

张志荣,陈勇,崔艳荣,等. 基于改进 ConvNeXt_Tiny 模型的水稻叶片病害识别方法[J]. 江苏农业科学,2025,53(20):308-315.
doi:10.15889/j.issn.1002-1302.2025.20.034

基于改进 ConvNeXt_Tiny 模型的水稻叶片病害识别方法

张志荣,陈勇,崔艳荣,胡蓉华

(长江大学计算机科学学院,湖北荆州 434000)

摘要:针对水稻叶片病害类别间差异较小、容易误识别以及病害重要特征难以提取等问题,提出了一种基于改进的 ConvNeXt_Tiny 模型的水稻叶片病害识别方法。首先收集细菌性条斑病、云纹叶枯病、窄条斑病 3 类病害以及健康叶片的图像,通过对比度增强、亮度增强、图像旋转和添加椒盐噪声等数据增强方式增加样本多样性以及增强模型泛化能力。接着以 ConvNeXt_Tiny 模型为基础,在降采样层后面添加 RFB 多尺度特征提取模块,通过扩张卷积对水稻叶片病害图像产生不同大小的感受野,增强网络对叶片病害中特征信息的学习能力。然后,在 stages 模块后面添加 Triplet 注意力机制模块,用于捕捉输入张量的空间和通道维度之间的跨维度交互信息,使模型着重关注水稻叶片表面的病害特征。试验结果表明,改进后的 ConvNeXt_Tiny 模型在水稻叶片病害数据集上的 Top-1 准确率可以达到 98.42%,较原模型的 97.14% 提升 1.28 个百分点,优于 ResNet34 (97.64%)、GhostNet (94.98%) 和 EfficientNet v2 (97.64%) 等 CNN 网络模型。综上所述,本研究提出的改进后的 ConvNeXt_Tiny 模型在水稻叶片病害识别方面具有更好的特征提取能力和特征表达能力,为水稻叶片病害识别提供了一种有效的方法。

关键词: ConvNeXt_Tiny; 水稻; 病害识别; 数据增强; RFB; Triplet

中图分类号: S126; TP391.41 **文献标志码:** A **文章编号:** 1002-1302(2025)20-0308-07

水稻是我国第一大口粮作物^[1]。2011 年以来,我国稻谷总产一直保持在 2 亿 t 以上,人均占有量稳定在 150 kg 左右,确保了国家口粮安全^[2]。然而,在水稻种植过程中,水稻叶片会受到各种细菌、霉菌的侵袭,导致稻谷的产量减少。2022 年,水稻纹枯病在我国稻田发生面积达 2.4 亿亩次(1 亩 = 1/15 hm²),稻瘟病发生面积达 6 000 万亩次,部分地区存在重发和扩散的风险^[3]。早期的水稻叶片病害检测可以让水稻叶片得到及时且有效的治疗,从而减少作物产量的损失。因此,本研究对水稻叶片病害的检测和防治有着重要的意义。

传统的水稻病害检测方法主要依靠人工识别,通过观察水稻叶片表面病斑的特征,再根据积累的经验进行识别,会受到个人主观的影响^[4]。随着人工智能技术在农业领域的广泛应用,基于深度学习和机器学习的农作物病害识别技术目前取得了比较显著的成果。翟梦婷等在 ConvNeXt 网络模型中

加入通道混洗模块,使通道特征之间的关联更加密切,以提升人参分级精确度,在 Block 中加入结构重参数化模块,提高模型对人参特征之间的表达能力,在人参数数据集上平均分类准确率达到 94.44%^[5]。田甜等在 ConvNeXt 模型中加入 SimAM 模块并采用 Focal Loss 函数作为模型中的损失函数,采用迁移学习的方式对茶叶病害图像进行微调,减少过拟合带来的影响,在茶叶病害数据集上准确率达 96.48%^[6]。Wu 等在 ConvNeXt 模型中引入了并行注意力机制(PAM)模块和多尺度特征融合(MFF)模块,以提高特征提取能力并减少复杂背景的影响,并使用迁移学习来提高模型的分类性能,提出的模型在自建的柑橘数据集上达到了 98.07% 的平均准确率^[7]。张澳雪等以玉米叶片病害为研究样本,以 RegNet 网络为主体,将 stem 中的卷积替换为 InceptionA 模块,并引入了金字塔池化模块,凸显玉米叶片病害中的重要特征,准确率最终提升了 1.26 百分点^[8]。马晓等在 ConvNeXt 模型中加入了 CBAM 注意力机制,使网络聚焦于玉米叶片病害的重要特征,以减少图像背景的干扰,在 CBAM 注意力模块中还使用了 LeakyReLU 激活函数,避免网络中出现神经元死亡的问题,在具有 3 种玉米常见叶片病害和健康叶片的图像数据集上,平

收稿日期:2024-08-16

基金项目:国家自然科学基金面上项目(编号:62077018)。

作者简介:张志荣(2002—),男,山西忻州人,硕士研究生,主要从事机器学习与人工智能研究。E-mail:3217126722@qq.com。

通信作者:陈勇,硕士,高级工程师,硕士生导师,主要从事 WEB 信息处理、人工智能应用研究。E-mail:285527563@qq.com。

均分类准确率达到 91.77%^[9]。针对水稻叶片病害识别问题,袁培森等使用 ECA 注意力机制改进 MobileNet v3Small 模型,并采用迁移学习的方法对模型进行优化,试验结果表明,在非迁移学习下,识别准确率达到 97.47%,引入迁移学习以后,准确率提升 0.09 百分点^[10]。崔金荣等将 MobileNet v3 中的 SE 注意力机制模块替换为 CA 注意力机制模块,加入 SVM 多分类器,改进后的模型在水稻叶片病害数据集上的识别准确率达 97.17%^[11]。卫雅娜等在 EfficientNet 模型中引入 CBAM 注意力机制,使用 Ghost 模块对网络中的卷积层进行优化,提升网络的整体性能并降低网络的参数量和计算量,最终在水稻病害数据集上识别准确率达到 95%,网络模型的参数量降低了 40%^[12]。

