

doi:10.6041/j.issn.1000-1298.2022.11.028

# SMS和双向特征融合的自然背景柑橘黄龙病检测技术

曾伟辉<sup>1,2</sup> 陈亚飞<sup>1,3</sup> 胡根生<sup>3</sup> 鲍文霞<sup>3</sup> 梁栋<sup>1,3</sup>

(1. 安徽大学互联网学院, 合肥 230039; 2. 科大国创软件股份有限公司中央研究院, 合肥 230088;

3. 安徽大学农业生态大数据分析与应用技术国家地方联合工程研究中心, 合肥 230601)

**摘要:** 柑橘黄龙病严重影响柑橘的产量和品质。在自然背景下, 柑橘叶片之间存在相互遮挡以及尺寸变化大的问题, 使得遮挡及小尺寸的黄龙病叶片容易漏检, 而且由于黄龙病叶片的颜色、纹理特征与柑橘其他病害十分相似, 容易存在误检的问题, 导致现有的算法对自然背景柑橘黄龙病检测的精度不高。本研究提出了一种结合剪切混合拼接(Shearing mixed splicing, SMS)增广算法和双向特征融合的自然背景柑橘黄龙病检测方法, 该方法通过 SMS、镜像翻转和旋转方法对训练集和验证集进行了增广, 增加了训练集和验证集图像中背景目标的数量和多样性; 为了自适应地改变柑橘黄龙病检测中的局部采样点, 增大有效感受野, 使用可变形卷积替换骨干网络后 3 个卷积层中所有的标准卷积; 为了减小自然背景的影响, 使用全局上下文模块对骨干网络后 3 个卷积层输出的特征图进行特征增强, 来建立有效的长距离依赖, 以便更好的学习到全局上下文信息; 使用双向融合特征金字塔, 改善浅层特征和深层特征的信息交流路径, 用以降低因柑橘黄龙病叶片尺寸变化大导致的漏检, 提高小尺寸的柑橘黄龙病叶片的检测精度。实验结果表明, 本研究提出的方法用于自然背景柑橘黄龙病的检测, 平均精度可达 84.8%, 性能优于 SSD、RetinaNet、YOLO v3、YOLO v5s、Faster RCNN、Cascade RCNN 等目标检测方法。

**关键词:** 柑橘; 黄龙病检测; 自然背景; 全局上下文模块; 可变形卷积; 双向特征融合

**中图分类号:** TP391.4 **文献标识码:** A **文章编号:** 1000-1298(2022)11-0280-08

**OSID:**

## Detection of Citrus Huanglongbing in Natural Background by SMS and Two-way Feature Fusion

ZENG Weihui<sup>1,2</sup> CHEN Yafei<sup>1,3</sup> HU Gensheng<sup>3</sup> BAO Wenxia<sup>3</sup> LIANG Dong<sup>1,3</sup>

(1. School of Internet, Anhui University, Hefei 230039, China

2. Central Research Institute, Guo Chuang Software Co., Ltd., University of Science and Technology of China, Hefei 230088, China

3. National Engineering Research Center for Agro-Ecological Big Data Analysis and Application, Anhui University, Hefei 230601, China)

**Abstract:** Citrus Huanglongbing is known as the “cancer” of citrus, which seriously affects the yield and quality of citrus. Therefore, accurate detection of citrus Huanglongbing is of great significance for timely protection and management of citrus. However, in the natural background, there are problems of mutual occlusion and large size changes among citrus leaves, which makes the occlusion and small-sized leaves of Huanglongbing easy to miss. In addition, because the color and texture characteristics of the leaves of Huanglongbing are very similar to other diseases of citrus, there is a problem of false detection. Therefore, when the background is complex, it is difficult for the existing algorithms to accurately detect and identify the leaves of Huanglongbing. In response to the above problems, a natural background citrus Huanglongbing detection method was proposed based on shearing mixed splicing and two-way feature fusion. The method proposed used Cascade RCNN as the baseline network and used LabelImg to manually label the Huanglongbing samples in training and validation images. Firstly, in order to reduce the impact of complex background on the detection of Huanglongbing, the training set and validation set were augmented with the shearing mixed splicing method, mirror flips and rotations, which increased the number and diversity of background objects in the training set and validation set images. Secondly,

收稿日期: 2022-07-06 修回日期: 2022-08-06

**基金项目:** 安徽省自然科学基金项目(2108085MC95)、安徽省科技重大专项(202003a06020016)、安徽省高校自然科学基金项目(KJ2020ZD03、KJ2020A0039)和农业生态大数据分析与应用技术国家地方联合工程研究中心开放项目(AE202004)

