北京林业大学学报
JOURNAL OF BEIJING FORESTRY UNIVERSITY

基于树皮纹理的轻量化YOLOv11树种识别方法

张政银 向玮 刘子锋 王俊文 张咪 杨俊俐 黄泽园

Lightweight YOLOv11 tree species recognition method based on bark texture

Zhang Zhengyin, Xiang Wei, Liu Zifeng, Wang Junwen, Zhang Mi, Yang Junli, Huang Zeyuan

引用本文:
张政银, 向玮, 刘子锋, 王俊文, 张咪, 杨俊俐, 黄泽园. 基于树皮纹理的轻量化YOLOv11树种识别方法[J]. 北京林业大学学报, 2025, 47(8):134–148. doi: 10.12171/j.1000–1522.20250151
Zhang Zhengyin, Xiang Wei, Liu Zifeng, Wang Junwen, Zhang Mi, Yang Junli, Huang Zeyuan. Lightweight YOLOv11 tree species recognition method based on bark texture[J]. Journal of Beijing Forestry University, 2025, 47(8):134–148. doi: 10.12171/j.1000–1522.20250151

在线阅读 View online: https://doi.org/10.12171/j.1000–1522.20250151

您可能感兴趣的其他文章

Articles you may be interested in

MAFF–YOLO: 面向造林验收的明穴目标检测模型

MAFF–YOLO: a target detection model for planting holes in afforestation acceptance

北京林业大学学报, 2025, 47(4): 142–154. https://doi.org/10.12171/j.1000–1522.20240353

基于无人机影像和MDIEA–YOLO苗木识别模型的造林验收智能系统

An intelligent system for afforestation acceptance based on UAV images and MDIEA–YOLO seedling recognition model

北京林业大学学报, 2025, 47(5): 14–25. https://doi.org/10.12171/j.1000–1522.20240354

改进YOLOv8n的林业害虫检测方法

Forestry pest detection method based on improved YOLOv8n

北京林业大学学报, 2025, 47(2): 119–131. https://doi.org/10.12171/j.1000–1522.20240326

结合多尺度纹理特征的高光谱影像面向对象树种分类

Object–oriented tree species classification with multi–scale texture features based on airborne hyperspectral images

北京林业大学学报, 2020, 42(6): 91–101. https://doi.org/10.12171/j.1000–1522.20190155

基于YOLOv7的红外相机野生动物图像筛选

Wildlife image screening for infrared cameras based on YOLOv7

北京林业大学学报, 2024, 46(2): 143–154. https://doi.org/10.12171/j.1000–1522.20230112

基于域对抗网络的林火烟雾图像跨域识别方法

Cross–domain recognition methods of forest fire smoke images based on domain adversarial network

北京林业大学学报, 2025, 47(6): 130–140. https://doi.org/10.12171/j.1000–1522.20250006

DOI:10.12171/j.1000−1522.20250151

基于树皮纹理的轻量化 YOLOv11 树种识别方法

张政银[1,2] 向 玮[1] 刘子锋[3] 王俊文[4] 张 咪[1] 杨俊俐[5] 黄泽园[6]

（1. 北京林业大学林学院，北京 100083；2. 中国矿业大学（北京）理学院，北京 100083；3. 北京化工大学信息科学与技术学院，北京 100029；4. 中国矿业大学（北京）人工智能学院，北京 100083；5. 北京邮电大学国际学院，北京 100876；6. 北京邮电大学叶培大创新创业学院，北京 100876）

摘要:【目的】为解决现有树种识别方法在多变光照条件下，因模型计算复杂度高而难以在硬件条件极端受限的移动端或边缘设备部署的问题，提出一种基于树皮纹理的轻量化树种识别方法。**【方法】**本研究通过改进 YOLOv11 构建 YOLOv11-SWER 模型。首先，引入轻量化特征提取网络 StarNet 作为主干网络，结合深度可分离卷积与通道混洗机制，显著降低网络的参数量和特征提取过程中的计算量。其次，采用多分支特征融合模块 RepNCSPELAN4，结合分组卷积与参数共享策略，兼顾全局特征与局部特征，提升多尺度特征融合效率。然后，设计小波池化（WaveletPool）层，减少噪声干扰并保留高频纹理细节，增强模型对树皮纹理微小特征的捕捉能力。最后，优化检测头结构 Detect_Efficient，使用双分支分组卷积结构提高计算效率。同时，基于自建的 70 类树种、6 681 张树皮图像数据集，通过消融实验和对比实验对改进的模型性能进行充分的评估验证。**【结果】**该模型的检测精确率、召回率、平均精度（mAP50）、平均精度均值（mAP50-95）以及精确率和召回率的调和平均数 F_1 分数分别达到 98.1%、98.4%、0.993、0.750 和 0.982，同时，相较于 YOLOv11 模型，其参数量和计算量分别降低 46.92% 和 51.5%，大幅降低了模型的空间复杂度和计算复杂度。在不同光照场景下保持稳定的识别性能，展现出良好的光照鲁棒性。**【结论】**本研究提出的 YOLOv11-SWER 模型通过轻量化设计与多尺度特征优化，在参数量减少近半的情况下，仍能保持高检测精度，实现了高检测精度与高效率的良好平衡。此方法有望应用于智能林业检测、城市林业资源管理等场景中。

关键词：树皮纹理识别；YOLOv11；StarNet 主干网络；小目标检测；轻量化

中图分类号：TP391.41；S718.4 **文献标志码：**A **文章编号：**1000−1522(2025)08−0134−15

引文格式：张政银，向玮，刘子锋，等. 基于树皮纹理的轻量化 YOLOv11 树种识别方法［J］. 北京林业大学学报，2025，47(8)：134−148. Zhang Zhengyin, Xiang Wei, Liu Zifeng, et al. Lightweight YOLOv11 tree species recognition method based on bark texture［J］. Journal of Beijing Forestry University, 2025, 47(8): 134−148.

Lightweight YOLOv11 tree species recognition method based on bark texture

Zhang Zhengyin[1,2] Xiang Wei[1] Liu Zifeng[3] Wang Junwen[4] Zhang Mi[1]
Yang Junli[5] Huang Zeyuan[6]

(1. School of Forestry, Beijing Forestry University, Beijing 100083, China;
2. School of Science, China University of Mining and Technology (Beijing), Beijing 100083, China;
3. College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China;
4. School of Artificial Intelligence, China University of Mining and Technology (Beijing), Beijing 100083, China;
5. International School, Beijing University of Posts and Telecommunications, Beijing 100876, China;
6. Ye Peida School of Innovation and Entrepreneurship, Beijing University of Posts and Telecommunications, Beijing 100876, China)

收稿日期: 2025−04−27 **修回日期:** 2025−08−01

基金项目:“十四五”国家重点研发计划（2022YFD2200502-04），北京高校大学生创新创业训练项目（202498117）。

第一作者: 张政银。主要研究方向：计算机视觉。Email: 2210730330@student.cumtb.edu.cn 地址：100083 北京市海淀区学院路丁 11 号中国矿业大学（北京）理学院。

责任作者: 向玮，教授，博士生导师。主要研究方向：天然林生长模型及模拟研究、树轮生态学等。Email: wxiang@bjfu.edu.cn 地址：100083 北京市海淀区清华东路 35 号北京林业大学林学院。

本刊网址: http://j.bjfu.edu.cn；http://journal.bjfu.edu.cn

Abstract: [**Objective**] To address the issue of existing tree species identification methods being difficult to deploy on mobile or edge devices with extremely limited hardware due to high computational complexity under varying lighting conditions, this study proposes a lightweight tree species identification method based on bark texture. [**Method**] This research improves YOLOv11 to construct the YOLOv11-SWER model. First, the lightweight feature extraction network StarNet was introduced as the backbone network, combining depthwise separable convolution and channel shuffle mechanisms to significantly reduce the model's parameter count and computational load during feature extraction. Second, a multi-branch feature fusion module, RepNCSPELAN4, was adopted, integrating group convolution and parameter-sharing strategies to balance global and local features, thereby enhancing multi-scale feature fusion efficiency. Then, a wavelet pooling (WaveletPool) layer was designed to reduce noise interference while preserving high-frequency texture details, improving the model's ability to capture subtle bark texture features. Finally, the detection head structure Detect_Efficient was optimized using a dual-branch group convolution architecture to improve computational efficiency. Additionally, based on a self-built dataset of 70 tree species with 6 681 bark images, ablation studies and comparative experiments were conducted to thoroughly evaluate the performance of improved model. [**Result**] The model achieved a detection precision of 98.1%, recall of 98.4%, mean average precision (mAP50) of 0.993, mean average precision across different IoU thresholds (mAP50-95) of 0.750, and an F_1 score (harmonic mean of precision and recall) of 0.982. Compared with the YOLOv11 model, the parameter count and computational load were reduced by 46.92% and 51.5%, respectively, significantly lowering the model's spatial and computational complexity. It maintained stable identification performance under varying lighting conditions, demonstrating strong illumination robustness. [**Conclusion**] The proposed YOLOv11-SWER model, through lightweight design and multi-scale feature optimization, achieves high detection accuracy while reducing parameters by nearly half, striking a good balance between high accuracy and efficiency. This method holds promise for applications in intelligent forestry monitoring and urban forestry resource management.

