

基于随机增强 Swin-Tiny Transformer 的 玉米病害识别及应用

吴叶辉, 李汝嘉, 季荣彪, 李亚东, 孙晓海, 陈娇娇, 杨建平

(云南农业大学 大数据学院, 昆明 650201)

摘要: 针对图像识别中获取全局特征的局限性及难以提升识别准确性的问题, 提出一种基于随机增强 Swin-Tiny Transformer 轻量级模型的图像识别方法. 该方法在预处理阶段结合基于随机数据增强(random data augmentation based enhancement, RDABE)算法对图像特征进行增强, 并采用 Transformer 的自注意力机制, 以获得更全面的高层视觉语义信息. 通过在玉米病害数据集上优化 Swin-Tiny Transformer 模型并进行参数微调, 在农业领域的玉米病害上验证了该算法的适用性, 实现了更精确的病害检测. 实验结果表明, 基于随机增强的轻量级 Swin-Tiny+RDABE 模型对玉米病害图像识别准确率达 93.586 7%. 在参数权重一致, 与性能优秀的轻量级 Transformer、卷积神经网络(CNN)系列模型对比的实验结果表明, 改进的模型准确率比 Swin-Tiny Transformer, Deit3_Small, Vit_Small, Mobilenet_V3_Small, ShufflenetV2 和 Efficientnet_B1_Pruned 模型提高了 1.187 7%~4.988 1%, 且能迅速收敛.

关键词: Swin-Tiny Transformer 模型; 数据增强; 迁移学习; 玉米病害识别; 图像分类

中图分类号: TP391 **文献标志码:** A **文章编号:** 1671-5489(2024)02-0381-10

Maize Disease Recognition and Application Based on Random Augmentation Swin-Tiny Transformer

WU Yehui, LI Rujia, JI Rongbiao, LI Yadong, SUN Xiaohai, CHEN Jiaojiao, YANG Jianping

(College of Big Data, Yunnan Agricultural University, Kunming 650201, China)

Abstract: Aiming at the problems of the limitation of obtaining global features in image recognition and the difficulty in improving recognition accuracy, we proposed an image recognition method based on the lightweight model of random augmentation Swin-Tiny Transformer. The method combined the random data augmentation based enhancement (RDABE) algorithm to enhance image features in the preprocessing stage, and adopted the Transformer's self-attention mechanism to obtain more comprehensive high-level visual semantic information. By optimizing the Swin-Tiny Transformer model and fine-tuning the parameters on a maize disease dataset, the applicability of the algorithm was verified on maize diseases in the agricultural field, and more accurate disease detection was achieved. The experimental results show that the lightweight Swin-Tiny+RDABE model based on stochastic enhancement has an accuracy of 93.586 7% for maize disease image recognition. The experimental

收稿日期: 2023-04-24.

第一作者简介: 吴叶辉(1999—), 男, 汉族, 硕士研究生, 从事计算机视觉和自然语言处理的研究, E-mail: 1617899733@qq.com.

通信作者简介: 杨建平(1977—), 男, 汉族, 博士, 教授, 从事信息技术的研究, E-mail: yangjypyn@163.com.

基金项目: 云南省重大科技专项基金(批准号: 202302AE09002003).

results compared with the excellent performance lightweight Transformer and convolutional neural network (CNN) series models with consistent parameter weights show that the accuracy of the improved model is higher than that of the Swin-Tiny Transformer, Deit3_Small, Vit Small, Mobilenet_V3_Small, ShufflenetV2 and Efficientnet_B1_Pruned models by 1.187 7% to 4.988 1%, and can converge rapidly.

Keywords: Swin-Tiny Transformer model; data augmentation; transfer learning; maize disease recognition; image classification

图像分类是计算机视觉领域中的一项重要任务. 近年来, 随着卷积神经网络(CNN)^[1]在图像分类问题中的广泛应用和不断发展, 涌现出了许多性能优异的卷积神经网络模型. 例如, Krizhevsky 等^[2]在数据集 ImageNet 上提出的 AlexNet 模型, 以其 15.3% 的 top5 错误率远超其他传统特征提取方法, 引起了深度学习领域的广泛关注; 陈智超等^[3]改进了 MobileNetv2 算法并应用于垃圾图像分类, 通过多种优化策略在自建数据集上实现了 94.6% 的准确率; 邵伟平等^[4]提出了一种轻量化卷积神经网络 YOLO-Slim, 该模型将模型大小减小了 90%, 在测试数据集 VOC2007 上获得了 76.42% 的平均精度均值(mAP), 识别速度为 16 ms, 适用于小型图像处理平台. 卷积神经网络擅长解构图像信息和提取底层特征, 但其模型存在缺陷, 如不擅长捕捉图像全局信息、卷积操作感受野的局限、梯度消失问题和信噪比下识别效果易受影响等^[5].

Transformer 的出现为解决 CNN 的这些缺点提供了一种有效方法. Vaswani 等^[6]引入了 Transformer 架构作为一种新的序列建模方法, 对输入序列中的全局依赖关系进行建模, 取得了在翻译任务中优异的性能, 并在计算机视觉领域引起了广泛关注. 而 Swin Transformer 技术作为其改进模型, 不仅具有全局建模能力, 还通过层级化网络结构和移动窗口信息交互方式扩大感受范围、降低计算量, 适用于不同尺度目标的特征提取^[7].

