doi:10.6041/j.issn.1000-1298.2023.03.030

基于 YOLOX 的复杂背景下木薯叶病害检测方法

宋　玲[1,2]　曹　勉[1]　胡小春[3]　贾沛沅[1]　陈　燕[1,2]　陈宁江[1,2]

(1.广西大学计算机与电子信息学院,南宁 530004;2.广西多媒体通信与网络技术重点实验室,南宁 530004;
3.广西财经学院大数据与人工智能学院,南宁 530007)

摘要:为解决田间环境下由于叶片间遮盖和堆叠等因素引起的木薯叶病害识别困难的问题,本文提出一种基于改进 YOLOX 网络的木薯叶病害检测(Cassava leaf disease detection,CDD)模型。首先,对复杂背景下木薯叶病害图像数据集进行数据增强,以减少环境影响造成的识别困难。其次,在 YOLOX 网络的基础上,使用多尺度特征提取模块加强细粒度特征提取并降低模型计算量,同时嵌入通道注意力机制,提高网络的表征能力。最后,结合质量焦点损失函数作为分类损失函数辅助网络收敛,提高目标分类的准确性。实验结果表明,提出的 CDD 模型对复杂背景下木薯叶病害进行检测,网络参数量为 5.04×10^6,平均精度均值达 93.53%,比基础模型高 6.02 个百分点,综合检测能力优于多种主流模型。因此,本文提出的 CDD 模型对田间木薯叶病害具有更快更准确的检测能力,为实现农作物病害检测提供了可借鉴的方法。

关键词:木薯叶病害;复杂背景;YOLOX;目标检测;多尺度特征

中图分类号:TP391.4　　**文献标识码**:A　　**文章编号**:1000-1298(2023)03-0301-07　　**OSID**:

Detection of Cassava Leaf Diseases under Complicated Background Based on YOLOX

SONG Ling[1,2]　CAO Mian[1]　HU Xiaochun[3]　JIA Peiyuan[1]　CHEN Yan[1,2]　CHEN Ningjiang[1,2]

(1. *College of Computer and Electronic Information*, *Guangxi University*, *Nanning* 530004, *China*
2. *Guangxi Key Laboratory of Multimedia Communications and Network Technology*, *Nanning* 530004, *China*
3. *College of Big Data and Artifical Intelligence*, *Guangxi University of Finance and Economics*, *Nanning* 530007, *China*)

Abstract: The present method has some difficulties in recognizing cassava leaf diseases in a field environment, such as covering and stacking between leaves. Based on the YOLOX network, cassava leaf disease detection (CDD) model was proposed. Firstly, the cassava leaf disease image data under complex background was augmented to reduce the recognition difficulty caused by environmental impact. Secondly, built on the YOLOX network, the lightweight multi-scale feature extraction (LME) module was used to strengthen fine-grained feature extraction and reduce the amount of model calculation. At the same time, the channel attention mechanism was embedded to improve the representation ability of the network. Finally, the quality focal loss was used as a part of the classification loss to assist the network convergence and improve the accuracy of target classification. In conclusion, the proposed CDD model can detect cassava leaf disease under complex background. The amount of network parameters was 5.04×10^6 and the mean average precision was 93.53%, which was 6.02 percentage points higher than that of the non-optimized network model. Comprehensive detection ability was better than that of previous models. Therefore, the proposed method CDD had faster and more accurate detection ability for cassava leaf diseases in the field, and provided a reference method for realizing intelligent field detection.

Key words: cassava leaf disease; complicated background; YOLOX; object detection; multi-scale feature

收稿日期:2022-04-14　修回日期:2022-05-26

基金项目:国家自然科学基金项目(62162003)、广西重点研发计划项目(桂科 AB19110050)和南宁市科技重大专项(20211005)

作者简介:宋玲(1967—),女,教授,主要从事物联网及大数据计算研究,E-mail:731486203@qq.com

通信作者:胡小春(1974—),男,副教授,主要从事智能算法及应用研究,E-mail:huxch999@163.com

0 引言

木薯是世界上主要的薯类作物之一,在生长过程中易受到多种病害的影响,导致产量大幅减少,准确快速地识别其病害种类是当务之急。随着计算机技术的发展,精确率高、实时性强的病害图像智能识别方法取得较大进展,是精准农业发展的关键技术之一[1]。图像识别方法可以自动高效地识别病斑的形状、颜色、纹理和位置,大幅缩减识别耗费的时间和人力。但此类研究多用于处理简单背景的病害数据,对识别田间杂乱背景的数据泛化能力不足[2];集成识别多种作物病害的方法反而出现作物品种的错误判断[3];无法进行同一图像上不同病害的识别[4]。

