

# 基于 Faster R-CNN 的小麦病虫害识别应用

王思平 李丹

四川大学锦城学院计算机与软件学院 四川 成都 611731

**【摘要】**本文探究了 Faster R-CNN 系列网络对复杂背景小麦病虫害数据集的识别能力,对比了 ResNet50-Faster R-CNN、ResNet101-Faster R-CNN、ResNet101-Faster R-CNN + FPN 三种网络对复杂背景小麦病虫害数据集在上的检测效果。实验结果三种网络的 Average Precision 均值可以达到 0.82 以上。ResNet101-Faster R-CNN + FPN 网络 (Average Precision 达到 0.91) 是三种训练网络中对小麦病虫害数据集检测效果最优的网络结构。实验评估表明:两阶段检测 ResNet-Faster R-CNN 在小麦病虫害识别当中有一定应用能力。

**【关键词】**Faster R-CNN; 目标检测; 农业病虫害识别

随着计算机发展,人工智能发展前景广阔,模型,算力,数据集已经成三大核心要素。对于人工智能的应用,业界正在不断发掘,从最普遍的人脸识别到现在的工业,农业,金融业等, AI 将在一定程度上助力企业转型升级,提高生产效率和质量。

中国是一个农业大国,2019 年农作物达到总播种面积 (千公顷)165930.66<sup>[1]</sup>,粮食作物播种面积 (千公顷)116063.60,其中小麦种植面积达 23730.00(千公顷)。据统计,近几年我国农作物病虫害呈多发重发态势,小麦、水稻等主要作物的病虫害发生面积在全国超过千万公顷。

中国目前的农业病虫害防治主要还是依托人工巡检,生物防治,喷洒农药,有部分地区开始引入了无人机大规模喷洒农业。随着计算机技术的发展,研究人员将计算机视觉相关技术应用在农作物病虫害识别中。王克如<sup>[2]</sup>使用数字图像处理技术对进行玉米叶片等病虫害图像进行去噪和其他操作,并且应用多个纹理特征,对有病斑叶片检测的正确率达 100%。任东等人<sup>[3]</sup>利用支持向量机算法对温室甜瓜病虫害图像进行分类最高准确率可达 96.7%。Bravo C<sup>[4]</sup>等人利用健康和小麦黄锈病感染后光谱反射率的差异,建立了基于光谱技术分类模型。温芝元<sup>[5]</sup>等人利用改进型分水岭算法提取桉柑病虫害的图像识别特征进行桉柑病虫害识别,5 种病虫害的平均识别正确率为 87%。随着人工智能发展,复杂的神经网络结构开始应用,姜敏<sup>[6]</sup>比较 FasterR-CNN Inception v2 和 SSD MobileNet v1 在不同预处理方式下(截取感兴趣区域,归一化尺寸,病理分割)对病虫害的检测效果。李想<sup>[7]</sup>将 SSD512 轻量级目标检测模型用于森林中红脂大小蠹进行病虫害检测。从总体来看,在计算

机视觉中,对农作物病虫害检测大多数采用的是收集病虫害叶片并且在相对纯净背景下重新制作病虫害数据集,实验室和近距离手机拍照识别可以达到很好的检测效果。但是生产环境中远距离相机拍摄由于光线干扰,背景复杂等因素识别效果不理想。本文将 FasterR-CNN 网络做为小麦病虫害检测的基础网络,探究 FasterR-CNN 系列网络在复杂背景环境下对小麦病虫害数据集的检测效果。

## 1 数据集

由于国内外开源的植物目标检测数据较少,所以从网上公开爬虫和 kaggle 中收集了一些小麦病虫害的数据集。数据集中一共有三个类别包括小麦白粉病,小麦锈病,小麦蚜虫。数据集中 70% 图片是田间拍摄图片,30% 图片是收集了小麦病虫害叶片在纯色背景下拍摄而成。实验中对数据进行了随机旋转,随机裁剪,随机对比度等数据增强处理,然后用 labelme 工具标注感兴趣的区域(病虫害特征部位),并且转换成 coco 数据集格式。

All/class	Allclass	powdery_mildew	rust	aphid
Train	290	100	90	100
Test	45	20	15	10

表 1 数据集种类表

## 2 目标检测网络模型

FasterR-CNN<sup>[8]</sup>是学者 Shaoqing Ren 在 2016 年正式提出的两阶段目标检测网络,相对于以前的 RCNN 网络有着较大的改进。FasterR-CNN 网络相对于 Fast-CNN 主要体现在前置卷积共享, rpn 网络, ROIpooling, 多分类和 anchors 二次回归修正。