Key words: bark texture recognition; YOLOv11; StarNet backbone network; small object detection; lightweight

城市林业资源的高效管理是生态城市建设的基石，而准确识别树种则是实现精细化保护与可持续利用的前提。随着移动终端算力的提升，将高精度且轻量化的树种识别模型部署到手机和平板等便携设备，不仅能够把专业知识即时推送给公众，显著提升科普互动效果，还能为生态文明建设提供实时、便捷的科技支撑。

在树种特征的选取上，已有研究主要聚焦于叶片和年轮两条路径。徐竞怡等[1]提出基于 GAN-DCNN 的树叶识别方法，使用对抗生成网络生成的样本，模拟不同季节和健康状况的叶片，使模型准确率提升到 99.8%；高欣等[2]则利用年轮径切面图像特征，实现了树种的快速初筛。然而，叶片形态易受季节、病虫害及环境胁迫影响，难以保证全年稳定获取；年轮虽具物种特异性，却需严格遵循年轮学采样规范，获取成本高，周期长，难以满足移动场景的即时需求。相比之下，树皮作为树木全年稳定可见的器官，其纹理、颜色与宏观形态在种间差异显著[3]，且受外界干扰较小。与叶片需综合色彩、轮廓、病斑等多维信息相比，树皮纹理以局部灰度变化为主，特征维度低、计算量小，既可通过手持设备近距拍摄，也能由无人机或监控摄像头远程采集。因此，树皮纹理已成为在移动端实现轻量化、高精度树种识别的理想特征来源。

传统的基于树皮的树种识别主要依赖人工特征提取和机器学习算法。例如：多化豫等[4]以均方差比为识别参数，在支持向量机和 BP 神经网络中取得较高的树种识别率；李晨光等[5]利用灰度共生矩阵提取特征并结合遗传规划分类算法，实现了 98.84% 的平均准确率。然而，均方差比对光照变化敏感，在户外多变光照场景中泛化性能不足；而灰度共生矩阵仅能反映纹理的宏观统计规律，对裂纹、孔洞等细粒度信息刻画不足。人工特征以经验为先导，难以兼顾纹理的多尺度结构，鲁棒性欠佳。

得益于无需人工定义图像纹理特征，并且识别准确率较高，深度学习近年来发展迅速[6]。刘嘉政[7]等改进 CNN 并采用 Maxout-ELU 激励函数，5 种树皮的识别准确率达到 97.70%，但该模型参数

量较大，无法满足移动端轻量化需求；肖恒玉等[8]提出的 FFTBarkNet 通过门控通道注意力模块（gated channel transformation，GCT）和快速傅里叶变换（faster Fourier transform，FFT），在减少参数量的同时保持精度，然而 FFT 频域变换会使树皮纹理高频细节丢失，难以区分形态相似的树种。YOLO 系列算法[9]因其高效的单阶段检测特性而被广泛引入到林业领域：陈万志等[10]通过多尺度级联注意力网络使模型精度达 98.3%；王婧怡等[11]在 YOLOv7 中引入 LSK 注意力机制和 GSConv，实现 96.39% 的平均精度；杨灿等[12]设计多尺度特征增强自适应网络 FastScaleNet 替换 YOLOv8 中的 C2f 结构，有效降低了复杂森林场景中树干的误检率和漏检率。上述的研究虽然识别精度高，但模型的参数量和计算复杂度仍然相对较大，面对一些移动设备或边缘设备（如野外林业监测相机、无人机边缘计算模组）时，受限于设备的存储空间、内存、功耗等问题而难以实现部署。马宏兴等[13]在 YOLOv5s 中嵌入 BoTNet 特征提取网络，结合深度可分离卷积和自注意力机制，在模型精度达到 95.7% 的同时，实现了在移动端上的部署；齐建东等[14]基于 MicroBlock 改进 YOLOv7 的主干网络，降低了 17.8% 的参数量且推理速度提升 14.3%，为在林业红外监测相机上实现轻量化检测模型部署提供了可行方案；韩誉等[15]将 Ghost 卷积引入 YOLOv10 主干网络中，模型在 Jetson Nano 4 GB 平台上帧速率（frames per second，FPS）达到 43.5，轻量化效果显著。然而这些研究仍存在涉及的树种覆盖度低、光照场景单一等不足。

面对上述挑战，本研究提出基于 YOLOv11 的树皮纹理识别改进模型 YOLOv11-SWER，并以自建的 70 类树种，共 6 681 张树皮纹理图像构建数据集开展研究。相较前代 YOLO，YOLOv11 在 YOLOv8 框架中引入更轻量的模块与高效学习机制，兼顾精度与速度。为了解决模型在硬件条件极端受限的设备上部署时存储空间上的限制问题，引入轻量化特征提取网络 StarNet 作为主干网络，与现有轻量化网络如 MobileNet 与 ShuffleNet 等主要依赖通道拆解与特征重排的方式不同，其核心 StarBlock 使用了星型连接结构，在提升特征的提取能力的同时，显著降低了参数量和计算量；针对树皮纹理的特殊性，采用多分支特征融合模块 RepNCSPELAN4，兼顾全局特征与局部特征并提升特征融合效率；对于树皮纹理高频细节丢失的问题，设计小波池化（WaveletPool）层，保留高频纹理细节以增强模型对树皮纹理微小特征的捕捉能力；最后，考虑到现有模型在边缘设备上内存占用过高的难点，优化检测头结构 Detect_Efficient，使用双分支分组卷积结构降低内存占用。之后，将 YOLOv11-SWER 与主流目标检测模型和轻量化网络进行对比，验证其在极端受限硬件上实现高精度、高效率树种识别的可行性。

1 研究数据

1.1 数据采集与预处理

本文研究所用的数据集涵盖了银杏（*Ginkgo biloba*）、白皮松（*Pinus bungeana*）、水杉（*Metasequoia glyptostroboides*）、雪松（*Cedrus deodara*）、玉兰（*Magnolia denudata*）等共计 70 种常见园林绿化树种的树皮图像数据。数据采集自北京市内多个代表性公园，包括国家植物园（南园、北园）、海淀公园、园林博物馆等地。在数据采集过程中，为确保数据的多样性和代表性，对同一树种在不同时间段（早晨 8：00—9：00，中午 12：00—13：00，傍晚 16：00—17：00）的多种光照条件下进行拍摄。此外，为保证数据一致性，规定所有图像均在距离树干 0.5 m 处以平视角度拍摄。

数据预处理包括以下步骤：（1）筛选并剔除样本中不符合拍摄要求、冗余重复的样本；（2）统一将所有图片尺寸调整为 3 468 × 4 624 像素，格式调整为 JPG；（3）使用 LabelImg 标注工具对清洗后的样本图像进行标注，并采用 YOLO 格式存储标注信息。经过预处理与标注，共得到 6 681 个有效树皮图像样本，按照 8∶1∶1 比例划分为训练集、验证集、测试集。表 1 展示了各树种的图片数量分布。

1.2 数据增强

为增强模型在不同光照、角度及环境条件下对树皮特征的适应能力，本研究采用数据增强技术对现有数据集进行扩充。具体方案为：基于 Albumentations 库实施包括亮度和对比度调整、随机裁剪、水平/垂直翻转等增强操作，以模拟真实场景中的纹理变化[16]。具体增强策略包括以下 4 点。

（1）随机亮度和饱和度变换（random brightness contrast）：通过随机调整亮度和对比度（幅度均为 ± 20%），增强模型对光照变化的适应能力。

（2）几何变换：利用 D4 变换（D4 group transformation）、水平/垂直变换（flip）、转置（transpose）模拟多角度观测树皮纹理，提升模型在不同视角下的稳定性。

（3）边框安全裁剪（bounding box safe random crop）：采用侵蚀率为 0.1 的裁切机制，在确保目标框完整的前提下，生成局部特征样本，增强模型对局部纹理信息的提取能力。

（4）概率控制：所有增强操作均以 5% 的概率随

表 1　数据集中的树种及其图片张数

Tab. 1　Tree species and number of images in dataset

序号	中文名称	学名	图片张数			序号	中文名称	学名	图片张数		
			早	中	晚				早	中	晚
1	梧桐	*Firmiana simplex*	25	30	37	36	洋白蜡	*Fraxinus pennsylvanica*	50	10	40
2	梓	*Catalpa ovata*	28	50	22	37	英桐	*Platanus acerifolia*	50	20	30
3	榔榆	*Ulmus parvifolia*	30	14	55	38	山楂	*Crataegus pinnatifida*	65	15	20
4	杂交鹅掌楸	*Liriodendron* × *sinoamericanum*	20	50	30	39	红花刺槐	*Robinia pseudoacacia*	40	10	50
5	元宝枫	*Acer truncatum*	35	42	23	40	脱皮榆	*Ulmus lamellosa*	10	30	30
6	圆柏	*Juniperus chinensis*	30	25	45	41	华东椴	*Tilia japonica*	40	20	42
7	玉兰	*Magnolia denudata*	25	55	20	42	欧洲白榆	*Ulmus laevis*	16	10	32
8	油松	*Pinus tabuliformis*	21	40	39	43	北美圆柏	*Sabina virginiana*	30	20	50
9	银杏	*Ginkgo biloba*	35	25	40	44	晚花杨	*Populus* × *canadensis* ‘serotina’	27	15	60
10	雪松	*Cedrus deodara*	40	19	40	45	黑松	*Pinus thunbergii*	25	60	15
11	水杉	*Metasequoia glyptostroboides*	20	60	20	46	山皂荚	*Gleditsia japonica*	30	40	30
12	柿	*Diospyros kaki*	40	25	35	47	旱柳	*Salix matsudana*	20	55	25
13	锐齿槲栎	*Quercus aliena* var. *acuteserrata*	45	15	40	48	加杨	*Populus* × *canadensis*	40	22	47
14	青檀	*Pteroceltis tatarinowii*	20	50	30	49	白皮松	*Pinus bungeana*	30	30	40
15	美桐	*Platanus occidentalis*	15	60	25	50	北京丁香	*Syringa pekinsis*	12	20	65
16	毛白杨	*Populus tomentosa*	30	50	20	51	榉树	*Zelkova serrata*	30	33	18
17	栾树	*Koelreuteria paniculata*	60	20	20	52	蒙古栎	*Quercus mongolica*	40	20	40
18	连香树	*Cercidiphyllum japonicum*	35	15	50	53	杏	*Armeniaca vulgaris*	10	20	25
19	黄连木	*Pistacia chinensis*	40	40	20	54	栗	*Castanea mollissima*	50	10	40
20	华山松	*Pinus armandii*	16	40	42	55	黄檗	*Phellodendron amurense*	20	11	10
21	胡桃	*Juglans regia*	30	20	49	56	盐麸木	*Rhus chinensis*	20	10	22
22	黑弹树	*Celtis bungeana*	20	30	50	57	黄栌	*Cotinus coggygria*	15	10	20
23	构树	*Broussonetia papyrifera*	15	25	60	58	白杜	*Euonymus bungeanus*	55	25	20
24	杜仲	*Eucommia ulmoides*	30	40	30	59	水曲柳	*Fraxinus mandschurica*	20	20	60
25	大果榆	*Ulmus macrocarpa*	10	20	70	60	君迁子	*Diospyros lotus*	35	25	40
26	刺楸	*Kalopanax septemlobus*	40	10	50	61	花曲柳	*Fraxinus rhynchophylla*	40	20	40
27	臭椿	*Ailanthus altissima*	60	15	25	62	白花泡桐	*Paulownia fortunei*	35	20	45
28	侧柏	*Platycladus orientalis*	20	15	65	63	南京椴	*Tilia miqueliana*	15	60	25
29	榆树	*Ulmus pumila*	10	25	65	64	金叶丝棉木	*Euonymus maackii* ‘Aureus’	30	20	50
30	黑榆	*Ulmus davidiana*	20	50	30	65	漆	*Toxicodendron vernicifluum*	35	20	45
31	流苏树	*Chionanthus retusus*	26	20	55	66	细裂槭	*Acer stenolobum*	20	56	24
32	乔松	*Pinus griffithii*	13	57	32	67	橙桑	*Maclura pomifera*	35	15	50
33	杉松	*Abies holophylla*	65	20	15	68	桑	*Morus alba*	35	35	30
34	白杆	*Picea meyeri*	30	50	20	69	健杨	*Populus* × *canadensis* ‘Robusta’	40	10	50
35	七叶树	*Aesculus chinensis*	30	30	40	70	枣	*Ziziphus jujuba*	40	19	20

