 (support vector machines, SVM)^[2]、K 近邻 (k-nearest neighbor classification, KNN)^[3] 等分类算法进行检测。基于机器学习的蚊虫检测技术需要人工设计大量的图像特征来作前提支撑。然而在实际生产过程中,

收稿日期: 2024-03-24

基金项目: 本研究成果受国家新闻出版署智能与绿色柔版印刷重点实验室招标课题资助 (ZBKT202301)

作者简介: 王晓红 (1971-), 女, 陕西汉中, 上海理工大学教授, 主要从事色彩学与色彩应用、印刷质量检测与控制、数字印刷技术研究, E-mail: wang_keyan@163.com

蚊虫所处背景复杂多样,无法获得准确适用的图像特征。因此,机器学习无法实现高精度的蚊虫自动识别。

近年来,深度学习技术迎来了极大的发展。将深度学习用于病虫害识别成为研究热点^[4-9]。相关算法多是特定害虫种类在相对简单背景下的检测,而针对各种所处复杂背景、形态各异的蚊虫难以发挥作用。目前,目标检测算法一般分为双阶段检测和单阶段检测两大主流。双阶段检测方面,Ding W. G.等^[10]提出了一种滑动窗口算法,首先找到候选区域,然后再在此区域内提取特征图来检测昆虫。Liu Z. Y.等^[11]则提出了一种基于R-CNN^[12]的水稻害虫分类算法,以提高对于多种虫类的识别准确率。双阶段检测方法拥有相对较高的检测精度,但其检测速度不能满足实际工业生产要求。以YOLO^[13]系列算法为代表的单阶段检测网络在检测速度方面更具优势。Li K. P.等^[14]在YOLOv3目标检测模型的基础上引入CSPDarknet-53作为主干特征提取网络,以缓解微小目标所带来的问题,同时采用CIOU作为回归预测的损失函数,以有效反映先验框与真实框的相对位置。但该算法准确率有待提高。Song L. M.等^[15]则提出了一种

改进的YOLOv4害虫物种识别算法。在网络中添加DenseNet加速梯度下降,提高了算法对害虫的识别准确率。Dai M.等^[16]设计了一种新的神经网络模型,将SwinTR^[17]和C3TR^[18]机制引入YOLOv5m网络,以便能够捕获更多全局特征,增加感受野,加强网络的特征融合能力。然而该方法在处理某些复杂背景下的微小昆虫时存在性能不足的问题。

综上,基于实际食品包装生产质检的需要,为实现复杂背景下蚊虫小目标的高效准确检测,本文提出一种基于SimAM无参数注意力机制的单阶段特征融合检测网络(MNTH-YOLOv8)。

2 MNTH-YOLOv8 网络模型设计

2.1 网络结构

食品包装纸上蚊虫大多尺寸较小并且所处背景复杂,YOLOv8网络不能满足检测需求。因此,本文以YOLOv8作为基础模型,改进优化卷积模块、特征融合模块,引入SimAM注意力机制,构建MNTH-YOLOv8网络。MNTH-YOLOv8的具体结构如图1所示。

