

李康顺, 杨振盛, 江梓锋, 等. 基于改进 YOLOX-Nano 的农作物叶片病害检测与识别方法 [J]. 华南农业大学学报, 2023, 44(4): 593-603.
LI Kangshun, YANG Zhensheng, JIANG Zifeng, et al. Detection and identification of crop leaf diseases based on improved YOLOX-Nano[J]. Journal of South China Agricultural University, 2023, 44(4): 593-603.

基于改进 YOLOX-Nano 的农作物叶片 病害检测与识别方法

李康顺^{1,2}, 杨振盛¹, 江梓锋¹, 王健聪¹, 王 慧³

(1 华南农业大学 数学与信息学院, 广东 广州 510642; 2 东莞城市学院 人工智能学院, 广东 东莞 523419;

3 深圳信息职业技术学院, 广东 深圳 518172)

摘要:【目的】实现精确迅速的农作物病害检测, 减少人工诊断成本, 降低病害带来的农作物产量和品质影响。【方法】根据对农作物病害和病斑特征的分析, 提出一种基于卷积注意力机制改进的 YOLOX-Nano 智能检测与识别模型, 该模型采用 CSPDarkNet 作为主干网络, 将卷积注意力模块 CBAM 引入到 YOLOX-Nano 网络结构的特征金字塔 (Feature pyramid network, FPN) 中, 并在训练中引入 Mixup 数据增强方式, 同时将分类的损失函数由二分类交叉熵损失函数 (Binary cross entropy loss, BCE Loss) 替换为焦点损失函数 Focal Loss、回归损失函数由 GIOU Loss 替换为本文设计的 CenterIOU Loss 函数, 采用迁移学习策略训练改进的 YOLOX-Nano 模型, 以此提升农作物病害检测的精度。【结果】改进后的 YOLOX-Nano 模型仅有 0.98×10^6 的参数量, 在移动端测试单张图片检测时间约为 0.187 s, 平均识别精度达到 99.56%。实践结果表明, 其能快速有效地检测与识别苹果、玉米、葡萄、草莓、马铃薯和番茄等农作物的常见病害, 且达到了精度与速度的平衡。【结论】改进后的模型不仅对农作物叶片病害识别具有较高的精度和较快的检测速度, 参数量和计算量较少, 还易于部署在手机等移动设备。该模型实现了在田间复杂环境对多种农作物病害精准定位与识别, 对于指导早期农作物病害的防治具有重要的现实意义。

关键词: YOLOX-Nano 网络; 病害识别; Focal Loss; 注意力机制; 农作物病害

中图分类号: TP391.41; S435.672

文献标志码: A

文章编号: 1001-411X(2023)04-0593-11

Detection and identification of crop leaf diseases based on improved YOLOX-Nano

LI Kangshun^{1,2}, YANG Zhensheng¹, JIANG Zifeng¹, WANG Jiancong¹, WANG Hui³

(1 College of Mathematics and Informatics, South China Agricultural University, Guangzhou 510642, China;

2 School of Artificial Intelligence, Dongguan City College, Dongguan 523419, China;

3 Shenzhen Institute of Information Technology, Shenzhen 518172, China)

Abstract: 【Objective】To identify crop diseases accurately and quickly, reduce the cost of artificial diagnosis, and reduce the impacts of crop diseases on crop yield and quality. 【Method】Based on the analysis of the characteristics of crop diseases and spots, an improved YOLOX-Nano intelligent detection and recognition model based on convolution attention mechanism was proposed. The model employed CSPDarkNet as the

收稿日期: 2022-07-26 网络首发时间: 2023-04-07 11:05:24

首发网址: <https://kns.cnki.net/kcms/detail/44.1110.S.20230406.1811.002.html>

作者简介: 李康顺, 教授, 博士, 主要从事计算机视觉、智能计算研究, E-mail: likangshun@sina.com

基金项目: 国家自然科学基金 (61573157); 广东省基础与应用基础研究基金 (2022A1515011447); 广东省教育厅重点领域专项 (2021ZDZX1029)

backbone network, added convolutional attention module CBAM to the feature pyramid network (FPN) of the YOLOX-Nano network structure, and then introduced the mixup data enhancement method in the training. At the same time, the classification loss function was replaced by the binary cross entropy loss function (BCE Loss) with the focus loss function, the regression loss function of GIOU Loss was replaced by the CenterIOU Loss function designed in this paper, and a transfer learning strategy was also used to train the modified YOLOX-Nano model so as to improve the accuracy of crop disease detection. 【Result】 The improved YOLOX-Nano model had parameters of 0.98×10^6 , and the detection time of a single sheet was about 0.187 s at the mobile end, with a mean average precision of 99.56%. The practical results of introducing this method into mobile terminal deployment showed that it could quickly and effectively identify common diseases of crops such as apples, corns, grapes, strawberries, potatoes and tomatoes, and achieve the balance of accuracy and speed. 【Conclusion】 The improved model not only has higher accuracy and detection speed for crop leaf disease identification, but also has less parameters and calculation amount. The model was easy to be deployed on mobile devices such as mobile phones. In addition, the model achieves accurate positioning and identification of a variety of crop diseases in complex field environment, which is of great practical significance to guide the prevention and control of early crop diseases.

Key words: YOLOX-Nano network; Disease identification; Focal Loss; Attention mechanism; Crop disease

中国是一个农业大国，作为国家发展的根本，农业在国家经济总量中占有很大的比重。如今全球粮食供给偏紧，农业是国家战略性、基础性核心产业，在农作物的生长发育过程中，气候环境、病虫害、病菌侵染等严重影响其产量和品质。目前常见的农作物病害有上千种，传统的依靠人工识别农作物病害，不但耗费时间精力，而且效率低。精确迅速地辨识出农作物病害有助于及时救助，对降低农作物病害带来的产量和品质影响具有重大意义。

随着深度学习技术在物体分类方面的深入研究与计算机图像处理能力的增强，计算机视觉技术也应用在农作物病害识别上，并且取得许多成果。刘翱宇等^[1]提出一种基于深度残差网络的植物病害检测网络，引入 Focal Loss 损失函数，对植物病害图像的检测平均准确率高达 97.96%。Srdjan 等^[2]将深度卷积神经网络用于植物病害识别，能够区分植物叶片与周围环境，并能够区分健康叶片与 13 种不同的植物病害，在开发模型测试中，单类病害的识别平均准确率高达 96.3%。Albattah 等^[3]提出了一种改进的 CenterNet 网络，以 DenseNet-77 为基础网络对病害样本图片特征进行提取，在 Plant Village 数据集上训练，然后分别对番茄、葡萄等在内的 14 种植物 26 类病害及健康叶片进行识别，检测精度比当时精确率最高的 EfficientNet 网络更高。刘洋等^[4]针对 PlantVillage 数据集 14 种植物中的 26 种病害对比研究了 2 个轻量级网络系统 MobileNet 和 Inceptionv3，结果表明 MobileNet 模