机触发，以在数据多样性和样本真实性之间保持平衡，防止过度增强导致的样本失真。

最后，将增强后的样本统一缩放到 640 × 640 像素，输入目标检测网络。

2　基于树皮纹理的轻量化 YOLOv11 树种识别方法

2.1　YOLOv11-SWER 模型框架

YOLOv11-SWER 模型以 YOLOv11[17]作为基准模型，针对树皮纹理识别任务进行网络改造与优化。原始 YOLOv11 模型由主干(Backbone)、颈部(Neck)、头部(Head)网络组成。图像先经由 Backbone 层下采样提取特征，再通过 Neck 上采样融合特征，最后将特征图送入 Head 中进行预测。YOLOv11-SWER 模型在 YOLOv11 基础上，引入轻量化特征提取网络 StarNet、多分支特征融合模块 RepNCSPELAN4、小波池化 WaveletPool、高效检测头 Detect_Efficient 等关键模块。其中，StarNet 作为主干网络，结合深度

可分离卷积与通道混洗机制，显著降低网络的参数量和特征提取过程中的计算量。RepNCSPELAN4 模块结合分组卷积与参数共享策略，兼顾树皮纹理图像的全局特征与局部特征，提升多尺度特征融合效率。WaveletPool 减少噪声干扰，并保留高频纹理细节，增强模型对树皮纹理微小特征的捕捉能力。Detect_Efficient 检测头使用双分支分组卷积结构，提高计算效率。通过上述模块的引入，能够在保证模型高检测精度的同时，减少计算量，提高轻量化程度。改进模型的结构示意图如图 1。

2.2 轻量化特征提取网络 StarNet

StarNet[18]是一种轻量化的特征提取网络，其主干网络可以分为 4 个阶段，每个阶段均由卷积层（convolutional layer，Conv）和 Star Blocks 模块组合而成。其中的 Star Blocks 模块利用的深度可分离卷积（depthwise separable convolution，DWConv）和通道混洗（channel shuffle）对输入特征图进行下采样，这两者能大幅降低模型的参数量和计算量。通过阶段的逐层堆叠，模型也能够提取出较深层次的图像特征。并且 Star Blocks 中使用的轻量级注意力模块可以将深层与浅层的特征进行动态融合，使得模型的表征能力和鲁棒性得到提升。这些设计使得 StarNet 利用低参数量就能够高效获取图像的深层信息，兼顾模型的精度要求和低计算量要求[19]。

本研究采用 StarNet 系列中的轻量化模型 StarNet_s050 作为基础框架，该主干网络的 4 个阶段分别包含 1、1、3、1 个 Star_Block。输入图像首先通过 Stem Layer 层进行初步的特征提取，其中的批归一化层和激活函数层可以有效提高训练速度和模型性能。随后，特征图依次通过 Stage 1 至 Stage 4 共 4 个阶段，进行下采样和特征提取。DWConv 的使用减少了 StarNet_s050 的计算复杂度，而 Star_Block 的堆叠保证了模型对特征的高效提取能力。该主干网络大幅增加了高精度模型轻量化的可行性。本研究采用的 StarNet_s050 架构如图 2 所示。

2.3 多分支特征融合模块 RepNCSPELAN4

RepNCSPELAN4 融合了 CSP（crossstagepartial）[20]和 ELAN（efficient layer aggregation network）[21]的核心设计思想，通过多分支的高效特征融合策略提升特征提取质量，同时实现网络的结构压缩与计算量降低。在 RepNCSPELAN4 中，多分支学习架构由 RepNCSP 模块和卷积层的组合实现，使模型既能够学习全局特征，也能关注到局部特征，保证模型的高识别精度。在结构实现上，RepNCSPELAN4 模块首先接收来自前一层的输入特征图，并通过一个 1*1 卷积进行初步通道压缩，以适配后续分支操作所需的维度配比与计算资源控制。随后，压缩后的特征图被划分为两个通道数相等的子特征图，分别命名为 $\boldsymbol{y}_1$ 与 $\boldsymbol{y}_2$，其中 $\boldsymbol{y}_1$ 直接保留用于后续与深层特征拼接，承担浅层细节信息的传递功能，而 $\boldsymbol{y}_2$ 则作为深度纹理特征提取的主干路径。该路径首先通过一个 RepNCSP 模块与 3*3 标准卷积的组合进行处理，并在得到中间特征后保留一份，再将其继续输入至另一个相同结构的 RepNCSP 模块中，以逐层堆叠方式提取更深层次的纹理特征。每个 RepNCSP 模块内部均采用基于 CSP 思想的结构设计，通过在分支间

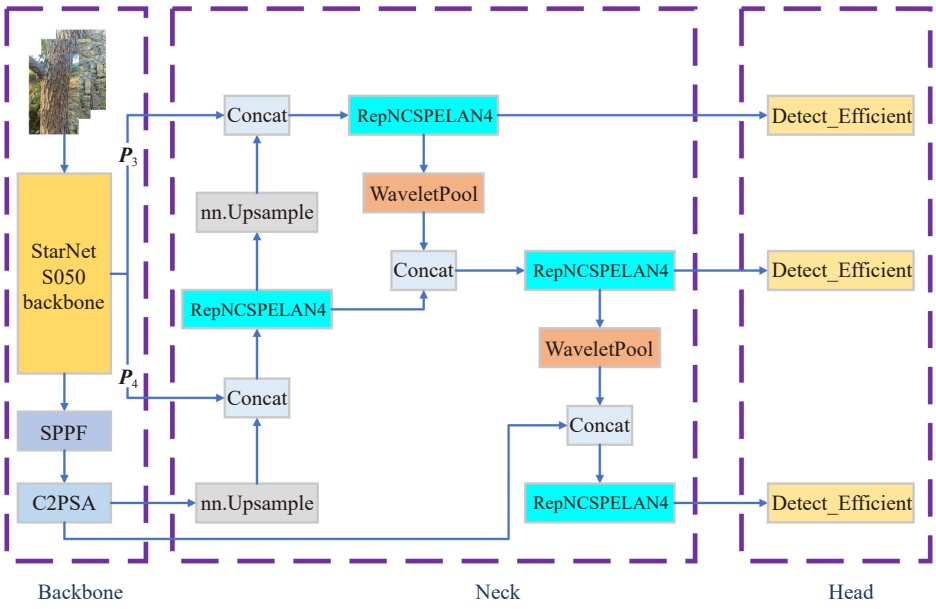


图 1 YOLOV11-SWER 树皮纹理识别模型结构示意图

Fig. 1 Schematic diagram of YOLOV11-SWER bark texture recognition model structure

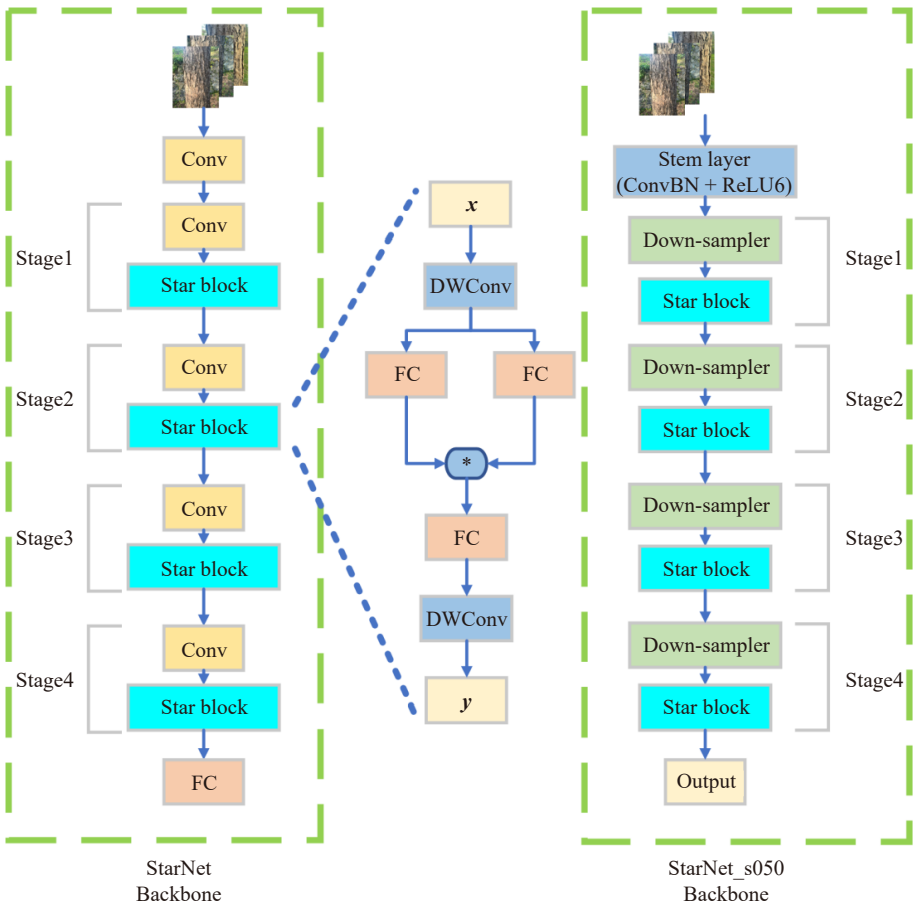


图 2　轻量化特征提取网络 StarNet 及 StarNet_s050 结构示意图

Fig. 2　Schematic diagram of structure of lightweight feature extraction network StarNet and StarNet_s050

