

熊梦园,詹 炜,桂连友,等. 基于 ResNet 模型的玉米叶片病害检测与识别[J]. 江苏农业科学,2023,51(8):164-170.

doi:10.15889/j.issn.1002-1302.2023.08.024

# 基于 ResNet 模型的玉米叶片病害检测与识别

熊梦园<sup>1</sup>,詹 炜<sup>1</sup>,桂连友<sup>2</sup>,刘 虎<sup>1</sup>,王佩文<sup>1</sup>,韩 涛<sup>1</sup>,李伟豪<sup>1</sup>,孙 泳<sup>1</sup>

(1. 长江大学计算机科学学院,湖北荆州 434023; 2. 湖北省农林病虫害预警与调控工程技术研究中心/长江大学农学院,湖北荆州 434025)

**摘要:**针对当前玉米病害发生量大、病情复杂、难以防治,严重影响玉米产量和质量的问题,提出了一种基于卷积神经网络和迁移学习的玉米叶片病害检测与识别方法。首先收集了 3 827 张玉米健康叶片图像和 3 种不同的玉米病害叶片图像样本,为了使模型拥有更好的泛化能力,使用生成对抗网络对样本进行处理,得到分辨率更高的样本,再对样本进行平移旋转,使样本数量达到 5 153 张。然后构建 ResNet 模型,分别对 ResNet34、ResNet50 及其添加 CBAM 注意力机制和 FPN 特征金字塔网络,并通过迁移学习方法将预训练权重迁移到训练模型中。试验结果表明,ResNet50 结合 CBAM 注意力机制模型的准确率达到 97.5%,相比 ResNet50 模型准确率提升了 4.2 个百分点,相比 ResNet34 模型准确率提升了 4.9 个百分点。本研究表明,提出的 ResNet50 结合 CBAM 注意力机制模型能够较精准地检测识别玉米枯萎叶、锈病叶、灰斑病叶和健康叶。并可将模型安装在无人机等移动设备上,实现对玉米叶片病害智能化防治,而且后期还会扩充更多的植物病害数据,实现对多类植物病害的检测,为智慧农业添砖加瓦,促进农业防治现代化。

**关键词:**玉米病害识别;ResNet;CBAM 注意力机制;智慧农业;计算机视觉

**中图分类号:**TP391.41 **文献标志码:**A **文章编号:**1002-1302(2023)08-0164-07

玉米作为我国粮食种植的重要部分,种植面积最大,总产最多,是重要的饲料和工业原料,对保障我国粮食安全具有重要战略作用,因而玉米病害防治也尤为重要。玉米病害主要发生在叶片,如果在大面积种植地通过肉眼去观察病害,或者通过以往种植经验盲目采取防治措施,不仅不能清楚地把握病情,而且会造成防治成本大量浪费<sup>[1-3]</sup>。

随着人工智能和计算机视觉技术的不断发展,迎来了智慧农业时代,植物病害检测方法已由传统人工观察方法向深度学习检测方法过渡。传统方法需要提取病害纹理、颜色特征<sup>[4]</sup>,或是用肉眼直接观察,这样效率和准确度都不高,难以真正投入应用,而深度学习检测方法,能够高效准确地检测植物病害情况。在卷积神经网络的研究开始之前,全连接神经网络被广泛用于图像处理,但是由于全连接神经网络的参数多,因此网络模型较大,运算

速度往往达不到实时要求,而且容易产生过拟合现象,随着数值计算效率得到巨大提升,卷积神经网络处理数字图像数据的巨大潜能被充分发现挖掘<sup>[5-7]</sup>。如陶国柱提出了一种基于改进 MobileNetV2 的茉莉花病害识别网络模型,在兼顾速度的同时,精度达到了 94.34%<sup>[8]</sup>。刘君等提出了一种基于 YOLO 卷积神经网络的番茄病害识算法,对 8 类番茄病害的平均检测精度达到了 85.09%<sup>[9]</sup>。戴子兵提出了基于改进 U-Net 的水稻病害检测方法<sup>[10]</sup>。郭利进等提出一种改良的 LeNet 深度卷积神经网络用于玉米叶部病害识别和分类,平均准确率达到 95.39%<sup>[11]</sup>。上述网络虽然能够在复杂的环境中取得不错的效果,但是精度还有很大的提升空间。本研究基于 PyTorch 框架及 ResNet 模型,对玉米叶片病害数据集进行识别分类,并将改进算法与其他网络模型方法进行对比,可将训练好的模型安装在无人机上,对玉米病害进行实时检测和针对性防治。