基于深度学习的水稻病害识别方法可以有效地解决人工识别方法的缺点,但是依旧存在一些不

足之处:相似病害很容易造成误识别,水稻病害识别容易受到图像背景的干扰。为解决此问题,本研究提出一种改进的 ConvNeXt 模型,通过使用不同大小的卷积核和扩张卷积捕获图像中的多尺度特征,添加 Triplet 注意力机制模块,实现对水稻健康叶片和细菌性条斑病、云纹叶枯病、窄条斑病 3 类病害叶片的准确分类。

1 试验材料

1.1 数据来源

本研究所使用的数据集来自 Kaggle 官方网站,包括水稻叶片病害细菌性条斑病、云纹叶枯病、窄条斑病 3 类病害以及健康叶片,总计 5 093 张,其中细菌性条斑病 1 284 张、云纹叶枯病 1 312 张、窄条斑病 1 266 张、健康叶片 1 231 张,图片大小均为 256 像素 × 256 像素,具体如图 1 所示。

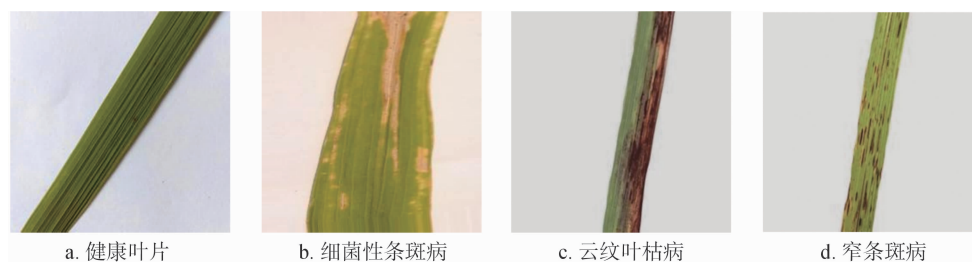


图1 部分水稻叶片病害示意

1.2 数据预处理

为提高模型的泛化能力,防止模型在训练时出现过拟合现象,本研究对原始水稻叶片数据集进行数据增强。为了确保不对测试数据集造成影响,采用离线数据增强的方法,先将数据集按照 6 : 2 : 2 的比例划分为训练集、验证集和测试集,然后只对训练集中的数据进行数据增强,增强后的图像见图

2,具体的增强方法为对比度增强、亮度增强、图像旋转以及添加椒盐噪声 4 种数据增强方法,其中旋转的角度为 90°,添加的噪声数目比例为 0.5,像素数目为 0.04 × 256 × 256 × 0.5。经过离线数据增强以后,数据集中共有图片 14 267 张,每类水稻病害图片的具体张数见表 1。

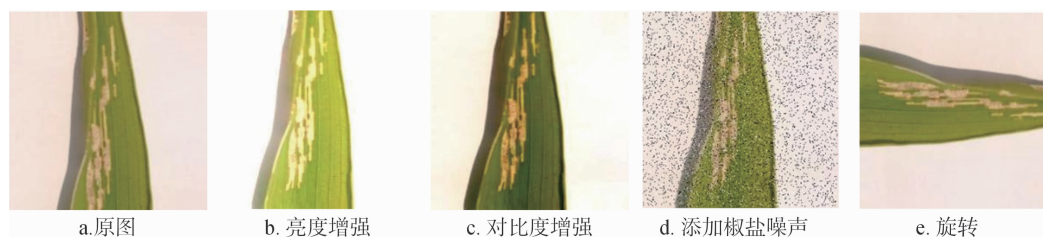


图2 数据预处理

2 试验方法

2.1 ConvNeXt_Tiny 基本模型

ConvNeXt 是一种计算机视觉模型,由 Facebook

AI 的研究人员在 2022 年提出,核心目标是结合 Transformer 模型中的设计理念,来改进传统的神经网络结构^[13]。ConvNeXt 网络根据模型的大小共有 4 个版本,分别是 ConvNeXt_Tiny、ConvNeXt_Small、

表 1 水稻叶片病害数据集

类别名称	原始图像数 (张)	增强后的图像数 (张)
细菌性条斑病	1 284	3 597
云纹叶枯病	1 312	3 676
窄条斑病	1 266	3 546
健康叶片	1 231	3 448
总计	5 093	14 267

ConvNeXt_Base 和 ConvNeXt_Large。本研究使用的是 ConvNeXt_Tiny 模型,该模型的结构主要由 3 个部分构成,分别是 downsample_layers、stages 和 head。downsample_layers 的第 1 个模块是由 4 × 4 的二维卷积和归一化层构成,先进行卷积再进行归一化;后面的 3 个模块都是由归一化层和 2 × 2 的二维卷积构成,先进行归一化再进行卷积。这里的归一化操作使用的是 LN 归一化,与 BN 归一化的不同之处在于,BN 归一化是对输入的数据进行批量的处理,而 LN 归一化是对输入数据的每个特征进行归一化,这样处理有助于提升模型的性能。stages 里面共有 4 个模块,每个模块都是由多个 Block 块组成的,Block 的堆叠次数分别为 3、3、9、3。head 里面包含全局平均池化、归一化层以及全连接层,对最后的输出结果进行分类。