随着计算机技术的深入应用和快速发展, 机器学习已成为农作物病害识别领域的重要研究方向, 同时, 图像识别技术的不断成熟也促进了机器学习技术与图像识别技术的结合越来越普及^[8]. 玉米作为我国种植面积和总产量最大的粮食作物之一, 对其进行病害检测以保障农作物产量具有重要意义^[9-10].

本文采用基于随机数据增强的增强算法对数据集进行预处理, 然后在增强的玉米叶片病害数据集上采用迁移学习进行训练. 在训练过程中, 使用轻量级的 Swin-Tiny Transformer 模型框架进行优化, 以分析玉米病害叶片图像的特征, 并进行鉴定.

1 随机增强的 Swin-Tiny Transformer

本文提出一种基于 Swin Transformer 模型中的轻量版本 Swin-Tiny Transformer 为骨干网络的图像识别模型, 该模型的结构如图 1 所示.

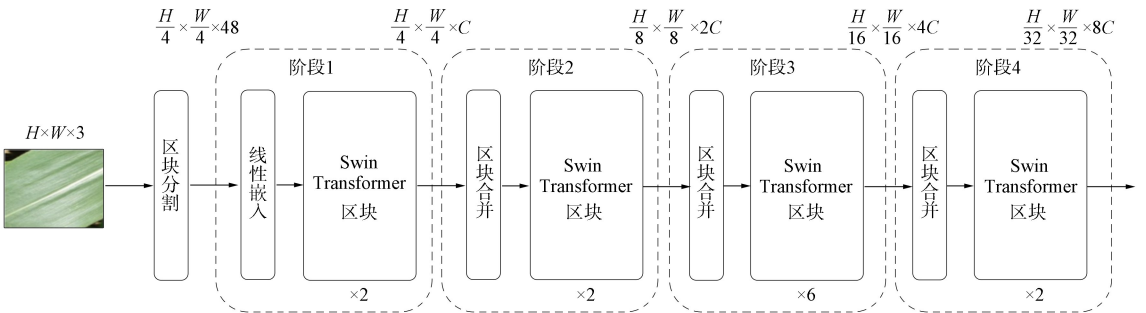


图 1 Swin-Tiny Transformer 网络架构

Fig. 1 Architecture of Swin-Tiny Transformer network

首先, 利用基于随机数据增强算法(random data augmentation based enhancement, RDABE)对已

有数据集进行数据增强, 然后在数据集 ImageNet 上预先训练的 Swin Transformer 图像分类模型的 tiny 版本, 在经过增强后的图像数据集上进行调整, 通过玉米病害图像实现对其疾病分类。

该模型的总体框架包括数据预处理、分类核心网络和图像预测. 与传统卷积神经网络模型相比, 基于随机增强 Swin-Tiny Transformer 的图像识别模型主要进行了如下优化:

1) 对玉米叶片病害图片数据集中的训练集在传入后进行数据预处理和数据增强. 数据处理时, 用 RDABE 算法进行数据增强, 与现有方法相反, RDABE 算法是在定义的范围内随机生成增强超参数范围, 使得其很健壮, 不但提高了模型的泛化能力, 而且防止过度拟合到特定数据集。

2) 为得到玉米病害图像的整体信息及上层视觉语义信息, 本文用 Swin-Tiny Transformer 作为分类核心网络. 该主干网络的核心采用类似于 CNN 的分层建构方法, 使用不同的下采样比例对图像进行下采样. 在特征提取的整个过程中, 选择 Swin Transformer 的轻量版本 Tiny, 它具有较小的计算量, 但可以加强各分割模块之间的信息联系, 从而更好地提取玉米病害图像的综合病害特征, 进而能更准确地识别和划分玉米病害的种类。

3) 通过迁移学习技术, 将预训练的 Swin-Tiny Transformer 模型的权重信息迁移到改进玉米病害检测模型中, 从而降低了对数据集大小的要求. 该技术的最大优点是在轻量级模型已经拥有较小计算量的基础上, 进一步缩短了训练时间并增加了玉米病害识别的准确率。

改进的 Swin-Tiny Transformer 玉米病害识别模型主要通过数据预处理阶段进行数据提升, 然后进行玉米病害分类识别, 并对模型权重进行调整, 从而构建了基于 Swin-Tiny Transformer 的玉米病害分类模型。

1.1 基于随机数据增强的增强算法

在复杂的深度学习(DL)架构和使用依赖数据集的静态超参数提高数据质量方面, 由于数据质量和超参数对特定数据集的过度拟合, 导致性能依旧受限^[11], 如图 2 所示. 而 RDABE 算法是一种生成化、数据独立和计算高效的增强方法, 以此提高 DL 的植物病害数据质量. 通过提高图像的亮度和对比度提高质量, 与现有方法相反, RDABE 算法在定义的范围内随机生成增强超参数范围, 使得其很健壮, 并防止过度拟合到特定数据集, 如图 3 所示. 传统 RDABE 算法尝试使用直方图均衡化方法对颜色对比度进行数据增强, 以提高图像质量. 但实验发现, 这种方法与颜色对比度之间的相关性较小, 且增强后的数据无法很好地突出图像特征. 数据增强技术是数据预处理的一种重要方法, 通过使用已有的玉米病害图像训练样本, 创造出更多不同风格的训练数据, 为使经过增强后的图像样本更接近真实情况, 从而提升检测的准确度并增强模型的泛化能力, RDABE 算法以数据集特征为依据, 限制增强参数生成范围, 并综合考虑数据集的统计特征和数据增强策略, 以确保图像在增强过程中保持语义一致性. 从而可确保数据增强过程不会对图像的语义内容造成显著改变, 同时提高了算法在实际应用中的可靠性和稳定性。