## 1 材料与方法

### 1.1 数据来源

试验数据主要收集了 3 827 张玉米叶病害图片。分别是枯萎病叶、锈病叶、灰斑病叶和健康叶,样图如图 1 所示。

收稿日期:2022-06-17

基金项目:国家自然科学基金面上项目(编号:31772206、31972274);  
中国高校产学研创新基金新一代信息技术创新项目(编号:  
2020ITA03012)。

作者简介:熊梦园(1998—),男,湖北枝江人,硕士,主要从事计算机  
视觉技术应用研究。E-mail:2021710574@yangtzeu.edu.cn。

通信作者:詹 炜,博士,教授,主要从事计算机视觉技术应用研究。  
E-mail:zhanwei814@yangtzeu.edu.cn。

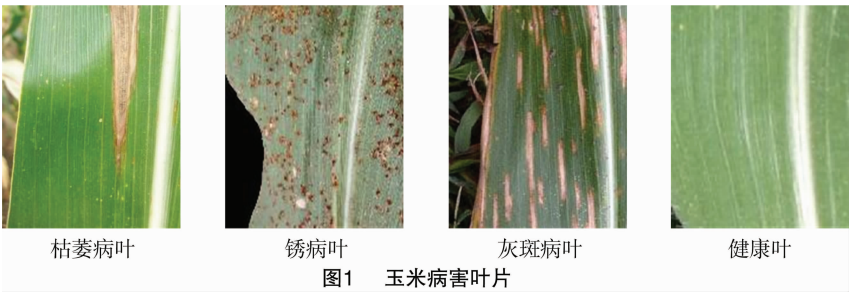


图1 玉米病害叶片

1.2 数据预处理

由于各类病害样本数不是十分充足,为了防止模型过拟合,能够有更好的训练效果,因此在训练之前通过生成对抗网络对样本进行处理,使得到的

新样本分辨率更高,并且对生成对抗网络处理得到的样本平移和旋转,使样本图片扩充到了 5 153 张,如图 2 图像预处理所示。

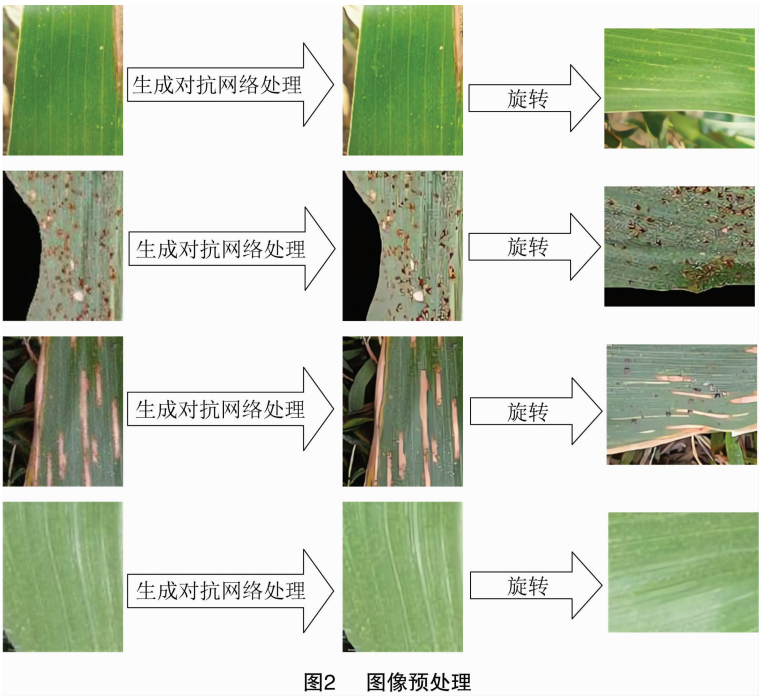


图2 图像预处理

1.3 方法

1.3.1 网络模型 本研究采用的网络模型是 ResNet50,通过残差模块在增加模型深度的同时防止梯度消失<sup>[12]</sup>,为了提高模型的精准度,在 ResNet50 的每一层上添加了 CBAM 注意力模块,在空间和通道 2 个不同的维度上添加注意力,然后与输入的特征映射相乘,对玉米叶片病害特征图进行特征提取<sup>[13]</sup>。本研究整体方案如图 3 所示,网络结构如图 4 所示。

1.3.2 残差结构 ResNet 在 2015 年由何恺明等提出,并且该网络在多个比赛中获得了不错的成绩,ResNet 网络中提出的残差结构在如今应用越来越广泛。深度卷积网络在图像分类上取得了很大的成功,例如 LeNet、VGGnet 和 GoogLeNet 都通过增加

