

# 基于改进YOLOv5s对苹果叶片病害的检测识别

许龙\*, 吴丽丽#

甘肃农业大学信息科学技术学院, 甘肃 兰州

收稿日期: 2024年4月9日; 录用日期: 2024年6月21日; 发布日期: 2024年6月30日

## 摘要

苹果是一种备受欢迎且营养丰富的水果, 市场需求也在持续增长。然而, 苹果叶片病害严重威胁着苹果的产量和质量。准确、快速地识别病害对苹果产业的发展尤为重要, 因此, 本文利用YOLOv5s模型构建了一个改进的病害检测识别模型, 通过引入SE注意力机制、Res2Net和解耦检测头结构, 改进了模型的注意力、感受野和检测效果。探索了在苹果叶片病害检测中提高准确性和性能的方法。构建出YOLOv5s-SRD模型, 准确率、召回率、平均检测精度分别达到84.8%、83.3%、86.0%, 相比原模型分别提高了1.3%、1.5%、1.8%, 提升效果显著, 模型检测效果较好。有助于苹果叶片病害的检测识别, 从而提高苹果产量。

## 关键词

苹果叶片病害, 目标检测, YOLOv5, 模型改进

# Detection and Identification of Apple Leaf Disease Based on the Improved YOLOv5s

Long Xu\*, Lili Wu#

College of Information Science and Technology, Gansu Agricultural University, Lanzhou Gansu

Received: Apr. 9<sup>th</sup>, 2024; accepted: Jun. 21<sup>st</sup>, 2024; published: Jun. 30<sup>th</sup>, 2024

## Abstract

Apples are a popular and nutritious fruit, and the market demand continues to grow. However, apple leaf disease seriously threatens the yield and quality of apples. Accurate and rapid identification of diseases is particularly important for the development of apple industry. Therefore, this paper

\*第一作者。

#通讯作者。

uses YOLOv5s model to build an improved disease detection and identification model. By introducing SE attention mechanism, Res2Net and decoupling detection head structure, the attention, receptive field and detection effect of the model are improved. Methods to improve accuracy and performance in apple leaf disease detection were explored. The YOLOv5s-SRD model was constructed, and the accuracy, recall rate and average detection accuracy reached 84.8%, 83.3% and 86.0% respectively, which were 1.3%, 1.5% and 1.8% higher than the original model respectively, with the significant improvement effect and the model detection effect was good. It is helpful for the detection and identification of apple leaf disease, so as to increase the apple yield.

## Keywords

Apple Leaf Disease, Target Detection, YOLOv5, Model Improvement

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

苹果叶片病害的种类繁多[1], 不同病害的特征差异较小, 这给病害的准确识别带来了困难。其次, 苹果叶片病害的发生和发展受到环境、气候等多种因素的影响, 使得病害的表现形式和程度存在较大的差异。为了解决这些问题, 学者们不断探索新的技术和方法[2]。

李辉等[3]提出了一种基于改进 SSD 的苹果叶片病害检测系统, 通过引入注意力机制, 改进损失函数, 采用轻量级特征融合结构对 SSD 进行改进, 改进后的模型对病害识别的精度提高, mAP 达 81.95%。李书琴等[4]基于轻量级残差网络提出的植物叶片病害识别方法, 在自建苹果叶片病害数据集上获得 1.52% 的低识别错误率。Li 等[5]提出的 BTC-YOLOv5s 模型(模型大小为 15.8 M)能够有效检测自然场景中的 4 种苹果叶病, 平均精度(mAP)为 84.3%。曾晏林等[6]提出了一种融合自注意力机制和 Transformer 模块的目标检测算法——BCE-YOLOv5, 实现对自然环境下对苹果叶片病虫害的自动识别与检测。Mostafa 等[7]提出了 YOLOv4 模型, 融合了 FPN 和 PAN, 大大提高了多尺度特征提取能力, 添加了 CSP 模块, 减少了计算量的同时, 可以保证准确率, 取得了较好的成果。Al-Wesabi 等[8]提出了一种用于精准农业的新型人工智能苹果叶病分类(AIE-ALDC)技术。所提出的 AIE-ALDC 技术涉及基于方向的数据增强和基于高斯滤波的噪声去除过程。