**作者简介:** 曾伟辉(1982—),女,讲师,博士,主要从事计算机视觉和模式识别研究,E-mail: whzeng@ahu.edu.cn

**通信作者:** 胡根生(1971—),男,教授,博士,主要从事农业信息处理与应用研究,E-mail: hugs2906@sina.com

deformable convolution was used to replace all standard convolutions in the backbone network Conv3 ~ Conv5 to reduce the influence of irregular leaf shape and increase the effective receptive field and adaptively change the local sampling points in the detection of citrus Huanglongbing. Thirdly, in order to reduce the influence of the natural background on the detection results of citrus Huanglongbing and enhance the ability of the backbone network to extract the detailed features of the citrus Huanglongbing disease area, the global context block was used to enhance the feature map output by Conv3 ~ Conv5 to establish an effective long-term distance dependence, so that the network can better learn the global context information. Finally, in order to reduce the influence of large changes in the size of the leaves of Huanglongbing on the detection results, two-way fusion feature pyramid networks was used to improve the information exchange path between shallow features and deep features, thereby improving the detection accuracy of small-sized blades. To verify the rationality and effectiveness of the method, in the training phase, the stochastic gradient descent strategy was adopted to train the network model. The initial learning rate was 0.02, the momentum was 0.9, the weight decay was 0.0001, and the number of iterations was 500. During the testing phase, the method proposed achieved 85.0% recall, 86.4% precision, and 84.8% average precision on the test set. The proposed method was compared with other detection algorithms (SSD, RetinaNet, YOLO v3, YOLO v5s, Faster RCNN, Cascade RCNN). Comparative experiments showed that the mean average precision of this method was 30.5 percentage points higher than that of SSD, 21.9 percentage points higher than that of RetinaNet, 13.2 percentage points higher than that of YOLO v3, 6.8 percentage points higher than that of YOLO v5s, and 20.1 percentage points higher than that of Faster RCNN, which was 3.2 percentage points higher than that of Cascade RCNN, and the detection result of this method was better than other classical deep learning methods.

**Key words:** citrus; Huanglongbing detection; natural background; global context block; deformable convolution; two-way feature fusion

## 0 引言

柑橘是世界第一大类水果,是全球最重要的经济作物之一<sup>[1]</sup>。中国是柑橘生产和消费大国,近年来,我国柑橘产业迅速发展,柑橘的种植面积不断扩大。然而,柑橘黄龙病严重影响柑橘的产量和质量,被称为柑橘癌症。我国 19 个种植柑橘的省份已有 11 个省份遭受柑橘黄龙病的危害,受灾面积已超过种植总面积的 80%<sup>[2]</sup>。因此,尽早发现受感染的病树,并采取相应的防治措施对提高柑橘产量十分重要。

柑橘黄龙病的病害特征大多是通过叶片表现出来,传统的柑橘黄龙病诊断依靠有经验的果农观察柑橘叶片进行病害判断,但这项工作不仅耗时耗力,而且容易出现人工误判。随着科学技术的飞速发展,使用新一代信息技术对柑橘黄龙病进行检测,并及时对柑橘的病害做出相应的判断和防治,提高柑橘的产量,是柑橘信息化、科学化种植的必然趋势。

目前,基于卷积神经网络的目标检测模型主要分为 One-stage 检测模型和 Two-stage 检测模型。One-stage 检测模型主要包括 SSD (Single shot multibox detector)<sup>[3]</sup>、RetinaNet<sup>[4]</sup>、EfficientDet<sup>[5]</sup>、YOLO (You only look once)<sup>[6]</sup> 系列等,已被广泛地应

用于智慧农业的植物识别、病虫害检测、水果识别等研究中<sup>[7-11]</sup>。

Two-stage 检测模型主要包括 Faster RCNN (Faster region with CNN)<sup>[12]</sup>、RCNN<sup>[13]</sup>、Cascade RCNN<sup>[14]</sup> 等。Two-stage 检测模型相较于 One-stage 检测模型具有更高的检测精度,因此它在农业方向的应用也十分广泛<sup>[15-16]</sup>。

在柑橘黄龙病的检测研究方面,已有不少学者进行了相应的研究工作<sup>[17-22]</sup>。但是直接应用到自然背景下的柑橘黄龙病检测技术仍然存在以下问题:自然场景中拍摄时,图像背景中存在大量的健康叶片、其他病害叶片、杂草等干扰因素,导致上述算法的检测精度不高;图像中经常出现黄龙病叶片被遮挡的现象,导致部分黄龙病叶片形状不规则,这对检测带来了挑战;柑橘黄龙病叶片尺寸变化大,极易出现小尺寸目标漏检的情况。

针对自然背景干扰的问题,本研究使用 SMS、镜像翻转和旋转方法对数据集进行增广,同时将全局上下文模块 (Global context block, GC-Block)<sup>[23]</sup> 嵌入到骨干网络的后 3 个卷积层中,实现自然背景下柑橘黄龙病的检测。针对黄龙病叶片形状不规则和尺寸变化大的问题,本研究采用可变形卷积<sup>[24]</sup> 和双向融合的特征金字塔网络 (Two-way fusion feature pyramid networks, TFFPN)<sup>[25]</sup>, 实现对不规则叶片和

小尺寸叶片特征的准确提取。最后,通过实验测试其对柑橘黄龙病检测的效果。

## 1 研究方法

### 1.1 数据来源及预处理

本研究数据一部分来源于科大讯飞国创软件股份有限公司举办的柑橘病虫害识别挑战赛提供的数据集,共有 683 幅柑橘叶片图像,其中 428 幅为柑橘黄龙病叶片图像,其余为柑橘健康叶片或其他柑橘病害图像,其中其他柑橘病害包括柑橘脂斑病、柑橘溃疡病、柑橘炭疽病和柑橘黑星病;另一部分来源于网络搜索到的 100 幅柑橘黄龙病叶片图像。由于比赛主办方要求数据集保密,所以本研究展示的实验图像全部来源于网络。最终本研究的数据集原图包含 528 幅图像(428 幅科大讯飞比赛图像 + 100 幅网络图像),按照比例 6:1:3 将数据集划分为训练集、验证集和测试集,其中训练集 317 幅图像,验证集 53 幅图像,测试集 158 幅图像。