网络深度,取得了较好的性能。但是后面发现,随着网络的加深,会出现梯度消失或者梯度爆炸,反而会导致准确率下降,而 ResNet 就可以解决这一问题,ResNet 学习的是残差函数  $G(x) = F(x) - x$ ,当  $G(x) = 0$  时,  $F(x) = x$ ,就是恒等映射。而 ResNet 没有加入其他的计算复杂度和参数,当优化目标无限接近恒等映射时,相比学习一个新的映射函数,学习找到对恒等映射的干扰会更加简单。完成这一功能的就是如图 5 所示的残差块结构,图 5 - a 是用于 50 层以下网络的结构,图 5 - b 是用于 50 层及更高层网络的结构。

随着网络层数的增加,计算量会大量增加,为了解决这个问题,对残差块做了计算优化,即将图 5 - a 中 2 个  $3 \times 3$  的卷积层替换为  $1 \times 1 + 3 \times 3 + 1 \times$

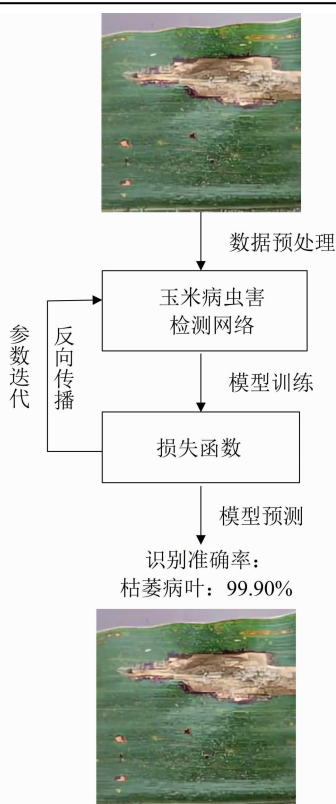


图3 研究方案流程

1。图 5-b 中  $3 \times 3$  的卷积层首先在一个降维  $1 \times 1$  卷积层下减少了计算,然后在另一个  $1 \times 1$  的卷积层下做了还原,既保持了精度又减少了计算量。

1.3.3 CBAM 注意力机制 注意力机制类似人的注意力,通过观察目标所有特征,得到比较重要的目标特征,然后投入更多的注意力资源,获取更多

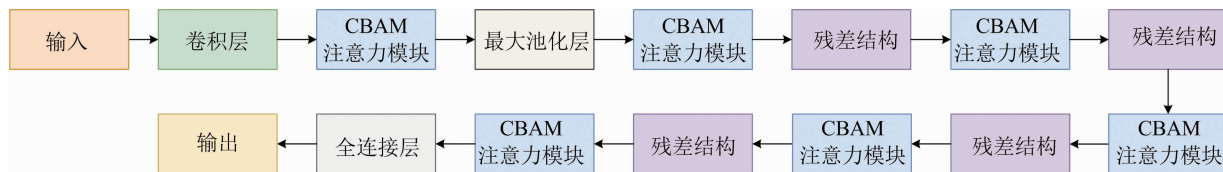
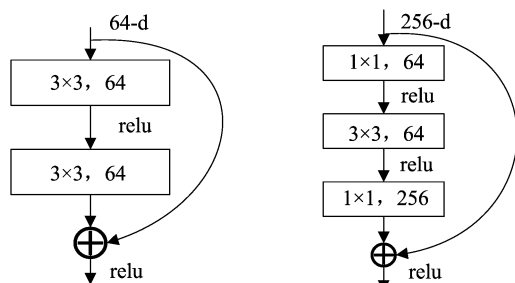


图4 网络流程



a. 用于 50 层以下网络

b. 用于 50 层及更高网络

通过输入不同通道的特征图,经过卷积、激活操作,然后输出

图5 残差结构

更详细的信息,减少甚至忽略对其他特征的关注。即当模型计算能力一定时,注意力机制可以将算力分配到需要注意的特征上,这样就可以更好地学习特征,并且可以减少资源浪费。CBAM 结合了通道注意力模块和空间注意力模块,分别在通道和空间位置上设置注意力机制,并且合并起来对特征进行优化,其结构如图 6 所示。通道注意力模块关注输入图像中有意义的内容,空间注意力模块专注信息的位置,这 2 个注意力模块可以按照并行或不同顺序串行连接<sup>[14-15]</sup>。

通道注意力机制输入特征是  $H \times W \times C$ ,经过最大池化和平均池化可以得到 2 个  $1 \times 1 \times C$  的特征图,然后通过全连接神经网络,再与权重系数相乘,其结构如图 7 所示。通道注意力机制可以表达为:

$$M_c(F) = \sigma \{ MLP [ \text{avgPool}(F) ] + MLP [ \text{maxPool}(F) ] \} = \sigma \{ W_1 [ W_0 (F_{\text{avg}}^C) ] + W_1 [ W_0 (F_{\text{max}}^C) ] \}.$$