虽然现有的深度学习算法在苹果叶片病害检测中取得了一定的成果, 但仍存在一定的局限性, 需要进行进一步的改进和完善。在众多深度学习技术中, YOLOv5s 以其高效性、轻量级、高准确性等优点, 在计算机视觉领域具有莫大的潜力[9]。因此, 本研究以 YOLOv5s 为基础模型进行改进, 提高其在苹果叶片病害检测识别中的准确性和效率, 为解决苹果叶片病害问题提供有效的解决方案。

## 2. 材料与方法

### 2.1. 苹果叶片病害图像数据

#### 2.1.1. 数据来源及数据集构成

本次实验研究数据集来源于飞桨 AI Studio 深度学习平台的公开数据集, 共有 8 种不同类别的苹果叶片图像, 其中包含健康(Health)的苹果叶片图像和患有斑点落叶病、褐斑病、蛙眼叶斑病、灰斑病(Grey spot)、

花叶病、锈病、黑星病的苹果叶片图像。苹果叶片病害样例如图 1 所示。



Figure 1. Images of apple leaf disease

图 1. 苹果叶片病害图像

由于原始数据集存在不同种类的病害图像分布不均、图像质量较低等情况，首先对图像进行了筛选，去除了目标特征模糊以及相似性高的图像，最后保留原数据集图像总计 10,780 张，其中健康叶片图像 1673 张，患有斑点落叶病图像 1428 张，患有褐斑病图像 2016 张，患有病青蛙眼病图像 847 张，患有灰斑病图像 1113 张，患有花叶病图像 1204 张，患有锈病图像 1561 张，患有黑星病的苹果叶片图像 938 张。

### 2.1.2. 数据预处理

在本次实验中，为了提高模型的泛化能力和准确性，丰富苹果叶片病害数据集的多样性和复杂性，对样本数量进行了有效的扩增。采用的数据增强操作主要有随机旋转、平移、倍增亮度、拉普拉斯噪声和图像翻转等。最终构建了本实验所用的苹果叶片病害数据集，总计 20,020 张图像。选用 LabelImg 对图像数据进行标签标注。通过使用 LabelImg 对处理好的数据集图像进行标注后得到了 8 种类别标签，分别是健康(Healthy)、链格孢属叶斑病/斑点落叶病(*Alternaria blotch*)、褐斑病(Brown spot)、蛙眼叶斑病(也叫青蛙眼病)(Frogeye leaf spot)、灰斑病(Grey spot)、花叶病(Mosaic)、锈病(Rust)、黑星病(Scab)。

## 2.2. 基于改进 YOLOv5 的方法

### 2.2.1. YOLOv5

YOLOv5 是目前应用较为广泛的一种目标检测的模型[10]，它将目标检测转变成了一个回归问题，把整张图作为网络的输入，仅仅经过一个神经网络，就可以得到 bounding box 的位置及其所属的类别。YOLOv5 的结构[11]主要分为输入端、主干网络(Backbone)、颈部网络(Neck)、Prediction 四个部分。

### 2.2.2. SE 注意力

SE 注意力机制[12] (Squeeze-and-Excitation Networks)是一种在通道维度上增加的注意力机制，它最主要的操作是压缩(squeeze)和激励(excitation)。具体计算步骤如下。

首先进行转型，如图 2 左侧白图，给定一个输入特征图，让其经过  $F_r(\cdot, \theta)$  操作得到一个新的特征图。

$F_{tr}(\cdot; \theta)$  可以看作一个标准的卷积算子。公式定义如(1)所示:

$$U_c = V_c * X = \sum_{s=1}^{c_1} V_c^s * X^s \tag{1}$$

其中,  $X \in R^{h \times w \times c_1}$  为输入特征图;  $U \in R^{h \times w \times c_2}$  为输出特征图;  $V$  表示学习到的一组滤波器核;  $V_c$  指的是第  $c$  个滤波器的参数;  $V_c^s$  表示一个 2D 的空间核;  $*$  表示卷积操作。