由于柑橘病虫害识别挑战赛提供的 683 幅柑橘叶片图像中,有 255 幅图像为柑橘健康叶片或者其他柑橘病害图像,为了增加训练过程中检测目标的图像数量和多样性,本研究对其中不含柑橘黄龙病叶片的 255 幅图像进行了剪切混合拼接处理。剪切混合拼接是将病害图像中的部分柑橘黄龙病叶片剪切下来,再和健康的或其他病害的柑橘叶片图像拼接起来,来增加训练集图像中检测目标的数量和多样性。经过剪切混合拼接后,训练集和验证集的图像分别扩充为 536 幅和 89 幅。为了进一步增加训练数据的多样性和数量,本文对训练集和测试集进行了进一步的扩充。增广方式包括镜像翻转、顺时针旋转  $180^\circ$ 、顺时针旋转  $90^\circ$ 、顺时针旋转  $270^\circ$ ,图像增广流程图如图 1 所示。经过增广后,训练集共有 2 680 幅图像、验证集共有 445 幅图像,增广前后训练集、验证集和测试集的图像数量和黄龙病叶片数量如表 1 所示。本研究使用 LabelImg 对该数据集进行标注,生成 .xml 文件,该文件中包含柑橘黄龙病病害叶片在图像中的位置信息。

### 1.2 柑橘黄龙病检测网络

#### 1.2.1 改进的 Cascade RCNN 模型

为了能够实现自然背景的柑橘黄龙病准确检测,本研究提出一种改进的 Cascade RCNN 网络。该网络由 3 部分组成:骨干网络、双向融合的特征金字塔、级联检测器。改进后的 Cascade RCNN 结构如图 2 所示,图像先输入到改进后的骨干网络进行特征提取,随后将 Conv2 输出的特征图和经过 GC-Block 增强后的特征图一同输入到双

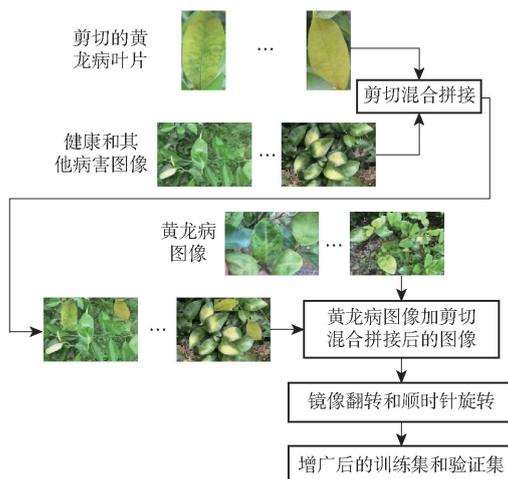


图 1 图像增广流程图

Fig. 1 Flowchart of image augmentation

表 1 增广前后的数据集

Tab. 1 Data sets before and after augmentation

数据集	训练集		验证集		测试集	
	图像数量/幅	黄龙病叶片数	图像数量/幅	黄龙病叶片数	图像数量/幅	黄龙病叶片数
原始数据集	317	1 017	53	146	158	539
剪切混合拼接	536	1 652	89	239	158	539
旋转和镜像	2 680	8 260	445	1 195	158	539

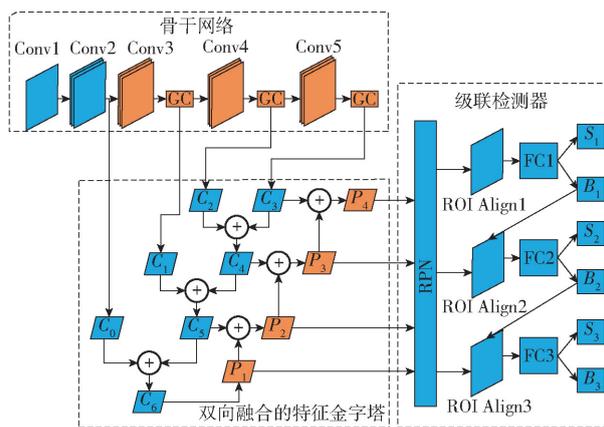


图 2 改进的 Cascade RCNN 结构图

Fig. 2 Improved Cascade RCNN structure diagram

向融合的特征金字塔中进行特征图的多尺度融合,最后将融合后的特征图输入到级联检测器中,得到最终的检测结果。

图 2 中蓝色模块为原网络,橙色模块为改进部分。Conv1 ~ Conv5 表示骨干网络 ResNetXt101<sup>[26]</sup> 的 5 个卷积层,其中蓝色卷积层是由标准卷积组成的卷积层,橙色卷积层是由可变形卷积组成的卷积层。 $C_0$  表示 Conv2 输出的特征图经  $1 \times 1$  卷积后得到的特征图, $C_1$ 、 $C_2$ 、 $C_3$  表示 Conv3、Conv4、Conv5 输出的特征图依次输入到全局上下文模块后再经  $1 \times 1$  卷积得到的特征图, $C_4$ 、 $C_5$ 、 $C_6$  表示单向融合后的特征

