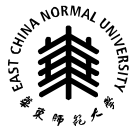


2025届硕士专业学位研究生学位论文

分 类 号：_____
密 级：_____

学校代码：10269
学 号：71255902007



華東師範大學

East China Normal University

硕士专业学位论文

Master' s Degree Thesis (Professional)

论文题目： 基于改进Y0L0v8的植物
叶表气孔检测系统研究

院 系：	软件工程学院
专业学位类别：	电子信息
专业学位领域：	软件工程
学位 申 请 人：	郭东方
指 导 教 师：	刘垚（副） 教授

2025年4月

Thesis for Master's Degree (Professional) in 2025

University code: 10269

Student ID: 71255902007

East China Normal University

Title:

Research on Plant Leaf Stomata Detection

System Based on Improved YOLOv8

Department/School: Software Engineering Institute

Category: Electronic and Information Engineering

Field: Software Engineering

Candidate: DongFang Guo

Supervisor: A/Prof Yao Liu

Apr, 2025

华东师范大学学位论文原创性声明

郑重声明：本人呈交的学位论文《基于深度学习的智能车辆异常检测系统研究》，是在华东师范大学攻读~~硕士~~/博士（请勾选）学位期间，在导师的指导下进行的研究工作及取得的研究成果。除文中已经注明引用的内容外，本论文不包含其他个人已经发表或撰写过的研究成果。对本文的研究做出重要贡献的个人和集体，均已在文中作了明确说明并表示谢意。

作者签名 徐杰

日期2024年5月9日

华东师范大学学位论文著作权使用声明

《基于深度学习的智能车辆异常检测系统研究》系本人在华东师范大学攻读学位期间在导师指导下完成的~~硕士~~/博士（请勾选）学位论文，本论文的著作权归本人所有。本人同意华东师范大学根据相关规定保留和使用此学位论文，并向主管部门和学校指定的相关机构送交学位论文的印刷版和电子版；允许学位论文进入华东师范大学图书馆及数据库被查阅、借阅；同意学校将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于（请勾选）

（ ）1. 经华东师范大学相关部门审查核定的“内部”或“涉密”学位论文*，于 年 月 日解密，解密后适用上述授权。

（☒）2. 不保密，适用上述授权。

导师签名

作者签名

徐杰

日期2024年5月9日

* “涉密”学位论文应是已经华东师范大学学位管理办公室或保密委员会审定过的学位论文（需附获批的《华东师范大学研究生申请学位论文“涉密”审批表》方为有效），未经上述部门审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权）。

郭东方 硕士学位论文答辩委员会成员名单

姓名	职称	单位	备注
吴亚栋	副教授	夏普（上海）高科技研发中心	主席
陈志云	教授	华东师范大学	
江红	副教授	华东师范大学	

摘要

气孔是植物与大气进行气体交换的关键通道，其形态特征与调控机制直接影响植物碳吸收和水分利用效率，对理解全球碳-水循环及植物环境适应策略具有重要意义。传统气孔形态检测依赖人工测量，存在效率低、主观性强等局限，难以满足现代植物生理生态研究需求。本研究基于深度学习技术，提出改进的 YOLOv8 算法，开发多模态数据兼容的自动化气孔分析系统。通过融合自主采集与公开数据集融合多物种气孔图像库，结合轻量化网络设计、多尺度注意力机制和边界框损失函数优化，显著提升微小气孔检测精度与模型泛化能力。系统集成参数自适应调节与可视化功能，实现气孔数量、分布及面积的快速解析，检测效率较传统方法提升数倍。结果表明，该系统在复杂背景下仍能保持高精度检测，为植物功能性状研究提供了高效工具。主要工作如下：

1. 构建数据集，夯实植物气孔检测数据基础。建立覆盖 72 个物种的气孔图像数据库，通过标准化采集流程获取多尺度显微影像。采用多阶段质量筛选机制，结合专业设备拍摄与公开数据融合策略，最终形成包含 665 张高质量图像的数据集。基于半自动标注工具完成 18482 个气孔实例的精细标注，为模型训练提供可靠的数据集。
2. 算法优化，提升植物气孔检测系统性能。面向移动端部署需求，重构 YOLOv8 网络架构，采用 MobileNet4 轻量化主干网络，检测速度提升 32%，模型大小降低 18%。替换 EMA 注意力模块，增强模型对小气孔特征的捕获能力，小目标检测精度提升 4.3%。引入 SIoU 损失函数优化边界框回归，有效解决复杂背景下气孔重叠检测难题。经对比实验验证，改进模型在 mAP@50、召回率等关键指标上均有所提升，实现检测速度与精度的最佳平衡。
3. 系统开发，打造高效植物气孔检测平台。基于优化的 YOLOv8 算法，开发植物叶表气孔检测系统。该系统大幅优化气孔统计流程，相比传统人工统计，显著提高气孔密度统计、平均面积计算等工作的标准化程度，为植物生理生态研究提供可靠的数字化工具支撑。