然后进行全局信息嵌入, 将特征图进行全局平均池化, 生成一个  $1 \times 1 \times c_2$  的向量, 这样每个通道让一个数值表示。公式如(2)所示:

$$z_c = F_{sq}(u_c) = \frac{1}{h \times w} \sum_{i=1}^h \sum_{j=1}^w u_c(i, j) \tag{2}$$

其中,  $z_c$  为  $z$  的第  $c$  个元素。

再然后做自适应重新校正, 通过两层全连接层完成, 通过权重  $W$  生成我们所需要的权重信息, 其中  $w$  是通过学习得到的, 用来显示建模需要的特征相关性, 公式如(3)所示:

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1, z)) \tag{3}$$

通过两个全连接层  $W_1$ 、 $W_2$  对上一步得到的向量  $z$  进行处理, 得到我们想要的通道权重值  $s$ , 经过两层全连接层后,  $s$  中不同的数值表示不同通道的权重信息, 赋予通道不同的权重。

最后进行重新加权, 将前面得到的注意力权重加权到每个通道的特征上。生成权重向量  $s$  对特征图  $U$  进行权重赋值, 得到我们想要的特征图  $\tilde{x}$  其尺寸大小与特征图  $U$  完全一样, SE 模块不改变特征图的大小, 公式如(4)所示:

$$\tilde{x}_c = F_{scale}(u_c, s_c) = s_c u_c \tag{4}$$

通过生成的特征向量  $s(1 \times 1 \times c_2)$  与特征图  $U(h \times w \times c_1)$ , 对应通道相乘, 即特征图  $U$  中每个通道的  $h \times w$  个数值都乘  $s$  中对应通道的权值。

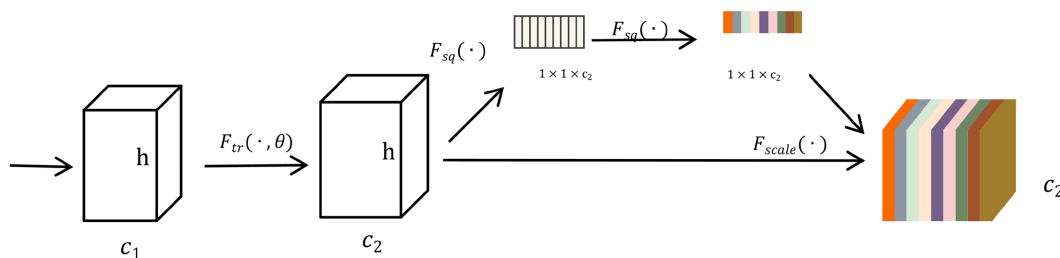


Figure 2. SE attention  
图 2. SE 注意力

### 2.2.3. Res2Net

Res2Net 是在 ResNet 模型基础上提出的一种新的模块, 它的核心思想是将 ResNet 中的卷积分解成多个子模块, 每个子模块可以利用更多的特征信息, 从而增强了网络的表达能力。与 ResNet 相比, Res2Net 的主要优势在于它能够更好地捕捉不同尺度的特征信息。如图 3 所示, 在 ResNet 中, 每个卷积层都只能捕捉一定范围内的特征, 而 Res2Net 通过将卷积分解成多个子模块, 并且将这些子模块连接起来, 可以有效地扩展感受野的范围, 从而捕捉到更丰富的特征信息。

如图 3 所示, Res2Net 将原来的 ResNet 中间的卷积换成了一种拆分混合连接结构, 如图右侧红色部分, 该部分最少是不经过任何操作, 直接连接。最多则会经过 3 个的卷积, 获得不同的感受野, 这就是

感受野比原结构多的原因。而里面的结构是经过卷积之后，又把特征图分成了 4 个部分。第一部分线路很简单，不做处理，直接传到；第二部分线路，经过卷积之后分为两条线路，一条继续向前传播给，另一条传到，这样第三条线路就获得了第二条线路的信息；第三条线路、第四条线路以此类推。每条线路的通道数为  $n/s$ 。