空间注意力机制将输入的特征图分别进行最大池化和平均池化,得到 2 个  $H \times W \times 1$ ,然后拼接起来得到  $H \times W \times 2$  的特征图,再经过卷积核为  $7 \times 7$  的卷积降为 1 通道,然后再通过 Sigmoid 函数生成空间权重系数,其结构如图 8 所示,表达式如下:

$$M_s(F) = \sigma (f^{7 \times 7} \{ [ \text{avgPool}(F); \text{maxPool}(F) ] \}) = \sigma \{ f^{7 \times 7} [ (F_{\text{avg}}^S; F_{\text{max}}^S) ] \}.$$

由于 CBAM 是轻量级模块,将其运用到 ResNet 网络中,在保留图像信息并提升小样本玉米病害识别准确率的同时,不会增加过多的网络参数。玉米

病害图像特征经过 CBAM 处理后,输出更具病害本质特性的精细特征。

1.3.4 迁移学习 迁移学习是一项允许对现有模型算法进行微调,以应用于新领域或新功能的技术。迁移学习中可以将数据分成源数据和目标数据。源数据是与本项目无直接关系的数据,通常具有较大的数据集,目标数据是与任务直接相关的数据,数据量一般较小。其主要思想是从其他相关领域中迁移标注数据或模型结构,运用到现有的模型中加以改进,完成目标域的任务<sup>[16]</sup>。迁移学习改变