型的程序数量更小，执行速率也更快，可部署于安卓操作系统中，在手机端也实现了图像识别。Saleem 等^[5]对 Xception 网络精修后，选择性能提升最明显的 Adam 优化器，对 PlantVillage 数据集中的 26 种植物病害进行检测，准确率达到了 97.81%。孙俊等^[6]使用 AlexNet 作为基础网络，然后通过批归一化方法、引入全局池化层和减少特征图量等方式，共获得了 8 种改进模式，对 PlantVillage 数据集 14 种植物的 26 种病害特征进行了鉴别，效果最好的检测模式平均准确度超过了 98.1%。李书琴等^[7]提出了一种基于轻量级残差网络，通过缩减网络卷积核数目和轻量级残差模块减少网络参数，网络参数大幅减少的同时，在 PlantVillage 数据集上的检测精度达到 98.45%。

上述深度学习方法不仅在检测准确度方面有了较高的提升，也有向低内存、少参数量改进的趋势，从而达到高精度识别、迅速检测、易部署移动端的目的。然而，检测精度提升的同时，网络参数量也在不断增多，网络深度和宽度也不断加深，浮点运算量加大从而带来较高的延迟，实时检测的效果较差。人工智能公司旷视科技提出的 YOLOX-Nano^[8]，有着参数量小、浮点运算快、延迟低、易部署移动端的优势且检测精度保持着中高水平。本文在原始 YOLOX-Nano 网络结构的基础上进行改进，在不显著增加网络内存和运算参数数量的基础上，提出一种混合卷积注意力模块 CBAM^[9]、Mixup 数据增强策略^[10]、Focal Loss 分类损失函数^[11]和本

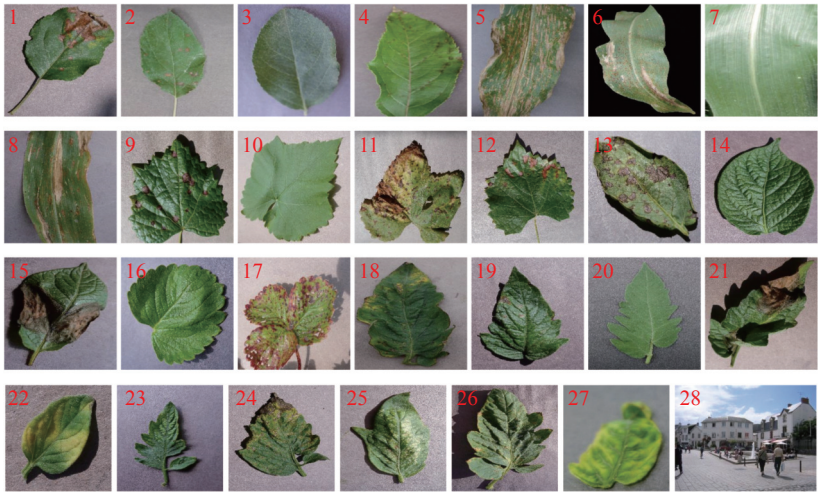
文设计的基于 CenterIOU 回归损失函数的改进 YOLOX-Nano 模型。

1 材料与方法

1.1 数据集的获取

试验所需要的所有病害图片数据来自于公开数据集 PlantVillage 网站 (<http://plantvillage.psu.edu>), 数据集记录包含 54 309 张图像, 这些图像涵盖了 14 种作物: 苹果、蓝莓、樱桃、玉米、葡萄、柑橘、桃、甜椒、马铃薯、覆盆子、大豆、南瓜、草莓和番茄。考虑到应尽可能研究某一类农作物的多种病害 (3 种及以上), 因此剔除数据集内仅含健康叶片的农作物数据和仅有 2 类病害的农作物数据, 最终在公开数据集上精心挑选苹果叶片图像 4 645 张, 其中, 苹果黑星病 1 000 张、苹果黑腐病 1 000 张、苹果松锈病 1 000 张、健康叶片 1 645 张; 玉米叶片

4354 张, 其中, 玉米灰斑病 1 000 张、玉米锈病 1 192 张、玉米大斑病 1 000 张、健康叶片 1 162 张; 葡萄叶片 4 369 张, 其中, 葡萄黑腐病 1 180 张、葡萄轮斑病 1 076 张、葡萄褐斑病 1 383 张、健康叶片 1 000 张; 马铃薯叶片 3 000 张, 其中, 马铃薯早疫病、马铃薯晚疫病、健康叶片各 1 000 张; 草莓叶片 4 218 张, 其中, 草莓叶焦病 2 218 张, 健康叶片为 2 000 张; 番茄叶片 18 835 张, 其中, 番茄疮痂病 1 404 张、番茄早疫病 1 000 张、番茄晚疫病 1 909 张、番茄叶霉病 1 000 张、番茄斑枯病 1 771 张、番茄红蜘蛛损伤 1 676 张、番茄斑点病 2 127 张、番茄黄叶曲叶病 5 357 张、番茄花叶病毒病 1 000 张和健康叶片 1 591 张, 共计 6 类农作物 27 种病害, 保证各类样本间数量均衡。此外, 考虑到真实场景背景复杂的情况, 数据集添加了无叶片背景图像 1 143 张。试验样本图像如图 1 所示。



1~4 依次为苹果黑星病、黑腐病、松锈病和健康叶; 5~8 依次为玉米灰斑病、锈病、健康叶和大斑病; 9~12 依次为葡萄黑腐病、轮斑病、褐斑病和健康叶; 13~15 依次为马铃薯早疫病、晚疫病和健康叶; 16~17 依次为草莓叶焦病和健康叶; 18~27 依次为番茄疮痂病、早疫病、晚疫病、叶霉病、斑枯病、红蜘蛛损伤、斑点病、黄叶曲叶病、花叶病毒病和健康叶; 28 为无叶片背景图

1~4 are apple scab, apple black rot, cedar apple rust and apple healthy leaves in turn; 5~8 are corn *Cercospora zeaemaydis* Tehon and Daniels, corn *Puccinia polysora*, corn healthy leaves and corn *Curvularia* leaf spot fungus; 9~12 are grape black rot, grape black measles, grape leaf blight fungus and grape healthy leaves in turn; 13~15 are potato early blight, potato late blight and potato healthy leaves in turn; 16~17 are strawberry leaf scorch and healthy leaves in turn; 18~27 are tomato bacterial spot bacteria, tomato early blight, tomato late blight water mold, tomato leaf mold, tomato septorial leaf spot, tomato spider mite damage, tomato target spot bacteria, tomato YLCV virus, tomato ToMV disease and tomato healthy leaves in turn; 28 is bladeless background image

图 1 PlantVillage 部分数据集
Fig. 1 Partial dataset of PlantVillage

1.2 数据预处理网络训练参数设置

首先将所有样本图像的分辨率转变成 416×416, 使用标注工具 LabelImg 将所有样本按 Pascal VOC 数据集的标注, 生成.xml 类型的标注文件。此外, 从试验数据集图像中随机选择 60% 样本 (约 23235 张) 用于训练模型, 20% 的样本 (约 7 675 张) 用来验证, 20% 样本 (约 7 675 张) 用于测试。

