doi:10.6041/j.issn.1000-1298.2026.13.012

基于机器视觉与深度学习的猴头菇质量检测方法

黎　强　汪学勇　孟二从
(西南大学工程技术学院, 重庆 400715)

摘要: 猴头菇作为我国传统的食药两用真菌,其商品化过程中对质量的快速、客观评估提出了迫切需求。针对传统人工感官判定方法效率低、标准不一的问题,提出了一种基于机器视觉与深度学习的猴头菇质量智能检测方法。首先,构建了一种多尺度优化的分水岭图像分割算法,以精准提取猴头菇边界并抑制复杂背景下的过分割现象。随后,通过图像增强与数据增强策略提升图像质量与模型鲁棒性。在分类阶段,引入 Swin Transformer Tiny 网络结构,结合局部窗口注意力与移位窗口机制,实现对猴头菇正常与异常样本的高精度判别。实验结果表明,所提出方法在实际采集图像上取得了 94.3% 的分类准确率,相较于 ResNet－18、EfficientNet、DilateFormer、MambaVision 等主流模型分别提高了 3.5、2.2、1.5、1.2 个百分点,具备良好的工程实用性与部署价值。

关键词: 猴头菇; 机器视觉; 图像分割; Swin Transformer; 质量检测; 深度学习
中图分类号: TP391.41; S646　　**文献标识码**: A　　**文章编号**: 1000-1298(2026)13-0160-06　　**OSID**:

Quality Detection Technology of *Hericium erinaceus* Based on Machine Vision

Li Qiang　Wang Xueyong　Meng Ercong
(*College of Engineering and Technology*, *Southwest University*, *Chongqing* 400715, *China*)

Abstract: *Hericium erinaceus* is a traditional edible and medicinal fungus with high nutritional and economic value, and its commercial circulation requires rapid, objective, and stable quality evaluation. However, manual sensory grading is easily affected by evaluator experience, illumination conditions, and fatigue, which leads to low efficiency and inconsistent judgment results in large-scale production scenarios. Aiming to address these problems, an intelligent quality detection method for *H. erinaceus* based on machine vision and deep learning was proposed. Firstly, images of normal and defective samples were collected under practical acquisition conditions, and preprocessing operations were used to reduce background interference and improve the visibility of surface texture features. A multi-scale optimized watershed segmentation algorithm was then constructed to extract the target region accurately, suppress over-segmentation in complex backgrounds, and provide reliable input for subsequent feature learning. On this basis, image enhancement and data augmentation strategies were introduced to enrich sample diversity, alleviate the influence of limited datasets, and strengthen the generalization ability of the model. In the classification stage, a Swin Transformer Tiny network was adopted to combine local window attention with shifted window attention, so that both fine-grained local texture information and broader contextual relationships can be effectively captured. The model was used to discriminate between normal and defective *H. erinaceus* samples, and its performance was compared with several mainstream deep learning models under the same experimental conditions. Experimental results showed that the proposed method achieved a classification accuracy of 94.3% on real-world collected images, outperforming ResNet－18, EfficientNet, DilateFormer, and MambaVision by 3.5, 2.2, 1.5, and 1.2 percentage points, respectively. The results indicated that the combination of accurate segmentation and attention-based feature extraction can improve the reliability of *H. erinaceus* quality detection. The proposed system demonstrated strong engineering practicability and deployment potential, providing a feasible technical route for intelligent quality inspection of structurally complex edible fungi and supporting the standardization and intelligent development of the edible and medicinal fungus industry.

Key words: *Hericium erinaceus*; machine vision; image segmentation; Swin Transformer; quality detection; deep learning

收稿日期: 2025－09－24　修回日期: 2025－12－03
基金项目: 2025 年西南大学研究生教育教学改革研究重点项目(SWUYJS256102)
作者简介: 黎强(1972—),男,副教授,主要从事计算流体力学与传热学研究,E-mail: xnliqiang@163.com

0　引言

猴头菇(*Hericium erinaceus*)是我国重要的食药两用真菌,在免疫调节、抗氧化和神经保护等方面具有较高研究价值,其营养与药用特性已得到较多研究证实[1-2]。随着食用菌产业规模化生产和商品化流通加快,猴头菇外观品质、完整度及缺陷状态的快速检测逐渐成为生产分级和质量控制中的关键环节。

目前,猴头菇质量判定仍以人工感官经验为主,评价结果易受人员经验、光照条件和主观标准影响,存在效率低、重复性差和标准不统一等问题,难以满足连续化、批量化生产需求[3]。机器视觉具有非接触、速度快和结果可量化等特点,可通过图像采集与特征分析提取目标的形态、颜色和纹理信息,为农产品品质检测提供客观依据[4]。