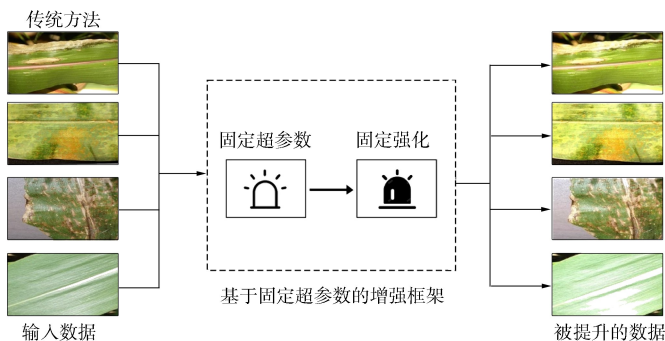


图 2 传统固定超参数提高数据质量框架

Fig. 2 Traditional fixed hyperparameters improve data quality framework

使用数据增强技术可增大深度学习网络的训练数据集, 以此处理网络的过拟合问题和数据样本分布不均衡的困难, 因为其直接影响模型的泛化能力和鲁棒性. 单样本数据增强和多样本数据增强是两

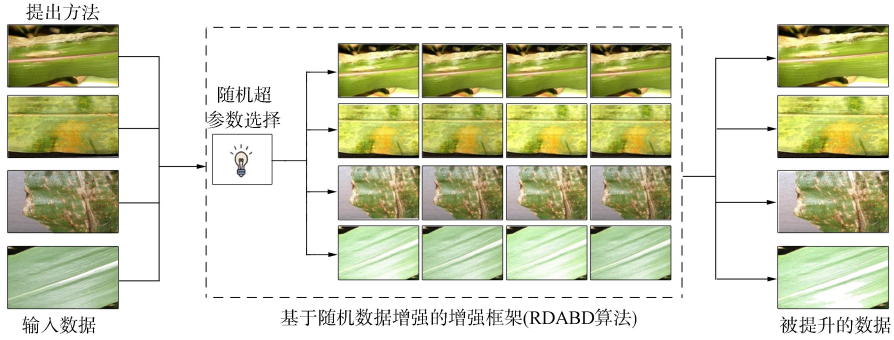


图 3 RDABE 算法提高数据质量框架

Fig. 3 RDABE algorithm improves data quality framework

种最常用的数据增强方法. 其中, 单样本数据增强方法可基于颜色空间变换或基于几何变换, 本文使用的 RDABE 算法即为前者. 相反, 多样本数据增强方法则不是对单个样本进行变换, 而是使用多个图像样本生成新的样本. 该方法通过将离散的样本点连续化拟合真实样本分布.

本文引入 RDABE 数据增强算法对玉米病害数据进行处理, 以提高玉米病害识别模型的泛化能力. RDABE 是一种计算高效的基于随机数据增强的增强方法, 除数据的选择是随机的, 增强超参数同样不是根据数据手动选择的, 因此增强方法不会过度拟合特定数据集. 通过 RDABE 算法可以改进网络架构的泛化能力, 并提高网络模型对图像样本识别的鲁棒性, 与其他先进的算法相比, RDABE 算法在准确性和执行时间方面都性能优异. RDABE 算法的主要原理如下:

1) 设 $f(i)$ 是输入图像, $g(i)$ 是增强后的图像, 则

$$g(i) = \alpha f(i) + \beta, \tag{1}$$

其中 α 和 β 为增益和偏置, 用这些参数调节对比度(α)和亮度(β).

2) 与其他使用固定超参数进行增强并容易在特定数据集上过度拟合的增强方法不同, RDABE 算法从定义的集合中随机选择 α 和 β 的值. 为防止过度拟合, 使用的集合由 α 和 β 的正负值组成, 这些值从该集合中随机选取, 然后用于增强数据.

1.2 基于迁移学习的 Swin-Tiny Transformer 模型

迁移学习能使模型在训练初期就达到较好的性能, 从而缩短训练时间并降低计算资源的使用量. 此外, 迁移学习还有助于提高训练后模型的鲁棒性和泛化能力, 使其性能更好^[12].

Transformer 模型^[6] 主要用于自然语言处理. 由于该模型的优异性能, 研究人员开始尝试将 Transformer 应用于计算机视觉领域^[13], 并取得了很好的效果. Vision Transformer 在计算机视觉领域性能优异, 但其参数量较大. 文献^[14]提出的 Swin Transformer 展现了一种滑动窗口自注意力机制, 能在局部窗口上执行自注意力机制 (self-attention), 从而有效降低了参数量, 并且结合了卷积神经网络的思想, 采用层次化构建方式堆叠 Transformer 模型, 可同时提取全局特征并高效并行计算, 而 CNN 只能提取局部特征. Swin Transformer 分层特征图如图 4 所示.