图,  $P_1, P_2, P_3, P_4$  表示双向融合后的特征图。RPN 为区域建议网络, ROI Align1、ROI Align2、ROI Align3 为 3 个不同 IoU 阈值的感兴趣区域对齐 (Region of interest align, ROI Align)<sup>[27]</sup>; FC1、FC2、FC3 表示 3 个全连接层;  $S_1, S_2$  和  $B_1, B_2$  分别表示两个阶段的分类分数和候选框;  $S_3$  和  $B_3$  分别表示最终的分类分数和边界框。

1.2.2 骨干网络

本研究将 Cascade RCNN 中的骨干网络

ResNet101<sup>[28]</sup> 替换为 ResNetXt101, ResNetXt101 和 ResNet101 的网络结构图如表 2 所示, 括号内是残差结构, 乘号后的数字表示堆叠块数,  $c$  表示每组的卷积数。ResNetXt101 与 ResNet101 的本质区别是 ResNetXt101 采用了通道分组, 以 Conv2 为例, ResNet101 是 64 个  $1 \times 1$  的卷积核, 而 ResNetXt101 是 32 组, 每组是 4 个  $1 \times 1$  的卷积核。与 ResNet101 相比, ResNetXt101 的网络结构简明, 更加模块化; 需要手动调节的超参数少; 具有更好的特征提取能力。

表 2 ResNet101 和 ResNetXt101 的网络结构

Tab.2 Network structure of ResNet101 and ResNetXt101

阶段	ResNet101	ResNetXt101
Conv1	$7 \times 7, 64$ , 步长为 2, $3 \times 3$ , maxpool, 步长为 2	$7 \times 7, 64$ , 步长为 2, $3 \times 3$ , maxpool, 步长为 2
Conv2	$(1 \times 1, 64 \ 3 \times 3, 64 \ 1 \times 1, 256) \times 3$	$(1 \times 1, 128 \ 3 \times 3, 128 \ 1 \times 1, 256, c = 32) \times 3$
Conv3	$(1 \times 1, 128 \ 3 \times 3, 128 \ 1 \times 1, 512) \times 4$	$(1 \times 1, 256 \ 3 \times 3, 256 \ 1 \times 1, 256, c = 32) \times 4$
Conv4	$(1 \times 1, 512 \ 3 \times 3, 512 \ 1 \times 1, 1024) \times 23$	$(1 \times 1, 512 \ 3 \times 3, 512 \ 1 \times 1, 1024, c = 32) \times 23$
Conv5	$(1 \times 1, 512 \ 3 \times 3, 512 \ 1 \times 1, 2048) \times 3$	$(1 \times 1, 1024 \ 3 \times 3, 1024 \ 1 \times 1, 2048, c = 32) \times 3$

1.2.3 可变形卷积

为了减少叶片形状不规则的影响, 本研究用可变形卷积代替骨干网络的后 3 个卷积层中的标准卷积, 根据黄龙病叶片的形状自适应改变局部采样点, 增强网络对目标几何形变的建模能力。

可变形卷积的基本思想是用带有偏移的采样来代替固定位置采样, 并根据学习到的偏移量, 来调整可变形卷积核的大小和位置, 不同位置卷积核的采样点会根据检测目标的位置和大小自适应调整, 从而适应柑橘黄龙病叶片的几何形变。可变形卷积的结构如图 3 所示, 在原始特征图中, 先通过一个卷积层来计算特征图的 2 维偏移量, 得到偏移量, 然后通过双线性插值算法将带有偏移量的采样点集中到一起, 使得输出特征图与输入特征图的空间分辨率一致, 再将输出特征图传递给下一网络层, 从而实现形状不规则目标的特征提取。

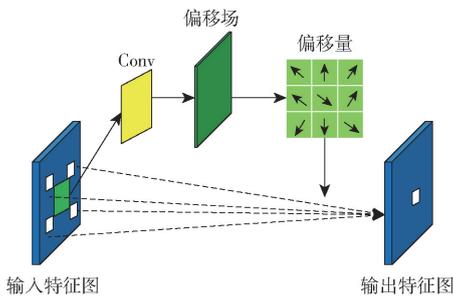


图 3 可变形卷积结构示意图

Fig.3 Structure diagram of deformable convolution

1.2.4 全局上下文模块

为了减小背景对检测结果的干扰, 本研究将 GC-Block 嵌入到骨干网络的 Conv3、Conv4、Conv5

中, 构建有效的长距离依赖, 增强骨干网络的全局上下文建模能力。GC-Block 结合了非局部模块 (Non-local block, NL-Block)<sup>[29]</sup> 和挤压激励模块 (Squeeze excitation block, SE-Block)<sup>[30]</sup> 的优点, 既具有 NL-Block 的全局上下文建模能力, 又具有 SE-Block 轻量级的优点。

GC-Block 可以抽象为 3 个步骤: 首先通过  $1 \times 1$  卷积和 Softmax 函数获取特征图的注意力权重, 再通过矩阵乘法获取全局上下文特征; 然后通过  $1 \times 1$  卷积进行特征优化, 从而捕获通道间的依赖, 为了降低优化难度, 在  $1 \times 1$  卷积后面增加了归一化层 (Layer normalization, LN) 和线性整流函数 (Rectified linear unit, ReLU); 最后采用对应元素相加的方法将全局上下文特征聚合到每个位置的特征上, 从而实现对原始特征图的增强。GC-Block 结构如图 4 所示, 其中  $H, W$  分别表示特征图的高和宽,  $C$  表示特征图的通道数,  $C \times HW$  是原始特征图经过线性变换后的二维矩阵,  $r$  表示瓶颈比率。