在农产品质量检测领域,机器视觉已被用于水果分级、茭白自动分级包装、茶叶等级识别和哈密瓜综合品质评估等任务[5-8]。这些研究表明,视觉特征能够有效反映农产品外观品质差异,深度学习模型也能够在复杂纹理和非规则形态识别中表现出较好的判别能力。近年来,图像增强、目标检测、可见/近红外光谱和机器视觉融合方法也被用于苹果、皮蛋、枇杷、板栗、葡萄、茶叶及肉品等对象的检测与分级研究[9-15],为猴头菇质量检测提供了方法参考。

近年来,卷积神经网络和 Transformer 结构不断推动图像分类与目标识别方法发展。ResNet 通过残差连接缓解深层网络训练退化问题,EfficientNet 通过复合缩放策略提升参数利用效率,Swin Transformer 则利用局部窗口注意力与移位窗口机制兼顾局部细节和跨区域上下文建模能力[16-18]。这些方法为猴头菇复杂表面结构的自动识别提供了模型基础。

尽管相关方法已在多类农产品检测中得到应用,但猴头菇具有菌刺密集、边缘不规则、表面纹理细碎且个体形态差异较大的特点,普通分割方法容易受到背景、阴影和局部粘连区域干扰。目前,面向猴头菇质量检测的专用图像分割与分类模型仍相对不足,产业化分级环节仍缺少稳定、可部署的视觉检测方案。

因此,本文围绕猴头菇质量智能检测问题,构建融合图像分割优化、图像增强、数据增强与 Swin Transformer Tiny 分类模型的检测流程。首先,针对猴头菇边界模糊和纹理复杂问题,设计多尺度优化的分水岭分割算法以获得较稳定的目标区域;其次,通过图像增强与数据增强提升样本质量和模型鲁棒性;最后,利用 Swin Transformer Tiny 对正常样本与异常样本进行分类识别,为猴头菇自动化质量分级提供技术参考。

1　图像分割模型

分水岭算法是一种基于数学形态学和拓扑理论的图像分割方法,其基本思想是将灰度图像看作地形表面,将局部极小值视为汇水盆地,并通过模拟水位上升过程确定不同区域之间的分界线。该方法适合处理边界相互接近或局部粘连的目标,但在噪声较多、纹理细碎和背景复杂的图像中容易产生过分割。针对猴头菇菌刺密集、边界模糊和外形不规则等特点,本文在传统分水岭算法基础上引入多尺度处理与投票融合机制,以提升分割结果的稳定性。

具体流程为:首先对原始图像进行高斯模糊处理,以削弱采集噪声和局部纹理干扰;随后采用 Otsu 自适应阈值法完成二值化,增强前景与背景的灰度区分;再利用形态学开运算去除残余噪点,并通过距离变换强化目标前景区域。算法分别在原始尺度及两个缩放尺度下执行分水岭分割,得到多组初始分割结果;随后将不同尺度结果统一重采样至原图尺寸,并采用像素级投票方式进行融合,形成最终分割掩膜,从而缓解单一尺度选择导致的过分割和边界缺失问题。

分水岭算法的数学原理可表述如下。设输入灰度图像为 $f(x,y)$,其中 (x,y) 表示像素坐标,$f(x,y)$ 表示该像素的灰度。算法首先在灰度地形中搜索局部极小值点,并将其作为候选汇水盆地的初始标记集合

$$M = \{m_1, m_2, \cdots, m_k\} \tag{1}$$

式中　M——局部极小值集合

m_i——第 i 个局部极小值点

k——局部极小值点的数量

在此基础上,图像区域 Ω 可划分为 k 个互不相交的子区域 C_i,即

$$\Omega = \bigcup_{i=1}^{k} C_i \quad (C_i \cap C_j = \varnothing, i \neq j) \tag{2}$$

式中　C_i——由第 i 个标记点扩展得到的分割区域

对于二值化图像 $B(x,y)$,采用距离变换定义前景区域的种子点,距离函数可表示为

$$D(x,y) = \min(((x-x')^2 + (y-y')^2)^{1/2}) \tag{3}$$

$$S_f = \{(x,y) \mid D(x,y) > \theta\} \tag{4}$$

式中　S_f——前景像素集合

$D(x,y)$——像素点 (x,y) 到最近背景像素的欧氏距离

(x',y')——邻域像素坐标

θ——种子点筛选阈值

根据距离变换结果选取置信度较高的前景区域作为种子点,并结合连通域标记生成初始标记矩阵。最终,将初始标记矩阵输入分水岭变换,不同标记区域在扩展过程中相遇的位置即为分水岭线,也就是目标与背景或相邻区域之间的分割边界

$$W = \{(x,y) \mid (x,y) \in \partial C_i \cap \partial C_j, i \neq j\} \tag{5}$$