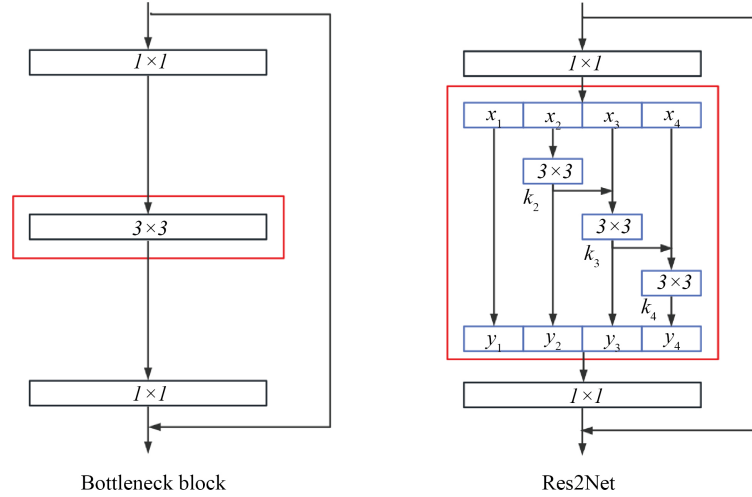


Figure 3. Bottleneck block and Res2Net  
图 3. Bottleneck block 与 Res2Net

### 2.2.4. 解耦头

在目标检测的定位和分类任务中，存在空间错位(spatial misalignment)的问题，即两个任务所聚焦和感兴趣的地方不同[13]，分类更加关注的是所提取的特征与已经存在的类别中的哪一类最为相近，也就是更关注目标的纹理内容，而定位更加关注 GT Box 的位置坐标，从而进行边界框参数修正，即更关注目标的边缘信息。因此，若是采取用同一个特征图进行分类和定位，效果会不好，而当采用解耦检测头(Decoupled Head)结构时，会同时综合考虑到分类和定位所关注的内容的不同，采用不同的分支来进行运算，有利于效果的提升。

如图 4 所示，对于 FPN 特征的每个级别，首先采用一个  $1 \times 1$  的 conv 层，将特征通道减少到 256，然后添加两个平行分支，每个分支有两个  $3 \times 3$  的 conv 层，分别用于分类和回归任务，并且在回归分支上添加了 IOU 分支。

### 2.3. 评价指标

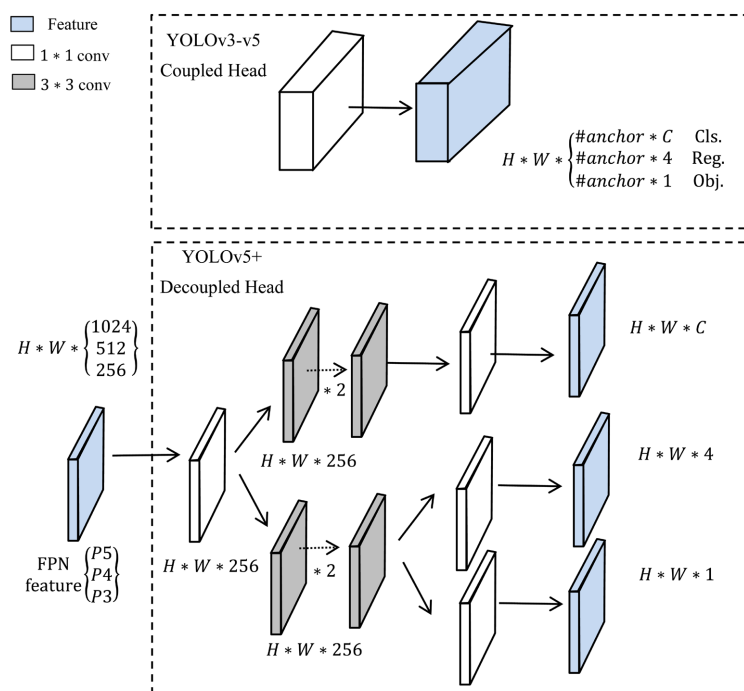
为了评估模型的性能，本次实验使用了精度(P)、召回率(R)和平均精度(mAP)作为判断模型性能的评价指标。以上指标的计算方法如下：