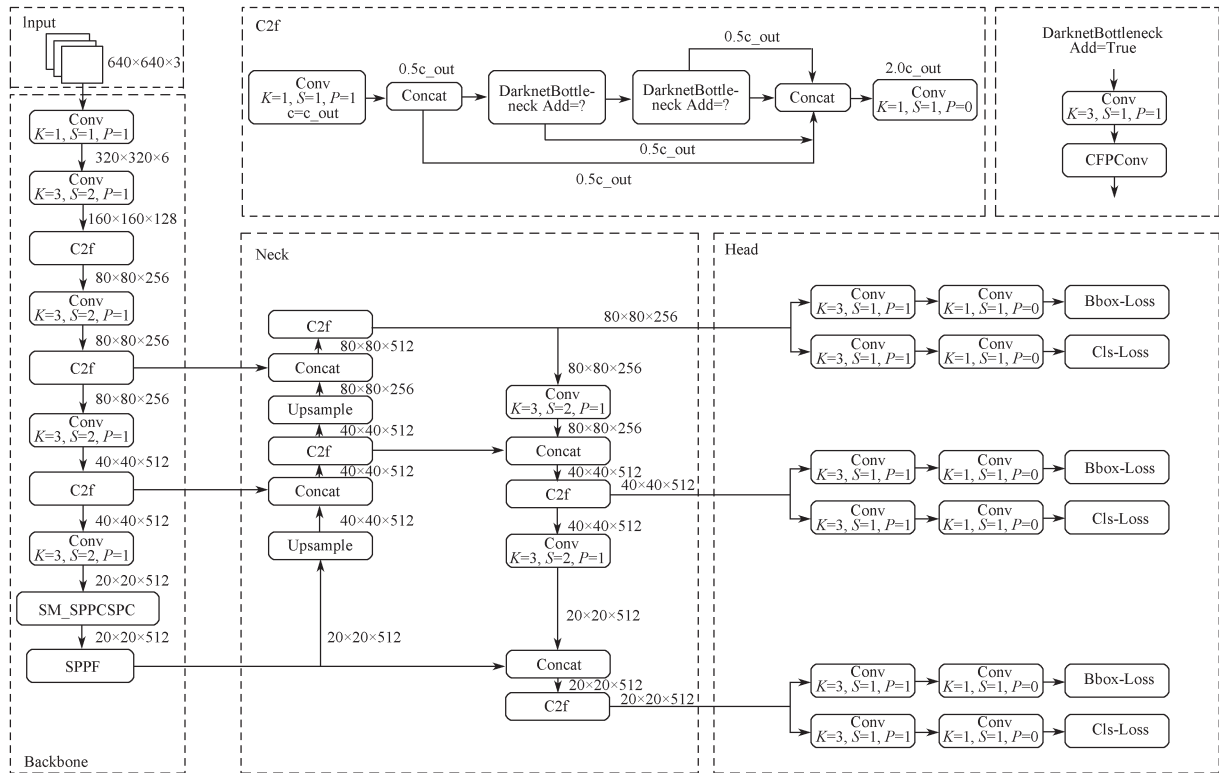


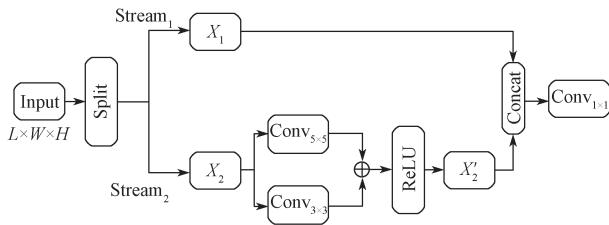
图1 MNTH-YOLOv8 整体网络结构图

Fig. 1 Overall network architecture diagram of MNTH-YOLOv8

首先, 提出了一种通道特征部分卷积模块 CFPCConv^[19], 基于此模块重新设计了 C2f 模块中的 DarknetBottleneck 结构; 基于 SimAM^[20] 注意力机制, 改进 YOLOv8 的 SPPCSPC 结构, 得到 SM_SPPCSPC 模块。然后, 添加跨层信息连接以及多尺度特征自适应融合, 以适应复杂背景下的蚊虫检测。最后, 引入 NWD^[21] 损失函数, 将其与 YOLOv8 的 CIoU 结合作为定位回归损失函数, 以优化小目标的检测性能。

2.2 CFPCConv 模块

一般来说, 卷积神经网络随着网路层数的增加, 特征图经过一系列的处理后其尺寸会逐渐减小, 模型感受野增大的同时也会逐渐丢失位置信息^[22], 这不利于蚊虫小目标的检测。针对此问题, 本文设计了通道特征部分卷积模块 CFPCConv, 并将其整合进 YOLOv8 中的 C2f 模块当中, 减少计算参数的同时有效地整合低级和高级特征信息, 改善模型的性能。CFPCConv 结构如图 2 所示。



注: ⊕ 表示 Concat 操作。

图 2 CFPCConv 结构图

Fig. 2 CFPCConv structure diagram

CFPCConv 模块收到尺寸为 $L \times W \times H$ 的特征图后, 会进行一个二分流处理, X_1 保持不变, 对 X_2 进行相关卷积处理得到 X_2' (见式 (1)); 最后将 X_1 与 X_2' 进行拼接 (见式 (2)), 保留有用的局部信息, 并通过一个逐点卷积来提取未处理分支通道的有用信息, 丰富特征图的多尺度信息。食品包装质检过程中所拍摄的图像会存在大量的冗余信息, 而分流处理不仅减少了计算的梯度压力, 还将不同的特征子空间进行融合, 有助于模型学习更加丰富多元化的特征表示, 使得模型能够有效提取到小目标物体的特征信息。

$$\text{Output} = \text{Conv}(\text{Stream}_1(X_1) + \text{Stream}_2(X_2)), \quad (1)$$

$$X_2' = \gamma(\text{Conv}_{3 \times 3}(X_1) + \text{Conv}_{5 \times 5}(X_2)). \quad (2)$$

