

# 改进 Resnet50 的葡萄叶片病害识别

兰博文

广西民族大学, 广西南宁, 530007;

**摘要:** 葡萄叶片病害在早期病变时具有病斑不明显, 病斑形状不规则等特点, 使得现有卷积网络识别困难。本文提出了一种改进的 ResNet50 模型用于葡萄叶片病害识别。该模型在 ResNet50 的基础上, 添加了 SimAM 注意力机制, 增加模型对病斑的敏感性, 引入多尺度卷积结构, 以捕捉病害斑点的不同尺度信息。在 PlantVillage 上进行训练, 结果表明, 改进的模型相比原始 ResNet50 取得了更高的识别准确率, 达到了 99.63%, 为建立自动化病害诊断系统提供重要应用价值。

**关键词:** 葡萄叶片病害; 残差网络; SimAM; 多尺度卷积

**DOI:**10.69979/3041-0673.25.03.044

随着全球变暖, 葡萄作物遭受各种病害的威胁日益严重, 给葡萄产业带来重大损失。手工病害识别低效无法满足大面积种植的需要<sup>[1]</sup>。开发自动化的葡萄病害图像识别技术, 对保障葡萄产量和质量意义重大。

基于深度学习的图像识别技术对农业生产的病害识别产生革命性影响。张彦通<sup>[2]</sup>等使用迁移学习方法对 MobilNetV2 模型训练来识别玉米病害, 精确率达到 96.83%; 谢圣桥等<sup>[3]</sup>改变了原始 ResNet50 中残差模块的卷积顺序, 改进模型识别准确率达 97.87%; 何前等<sup>[4]</sup>改进 AlexNet 模型, 在部分卷积层后加入池化层, 准确率达到 99.1%。Bahdanau<sup>[5]</sup>等人在机器翻译中首次提出了注意力机制。陈浪浪等<sup>[6]</sup>提出基于改进深度卷积神经网络的水稻病虫害识别模型, 将 DenseNet121 与坐标注意力机制结合, 识别准确率达 98.95%; 贾璐<sup>[7]</sup>基于 SE 和多尺度融合提出了葡萄病害识别模型 MANet, 识别准确率为 87.93%。研究表明, 注意力模块可以集成到卷积神经网络中, 帮助网络学习到更加聚焦的特征, 但加入注意力模块的同时也引入了额外的参数, 使模型复杂化, 同时葡萄叶片病斑尺度不一, 小尺度病斑识别困难, 识别准确率仍有待提升。

针对上述问题, 本文提出了改进的 ResNet50 网络结构。该模型引入了多尺度特征提取模块, 葡萄叶片病害症状出现在不同尺度, 采用多尺度卷积核可以充分表示病斑等症状。在卷积层中嵌入 SimAM 注意力机制。实验结果表明, 改进的模型取得了显著的性能提升。

## 1 数据

### 1.1 数据来源

PlantVillage 是公开的植物病害数据集, 包含多种植物的多种病害类别, 如葡萄、番茄等作物。本文数据

来自 PlantVillage 数据集中的 Grape, 包含四种分类, 黑腐病 1179 张、埃斯卡病 1383 张、褐斑病 1076 张和健康 423 张。

### 1.2 数据预处理与数据增强

根据模型结构, 将图片统一调整为  $3 \times 224 \times 224$ , 由于数据量较少且种类不均衡, 通过水平翻转、垂直翻转等方式扩充数据集, 提升模型的泛化能力, 最终数据集包含 6248 张图片。将数据集按 4:1 划分成训练集与测试集。