$$P = \frac{TP}{TP + FP} \tag{5}$$

$$R = \frac{TP}{TP + FN} \tag{6}$$

$$AP = \int_0^1 P(R) dR \tag{7}$$

$$mAP = \frac{\sum_{i=1}^n AP_i}{n} \tag{8}$$



**Figure 4.** Difference between Coupled Head and Decoupled Head of YOLOv3~v5  
**图 4.** YOLOv3~v5 Coupled Head 和 Decoupled Head 的区别

其中, TP 是预测为正类的阳性样本数; FP 是预测为正类的负样本数; FN 是预测为负类的正样本数; P 表示准确率(precision, P), 表示预测为正确的正样本数(TP)与模型预测的总样本数(TP + FP)之比; R 表示召回率(recall, R), 表示预测为正确的正样本数与总标签数(TP + FN)之比; AP 是 PR 曲线下方的区域。mAP 是每个类别的平均 AP 值, 即平均精度, mAP 值越大表示模型性能越好(本文中的 mAP 均指平均精度)。n 表示对象检测中的类别数。在本文中, n = 3。

## 2.4. 实验环境和训练参数

### 2.4.1. 实验环境

实验使用的计算机的 GPU 是 RTX4 090 (24 GB)\*1。CPU 是 12vCPU Intel(R) Xeon (R)Platinum8 352V CPU@2.10 GHz。PyTorch 版本为 1.11.0, Python 版本为 3.8, CUDA 版本为 11.3。

### 2.4.2. 训练参数

模型训练时选用官方提供的 YOLOv5s.pt 作为预训练权重进行训练, 数据集选用自己筛选、标注并进行增强操作后的葡萄叶片病害数据集, 标签格式为 TXT 格式, 输入图像大小为 640\*640, 批量大小为 16, 迭代次数为 600, 设置初始学习率为 0.01。

## 3. 实验结果与分析

### 3.1. 适用于改进 YOLOv5 的方法

#### 3.1.1. 引入注意力机制对模型的影响

SE 注意力通过对通道维度进行自适应的特征重标定, 可以使网络更加关注重要的特征, 从而提高模型的表达能力和性能。所以本文将其引入到 YOLOv5s 算法中。图 5 展示了 SE 注意力机制在 YOLOv5



模型的 Backbone 部分的网络结构设计。

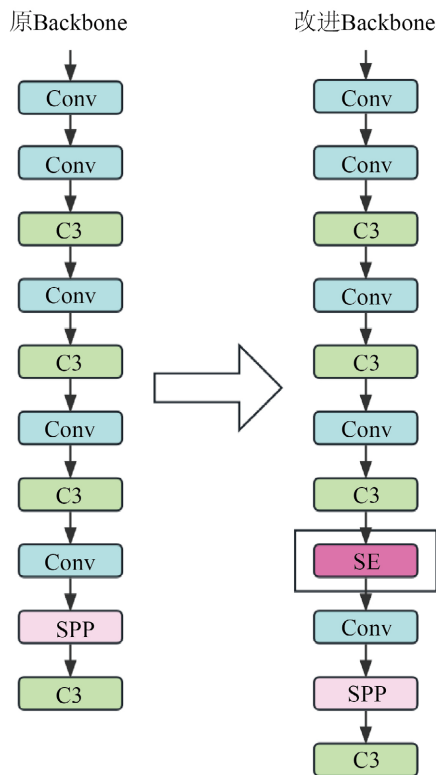


Figure 5. YOLOv5 improved SE structure diagram  
图 5. YOLOv5s 改进 SE 结构图

通过实验发现，改进后的模型效果优于 YOLOv5s 原始模型。实验结果如表 1 所示。

Table 1. Effect of SE attention on the model  
表 1. SE 注意力对模型的影响

模型	精确率(%)	召回率(%)	平均精度(%)
YOLOv5s	83.5	81.8	84.2
YOLOv5s-SE	88.3	79.2	85.7

从表 1 中可以看出，改进后的网络模型虽然在召回率上下降。但是在精确度、平均精度上有明显的提升，分别提升了 4.8%、1.5%。从检测结果来看，引入 SE 模块后的模型，其检测精度有所提升。