式中: + 表示拼接操作; Stream_i 表示对输入的特征图执行通道分流操作后得到的通道特征分支结构; γ 表示 $\text{ReLU}^{[23]}$ 激活函数。

2.3 SM_SPPCSPC 模块

复杂背景的干扰往往会降低模型对于蚊虫小目标物体所在区域的关注度, 从而降低了检测准确率。为解决此问题, 本文通过引入 SimAM 注意力机制设计了 SM_SPPCSPC 模块。一维注意力机制强调不同通道之间的关系, 二维注意力机制强调不同位置之间的关系, 而 SimAM 注意力机制通过神经元感知生成三维注意力机制权重, 同时考虑了特征图中的通道信息和空间位置信息, 根据不同通道和不同位置之间的关系来动态调整注意力权重, 增强模型对于小目标特征信息的感知能力。通过对输入特征图 X 中的每个元素进行加权, 允许模型更加关注小目标区域, 如式 (3) 所示。SimAM 注意力机制结构如图 3 所示。

$$X' = \text{sigmoid}\left(\frac{1}{E}\right) \odot X, \quad (3)$$

式中: E 表示注意力权重; $\odot$ 表示逐元素相乘。

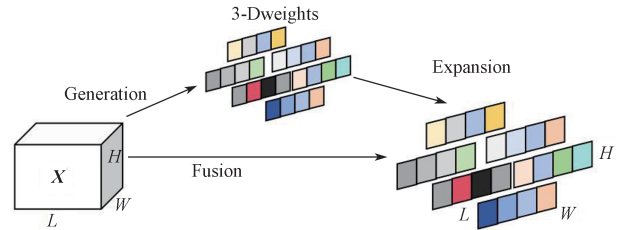


图 3 SimAM 注意力机制结构图

Fig. 3 Structure diagram of SimAM attention mechanism

SM_SPPCSPC 模块如图 4 所示。首先, 利用 SimAM 注意力机制对输入特征图进行处理, 重点强调小目标的关键特征, 并将复杂背景与小目标区域进行细分。其次, 通过不同尺寸 (5×5 、 7×7 、 9×9) 的池化核, 对特征图进行多尺度提取, 以更全面地捕获小目标的特征信息。再次, 对分支流进行 Concat 融合, 增强特征表示。最后, 再利用 SimAM 注意力机制进一步聚焦于密集小目标区域, 减弱复杂背景的干扰, 提升模型的鲁棒性。

2.4 GFPN 特征融合模块

Yolov8 原始网络使用 PANet^[24] 结构来合并特征, 以此来解决特征图逐步高度抽象过程中可能出现的特征信息丢失问题。然而, 该结构仅考虑了相邻层级之间同一尺度和相邻尺度之间同一层级的影响, 导致在特征融合过程中存在较多的信息丢失, 特别是对于小目标的特征。本文引入改进的广义特征金字塔网络 (generalized feature pyramid network, GFPN)^[25], 通过跨尺度连接不同层级与不同尺度, 从而促进深层和