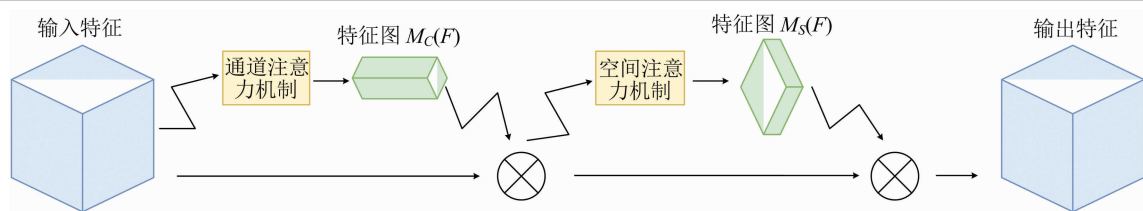


图6 CBAM 注意力机制

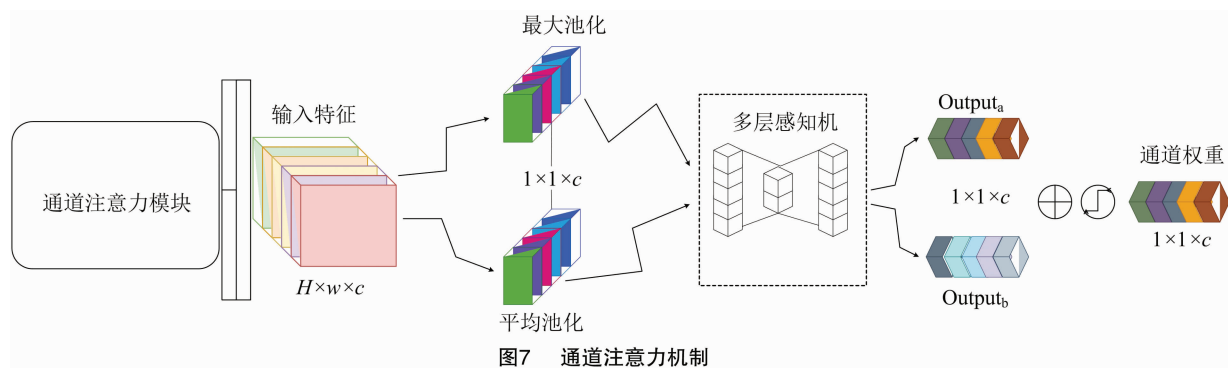


图7 通道注意力机制

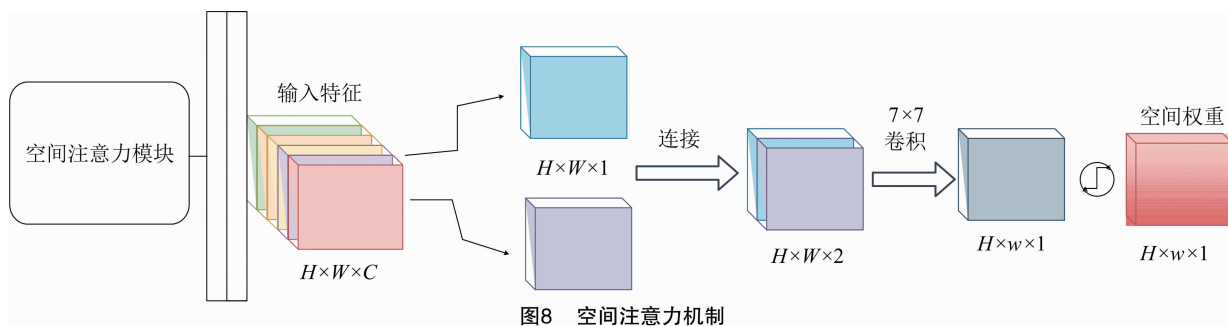


图8 空间注意力机制

了传统训练神经网络模型的固有模式,增加了使用其他域进行预训练的步骤,降低了资源损耗,提高了训练效率<sup>[17]</sup>。本项目导入 ResNet50 在其他域数据上的预训练模型,然后对玉米病害分类模型进行训练。如图 9 所示,其中微调的卷积层和模型的修改都会影响迁移训练的结果。

**1.3.5 学习率** 在训练过程中,学习率表示网络参数在每次迭代时改变的幅度。当学习率过大时,损失函数会产生震荡而不收敛,当学习率过小时,损失函数收敛速度慢。因此在神经网络训练过程中,需要不断调整网络的参数以使模型达到更好的预测效果,其中参数更新的公式如下:

$$W_{n+1} = W_n - L \times \nabla$$

式中:  $W_{n+1}$  是更新后的参数;  $W_n$  是当前参数;  $L$  代表学习率;  $\nabla$  是损失函数的梯度。损失函数就是网络对于单个样本的预测值与真实值的差的函数,不同的学习率将会对损失函数产生不同的影响<sup>[18]</sup>。

试验中尝试了指数衰减、余弦退火和余弦退火重启机制,最终发现余弦退火学习率收敛更快,精确度更高,更适合本试验。

## 2 结果与分析

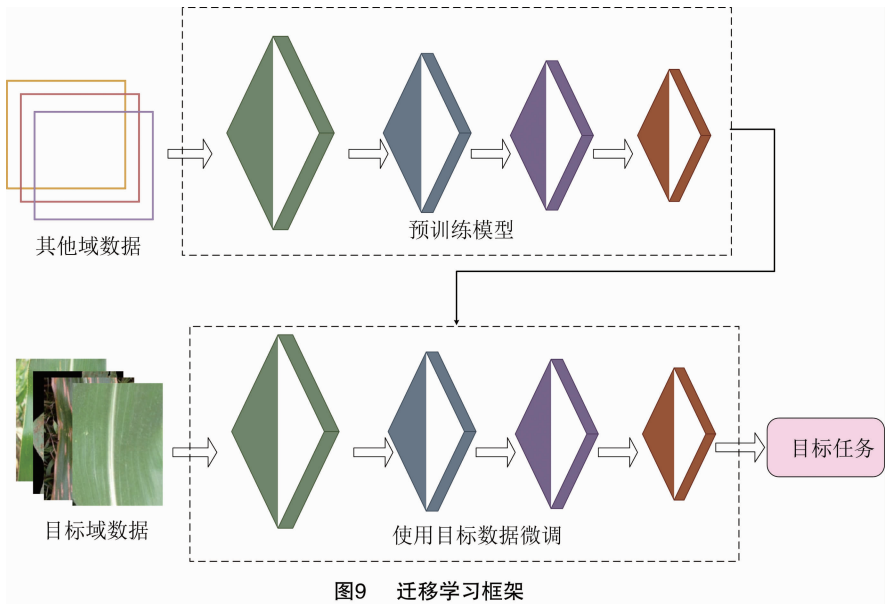
### 2.1 试验条件

所有试验均使用同一平台,使用的计算机设备的处理器为 Intel® Core™ i9 - 9900K CPU @ 3.60 GHz,显卡是 NVIDIA GeForce RTX 2080Ti,显存为 12 G。用到的软件开发环境为 Windows10 企业版操作系统,Python 环境为 3.7,Cuda 10.0,深度学习框架为 Pytorch 1.7.1。

### 2.2 评价指标

为了能够比较全面地评价模型的性能,笔者所在课题组将从准确率(A)、损失值(L)、预测时间(T)和模型大小(M)4个指标评价模型。准确率<sup>[19]</sup>反映的是识别结果中识别正确的图像数量占全部