试验运用迁移学习, 让 YOLOX-Nano 在大规模公开数据集 Image Net 上进行训练, 获得一个收

敛的预训练权重, 然后利用这个预训练权重迁移到改进的 YOLOX-Nano 网络模型中进行参数调优。训练过程中超参数设置为每批量样本数为 32, 1 次遍历完全部训练集数据称为 1 次迭代, 批量设置为 100。采用随机梯度下降 (Stochastic gradient descent, SGD) 优化算法调优网络参数, 初始学习率设置为 0.01, 动量因子为 0.9, 每经过 10 次迭代训练, 将学习率降低 1/10, 模型迭代 1 次保存 1 次权重。

1.3 基于改进 YOLOX-Nano 的农作物叶片病害检测与识别

1.3.1 YOLOX-Nano 网络 YOLOX-Nano 作为 YOLOX 系列的最轻量化版本, 对比 YOLOX-X、YOLOX-L、YOLOX-M、YOLOX-S、YOLOX-Darknet 和 YOLOX-Tiny 高性能版本, 其通过降低网络的宽度和深度、减少模型运算的参数数量、取消训练时 Mixup 数据增强方式等改进, 使网络结构更

简单, 速度更快。

YOLOX-Nano 的结构如图 2a 所示, 整个 YOLOX-Nano 可以分为 3 个部分, 分别是 CSPDarkNet 主干提取网络^[12]、特征金字塔网络 (Feature pyramid network, FPN)^[13] 和路径聚合网络 (Path aggregation network, PAN)^[14] 中间层以及 YoloHead 输出层。

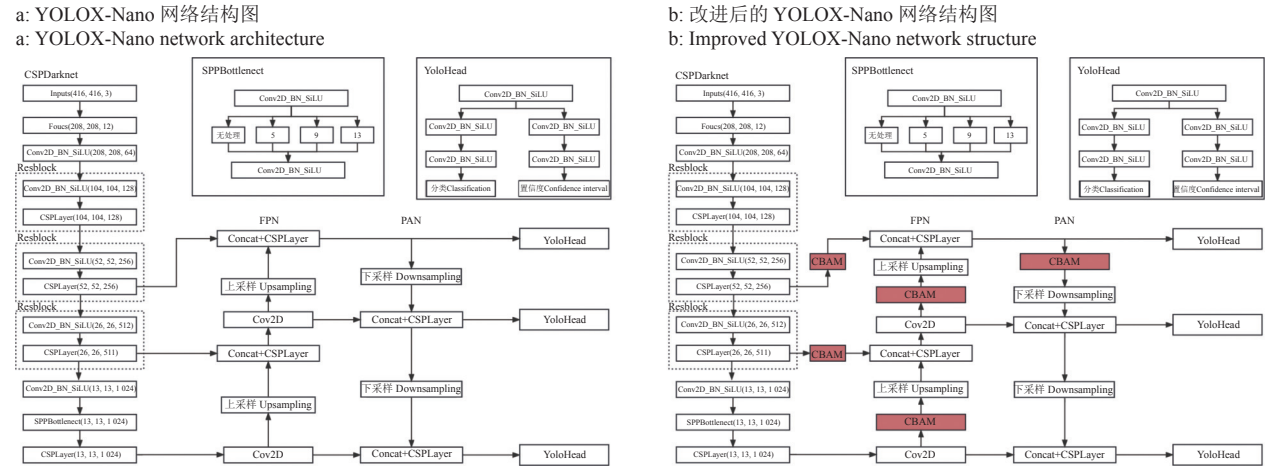


图 2 YOLOX-Nano 网络结构和改进后的 YOLOX-Nano 网络结构

Fig. 2 Network structure of YOLOX-Nano and improved YOLOX-Nano

输入的图片首先在 CSPDarkNet 进行特征提取, 提取到 3 个特征层进行下一步网络的构建, FPN 是 YOLOX-Nano 的加强特征提取网络, 在主干部分获得的 3 个特征层会在这一部分进行特征融合, 获取不同尺度的特征信息。在 FPN 部分, 已经获得的有效特征层被用于继续提取特征, 在 PAN 层中, 网络会对特征进行阶段上采样和下采样操作实现特征融合。YoloHead 是 YOLOX-Nano 的分类器与回归器, 输入图片通过 CSPDarkNet、FPN 和 PAN 层已经获得了 3 个加强特征层, 每一个特征图看作每一个特征点的集合, 而 YoloHead 会对特征点进行判断, 判断特征点是否有物体与其对应。YOLOX-Tiny 的 YoloHead 被分为了 2 部分, 分别实现检测和分类操作, 最后在预测阶段将两者整合输出结果。

由于 YOLOX-Nano 宽度和深度较浅, 没有充分学习输入特征向量通道与空间之间的重要程度关系, 对近似的难分辨目标预测效果较差, 其次数据集内多种农作物病害图像样本复杂多样, 二分类交叉熵损失 (Binary cross entropy loss, BCELoss) 作为分类损失函数将困难样本和容易样本赋予相同的权重进行训练, 这样造成一些困难样本分类效果不理想, GIOU 损失函数对 2 个候选框没有相交时,

IOU=0, 不能反映两者的距离大小, 也无法精确反映两者的重合度大小, 不利于梯度回传和学习训练, 致使密集病害定位缺失。

综合以上问题, 本文在 YOLOX-Nano 模型混合了卷积注意力 CBAM 模块, 引入了 Mixup 数据增强方式来充实样本集, 让网络训练地更充分, 最后分类损失函数将 BCELoss 替换为 Focal Loss, 回归损失函数将 GIOU Loss 替换为 CenterIOU Loss, 改进后的 YOLOX-Nano 网络结构图如图 2b 所示。

1.3.2 混合卷积注意力机制 CBAM 模块 在农作物病害检测与识别任务中, 农作物叶片病害特征大体相似, 只有细节部分略有差异。YOLOX-Nano 作为浅层模型, 检测精度不够高, 这是需要注意的地方。

注意力机制是运用在深度学习模型中常见的小技巧, 是实现网络自适应注意的一个方式, 核心重点是让网络关注到它需要关注的特征, 忽略不重要的特征, 主要分为通道注意力机制和空间注意力机制, 卷积注意力机制结构如图 3 所示。卷积注意力机制的前半部分为通道注意力机制, 通道注意力机制先针对输入进来的单个特征层, 依次执行全局平均池化方法和全局最大池化方法, 之后再针对平均池化和最大池化的结果, 利用共享