式中 W——分水岭线集合

图 1 展示了改进分水岭算法在猴头菇图像分割过程中的关键处理流程,包括原始图像、二值化处理、去噪操作、距离变换、标记提取以及最终分割结果。可以看出,多尺度融合后的分割结果能够较完整地保留猴头菇主体边缘,并降低复杂纹理造成的误分割影响。

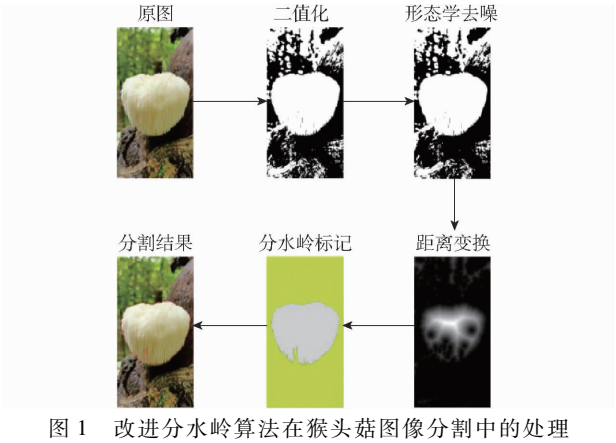

图 1 改进分水岭算法在猴头菇图像分割中的处理流程示意图

Fig. 1 Illustration of multi-scale watershed segmentation process on *Hericium erinaceus* images

2 猴头菇质量检测模型

2.1 图像增强

图像增强可提升原始图像的视觉清晰度与结构特征表达能力,从而提高模型对关键属性的识别性能。本文采用直方图均衡化与线性灰度拉伸两种方法对猴头菇图像进行增强处理。直方图均衡化通过拉伸灰度分布提升对比度。设图像灰度为 r,对应的概率密度函数为

$$p(r) = \frac{n_r}{n} \tag{6}$$

式中 $p(r)$——灰度 r 出现的概率

n_r——灰度为 r 的像素数

n——图像总像素数

映射变换函数定义为

$$s = T(r) = \int_0^r p(w)\,\mathrm{d}w \tag{7}$$

式中 s——线性拉伸后的灰度输出值

线性灰度拉伸将原灰度区间$[a,b]$映射至目标范围$[c,d]$,增强图像局部细节的可辨性。其映射函数为

$$s = \frac{r-a}{b-a}(d-c) + c \tag{8}$$

图 2 展示了原始图像与增强图像的视觉对比效果。处理后的图像呈现出更清晰的结构边界和纹理特征,有助于后续的特征提取与判别建模。

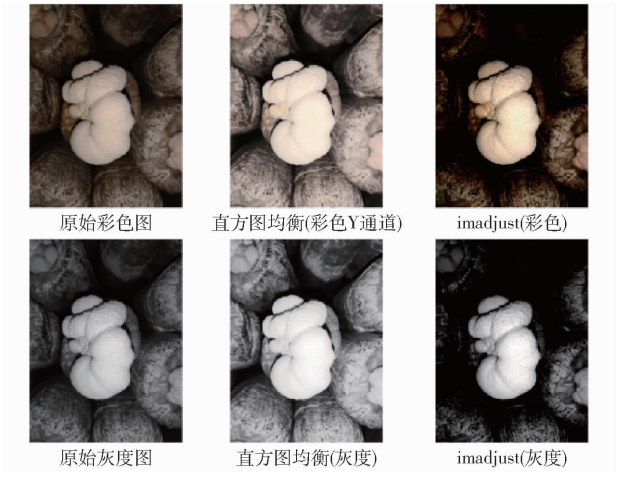

图 2 猴头菇图像增强对比

Fig. 2 Comparison of *Hericium erinaceus* images before and after enhancement

2.2 数据增强

为提高模型对样本分布变化的适应能力,增强其在不同采集条件下的鲁棒性,本文设计了包含几何变换与像素扰动在内的多种数据增强方法。

在几何空间中,通过随机执行水平翻转与垂直翻转操作,以设定概率 P_{aug} 对图像进行镜像变换,增强模型对结构对称性的识别能力。随机旋转将图像旋转角度 θ_r 限制在设定范围内,以提高模型对不同朝向目标的判别效果。随机裁剪操作生成不同视野的子图像,其输出尺寸按原图比例缩放,模拟视野遮挡与结构缺失场景。图像平移扰动用于模拟位置偏移,其最大幅度不超过图像宽度或高度的 20%。透视变换通过控制变换矩阵改变图像几何关系,增强模型对复杂形变的感知能力。在像素域处理方面,采用亮度扰动方法对图像亮度 l 增加扰动量 Δl,使变换后的亮度 l' 满足