识别图像数量的比率,可以反映模型的训练结果,表达式如下:

$$A = TP / (TP + FP)。$$

其中,  $TP$  指实际是正类且预测为正类的样本数量,  $FP$  指实际为负类但预测为正类的样本数量。

损失值可以描述模型所预测的值与其真实值的差距,损失值越小,表明模型越准确。本模型采用的是交叉熵损失函数,它在作分类训练时是非常有用的。交叉熵主要描述的是实际值与期望值的距离,即交叉熵的值越小,这 2 个值就越接近。假设概率分布  $p$  为期望输出,概率分布  $q$  为实际输出,  $H(p, q)$  为交叉熵,则

$$H(p, q) = - \sum [p(x) \log q(x)]。$$

训练时间也是评价模型优良程度的一方面,因为训练时间更短,可以提高模型使用效率。而且在后期拓展到检测其他植物病害时,会投入大量时间

去训练模型,训练时间短就可以节省大量时间。另外就是检测每一张图片的速度,这关系到网络模型是否具有较好的效率和时效性,能否真正投入到实际运用中。因为将模型装载到无人机上进行实时检测识别并喷洒农药时,需要及时识别到玉米叶和具体病害,然后针对性喷洒农药。

模型大小在这里是指模型所占存储空间和每一次训练所占用的存储空间大小,如果模型太大,设备内存小,也不能完整使用。

2.3 试验结果

通过基于 ResNet50 网络加入 CBAM 注意力的模型,将预处理的数据集划分成 80% 训练集和 20% 验证集,将模型迭代训练 500 次,最终虫病害分类识别率达到了 97.5%,测试样图如图 10 所示。训练过程中准确率、损失值和学习率的变化如图 11 所示。

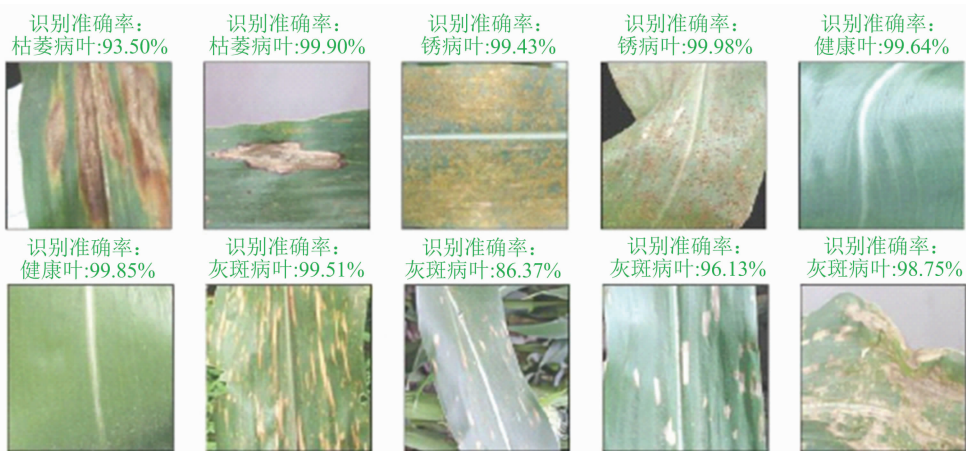
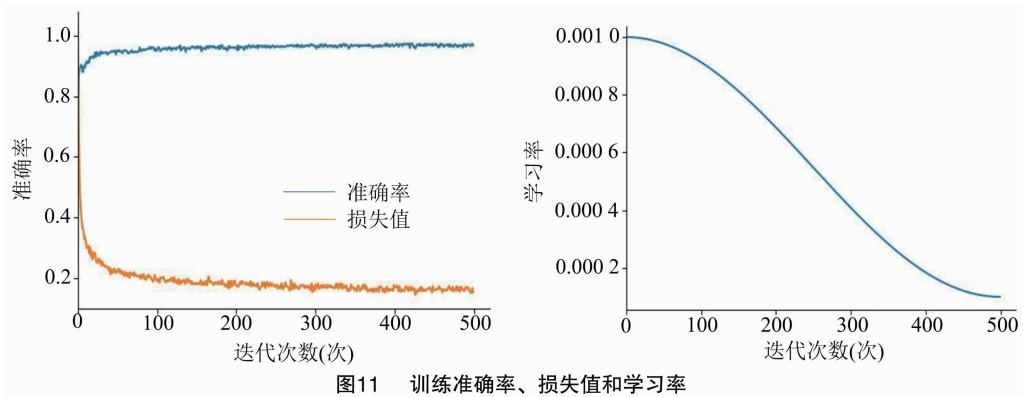


图10 测试效果



为了使试验更具科学性和合理性,在试验过程中增加了消融试验,在控制变量的情况下,分别对

网络、改进方法、数据集处理 3 项进行调整,得到的结果如表 1 所示。

表 1 各类消融试验结果对比

网络	改进方法	数据集	测试精度 (%)	测试时间 (s)	模型大小 (kb)
ResNet50	添加 CBAM 注意力机制	数据预处理	97.5	25	94 911
ResNet50		数据预处理	93.8	14	92 177
ResNet50	添加 CBAM 注意力机制		94.6	18	94 911
ResNet50			93.3	12	92 177
ResNet34			92.6	11	91 477