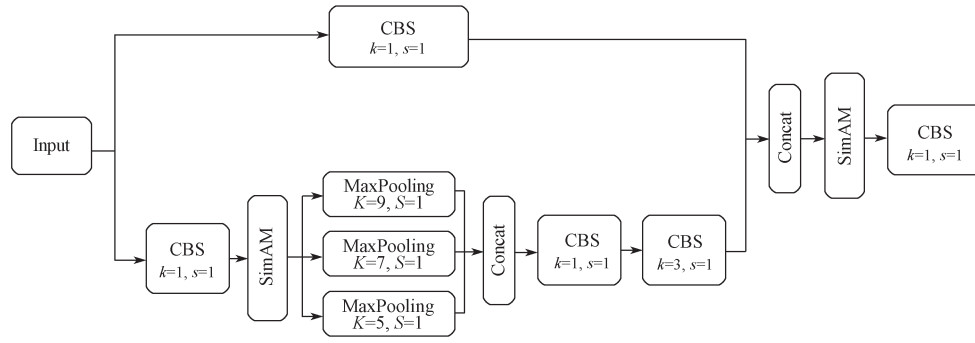


图4 SM_SPPCSPC 结构图

Fig. 4 Structure diagram of SM_SPPCSPC

浅层、大尺度和小尺度之间信息的充分交流。GFPN 结构如图5所示。

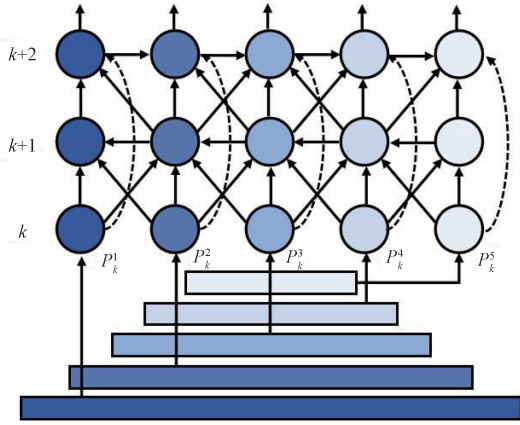


图5 GFPN 结构图

Fig. 5 Structure diagram of GFPN

图5中, k 表示特征金字塔的层级, P_k^i 表示第 k 层的第 i 个尺度特征, 虚线连接表示跨层连接。改进的 GFPN 在 PANet 的基础上引入了跨尺度连接, 在特征金字塔逐步聚拢特征信息的同时充分融合高层语义信息与底层空间信息, 以更有效地传递小目标的特征信息。跨层连接信息融合过程如式(4)所示。

$$P_k^i = \text{Conv}(\text{Concat}(P_1^i, P_2^i, \dots, P_{k-1}^i)), \quad (4)$$

式中: $\text{Concat}()$ 表示前面所有层所产生的特征映射连接; $\text{Conv}()$ 表示 3×3 卷积。

在跨尺度融合中, 考虑到不同层级特征图的尺度大小差异, 为确保特征信息的有效融合, 采用自适应特征融合策略, 具体融合过程见图6。

在进行大尺寸特征图与小尺寸特征图的跨层匹配时, 采用最大池化操作对大尺寸特征图进行下采样处理。相应地, 小尺寸特征图则通过双线性插值进行上采样, 以满足相邻层级之间的跨尺度连接需求, 具体如式(5)所示。

$$P_k^i = \text{Conv}(\text{Concat}(\text{MaxPool}(P_k^{i-1}), P_k^i, \text{Bilinear}(P_k^{i+1}, P_k^{i+1}))), \quad (5)$$

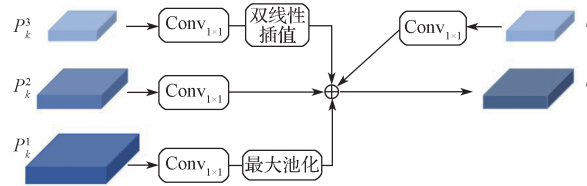


图6 自适应特征融合过程

Fig. 6 Adaptive feature fusion process

2.5 位置损失函数

损失函数用来衡量模型预测结果和真实结果之间的差距。YOLOv8 采用 CIoU 和 DFL 相结合的方法来计算边界回归损失。CIoU 的计算公式

$$L_{\text{CIoU}} = 1 - \text{IoU} + \frac{\rho^2(b, b_{\text{gt}})}{c_w^2 + c_h^2} + \frac{4}{\pi^2} \left(\arctan\left(\frac{w_{\text{gt}}}{h_{\text{gt}}}\right) - \arctan\left(\frac{w}{h}\right) \right)^2, \quad (6)$$

式中: w_{gt} 和 h_{gt} 分别表示真实框的宽和高; $\rho(b, b_{\text{gt}})$ 表示预测框与真实框之间的欧式距离; c_w 和 c_h 表示由预测框和真实框组成的最小包围框的宽和高; IoU 表示预测框与真实框相交比。

然而当目标发生位置偏移时, CIoU 容易产生剧烈震荡。为解决此问题, 引入归一化 Wasserstein 距离 (normalized wasserstein distance, NWD)。NWD 利用二维高斯分布计算预测框与真实框的相似度, 不受尺度影响, 适用于小目标检测, 且计算效率更高。NWD 计算公式如下:

$$\text{NWD}(N_a, N_b) = \exp\left(-\frac{\sqrt{W_2^2(N_a, N_b)}}{C}\right), \quad (7)$$

$$W_2^2(N_a, N_b) = \left\| \left[\begin{matrix} cx_a, cy_a, w_a, h_a \end{matrix} \right]^T, \left[\begin{matrix} cx_b, cy_b, \frac{w_b}{2}, \frac{h_b}{2} \end{matrix} \right]^T \right\|_2^2, \quad (8)$$