$$l' = l + \Delta l \tag{9}$$

该策略提高了模型对光照变化的适应性。高斯模糊使用高斯核函数对图像卷积处理,模拟模糊或低质量采集环境。

上述增强策略在不改变图像标签的前提下显著增加了样本多样性,并在实验证明中有效提升了模型泛化性能。图 3 展示了典型增强方式下猴头菇图像的可视化结果,包括水平翻转、亮度调整、随机旋转、高斯模糊、平移、裁剪与透视变换等典型增强手

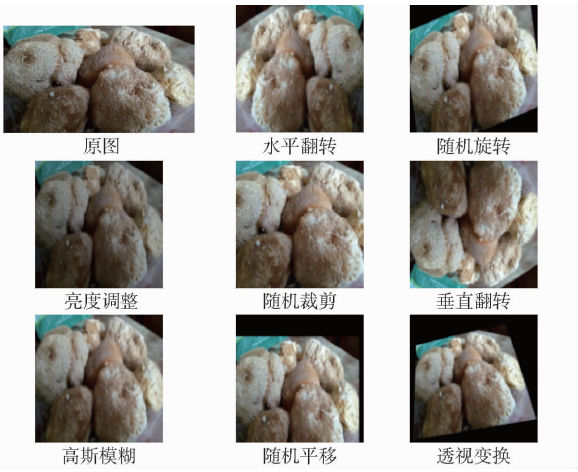

图 3　猴头菇图像主要数据增强方式示意图

Fig. 3　Demonstration of major data augmentation techniques for *Hericium erinaceus* images

段生成的图像样本。

2.3　基于 Swin Transformer 的分类器

为提升对猴头菇质量等级的识别性能，本文引入 Swin Transformer Tiny 网络作为图像分类模型的主干结构。该模型结合局部建模与窗口迁移机制，在提取表面纹理、结构边界等多尺度视觉特征方面具有显著优势。

4 ×4C 输入图像首先被划分为 4 ×4 大小的非重叠 patch，每个 patch 经展平后通过线性映射投影至嵌入维度 C，公式为

$$X = \text{PatchEmbedding}(I) \tag{10}$$

嵌入后的图像特征依次通过 4 个层级 Stage，每个 Stage 由多个 Swin Transformer Block 构成。每个 Block 采用窗口多头自注意力（W - MSA）与移位窗口多头自注意力（SW - MSA）交替堆叠。第 l 层 Block 的前向传播公式为

$$z^l = \text{W - MSA}(\text{LN}(z^{l-1})) + z^{l-1} \tag{11}$$

$$z^l = \text{MLP}(\text{LN}(z^l)) + z^l \tag{12}$$

式中　z^l——第 l 层输出特征

LN——层归一化操作

MLP——两层全连接网络

该结构既保留了图像的局部细节，又实现了跨区域的信息整合。

4 个 Stage 输出的特征 F 经过全局平均池化处理，得到特征向量 $\boldsymbol{f}$，最终通过全连接层和 Softmax 函数输出预测类别 y，即

$$\boldsymbol{f} = \text{GlobalAveragePooling}(F) \tag{13}$$

$$y = \text{Softmax}(\boldsymbol{W}\boldsymbol{f} + \boldsymbol{b}') \tag{14}$$

式中　$\boldsymbol{W}$——全连接层权重矩阵

$\boldsymbol{b}'$——偏置项

Swin Transformer 结构在特征提取效率与判别能力之间取得了良好平衡，适应猴头菇图像在复杂纹理、密集菌刺与形态差异方面的表达需求，验证了其在食用菌质量检测任务中的可行性与有效性。

如图 4 所示，本文所采用的 Swin Transformer Tiny 模型由 3 个主要模块构成：Patch Embedding 模块、编码器与分类头，其中编码器采用多层次的局部窗口与移位窗口注意力结构。