### 3.1.2. Res2Net 对模型的影响

Res2Net 能够更好地捕捉不同尺度的特征信息，因此本文将 Res2Net 加入 YOLOv5 的颈部网络中来提升模型检测效果。Neck 部分改进对比见图 6。

为了验证 Neck 部分改进的效果，使用 Res2Net 结构，改进后的算法称为 YOLOv5s-Res2Net。将改进后的算法和原来的 YOLOv5s 算法进行比较，实验结果如表 2 所示。





来提取特定的类别或属性信息, 将输出特征图分解为不同的通道。为 YOLOv5 模型添加解耦头后判断其对模型检测性能的影响, 结果如表 3 所示。

**Table 3.** Effect of the decoupling head on the model

**表 3.** 解耦头对模型的影响

模型	精确率(%)	召回率(%)	平均精度(%)
YOLOv5s	83.5	81.8	84.2
YOLOv5s-Decoupled head	87.7	80.5	85.5

实验结果表明, 对模型添加解耦头结构, 对 YOLOv5 模型的检测性能有所提升。与原模型相比, 精确度和平均精度分别提高了 4.2% 和 1.3%。虽然召回率有所下降, 但是检测精度有所提升, 证明对模型检测性能有所提升。

### 3.2. 消融实验

由前面的实验结果可知, YOLOv5s 添加 SE 注意力、引入 Res2Net、增添解耦头都使得模型的检测精度提升。因此, 进行消融实验, 以探索这些方法的融合是否能进一步提升 YOLOv5s 网络的性能。以原来的 YOLOv5s 为基准, 分别在网络中引入了 SE 注意力、Res2Net 和 Decoupled 并进行不同组合。实验结果如表 4 所示。

**Table 4.** Ablation experiments

**表 4.** 消融实验

SE 注意力	残差分辨率网络	解耦头	精确率(%)	召回率(%)	平均精度(%)	参数量(M)
×	×	×	83.5	81.8	84.2	7.03
√	√	×	87.1	80.7	85.0	7.05
×	×	√	87.7	80.5	85.5	14.3
√	√	×	86.9	78.5	83.7	6.8
√	×	√	84.6	79.9	84.8	14.3
×	√	√	85.8	81.1	85.4	14.1
√	√	√	84.8	83.3	86.0	14.1

从表结果可知, 除 SE 注意力和 Res2Net 两种改进方法相互融合时模型检测性能较原模型降低, 其他任意组合都对模型检测性能有明显提升。其中, 同时使用三种改进方法改进后的模型检测精度最高, 与原模型相比, 精确率、召回率和平均精度分别提高了 1.3%、1.5% 和 1.8%。由此证明本文对 YOLOv5s 网络模型的改进有效, 模型检测精度提升。

### 3.3. 模型结果对比实验

为验证改进后的 YOLOv5s 在苹果叶片病害数据集上的检测效果, 本文分别将 Faster R-CNN、YOLOv5s、YOLOv7, YOLOv7-Tine、YOLOv8 及改进后的 YOLOv5s 模型在苹果叶片病害数据集上进行训练, 实验结果如表 5 所示。通过表 5 对比结果可以看出, 在相同迭代次数内, 改进后 YOLOv5s 模型在苹果叶片病害数据集上的 mAP 值最高, 比 YOLOv5s、YOLOv7、YOLOv7-Tine 及 YOLOv8 四个不同的模型分别高 1.8%、0.8%、2.6% 和 0.7%。

**Table 5.** Detection results of different target detection models on apple leaf disease data set  
**表 5.** 不同目标检测模型在苹果叶片病害数据集上检测结果

模型	精确率(%)	召回率(%)	平均精度(%)	参数量(M)
YOLOv5s	83.5	81.8	84.2	7.03
YOLOv5s-SRD	84.8	83.3	86.0	14.1
YOLOv7	88.1	78.6	84.2	37.2
YOLOv7-Tine	82.2	79.4	83.5	6.03
YOLOv8	85.5	80.2	85.3	30.7