随着深度学习的进一步发展,更精准的图像检测方法解决了单图像多目标分类与定位的问题。该算法分为两类,一类是基于启发式方法或卷积神经网络产生建议区域的两阶段算法[5]。此类算法效果较好,但速度普遍较慢,代表算法为 R-CNN 系列算法[6-7]等。此类算法已经成功应用在病虫害检测[8]、植株检测与计数[9]、果实采摘[10]等方面,检测精度获得较大的提升。张乐等[11]在 Faster R-CNN 网络中引入共享卷积,利用多种特征提取网络对病害特征进行提取,实现油菜与杂草的高效目标识别。李鑫然等[12]使用特征金字塔网络融合丰富的深浅层特征,对精确感兴趣区域进行池化操作以减少像素偏差,对自然条件的苹果叶片病害检测效果良好。另一类是使用单网络直接预测不同目标的类别与位置的单阶段算法。此算法效果稍低于两阶段,但速度快且实时性好,代表算法为 YOLO 系列[13]、SSD[14]等算法。在实时性较强田间作物监控[15]和农业机器人采摘与筛选[16]等方面已有研究。刘君等[17]使用改进 YOLO v3 的番茄病虫害检测算法,平均精度有较大提升。RAMCHARAN 等[18]使用 SSD 对木薯多种病害进行检测,实现对真实图像轻重病症的分类。

农作物病害检测存在许多难点,因光线变化、叶片相互遮挡、相似病斑等前景因素,以及土壤和杂草等复杂的背景因素对病害检测造成的干扰,导致上述研究难以提取有效病害特征,影响检测效果。为解决以上难点,本文以 4 种木薯病害为研究对象,选用实时性强的轻量化单阶段检测方法,提出一种改进 YOLOX 的木薯叶病害检测(CDD)模型。将自然环境中的木薯叶片病害图像进行数据增强等处理后,利用多尺度特征提取模块、SE 注意力模块和质量焦点损失函数改进模型,增强模型特征提取能力,使模型获得高精确率和强鲁棒性。

1 数据与处理

本文采用 Kaggle 比赛的复杂背景下木薯叶病害图像数据进行研究,数据集由木薯细菌性枯萎病(Cassava bacterial blight, CBB)、木薯褐条病(Cassava brown streak disease, CBSD)、木薯绿斑驳病(Cassava green mottle, CGM)、木薯花叶病(Cassava mosaic disease, CMD)4 种病害图像组成,共计获取 2 080 幅 RGB 病害图像。为了区分病害的严重程度,参考 NY/T 3005—2016《热带作物病虫害监测技术规程》规定的病情分级标准对病害程度分为Ⅰ、Ⅱ两个等级,Ⅰ级表示病斑面积占叶片面积小于 1/2,Ⅱ级表示病斑面积占叶片面积大于 1/2。经人工筛选后,使用 Labelme 软件对数据集进行标注,对每个完整的木薯带病叶片用矩形框进行框选,一个矩形框对应一个病害样本,则每幅图像至少包含一种病害样本。从数据集挑选的 4 种病害典型样例如图 1 所示。



图 1 木薯叶病害数据集

Fig. 1 Dataset of cassava leaves disease

图像检测效果往往受到拍摄角度、叶片遮盖、光照、复杂背景等因素影响。为了让模型在复杂环境下达到更好的检测效果,增强模型的鲁棒性和泛化性,将原始木薯叶病害数据集的 80%(共 1 677 幅图像)作为数据集,使用数据增强工具库 Albumentation 进行处理,原始木薯叶病害数据集剩余的 20%(共 403 幅图像)作为测试集。随机选取数据集的 80% 作为训练集,在训练集中随机选取 20% 作为验证集,用于在训练过程中调整模型参数。

为了模拟不同的拍摄视角,对数据集图像进行垂直翻转、水平翻转以及多次随机顺时针 90°旋转处理;加入高斯噪声和动态模糊模拟干扰环境;使用伽马变换、自适应直方图变换、随机亮度等方法增强

图像对比度，用于提高模型获取目标信息的能力。对数据集进行数据增强后，共获得 5 970 幅叶片病害图像，总计 9 355 个病害样本，具体的病害分类等级及其样本数如表 1 所示。

表 1　数据集样本数量

Tab. 1　Number of dataset samples　　个

病害名称	等级	原始样本数量	增强后样本数量	测试集样本数量
CBB	Ⅰ	268	915	51
	Ⅱ	461	1 383	99
CBSD	Ⅰ	400	1 200	68
	Ⅱ	326	978	59
CGM	Ⅰ	367	1 101	76
	Ⅱ	440	1 320	96
CMD	Ⅰ	349	1 058	69
	Ⅱ	500	1 400	116
总计		3 111	9 355	634