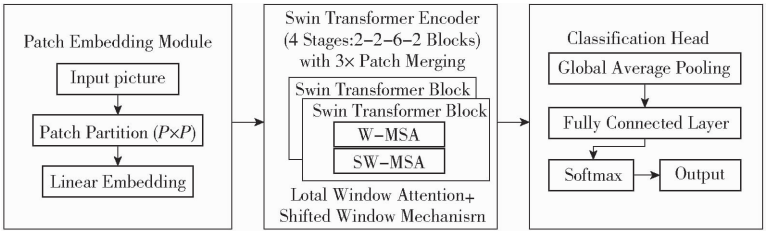

图 4　Swin Transformer Tiny 模型结构图

Fig. 4　Architecture of Swin Transformer Tiny model

3　结果分析

3.1　数据集构建

为验证所提出模型的有效性，本文构建了一套包含猴头菇图像的专用数据集。数据来源于实际生产场景下的采集，图像共计 1 000 幅，涵盖正常猴头菇和存在质量缺陷的样本各 500 幅。

图 5 是基于图像可视质量特征展示猴头菇图像数据集中样本的分类标准。图像在清洗与预处理后，根据是否存在明显缺陷划分为正常类与异常类。异常类包括霉变、损伤和表面污染 3 种典型可视缺陷类型。分类结果用于构建训练标签，模型训练采用二分类方式。

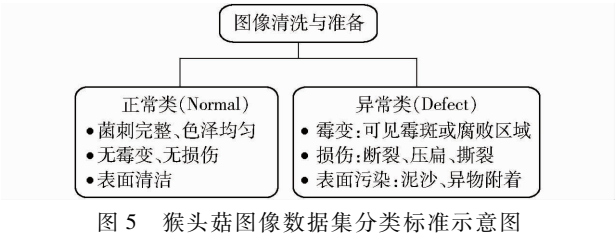

图 5　猴头菇图像数据集分类标准示意图

Fig. 5　Classification criteria diagram of *Hericium erinaceus* image dataset

所有图像分辨率一致，经过清洗与标注后，按照 80% 与 20% 的比例划分为训练集与测试集。数据分布情况如表1所示，确保模型训练与评估过程中

表 1 猴头菇图像数据集构成

Tab. 1 Composition of *Hericium erinaceus* image dataset

数据类别	图像数量	训练集数量	测试集数量
正常猴头菇	500	400	100
异常猴头菇	500	400	100
合计	1 000	800	200

的样本均衡性与代表性。

3.2 实验环境配置及参数设置

实验在性能充足的图像识别工作站上进行。硬件平台包括 Intel Core i7 - 13700K 处理器,主频 3.4 GHz,搭载 NVIDIA RTX 4080 GPU,配备 16 GB 显存与 32 GB DDR5 内存。数据存储采用 1 TB SSD,以确保大规模图像在训练过程中的高速访问能力。

软件环境构建于 Windows 11 操作系统,使用 Python 3.10 作为开发语言,深度学习任务通过 PyTorch 2.3 框架实现。图像预处理与增强过程主要依赖 OpenCV 4.8 与 NumPy 1.26,训练与结果可视化则使用 Matplotlib 3.8。

模型训练过程中采用 Adam 优化器,初始学习率设为 0.001,批次大小为 32,最大迭代轮数为 50 个 epoch,损失函数为交叉熵损失。在训练过程中,所有模型使用相同数据划分与数据增强策略,确保对比分析的公正性与一致性。

3.3 消融实验分析

为评估图像增强与数据增强策略对模型性能的贡献,设计了 3 组消融实验,依次引入不同的增强机制,观察模型分类能力的变化趋势。

第 1 组实验采用未经任何增强处理的原始图像作为输入,构成 Baseline 方案;第 2 组实验在 Baseline 基础上引入直方图均衡化与线性灰度拉伸增强方法;第 3 组实验进一步叠加了数据增强操作,包括旋转、翻转、亮度调整与模糊处理等策略,构成完整增强模型。3 组实验在相同网络架构下进行训练与测试,其分类准确率、F1 分数与召回率结果如表 2 所示。

表 2 消融实验结果对比

Tab. 2 Comparison of ablation experiment results

实验设置	准确率/%	F1 分数	召回率/%
Baseline	85.50	0.854	84.20
Baseline + 图像增强	89.80	0.896	88.70
Baseline + 图像增强 + 数据增强	93.70	0.934	93.10

实验结果表明,图像增强显著提升了特征清晰度和模型判别力,而数据增强进一步提升了模型在测试集上的泛化能力,最终使准确率较 Baseline 方案提高 8 个百分点以上,验证了增强策略的有效性。