#### 4. 结束语

本文基于 YOLOv5s 模型构建了一个改进的病害检测识别模型,用于检测识别健康的苹果叶片和常见的 7 种苹果叶片病害,实现对病害类型的准确判断,从而提高苹果叶片病害的检测精度。本文所做的主要工作总结为: 1) 对苹果叶片病害图像进行刷选,并使用 LabelImg 标注工具对叶片上的病害进行标注; 2) 改进 YOLOv5s 模型,在 YOLOv5s 模型的 Backbone 部分引入 SE 注意力机制;在 Neck 部分引入 Res2Net;采用解耦检测头结构。构建出了 YOLOv5s-SRD 模型,该模型准确率、召回率、平均精度分别达到 84.8%、83.3%、86.0%,相比原模型分别提高了 1.3%、1.5%、1.8%,通过消融实验和对比实验证明改进模型的提升效果显著。下一步的工作是研究如何将改进后的模型方便、快捷地应用到实践中。

#### 基金项目

项目名称: 甘肃省科技计划项目(定西地区农村电子商务营销综合能力提升);项目编号: 20CX9NA095。

#### 参考文献

- [1] 刘莎. 基于轻量级网络的苹果叶片病害检测研究[D]: [硕士学位论文]. 咸阳: 西北农林科技大学, 2023.
- [2] 闫甄. 基于改进 YOLOv5s 的苹果叶片病害识别研究与应用[D]: [硕士学位论文]. 合肥: 安徽农业大学, 2023.
- [3] 李辉, 严康华, 景浩, 等. 基于改进 SSD 的苹果叶部病理检测识别[J]. 传感器与微系统, 2022, 41(10): 134-137.
- [4] 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.
- [5] Li, H., Shi, L., Fang, S. and Yin, F. (2023) Real-Time Detection of Apple Leaf Diseases in Natural Scenes Based on YOLOv5. *Agriculture*, **13**, Article 878. <https://doi.org/10.3390/agriculture13040878>
- [6] 曾晏林, 贺壹婷, 蔺瑶, 等. 基于 BCE-YOLOv5 的苹果叶部病害检测方法[J]. 江苏农业科学, 2023, 51(15): 155-163.
- [7] Mostafa, A.M., Kumar, S.A., Meraj, T., Rauf, H.T., Alnuaim, A.A. and Alkhayyal, M.A. (2021) Guava Disease Detection Using Deep Convolutional Neural Networks: A Case Study of Guava Plants. *Applied Sciences*, **12**, Article 239. <https://doi.org/10.3390/app12010239>
- [8] Al-Wesabi, F.N., Abdulrahman Albraikan, A., Mustafa Hilal, A., Eltahir, M.M., Ahmed Hamza, M. and Sarwar Zamani, A. (2022) Artificial Intelligence Enabled Apple Leaf Disease Classification for Precision Agriculture. *Computers, Materials & Continua*, **70**, 6223-6238. <https://doi.org/10.32604/cmc.2022.021299>
- [9] 赵子秋. 基于卷积神经网络和注意力模块的恶意软件检测研究[D]: [硕士学位论文]. 石家庄: 河北师范大学, 2022.
- [10] 韦强, 胡晓阳, 赵虹鑫. 改进 YOLOv5 的交通标志检测方法[J]. 计算机工程与应用, 2023, 59(13): 229-237.
- [11] 王梦妮, 顾寄南, 王化佳, 等. 基于改进 YOLOv5s 模型的茶叶嫩芽识别方法[J]. 农业工程学报, 2023, 39(12): 150-157.
- [12] 黄英来, 李宁, 刘镇波, 等. 改进轻量卷积网络在葡萄病害叶片的分类方法[J]. 哈尔滨理工大学学报, 2023, 28(3): 1-9.
- [13] 葛录录. 基于特征增强的改进型 YOLO 算法研究[D]: [硕士学位论文]. 唐山: 华北理工大学, 2022.