## 2 模型结构

### 2.1 残差网络结构

ResNet 是由 Kaiming He 等提出的架构, 可以训练非常深的网络而不会产生过拟合问题。该模型通过堆叠多个残差单元来构建网络, 残差网络单元假设输入特征图为  $x$ , 经过卷积计算得到  $F(x)$ , 然后与原输入  $x$  相加, 得到模块的最终输出:  $H(x) = F(x) + x$  这里  $F(x)$  表示连续卷积层的计算功能。网络层次深, 梯度在反向传播时可能会出现消失或爆炸。残差结构可以让梯度直接跳过数个层, 从而避免这个问题, 使得极深网络的训练成为可能。相比普通的卷积网络, ResNet 具有更好的性质, 可以训练超过 100 层以上的深度模型。整个 ResNet 由多个模块构成, 每个模块就是由图 1 所示的残差结构堆叠而成, 不同模块之间通过步长卷积进行下采样, 扩大感受野。

### 2.2 SimAM 注意力机制

本文在原模型中嵌入了 SimAM 注意力机制, 用来强化病斑区域的特征表达, 增强易被忽略的早期病斑特征。SimAM 是由 LingXiao Yang 等人提出的一种基于能量函

数的注意力机制，模型参考神经科学的理论，信息丰富的神经元通常表现出与周围神经元不同的放电模式。激活神经元通常会抑制周围神经元，该类神经元应当赋予更高的重要性。最简单的寻找重要神经元的方法：度量神经元之间的线性可分性。因此，定义了如下能量函数：

$$e_t(w_t, b_t, y, x_i) = (y_t - \hat{t})^2 + \frac{1}{M-1} \sum_{i=1}^{M-1} (y_0 - \hat{x}_i)^2 \#(1)$$

为简单起见，我们采用二值标签<sup>1</sup>，并添加正则项，

$$e_t(w_t, b_t, y, x_i) = \frac{1}{M-1} \sum_{i=1}^{M-1} (-1 - (w_t x_i + b_t))^2 + (1 - (w_t$$

每个通道有  $M = H * W$  个能量函数。上述公式具有如下解析解：

$$w_t = -\frac{2(t - u_t)}{(t - u_t)^2 + 2\sigma_t^2 + 2\lambda} b_t = -\frac{1}{2}(t + u_t)w_t \#(3)$$

$$其中, u_t = \frac{1}{M-1} \sum_{i=1}^{M-1} x_i, \sigma_t^2 = \frac{1}{M-1} \sum_{i=1}^{M-1} (x_i - u_t)^2。$$

因此，最小能量可以通过如下公式得到：

$$e_t^* = \frac{1}{(t - \hat{u})^2 + 2\sigma_t^2 + 2\lambda} \#(4)$$

上述公式意味着：能量越低，神经元  $t$  与周围神经元的区别越大，重要性越高。因此，神经元的重要性可以通过  $1/e_t^*$  得到。最后对特征进行增强处理：

$$\tilde{X} = \text{sigmoid}\left(\frac{1}{E}\right) * X$$

## 2.3 多尺度特征结构

多尺度卷积最早是在 GoogleNet 中被提出的。针对葡萄叶片病害存在病斑尺寸变化大，形状不规则的特点，多尺度卷积通过多个尺度卷积核并行运算，感受不同尺度的病斑。本文使用三个卷积核并行运算替代原始模型中的单个卷积核，减少参数同时充分利用不同尺度特征的互补信息。

## 2.4 模型整体结构

该模型以 resnet50 为主干网络，将 resnet50 的 7x7 卷积核替换成改进的多尺度特征提取结构；在原模型中加入 simAM 注意力机制，在每个 Bottleneck 中，输入经过 3x3 卷积核之后，进入 simAM 注意力层，以强化对病斑区域的特征表达。

## 3 实验结果与分析

### 3.1 实验环境及参数设置

实验在 linux 上进行，gpu 为两张 Tesla T4，内存共 32gb，cuda 驱动版本 11.2，语言为 python，深度学习框架 pytorch，版本 1.10。训练及测试时采用批处理的方式，设定 epoch 为 30，每批数据量为 32 张图片。