2　模型结构设计

木薯叶病害检测属于图像目标检测任务，需要准确地对图像上不同位置的木薯叶病害进行分类与定位。本文以轻量化的目标检测 YOLOX 模型[19]为原始模型，在此基础上加强有效特征提取以改进模型结构，提出一种改进的木薯叶病害检测模型（CDD）。

2.1　YOLOX 模型

YOLOX 是不使用先验框的单阶段目标检测模型，由主干网络（Backbone）、加强特征融合网络（Neck）和分类回归网络（Head）组成。主干网络使用 CSPDarkNet53 作为特征提取器，该网络使用残差结构在一定程度上解决反向传播时的权重矩阵退化问题[20]。输入图像在主干网络先经过通道拆分重组（Focus）结构，目的是在减少计算量的情况下，减少低层特征提取的损失。经过主干网络输出由低到高层级的特征图，低级特征图反映诸如边距和纹理之类的详细信息，高级特征图反映概括性的全局信息。通过空间金字塔池化（Spatial pyramid pooling，SPP）模块使用不同核大小的空间金字塔池化层对最高层信息进行融合。Neck 采用 PANet 进行上下采样并将提取的特征再次融合。分类回归网络使用解耦头（Decoupled head）结构，分别实现预测分类和回归。

YOLOX 使用了 Mosaic 数据增强和 SimOTA 自动标签分类，前者拼接 4 幅图像实现数据训练时增强，丰富图像背景并提升训练速度；后者可为不同大小的目标动态匹配正样本。使用了标签平滑方法和余弦退火算法对损失和学习率进行自动调整。针对该模型在复杂背景中多目标的特征提取能力略有不足、难分类样本界定不够准确的问题进行改进，提高检测精度。

2.2　CDD 模型

2.2.1　融合多尺度特征提取模块

由于单一尺度卷积核限制了特征提取的感受野大小，导致丢失部分有效特征。为了有效地提取不同尺度和位置的木薯叶病害信息，将多尺度特征提取模块添加到 CDD 主干网络中的 CSP 层的残差结构部分。普通的残差结构由卷积层和恒等映射叠加构成，目的是加深网络的同时减少参数量，防止深层网络退化导致精确度下降。GAO 等[21]提出的 Res2Net 模块在单个残差块内构造分层的残差连接，通过增加残差块内的感受野提取多尺度的特征表示。本文对 Res2Net 模块进行缩减并进行特征分组，得到参数量少且更高细粒度的轻量化多尺度特征提取（Lightweight multi-scale feature extraction，LME）模块。LME 模块首先使用 1×1 卷积层将特征图 m 的通道维度变为 C（图 2 中①表示卷积层，由卷积、正则化和 ReLU 激活函数构成）。其次将特征图 m 均分为通道数均为 $C/3$ 的 m_1、m_2、m_3，再将特征图 m_1、m_2 输入到对应的 3×3 卷积层中获得 n_1 和 n_2。其中 3×3 卷积层使用小尺寸卷积核替换大尺寸卷积核，能有效地融合不同小特征图的信息。利用组卷积增加卷积的对角相关性，学习特征多分组的内容以加强表征能力[22]。最后将特征图拼接到通道数 C，并使用 SE 注意力模块分配特征权重，自适应校准各通道的特征响应。

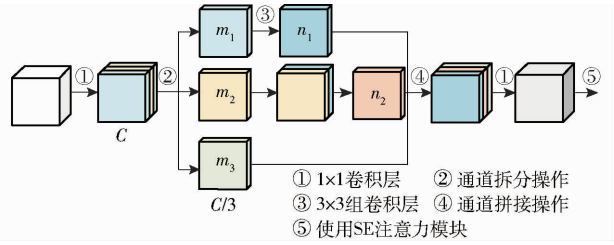

图 2　LME 模块结构图

Fig. 2　Architecture of LME

2.2.2　融合通道注意力模块

为了加强网络对不同病斑特征的提取和区分能力，减少复杂背景产生的噪声因素影响，本文在主干网络的多尺度特征提取模块中与各级特征图后添加 SE 注意力模块（Squeeze-and-excitation block，SE-Block）。注意力机制是一种实现自适应注意的方式，通过模仿人的视觉感知，根据需求关注并学习特定部分内容。在新场景下，注意力仍会集中于学习过的特征。该模块通过调整有效和无效特征的权重，增强模型识别有效信息的能力，进而提升模型识