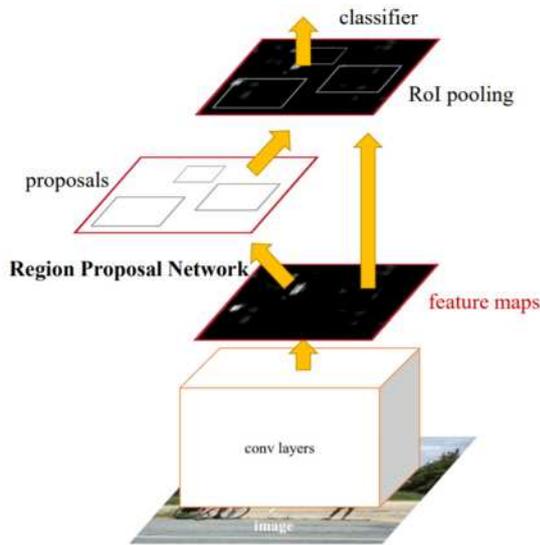


图 1 FasterR-CNN 网络结构图 [8]

### 2.1 前置卷积共享

FasterR-CNN 网络中将特征提取网络学习的内容传递给 rpn 网络，极大地减少了目标检测的时间。FasterR-CNN 由于骨干网络和 rpn 属于两个阶段，所以骨干网络 vgg16 可以替换成很多分类网络如 ResNet 模型，DarkNet 模型，MobileNet 模型。

### 2.2 rpn 网络

rpn 网络首先在 feature map 基础上生成了  $h*w*9$  个默认大小的候选框。Anchors 是在训练前通过对目标框进行聚类计算得到的。rpn 网络主要分成两条线路：

① 对特征提取网络 (feature map) 进行边框回归：边框回归微调使得默认大小的检测框通过一定函数变换后接近真实框。令  $P_x, P_y, P_w, P_h$  是默认检测框生成的 anchors,  $G_x, G_y, G_w, G_h$  是真实标注的 anchors,  $G_x, G_y, G_w, G_h$  是边框微调后的结果。通过一定变换  $F(P_x, P_y, P_w, P_h) = (G_x, G_y, G_w, G_h) \approx (G_x, G_y, G_w, G_h)$ ，rpn 框回归网络的策略是在边框修正过程中是先平移后放缩修正过程：

$$t_x = (G_x - P_x) / P_w \quad t_w = (G_w - P_w) / P_w$$

$$t_y = \log(G_y / P_y) \quad t_h = \log(G_h / P_h)$$

②对默认生成的候选 anchors 进行二分类：通过 cnn 将预生成的 anchors 二分类成有目标的 positive anchors 以及无目标的 anchors。

### 2.3 POIPooling 网络

POIPooling 的输入有两个，一个是 rpn 网络产生的 proposal，另一个是 feature map。feature map 通过前置主干网络 (vgg16) 缩小为原来的 1/16，POIPooling 池化首先将 feature map 放大映射成 resized 时的尺寸，然后将其划

分为  $7*7$  的网格，并且每个网格进行平均池化或最大池化。理论上 POIPooling 网络可以接受任意大小尺寸的图片输入。

## 3 ResNet 网络模型

ResNet<sup>[9]</sup> 是何凯明等学者在 2015 年提出的图像分类网络模型，其中最主要的思想是引入了残差神经网络，在一定程度上缓解了训练时神经网络层数过深容易梯度消失得情况。设  $D(x)$  是经过堆积块学习到的特征，残差学习特征表示  $F(x)=D(x)+x$ ，在好的情况下经过堆积块学到的特征等于恒等映射网络特征和残差网络学习到的特征，即使在最差的情况下  $F(x)=x$ ，堆积块网络也是做了恒等网络映射。通过残差神经网络的设计可以极大程度的提高网络训练的深度，让神经网络有着更多的表现特征。

## 4 fpn 网络

fpn<sup>[10]</sup> 网络即特征映射金字塔。在传统的目标检测中大多数算法都是在顶层做直接预测，顶层做直接预测的特点在于特征语义表达信息较多，目标位置比较粗略。这种方式容易忽略检测中的小目标即漏检，使得模型在有较多小目标的数据集上表现效果不理想。fpn 主要分为两条路线：一个自底向上即正常的 cnn 前向传播计算，随着不断地卷积运算，图像的尺寸不断地缩小，在缩小过程中，存在部分卷积前后图像大小尺寸不变的时候，作者选取了几个这样的阶段作为金字塔特征图。另一条路线是自下向上横向连接，经过正常前向传播得到的高特征图被上采样，高特征图的尺寸扩大并且和前向传播得到的 feature map 融合。