引入分组卷积和权重共享策略，有效降低通道维度所带来的计算压力，同时保留关键的非线性表达能力。值得注意的是，经过这两个模块的深层分支在通道与空间尺度上均保持一致，从而使得与浅层特征 $\boldsymbol{y}_1$ 的拼接在语义与结构上更加兼容。最终，3 路输出特征在通道维度上进行拼接，并通过一个 1*1 卷积再次调整通道数，使融合后的输出特征在维度与分布上适配下游网络需求，实现深浅纹理特征的协同增强与统一表达。

在 RepNCSP 中的 RepNBottleneck 模块中，特征图的通道维度将进行压缩–扩展变换，显著增强了树皮特征的表征能力。具体来说，对应输入到 RepNBottleneck 的大小为 [N, C_1, H, W] 的特征图，首先经过 RepConvN 模块，压缩通道数变为 [N, C, H, W]，其中 $C = e*C_2$，e 为压缩比，C_2 为目标通道数；再通过 3*3 标准卷积层进行树皮特征提取的同时将通道数恢复到目标通道数 C_2，判断是否满足残差连接条件（$C_1 = C_2$），若满足条件则拼接残差，否则不拼接，最后通过 SiLU 激活函数输出特征图，大小为 [B, C_2, H, W]。

RepNCSPELAN4 多分支学习架构及其相互之间的独立性，使得模型在保证高识别精度的情况下降低计算复杂度，大大提高将其部署在移动设备上的可能性，其结构如图 3 所示。

2.4　小波池化 WaveletPool

Wavelet Pooling 是一种基于小波变换的池化方法，不同于传统的最大池化和平均池化，它能够更有效地利用图像的高低频信息减少特征图的维度，同时增强模型对细微特征的注意力，保留更多的关键特征信息[22]。

本研究在 WaveletPool 模块中定义了 4 种 Haar 小波滤波器核：低通滤波（LL）通过增强平滑区域特征进而保留树皮纹理的整体近似特征，水平高通滤波（LH）通过增强垂直边缘响应进而捕获垂直方向的树皮裂纹细节，垂直高通滤波（HL）通过增强水平边缘响应进而提取水平方向的树皮沟壑特征，对角高通滤波（HH）通过角点特征提取进而检测树皮表面的斑点与对角纹理，对应的 4 个小波滤波器核在 WaveletPool 结构图中体现。这些小波滤波器通过复制堆叠 C 次（C 为输入张量的通道数）形成的形状为（4*C, 1, 2, 2）的滤波器组 filters 将作为卷积核组成后续的卷积操作，这确保模型在池化过程中能够同时兼顾全局特征与局部特征，减少关键特征信息的损耗。此外，这些滤波器在初始化后便保持固定，避

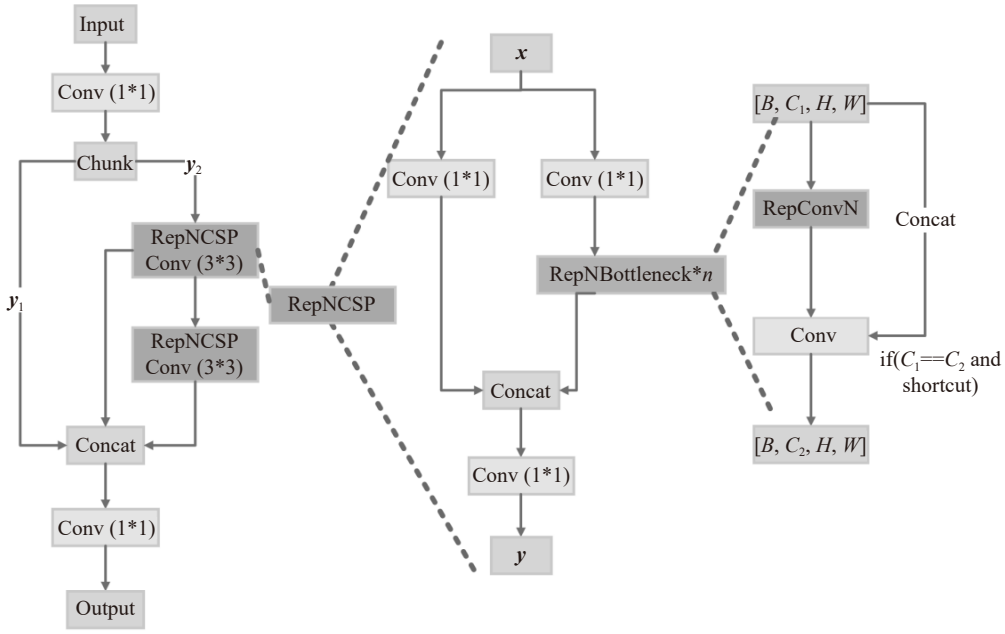


图 3 RepNCSPELAN4 多分支特征融合模块结构示意图

Fig. 3 Schematic diagram of RepNCSPEAN4 multi branch feature fusion module structure

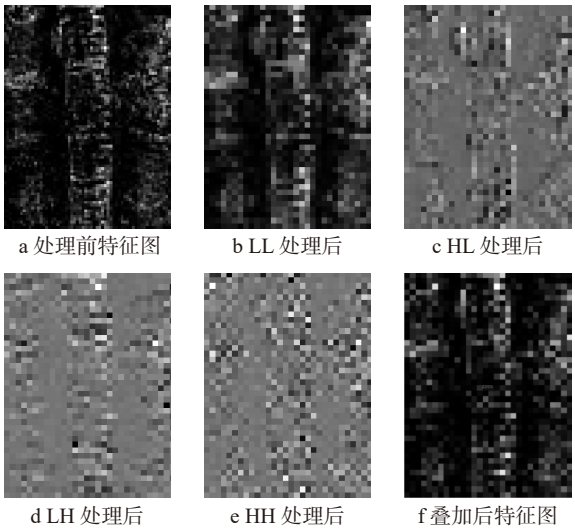
免了额外的梯度计算和参数更新，从而降低计算成本。

在前向传播过程中，输入特征图 $\boldsymbol{x}(B, C, H, W)$ 将通过分组卷积操作，使得每个输入通道独立地与对应的小波滤波器进行卷积，之后通过步幅为 2 的下采样模块进一步实现高效降维，输出特征图 $\boldsymbol{y}$ 的形状为$(B, 4*C, H/2, W/2)$。Wavelet Pooling 通过高通滤波和低通滤波捕捉图像的高低频信息，使模型能够同时关注图像中的全局结构和局部细节，从而强化高低频纹理细节的提取能力，提高模型对树皮细微特征的识别精度。此外，小波变换在增强多尺度特征提取能力的同时，有效降低了参数量[23]。特征图经过 LL、LH、HL、HH 这 4 个滤波器的前后变化如图 4 所示，WaveletPool 结构图如图 5 所示。

2.5 高效检测头 Detect_Efficient

Detect 检测头作为 YOLOv11 网络的最后一层，主要用于接收 neck 端输出的 $\boldsymbol{p}_3$、$\boldsymbol{p}_4$、$\boldsymbol{p}_5$ 特征图，并利用 Backbone 端与 Neck 端提取的特征进行边界框回归和类别概率预测。传统的 Detect 检测头采用深度可分离卷积以减少参数量和计算量，而本研究提出的 Detect_Efficient 可以进一步优化模型的轻量化效果，使其计算开销更低，同时提升检测效率。

在 Detect_Efficient 结构中，Neck 端输出的 $\boldsymbol{p}_3$、$\boldsymbol{p}_4$、$\boldsymbol{p}_5$ 特征图依次经过 stem 模块进行进一步的特征提取。随后，特征图进入两支相互独立的 1*1 卷积分支，一条分支负责预测界框的偏移量，另一条分支负责预测类别概率。最终，拼接两个分支输出的特征图以完成目标检测任务。其中，stem 模块由两层分组卷积（Group Conv）组成，相较于传统的 DWConv



a 处理前特征图 b LL 处理后 c HL 处理后

d LH 处理后 e HH 处理后 f 叠加后特征图

LL 为低通滤波，LH 为水平高通滤波，HL 为垂直高通滤波，HH 为对角高通滤波。

图 4 特征图小波处理前后示意图

Fig. 4 Schematic diagram of feature map before and after wavelet processing

和普通卷积（Conv），Group Conv 通过将特征图的通道划分为多个独立组，并在组内进行卷积计算，使得计算量和参数量降低，计算效率得以提高。Detect_Efficient 检测头简化了 Detect 检测头的结构，去掉了端到端模式及额外的分支预测，进一步降低了计算复杂度。Detect_Efficient 的结构图如图 6 所示。