别效果。

模块的基本结构如图 3 所示。SE 模块由压缩操作 F_{sq}(Squeeze)、激发操作 F_{ex}(Excitation)和分配操作 F_{scale}(Scale)3 部分组成,首先调整尺寸为 $H\times W\times C$ 的特征图,将其特征映射为 $H\times W\times C'$。其次通过 F_{sq} 将每层全局平均池化的信息压缩至 1×1,在 F_{sq} 中,经过全连接层放缩特征映射后,加入 ReLU 激活和 Sigmoid 函数获得最终的特征权重。F_{scale} 为特征权重与原特征图逐通道相乘,完成权重分配。

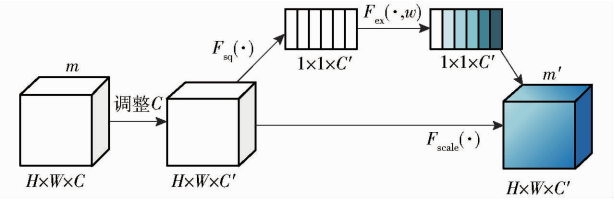

图 3 SE - Block 结构图

Fig. 3 Architecture of SE - Block

2.2.3 分类损失函数改进

为了解决原模型存在难分类样本界定不准确的问题,本文对分类损失函数进行改进。由于 YOLOX 为单阶段的无锚目标检测算法,该算法虽可加速目标分类和坐标回归,但同时会产生大量冗余的锚框,从而容易导致预测框包含目标的正样本与不包含目标的负样本在数量上差距悬殊。LI 等[23]提出的质量焦点损失(Quality focal loss,QFL)有效缓解了正负样本差距大与难分类样本界定的问题。QFL 使用质量估计联合表示预测框的置信度得分和分类得分,质量是指预测的回归框与真实框的交并比(Intersection over union,IoU),直观展示相应目标的准确度。QFL 对不同质量的正负样本赋予对应的权重,有效缓解了病害相似难分类的问题。YOLOX 损失由分类损失、回归损失、定位损失组成,本文在此基础上加入质量焦点损失辅助分类分支收敛,分类损失函数 L_{cls} 的计算公式为

$$L_{cls}=-\sum\alpha_i[(1-y_i)\lg(1-\sigma_i)+y_i\lg\sigma_i] \quad (1)$$

其中

$$\alpha_i=\begin{cases}|y_i-\sigma_i|^{\beta}+\omega & (x_i\in X)\\ |y_i-\sigma_i|^{\beta} & (x_i\notin X)\end{cases} \quad (2)$$

$$\sigma_i=\mathrm{sigmoid}(x_i) \quad (3)$$

式中 x_i——样本 i 的类别分数

y_i——样本 i 的质量分数

X——simOTA 方法选择的正样本

ω、β——调整因子,取 $\omega=1$,$\beta=2$

本文总损失函数 L 由分类损失 L_{cls}、目标损失 L_{obj} 和回归损失 L_{reg} 构成,在 L_{reg} 中使用广义交并比损失 GIOU Loss[24]。总损失计算公式为

$$L=L_{cls}+L_{obj}+L_{reg} \quad (4)$$

其中

$$L_{reg}=\sum\left(\frac{|A\cap B|}{|A\cup B|}-\frac{\left|\frac{S_{AB}}{A\cap B}\right|}{|S_{AB}|}\right) \quad (5)$$

$$L_{obj}=-\sum[(1-y_i)\lg(1-\sigma_i)+y_i\lg\sigma_i] \quad (6)$$

式中 A——真实框坐标 B——预测框坐标

S_{AB}——A 与 B 的最小闭包区域面积

2.2.4 模型整体结构

CDD 模型结构图如图 4 所示,图中 CBS 表示卷积层,CSP 表示残差块,SPP 表示空间金字塔池化层,LME 表示多尺度特征提取模块,Decoupled head 表示分类回归层。将图像通道调整为 $640\times 640\times 3$ 后输入网络,使用 CSPDarkNet53 作为主干特征提取网络,使用 LME 模块替代 CSP 的部分卷积层。在改进网络中,输入图像经过主干网络提取特征后,输出尺度分别为 $20\times 20\times 32$、$40\times 40\times 16$、$80\times 80\times 8$ 的特征图,分别通过 PANet 结构融合高低层信息获得 3 种不同尺度的特征图,分别经过解耦头后获得预测信息,用损失函数计算预测值与真实值的偏差,实现木薯叶病害的自动检测。