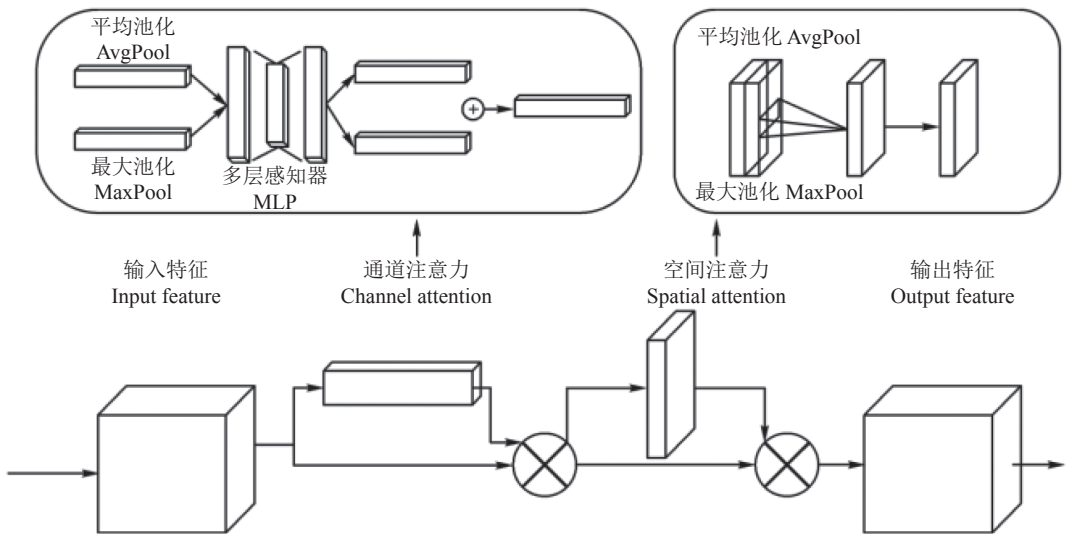


图 3 CBAM 卷积注意力机制结构

Fig. 3 Structure of CBAM convolutional attention mechanism

的全连接层进行处理, 然后再对处理后的 2 个结果进行相加, 之后再通过 Sigmoid 激活函数, 得到输入特征层中每一条通道的加权值 (0~1)。将这个权值乘上原输入特征层即可得到经过通道注意力机制处理后的特征图。卷积注意力机制的后半部分为空间注意力机制, 首先对通过通道注意力机制处理的特征层每一个特征点的通道取最大值和平均值, 之后将这 2 个结果进行堆叠, 利用通道数为 1 的卷积核卷积调整通道数, 再经过 Sigmoid 激活函数, 获得输入特征层每一个特征点的权值 (0~1), 然后将这个权值乘上原输入特征层即可得到最终处理的特征层。

本文引入 CBAM 注意力机制模块, 结合通道注意力机制和空间注意力机制, 对输入的特征层分别进行通道注意力机制和空间注意力机制的处理, 卷积注意力机制对不同通道的特征图进行权重的重新分配, 加强网络结构深层信息提取, 让网络关注到它需要关注的特征, 忽略不重要的特征。

输入图像经过主干特征提取网络后获得 3 个有效特征层, 在构建 FPN 层上采样前首先将经由加强特征提取网络的 3 个有效特征层送入 CBAM 模块, 自动学习特征通道之间的相关性, 输出 3 个尺度的检测结果, 再经过卷积层提取特征。随后进行 PAN 的构建, 在 PAN 融合的下采样前, 把主干特征网络获得的 3 个有效特征层, 再经过 CBAM 注意力机制模块进行融合, 然后经过另一个卷积模块再一次提取特征, 将最后一个有效特征层卷积, 最终获得 3 个特征。

1.3.3 引入 Mixup 数据增强策略 Mixup 是一种混类增强的数据增强策略, 它可以将不同类的图像

进行混合, 从而扩充训练数据集。Mixup 数据增强策略原理是假设 batch_{x1} 是一个批次的图像样本, batch_{y1} 是该批次图像样本对应的标签; batch_{x2} 是另一个批次图像样本, batch_{y2} 是另一个批次图像样本对应的标签, λ 是由参数为 α, β 的 Beta 分布计算出来的混合系数。

由此可以得到 Mixup 原理公式为

$$\lambda = \text{Beta}(\alpha, \beta), \tag{1}$$
$$\text{mixed_batch}_x = \lambda \times \text{batch}_{x1} + (1 - \lambda) \times \text{batch}_{x2}, \tag{2}$$
$$\text{mixed_batch}_y = \lambda \times \text{batch}_{y1} + (1 - \lambda) \times \text{batch}_{y2}. \tag{3}$$

以玉米病害图像为例, 在训练时先读取 1 张样本图像, 图像两侧填充, 放大图片到 416×416 (图 4a), 再随机选取 1 张样本图像, 也放大到 416×416 (图 4b)。然后设置 1 个融合系数, 如 0.5, 将 2 张图像加权融合, 最终得到完整的图像 (图 4c)。

在 YOLOX 中以 Mosaic 为基础, Mixup 策略作为一种额外的数据增强策略, 在浅层模型用 Mixup 数据增强策略会降低平均精确率 (Average precision, AP), 所以 YOLOX-Nano 没有使用 Mixup 数据增强策略。本文使用 Mixup 数据增强策略, 加大样本数据量防止模型欠拟合。根据图 5 中的 Beta 分布概率密度曲线, 当 $\alpha = \beta = 1$ 时, 等于 (0, 1) 均匀分布; 当 $\alpha = \beta < 1$ 时, 表现为两头的概率大, 中间的概率小, 当 $\alpha = \beta \rightarrow 0$ 时, 相当于 $\{0, 1\}$ 二项分布, 要么取 0, 要么取 1, 等于原始数据没有增强, 本试验随机设置 Mixup 参数 $\alpha = \beta \in [0.1, 0.9]$, 就可以得到多样化的 $\lambda \in (0, 1)$ 区间内的概率分布, 使数据增强更具有随机性, 模型更具有鲁棒性。

1.3.4 分类损失函数的改进 一个好的损失函数

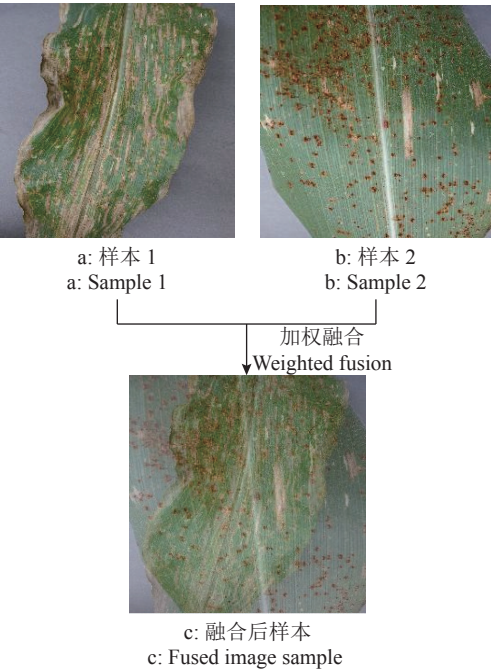


图 4 Mixup 数据增强策略示意图 (以玉米灰斑病为例)

Fig. 4 Schematic diagram of mixup data enhancement strategy(Take corn *Cercospora zeae-maydis* as an example)