3 树种识别方法的实验验证

3.1 实验环境

本实验采用 Ubuntu 20.04 操作系统，CPU 为 Intel（R）Xeon（R）Platinum 8481C CPU，GPU 为 RTX

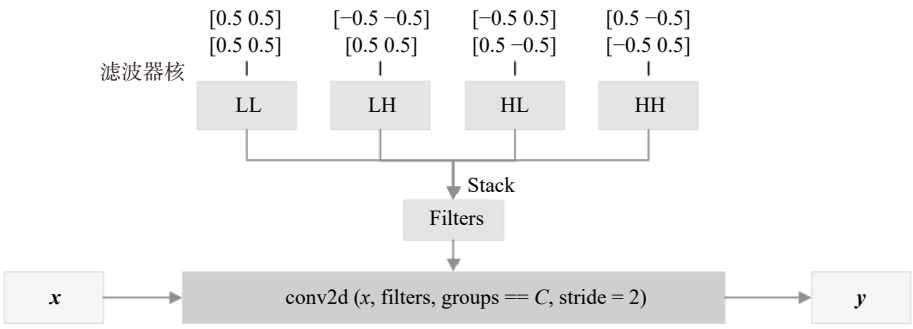


图 5　WaveletPool 模块结构示意图

Fig. 5　Schematic diagram of WaveletPool module structure

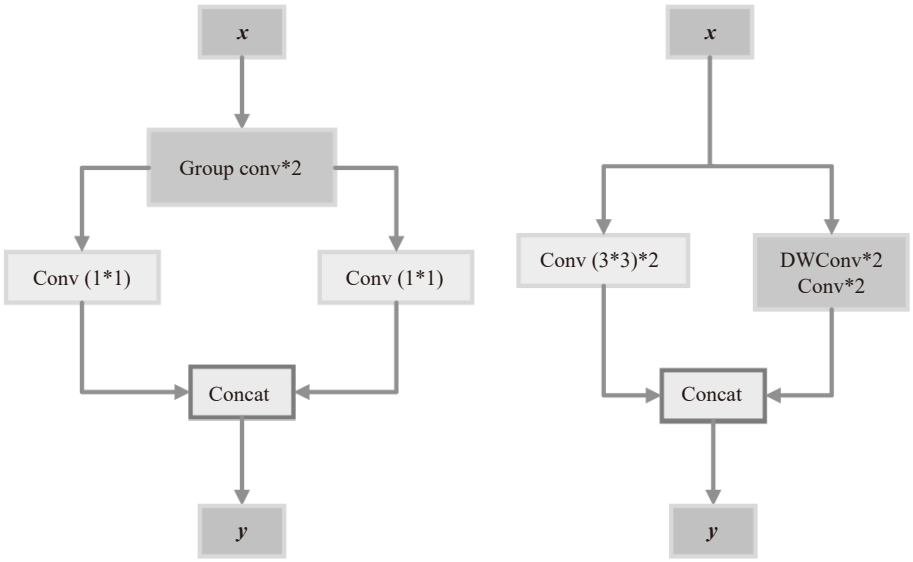


图 6　高效小目标检测头 Detect_Efficient 结构示意图

Fig. 6　Schematic diagram of structure of efficient small target detection head Detect_Efficient

4090D，显存 24 GB，内存 80 GB 的云服务器环境，采用 Python3.8 编程语言，CUDA 版本为 11.8, 深度学习框架为 Pytorch2.0.0。

实验训练参数设置如下：输入图像统一调整为 640 × 640 像素的 RGB 彩色图像，训练轮数（epoch）设为 300，批量尺寸（batch size）设为 8。优化器使用随机梯度下降（SGD），初始学习率为 0.01，动量系数为 0.937，权重衰减系数为 0.000 5。

3.2　验证实验

为验证各模块对模型性能的贡献程度及其组合对模型的优化效果，本文对上述模块展开消融实验。并通过对比试验将优化后的模型与主流轻量化模型进行对比，衡量其检测精度和轻量化程度。为了清晰展示模型对不同树种的识别能力差异、不同模型的识别能力差异以及模型对特征的专注区域差异，后期进行可视化分析。最后，为检验模型对光照条件的鲁棒性，进行不同光照场景下的模型性能分析实验。

3.2.1　模块有效性分析

本消融实验设计主要围绕单模块替换和整体融合两种方法展开。

（1）单模块替换实验：本研究在原始 YOLOv11 模型网络上分别将主干网络、特征提取与融合模块、下采样模块以及检测头替换为相应的改进模块，明确各改进模块在改进中的具体作用，同时验证其对检测性能的波动影响，为后续整体方案的设计提供充分依据。

（2）整体融合实验设计：在单模块替换实验验证各模块改进效果的基础上，进一步将单个改进模块逐步融合到原始模型中，通过联合优化各个模块，实现模型结构的整体简化和计算资源的最大化节约。

3.2.2　不同模型性能对比

为衡量改进后模型检测精度和轻量化程度，选取单阶段和双阶段两大类典型目标检测算法进行不同目标检测模型的对比实验，以测试本文提出的 YOLOv11-SWER 模型在树皮目标检测中的表现。此外，选取 2 种轻量化网络结构，分别替换 YOLOv11 的主干网络，以测试 YOLOv11-SWER 模型的轻量化程度。选取的单阶段检测算法包括 YOLOv5、YOLOv8、YOLOv9 和 YOLOv10 和 SSD，选择 YOLO

系列算法用于研究 YOLO 算法从早期到最新的目标检测算法优化趋势，选择 SSD 用于研究 YOLOv11-SWER 模型在多尺度目标检测方面的性能。双阶段检测算法则选择 Faster R-CNN[24]，它在检测精度方面具有较高的代表性[25]。通过这种多模型对比，可以全面展示 YOLOv11-SWER 在树皮目标检测任务中的适用性和领先优势。在统一训练与测试条件下，通过对各模型在参数量、计算复杂度和检测精度等指标的横向比较，系统性地验证 YOLOv11-SWER 在树皮目标检测任务中的优越性。选取的轻量化网络结构包括 MobileNetv2[26]和 ShuffleNetv2[27]。MobileNetv2 采用倒置残差块结构，并使用深度可分离卷积，能够在低维空间进行轻量计算后，再通过升维保留更多有效信息，拥有相对较好的精度表现；ShuffleNetv2 采用通道混洗的机制，可以解决传统分组卷积操作中信息在通道间不流通的问题，结构简单高效，更适合追求检测速度的场景。通过与这两种轻量化网络结构的对比，可以测试以 StarNet_s050 作为主干网络的 YOLOv11-SWER 模型轻量化改进对模型性能的影响。

3.2.3 可视化分析

模型对不同树种的识别能力差异可以通过绘制模型对 70 种树种的识别结果指标的雷达图来清晰表示。为可视化不同模型的识别能力差异，可以选取具有代表性的树皮图片，绘制出不同模型对其检测出的检测框图，通过对比检测框与真实树皮位置来比较不同模型的识别能力。为跟踪模型优化前后对特征的专注区域差异，可通过绘制热力图来突出模型关注程度较大的区域，通过对比区域大小的变化来判断模型识别能力是否有所提高。

3.2.4 不同光照场景下的模型性能分析

自然场景中的图像采集常常伴随不同自然干扰带来的变化，这些变化可能影响目标检测模型的泛化能力和性能稳定性。Premakumara 等[28]通过合成数据增强的方法模拟自然干扰，研究模型在不同干扰下的鲁棒性。在基于树皮纹理的树种识别中，不同光照场景是重要干扰因素，因此，本研究基于原始测试集图像，采用基于 Albumentations 库的增强方法，保持图像内容与目标标注完全一致，仅通过亮度、色温、对比度和阴影等调整，模拟设计早晨、中午、傍晚 3 个不同时间段的光照场景，通过对比不同光照场景下模型的精度来衡量模型的鲁棒性。

3.3 评价指标

为了全面评估 YOLOv11-SWER 模型的检测性能，本研究从检测精度和计算效率两个方面进行评估。在精度方面，使用精确率、召回率、平均精度、平均精度均值、精确率和召回率的调和平均数（F_1）作为评价指标；在计算效率方面，使用模型参数量（parameters）、浮点运算数（FLOPs）和帧速率（FPS）评估模型计算量和对硬件的算力要求。

精确率（precision，P）是指在所有被模型预测为正类的样本中，实际也为正类的样本所占的比例。精确率反映的是模型的误检情况[29]，计算公式为

$$P = \frac{T_P}{T_P + F_P} \times 100\% \tag{1}$$

式中：T_P 为模型正确预测为正类的样本数，F_P 为模型错误预测为正类的样本数。

召回率（recall，R）是指所有实际为正类的样本中，被模型正确预测为正类的样本所占的比例。召回率反映的是模型的漏检情况，计算公式为

$$R = \frac{T_P}{T_P + F_N} \times 100\% \tag{2}$$

式中：F_N 为模型错误预测为负类的样本数。

平均精度（P_A）是在以召回率为横轴，精确率为纵轴，绘制的精确率−召回率（precision-recall, PR）曲线下的面积，即对 PR 曲线的积分。计算公式为

$$P_A = \int_0^1 P(R)\,\mathrm{d}R \tag{3}$$

平均精度均值（mean average precision，mAP）是指所有预测类别计算得到的平均精度的均值[30]。mAP50 指当预测框与真实框的交集面积和并集面积之比，即交并比（intersection over union，IoU）阈值为 0.5 时的平均精度；mAP50-95 指在 IoU 阈值从 0.5 到 0.95 之间的平均精度。计算公式为

$$\mathrm{mAP} = \frac{1}{M}\sum_{k=1}^{M} P_{Ak} \tag{4}$$

式中：P_{Ak} 代表第 k 个树种的平均精度。当计算 mAP50 时，平均精度为在 IoU = 0.5 时的平均精度；当计算 mAP50-95 时，平均精度为 IoU 阈值从 0.5 到 0.95 之间下的平均精度。

F_1 分数即为精确率和召回率的调和平均数，用于平衡精确率和召回率，反映了模型的综合性能。F_1 分数的取值范围为 0 ~ 1，越接近 1，模型的综合性能越好。计算公式为

$$F_1 = \frac{2PR}{P+R} \tag{5}$$

4 结果与分析

4.1 模块有效性

分析可知，各改进模块均能降低模型参数量与计算复杂度，并在大多数情况下保持或提升检测精度（表 2）。

表 2　消融实验结果

Tab. 2　Results of ablation experiments

YOLOv11	WaveletPool	Detect_Efficient	RepNCSPELAN4	Starnet_s050	参数量	FLOPs	mAP50	mAP50-95
√					2.60×10^6	6.4×10^9	0.992	0.750
√	√				2.19×10^6	5.5×10^9	0.993	0.751
√		√			2.34×10^6	5.2×10^9	0.990	0.744
√			√		2.14×10^6	5.0×10^9	0.992	0.748
√				√	1.96×10^6	5.1×10^9	0.993	0.753
√	√		√		1.67×10^6	4.2×10^9	0.991	0.740
√	√	√	√		1.41×10^6	3.0×10^9	0.993	0.751
√	√		√	√	1.64×10^6	4.5×10^9	0.993	0.756
√	√	√	√	√	1.38×10^6	3.3×10^9	0.993	0.750

注: FLOPs表示模型的浮点运算数, 用于衡量模型计算量的大小; mAP50表示IoU 阈值为 0.5 时, 计算各类别PA后的平均值; mAP50-95 表示在 IoU 阈值从 0.5 到 0.95(步长 0.05) 之间共 10 个阈值下, 计算每个阈值下的 mAP的均值。mAP50与mAP50-95均用于衡量模型精度。