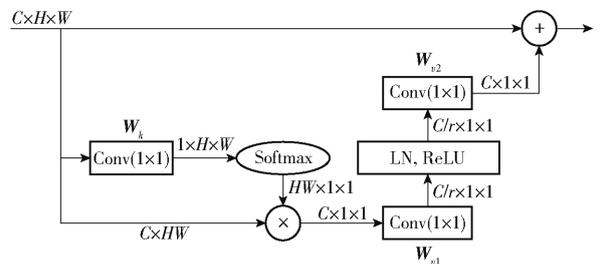


图 4 GC-Block 结构图

Fig.4 Structure diagram of GC-Block

GC-Block 定义为

$$Z_i = x_i + W_{i2} \text{ReLU} \left( \text{LN} \left( W_{i1} \sum_{j=1}^{N_p} \frac{e^{W_k x_j}}{\sum_{m=1}^{N_p} e^{W_k x_m}} \right) \right) \quad (1)$$

式中  $Z_i$ ——查询位置输出矩阵  
 $x_i$ ——查询位置输入矩阵  
 $i$ ——查询位置索引  
 $j$ ——其他位置索引  
 $x_j$ ——其他位置输入矩阵  
 $x_m$ ——注意力位置输入矩阵  
 $W_{v1}$ 、 $W_{i2}$ 、 $W_k$ ——线性变换矩阵  
 $N_p$ ——特征图中的位置数  
 $m$ ——注意力位置索引

1.2.5 双向融合的特征金字塔

为了解决柑橘黄龙病叶片的小尺寸目标的漏检问题,本研究在 Cascade RCNN 的基础上,使用 TFFPN 替换原有的特征金字塔网络 (Feature pyramid networks, FPN)<sup>[31]</sup>。从图 5 中可以看到, TFFPN 与 FPN 不同, FPN 是一个自上而下的单向融合,而 TFFPN 是先由上而下再由下而上的双向融合。 $C_0$  表示 Conv2 输出的特征图经  $1 \times 1$  卷积后得到的特征图,  $C_1$ 、 $C_2$ 、 $C_3$  表示 Conv3、Conv4、Conv5 输出的特征图依次输入到全局上下文模块后再作  $1 \times 1$  卷积得到的特征图,这里  $1 \times 1$  卷积的作用是用来改变特征图的通道数。随后,再对  $C_3$  进行 2 倍上采样,来增大  $C_3$  的分辨率,使其分辨率与  $C_2$  保持一致,这样深层特征图和浅层特征图就可以通过简单地加法实现融合,得到  $C_4$ ,再通过相同的操作得到特征图  $C_5$ 、 $C_6$ 。然后对单向融合后的特征图作一个步长为 2 的  $3 \times 3$  卷积,来降低  $P_1$  的分辨率,使其分辨率与  $C_5$  保持一致,随后对  $C_5$  进行  $3 \times 3$  卷积来消除上采样产生的混叠效应,

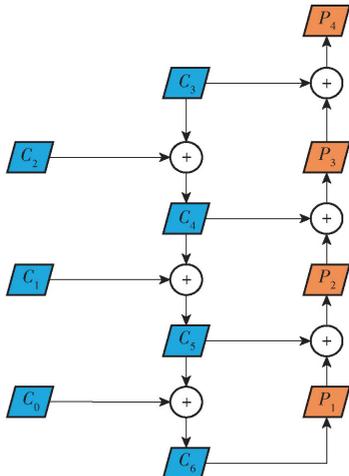


图 5 TFFPN 结构图

Fig. 5 Structure diagram of TFFPN

然后通过简单的加法实现  $P_1$  与  $C_5$  的融合,再对融合后的特征图作一个  $3 \times 3$  卷积得到特征图  $P_2$ ,同理得  $P_3$  和  $P_4$ 。TFFPN 不仅能够丰富每层特征图的位置信息,还能够保留每层特征图的语义信息,增强了模型的多尺度信息融合能力,有利于小尺寸目标的检测。

1.2.6 级联检测器

级联检测器是由 3 个具有不同交并比 (Intersection over union, IoU) 的 ROI Align 级联组成,其中 ROI Align1、ROI Align2、ROI Align3 的 IoU 阈值是递增的,可以使每一阶段为下一阶段提供更好的正样本,从而逐渐获得更好的检测效果。

级联检测器结构如图 6 所示,将 TFFPN 得到的特征图  $P_1$ 、 $P_2$ 、 $P_3$ 、 $P_4$  一同输入到 RPN 中,得到候选框,然后将候选框一同输入到 ROI Align1 中,再将 ROI Align1 输出的结果输入到全连接层 FC1 中,得到阶段一的分类分数  $S_1$  和候选框  $B_1$ 。将特征图和  $B_1$  一同输入到 ROI Align2 中,再将 ROI Align2 输出的结果输入到 FC2 中,得到阶段二的分类分数  $S_2$  和候选框  $B_2$ 。最后将特征图和  $B_2$  一同输入到 ROI Align3 中,再将 ROI Align3 输出的结果输入到 FC3 中,得到最终的分类分数  $S_3$  和边界框  $B_3$ 。