3.4 模型对比实验

在综合评估所提出方法的性能基础上,本文将 Swin Transformer Tiny 分类器与主流卷积神经网络模型进行了对比分析,涵盖结构复杂度、分类准确率、F1 分数以及推理时间 4 项指标。对比模型包括传统的 Simple - CNN、作为经典残差网络结构且在图像识别任务中被广泛应用的 ResNet - 18[16]、基于复合缩放策略并在参数效率与精度之间实现良好平衡的 EfficientNet - B0[17]、引入多尺度膨胀注意力机制并兼顾局部与全局建模能力的 DilateFormer[19],以及融合状态空间建模结构与 Transformer 自注意力模块的 MambaVision[20]。为确保公平性,全部模型均在统一实验环境与数据集上进行训练测试,并使用与本文 Swin - T 相同的数据划分、图像增强和数据增强策略。

表 3 汇总了各模型的对比结果。从结果可见,Swin - T 在准确率与 F1 分数方面均优于其他模型,推理时间略高但仍保持在毫秒每幅级别,适用于实际生产场景下的在线检测任务。

表 3 不同模型结构性能对比

Tab. 3 Performance comparison of different model architectures

模型结构	参数量	准确率/%	F1 分数	推理时间/(ms·f^{-1})
Simple - CNN(Baseline)	1.20×10^6	85.50	0.854	0.65
ResNet - 18	1.17×10^7	90.80	0.908	1.20
EfficientNet - B0	5.30×10^6	92.10	0.920	1.35
DilateFormer	1.70×10^7	92.80	0.931	1.39
MambaVision	3.18×10^7	93.10	0.935	1.45
Swin - T(本文)	2.83×10^7	94.30	0.945	1.58

对比结果表明,尽管 Swin - T 模型参数量相对较大,但在保持较高准确率与鲁棒性的同时,仍具备较快的推理效率,具有较强的工程实用性,适合部署于农产品自动分级系统中进行高精度质量检测。

4 结论

(1)针对猴头菇商品化过程中人工评定效率低、标准不统一的问题,本文构建了融合图像分割优化、图像增强、数据增强与 Swin Transformer Tiny 分类模型的智能检测方法,形成了从图像预处理到质量分类的完整技术流程。

(2)在图像分割阶段,本文设计了多尺度优化分水岭算法,结合高斯模糊、Otsu 自适应阈值、距离变换与尺度融合机制,改善了复杂背景下猴头菇边缘提取不完整和过分割问题,提高了目标区域获取

的稳定性。

(3)在图像分类阶段,Swin Transformer Tiny 通过局部窗口建模与跨区域特征整合提取猴头菇表面纹理、菌刺和结构边界等关键特征;同时结合直方图均衡化、灰度拉伸、旋转、翻转、亮度调整、模糊和透视变换等增强策略,提高了模型对不同采集条件的鲁棒性。

(4)本文构建的数据集包含 1 000 幅猴头菇图像,其中正常样本和异常样本各 500 幅,并按比例 8∶2划分为训练集和测试集。消融实验结果表明,基础模型准确率为 85.5%;引入图像增强后提升至 89.8%;同时采用图像增强与数据增强后提升至 93.7%,说明增强策略能够有效改善模型泛化能力。

(5)在完整增强策略与 Swin Transformer Tiny 支持下,最终模型分类准确率达到 94.3%,F1 分数为 0.945,召回率为 93.1%。与 ResNet－18、EfficientNet－B0、DilateFormer 和 MambaVision 相比,准确率分别提升 3.5、2.2、1.5、1.2 个百分点,同时保持毫秒每幅级推理时间,具有较好的工程部署潜力。

参　考　文　献

[1] 张宝翠, 刘晓鹏, 朱玉昌, 等. 猴头菇的研究进展[J]. 食品安全质量检测学报, 2019, 10(8): 2285－2292.
Zhang Baocui, Liu Xiaopeng, Zhu Yuchang, et al. Research progress on *Hericium erinaceus*[J]. Journal of Food Safety and Quality Detection, 2019, 10(8): 2285－2292. (in Chinese)

[2] 张微思. 猴头菇的营养药用价值及产品研究现状[J]. 食品与发酵科技, 2018, 49(1): 104－108.
Zhang Weisi. Nutritional and medicinal value of *Hericium erinaceus* and current status of product research[J]. Food and Fermentation Science and Technology, 2018, 49(1): 104－108. (in Chinese)