2.4 分析

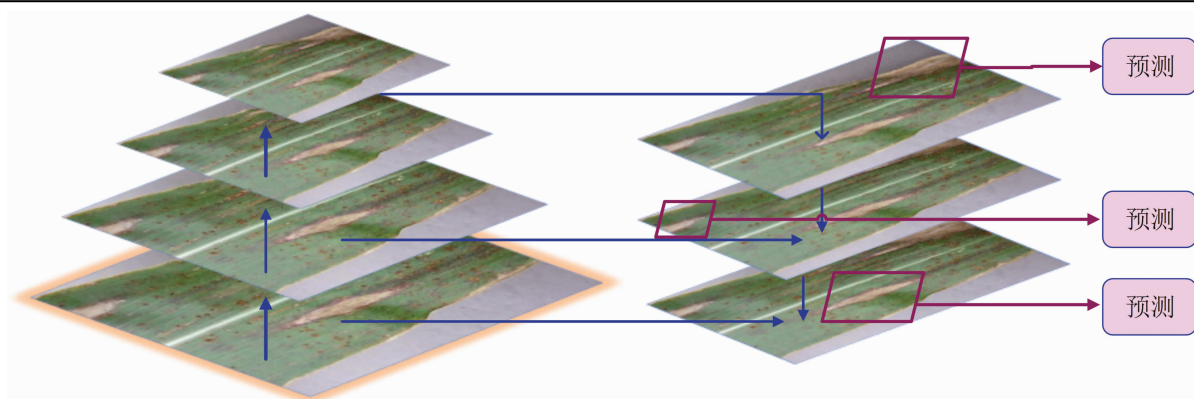
为了寻求更精确的试验结果,由于锈病和灰斑病都是小目标,因此尝试了 ResNet + FPN 结构<sup>[20]</sup>。小目标检测是语义分割和目标检测任务中经常遇到的难题。当目标非常小时,例如玉米灰斑病和锈病的一个斑点才几十个像素,卷积网络在实现卷积操作时,从一个特征层到下一个特征层,无论步长是大还是小,卷积核都会占满整个图片。然而大目标所占的像素会远大于小目标,因此大目标就会被卷积很多次,所以就会学习到很多关于大目标的特征,而小目标很容易被卷积操作跳过,因此学习到的小目标特征就会很少,经过很多层的卷积之后,小目标的特征就会越来越少。如图 12 所示,FPN 的提出是为了实现更好的特征图融合,一般的网络都是直接使用最后一层的特征图,虽然最后一层的特征图语义强,但是位置和分辨率都比较低,容易检测不到比较小的物体。FPN<sup>[21]</sup>通过自上向下的路径及横向连接将分辨率低但语义强的图像上层特征和语义弱但分辨率高的图像下层特征进行融合,进而将玉米叶片病害图像的语义信息和空间信息相结合,使得神经网络对于特征的检测性能有了很大的提升。但是通过实践来看,特征金字塔计算量

特别大,运行起来也很慢,并且准确率也只有 85%,该方法在准确度、训练时间和对机器性能要求上都不太符合本次项目需求,因此综合来看还是 ResNet50 + CBAM 更符合需求。

3 结论

针对目前玉米病害泛滥、受灾损失大的问题,为了能够精准喷洒农药,高效防治,降低污染,本研究采用 ResNet50 + CBAM 注意力机制的方法,对玉米 3 种常见且难以防治的病害和健康叶片进行识别研究。模型将 ResNet50 网络每一层融入 CBAM 注意力机制,更有利于网络学习到病害特征,平均识别精度达到了 97.5%,相比 ResNet50 模型准确率提升了 4.2 百分点,相比 ResNet34 模型准确率提升了 4.9 百分点。

在其他条件相同的情况下,本研究比较了 ResNet34、ResNet50、ResNet50 + CBAM 和 ResNet50 + FPN 方法,证实了 ResNet50 + CBAM 方法在准确率和速度 2 个方面综合实力最强。同时,也探究了不同大小的数据集对不同模型性能的影响,并且还探究了不同优化算法和学习率对模型收敛速度和准确率的影响。研究结果表明,迁移学习技术能在一



通过左下角输入原图，自底向上构建不同尺寸特征图，然后在右边自顶向下学习病害特征从而进行预测。

图12 特征金字塔结构

定程度上加快模型收敛速度；随着数据集的扩充增强，试验数据更加准确；余弦退火学习率效果最好，准确率最高，模型收敛速度也较快<sup>[22-24]</sup>。

在后期投入使用中，还将不断完善模型，而且计划将更多植物病害数据导入模型训练，得到可以检测多类植物病害的模型，这样可以更好地投入到生产实践中，节省人力物力，提高生产量，为智慧农业添砖加瓦<sup>[25-26]</sup>。