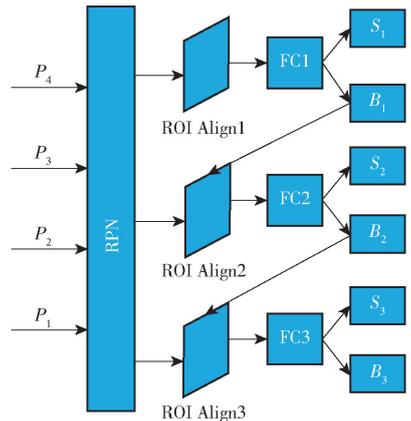


图 6 级联检测器结构示意图

Fig. 6 Cascade detector schematic

2 结果与分析

2.1 参数设置与评价指标

柑橘黄龙病检测网络用 Pytorch 框架搭建,模型优化使用随机梯度下降法 (Stochastic gradient descent,SGD),初始学习率为 0.02,动量为 0.9,权重衰减为 0.0001,迭代次数为 500。

采用平均精度均值 (Mean average precision, mAP)、精确率 (Precision,  $P$ )、召回率 (Recall,  $R$ ) 作为评价指标,mAP 越高,说明检测效果越好。

2.2 实验结果分析

选取 One-stage 的 SSD、RetinaNet、YOLO v3、

YOLO v5s 和 Two-stage 的 Faster RCNN、Cascade RCNN 与本研究方法进行对比实验。从表 3 可以看出,图像没有经过增广处理时,本文方法的 mAP、 $P$  和  $R$  均高于其它方法,mAP 达 82.5%,与

RetinaNet、YOLO v5s 和 Cascade RCNN (ResNetXt101)相比,mAP 分别提高 22.1、5.9、5.4 个百分点。

为了降低柑橘其他病害叶片和健康叶片对检测

表 3 不同检测算法的对比结果

Tab. 3 Comparison results of different detection algorithms

算法	无增广处理			增广处理			内存占用 量/MB	平均测试 时间/s
	$R$ /%	$P$ /%	mAP/%	$R$ /%	$P$ /%	mAP/%		
SSD	48.2	62.8	51.7	50.8	65.9	54.3	249.5	0.197
RetinaNet	61.0	60.6	60.4	63.1	62.6	62.9	276.7	0.215
YOLO v3	63.6	70.9	69.3	64.7	72.3	71.6	496.1	0.146
YOLO v5s	75.1	77.6	76.6	75.7	78.3	78.0	14.4	0.048
Faster RCNN	65.1	63.1	61.8	65.3	62.3	64.7	315.0	0.253
Cascade RCNN(ResNet101)	77.6	78.3	77.1	78.8	79.4	79.2	672.3	0.366
Cascade RCNN(ResNetXt101)	78.5	80.7	79.1	80.0	82.1	81.6	669.7	0.351
本文方法	83.3	84.9	82.5	85.0	86.4	84.8	787.4	0.485

结果的影响,本研究通过 SMS、镜像翻转和旋转的方法对训练集和验证集进行了增广,增加了训练集和验证集图像中背景目标的数量。从表 3 可以看出,图像经过增广后,不同检测方法的检测精度都有所提升。

从表 3 可以看出,本文方法 mAP 达到 84.8%,高于其他方法。但随着可变形卷积、GC-Block 和 TFFPN 的加入,本文方法的模型大小也略大于其他方法,达到 787.4 MB,本文方法的平均测试时间相应的有所增加。本文方法的平均测试时间为 0.485 s,虽然略高于其他目标检测方法,但与其他检测方法的平均测试时间仍处于同一数量级。相较于表 3 中其他对比方法,本文方法在保证高检测精度的同时,兼顾了检测的实时性,具有较好的性能。

为了说明本文方法适用于自然背景下的柑橘黄龙病检测,本研究根据图像背景复杂度的不同,将测试集 158 幅图像划分成 5 个级别。1 级:仅含一片柑橘黄龙病叶片,共 17 幅;2 级:含有 2 片柑橘黄龙病叶片,共 23 幅;3 级:含有 3 片柑橘黄龙病叶片,共 18 幅;4 级:含有 3 片柑橘黄龙病叶片,并存在遮挡现象,共 35 幅;5 级:含有 3 片以上的柑橘黄龙病叶片,并存在遮挡现象,共 65 幅。不同级别的 mAP 如图 7 所示。从图 7 可以看出,当图像背景的复杂度为 5 级时,本文方法相较于 Cascade RCNN,mAP 提高 5.62 个百分点。

### 2.3 消融实验

消融实验结果如表 4 所示,在骨干网络为 ResNetXt101 的前提下,本研究利用可变形卷积增大有效感受野,使用 GC-Block 进行特征增强,并使用 TFFPN 来减少小尺寸目标的漏检,提高了模型的召回率、精确率和平均精度均值。