## 5 FasterR-CNN 网络评估标准

### 5.1 混淆矩阵

表 2 混淆矩阵表

		True class	
		T	F
Pre_class	P	True positives	False positives
	N	True negatives	False negatives

① True positives: 正样本被正确识别的个数；② True negative: 负样本被正确识别的个数；③ False positive: 负样本被错误识别的个数；④ False negatives: 正样本被错误识别的个数。

Precision(精确率) =  $TP / (TP+FP)$  表示正确识别类别的个数占总预测类别的个数的比例。

Recall(召回率) =  $TP / (TP+FN)$  表示正确预测为正的占

全部实际为正的比列。

### 5.2 map

实验评估采用的是 coco 数据集指标, AP is averaged over all categories。

## 6 ResNet FasterR-CNN 训练

实验选用 ResNet 作为主干网络搭建 FasterR-CNN 模型。选用了三种经典改进的 FasterR-CNN 网络:

- ① ResNet50-Faster R-CNN, ② ResNet101-Faster R-CNN, ③ ResNet101-Faster R-CNN + FPN。

### 6.1 训练硬件

实验采用 paddlepaddle 框架作为深度学习框架, paddlepaddle 是百度 2016 年正式开源的功能完备的深度学习框架具有易用和高性能等优点。实验硬件: Tesla V100 显卡 16G, 32G 运行内存, 处理器: Intel(R) Xeon(R) Gold 6271C CPU @ 2.60GHz。

### 6.2 训练方法

由于数据集较少, 所以采用模型预训练权重初始化网络(迁移学习)来加速 FasterR-CNN 模型快速收敛。

#### 6.2.1 迁移学习

迁移学习是指在先将神经网络模型放在大型多分类数据集上进行训练, 保留在大型多分类数据集上的权重, 然后将神经网络模型的权重冻结一部分放到其他小数据集上进行二次训练, 有着加速网络模型和减弱模型对硬件训练的要求。卷积神经网络本质是提取图像的特征信息, 当卷积神经网络学习过的数据越多时, 卷积得到的高特征图学习到的特征也越多。实验中三个目标检测模型均先在 COCO 数据集上进行预训练, 然后将其应用在小麦病虫害数据集上。在训练中固定 norm layer 的权重, 从第 2 阶段开始固定。

#### 6.2.2 边框聚类

实验中使用 kmeans 聚类算法对标注框进行了大小和比例的聚类。并且对数据集目标框进行聚类, 调节数据如下:

anchor\_sizes: [64,128,394,264,300,438,235,329]

aspect\_ratios: [0.3, 0.34, 0.51, 0.53, 0.6, 0.77, 1.01, 1.14, 2.32, 2.9, 4.13, 4.33]

#### 6.2.3 读取处理和迭代次数

对一个 batch 中的图片统一做的数据增强, 包括对图片进行随机翻转, 均值归一化, 随机裁剪。迭代次数 68000, batch\_size=5。每迭代 2000 次评估一次。

## 6.3 训练结果

表 3 训练集评估指标表

网络结构	Average Precision			Average Recall	Ave-fps
	IoU=0.50:0.95	IoU=0.50	IoU=0.75	IoU=0.50:0.95	
ResNet50-Faster R-CNN	IoU=0.50:0.95	IoU=0.50	IoU=0.75	IoU=0.50:0.95	20fps
	0.901	0.91	0.903	0.82	
ResNet101-Faster R-CNN	IoU=0.50:0.95	IoU=0.50	IoU=0.75	IoU=0.50:0.95	17 fps
	0.946	0.95	0.941	0.86	
ResNet101-Faster R-CNN + FPN	IoU=0.50:0.95	IoU=0.50	IoU=0.75	IoU=0.50:0.95	30 fps
	0.951	0.971	0.969	0.91	

表 4 测试集评估指数

网络结构	Average Precision			Average Recall	Ave-fps
	IoU=0.50:0.95	IoU=0.50	IoU=0.75	IoU=0.50:0.95	
ResNet50-Faster R-CNN	IoU=0.50:0.95	IoU=0.50	IoU=0.75	IoU=0.50:0.95	20fps
	0.821	0.830	0.826	0.769	
ResNet101-Faster R-CNN	IoU=0.50:0.95	IoU=0.50	IoU=0.75	IoU=0.50:0.95	17 fps
	0.841	0.866	0.856	0.828	
ResNet101-Faster R-CNN + FPN	IoU=0.50:0.95	IoU=0.50	IoU=0.75	IoU=0.50:0.95	30 fps
	0.906	0.921	0.916	0.861	