用 WaveletPool 模块替换原始 YOLOv11 下采样模块后, 模型参数量减少了 15.8%(由 2.60×10^6 降低至 2.19×10^6), 计算量减少了 14.06%(由 6.4×10^9 次浮点运算降低至 5.5×10^9 次浮点运算), 而 mAP50 与 mAP50-95 均上升了 0.001, 其原因是 WaveletPool 模块利用小波变换进行多尺度特征提取, 减少了冗余计算, 同时提高了模型对不同分辨率目标的适应能力。用 RepNCSPELAN4 替换 C3k2 模块后, 模型参数量减少了 17.7%, 计算量减少了 3.85%。这是因为 RepNCSPELAN4 采取多分支学习与特征融合策略, 提高特征提取能力的同时减少了计算冗余。用 StarNet_s050 替换原来 YOLOv11 骨干网络 backbone 后, 由于 StarNet 使用了深度可分离卷积和通道混洗, 优化了计算效率, 使参数量减少了 24.6%, 计算量减少了 2.00%(浮点运算由 6.4×10^9 次降低至 5.1×10^9 次), 模型更加轻量化, 同时使 mAP50-95 增加了 0.003。用 Detect_Efficient 检测头替换 Detect 检测头后, 模型参数量减少了 10.0%, 计算量减少了 17.65%(浮点运算由 6.4×10^9 次降低至 5.2×10^9 次), mAP50 与 mAP50-95 分别降低了 0.002 与 0.006。这是因为 Detect_Efficient 检测头使用分组卷积代替原来的深度可分离卷积和普通卷积, 独立组别下的卷积操作进一步减少了大量不必要的计算量, 但由于组间信息独立处理, 略微影响了模型检测精度。最后, 将所有改进的模块融合到 YOLOv11 后, 尽管相较于原始模型, YOLO v11-SWER 模型的 mAP50 仅上升了 0.001, mAP50-95 维持在 0.75。但其参数量仅为原来的模型的 52.9%, 计算量仅为原来的模型的 51.5%, 轻量化效果显著, 在高检测精度和高效率之间取得了更好的平衡。综上所述, 本文提出的改进方法在保持高检测精度的同时, 有效降低了模型计算复杂度, 实现了模型轻量化, 提高了模型在硬件资源极端受限的部署场景中的适用性。

4.2　不同模型性能对比

在轻量化方面, YOLOv11-SWER 模型参数量仅为 1.38×10^6, 相较于原始 YOLOv11 模型, 参数量减少了 46.92%, 而相较于对比实验中最轻量化的 YOLOv5n, 参数量减少了 25.4%。在检测精度方面, 尽管模型规模显著缩小, 改进模型仍保持高检测精度: mAP50 达 0.993, 与最高检测精度 0.994 基本持平, 其 mAP50-95 为对照组中的最大值。可见, YOLOv11-SWER 在轻量化与检测精度之间达到了良好平衡。与现有目标检测模型相比, YOLOv11-SWER 在大幅降低参数量的同时, 仍能维持高检测精度, 具备更强的部署适应性和计算效率(表 3)。

通过不同轻量网络的模型对比实验, 进一步验证 YOLOv11-SWER 模型在轻量化与检测精度之间的综合优势。YOLOv11-SWER 相较于 MobileNetv2 与 ShuffleNetv2, 在模型规模方面, 模型体积、参数量和 FLOPs 这 3 项指标上均为最小; 相较于 MobileNetv2, 缩小了 77.9% 的模型体积与 60.1% 的参数量, FLOPs 减少了 62.9%。而在推理速度方面, YOLOv11-SWER(86 帧/s)与 StarNet(88 帧/s)略低于采取了更高效的通道拆解与重排的策略的 ShuffleNetv2(91 帧/s)。从检测精度方面, MobileNetv2 在 mAP50 的指标上取得了最大值 0.994, 但在更具判别能力的 mAP50-95 指标上, StarNet 以 0.753 领先, 而 ShuffleNetv2 在精度对比上相对欠缺。综上所述, MobileNetv2 虽然在精度指标 mAP50 上表现最优, 但模型体积和推理速度限制了其应用范围。ShuffleNetv2 在推理速度方面表现突出, 但精度略有欠缺。YOLOv11-SWER 在各项性能中实现了较优

表 3 不同目标检测模型对比实验结果

Tab. 3 Comparative experimental results of different object detection models

模型	参数量	FLOPs	精确率(P)/%	召回率(R)/%	F_1分数	mAP50	mAP50-95
YOLOv5n	1.85×10^6	4.4×10^9	90.9	91.3	0.911	0.963	0.693
YOLOv8n	3.01×10^6	8.1×10^9	97.5	98.5	0.980	0.994	0.743
YOLOv9t	3.67×10^6	16.1×10^9	90.4	90.6	0.905	0.967	0.703
YOLOv10n	2.73×10^6	8.4×10^9	97.2	97.0	0.971	0.994	0.743
YOLOv11	2.60×10^6	6.4×10^9	97.7	98.4	0.980	0.992	0.750
FasterRCNN	28.98×10^6	250.4×10^9	84.36	92.31	0.878	0.911	0.501
SSD	32.97×10^6	33.86×10^9	91.56	91.76	0.915	0.918	0.550
YOLOv11-SWER	1.38×10^6	3.3×10^9	98.1	98.4	0.982	0.993	0.750

的均衡表现，凸显了其良好的部署潜力(表 4)。尽管本研究未涉及实际部署测试，但上述指标可为条件受限的场景提供理论参考与结构设计依据。

4.3 可视化分析

4.3.1 雷达图

YOLOv11-SWER 在树皮识别任务中表现出色：所有树种的检测指标均保持较高水平，其中 65 个树种(占比 92.8%)的 F_1 分数超过 0.9(图 7)，充分满足实际应用的精度需求。但模型对乔松(*Pinus griffithii*，树种编号 32)的识别效果较差，其精确度和召回率分别为 87.0% 和 72.7%，均为所有树种中的最低值，其次是黑松(*Pinus thunbergii*，树种编号 45)，其精确率为 82.7%，略低于其他树种。实验的混淆矩阵分析表明，这两者间存在较高混淆度，其主要原因可能是乔松与黑松同为松属，树皮纹理的形态特征相似：黑松的老树树皮呈灰褐色至黑褐色，深纵裂成龟甲状厚块[31]；乔松的老树树皮呈暗灰褐色，裂成不规则鳞片状[32]。两者的树皮颜色相似，裂纹差异因本研究数据集拍摄规定显现不明显。尽管个别树种的检测效果存在一定局限性，但整体检测性能仍能达到实际应用标准。

4.3.2 检测框

为验证不同模型的检测性能差异，选取梧桐(*Firmiana simplex*)中的典型样本进行对比测试，并绘制检测框可视化结果(图 8)。除检测框外，还一并输出了类别标签与置信度。结果表明，所有对比模型均能有效完成树皮的检测与分类任务，这一结果与消融实验和对比实验的定量分析结果一致。然而，在检测框的定位精度方面，不同模型之间存在显著差异：YOLO v11-SWER 在所有轻量化模型中表现最佳，仅次于极高参数量和计算量的 FasterCNN 和 SSD 网络；此外，YOLOv11-SWER 的检测框类别置信度达到 0.89，相比基准模型(YOLOv11)有显著提升；其生成的检测框能够更完整地覆盖树皮区域，减少背景干扰。这表明 YOLOv11-SWER 经过轻量化设计后仍能保持优异的特征提取和目标定位能力。

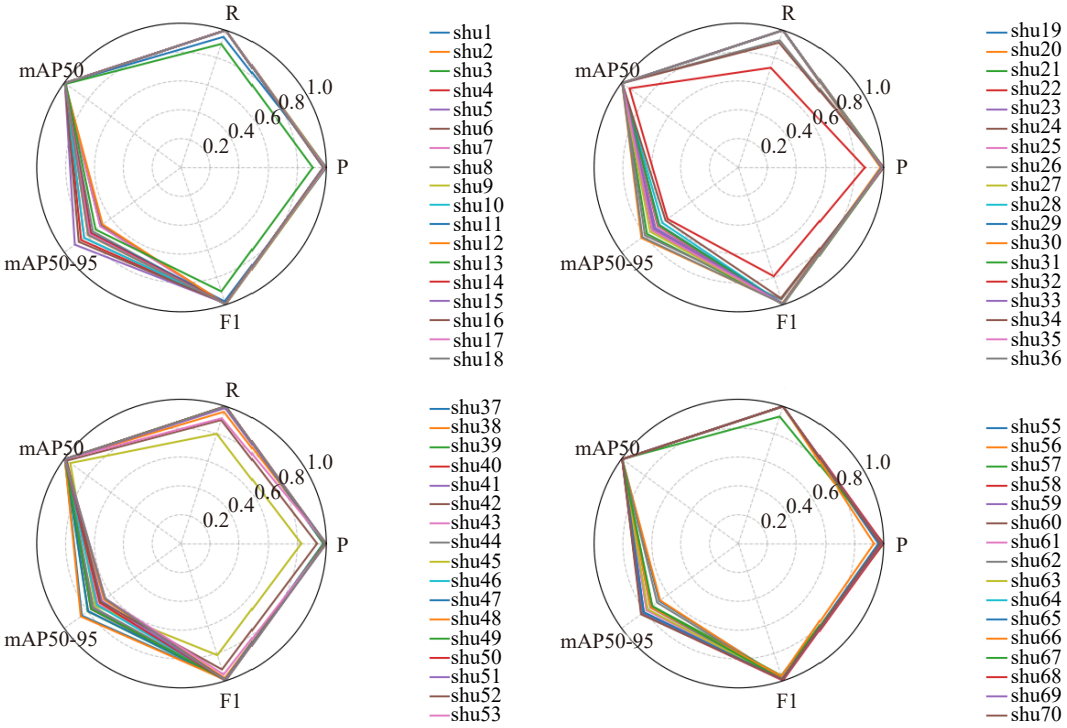
4.3.3 热力图

为进一步探究 YOLOv11-SWER 的特征提取能力，本文基于 Grad-CAM++ 方法[33]，选取数据集中的 3 种不同光照场景下的梧桐拍摄样本进行了特征关注度分析，并生成热力对比图(图 9)。可视化分析结果表明，改进后的模型表现出以下优势：(1)检测框内高响应区域(红色)显著增加，改进模型能更精准地聚焦树皮纹理特征，相比基准模型，避免了对无关背景的误关注；(2)检测框与真实标注框的空间重合度明显提高，说明模型不仅能准确识别目标，还能更精确地定位目标边界，减少误检和漏检现象；