综上所述，本研究通过构建多物种气孔数据库、优化轻量化检测算法及开发智能分析系统，实现了气孔特征的高效精准解析。该系统检测效率较传统方法大幅提升，为植物碳-水调控机制研究提供了可靠工具。在全球气候变化背景下，该成果不仅助力解析植物环境适应策略，更可为生态系统建模和抗旱作物选育提供技术支撑。

关键词：YOLOv8 算法；叶表气孔；目标检测

ABSTRACT

Stomata are key channels for gas exchange between plants and the atmosphere. Their morphological characteristics and regulatory mechanisms directly influence the carbon absorption and water-use efficiency of plants, which is of great significance for understanding the global carbon - water cycle and plant environmental adaptation strategies. Traditional detection of stomatal morphology relies on manual measurement, suffering from limitations such as low efficiency and strong subjectivity. Moreover, dynamic observation technologies also have obvious deficiencies, making it difficult to meet the needs of modern plant physiological and ecological research.

In this study, based on deep - learning technology, an improved YOLOv8 algorithm is proposed, and an automated stomatal analysis system compatible with multimodal data is developed. By integrating self - collected and public datasets, a multi - species stomatal image library is constructed. Combining lightweight network design, multi - scale attention mechanisms, and optimization of the bounding - box loss function, the detection accuracy of small - target stomata and the generalization ability of the model are significantly enhanced. The system integrates parameter adaptive adjustment and visualization functions, enabling rapid analysis of the number, distribution, and morphological characteristics of stomata. The detection efficiency is dozens of times higher than that of traditional methods. The results show that the system can still maintain high - precision detection under complex backgrounds, providing an efficient tool for the study of plant functional traits. The main work is as follows:

- 1 Constructing the dataset to lay a solid foundation for plant stomatal detection data: A stomatal image database covering 72 plant species is established. A multi - stage quality - screening mechanism is adopted, and by combining the shooting with professional equipment and the integration strategy of public data, a core dataset containing 665 high - quality images is finally formed. Based on semi - automatic annotation tools, fine annotation of 18,482 stomatal instances is completed, constructing a data base with

morphological diversity and environmental adaptation characteristics, providing a reliable guarantee for deep - model training.

- 2 Optimizing the algorithm to improve the performance of the stomatal detection system: To meet the deployment requirements for mobile devices, the YOLOv8 network architecture is reconstructed. The lightweight MobileNet4 backbone network is adopted, which increases the detection speed by 32% and reduces the model size by 18%. A cross - scale EMA attention module is designed to enhance the model's ability to capture the features of small stomata, improving the detection accuracy of small targets by 4.3%. The SIoU loss function is introduced to optimize bounding - box regression. Through comparative experiments, it is verified that the improved model has improved in key indicators such as mAP@50 and recall rate, achieving the best balance between detection speed and accuracy.
- 3 Developing the system to create an efficient plant stomatal detection platform: Based on the optimized YOLOv8 algorithm, a stomatal detection system for plant leaf surfaces is developed. The system has user - login, personalized - configuration, and result - display modules. The system greatly optimizes the stomatal statistics process. Compared with traditional manual statistics, it significantly improves the standardization of tasks such as stomatal density statistics and average - area calculation, providing reliable digital tool support for plant physiological and ecological research.

In summary, this research achieved efficient and precise analysis of stomatal characteristics by constructing a multi-species stoma database, optimizing a lightweight detection algorithm, and developing an intelligent analysis system. providing a reliable tool for studying the carbon-water regulation mechanism in plants. In the context of global climate change, this achievement not only helps to analyze the environmental adaptation strategies of plants but also provides technical support for ecosystem modeling and the breeding of drought-resistant crops.