Block 块中包含 1 个深度可分离卷积块、2 个逐点卷积块、1 个归一化层以及 1 个 GELU 激活函数。其中,深度可分离卷积块的核 K 为 7×7 ,填充 P 为 3,步长 S 为 1,根据公式 $Out = (In - F + 2P)/S$ (式中:Out 为输出特征图尺寸;In 为输入特征图尺寸; F 为可分离卷积核尺寸),可以得到,经过这个卷积的特征图,其高度 H 、宽度 W 的值不发生变化。在这个卷积中,groups = dim,其中 dim 为输入数据的通道数,对于输入数据的各个通道都会单独地与深度可分离卷积块中的一个 7×7 的卷积核进行卷积,共有 dim 个卷积核,每个卷积核都会与输入数据的 dim 个通道一一对应,因此,输出的特征图的通道信息都与对应的输入特征图的通道信息一一关联。这样做可以减少网络模型的计算量和参数数量,并使输出的特征图保持一定的特征表达能力,为后续的水稻叶片病害识别提供一定的帮助。逐点卷积会把输入的 dim 个通道增加到 $4 \times dim$ 个通道,可以使图片的特征之间有更多的组合以及通道之间有更多的交互,让模型学习到更丰富的特征,提高模型的泛化能力。在第一个深度可分离卷积块后面

是一个 LN 归一化层,通过 LN 归一化层对深度卷积之后的特征图进行归一化,可以进一步提升模型的稳定性。在第一个逐点卷积块后面使用 GELU 激活函数,可以减少网络中的梯度消失和爆炸现象。Block 块的结构见图 3。

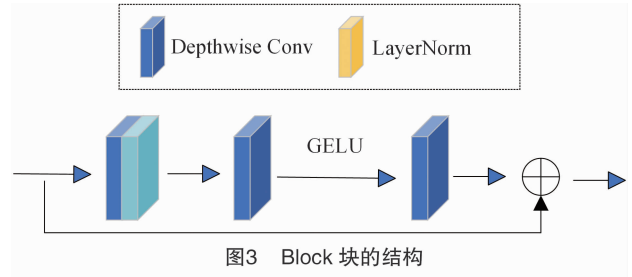


图3 Block 块的结构

2.2 Triplet 注意力机制

为了捕捉输入张量的空间和通道维度之间的跨维度交互,并在微小的计算开销下提供显著的性能提升,Misra 等在 2020 年提出了 Triplet 注意力机制^[14]。对于现有的注意力机制 SE、CBAM 等并没有完全考虑通道之间的跨维度交互,并且会增加网络模型的计算开销。因此,本研究引入了 Triplet 注意力机制,可以从输入张量的 3 个不同的分支之间建立交叉依赖,并且不会减少网络对维度的关注度,使得通道与权重之间的对应关系更为直接。Triplet 注意力机制,结合 3 种不同方向的注意力门,从宽度、高度、空间 3 个方面与通道维度建立交互,将输入的张量从 0 维变为 2 维,保留输入张量的重要特征,缩小深度,并且使其计算量进一步减少。根据特征生成对应的注意力权重,再将其进行 90° 旋转,保持与原本的形状相同,不改变输入特征的形状,然后将每个维度对应的注意力机制通过平均的方式进行聚合作为最终的输出,使得模型更好地关注水稻叶片上面病害之间的区别,有助于更快地识别不同的水稻病害,Triplet 结构见图 4。

假设输入张量 F 的大小为 $C \times H \times W$,其中 C 为输入图像的通道数, H 为高度, W 为宽度,然后传递到 3 个分支上面。首先,在第 1 个分支上面,面向高度 H 和通道 C 之间建立交互,将 F 在 H 轴方向上逆时针旋转 90°,将其变为 $W \times H \times C$,通过 ZPool 模块,进行最大池化和平均池化,变为 $2 \times H \times C$,之后通过 $k \times k$ 的标准卷积层和批归一化层,得到 1 个 $1 \times H \times C$ 的中间输出,再经过激活函数 Sigmoid 获得相应的权重,最后再沿着 H 轴方向上顺时针旋转 90°,与输入时的张量形状保持一致。然后,在第 2 个分支上面,面向宽度 W 和通道 C 之间建立交互,