表 4 使用不同轻量网络的模型对比实验结果

Tab. 4 Results of model comparison experiments using different lightweight networks

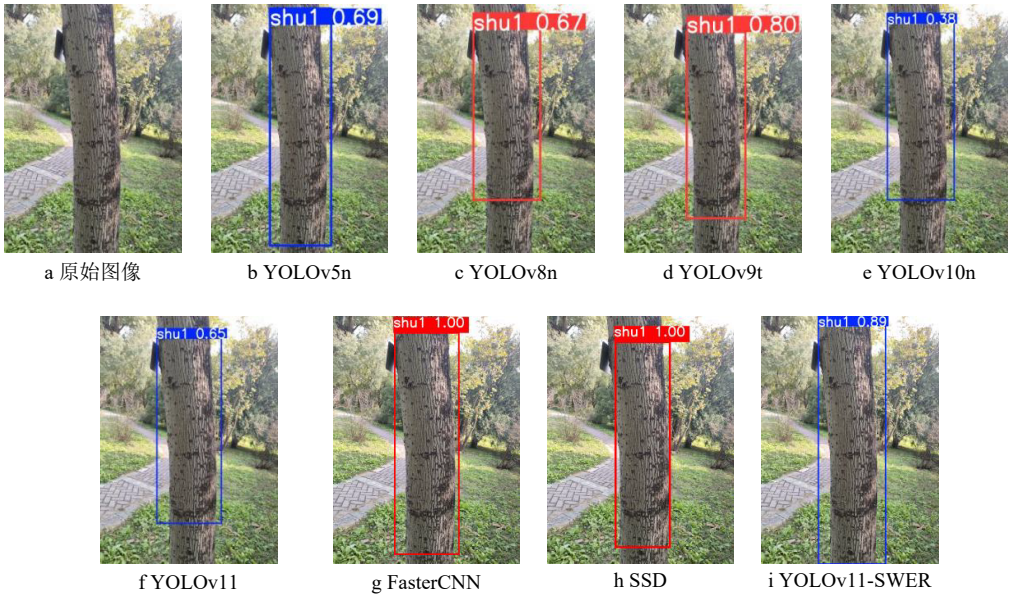
模型	模型体积/MB	参数量	FLOPs	帧速率/(帧·s^{-1})	mAP50	mAP50-95
MobileNetv2	13.62	3.46×10^6	8.9×10^9	77	0.994	0.751
ShuffleNetv2	9.91	2.49×10^6	6.3×10^9	91	0.992	0.733
StarNet	4.04	1.96×10^6	5.1×10^9	88	0.993	0.753
YOLOv11	5.28	2.60×10^6	6.4×10^9	83	0.992	0.750
YOLOv11-SWER	3.01	1.38×10^6	3.3×10^9	86	0.993	0.750



shu1 ~ shu70 代表树种编号，对应树种见表 1。

图 7　YOLOv11-SWER 对 70 种树种的识别性能雷达图

Fig. 7　Radar chart of YOLOv11-SWER's recognition performance for 70 tree species



a 原始图像　b YOLOv5n　c YOLOv8n　d YOLOv9t　e YOLOv10n

f YOLOv11　g FasterCNN　h SSD　i YOLOv11-SWER

检测框附带的信息内容为类别标签与对应的置信度，即模型对于检测框内存在该类别对象的确信程度。

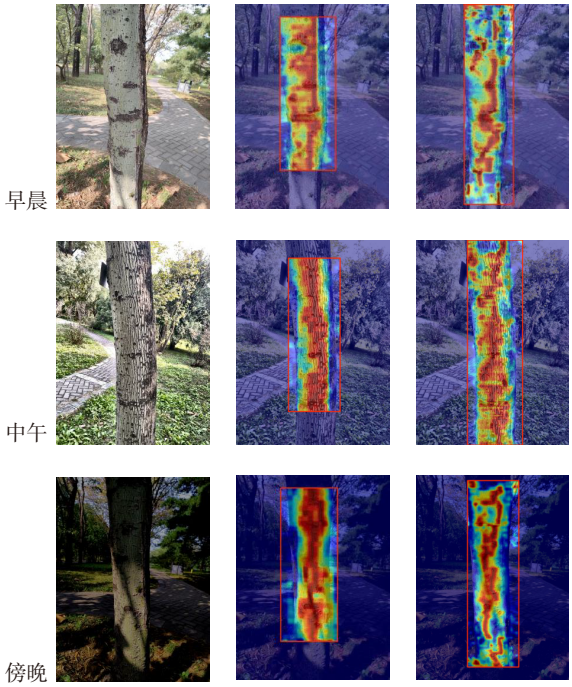
图 8　不同模型对梧桐的检测对比

Fig. 8　Detection comparison for *Firmiana simplex* by different models

(3)在 3 种不同光照场景下，改进模型均对树皮纹理特征保持更高的注意力。这验证了模型在实现模型轻量化的同时，仍能精准捕捉判别性特征，确保高效且高精度的模型性能，且对光照条件变化具有一定的鲁棒性。

4.4　不同光照场景下的模型性能

为进一步评估 YOLOv11-SWER 在不同光照场景下的鲁棒性，设计了早晨、中午、傍晚 3 个不同光照场景，通过对比模型在各场景下对数据集中所有树种的识别精度进行分析，实验结果如表 5。由表 5 可知，尽管 YOLOv11-SWER 在部分光照条件下的精确率和召回率略低于 YOLOv11，但其 mAP50 在 3 个场景下的波动幅度与 YOLOv11 表现相同，均为 0.015，且尽管在傍晚场景下 mAP50-95 指标稍



热力图采用颜色梯度表征模型的特征关注分布：红色区域代表高响应特征，即模型认为该部分特征对分类决策最重要；黄色区域为中等响应特征；蓝色区域则对应低响应特征，模型认为该部分对决策贡献较小。

图 9 不同光照场景下对梧桐的检测热力对比图

Fig. 9 Test thermal comparison chart of *Firmiana simplex* under different lighting scenes

逊于原始模型，但在早晨与中午场景下的 mAP50-95 指标上更优，说明 YOLOv11-SWER 在目标框定位精度上具有一定优势。选取银杏（树种编号 9）分别在早晨、中午、傍晚 3 个场景下的图片进行检测对比，结果如图 10 所示。3 种光照条件下模型检测框均能准确框定银杏树皮区域，且可信度都保持在 90% 以上。总体来说，YOLOv11-SWER 在引入一系列轻量化改进基础上，在不同光照场景下仍然能具有较好的鲁棒性，符合模型设计预期。

5 结 论

针对现有树种识别方法在复杂光照环境中计算复杂度高，难以满足在硬件资源极端受限的移动端实现部署需求的问题，本研究提出了一种基于树皮纹理的 YOLOv11-SWER 树皮纹理识别方法。YOLOv11-SWER 通过引入 StarNet 主干网络，将主干网络部分轻量化；采用 RepNCSPELAN4 多分支特征融合模块，在不同尺度下提升模型对复杂背景的鲁棒性；结合 Wavelet Pool 小波池化机制，提高对树皮纹理高频信息的捕捉能力；利用 Detect_Efficient 高效检测头，优化类别预测。这些设计使得模型仍保持 99.3% 的 mAP50 精度的同时，使参数量减少了 47.1%，计算量降低了 51.5%，并达到 86 帧/s 的检测速度。

本文提出的 YOLOv11-SWER 模型在轻量化与检测精度之间取得了良好平衡，但是仍存在一定的局限性。在复杂场景下，模型在形态高度相似的树种（如黑松与乔松）上仍存在一定的误检现象。为了提升其对相似树种的区分能力，后续考虑加入多模态数据信息，如树种的叶片形态、树枝形态、树冠纹理等，针对每一种树种信息训练检测模型，在识别树种时以模型投票的形式综合多个模型的判定结果，进而判定树种类型。此外，本研究未开展具体的实际部署实验，模型在边缘设备和移动端上的适应性和稳定性尚需进一步验证，未来将研究网络剪枝、量化和知识蒸馏等技术，在保持检测精度的同时，进一步压缩模型体积，并降低计算负载，以适应移动端与边缘设备等资源受限平台。最终，该模型可迁移至无人机、智能检测设备等低功耗设备场景进行进一步的研究。

本研究提出的 YOLOv11-SWER 不仅在树皮识别任务中展现了较高精度的识别能力，同时还大幅增加了将其在硬件资源极端受限的移动设备或边缘设备上进行部署的可能性。YOLOv11-SWER 为基于树皮纹理的树种识别提供了一种兼具精度与效率的技术路径，具备良好的边缘设备和移动端部署潜力，后续工作将进一步推动其在智能林业检测、城市林业资源管理等场景中的实际应用。

表 5 不同光照场景下模型在全部树种上的性能对比实验结果

Tab. 5 Comparative experimental results of model performance on all tree species under different lighting scenarios

模型	场景	*P*/%	*R*/%	F_1	mAP50	mAP50-95
YOLOv11	早晨	95.4	94.7	0.950	0.980	0.723
	中午	96.9	97.6	0.972	0.990	0.738
	傍晚	96.4	94.0	0.952	0.975	0.725
YOLOv11-SWER	早晨	93.9	92.8	0.933	0.978	0.724
	中午	97.2	96.3	0.967	0.991	0.742
	傍晚	94.4	92.7	0.935	0.976	0.723



a 早晨　　b 中午　　c 傍晚

图 10　不同场景下 YOLOv11-SWER 对银杏的检测对比

Fig. 10　Detection comparison for *Ginkgo biloba* under different scenarios by YOLOv11-SWER

参　考　文　献

[1] 徐竞怡, 张志, 闫飞, 等. 基于 GAN-DCNN 的树叶识别 [J]. 林业科学, 2024, 60(4): 40−51.
Xu J Y, Zhang Z, Yan F, et al. Leaf identification based on GAN-DCNN[J]. Scientia Silvae Sinicae, 2024, 60(4): 40−51.