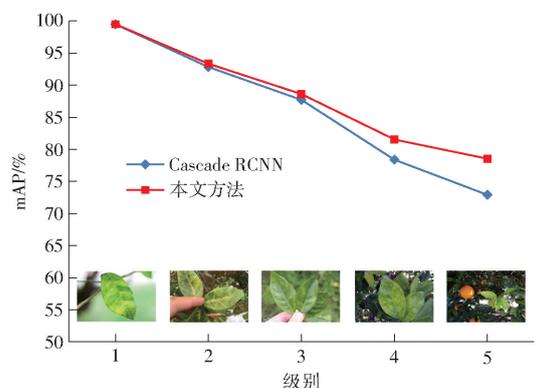


图 7 Cascade RCNN 和本文方法的 mAP 对比

Fig. 7 Comparison of mAP between Cascade RCNN and the proposed method

表 4 消融实验结果

Tab. 4 Results of ablation experiments %

网络	$R$	$P$	mAP
Cascade RCNN	80.0	82.1	81.6
Cascade RCNN + TFFPN	82.6	82.7	82.5
Cascade RCNN + GC-Block	81.1	84.0	82.7
Cascade RCNN + TFFPN + GC-Block	83.3	84.5	84.1
Cascade RCNN + 可变形卷积	80.9	83.7	82.5
Cascade RCNN + TFFPN + 可变形卷积	83.9	84.6	83.9
Cascade RCNN + GC-Block + 可变形卷积	81.6	85.4	83.3
Cascade RCNN + TFFPN + GC-Block + 可变形卷积	85.0	86.4	84.8

由 2.2 节和表 3 可知,对训练集和验证集进行增广后,增加了训练样本的多样性,在一定程度上提高了网络检测效果。在上述条件下,本研究分别在 Cascade RCNN 网络中加入可变形卷积和 GC-Block,显著提高了精确率,这是因为黄龙病叶片存在形状不规则和背景复杂的问题,可变形卷积和 GC-Block 可以有效抑制它们的影响。最后本研究将 FPN 换成了 TFFPN,对召回率提升明显,这是因

为 TFFPN 可以有效提取特征图的多尺度信息,从而降低了小尺度叶片的漏检率。可变形卷积、GC-Block 和 TFFPN 一起作用于网络时,mAP 提高 3.2 个百分点。

Cascade RCNN(ResNetXt101)和改进的 Cascade RCNN 的实际检测效果对比如图 8 所示。图中红框为模型预测的边框,红框上方的 HLB 为黄龙病的标签名,HLB 后的数字为置信度,蓝圈内为 Cascade RCNN 漏检或误检的目标。由图 8 可知,Cascade RCNN(ResNetXt101)在黄龙病的检测中存在漏检和误检现象,本研究使用了可变形卷积、GC-Block 和 TFFPN 对模型进行了改进,改进后的 Cascade RCNN

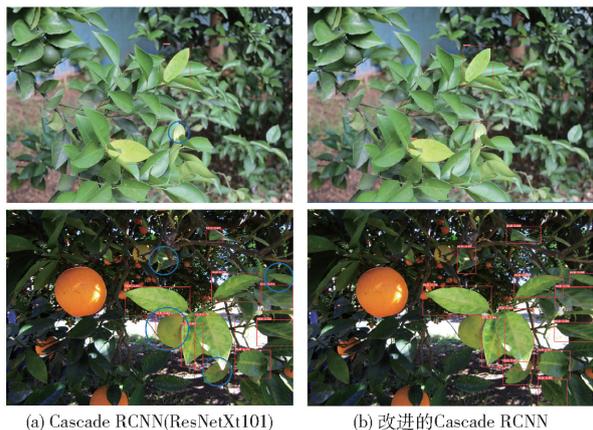


图 8 检测效果对比

Fig. 8 Comparison charts of detection effect

能够正确检测出 Cascade RCNN(ResNetXt101)漏检和误检的目标,这说明改进后的网络具有更强的病害特征提取能力。

### 3 结论

(1) 由于野外拍摄的柑橘黄龙病图像背景复杂,黄龙病叶片存在遮挡及形状尺寸变化大的问题,导致现有方法很难获得较高的检测精度,本研究提出了一种改进的 Cascade RCNN 模型用于自然场景下的柑橘黄龙病检测。该方法首先运用 SMS、镜像翻转和旋转来对数据集进行了增广,增加了训练图像中的背景目标数量和多样性。其次为了增强网络对目标几何形变的建模能力,在骨干网络中使用可变形卷积来增大有效感受野,再通过 GC-Block 对特征图进行增强,增强骨干网络对病害叶片特征的提取能力。最后通过 TFFPN 实现特征图的多尺度融合,以便网络学习到多尺度信息,提升模型的检测精度。

(2) 测试结果表明,该模型 mAP 为 84.8%,相较于 SSD、RetinaNet、YOLO v3、YOLO v5s、Faster RCNN、Cascade RCNN 等目标检测方法提高至少 3.2 个百分点。本文模型可以实现自然背景下柑橘黄龙病的精准检测,对实现农业智能化、降低劳动强度具有重要意义。