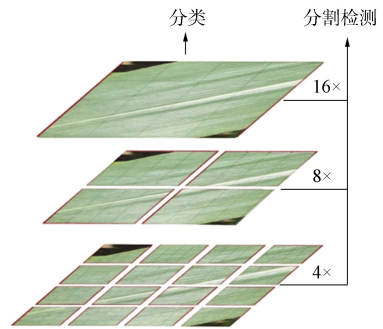


图 4 Swin Transformer 的分层特征图

Fig. 4 Hierarchical feature map of Swin Transformer

Swin Transformer 算法通过采用滑动窗口操作计算自注意力和层次化构建方式建立模型, 允许跨窗口连接, 以降低模型的复杂度, 提高模型的运行效率, 从而极大减少了模型的参数量, 更好地实现轻量化目标. Swin Transformer 模型组成: 窗口多头自注意力层 (W-MSA)、滑动窗口多头自注意力层 (SW-MSA)、标准化层 (LN) 和多层感知机 (MLP).

在对玉米病害图片进行训练时, Swin-Tiny Transformer 模型采用以下步骤进行图像处理: 首先, 使用 Patch Partition 模块将图片进行分块处理, 规定每个 Patch 为 4×4 相邻的像素, 并在 channel 方向进行展平; 其次, 图像通过 4 个 Stage 构建特征图, 在每个 Stage 中, 图像先经过 Linear Embedding 层, 然后剩下的 3 个 Stage 都要经过 Patch Merging 层^[15]; 最后, 在图像通过第 4 个 Stage 后, 会经过 Layer Norm 层、全局池化层以及全连接层的处理, 最终得到分类后的图像^[16]. Swin-Tiny Transformer 网络架构的结构如图 1 所示.

全局自注意力的计算复杂度以平方倍提升, 特别是在进行 CV 领域的下游任务中处理非常大尺寸的图片时, 其计算复杂度非常高. 针对该问题, Swin Transformer 采用了窗口计算自注意力的方法. 在每个窗口内进行自注意力计算, 但窗口之间无法进行信息传递, 因此提出了平移窗口多头自注意力机制(SW-MSA)模块, 使窗口之间能进行信息交流^[7], 如图 5 所示, 从而有效解决了不同窗口之间无法进行信息交流的问题.

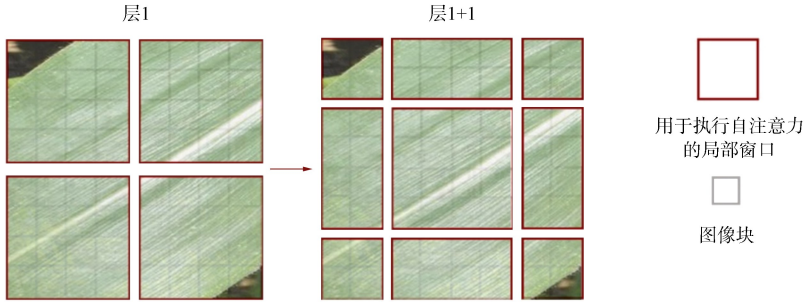


图 5 Swin Transformer 架构中的 SW-MSA

Fig. 5 SW-MSA in Swin Transformer architecture

1.3 模型评估

为科学衡量本文模型在玉米植株病害分类任务中的性能, 采用常用的经典性能衡量标准, 如准确率、精确度、召回率、 F_1 分数和混淆矩阵进行性能分析, 计算公式分别为

$$\text{准确率} = \frac{TP + TN}{TN + TP + FN + FP}, \tag{2}$$

$$\text{精确度} = \frac{TP}{TP + FP}, \tag{3}$$

$$\text{召回率} = \frac{TP}{TP + FN}, \tag{4}$$

$$F_1 \text{ 分数} = \frac{2 \times TP}{2 \times TP + FP + FN}, \tag{5}$$

其中: TP 表示准确预测为玉米病害叶片样本的数量, 也称为真阳性样本数量; TN 表示被正确预测为健康玉米叶片样本的数量, 也称为真阴性样本数量; FP 表示被错误预测为玉米病害叶片样本的数量, 也称为假阳性样本数量; FN 表示被错误预测为健康玉米叶片样本的数量, 也称为假阴性样本数量.

2 实验

2.1 图像数据集

本文使用源于 Kaggle 的公开可用的玉米叶片病害数据集^[17], 该图像数据集是用流行数据集 PlantVillage 和 PlantDoc 制作的^[18], 在数据集的形成过程中, 删除了一些无用的图像. 数据集共包含 4 个类别图像集合, 共 4 185 张玉米叶片图像, 其中健康玉米叶片图像 1 162 张, 感染枯萎病的玉米叶片病害图像 1 146 张、感染锈病的玉米叶片病害图像 1 306 张、感染灰斑病的玉米叶片病害图像 574 张. 玉米叶片病害图像数据样例如图 6 所示. 其中: 第一列为两张感染枯萎病的叶子, 叶端出现萎蔫、卷曲, 逐渐向下蔓延, 整个叶片逐渐干枯, 并出现褐色斑点和缺失区域; 第二列为两张感染锈病的叶子, 出现黄色或橙色的斑点或斑块, 这些斑点或斑块周围有明显的红色或棕色边缘; 第三列为两张

感染灰斑病的叶子, 出现小而圆的灰色或灰白色斑点, 斑点周围边缘会出现黄色或褐色的区域; 第四列为两张健康的玉米叶片, 呈现深绿色, 叶片质地柔软而富有光泽, 整体看起来较饱满. 玉米数据集的图像类别以及图像数量列于表 1.