[2] 高欣, 杨立新, 陈振举. 基于树木年轮径切特征的卷积神经网络树种识别 [J]. 应用生态学报, 2023, 34(1): 47−57.
Gao X, Yang L X, Chen Z J. Convolutional neural network tree species identification based on tree-ring radial section image features[J]. Chinese Journal of Applied Ecology, 2023, 34(1): 47−57.

[3] 谭灿灿, 聂稳, 刘逸夫, 等. 树皮表形与生理生化地理学研究进展 [J]. 林产工业, 2022, 59(11): 32−36.
Tan C C, Nie W, Liu Y F, et al. Research progress of bark morphology and physiological and biochemical geography[J]. China Forest Products Industry, 2022, 59(11): 32−36.

[4] 多化豫, 高峰, 李福胜, 等. 基于图像处理的木片与树皮的新识别参数研究 [J]. 西北林学院学报, 2015, 30(1): 207−210.
Duo H Y, Gao F, Li F S, et al. Approaching to the new identification parameteron wood and bark based on image processing[J]. Journal of Northwest Forestry University, 2015, 30(1): 207−210.

[5] 李晨光, 李亚敏, 孟腊梅. 基于遗传规划的树皮纹理图像识别 [J]. 农机化研究, 2012, 34(2): 185−188.
Li C G, Li Y M, Meng L M. Bark texture image recognition based on genetic programming[J]. Journal of Agricultural Mechanization Research, 2012, 34(2): 185−188.

[6] Zhao Y, Gao X, Hu J, et al. Tree species identification based on the fusion of bark and leaves[J]. Mathematical Biosciences and Engineering, 2020, 17(4): 4018−4033.

[7] 刘嘉政, 王雪峰, 王甜. 基于深度学习的 5 种树皮纹理图像识别研究 [J]. 北京林业大学学报, 2019, 41(4): 146−154.
Liu J Z, Wang X F, Wang T. Research on image recognition of five bark texture images based on deep learning[J]. Journal of Beijing Forestry University, 2019, 41(4): 146−154.

[8] 肖恒玉, 朱洪前, 杨滨, 等. 基于快速傅里叶增强深度网络的树皮种类识别 [J]. 林业工程学报, 2024, 9(4): 122−129.
Xiao H Y, Zhu H Q, Yang B, et al. Bark species recognition using fast Fouriertransform-enhanced deep network[J]. Journal of Forestry Engineering, 2024, 9(4): 122−129.

[9] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]//O'Coner L. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 779−788.

[10] 陈万志, 袁航. 改进 YOLOv8n 的林业害虫检测方法 [J]. 北京林业大学学报, 2025, 47(2): 119−131.
Chen W Z, Yuan H. Forestry pest detection method based on improved YOLOv8n[J]. Journal of Beijing Forestry University, 2025, 47(2): 119−131.

[11] 王婧怡, 许杰. 基于改进 YOLOv7 算法的无人机图像树种识别 [J/OL]. 林业工程学报, 1−9[2025−03−10]. https://doi.org/10.13360/j.issn.2096-1359.202405030.
Wang J Y, Xu J. Research on tree species recognition in drone images using an improved YOLOv7 algorithm[J/OL]. Journal of Forestry Engineering, 1−9[2025−03−10]. https://doi.org/10.13360/j.issn.2096- 1359.202405030.

[12] 杨灿, 范习健, 张九于. SSFYOLO: 一种面向复杂森林场景的树干检测算法 [J]. 北京林业大学学报, 2025, 47(2): 132−142.
Yang C, Fan X J, Zhang J Y. SSFYOLO: a trunk detection algorithm for complex forest scenarios[J]. Journal of Beijing Forestry University, 2025, 47(2): 132−142.

[13] 马宏兴, 董凯兵, 王英菲, 等. 基于改进 YOLO v5s 的轻量化植物识别模型研究 [J]. 农业机械学报, 2023, 54(8): 267−276.
Ma H X, Dong K B, Wang Y F, et al. Lightweight plant recognition model based on improved YOLO v5s[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(8): 267−276.

[14] 齐建东, 马鐘添, 郑尚姿. 基于 YOLOv7 的红外相机野生动物图像筛选 [J]. 北京林业大学学报, 2024, 46(2): 143−154.
Qi J D, Ma Z T, Zheng S Z. Wildlife image screening for infrared cameras based on YOLOv7[J]. Journal of Beijing Forestry University, 2024, 46(2): 143−154.

[15] 韩誉, 刘浩然, 林文树. 基于改进 YOLOv10 和无人机影像的树种识别 [J/OL]. 森林工程, 1−15[2025−05−28]. http://kns.cnki.net/kcms/detail/23.1388.s.20250425.1736.006.html.
Han Y, Liu H R, Lin W S. Tree species identification based on improved YOLOv10 and UAV RGB imagery[J/OL]. Forest Engineering, 1−15[2025−05−28]. http://kns.cnki.net/kcms/detail/23.1388.s.20250425.1736.006.html.

[16] 许超, 王云健, 刘洋, 等. 基于改进 Swin Transformer 的膝骨关节炎 X 光影像自动诊断 [J]. 电子测量技术, 2024, 47(19): 155−163.
Xu C, Wang Y J, Liu Y, et al. Knee osteoarthritis based on improved Swin Transformer X-ray image automatic diagnosis[J]. Electronic Measurement Technology, 2024, 47(19): 155−163.

[17] Khanam R, Hussain M. Yolov11: an overview of the key architectural enhancements[J/OL]. arXiv preprint arXiv: 2410.17725, 2024. [2025−04−04]. https://doi.org/10.48550/arXiv.2410.17725.

[18] Ma X, Dai X, Bai Y, et al. Rewrite the stars[C]//Ceballos C. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2024: 5694−5703.

[19] 方明, 张娇, 徐晶, 等. 基于改进 YOLOv8 的轻量化皮革缺陷检测方法 [J]. 电子测量技术, 2025, 48(1): 111−118.
Fang M, Zhang J, Xu J, et al. Lightweight leather defect detection method based on improved YOLOv8[J]. Electronic Measurement Technology, 2025, 48(1): 111−118.

[20] Wang C Y, Liao H Y M, Wu Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN[C]//O'Coner L. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Seattle: IEEE, 2020: 390−391.

[21] Wang C Y, Liao H Y M, Yeh I H. Designing network design strategies through gradient path analysis[J/OL]. arXiv preprint arxiv: 2211.04800, 2022. [2025−04−04]. https://doi.org/10.48550/arXiv.2211.04800.

[22] Williams T, Li R. Wavelet pooling for convolutional neural networks[EB/OL]. (2018−02−16)[2025−04−15]. https://openreview.net/forum?id=rkhlb8lCZ.

[23] 王泉, 叶广飞, 陈祺东. YOLO-SWR: 无人机视角下轻量级交通车辆检测算法 [J/OL]. 计算机工程与应用, 1−15[2025−04−04]. http://kns.cnki.net/kcms/detail/11.2127.tp.20250403.1521.012.html.
Wang Q, Ye G F, Chen Q D. YOLO-SWR: lightweight traffic vehicle detection algorithm from UAV perspective[J/OL]. Computer Engineering and Applications, 1−15[2025−04−04]. http://kns.cnki.net/kcms/detail/11.2127.tp.20250403.1521.012.html.

[24] Ren S, He K, Grishick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(6): 1137−1149.

[25] Mahesha Y, Nagaraju C. Spotting congenital heart diseases using palm print based on faster R-CNN and spatial method[J]. International Journal of Medical Engineering and Informatics, 2024, 16(1): 56−70.

[26] Sandler M, Howard A, Zhu M L, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]//O'Coner L. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 4510−4520.

[27] Ma N, Zhang X, Zheng H T, et al. ShuffleNetV2: practical guidelines for efficient CNN architecture design[C]//Shen L L, Asari V, Suzuki K, et al. Proceedings of the European Conference on Computer Vision. Munich: Springer, 2018: 116−131.

[28] Premakumara N, Jalaian B, Suri N, et al. Improving object detection robustness against natural perturbations through synthetic data augmentation[C]//Shen L L, Asari V, Suzuki K, et al. Proceedings of the 2023 Asia Conference on Computer Vision, Image Processing and Pattern Recognition. Phuket: Association for Computing Machinery, 2023: 1−6.

[29] 刘澳龄, 单武扬, 邱骏颖, 等. 基于图像恢复和空间通道注意力的下采样图像取证网络 [J/OL]. 计算机应用, 1−8 [2025−04−04]. http://kns.cnki.net/kcms/detail/51.1307.tp.20241105.1632.002.html.
Liu A L, Shan W Y, Qiu J Y, et al. Downsampled image forensic networks based on imagerecovery and spatial channel attention[J/OL]. Journal of Computer Applications, 1−8 [2025−04−04]. http://kns.cnki.net/kcms/detail/51.1307.tp.20241105.1632.002.html.

[30] 谢雨欣, 龚烨飞, 谷心浩, 等. 基于 RGB-D 特征融合的焊缝表面缺陷检测方法 [J]. 焊接学报, 2024, 45(12): 72−78.
Xie Y X, Gong Y F, Gu X H, et al. Research on weld surface defect detection method based on RGB-D feature fusion[J]. Transactions of the China Welding Institution, 2024, 45(12): 72−78.

[31] 郑万钧. 中国树木志 [M]. 北京: 中国林业出版社, 1985: 162−306.
Zheng W J. Chinese tree catalogue[M]. Beijing: China Forestry Publishing House, 1985: 162−306.

[32] 中国科学院中国植物志编辑委员会. 中国植物志 (第 7 卷 裸子植物门)[M]. 北京: 科学出版社, 1978: 225−270.
Editorial Committee of Flora of China, Chinese Academy of Sciences. Flora of China (Vol. 7: Gymnospermae) [M]. Beijing: Science Press, 1978: 225−270.

[33] Chattopadhyay A, Sarkar A, Howlader P, et al. Grad-CAM++: generalized gradient-based visual explanations for deep convolutional networks[C]//O'Coner L. Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision. Lake Tahoe: IEEE, 2018: 839−847.

（责任编辑 吴 娟
责任编委 陈志泊）