[3] 刘波. 计算机视觉研究综述[J]. 数字通信世界, 2019(12): 97.
Liu Bo. Review of computer vision research[J]. Digital Communication World, 2019(12): 97. (in Chinese)

[4] 李志伟, 霍静琦, 蒿晟昆. 机器视觉技术在农业智能装备中应用的研究进展[J]. 当代农机, 2021(7): 5－7.
Li Zhiwei, Huo Jingqi, Hao Chengkun. Research progress on application of machine vision technology in agricultural intelligent equipment[J]. Modern Agricultural Machinery, 2021(7): 5－7. (in Chinese)

[5] 杨再雄, 吴恋, 左建, 等. 基于人工智能的农产水果分级检测技术综述[J]. 科技创新与应用, 2021, 11(22): 41－43.
Yang Zaixiong, Wu Lian, Zuo Jian, et al. Review of fruit grading and detection technology for agricultural products based on artificial intelligence[J]. Science and Technology Innovation and Application, 2021, 11(22): 41－43. (in Chinese)

[6] 史勇, 郭俊先, 李雪莲, 等. 哈密瓜无损检测技术及分级装备的研究现状和发展趋势[J]. 包装与食品机械, 2021, 39(6): 75－82.
Shi Yong, Guo Junxian, Li Xuelian, et al. Status and development trend of nondestructive testing technology and grading equipment for Hami melon[J]. Packaging and Food Machinery, 2021, 39(6): 75－82. (in Chinese)

[7] 张文蓉, 王朋, 钟鸣, 等. 基于视觉的茭白自动分级包装设备研究与开发[J]. 中国农机化学报, 2021, 42(12): 244－250.
Zhang Wenrong, Wang Peng, Zhong Ming, et al. Research and development of vision-based automatic grading and packaging equipment for *Zizania latifolia*[J]. Journal of Chinese Agricultural Mechanization, 2021, 42(12): 244－250. (in Chinese)

[8] 俞文昌, 徐雅雯, 程英杰, 等. 卷积神经网络在茶叶分级中的应用[J]. 电脑知识与技术, 2021, 17(23): 102－104.
Yu Wenchang, Xu Yawen, Cheng Yingjie, et al. Application of convolutional neural networks in tea grading[J]. Computer Knowledge and Technology, 2021, 17(23): 102－104. (in Chinese)

[9] 宋怀波, 杨涵茹, 苏晓薇, 等. 基于 EnlightenGAN 图像增强的自然场景下苹果检测方法[J]. 农业机械学报, 2024, 55(8): 266－279.
Song Huaibo, Yang Hanru, Su Xiaowei, et al. Application of image enhancement technology based on EnlightenGAN in apple detection in natural scenes[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(8): 266－279. (in Chinese)

[10] 汤文权, 陈灼廷, 王东桥, 等. 基于改进 YOLO v5 的皮蛋裂纹在线检测方法[J]. 农业机械学报, 2024, 55(2): 384－392.
Tang Wenquan, Chen Zhuoting, Wang Dongqiao, et al. Online detection method for preserved egg cracks based on improved YOLO v5[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(2): 384－392. (in Chinese)

[11] 赵茂程, 邹涛, 齐亮, 等. 基于 MobileViT－CBAM 的枇杷表面缺陷检测方法[J]. 农业机械学报, 2024, 55(9): 420－427.
Zhao Maocheng, Zou Tao, Qi Liang, et al. Loquat surface defect detection method based on MobileViT－CBAM[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(9): 420－427. (in Chinese)

[12] 姜洪喆, 谭烽, 李兴鹏, 等. 转盘式板栗可见/近红外光谱检测分选系统研究[J]. 农业机械学报, 2024, 55(12): 462－469.
Jiang Hongzhe, Tan Feng, Li Xingpeng, et al. Research of rotary sorting system for chestnuts based on visible/near-infrared spectroscopy[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(12): 462－469. (in Chinese)

(下转第 199 页)