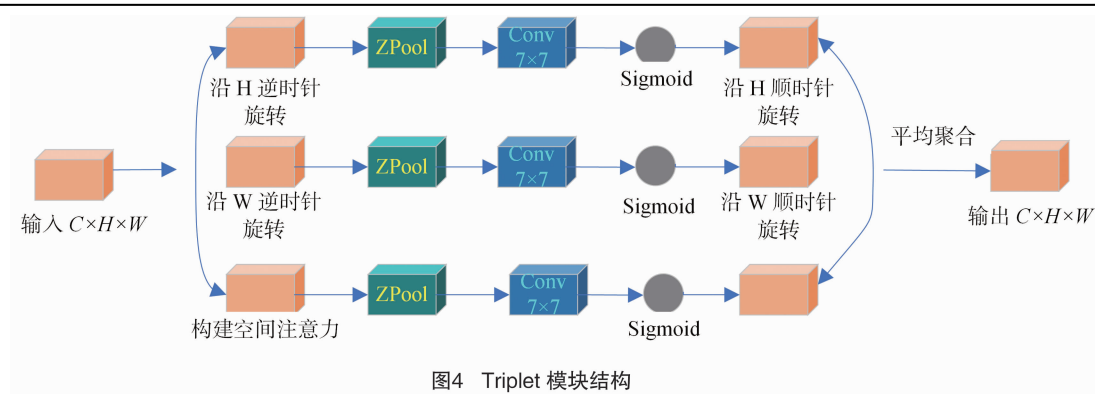


图4 Triplet 模块结构

将 F 在 W 轴方向上逆时针旋转 90° , 变为 $H \times C \times W$, 经过 ZPool 模块, 变为 $2 \times C \times W$, 之后通过标准卷积层和批归一化层, 得到 1 个 $1 \times C \times W$ 的中间输出, 再经过激活函数 Sigmoid 获得相应的权重, 最后再沿着 W 轴方向上顺时针旋转 90° 。最后, 对于第 3 个分支, 用于构建空间注意力, 输入的张量经由 ZPool 模块减少到 2 个, 经过与前 2 个分支相似的步骤得到 1 个 $1 \times H \times W$ 的张量。将 3 个分支上面的张量通过平均聚合得到形状为 $C \times H \times W$ 的张量 Q 。从输入张量 F 到输出张量 Q 的过程为:

$$Q = \frac{1}{3}(\overline{\hat{F}_1 w_1} + \overline{\hat{F}_2 w_2} + \overline{\hat{F}_3 w_3}) + \frac{1}{3}(\overline{Q_1} + \overline{Q_2} + \overline{Q_3})。$$

式中: F_1 、 F_2 、 F_3 为 3 个不同方向卷积层; w_1 、 w_2 和 w_3 是 Triplet 注意力机制中 3 个交叉维度的注意权重; Q_1 、 Q_2 、 Q_3 表示 3 个分支分别聚合的张量。

2.3 RFB 多尺度特征提取模块

目前, 在深度学习领域中, 为了更好地捕捉图像数据的重要特征, 研究人员提出了很多的特征提取网络, 例如 Inception、ASPP 等^[15-16]。特征提取网络可以有效地提取图像中的重要信息, 但是在提取过程中, 会带来很大的计算开销, 另一方面, 有些轻量化的特征提取网络, 在参数表现方面没有给模型增加计算量, 但是所带来的精度方面的提升微乎其微。为此, Liu 等通过人工特征加强这些轻量化模型对图片的特征提取能力, 以 RFs 的尺寸和偏心距之间的关系作为考量, 提出了 RFB 多尺度特征提取模块^[17]。RFB 参考了 Inception 结构, 在其基础上引入了扩张卷积, 设定各分支的扩张率分别为 branch1: $d=3$, branch2: $d=3$, branch3: $d=5$, 进而产生不同大小的感受野, 以不同大小的感受野为基础来增强网络对水稻叶片病害图片中特征信息的学习能力, 使其更快更好的学习到重要特征。

RFB 模块中有 3 个分支, 分别是 branch0、

branch1 和 branch2, 这 3 个分支结构中依次结合了 2、3、4 个卷积块, 这些卷积块中都包含着扩张卷积、BN 归一化层和 ReLU 激活函数。首先, 将图片依次经过 3 个分支结构, 对于分支结构中的每一个卷积块都是按照各自设定的卷积核大小以及扩张率进行卷积, 将输入通道数变为所设定的数值再输入, 然后经过 BN 归一化层进行归一化, 再将结果通过 ReLU 激活函数进行输出。ReLU 激活函数具有非线性特征, 对于神经网络中输出为负值的结果, 可以将其变为 0, 正值不发生变化, 可以帮助网络解决非线性问题, 提升了 RFB 模块在这方面的表达能力, 并可以适当的缓解训练过程中的过拟合现象。同时, RFB 模块中还使用了线性变换和残差连接, 在上述过程中, 输入的图片经过 3 个分支结构, 依次产生了 3 个不同的输出结果, 但是最终只需要 1 个输出, 所以还需要对这 3 个结果进行拼接, 从通道维度对其进行拼接, 然后再经过一个内核为 1×1 、步长为 1 的卷积层按照所设定的输出通道数进行线性变换。为了减少深层网络中梯度消失现象的发生, 在线性变换之后进行残差连接, 将具有相同尺寸和通道数的输入特征图与线性变化之后的特征图进行加权求和, 通过设置权重来分配残差项和直接变换项在总和中所占的比重, 让网络在不降低性能的基础上学习输入与输出之间的残差, 加快整体的训练过程。最后, 将求和后的特征经过 GELU 激活函数输出作为最终的结果。RFB 模块结构如图 5 所示。

具体的步骤为: 输入一个特征张量, 形状为 (N, C, H, W) , 分别表示单次处理样本的批量大小、通道数、宽度和高度。依次经过 3 个分支结构, 张量形状变为 (N, C_1, H, W) 、 (N, C_2, H, W) 和 (N, C_3, H, W) , 然后进行拼接, 结果为 $(N, C_1 + C_2 + C_3, H, W)$, 经由线性变换变为 (N, C_4, H, W) 。原始的张量 (N, C, H, W) 经过残差项变为 (N, C_5, H, W) , 再将 $(N, C_5, H,$