Keywords: YOLO algorithm; Leaf stomata; object detection

目录

摘要.....	I
ABSTRACT.....	III
第一章 绪论.....	1
1.1 研究背景及意义	1
1.2 国内外研究现状	2
1.3 研究内容	6
1.4 本文的组织结构	7
第二章 相关技术概述	9
2.1 卷积神经网络	9
2.1.1 卷积层	10
2.1.2 池化层	12
2.1.3 全连接层	13
2.1.4 激活函数	14
2.2 目标检测算法概述	16
2.2.1 传统目标检测算法	17
2.2.2 两阶段目标检测算法	18
2.2.3 单阶段目标检测算法	20
2.3 本章小结	25
第三章 植物叶表气孔数据集融合与预处理	26
3.1 数据采集	26
3.2 数据增强	29
3.3 数据标注	31
3.4 本章小结	33
第四章 基于 YOLOv8 的模型改进和实验	34
4.1 目标检测方法常用性能指标	35

4.2 主干网络改进	36
4.3 小目标检测优化	38
4.4 损失函数设计	40
4.5 实验结果和分析	41
4.5.1 实验环境	42
4.5.2 替换主干网络对模型检测效果的影响	42
4.5.2 添加注意力机制对模型检测效果的影响	43
4.5.3 损失函数对模型检测效果的影响	44
4.5.4 迁移实验	44
4.6 本章小结	45
第五章 植物叶表识别系统的设计与实现	47
5.1 开发环境	47
5.2 需求分析	47
5.3 系统设计	49
5.4 系统实现	51
5.5 本章小结	55
第六章 总结和展望	56
6.1 总结	56
6.2 展望	56
参考文献	58
致谢	64
攻读硕士学位期间所获科研成果	65

第一章 绪论

1.1 研究背景及意义

气孔多分布于植物叶片表面，一般包括 2 个保卫细胞以及中间的凸透镜状的小孔（图 1-1）。叶片气孔是植物与大气进行碳水交换的主要器官，约 98% 的二氧化碳吸收和水分流失通过气孔完成^[1]。叶片气孔通过开闭行为和形态变异来调控植物的光合作用与蒸腾作用，从而影响植物的水分利用效率。提高水分利用效率是促进作物改良、提升作物产量的核心策略之一^{[2][3]}。在短期内，植物通过精细调整保卫细胞的膨压，实现对气孔孔径的灵活控制，进而达到气孔开闭的精准调节（图 1-2）。从长期来看，气孔的形态直接影响气孔开闭的速度及其稳定性，从而实现对气孔功能的长期优化调节^{[4][5][6]}。气孔的特征包括气孔密度、面积、长度、宽度，是刻画气孔结构以及指征气孔功能的关键性状^[7]。在生理学、生态学以及农业科学的深入探索中，准确评估这些指标，对于揭示植物生理机制、理解生态过程以及优化农业实践均具有重要意义。