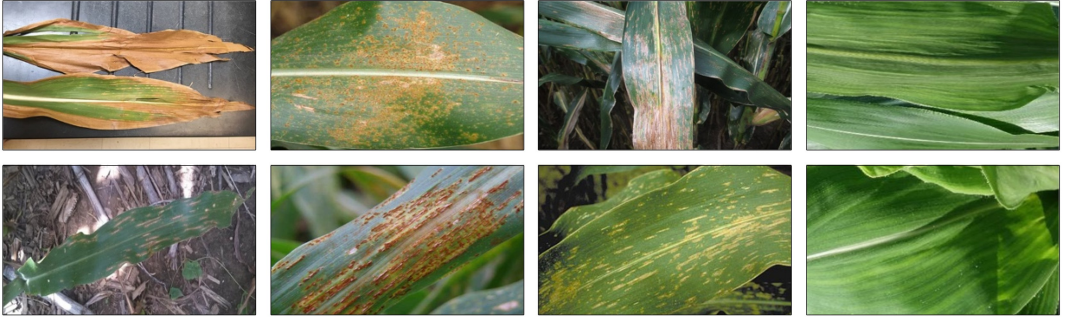


图 6 玉米叶片病害数据样例

Fig. 6 Data samples of maize leaf diseases

表 1 玉米数据集图像类别及数量

Table 1 Image categories and quantities of maize dataset

序号	病害类别名称	数量/张	序号	病害类别名称	数量/张
1	健康	1 162	3	感染锈病	1 306
2	感染枯萎病	1 146	4	感染灰斑病	574

数据集处理采用 Python 将数据集中的每个子目录作为一个类别, 使用随机采样的方式将每个类别的图像划分为训练集、验证集和测试集, 并将它们的路径和标签存储在相应的列表中, 生成类别名称和对应的数字索引存储于 JSON 文件中.

2.2 实验设备

轻量级模型需要更少的内存和计算资源, 更低的成本, 更容易部署在各种设备上, 为使实验研究更具对比性, 所有对比模型都选择轻量级版本. 其中 3 个轻量级卷积神经网络模型分别为 Mobilenet_V3_Small, ShufflenetV2, Efficientnet_B1_Pruned; 3 个轻量级 Transformer 系列模型为 Deit3_Small, Vit_Small, Swin-Tiny.

实验使用的硬件设备是一块 RTX 3090 显卡, 软件使用 jupyter, Pytorch 深度学习框架, torch-1.9.1+cu111, torchvision-0.10.1+cu111 构建所需的训练环境.

实验过程中, 将玉米病害数据集中的全部图像进行调整, 以符合所选模型框架对图像尺寸的要求, 将其像素大小转换为 224×224 , 以便在输入模型进行计算时使用. 按图像类别的 70%, 20%, 10%, 将其划分为训练集、验证集和测试集, 并完成了基于迁移学习的后续模型训练对比实验. 在进行模型训练时, 训练集和验证集被用来赋予模型的参数值, 而测试集则用于测试经训练后的模型并对其性能进行评估. 先加入 RDABE 算法作为数据增强技术, 进行与玉米病害相关的数据增强, 然后将其输入模型进行训练, 在进行训练数据前完成该过程. 采用迁移学习技术对模型进行训练, 利用预先在数据集 ImageNet 上训练的权重值, 从而显著提升了模型速度和性能.

为确保对比公平, 本文对每个模型的训练数据都单独分析并多次实验调整, 以确保每个模型的权重参数都被优化到其性能的最佳状态. 在使用这些参数计算测试模型时, 对准确率、精确度、召回率和 F_1 分数进行考察, 并且保存测试数据的模型预测结果及相应的标签, 然后用其绘制混淆矩阵. 其 Swin-Tiny Transformer 模型训练超参数设置如下: 优化器采用 SGD, 损失函数为 Cross-Entropy Loss, epochs 为 100, 批量大小为 64, 学习率设为 0.001, 权重衰减为 0.000 1.

2.3 实验结果分析

为评估深度学习训练模型在玉米病害分类识别任务中的性能, 本文选择表现优秀的 7 个模型进行基于迁移学习的实验训练和测试. 其中包括 3 个基于 CNN 架构的轻量级版本, 分别为 Mobilenet_V3_Small, ShufflenetV2 和 Efficientnet_B1_Pruned, 以及 3 个基于 Transformer 架构的轻量级版本, 分别

为 Deit3_Small, Vit_Small 和 Swin-Tiny, 此外还有经过提升的 Swin-Tiny+RDABE 模型. 使用 Pytorch 框架对这 7 个模型进行构建, 并进行基于迁移学习实验训练和测试, 从而全面评估这些模型在玉米病害图像分类识别任务中的分类性能.

图 7 和图 8 为本文各模型在玉米病害数据集训练过程中的结果. 这些曲线记录了每个训练周期的测试集准确率和训练集损失, 从而有助于评估每个模型的训练情况. 通过这些记录, 可确保所有模型都能在收敛状态下完成训练. 由图 7 和图 8 可见, 所有模型都是基于迁移学习技术, 并且都采用了轻量级版本的前提下, Swin-Tiny Transformer 模型在准确率和损失值上都取得了比其他模型更优的结果.