3 实验

3.1 实验方法和参数

实验环境采用 Pytorch 框架,CUDA11.1 的

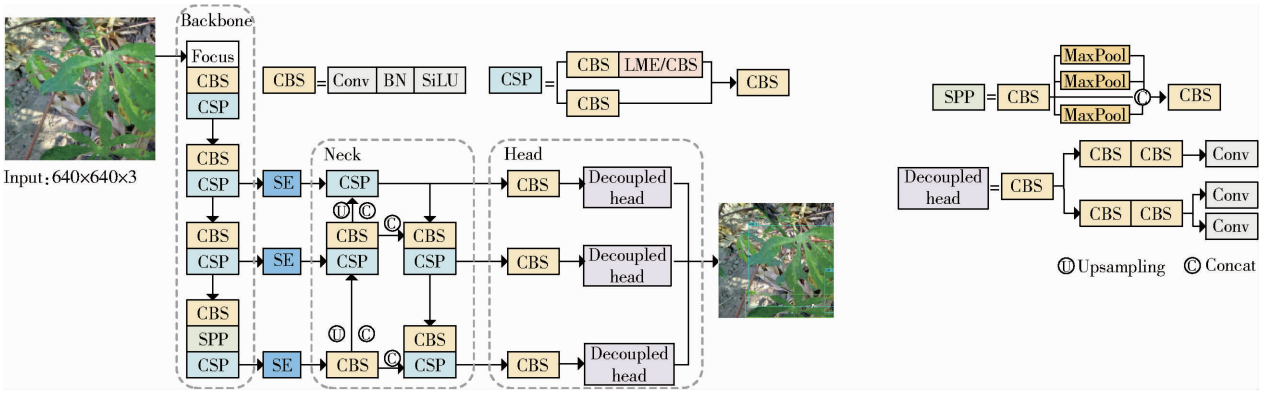

图 4 CDD 模型结构图

Fig. 4 Main architecture of CDD

Ubuntu 和 NVIDIA RTX3090。使用木薯叶病害数据集对模型进行训练；使用 Mosaic 增强提升训练速度；使用余弦退火学习率和标签平滑方法优化损失函数；通过 ImageNet 的均值和方差进行归一化，平衡几何变换的变化。

经多次调试后最优的网络参数如下：网络模型的批处理大小为 8，迭代次数为 500 次；网络通过 Xavier[25] 方法进行初始化，使用随机梯度下降方法优化参数。初始学习率为 1×10^{-4}，优化器动量大小为 0.9，衰减系数为 5×10^{-4}。使用基于 ImageNet 预训练的参数进行迁移学习。在后期处理中以原始图像大小对测试集进行检测并将输出结果与主流模型对比。

3.2　训练指标

为了客观评价病害检测效果，采用 F1 值（F1-score）、平均精度均值（mAP）、浮点计算量（Floating point operations，FLOPs）、参数量作为衡量指标。F1 值是分类问题最常见的评估指标，是精确率 P（Precision）和召回率 R（Recall）的调和平均数。F1 值越大说明模型质量越高。参数量、FLOPs 是与模型性能相关的重要参数，参数量过多占用大量的运行和存储资源，难以用于实际部署。FLOPs 用于衡量模型的计算复杂度，表示模型中乘法和加法的运算次数。

3.3　消融实验

为了验证改进模块的有效性，对模型进行消融实验，结果如表 2 所示。

由表 2 可知，添加 SE 模块后模型的 mAP 提高 2.27 个百分点，F1 值提高 0.95 个百分点。在增加可接受参数量的前提下，有效提升了模型的性能。SE 注意力模块对不同特征赋予不同的权重，提高病斑颜色和叶片边缘等特征权重，提升检测精确率。

表 2　不同模型消融实验结果

Tab. 2　Ablation experimental results of different models

SE	L_{cls}	LME	精确率/%	召回率/%	mAP/%	F1 值/%	浮点运算次数	参数量
			90.37	85.54	87.51	87.89	6.45×10^9	5.06×10^6
√			90.55	87.19	89.78	88.84	7.48×10^9	5.07×10^6
√	√		96.92	87.48	91.95	91.96	7.48×10^9	5.07×10^6
√	√	√	95.02	92.73	93.53	93.86	7.26×10^9	5.04×10^6

注：“√”表示使用对应模块。

使用改进分类损失函数后模型的 mAP 提高 2.17 个百分点，F1 值提高 3.12 个百分点。在不增加参数量和计算量的情况下，通过赋予正负样本不同权值，有效地提高了类别的分类精确率。

添加 LME 模块后模型的 mAP 提高 1.58 个百分点，F1 值提高 1.9 个百分点。浮点运算次数减少 2.2×10^8 次，参数量减少 3×10^4，反而加强了特征提取能力。精确率降低 1.9 个百分点，召回率提高 5.25 个百分点，说明模型更好地平衡了精确率和召回率。此模型相较于基础模型的 mAP 共提高 6.02 个百分点，F1 值共提高 5.97 个百分点。