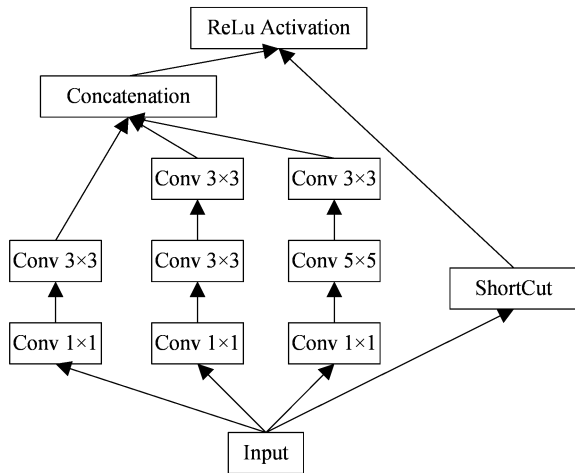


图5 RFB 模块结构

W)与线性变换后的 (N, C_4, H, W) 进行加权相加,得到 (N, C_7, H, W) ,然后经过激活函数,最终得到 (N, C_{out}, H, W) ,其中 C_{out} 的数值与初始输入维度 C 相同。

2.4 改进的 ConvNeXt_Tiny 模型

为了增强图像中的特征表达能力,更全面地理解输入数据的特征以及获得张量中不同维度之间

的交互信息,本研究在 ConvNeXt_Tiny 模型基础上进行了改进。首先,在降采样模块中的 LN 归一化层后面添加 RFB 多尺度特征提取模块。水稻叶片上的病斑由于颜色深浅、形状大小不同以及不同病害的类间差距小等原因会导致图片中的重要特征难以捕获,而 RFB 模块通过模拟人类视觉感受野的特性,使用多个不同尺寸的扩张卷积从多个尺度捕获图片中的重要特征,生成更高分辨率的特征图,得到水稻叶片表面病害的深层特征。然后,在 ConvNeXt_Tiny 模型的 stages 模块后面添加 Triplet 注意力机制,由于 stages 共有 4 个阶段,每个阶段的特征维度分别是 96、192、384 和 768,因此,共添加 4 次 Triplet 注意力机制。水稻叶片表面的病害在整个图像中所占的比例较小,通过使用 Triplet 注意力可以使模型着重关注水稻叶片病害的信息,降低对没有病害的叶片表面以及图像背景的关注度,可以从整张水稻叶片图片中更快地筛选出有利于识别的信息,提高水稻病害识别的效率和准确率。改进的 ConvNeXt_Tiny 模型结构图见图 6。

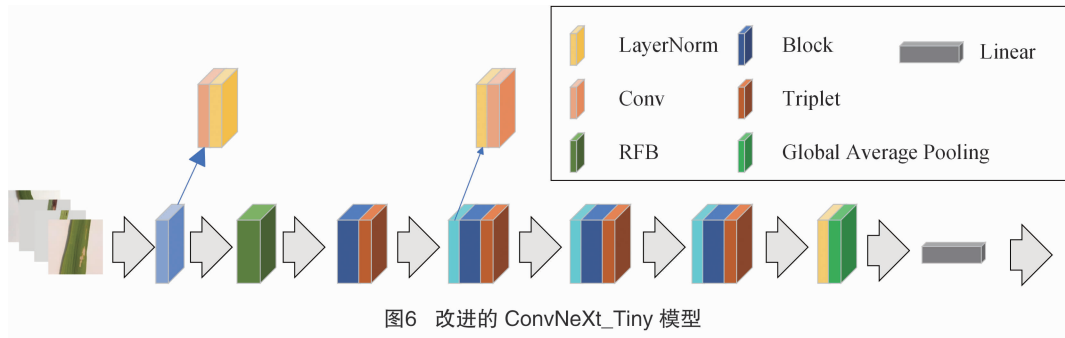


图6 改进的 ConvNeXt_Tiny 模型

3 试验结果与分析

3.1 试验环境与参数设置

本研究使用的 Python 版本为 3.10,深度学习框架为 PyTorch 2.1.0。CPU 型号为 16 vCPU Intel® Xeon® Gold 6430, GPU 型号为 RTX 4090,显存 24 G。操作系统为 Ubuntu 22.04。试验时间为 2024 年 6—8 月,试验地点为长江大学计算机科学学院。本研究均采用 CrossEntropyLoss 损失函数和 AdamW 优化器训练所有模型,动量设置为 0.9,权重衰减为 0.0005,初始学习率为 0.0001,采用余弦退火学习率策略, batch size 设置为 32,输入通道为 3,图片大小设置为 224 像素 × 224 像素,并设置 epoch 为 100 次。

3.2 评估指标

为了全面评估水稻叶片病害识别模型的性能,

本研究采用 Top - 1 准确率、Macro 平均精确率、Macro 平均召回率、Macro 平均 F_1 分数作为评价指标。本研究使用混淆矩阵描述模型的性能,将水稻叶片病害实际所属的类别与模型所预测的类别进行比较,矩阵中的行代表每个类别中实际的样本,列代表通过模型所预测的类别样本。通常包含 4 个指标,分别是真正类 TP、假正类 FP、假负类 FN、真负类 TN,其中:TP 代表将属于目标类别的样本正确预测;FP 代表将不属于目标类别的样本错误预测;FN 代表将属于目标类别的样本错误预测;TN 代表将不属于目标类别的样本正确预测。Top - 1 准确率表示预测类别与真实类别相同的样本在所有样本中所占的比例,计算公式为 $Accuracy = (TP + FN) / (TP + TN + FP + FN)$;精确率表示预测类别与真实类别相同的正样本数量在所有预测类别为正