采用 SGD 优化器，参数设置学习率为 0.0001，损失函数采用交叉熵损失函数。

### 3.2 模型评价指标

采用了 Accuracy、Precision、Recall、F1 四个指标。Accuracy: 反映模型总体分类正确的样本占全部样本的比例，作为模型首要评价指标，Precision: 对于每个类别，表示模型预测为正样本中实际为正样本的比例，评估模型的分类质量。Recall: 对于每个类别，表示实际为正样本中被正确预测为正样本的比例，评估模型的捕捉能力。F1: 考虑了 Accuracy 和 Recall，是两者的调和平均数，代表模型的整体分类性能。

### 3.3 消融实验

本实验中，为验证改进结构的有效性，分别与 resnet50、resnet50\_Multi、resnet50\_simAM 和 resnet50\_Multi\_simAM 进行比较。

#### 3.3.1 评价指标对比

将 Multi\_Scale 结构和 simAM 结构作为组件，分别测试结构加入模型之后的表现。结果如表所示。相比于原始模型，加入 Mutil\_Scale 和 SimAM 结构分别将准确率提升了 2.27% 和 2.47%，而融合两个组件的 SimAM\_MS C\_ResNet50 取得了最好的表现，相比于原始模型提升了 3.09% 的准确率。

Mod el	Resn et50	Multi _Scal e	simA M	Accu racy/ %	Preci sion/ %	Recal l/ %	F1/ %
1	✓			96.5	96.6	95.9	96.2
				4	7	1	6
2	✓	✓		98.8	99.0	98.5	98.7
				1	7	1	8
3	✓		✓	99.0	99.1	99.2	99.2
				1	9	1	0
4	✓	✓	✓	99.6	99.6	99.7	99.7
				3	7	0	0

#### 3.3.2 热力图可视化分析

为更直观的对比模型对于病害特征的捕捉，生成模型的热力图。红色和橙色是对类别预测最重要的区域。绿色和黄色属于中间水平的重要性。蓝色和黑色对图像预测是无关的。结果如图所示：在黑腐病中和埃斯卡病两种面积较小的病斑中，a、b 和 c 都捕捉了与分类无关的信息，而 d 能够更好的捕捉与分类相关的病斑特征。健康叶片四种模型都能较好的覆盖；黄腐病表现在整个叶片范围内变得枯黄，可以看到 d 能够更大范围的捕捉