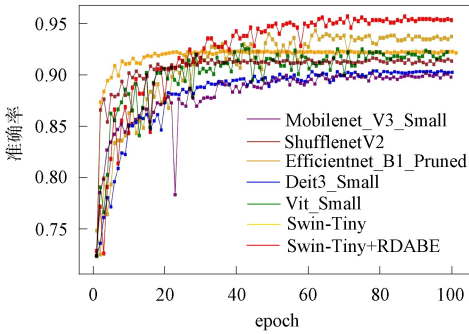


图 7 各模型准确率变化曲线

Fig. 7 Accuracy change curves of each model

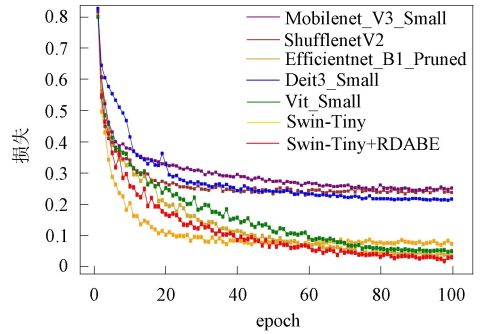


图 8 各模型训练损失曲线

Fig. 8 Training loss curves of each model

本文改进的 Swin-Tiny+RDABE 模型相比原 Swin-Tiny 模型, 无论是在测试集准确率还是在训练集损失中, 都获得了更优的训练结果. Swin-Tiny Transformer 训练对比曲线如图 9 和图 10 所示.

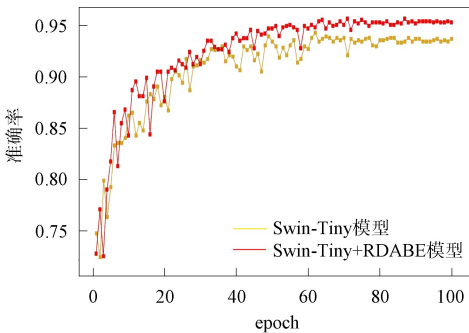


图 9 Swin-Tiny Transformer 模型准确率变化曲线

Fig. 9 Accuracy change curves of Swin-Tiny Transformer model

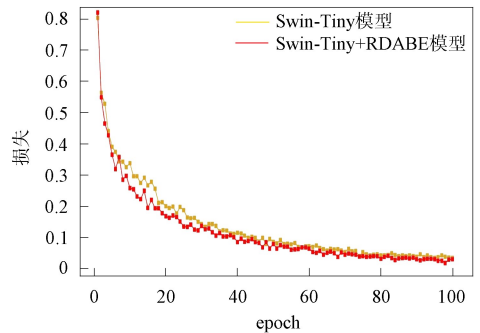


图 10 Swin-Tiny Transformer 模型训练损失曲线

Fig. 10 Training loss curves of Swin-Tiny Transformer model

在轻量级模型中无论是 CNN 系列还是 Transformer 系列都有较好性能, 其中性能最佳的是 Swin-Tiny Transformer 模型, 准确率高达 92.399%, F_1 分数也达到 89.9157%, 相比 MobileNet_V3_Small, ShuffleNetV2, Efficientnet_B1_Pruned, Deit3_Small, Vit_Small 模型具有显著优势. 不同模型实验结果参数指标列于表 2.

表 2 不同模型的准确率、精确度、召回率和 F_1 分数实验结果

Table 2 Experimental results of accuracy, precision, recall and F_1 scores of different models

模型	准确率	精确度	召回率	F_1 分数
MobileNet_V3_Small	0.885 986	0.855 740	0.835 289	0.840 672
ShuffleNetV2	0.912 110	0.883 080	0.881 906	0.882 476
Efficientnet_B1_Pruned	0.921 615	0.904 001	0.891 133	0.896 230
Deit3_Small	0.895 487	0.875 048	0.859 470	0.865 044
Vit_Small	0.916 865	0.900 375	0.882 246	0.888 922
Swin-Tiny	0.923 990	0.905 362	0.894 646	0.899 157
Swin-Tiny+RDABE	0.935 867	0.920 688	0.926 072	0.923 166

相对于原始的玉米病害数据集,本文比较了使用和不使用 RDABE 增强算法的 Swin-Tiny Transformer 模型的效果.结果表明,使用 RDABE 算法的模型在准确率、精确度、召回率和 F_1 分数等指标上都取得了相对较好的改进,准确率达到 93.586 7%,约提升 1.2%, F_1 分数达到 92.316 6%,约提升 2.4%,证明在加入了 RDABE 算法后的 Swin-Tiny+RDABE 模型提取特征效果优于之前.因此,通过 RDABE 增强算法改进了网络架构的泛化能力,防止网络模型过拟合,并提升了网络模型对图像样本识别的鲁棒性.

使用轻量级模型旨在希望拥有更快的训练速度、更低的计算资源、更好的模型泛化能力、更适合实时应用,更容易部署和维护.表 3 列出了模型的计算量(PLOPs)和参数量(parameter).由表 3 可见,基于随机增强的 Swin-Tiny+RDABE 模型的计算量和参数量与提升前没有变化,并且能提升各项指标.相比其他模型,Swin-Tiny+RDABE 模型在计算量上有一定优势,在参数量上有更多的自由度和灵活性,可以应对更多的数据变化和挑战.

表 3 不同模型的计算量与参数量

Table 3 Calculations and parameters of different models

模型	图像大小	参数量	计算量
Mobilenet_V3_Small	2 242	1.55	0.065 4
ShufflenetV2	2 242	2.30	0.146 0
Efficientnet_B1_Pruned	2 242	5.30	0.400 0
Deit3_Small	2 242	22.10	0.180 0
Vit_Small	2 242	6.60	0.460 0
Swin-Tiny	2 242	28.80	0.130 0
Swin-Tiny+RDABE	2 242	28.80	0.130 0

图 11 为本文 RDABE 算法通过训练得到的混淆矩阵在 Swin-Tiny Transformer 模型使用前后的预测识别结果.由图 11 可见,使用 RDABE 算法的模型在玉米病害类别锈病(Common_rust)的漏报和误差上性能明显改善.因此,可证明 RDABE 算法在提高 Swin-Tiny Transformer 模型性能方面有效.