样本的数量中所占比例,计算公式为 $Precision = TP / (TP + FP)$;召回率代表预测类别与真实类别相同的正样本数量在所有实际类别为正样本的数量中所占比例,计算公式为 $Recall = TP / (TP + FN)$; F_1 分数综合了精确率和召回率,计算公式为 $F_1 \text{ 分数} = 2 \times Precision \times Recall / (Precision + Recall)$ 。

3.3 试验结果与分析

3.3.1 数据增强对模型性能的影响 由于水稻叶片背景过于单一,为了提高新网络模型的泛化能力,本研究对水稻叶片数据集进行扩充。以改进后的 ConvNeXt_Tiny 模型为试验模型,在其他参数不变的情况下,分别对扩充前的水稻数据集(5 093 张)和扩充后的水稻数据集(14 267 张)进行对比试验,识别准确率分别为 94.00% 和 97.14%,提升了 3.14 百分点。数据增强前后的准确率曲线见图 7。

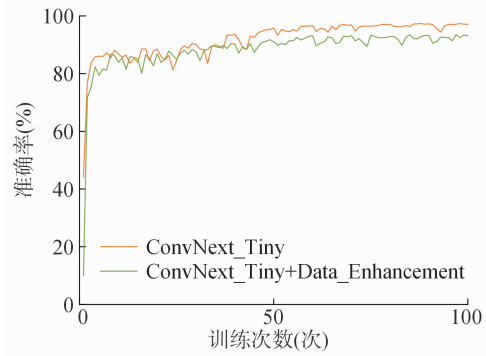


图7 数据增强前后准确率曲线

3.3.2 改进后的模型分析 本研究模型是在改进后的 ConvNeXt_Tiny 模型基础上得到的,训练轮数为 100 次,准确率为 98.42%,相比于改进之前的 97.14% 提升了 1.28 百分点。训练过程中对水稻叶片病害识别的准确率和损失值的变化以及混淆矩阵见图 8。

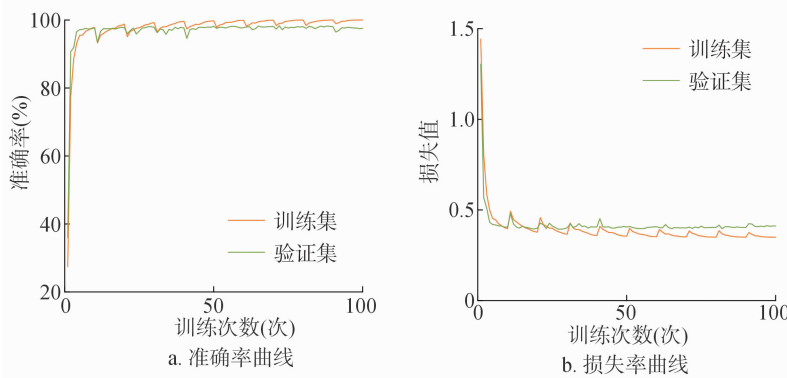


图8 准确率、损失值曲线和混淆矩阵

窄条斑病	0	0.008 1	0	0.944 7
云纹叶枯病	0	0	1	0
健康叶片	0	0.991 9	0	0.055 3
细菌性条斑病	1	0	0	0
	细菌性条斑病	健康叶片	云纹叶枯病	窄条斑病
	预测标签			
	c. 混淆矩阵			

3.3.3 不同分类模型性能对比试验 为了更直观地评估改进的 ConvNeXt_Tiny 模型在水稻叶片病害识别任务中的性能表现,在相同的试验环境、水稻数据集、训练批次(训练 100 次)下,设计对比试验,将改进的 ConvNeXt_Tiny 模型与 ResNet34、VGG19、GhostNet、EfficientNet v2、MobileNet v3 进行对比^[18-22]。由对比结果(表 2)可知,改进后的 ConvNeXt_Tiny 模型在 Top-1 准确率、Macro 平均精确率、Macro 平均召回率、Macro 平均 F_1 分数上均达到了最优,其中 Top-1 准确率为 98.42%、平均精确率为 98.43%、平均召回率 98.41%、平均 F_1 值为 98.39%。改进后的 ConvNeXt_Tiny 模型的 Top-1 准确率比 ResNet34、VGG19、GhostNet、EfficientNet v2、MobileNet v3 分别提高 0.78、1.37、3.44、0.78、1.86 百分点。在验证集上的准确率见图 9。

表 2 不同模型的性能对比

模型	准确率 (%)	精确率 (%)	召回率 (%)	F_1 分数 (%)
ResNet34	97.64	97.60	97.60	97.59
VGG19	97.05	97.04	97.01	96.99
GhostNet	94.98	95.23	94.98	94.91
EfficientNet v2	97.64	97.60	97.60	97.59
MobileNet v3	96.56	96.56	96.55	96.51
ConvNeXt_Tiny	97.14	97.13	97.13	97.10
改进的 ConvNeXt_Tiny	98.42	98.43	98.41	98.39

3.3.4 消融试验 本研究针对水稻叶片病害识别,在 ConvNeXt_Tiny 模型上做了 2 个方面的改进,第 1 个改进点是添加 RFB 多尺度特征提取模块,第 2 个改进点是添加 Triplet 注意力机制模块,为探究这 2 个改进点对 ConvNeXt_Tiny 模型的具体影响和准确率方面的提升表现,为此设计了消融试