式中: C 表示划分数据集的类别数; N_a 和 N_b 分别表示根据边界框 $A = (cx_a, cy_a, w_a, h_a)$ 和 $B = (cx_b, cy_b, w_b, h_b)$ 建模的高斯分布; $W_2^2(N_a, N_b)$ 表示距离度量。

将 CIoU 与 NWD 相结合作为位置损失函数 L , 即

$$L = \alpha \cdot L_{CIoU} + (1 - \alpha) \cdot NWD, \quad (9)$$

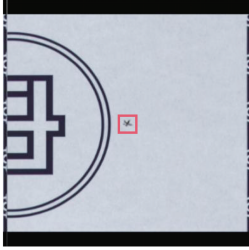
式中 α 表示 CIoU 的权重占比。

3 实验结果与分析

3.1 实验数据集

本文数据集来自于真实生产样本, 涵盖了多样化的包装品背景, 共计 1930 张。数据样本由工业生产视觉质检系统 AVT 所拍摄。表 1 为部分蚊虫数据样本。从表 1 可以看出, 待检测蚊虫尺寸小, 并且大多所处背景复杂。

表 1 数据样本
Table 1 Data samples

样本说明	样本 1	样本 2
薄纸、多色印刷、背面		
薄纸、多色印刷、正面		
硬壳纸、多色印刷、正面		
薄纸、单色印刷、正面		

3.2 评价指标

本文实验选用均值平均精度 (mean average

precision, mAP)、参数量以及误判率 (F) 作为模型性能评价指标。mAP 结合目标检测中的精确率 (P)

和召回率（ R ），能够反映出模型的准确性以及泛化能力；误判率是为了检测模型在实际生产质检过程中的性能表现。 mAP 以及误判率的计算公式如下：

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

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

$$AP = \int_0^1 p(r)dr, \quad (12)$$

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i, \quad (13)$$

$$F = \frac{FP}{FP + TN}, \quad (14)$$

式中： TP 为预测正确的正样本数量； TN 为预测正确的负样本数量； FP 为实际为负样本但被错误地预测为正样本的数量； FN 为实际为正样本但被错误预测为负样本的数量； AP 为单类别平均精度； $p(r)$ 为召回率 r 处的精度。

3.3 对比实验

为证明本文改进算法的优越性和有效性，将模型与其他 YOLO 系列算法进行了比较，结果如表 2 所示。YOLOv4 结合 CSPDarknet53 骨干网和 PAN 特征融合，实现速度和精度之间的良好平衡；YOLOv5 采用 Focus 结构增强感受野，使用 Mosaic 数据增强，网络的鲁棒性能更好；YOLOv6 采用一种新的自蒸馏策略来提高小模型性能，简化 YOLOv5 中的 SPPF 模块，采用辅助训练 AAT (anchor-aided training) 策略；YOLOv7 采用高效的网络架构 ELAN (efficient layer aggregation networks)。

表 2 YOLO 系列算法对比实验

Table 2 Comparative experiment of YOLO series algorithms

算法	参数量 / 10^6	mAP/%	F/%
YOLOv4	64.0	85.3	24
YOLOv5s	7.2	84.9	27
YOLOv5l	46.2	86.3	28
YOLOv6m	34.2	81.7	20
YOLOv7	37.2	94.6	17
YOLOv8s	11.1	90.1	13
本文算法	7.4	98.8	2

从表 2 可以看出，早期的 YOLO 系列算法，结构复杂，参数量大，对于蚊虫的检测精度也较

低；本文提出的 MNTH-YOLOv8 采用通道卷积模块 CFPCConv，结合 SimAM 注意力机制以及改进的特征融合结构 GFPN，降低参数数量的同时提高了检测精度。