3.4　与经典网络性能对比

为了验证该模型的有效性，在相同（或相似）的条件下，选取具有代表性的单阶段目标检测网络 SSD[14]、YOLO v4[26]、EffcientDet[27] 进行对比实验，如表 3 所示。

表 3　不同测试模型性能对比

Tab. 3　Performance comparison of different test models

模型	精确率/%	召回率/%	mAP/%	F1 值/%	浮点运算次数	参数量	检测时间/ms
SSD	83.41	69.60	82.96	75.88	3.23×10^9	4.47×10^6	177.75
YOLO v4	80.54	81.70	82.83	81.11	8.10×10^9	5.92×10^6	155.53
EfficientDet	88.51	83.99	91.01	86.19	6.00×10^9	6.56×10^6	515.33
CDD	95.02	92.73	93.53	93.86	7.04×10^9	5.04×10^6	162.65

在测试阶段，统一输入图像尺寸为 640 像素 × 640 像素图像。

由表 3 可知，本文提出的 CDD 模型精确率和召回率最高，分别为 95.02% 和 92.73%，F1 值为 93.86%，比 SSD、YOLO v4、EfficentDet 分别高 17.65、12.42、7.34 个百分点。平均检测时间为 162.65 ms，略低于 YOLO v4 模型，因为解耦头使用更多的参数，检测时间增加 7.12 ms。CDD 模型在检测的综合指标上有着更大的优势，能够高效完成木薯叶病害的检测。

SSD 模型参数量和计算量均为最低，但各项评价指标均低于其他模型，说明该网络因结构问题无法提取明确的病叶特征。

YOLO v4 模型在参数量和计算量不高的情况下，拥有高效的检测速度，mAP 仅为 82.83%。该模型在分类回归层使用单一的 Head 结构，导致精确率不高。而 CDD 模型使用解耦的头部结构，提升分类和回归的精度，同时加速了网络的收敛速度。

EfficientDet 模型需要最多的参数量和计算量，平均检测精度为 91%。理论上模型结构越深、宽度越大，其特征提取能力越强，但其性能低于参数量和计算量更小的 CDD 模型，说明该模型不能很好地泛

化到本文数据集上,未充分提取有效特征信息。

3.5 CDD 模型分析

为了明确本文的 CDD 模型对不同严重性的病害类别的预测结果,对 8 种病害类别和背景使用混淆矩阵对模型的预测情况进行精度评估,如图 5 所示。行坐标表示模型预测类别,列坐标表示图像真实类别。行列相交点上的数值表示列分类被预测成行分类的比例。

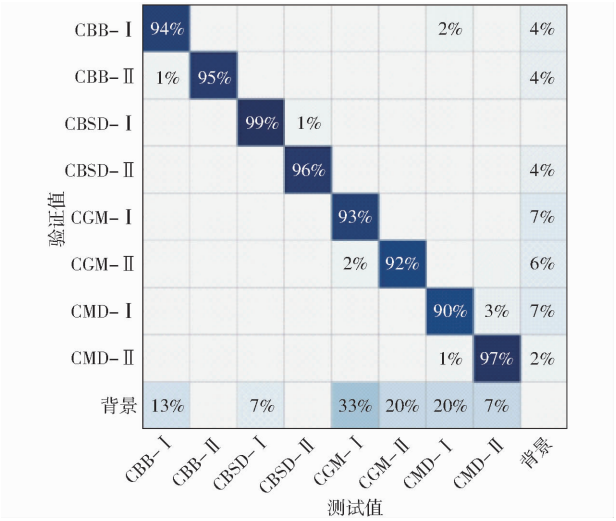

图 5 检测混淆矩阵

Fig. 5 Confusion matrix of detection

由图 5 可知,CDD 模型对患病叶片的症状有很好的区分,最高预测精确率为 99%,为 CBSD-Ⅰ类;最低的预测精确率为 90%,类别为 CMD-Ⅰ。其中,模型对 CMD-Ⅰ、CGM-Ⅰ的预测精确率较低,因为这两种轻症症状分别为黄白色小斑点和很小部分叶片扭曲缺绿,叶片病症特征不明显;或由于图像进行压缩等操作,导致模型无法学习丢失的部分小病斑特征。CGM 的平均精确率为 91%,可见模型有较小几率错误识别病害的程度,如图 6c 中轻型病被检测为重型病症。因为部分叶片感染面积接近于分级标准,较难区分轻重症。