#### 参考文献：

- [1] 张会敏, 谢泽奇, 张善文. 基于注意力胶囊网络的作物病害识别方法[J]. 江苏农业科学, 2022, 50(6): 101-106.
- [2] 李 颢, 余心杰, 郭俊先. 基于全卷积神经网络方法的玉米田间杂草识别[J]. 江苏农业科学, 2022, 50(6): 93-100.
- [3] 刘翱宇, 吴云志, 朱小宁, 等. 基于深度残差网络的玉米病害识别[J]. 江苏农业学报, 2021, 37(1): 67-74.
- [4] 万军杰, 祁力钧, 卢中奥, 等. 基于迁移学习的 GoogLeNet 果园病虫害识别与分级[J]. 中国农业大学学报, 2021, 26(11): 209-221.
- [5] 曾伟辉, 唐 欣, 胡根生, 等. 基于卷积块注意力胶囊网络的小样本水稻害虫识别[J]. 中国农业大学学报, 2022, 27(3): 63-74.
- [6] 张军锋, 尚展奎. 基于深度学习卷积神经网络的花生籽粒完整性检测[J]. 食品与机械, 2022, 38(5): 24-29, 36.
- [7] 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.
- [8] 陶国柱. 基于卷积神经网络的茉莉花病虫害识别算法研究[D]. 南宁: 广西民族大学, 2021: 1-16.
- [9] 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.
- [10] 戴子兵. 基于语义分割的水稻病害检测技术[D]. 成都: 西华大学, 2020: 1-25.
- [11] 郭利进, 李 言. 基于 DCNN 的玉米叶部病害识别方法[J]. 粮食与油脂, 2022, 35(3): 72-75, 98.
- [12] 王海燕, 张 渺, 刘虎林, 等. 基于改进的 ResNet 网络的中餐图像识别方法[J]. 陕西科技大学学报, 2022, 40(1): 154-160.
- [13] 肖成勇, 李 擎, 王 莉, 等. 基于 CBAM-Unet 的铁矿球团边缘分割实验方法[J]. 烧结球团, 2022, 47(2): 8-15, 23.
- [14] 张德春, 李海涛, 李 勋, 等. 基于 CBAM 和 BiFPN 改进 YoloV5 的渔船目标检测[J]. 渔业现代化, 2022, 49(3): 71-80.
- [15] 乔星星, 施文灶, 刘芑汐, 等. 基于 ResNet 双注意力机制的遥感图像场景分类[J]. 计算机系统应用, 2021, 30(8): 243-248.
- [16] 王思颖. 基于迁移学习的多领域图像内容审核研究[D]. 成都: 电子科技大学, 2022: 12-20.
- [17] 高 爽, 徐巧枝. 迁移学习方法在医学图像领域的应用综述[J]. 计算机工程与应用, 2021, 57(24): 39-50.
- [18] 张 爽. 基于深度学习的双目立体匹配方法研究[D]. 西安: 中国科学院大学(中国科学院西安光学精密机械研究所), 2020: 31-34.
- [19] 张睿哲. 基于机器学习的黄砂岩声发射平静期识别研究[D]. 太原: 太原理工大学, 2021: 39-40.
- [20] 李 莉, 乔 璐, 张浩洋. 结合 FPN 改进 R-FCN 的肺结节检测算法[J]. 计算机应用与软件, 2022, 39(4): 179-184.
- [21] 杨建华. 基于 FPN 的高分辨率建筑物图像的语义分割[J]. 遥感信息, 2021, 36(4): 133-141.
- [22] Huang H Z, Zhan W, Du Z Q, et al. Pork primal cuts recognition method via computer vision[J]. Meat Science, 2022, 192: 108898.
- [23] She J H, Zhan W, Hong S B, et al. A method for automatic real-time detection and counting of fruit fly pests in orchards by trap bottles via convolutional neural network with attention mechanism added[J]. Ecological Informatics, 2022, 70: 101690.
- [24] Hong S B, Zhan W, Dong T Y, et al. A recognition method of *Bactrocera minax* (Diptera: Tephritidae) grooming behavior via a multi-object tracking and spatio-temporal feature detection model[J]. Journal of Insect Behavior, 2022, 35: 67-81.
- [25] Zhang Z L, Zhan W, He Z Z, et al. Application of spatio-temporal context and convolution neural network (CNN) in grooming behavior of *bactrocera minax* (Diptera: Trypetidae) detection and statistics[J]. Insects, 2020, 11(9): 565.
- [26] Zhan W, Sun C F, Wang M C, et al. An improved YOLOv5 real-time detection method for small objects captured by UAV[J]. Soft Computing, 2022, 26(1): 361-373.