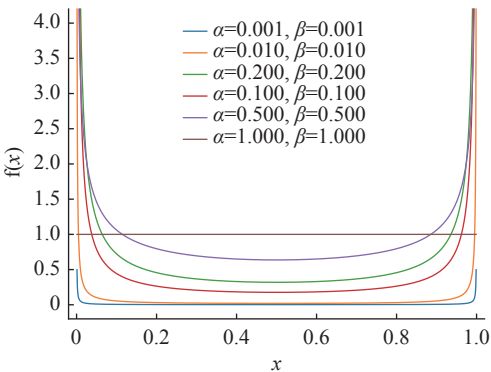


图 5 α 和 β 相同时的 Beta 分布概率密度曲线

Fig. 5 Probability density curve of beta distribution while α and β were equal

对于网络参数调优达到事半功倍的效果。由于多种农作物叶片病害纹理特征大致形似，模型对一些病害样本计算的分类置信度产生偏差，进而影响样本分类效果，造成错误分类病害的情况。

基于以上计算样本分类置信度偏差的问题，YOLOX-Nano 将分类损失函数由 BCE Loss 损失函数替换为 Focal Loss 损失函数，其能控制容易分类和难分类样本的权重，通过减少易分类样本的权重，使得模型在训练时更专注于难分类的样本，让容易样本对损失产生贡献小，难样本对损失产生贡献大。

Focal Loss 分类损失函数公式定义如式 (4) 所示， $\hat{p}[\hat{p} \in (0,1)]$ 表示预测值， y 表示真实值， $\gamma(\gamma \geq$

0) 表示调节因子。设 p_t 表示预测某真实标签的概率，定义如式 (5) 所示，范围为 $[0,1]$ 。由式 (4)、(5) 归纳后 Focal Loss (FL) 损失函数公式如式 (6) 所示：

$$f(x)=\begin{cases} -(1-\hat{p})^\gamma\ln(\hat{p}),\text{if }y=1\\ -\hat{p}^\gamma\ln\hat{p},\text{if }y=0 \end{cases}, \tag{4}$$

$$f(x)=\begin{cases} \hat{p},\text{if }y=1\\ 1-\hat{p},\text{其他} \end{cases}, \tag{5}$$

$$FL(p_t)=- (1-p_t)^\gamma\ln p_t。 \tag{6}$$

与交叉熵函数损失函数相比，公式 (6) 多了一个 $(1-p_t)^\gamma$ 系数，实质是对交叉熵损失函数的优化。如图 6 所示，可以通过改变 γ 调节因子的值，从而控制容易分类和难分类样本对损失的贡献。由公式 (6) 可知，若某样本类别概率 p_t 趋于 1，说明预测值 $\hat{p}$ 接近于真实值 y ，这类样本就属于容易区分的样本，此时系数 $(1-p_t)^\gamma$ 趋于 0，对损失贡献极少。

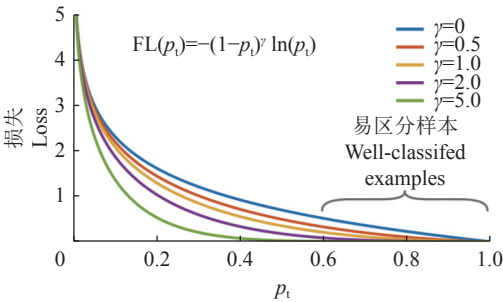


图 6 Focal Loss 函数曲线

Fig. 6 Focal Loss function curve

1.3.5 回归损失函数的改进 YOLOX-Nano 在回归候选框时仍使用 GIOU 损失函数，本文根据数据集图片样本特点，设计出利于回归候选框的 CenterIOU 损失函数。

如图 7 所示，已知候选框中心点 (x_{pre}, y_{pre}) 、真实框中心点 (x_{tag}, y_{tag}) 与最小外接矩形中心点 (x_{ct}, y_{ct}) 。根据公式 (7) 计算候选框和真实框的 IOU，式中， $BOX_{交集}$ 、 $BOX_{并集}$ 分别表示候选框和真实框的交集和并集；构造候选框中心点和真实框中心点两点的一般直线方程 $Y = Ax + By + C$ ；根据公式 (8) 计算最小外接矩形中心点 (x_{ct}, y_{ct}) 到两框中心点的直线距离 (Distance)；最后代入公式 (9)，得到候选框回归损失。

$$IOU=\frac{BOX_{交集}}{BOX_{并集}}, \tag{7}$$

$$Distance=\frac{|Ax_0+By_0+C|}{\sqrt{(A^2+B^2)}}, \tag{8}$$

$$CenterIOU_Loss=1-(IOU-Distance)。 \tag{9}$$

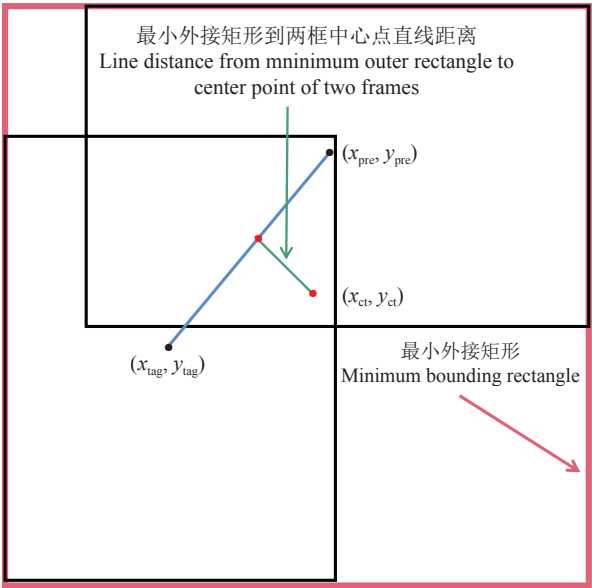


图 7 CenterIOU 损失函数图解

Fig. 7 Diagram of CenterIOU loss function

由于定位损失函数更换为 CenterIOU 损失函数, 较原 GIOU 函数让网络学习到预测框和目标框的相对位置, 网络在反向传播时能更好地调优回归参数。

1.4 病害检测系统的设计

为获取实际状态下病害检测所需时间, 对比网络改进前后的效果, 以训练好的模型为基础, 开发手机端病害识别系统。利用 IOS 官方的 Core ML Tools 模型转换库将训练好的 pth 格式模型转换为 IOS 平台使用的格式。前端界面提供图库、识别、保存 3 个按钮, 对应调用相册上传识别图片、快速识别作物病害、保存识别记录 3 种功能。编译后生成安装文件部署在 iPhone 8 手机, 显示如图 8 所示。