CDD 模型的部分测试结果如图 6 所示,在复杂背景下,大部分病叶都能成功检测,但存在以下 3 种问题:①部分漏检现象(对应图 5 中最后一列)。如图 6a 中顶部的叶子未被检测出。分析原因可能为训练集叶片形状不够齐全或测试图像不包含叶片中心点则易被判别为背景。②背景被检测为病叶的现象(对应图 5 中最后一行)。原因是测试图像背景中包含有症状的叶片,但背景叶片残缺无法确认等级,如图 6b 的左上角区域。③少量病害错误检测。原因是部分症状易混淆难以区分,如图 6c 的右下角区域,CGM-Ⅱ被误检测为 CGM-Ⅰ。或因为图像原因造成识别困难,如图 6d 左上角 CMD-Ⅰ被误检测为 CBB-Ⅰ。

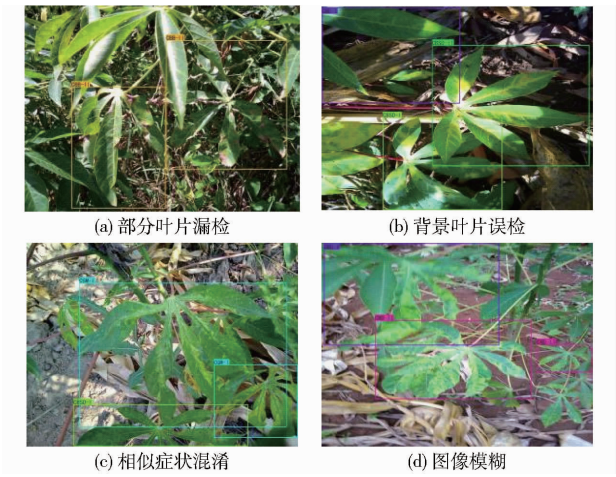

图 6 部分测试集结果

Fig. 6 Partial test set results

4 结论

(1)主干网络采用多尺度特征提取模块对木薯叶病害多尺度的特征进行提取,在减少模型计算量的情况下,更好地提取细粒度特征信息,加快检测速度;在主干网络不同层加入 SE 注意力模块,对各层输出特征进行调整,提高检测精度;在分类损失函数中结合质量焦点损失函数,在不增加参数量的情况下,提高了检测分类的精确率。

(2)CDD 模型在真实复杂背景木薯叶病害图像测试集下平均精度均值为 93.53%,F1 值为 93.86%,浮点运算次数为 7.04×10^9,网络参数量仅为 5.04×10^6,平均每幅图像的检测时间为 162.65 ms,平衡各项指标均优于 SSD、YOLO v4、EfficientDet 模型。

参 考 文 献

[1] 翟肇裕,曹益飞,徐焕良,等. 农作物病虫害识别关键技术研究综述[J]. 农业机械学报,2021,52(7):1-18.
ZHAI Zhaoyu, CAO Yifei, XU Huanliang, et al. Review of key techniques for crop disease and pest detection[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021,52(7):1-18. (in Chinese)

[2] BRAHIMI M,BOUKHALFA K,MOUSSAOUI A. Deep learning for tomato diseases: classification and symptoms visualization [J]. Applied Artificial Intelligence, 2017, 31(4): 299-315.

[3] 黄林生,罗耀武,杨小冬,等. 基于注意力机制和多尺度残差网络的农作物病害识别[J]. 农业机械学报,2021,52(10):264-271.
HUANG Linsheng, LUO Yaowu, YANG Xiaodong, et al. Crop disease recognition based on attention mechanism and multi-scale residual network[J]. Transactions of the Chinese Society for Agricultural Machinery,2021,52(10):264-271. (in Chinese)