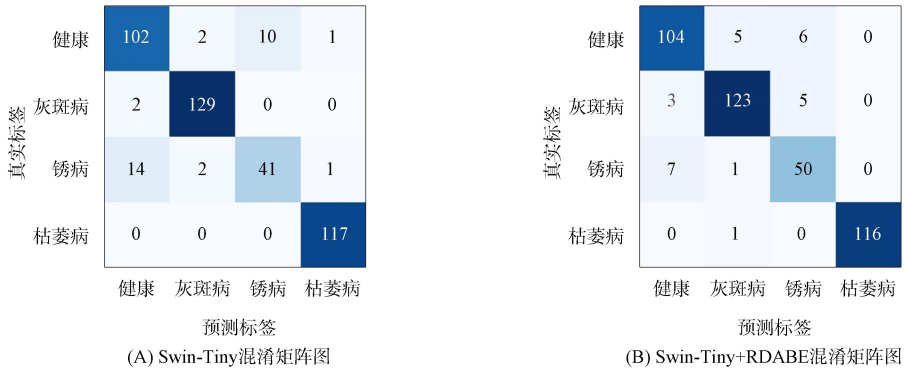


图 11 Swin-Tiny Transformer 混淆矩阵提升对比结果

Fig. 11 Comparison results of improvement of Swin-Tiny Transformer confusion matrix

2.4 小程序部署应用

为更快速、准确地识别玉米病害并提供更直观的体验,本文设计了一个玉米病害智能识别小程序系统,该系统集成了本文训练的玉米病害识别模型.

在小程序开发中,使用 HBuilderX 进行前端开发,包括创建页面、组件和样式,并编写业务逻辑.同时,在 PyCharm 中进行服务端开发,编写服务器端的业务逻辑、数据处理和接口逻辑,并进行本地测试和调试.使用微信开发者工具进行调试和发布,包括创建小程序并配置基本信息,如名称、AppID、服务器地址等.完成开发后,将前端和后端代码上传到微信服务器,并进行代码审核和发布.通过微信开发者工具中的真机调试功能,在 iPhone 12 mini 手机上进行测试和调试,如图 12 所示.

玉米病害识别小程序的主要功能是:通过上传本地图库或现场拍摄的玉米叶片图像,利用 Swin-Tiny+RDABE 模型进行图像检测,以预测该图像可能属于哪种玉米病害.一旦完成上传检测,该小

程序会立即提供关于该图像所属病害的名称、可能性以及相关特点描述信息。因 Swin-Tiny+RDABE 模型是一种轻量级模型, 故可以方便地进行部署和应用。目前, 该小程序已经完成制作并应用, 该算法模型的使用与部署取得了较好成果。

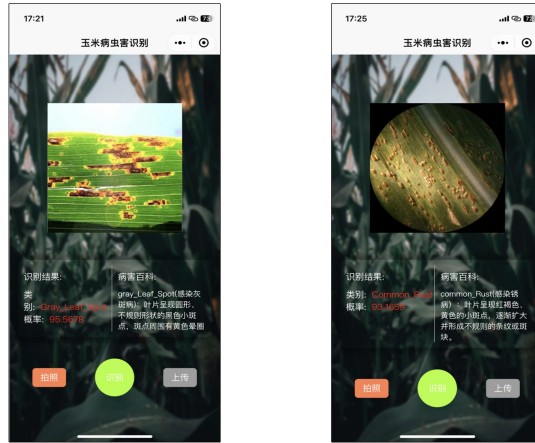


图 12 Swin-Tiny+RDABE 模型小程序的应用效果

Fig. 12 Application effect of Swin-Tiny+RDABE model applet

综上所述, 为解决在图像识别过程中提取全局特征的局限性, 并更好地捕捉整体信息, 本文提出了一种使用迁移学习基于随机增强的轻量级 Swin-Tiny Transformer 模型。通过在农业领域的玉米病害图像检测验证, 证明了该算法的适用性。该方法采用了基于随机数据增强的 RDABE 算法增强玉米叶片数据集, 并对 Swin-Tiny Transformer 网络模型进行微调。实验结果表明, 轻量级 Transformer 系列模型与其他优秀的轻量级 CNN 系列模型相比, Swin-Tiny Transformer 模型在准确率、精确度、召回率和 F_1 分数方面性能更优。在本文模型使用 RDABE 数据增强算法时, 准确率可达 93.586 7%, 相比于不使用该算法时有明显提高。此外, 该方法采用了轻量级设计, 并已成功应用于微信小程序中取得了较好的效果, 实现了实际的应用价值。