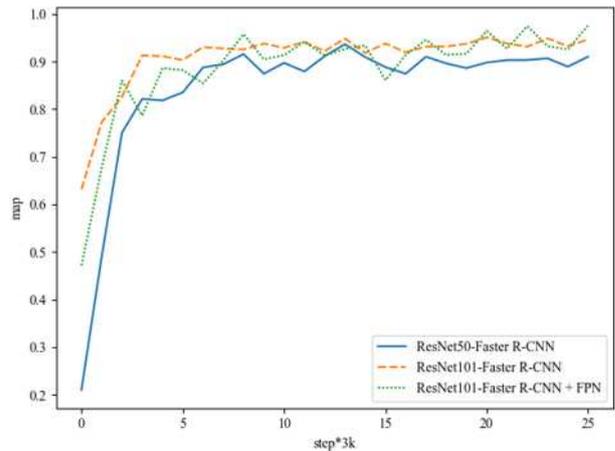


图 2 训练时 map 对比图

## 7 结果与分析

从实验结果和评价标准分析, 三种经典的 FasterR-CNN 系列网络在小麦数据集上的 Average Precision 都在 0.82 以上, 说明 FasterR-CNN 系列网络在复杂背景环境下对小麦病虫害检测有着良好的检测效果, 可以尝试将 FasterR-CNN 系列网络应用在小麦病虫害定点检测上。实验对比了三种 FasterR-CNN 网络。ResNet101-Faster R-CNN 网络比 ResNet50-Faster R-CNN 网络的 ap 值提高了约 4%, 说明随 ResNet 系列主干网络层数的加深, 对 Faster R-CNN 网络识别精度影响是积极的。随着主干网络层数加深, 基础主干网络学习到的图片有效特征越多, 信息表达能力越强, 在一定程度上能够有效提高网络识别的精确率和召回率。ResNet101-Faster R-CNN +

FPN 比 ResNet101-Faster R-CNN 的 ap 提高了约 1%，得益于 fpn 特征映射金字塔提高了 Faster R-CNN 网络对小目标的检测能力。

## 8 结束语

在开源农作物病虫害数据集比较缺乏情况下，网上公开收集了 500 张小麦病虫害图片，并且将其中 335 张图片进行病虫害特征标注，包括小麦白粉病，小麦锈病，小麦蚜虫三种病虫害类别。研究对于三种经典改进的 FasterR-CNN 网络对小麦病虫害数据集检测效果。在对比的三种网络模型中 ResNet101-Faster R-CNN + FPN 对复杂背景的小麦病虫害数据集得分最高。从检测准确效果来看，下一步可以尝试将模型部署在小麦田中做进一步的检测实验。

## 【参考文献】

- [1] 数据引用国家统计局官网.
- [2] 王克如. 基于图像识别的作物病虫害诊断研究 [D]. 北京: 中国农业科学院, 2005.
- [3] 任东, 于海业, 王纪华. 基于线性组合核函数支持向量机的病害图像识别研究 [J]. 农机化研究, 2007(09):41-43.
- [4] Bravo C, Moshou D, West J, McCartney A, Ramon H. Early disease detection in wheat fields using spectral reflectance. *Biosystems Engineering*. 2003 Feb;84(2):137-145. DOI: 10.1016/s1537-5110(02)00269-6.
- [5] 温芝元, 曹乐平. 基于为害状色相多重分形的椴柑病虫害图像识别 [J]. 农业机械学报, 2014, 45(03):262-267.
- [6] 姜敏, 沈一鸣, 张敬尧, 饶元, 董伟. 基于深度学习的水稻病虫害诊断方法研究 [J]. 洛阳理工学院学报 (自然科学版), 2019, 29(04):78-83.
- [7] 李想. 基于物联网的虫情监测系统 [D]. 北京: 北京林业大学, 2019.
- [8] S. Ren, K. He, R. B. Girshick, and J. Sun. FasterR-CNN: towards real-time object detection with region proposal networks. *CoRR*, abs/1506.01497, 2015.
- [9] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition, in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.
- [10] T. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie. Feature Pyramid Networks for Object Detection. *arXiv e-prints*, arXiv:1612.03144, Dec 2016.