243－251.
Huang Hua, Zhang Hao, Hu Xiaolin, et al. Recognition and localization method for peper clusters in complex environments based on improved YOLO v5[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(3): 243－251. (in Chinese)
[10] 张立杰,周舒骅,李娜.基于改进 SSD 卷积神经网络的苹果定位与分级方法[J]. 农业机械学报,2023,54(6):223－232.
Zhang Lijie, Zhou Shuhua, Li Na. Positioning and grading method of apples based on improved SSD convolutional neural network[J]. Transactions of the Chinese Society for Agricultural Machinery,2023,54(6):223－232. (in Chinese)
[11] 翟肇裕,张梓涵,徐焕良,等. YOLO 算法在动植物表型研究中应用综述[J]. 农业机械学报, 2024, 55(11): 1－20.
Zhai Zhaoyu, Zhang Zihan, Xu Huanliang, et al. Review of applying YOLO family algorithms to analyze animal and plant phenotype[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(11): 1－20. (in Chinese)
[12] Jiang K L, Hu W J, Han P. Multi-task segmentation of oilseed rape and weeds in agricultural fields: a hybrid approach combining enhanced UNet and unsupervised clustering[J]. Computers and Electronics in Agriculture,2025,238:110827.
[13] Ding Y, Yang W, Zhang J. An improved DeepLabV3 + based approach for disease spot segmentation on apple leaves[J]. Computers and Electronics in Agriculture,2025,231:110041.
[14] 张勤,庞月生,李彬. 基于实例分割的番茄串视觉定位与采摘姿态估算方法[J]. 农业机械学报, 2023, 54 (10): 205－215.
Zhang Qin, Pang Yuesheng, Li Bin. Visual positioning and picking pose estimation of tomato clusters based on instance segmentation[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(10): 205－215. (in Chinese)
[15] Li Z, Wang Z, Zhang H, et al. Light－AdaPSPNet: a novel lightweight adaptive weighted semantic segmentation network based on PSPNet for extracting *Spartina alterniflora* from dual-polarimetric SAR images[J]. Computers and Electronics in Agriculture,2026,242:111338.
[16] Meng K, Xu K, Cattani P, et al. Camouflaged cotton bollworm instance segmentation based on PVT and Mask R－CNN[J]. Computers and Electronics in Agriculture,2024,226:109450.
[17] Li Z, Shen Y, Tang J, et al. IMLL－DETR: an intelligent model for detecting multi-scale litchi leaf diseases and pests in complex agricultural environments[J]. Expert Systems with Applications,2025,273:126816.
[18] Zhao Z, Chen S, Ge Y, et al. RT－DETR－Tomato: tomato target detection algorithm based on improved RT－DETR for agricultural ssafety production[J]. Applied Sciences,2024,14(14):6287.
[19] Zhou H, Jiang F, Lu H. SSDA－YOLO: semi-supervised domain adaptive YOLO for cross-domain object detection[J]. Computer Vision and Image Understanding, 2023, 229: 103649.
[20] Oquab M, Darcet T, Moutakanni T, et al. Dinov2: learning robust visual features without supervision[PP/OL]. V1. arXiv (2023－04－14)[2026－03－08]. https://arXiv.org/abs/2304.07193.
[21] Qiu Z, Wang Z, Zheng B, et al. Gated attention for large language models: non-linearity, sparsity, and attention-ssink-free [PP/OL]. V1. arXiv (2025－05－10)[2026－03－08]. https://arXiv.org/abs/2505.06708.

(上接第 165 页)

[13] 刘文政, 周雪健, 平凤娇, 等. 基于可见-近红外光谱的鲜食葡萄成熟品质关键指标检测[J]. 农业机械学报, 2024, 55(2): 372－383.
Liu Wenzheng, Zhou Xuejian, Ping Fengjiao, et al. Detection of key maturity quality indicators of grapes based on visible-near infrared spectroscopy[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(2): 372－383. (in Chinese)
[14] Zou H, Lan T, Jiang Y, et al. Research on rapid detection methods of tea pigments content during rolling of black tea based on machine vision technology[J]. Foods, 2024, 13(23): 3718.
[15] Xu Z, Han Y, Zhao D, et al. Research progress on quality detection of livestock and poultry meat based on machine vision, hyperspectral and multi-source information fusion technologies[J]. Foods, 2024, 13(3): 445.
[16] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 770－778.
[17] Tan M, Le Q V. EfficientNet: rethinking model scaling for convolutional neural networks[C]// Proceedings of the 36th International Conference on Machine Learning. Cambridge, U.S.A.: PMLR, 2019: 6105－6114.
[18] Liu Z, Lin Y, Cao Y, et al. Swin Transformer: hierarchical vision transformer using shifted windows[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2021: 10012－10022.
[19] Jiao J, Tang Y M, Lin K Y, et al. DilateFormer: multi-scale dilated transformer for visual recognition[J]. IEEE Transactions on Multimedia, 2023,25:8906－8919.
[20] Hatamizadeh A, Kautz J. MambaVision: a hybrid Mamba-Transformer Vision backbone[C]// Proceedings of the Computer Vision and Pattern Recognition Conference. Piscataway: IEEE, 2025: 25261－25270.