### 参 考 文 献

- [1] 郭文武,叶俊丽,邓秀新. 新中国果树科学研究 70 年——柑橘[J]. 果树学报, 2019, 36(10): 1264-1272. GUO Wenwu, YE Junli, DENG Xiuxin. Fruit scientific research in new China in the past 70 years: citrus[J]. Journal of Fruit Trees, 2019, 36(10): 1264-1272. (in Chinese)
- [2] 陈凯男,李充璧. 柑橘黄龙病研究概况[J]. 浙江农业科学, 2015, 56(7): 1048-1050. CHEN Kainan, LI Chongbi. Overview of research on citrus Huanglongbing[J]. Zhejiang Agricultural Science, 2015, 56(7): 1048-1050. (in Chinese)
- [3] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//European Conference on Computer Vision, Springer, Cham, 2016: 21-37.
- [4] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017: 2980-2988.
- [5] TAN M, LE Q. Efficientnet: rethinking model scaling for convolutional neural networks[C]//International Conference on Machine Learning. PMLR, 2019: 6105-6114.
- [6] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 7263-7271.
- [7] LIU C, ZHU H, GUO W, et al. EFDet: an efficient detection method for cucumber disease under natural complex environments[J]. Computers and Electronics in Agriculture, 2021, 189: 106378.
- [8] LI X, PAN J, XIE F, et al. Fast and accurate green pepper detection in complex backgrounds via an improved YOLOv4-tiny model[J]. Computers and Electronics in Agriculture, 2021, 191: 106503.
- [9] SUN H, XU H, LIU B, et al. MEAN-SSD: a novel real-time detector for apple leaf diseases using improved light-weight convolutional neural networks[J]. Computers and Electronics in Agriculture, 2021, 189: 106379.
- [10] TIAN Y, YANG G, WANG Z, et al. Apple detection during different growth stages in orchards using the improved YOLO-V3 model[J]. Computers and Electronics in Agriculture, 2019, 157: 417-426.
- [11] LYU Z, JIN H, ZHEN T, et al. Small object recognition algorithm of grain pests based on SSD feature fusion[J]. IEEE Access, 2021, 9: 43202-43213.

- [12] REN S, HE K, GIRSHICK R, et al. Faster R - CNN: towards real-time object detection with region proposal networks[J]. *Advances in Neural Information Processing Systems*, 2015, 28: 91 - 99.
- [13] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014: 580 - 587.
- [14] CAI Z, VASCONCELOS N. Cascade R - CNN: delving into high quality object detection[C]//*2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018: 6154 - 6162.
- [15] GAO F, FU L, ZHANG X, et al. Multi-class fruit-on-plant detection for apple in SNAP system using Faster R - CNN[J]. *Computers and Electronics in Agriculture*, 2020, 176: 105634.
- [16] JIAO L, DONG S, ZHANG S, et al. AF - RCNN: an anchor-free convolutional neural network for multi-categories agricultural pest detection[J]. *Computers and Electronics in Agriculture*, 2020, 174:105522.
- [17] SHARIF M, KHAN M A, IQBAL Z, et al. Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection[J]. *Computers and Electronics in Agriculture*, 2018, 150: 220 - 234.
- [18] 兰玉彬, 朱梓豪, 邓小玲, 等. 基于无人机高光谱遥感的柑橘黄龙病植株的监测与分类[J]. *农业工程学报*, 2019, 35(3):92 - 100.  
LAN Yubin, ZHU Zihao, DENG Xiaoling, et al. Monitoring and classification of citrus Huanglongbing plants based on UAV hyperspectral remote sensing[J]. *Transactions of the CSAE*, 2019,35(3):92 - 100. (in Chinese)
- [19] 刘燕德, 肖怀春, 孙旭东, 等. 基于高光谱成像的柑橘黄龙病无损检测[J]. *农业机械学报*, 2016,47(11):231 - 238.  
LIU Yande, XIAO Huaichun, SUN Xudong, et al. Non-destructive detection of citrus Huanglong disease using hyperspectral image technique[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2016,47(11):231 - 238. (in Chinese)
- [20] BARMAN U, CHOUDHURY R D, SAHU D, et al. Comparison of convolution neural networks for smartphone image based real time classification of citrus leaf disease[J]. *Computers and Electronics in Agriculture*, 2020, 177: 105661.
- [21] 梅慧兰, 邓小玲, 洪添胜, 等. 柑橘黄龙病高光谱早期鉴别及病情分级[J]. *农业工程学报*, 2014,30(9):140 - 147.  
MEI Huilan, DENG Xiaoling, HONG Tiansheng, et al. Early identification and disease grading of citrus Huanglongbing by hyperspectral[J]. *Transactions of the CSAE*, 2014,30(9):140 - 147. (in Chinese)
- [22] 戴泽翰, 郑正, 黄莉舒, 等. 基于深度卷积神经网络的柑橘黄龙病症状识别[J]. *华南农业大学学报*, 2020,41(4):111 - 119.  
DAI Zehan, ZHENG Zheng, HUANG Lishu, et al. Symptom recognition of citrus Huanglongbing disease based on deep convolutional neural networks[J]. *Journal of South China Agricultural University*, 2020,41(4):111 - 119. (in Chinese)
- [23] CAO Y, XU J, LIN S, et al. Genet: non-local networks meet squeeze-excitation networks and beyond[C]//*Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, 2019: 1971 - 1980.
- [24] DAI J, QI H, XIONG Y, et al. Deformable convolutional networks[C]//*2017 IEEE International Conference on Computer Vision (ICCV)*, 2017:764 - 773.
- [25] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]//*2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018:8759 - 8768.
- [26] XIE S, GIRSHICK R, DOLLÁR P, et al. Aggregated residual transformations for deep neural networks[C]//*2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017: 5987 - 5995.
- [27] HE K, GKIOXARI G, DOLLAR P, et al. Mask R - CNN[C]//*Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2017: 2980 - 2988.
- [28] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//*2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2016: 770 - 778.
- [29] WANG X, GIRSHICK R, GUPTA A, et al. Non-local neural networks[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 7794 - 7803.
- [30] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 7132 - 7141.
- [31] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//*2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017: 936 - 944.