3.4 消融实验

为了验证本文提出的每种改进策略的有效性，使用自制数据集对基准模型 YOLOv8s 进行消融实验，并且在验证集上测试对小目标的检测性能，实验结果如表 3 所示。由表可知，与基准模型 YOLOv8s 相比，在实验 b 中，基准模型加入融合 SimAM 注意力机制的 SM_SPPCSPC 模块，mAP 值提升了 2.2%， F 值降低了 5%；在实验 c 中，基准模型引入改进特征融合模块 GFPN，mAP 值提高 3.4%；在实验 d 中，基准模型使用 CFPCConv 之后，参数量降低了 9.9%，同时 mAP 值提升了 1.6%；在实验 g 中，融合 NWD 参与模型训练，算法各项检测指标均得到了提升，参数量下降了 33.3%，mAP 值达到 98.8%，提升了 8.7%， F 值下降为 2%，这证明了本文算法各改进部分的有效性。

表 3 消融实验

Table 3 Ablation experiment

实验编号	模 型	参数量 / 10^6	mAP/%	F/%
a	YOLOv8s	11.1	90.1	13
b	YOLOv8s+SM_SPPCSPC	12.4	92.3	8
c	YOLOv8s+GFPN	13.8	93.5	10
d	YOLOv8s+CFPCConv	10.0	91.7	13
e	YOLOv8s+SM_SPPCSPC+ CFPCConv	6.7	95.9	8
f	YOLOv8s+SM_ SPPCSPC+CFPCConv+GFPN	7.1	96.8	3
g	YOLOv8s+SM_ SPPCSPC+CFPCConv+ GFPN+NWD (本文算法)	7.4	98.8	2

4 结语

本文针对在食品包装生产过程中蚊虫检测容易出现误检、漏检等问题，提出了 MNTH-YOLOv8 网络模型，并通过自制数据集来训练和验证算法在实际工业生产中的表现。实验结果表明，MNTH-YOLOv8 网络模型具有较高的检测精度和识别效率，参数量为 7.4×10^6 ，mAP 值为 98.8%， F 值为 2%。可见，MNTH-YOLOv8 应用于在线食品包装品生产质检是可行有效的。

参考文献:

- [1] 王耀南, 刘学兵, 张 辉, 等. 机器视觉技术在包装行业研究进展与应用综述 [J]. 包装学报, 2022, 14(2): 1-14, 107.
WANG Yaonan, LIU Xuebing, ZHANG Hui, et al. Research Progress and Application Review of Machine Vision Technology in Packaging Industry[J]. Packaging Journal, 2022, 14(2): 1-14, 107.
- [2] KURILOVÁ V, GOGA J, ORAVEC M, et al. Support Vector Machine and Deep-Learning Object Detection for Localisation of Hard Exudates[J]. Scientific Reports, 2021, 11: 16045.
- [3] TAUNK K, DE S, VERMA S, et al. A Brief Review of Nearest Neighbor Algorithm for Learning and Classification[C]//2019 International Conference on Intelligent Computing and Control Systems (ICCS). Madurai: IEEE, 2019: 1255-1260.
- [4] ALBANESE A, NARDELLO M, BRUNELLI D. Automated Pest Detection with DNN on the Edge for Precision Agriculture[J]. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 2021, 11(3): 458-467.
- [5] MEKHA J, PARTHASARATHY V. An Automated Pest Identification and Classification in Crops Using Artificial Intelligence: A State-of-Art-Review[J]. Automatic Control and Computer Sciences, 2022, 56(3): 283-290.
- [6] LIU J, WANG X W. Plant Diseases and Pests Detection Based on Deep Learning: A Review[J]. Plant Methods, 2021, 17(1): 22.
- [7] TURKOGLU M, YANIKOĞLU B, HANBAY D. PlantDiseaseNet: Convolutional Neural Network Ensemble for Plant Disease and Pest Detection[J]. Signal, Image and Video Processing, 2022, 16(2): 301-309.
- [8] WAHEED H, ZAFAR N, AKRAM W, et al. Deep Learning Based Disease, Pest Pattern and Nutritional Deficiency Detection System for “Zingiberaceae” Crop[J]. Agriculture, 2022, 12(6): 742.
- [9] 朱香元, 聂 轰, 周 旭. 基于 TPH-YOLOv5 和小样本学习的害虫识别方法 [J]. 计算机科学, 2022, 49(12): 257-263.
ZHU Xiangyuan, NIE Hong, ZHOU Xu. Pest Identification Method Based on TPH-YOLOv5 Algorithm and Small Sample Learning[J]. Computer Science, 2022, 49(12): 257-263.
- [10] DING W G, TAYLOR G. Automatic Moth Detection from Trap Images for Pest Management[J]. Computers and Electronics in Agriculture, 2016, 123(C): 17-28.
- [11] LIU Z Y, GAO J F, YANG G G, et al. Localization and Classification of Paddy Field Pests Using a Saliency Map and Deep Convolutional Neural Network[J]. Scientific Reports, 2016, 6: 20410.
- [12] GIRSHICK R. Fast R-CNN[C]//2015 IEEE International Conference on Computer Vision (ICCV). Santiago: IEEE, 2015: 1440-1448.
- [13] REDMON J, DIVVALA S, GIRSHICK R, et al. You only Look Once: Unified, Real-Time Object Detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016: 779-788.
- [14] LI K P, ZHU J S, LI N Q, et al. Insect Detection and Counting Based on YOLOv3 Model[C]//2021 IEEE 4th International Conference on Electronics Technology (ICET). Chengdu: IEEE, 2021: 1229-1233.
- [15] SONG L M, LIU M Y, LIU S H, et al. Pest Species Identification Algorithm Based on Improved YOLOv4 Network[J]. Signal, Image and Video Processing, 2023, 17(6): 3127-3134.
- [16] DAI M, DORJOY M M H, MIAO H, et al. A New Pest Detection Method Based on Improved YOLOv5m[J]. Insects, 2023, 14(1): 54.
- [17] SUBAKAN C, RAVANELLI M, CORNELL S, et al. Attention Is All You Need in Speech Separation[C]//2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Toronto: IEEE, 2021: 21-25.
- [18] KONDO Y, NUMADA M, KOSHIMIZU H. A Fast Hough Transform Based on C3TR Using Shift Operation[C]//2015 21st Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV). Mokpo: IEEE, 2015: 1-6.
- [19] THOMAS H, QI C R, DESCHAUD J E, et al. KPConv: Flexible and Deformable Convolution for Point Clouds[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul: IEEE, 2019: 6410-6419.
- [20] YANG L X, ZHANG R Y, LI L D, et al. SimAM: A Simple, Parameter-Free Attention Module for Convolutional Neural Networks[C/OL]//2021 International Conference on Machine Learning. [2023-12-03]. <http://proceedings.mlr.press/v139/yang21o/yang21o.pdf>.
- [21] WANG J W, XU C, YANG W, et al. A Normalized Gaussian Wasserstein Distance for Tiny Object Detection[EB/OL]. [2023-09-19]. <http://arxiv.org/>