参 考 文 献

- [1] LECUN Y, BOTOU L, BENGIO Y, et al. Gradient-Based Learning Applied to Document Recognition [J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [2] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet Classification with Deep Convolutional Neural Networks [J]. Communications of the ACM, 2017, 60(6): 84-90.
- [3] 陈智超, 焦海宁, 杨杰, 等. 基于改进 MobileNet V2 的垃圾图像分类算法 [J]. 浙江大学学报(工学版), 2021, 55(8): 1490-1499. (CHEN Z C, JIAO H N, YANG J, et al. Garbage Image Classification Algorithm Based on Improved MobileNet V2 [J]. Journal of Zhejiang University (Engineering Science), 2021, 55(8): 1490-1499.)
- [4] 邵伟平, 王兴, 曹昭睿, 等. 基于 MobileNet 与 YOLOv3 的轻量化卷积神经网络设计 [J]. 计算机应用, 2020, 40(增刊 1): 8-13. (SHAO W P, WANG X, CAO Z R, et al. Design of Lightweight Convolutional Neural Network Based on MobileNet and YOLOv3 [J]. Computer Applications, 2020, 40(Suppl 1): 8-13.)
- [5] 苏琮智, 杨承志, 邴雨晨, 等. 基于 CNN-Swin Transformer Network 的 LPI 雷达信号识别 [J]. 现代雷达, 2021, 43(2): 1-9. (SU C Z, YANG C Z, BING Y C, et al. LPI Radar Signal Recognition Based on CNN-Swin Transformer Network [J]. Modern Radar, 2021, 43(2): 1-9.)
- [6] VASWANI A, SHAZEER N, PARMAR N, et al. Attention Is All You Need [J]. Advances in Neural Information Processing Systems, 2017, 30(12): 5998-6008.
- [7] 高家军, 张旭, 郭颖, 等. 融合 Swin Transformer 的虫害图像实例分割优化方法研究 [J]. 南京林业大学学报(自然科学版), 2021, 45(5): 1-10. (GAO J J, ZHANG X, GUO Y, et al. Research on Optimization Method for Insect Image Instance Segmentation Based on Swin Transformer Fusion [J]. Journal of Nanjing Forestry University (Natural Sciences Edition), 2021, 45(5): 1-10.)

- [8] 何前, 郭峰林, 方皓正, 等. 基于改进 LeNet-5 模型的玉米病害识别 [J]. 江苏农业科学, 2022, 50(20): 35-41. (HE Q, GUO F L, FANG H Z, et al. Corn Disease Recognition Based on an Improved LeNet-5 Model [J]. Jiangsu Agricultural Sciences, 2022, 50(20): 35-41.)
- [9] 邓朋飞, 官铮, 王宇阳, 等. 基于迁移学习和模型压缩的玉米病害识别方法 [J]. 计算机科学, 2022, 49(增刊 2): 444-449. (DENG P F, GUAN Z, WANG Y Y, et al. Corn Disease Recognition Method Based on Transfer Learning and Model Compression [J]. Journal of Computer Science, 2022, 49(Suppl 2): 444-449.)
- [10] 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, 2021, 49(12): 1-8. (XIONG M Y, ZHAN W, GUI L Y, et al. Detection and Recognition of Maize Leaf Diseases Based on ResNet Model [J]. Jiangsu Agricultural Sciences, 2021, 49(12): 1-8.)
- [11] ALEEM S, KUMAR T, LITTLE S, et al. Random Data Augmentation Based Enhancement: A Generalized Enhancement Approach for Medical Datasets [EB/OL]. (2022-10-03)[2023-01-10]. <https://arxiv.org/abs/2210.00824>.
- [12] 张瑞青, 李张威, 郝建军, 等. 基于迁移学习的卷积神经网络花生荚果等级图像识别 [J]. 农业工程学报, 2020, 36(23): 171-180. (ZHANG R Q, LI Z W, HAO J J, et al. Peanut Pod Grading Image Recognition Based on Transfer Learning Using Convolutional Neural Networks [J]. Transactions of the Chinese Society of Agricultural Engineering, 2020, 36(23): 171-180.)
- [13] 高洪波, 殷俊爽, 李长英. 基于 Swin Transformer 和 CNN 模型融合的包装食品目标检测方法研究 [C]//2022 年中国家用电器技术大会论文集. 北京:《电器》杂志社, 2023: 52-57. (GAO H B, YIN J S, LI C Y. Research on Object Detection Method for Packaged Food Based on Fusion of Swin Transformer and CNN Model [C]// Proceedings of the 2022 China Household Electrical Appliances Technical Conference. Beijing: Electric Journal Press, 2023: 52-57.)
- [14] LIU Z, LIN Y T, CAO Y, et al. Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows [EB/OL]. (2021-08-17)[2023-01-15]. <https://arxiv.org/abs/2103.14030>.
- [15] 祝冰艳, 陈志华, 盛斌. 基于感知增强 Swin Transformer 的遥感图像检测 [J]. 计算机工程, 2021, 47(8): 1-9. (ZHU B Y, CHEN Z H, SHENG B. Remote Sensing Image Detection Based on Perception-Enhanced Swin Transformer [J]. Computer Engineering, 2021, 47(8): 1-9.)
- [16] 郑楚伟, 林辉. 基于 Swin Transformer 的 YOLOv5 安全帽佩戴检测方法 [J]. 计算机测量与控制, 2023, 31(3): 15-21. (ZHENG C W, LIN H. YOLOv5 Helmet Detection Method Based on Swin Transformer [J]. Computer Measurement & Control, 2023, 31(3): 15-21.)
- [17] SINGH D, JAIN N, JAIN P, et al. PlantDoc: A Dataset for Visual Plant Disease Detection [EB/OL]. (2019-11-23)[2023-02-01]. <https://arxiv.org/abs/1911.10317>.
- [18] GEETHARAMANI G, PANDIAN A. Identification of Plant Leaf Diseases Using a 9-Layer Deep Convolutional Neural Network [J]. Computer & Electrical Engineering, 2019, 76: 323-338.

(责任编辑: 韩 啸)