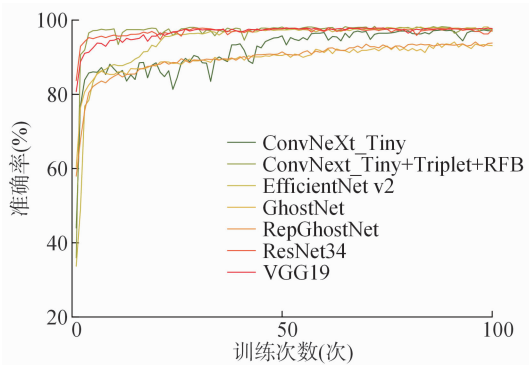


图9 各模型验证集准确率对比曲线

验。消融试验的结果见表3。由表3可知,ConvNeXt_Tiny + RFB 的 Top - 1 准确率可以达到 98.32%, 与原模型相比提升 1.18 百分点;ConvNeXt_Tiny + Triplet 的 Top - 1 准确率可以达到 97.83%, 比原模型提升 0.69 百分点。当 RFB 和 Triplet 同时添加到 ConvNeXt_Tiny 模型中, 构建 ConvNeXt_Tiny + RFB + Triplet 模型, 模型的 Top - 1 准确率可以达到 98.42%, 比原模型提升 1.28 百分点。由此可以得出, 分别添加 RFB 和 Triplet, 模型的准确率都有所上升, 当这 2 个改进点同时添加, 模型的准确率达到最大值, 效果最好。在验证集上的准确率见图 10。

表3 消融试验结果对比

模型	Top-1 准确率 (%)	精确率 (%)	召回率 (%)	F ₁ 分数 (%)
ConvNeXt_Tiny	97.14	97.13	97.13	97.10
ConvNeXt_Tiny + Triplet	97.83	97.86	97.81	97.79
ConvNeXt_Tiny + RFB	98.32	98.33	98.31	98.29
ConvNeXt_Tiny + Triplet + RFB	98.42	98.43	98.41	98.39

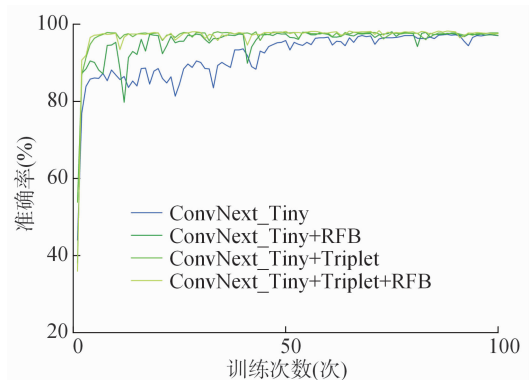


图10 消融试验验证集准确率对比曲线

4 结论

针对水稻叶片表面病害类别之间差异较小导致的误识别和识别准确率不高的情况, 本研究提出了一种改进的 ConvNeXt_Tiny 模型, 完成了对 3 类水稻病

害(细菌性条斑病、云纹叶枯病、窄条斑病)以及健康水稻叶片的识别研究。在 ConvNeXt_Tiny 模型的降采样层后面添加 RFB 多尺度特征提取模块, 能够更好地提取水稻叶片病害的重要特征, stages 模块后面添加 Triplet 注意力机制模块, 进一步增加水稻叶片病害特征不同维度之间的交互。试验结果表明, 改进后的 ConvNeXt_Tiny 模型对于水稻病害的识别有更好的表现效果, 相比于原 ConvNeXt_Tiny 模型, 提升了 1.28 百分点, 最终达到了 98.42% 的识别准确率; 比一些 CNN 网络模型 ResNet34、VGG19、GhostNet、EfficientNet v2、MobileNet v3 的识别效果更好, 可以让水稻叶片病害更快更准确地被识别, 使水稻病害在早期阶段得到有效防治。虽然, 改进后的 ConvNeXt_Tiny 模型在识别准确率方面有了提升, 但是整体的参数量并没有下降, 希望在未来的工作研究中, 让网络模型变得更加轻量化, 同时能够收集更多的水稻叶片病害图像, 进一步提升模型的预测能力。

参考文献:

[1]徐春春, 闻军清, 纪 龙, 等. 中国水稻种业发展现状、问题与展望[J]. 中国稻米, 2022, 28(5): 74 - 78.

[2]徐春春, 纪 龙, 李凤博, 等. 当前我国水稻产业发展形势与战略对策[J]. 华中农业大学学报, 2022, 41(1): 21 - 27.

[3]2022 年全国农作物重大病虫害发生趋势预报[J]. 中国植保导刊, 2022, 42(4): 107 - 108.

[4]周驰燕, 朱宇涵, 姚照胜, 等. 基于信息技术的水稻病害识别与检测研究进展[J]. 现代农业科技, 2019(7): 111 - 113.

[5]翟梦婷, 张丽娟, 朴欣茹, 等. 基于改进 ConvNeXt 网络的人参分级模型[J]. 吉林农业大学学报, 2023, 45(6): 791 - 802.

[6]田 甜, 程志友, 鞠 薇, 等. 基于 SimAM - ConvNeXt - FL 的茶叶病害小样本分类方法研究[J]. 农业机械学报, 2024, 55(3): 275 - 281.

[7]Wu Q H, Ma X, Liu H F, et al. A classification method for soybean leaf diseases based on an improved ConvNeXt model[J]. Scientific Reports, 2023, 13(1): 19141.

[8]张澳雪, 崔艳荣, 李素若, 等. 基于改进 RegNet 网络的玉米叶片病害识别研究[J]. 江苏农业科学, 2024, 52(11): 216 - 224.