图 8 病害检测系统界面

Fig. 8 Interface of disease detection system

1.5 评价指标

对于农作物叶片病害检测, 需要网络具有高精确率和实时性。本文采用 COCO mAP@[0.50:0.05:0.95]^[15-16] 作为农作物病害检测评价指标来综合评估模型对苹果、玉米、苹果、玉米、葡萄、马铃薯、草莓、番茄 6 类农作物的 27 种病害和无叶片背景图像的检测性能, 即分别在 IOU 阈值设置以 0.05 的步距从 0.50 开始遍历至 0.95, 取这些阈值条件下评估所得到的平均准确率 (Average precision, AP) 及其均值 (Mean average precision, mAP)。mAP 与准确率 (Precision, P)、召回率 (Recall, R) 有关, 其计算公式如下所示:

$$P = \frac{TP}{TP + AP} \times 100\%, \tag{10}$$

$$R = \frac{TP}{TP + FN} \times 100\%, \tag{11}$$

$$AP = \int_0^1 P(R) dR, \tag{12}$$

$$mAP = \frac{1}{M} \sum_{k=1}^M AP(k) \times 100\%, \tag{13}$$

式中, TP 表示被模型预测正确的正样本数, FP 为被模型预测错误的正样本数, FN 为被模型预测错误的负样本数, M 为类别总数, AP(k) 为第 k 类平均准确率。

将训练好的模型部署移动端系统后, 检测单张样本图像所消耗的平均时间 (包含程序加载图片所消耗时间和检测时间) 作为实时性指标, 单位为 s。参数量指标指深度网络模型中可训练参数的个数, 以百万为单位, 表示为 1×10⁶。

2 结果与分析

2.1 改进方法性能分析

使用 YOLOX-Nano 模型为基础网络, 分别引入 CBAM 卷积注意力机制、Mixup 数据增强策略、使用 Focal Loss 替换原分类损失函数 BCELoss, 本文设计 CenterIOU 损失函数替换原回归损失函数 GIOU 作为改进网络。

2.1.1 融合 CBAM 卷积注意力机制 YOLOX-Nano 对农作物病害叶片检测与识别的精度较低, 单张样本图片检测精准率仅为 97.97%, 其原因是 6 种农作物 27 类病害的叶片病态特征表现比较相似。通道注意力可对特征图的不同通道赋予不同特征, 选择性增大不同病害斑点特征通道的权重值; 空间注意力对同一特征图不同位置特征点给予不同权重, 区别对待特征图内部像素点。

引入 CBAM 卷积注意力机制模块后,虽然模型参数量由 0.92×10^6 增加到 0.96×10^6 、单张样本推断时间增加 18 ms,但平均精确率由 97.97% 提升到 98.89%,提升了 0.92 个百分点,在很大程度上提升了对两者叶片病害识别的准确率。两者结合可总体提升识别准确率,表明了 CBAM 模块区别相似物体的有效性。

利用 Grad-CAM^[17] 工具可视化进入 FPN 层前第一个 CBAM 卷积注意力机制^[18],明确对特征提取

的效果。由于玉米病害的样本特征更为接近,可视化效果更佳,本文以玉米灰斑病样本、玉米锈病样本、玉米大斑病样本和健康玉米样本为例进行展示,如图 9 所示,引入卷积注意力机制前,网络对样本特征提取较随机,对叶片病害区域特征点关注不够。引入 CBAM 卷积注意力机制后,网络在前向传播过程中,重要特征通道逐渐占有更大的比重,能让网络学习到需重点关注的部分,使改进的 YOLOX-Nano 模型可以更高效地提取图像难区分的特征信息。

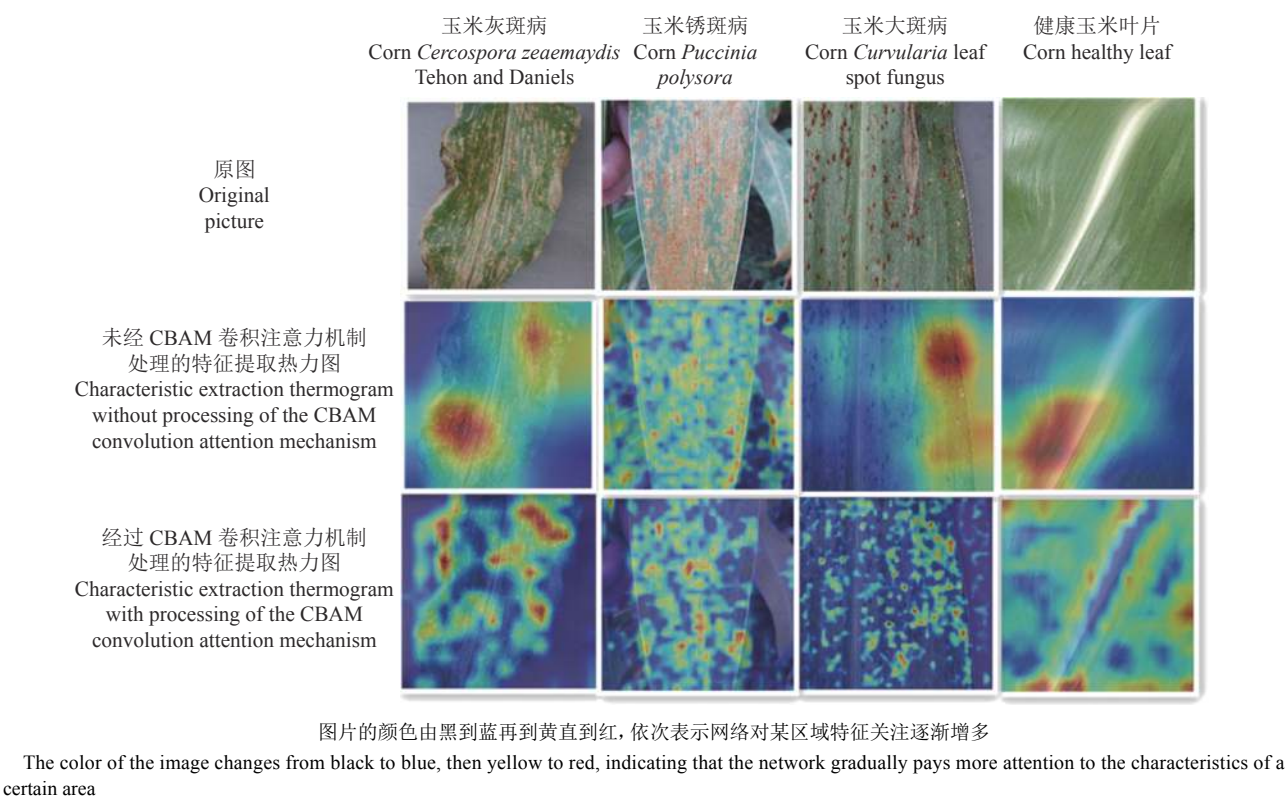


图 9 Grad-CAM 可视化添加 CBAM 前后玉米叶片特征提取的热力图

Fig. 9 Grad-CAM visual characteristic extraction thermal diagram of corn leaf before and after adding CBAM

2.1.1.2 训练引入 Mixup 数据增强策略 使用 Mixup 数据增强策略训练的网络模型单张样本图片检测时间较改进前平均减少了 4 ms,其检测准确率却增加了 1.02 个百分点,一方面减少了检测推断所用时间,另一方面可以提高对农作物病害检测精度,可见引入 Mixup 数据增强策略扩充数据集可以提升农作物病害检测的效率,模型可以训练更充分,减少推断时间。