[4] 樊湘鹏,周建平,许燕,等.基于改进卷积神经网络的复杂背景下玉米病害识别[J].农业机械学报,2021,52(3):210-217.
FAN Xiangpeng, ZHOU Jianping, XU Yan, et al. Corn disease recognition under complicated background based on improved convolutional neural network[J]. Transactions of the Chinese Society for Agricultural Machinery,2021,52(3):210-217. (in Chinese)
[5] 林猛,周刚,杨亚伟,等. 特殊天气条件下的目标检测方法综述[J].计算机工程与应用,2022,58(13):36-47.
LIN Meng, ZHOU Gang, YANG Yawei, et al. Survey of object detection methods under adverse weather conditions[J]. Computer Engineering and Application,2022,58(13):36-47. (in Chinese)
[6] 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 & Machine Intelligence, 2017, 39(6):1137-1149.
[7] HE K, GEORGIA G, DOLLAR P, et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2020,42(2):386-397.
[8] 候瑞环,杨喜旺,王智超,等.一种基于 YOLOv4-TIA 的林业害虫实时检测方法[J].计算机工程,2022,48(4):255-261.
HOU Ruihuan, YANG Xiwang, WANG Zhichao, et al. A real-time detection methods for forest pests based on YOLO V4-TIA algorithm[J]. Computer Engineering,2022,48(4):255-261. (in Chinese)
[9] 易诗,李俊杰,张鹏,等.基于特征递归融合 YOLOv4 网络模型的春见柑橘检测与计数[J].农业工程学报,2021,37(18):161-169.
YI Shi, LI Junjie, ZHANG Peng, et al. Detecting and counting of spring-see citrus using YOLOv4 network model and recursive fusion of features[J]. Transactions of the CSAE,2021,37(18):161-169. (in Chinese)
[10] 张晴晖,孔德肖,李俊萩,等.基于逆运动学降维求解与 YOLO v4 的果实采摘系统研究[J].农业机械学报,2021,52(9):15-23.
ZHANG Qinghui, KONG Dexiao, LI Junqiu, et al. Design of fruit picking system based on inverse kinematics dimension reduction and YOLO v4[J]. Transactions of the Chinese Society for Agricultural Machinery,2021,52(9):15-23. (in Chinese)
[11] 张乐,金秀,傅雷扬,等.基于 Faster R-CNN 深度网络的油菜田间杂草识别方法[J].激光与光电子学进展,2020,57(2):304-312.
ZHANG Le, JIN Xiu, FU Leiyang, et al. Recognition method for weeds in rapeseed field based on Faster R-CNN deep network[J]. Laser & Optoelectronics Progress,2020,57(2):304-312. (in Chinese)
[12] 李鑫然,李书琴,刘斌.基于改进 Faster R-CNN 的苹果叶片病害检测模型[J].计算机工程,2021,47(11):298-304.
LI Xinran, LI Shuqin, LIU Bin. Apple leaf diseases detection model based on improved Faster R-CNN[J]. Computer Engineering,2021,47(11):298-304. (in Chinese)
[13] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]// Computer Vision and Pattern Recognition. IEEE, 2016:779-788.
[14] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]// European Conference on Computer Vision, 2016: 21-37.
[15] 胡小平,户雪敏,马丽杰,等.作物病害监测预警研究进展[J].植物保护学报,2022,11(4):1-36.
HU Xiaoping, HU Xuemin, MA Lijie, et al. Research progress of crop diseases on monitoring and prediction[J]. Journal of Plant Protection,2022,11(4):1-36. (in Chinese)
[16] 刘芳,刘玉坤,林森,等.基于改进型 YOLO 的复杂环境下番茄果实快速识别方法[J]. 农业机械学报,2020,51(6):229-237.
LIU Fang, LIU Yukun, LIN Sen, et al. Fast recognition method for tomatoes under complex environments based on improved YOLO[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020,51(6):229-237. (in Chinese)
[17] 刘君,王学伟. 基于 YOLO 的番茄病虫害识别算法[J].中国瓜菜, 2020,33(9): 18-22,38.
LIU Jun, WANG Xuewei. Tomato disease and pest detection algorithm based on YOLO convolutional neural network[J]. China Cucurbits and Vegetables,2020,33(9): 18-22, 38. (in Chinese)
[18] RAMCHARAN A, MCCLOSKEY P, BARANOWSKI K, et al. A mobile-based deep learning model for cassava disease diagnosis[J]. Frontiers in Plant Science, 2019, 10(9):272.
[19] GE Z, LIU S, WANG F, et al. YOLOX: exceeding YOLO series in 2021[C]// Computer Vision and Pattern Recognition, 2021: 2107.08430.
[20] HE K, ZHANG X, REN S, et al. Identity mappings in deep residual networks[C]// European Conference on Computer Vision, 2016: 630-645.
[21] GAO S, CHENG M, ZHAO K, et al. Res2Net: a new multi-scale backbone architecture[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(2): 652-662.
[22] ALEX K, ILYA S, GEOFFREY E H. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
[23] LI Xiang, WANG Wenhai, WU Lijun, et al. Generalized focal loss: learning qualified and distributed bounding boxes for dense object detection[C]// Conference and Workshop on Neural Information Processing Systems (NeurIPS), 2020,2020.
[24] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: a metric and a loss for bounding box regression[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019.
[25] GLOROT X, BENGIO Y. Understanding the difficulty of training deep feedforward neural networks[C]// Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics. JMLR Workshop and Conference Proceedings, 2010: 249-256.
[26] BOCHKOVSKIY A, WANG C Y, LIAO H. YOLOv4: optimal speed and accuracy of object detection[C]//CVPR 2020,2020.
[27] TAN M, PANG R, LE Q V. EfficientDet: scalable and efficient object detection[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020.