[9]马 晓, 邢 雪, 武青海. 基于改进 ConvNext 的复杂背景下玉米叶片病害分类[J]. 江苏农业科学, 2023, 51(19): 190 - 197.

[10]袁培森, 欧阳柳江, 翟肇裕, 等. 基于 MobileNet v3Small - ECA 的水稻病害轻量级识别研究[J]. 农业机械学报, 2024, 55(1): 253 - 262.

[11]崔金荣, 魏文钊, 赵 敏. 基于改进 MobileNet v3 的水稻病害识别模型[J]. 农业机械学报, 2023, 54(11): 217 - 224, 276.

[12]卫雅娜, 王志彬, 乔晓军, 等. 基于注意力机制与 EfficientNet 的轻量化水稻病害识别方法[J]. 中国农机化学报, 2022, 43(11): 172 - 181.

[13]Liu Z, Mao H Z, Wu C Y, et al. A ConvNet for the 2020s[C]//

曾龙军,卢承方,陈勇,等. 基于 YOLO-FBA 的复杂环境下豆叶病害检测研究[J]. 江苏农业科学,2025,53(20):315-324.
doi:10.15889/j.issn.1002-1302.2025.20.035

基于 YOLO-FBA 的复杂环境下豆叶病害检测研究

曾龙军,卢承方,陈勇,崔艳荣,尹利华

(长江大学计算机科学学院,湖北荆州 434000)

摘要:针对大豆叶片病害在复杂环境下存在背景干扰难以准确识别的问题,提出一种基于改进 YOLO v8n 目标检测模型 YOLO-FBA。在 Backbone 端引入 FasterNet 模型主干,减少数据特征的冗余,有效提升模型的病害特征提取能力。在 Neck 端引入并重新设计 BiFPN 特征融合网络,显著降低模型的参数量和计算量。最后在 BiFPN 基础上引入语义与细节注入模块(SDI),并设计自适应多尺度融合模块(AMI)改进特征融合方式,以增强模型表征能力。结果表明,相较于基线模型 YOLO v8n,改进后模型的参数量为 2.77 M,计算量为 7.6 GFLOPs,两者分别下降 10.1%、6.2%;在 $mAP@0.5$ 方面,相较于基线模型提升 2.0 个百分点;与 YOLO v3、YOLO v5、YOLO v6、YOLO v7、YOLO v8n 目标检测算法相比,YOLO-FBA 的 $mAP@0.5$ 分别提高 10.3、3.5、1.7、2.4、2.0 百分点。本研究提出的大豆叶片病害识别方法在使用更少的参数和计算量的情况下达到了更高的精度,可为复杂环境下的大豆叶片病害检测提供更可行的解决方案,并为真实田间环境下大豆叶片病害检测系统的开发奠定基础。

关键词:大豆叶片病害;深度学习;目标检测;YOLO v8n;病害检测

中图分类号:S126;TP391.41 **文献标志码:**A **文章编号:**1002-1302(2025)20-0315-10

大豆是一种被广泛种植的植物,既是重要的油料作物,也是蛋白质的重要来源^[1]。作为我国重要

的农产品之一,大豆的产量和质量直接关系到粮食安全。在大豆种植过程中,因管理不善,病虫害频繁发生,最终对大豆的品质和产量产生了显著影响^[2]。病虫害造成的大豆产量损失通常在 15%~30% 之间,严重时甚至可能导致绝收^[3]。传统的病害识别和检测方法通常需要植保专家在现场进行人工观察和诊断,这种方法不仅效率较低,且会因人类主观判断及个人经验存在差异,对某些病害的判断结果不尽相同^[4]。此外,一些病害在发生初期

收稿日期:2024-10-27

基金项目:国家自然科学基金面上项目(编号:62077018)。

作者简介:曾龙军(2000—),男,湖南邵阳人,硕士研究生,主要研究方向为深度学习与目标检测。E-mail:2023710698@yangtzeu.edu.cn。

通信作者:陈勇,硕士生导师,高级工程师,从事 WEB 信息处理、人工智能应用研究。E-mail:285527563@qq.com。

IEEE. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, 2022:11966-11976.

[14] Misra D, Nalamada T, Arasanipalai A U, et al. Rotate to attend: convolutional triplet attention module [C]//2021 IEEE Winter Conference on Applications of Computer Vision (WACV). Waikoloa, 2021:3138-3147.

[15] Szegedy C, Liu W, Jia Y Q, et al. Going deeper with convolutions [EB/OL]. (2014-09-17) [2024-07-10]. <https://arxiv.org/pdf/1409.4842>.

[16] He K M, Zhang X Y, Ren S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9):1904-1916.

[17] Liu S T, Huang D, Wang Y H. Receptive field block net for accurate and fast object detection [M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer Vision - ECCV 2018. Cham;

Springer International Publishing, 2018:404-419.

[18] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [EB/OL]. (2015-12-10) [2024-07-10]. <https://arxiv.org/pdf/1512.03385>.

[19] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [J]. CoRR, 2014, abs/1409.1556.

[20] Han K, Wang Y H, Tian Q, et al. GhostNet: more features from cheap operations [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, 2020:1577-1586.

[21] Tan M X, Le Q V. EfficientNet v2: smaller models and faster training [EB/OL]. (2021-04-01) [2024-07-10]. <https://arxiv.org/pdf/2104.00298v1>.

[22] Howard A, Sandler M, Chu G, et al. Searching for MobileNet v3 [EB/OL]. (2019-11-20) [2024-07-10]. <https://www.semanticscholar.org/reader/5e19eba1e66447c83f607383d256deea71f87ae>.