特征。根据热力图可以看出，无论是对病斑较小的黑腐病和埃斯卡病，还是病斑范围较大的黄腐病，SimAM\_MS

C\_ResNet50 都能更精确地捕捉特征。

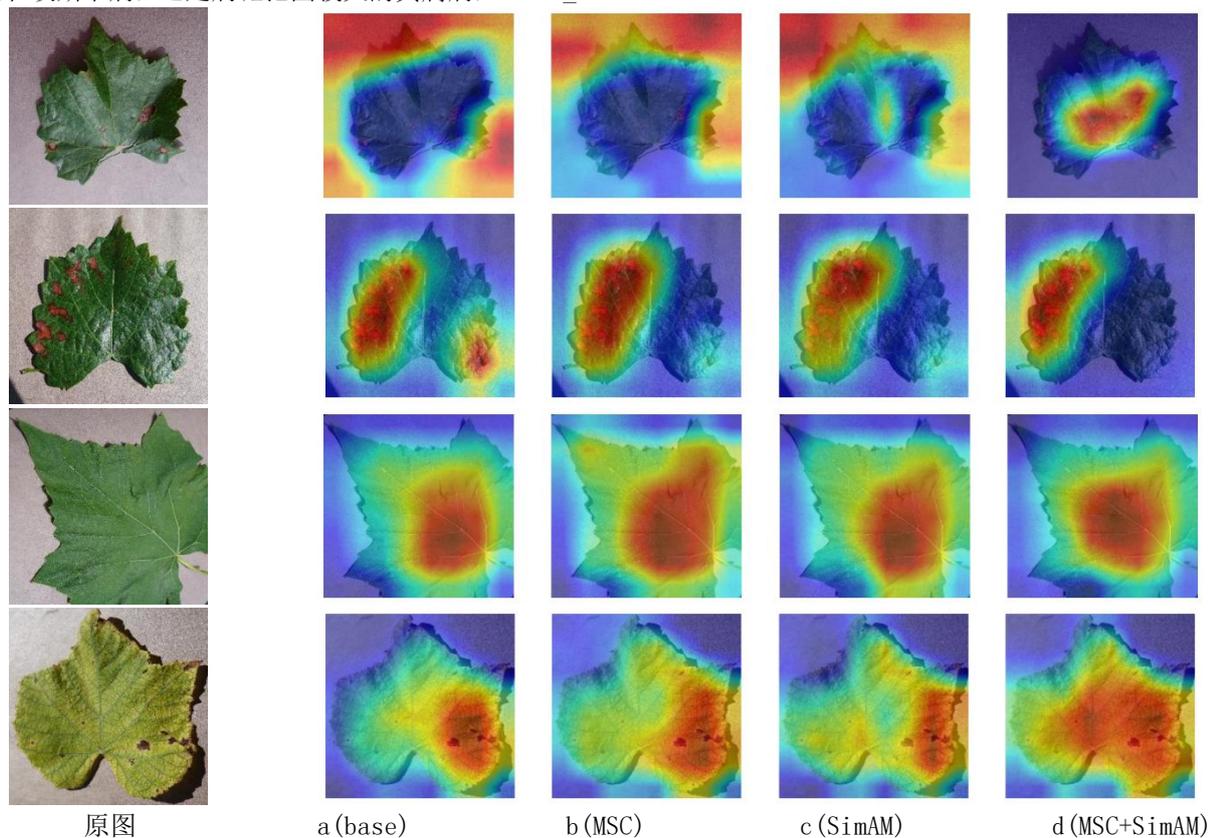


图 5 不同模型识别效果的可视化

## 4 结论

针对葡萄叶片病害中病斑不规则、不明显导致识别难度大的问题，提出一种改进的残差网络模型。以葡萄叶片中常见的病害为研究对象，以 ResNet50 为基础，加入多尺度卷积结构，以更好的捕捉不同尺度病斑；在每个 Bottleneck 的 3x3 卷积层后面添加 SimAM 注意力机制，提升对病害的敏感性。最终在 PlantVillage 上取得了 99.63% 的准确率。下一步将搜集更多类别的葡萄叶片病害数据，改进模型以适应真实场景下能准确识别葡萄叶片的各类病害，为构建自动化病害识别系统提供技术支持。

### 参考文献

[1] 潘鸿, 韩薇, 王曰军. 葡萄树常见病虫害的防治技术[J]. 农业灾害研究, 2023, 13(01): 10-12.  
[2] 张彦通, 苏前敏. 基于迁移学习的玉米病害图像识别[J]. 中国农业科技导报, 2023, 25(10): 119-125. DOI: 10.13304/j.ny.kjdb.2022.0218.

[3] 谢圣桥, 宋健, 汤修映等. 基于迁移学习和残差网络的葡萄叶部病害识别[J]. 农机化究, 2023, 45(08): 18-23+28. DOI: 10.13427/j.cnki.njyi.2023.08.048.  
[4] 何前, 郭峰林, 王哲豪等. 基于改进 AlexNet 的葡萄叶部病害分类算法[J]. 扬州大学学报(自然科学版), 2023, 26(02): 52-58. DOI: 10.19411/j.1007-824x.2023.02.008.  
[5] Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate[J]. arXiv preprint arXiv:1409.0473, 2014.  
[6] 陈浪浪, 张艳. 基于改进深度卷积神经网络的水稻病虫害识别[J]. 山农农业科学, 2023, 55(05): 164-172. DOI: 10.14083/j.issn.1001-4942.2023.05.024.  
[7] 贾璐, 叶中华. 基于注意力机制和特征融合的葡萄病害识别模型[J]. 农业机械学报, 2023, 54(07): 223-233.

作者简介: 兰博文(1999-)男, 汉族, 福建省龙岩市, 广西民族大学, 硕士, 研究方向: 计算机视觉