2.1.1.3 引入 Focal Loss 损失函数试验 由式 (6) 可知,调节因子 γ 可以控制样本对损失的贡献,从而影响到病害分类的性能。试验中通过调节 γ 值来检测病害分类的效果,试验结果表明, γ 取 0 时 mAP 为 97.97%, γ 取 1 时 mAP 为 98.69%, γ 取 2 时 mAP 为 99.02%, γ 取 3 时 mAP 为 98.42%, γ 取 4 时 mAP 为 97.32%,mAP 随 γ 值的增大呈现先升高后降低的趋势,当 γ 设置为 2 时,mAP 最大,检测效果最好。

图 10a、10b 是选取 γ 为 2 时 Focal Loss 损失函数和原 BCELoss 损失函数在训练和测试阶段的分类损失的变化曲线,可以发现,改进后的 YOLOX-Nano 分类损失总体趋势都是低于 YOLOX-Nano 的。

改进后的 YOLOX-Nano 网络训练到 100 个迭代结束,训练和测试阶段均呈现出持续下降的趋势,且还存在下降空间,说明网络仍在学习,分类精度仍有上升的可能性,而曲线逐渐趋于稳定,说明网络过拟合已经没有学习必要。

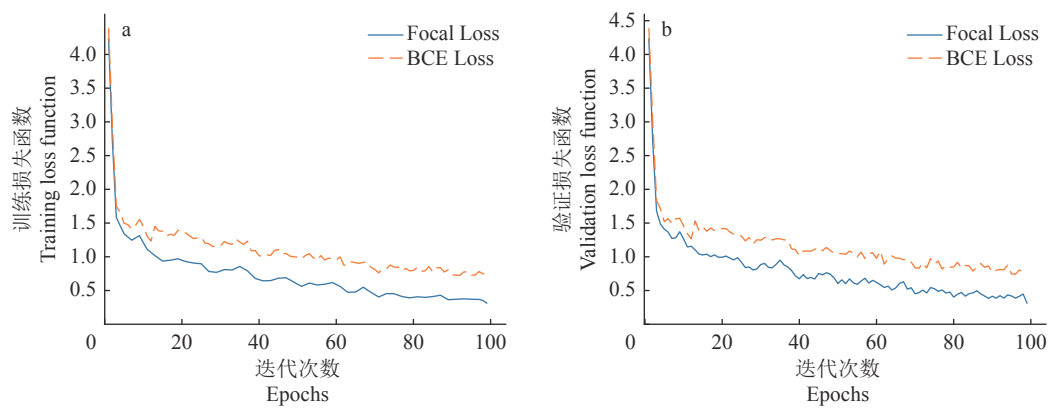


图 10 训练阶段 (a) 和测试阶段 (b) 分类损失变化曲线图
Fig. 10 Classification loss curve during training(a) and test(b) phases

2.1.4 引入 CenterIOU 损失函数 不同的回归框损失函数对检测精度带来不同程度的影响。试验通过比较 GIOU 和本文设计的 CenterIOU 损失函数的性能, 来确定最适合的回归损失函数, 最终试验表明使用 GIOU 的 YOLOX-Nano 在 mAP^{50} 、 $mAP^{50:95}$ 的检测精度分别为 98.23% 和 97.97%, 而使用 CenterIOU 损失函数的 YOLOX-Nano 在 mAP^{50} 、 $mAP^{50:95}$ 的检测精度分别为 98.53% 和 98.20%, 较改进前分别提高了 0.30 和 1.23 个百分点, 本文设计的 CenterIOU 损失函数与 YOLOX-Nano 网络组合后, 在数据集上对不同阈值表现得更好, 能够明显地提高农作物病害的检测精度。

2.2 YOLOX-Nano 改进前后总体性能分析

改进后的 YOLOX-Nano 模型参数量虽然增加了 0.06×10^6 , 单张图片检测时间增加了 11 ms, 但是改进后的 YOLOX-Nano 网络对所有农作物病害的平均检测精度达到了 99.56%, 较改进前精度 (97.97%) 提升了 1.59 个百分点。改进后的模型对苹果松锈病、葡萄黑腐病、马铃薯早疫病、马铃薯晚疫病、番茄叶霉病、番茄斑枯病和番茄黄叶曲叶病的识别精度都达到了 100%, 对玉米灰斑病识别精度 (98.66%) 稍低, 主要原因是玉米大斑病与灰斑病病害症状相似, 不易判断, 导致出现错分类现象。

同时, 无叶片背景图这一类别存在背景中包含叶片的情况, 模型能检测出叶片的存在, 进而导致 2 个模型对无叶片背景图识别精度都表现较低, 这也证明了本模型没有出现拟合的现象, 改进的网络有很好的农作物病害检测与识别能力。

图 11 是改进前后网络训练平均精度和训练迭代的效果图, 可见在训练 10 个迭代时, 改进的 YOLOX-Nano 就能比改进前产生整体较高的 mAP, 预训练权重很快适应改进网络, 趋于平稳, 直到 100 个迭代

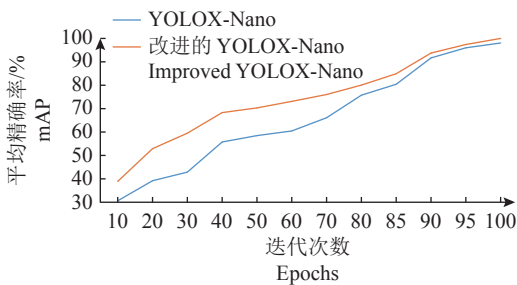


图 11 模型改进前后网络精度随迭代次数变化曲线图
Fig. 11 Curve of network accuracy changing with number of iterations before and after model improvement

训练结束, 最终改进后的 YOLOX-Nano 精度更高。
2.3 不同算法的对比试验

本文还与现阶段较成熟的轻量级模型 ResNet-18、MobileNet-v2、YOLOv4-Tiny、YOLOX-Tiny^[19-24] 进行对比, 试验结果如表 1 所示, 改进后的 YOLOX-Nano 检测准确率较其他轻量级网络高出 1.0~2.2 个百分点。在真机测试的检测时间对比中, 本文提出方法的优势更加明显, 单张样本图片检测时间仅需 0.187 s, 较其他各类模型节省大约 20~240 ms 的计算时间。不仅如此, 在参数量方面, 本文方法较 ResNet-18 减少了 91.3%、较 YOLOv4-Tiny 减少了 83.9%、较 YOLOX-Tiny 减少了 80.7%、较 MobileNet-v2 减少了 72.0%, 更小的参数量意味着更少的部署成本, 因此改进后模型的综合性能非常优秀, 对于农业应用具有明显的优势。