- abs/2110.13389v2.
- [22] CHEN G, WANG H T, CHEN K, et al. A Survey of the Four Pillars for Small Object Detection: Multiscale Representation, Contextual Information, Super-Resolution, and Region Proposal[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2022, 52(2): 936–953.
- [23] AGARAP A F. Deep Learning Using Rectified Linear Units (ReLU)[EB/OL]. [2023–08–10]. <http://arxiv.org/abs/1803.08375v2>.
- [24] LIU S, QI L, QIN H F, et al. Path Aggregation Network for Instance Segmentation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 8759–8768.
- [25] JIANG Y Q, TAN Z Y, WANG J Y, et al. GiraffeDet: A Heavy-Neck Paradigm for Object Detection[EB/OL]. [2023–05–17]. <https://arxiv.org/pdf/2202.04256.pdf>.
- (责任编辑: 邓 彬)

MNTH-YOLOv8: A Deep Learning Approach for Efficient Mosquito Detection in Food Packaging

WANG Xiaohong, ZHANG Wei

(School of Communication and Art Design, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: Food safety has always been a focus of social concern. However, the presence of mosquitoes and other insects during the food packaging and printing process poses a serious challenge to food safety. Aiming at the current situation of manual screening for mosquitoes and other insects during food packaging quality inspection, and account of the small size of insect targets and the complexity of their backgrounds, a fully automatic MNTH-YOLOv8 detection method based on deep learning was proposed. this method was based on the powerful object detection capability of YOLOv8, combined with channel-wise partial convolution modules and SimAM attention mechanism, with CIoU and normalized Wasserstein distance as the localization regression loss function. Experimental results demonstrated significant advantages of the proposed method in real datasets. It not only effectively improved the detection accuracy of small insect targets but also significantly reduced the parameter count while maintaining detection speed, indicating its great prospect in the application of real-time detection of mosquitoes and other insects in food packaging.

Keywords: food packaging safety; mosquito detection; YOLOv8; small target detection; SimAM attention mechanism; feature fusion; normalized Wasserstein distance