2.4 病害检测系统识别效果

为验证本网络移动端部署可行性和所提出网络的实际分类效果, 从试验数据集中, 随机取出 50 张图片, 录入手机端系统进行识别, 手机端检测部分结果如图 12a、12b 所示。经检测, 所有图片病害全部识别准确, 置信度保持在 99.4% 以上, 最高置信度表现在苹果、葡萄和马铃薯 3 种作物的病害检测, 均保持在 99.8% 以上, 实际运行最多耗时 0.324 s,

表 1 与主流轻量型网络性能对比

Table 1 Performance of the model versus mainstream lightweight network

模型 Model	参数量($\times 10^6$) No. of parameters	平均精确率/% Mean average precision	单张图片检测时间/s Detection time of single image
YOLOX-Nano	0.92	97.97	0.176
ResNet-18	11.24	98.60	0.429
MobileNet-v2	3.40	97.33	0.189
YOLOv4-Tiny	6.06	97.42	0.382
YOLOX-Tiny	5.06	98.58	0.286
改进YOLOX-Nano	0.98	99.56	0.187
Improved YOLOX-Nano			



图 12 病害检测结果展示

Fig. 12 Display of disease detection results

最少耗时 0.169 s, 平均耗时为 0.187 s。此外, 从苹果、玉米、葡萄、马铃薯、草莓、番茄 6 类农作物病

害样本中各筛选 5 张田间拍摄照片进行检测, 部分检测结果如图 12c、12d 所示, 所有病害依然识别准确, 仍保持着 99% 以上的置信度, 识别速度并未减弱, 也无漏检错检的情况。可见本改进模型的检测效果和运行速度都保持着较高的水平, 同时达到了精度和速度的平衡。

3 结论

相对于人工诊断农作物病害, 利用计算机视觉方法检测农作物病害具有成本低、准确率高、时延短的优点。本文通过在 YOLOX-Nano 分类模型中引入空间和通道注意力机制和 Mixup 数据增强策略, 将分类损失函数更换为 Focal Loss, 提出一种改进的 YOLOX-Nano 优化网络, 利用此网络对农作物病害数据集进行训练与测试, 得出以下结论。

- 1) 改进的 YOLOX-Nano 网络对 PlantVillage 数据集中 6 种农作物 27 种病害的检测总体 mAP 达到了 99.56%, 较 YOLO-Nano 基础网络提升了 1.59 个百分点, 单张图片平均检测时间为 0.187 s, 仅仅较基础网络增加 11 ms 的推断延迟, 达到了精度与速度的平衡, 而且网络参数量为 0.98×10^6 , 利于植保无人机设备和移动端设备的部署。
- 2) 卷积注意力机制可以对不同通道的特征图进行权重的重新分配, 加强网络结构深层信息的提取, 对于很多细粒度特征的提取具有很好的效果, 此外 Mixup 数据增强策略对一些小型网络也能起到促进作用; Focal Loss 能控制容易分类和困难分类样本的权重, 通过减少易分类样本的权重, 在训练时更专注于难分类的样本, 让容易样本对损失贡献小, 难样本对损失贡献大。
- 3) 本文提出的 CenterIOU 损失函数对本数据集训练具有促进作用, 也可以推广应用到其他数据集。

参考文献:

[1] 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.

[2] SRDJAN S, MARKO A, ANDRAS A, et al. Deep neural networks based recognition of plant diseases by leaf image classification[J]. *Computational Intelligence and Neuroscience*, 2016(6): 3289801. doi: 10.1155/2016/3289801.

[3] ALBATTAH W, NAWAZ M, JAVED A, et al. A novel deep learning method for detection and classification of plant diseases[J]. *Complex & Intelligent Systems*, 2021, 8(2): 507-524.

[4] 刘洋, 冯全, 王书志. 基于轻量级 CNN 的植物病害识别方法及移动端应用[J]. *农业工程学报*, 2019, 35(17): 194-204.

[5] SALEEM M H, POTGIETER J, ARIF K M. Plant disease classification: A comparative evaluation of convolutional neural networks and deep learning optimizers[J]. *Plants (Basel, Switzerland)*, 2020, 9(10): 1319. doi: 10.3390/plants9101319.

[6] 孙俊, 谭文军, 毛罕平, 等. 基于改进卷积神经网络的多种植物叶片病害识别[J]. *农业工程学报*, 2017, 33(19): 209-215.

[7] 李书琴, 陈聪, 朱彤, 等. 基于轻量级残差网络的植物叶片病害识别[J]. *农业机械学报*, 2022, 53(3): 243-250.

[8] GE Z, LIU S T, WANG F, et al. YOLOX: Exceeding YOLO series in 2021[EB/OL]. arXiv: 2107.08430. 2021: <https://arxiv.org/abs/2107.08430>.

[9] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]//European Conference on Computer Vision. Cham: Springer, 2018: 3-19.

[10] ZHANG H, CISSE M, DAUPHIN Y N, et al. Mixup: Beyond empirical risk minimization[EB/OL]. 2017: arXiv: 1710.09412. <https://arxiv.org/abs/1710.09412>.

[11] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017: 2999-3007.

[12] WANG C Y, MARKLAO H Y, WU Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN[C]//2020 IEEE/CVF Conerence on Computer VSION and Pattern Recognition Workshops(CVPRW). Seattle, WA, USA: IEEE, 2020: 1571-1580.

[13] LIN T, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE, 2017: 936-944.

[14] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA: IEEE, 2018: 8759-8768.

[15] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: Common objects in context[C]//European Conference on Computer Vision. Cham: Springer, 2014: 740-755.

[16] EVERINGHAM M, ALIESLAMI S M, VAN GOOL L, et al. The pascal visual object classes challenge: A retrospective[J]. *International Journal of Computer Vision*, 2015, 111(1): 98-136.

[17] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization[J]. *International Journal of Computer Vision*, 2020, 128(2): 336-359.

[18] 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 (TPAMI)*, 2015, 37(9): 1904-1916.

[19] 王立舒, 秦铭霞, 雷洁雅, 等. 基于改进 YOLOv4-Nano 的蓝莓成熟度识别方法[J]. *农业工程学报*, 2021, 37(18): 170-178.

[20] 赵志成, 罗泽, 王鹏彦, 等. 基于深度残差网络图像分类算法研究综述[J]. *计算机系统应用*, 2020, 28(1): 14-21.

[21] 孙云云, 江朝晖, 董伟, 等. 基于卷积神经网络和小样本的茶树病害图像识别[J]. *江苏农业学报*, 2019, 35(1): 48-55.

[22] 李晓振, 徐岩, 吴作宏, 等. 基于注意力神经网络的番茄叶部病害识别系统[J]. *江苏农业学报*, 2020, 36(3): 561-568.

[23] 王艳玲, 张宏立, 刘庆飞, 等. 基于迁移学习的番茄叶片病害图像分类[J]. *中国农业大学学报*, 2019, 24(6): 124-130.

[24] 邱靖, 刘继荣, 曹志勇, 等. 基于卷积神经网络的水稻病害图像识别研究[J]. *云南农业大学学报(自然科学)*, 2019, 34(5): 884-888.

【责任编辑 霍 欢】