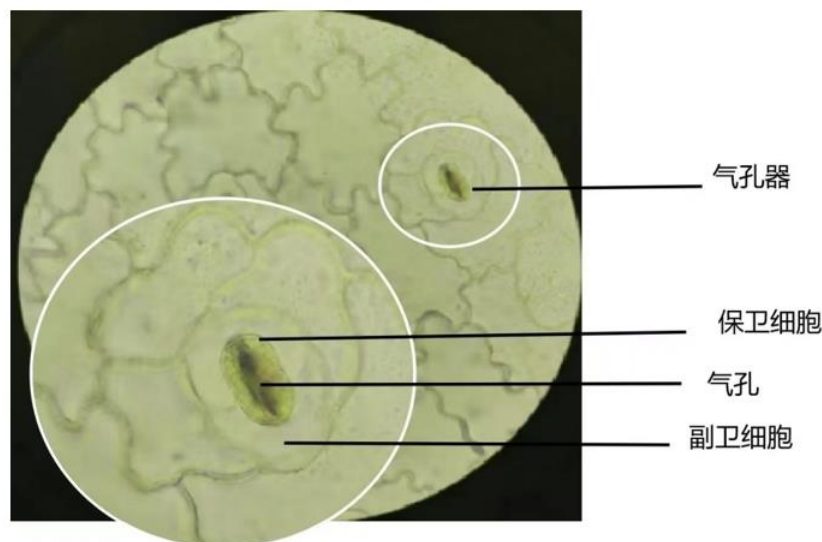


图 1-1 显微镜下的植物叶表气孔及其周围的细胞示意图

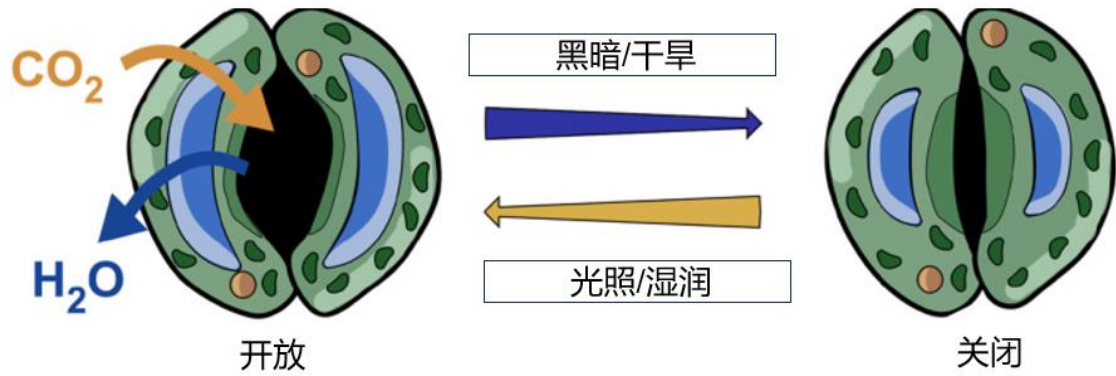


图 1-2 保卫细胞调整气孔开和闭的示意图

传统的气孔观测存在诸多弊端，不仅耗时耗力，而且成本高昂。在应对大规模样本筛选任务时，其局限性愈发显著，难以满足高效、精准的研究需求。以指甲油印记法为例，作为获取植物叶表气孔特征的主流技术之一，其操作流程颇为繁杂：首先在叶片表面涂抹透明指甲油以制备印模，待指甲油干燥后将印模剥离；接着借助光学或电子显微镜观察印模并进行图像捕捉；最后依靠人工手动测量相关特征。采用手动方式测量气孔特征效率低下，从气孔图像的获取到后续数据分析的整个流程极为繁琐，该方法无法满足大规模、多物种气孔观测的迫切需求。随着技术的持续革新，深度学习在气孔测定方面的应用已显著提升了传统方法的效率，但当前研究方法仍存在一定的局限性。例如，当前绝大多数研究方法仅测定气孔的相对面积大小，这一限制在很大程度上制约了气孔研究的深入进行。

本文紧扣植物叶片表面气孔检测的实际需求，深度关联应用场景，精心选取YOLOv8作为植物叶片气孔检测任务的算法框架。在此基础上，本研究创新性地提出了一系列改进策略，并成功构建起一套能够自动且高效识别植物叶片表面气孔的检测系统。该系统具备强大的兼容性，能够对各类数据源实现有效兼容。在显著提升气孔检测效率方面，此系统显现出了卓越不凡的性能，有望为相关研究与应用带来全新的突破。

1.2 国内外研究现状

气孔是植物由水生向陆生发展过程中产生的特化器官，大量分布在叶片表面。气孔具有不同的形状、大小、密度、分布等特征。其中，最常见的两个气孔特征是气孔

密度和气孔面积。气孔密度是衡量作物抗旱能力的一个重要指标。气孔密度越大，光合速率越快，对水分的利用率越高，但同时也会加剧植物对干旱胁迫的敏感性。气孔密度具有复杂的变异性：在植物内部，冠层顶部的气孔密度高于冠层底部，冠层外部的气孔密度高于冠层内部；在同一叶片的不同部位，与叶边缘相比，气孔密度在越靠近叶中脉的位置越低；与叶基部相比，越靠近叶片顶部的位置气孔密度越低。在对环境变化的响应上，升高的 CO_2 浓度会降低气孔密度，光照和干旱的加强会增加气孔密度^[7]。气孔面积的计算方式有多种，一种是将气孔的长看作气孔器的长度，宽看作气孔器宽度的长方形。或将其长轴作为气孔器长度，短轴作为气孔器宽度的椭圆形。气孔面积与植物的气体交换能力密切相关，较大的气孔面积会增加水分蒸腾速率，不利于植物在干旱环境中生存。因此，在高温干燥的环境中，植物会主动减小气孔面积，以减少水分的散失，尽管这一策略可能对产量提升产生不利影响^[8]。

气孔的观测包括气孔成像和指标测定两部分。其中，气孔成像是从叶片表面获取气孔的微观图像。依据叶片的采样方法，气孔成像分为破坏性和非破坏性两种类型。最传统的气孔成像方法是印记法，属于破坏性方法。该方法需要采集离体的叶片，使用固定剂固定气孔后带回实验室在显微镜下观察^[9]。尽管传统方法获得的图像清晰，但这些方法耗时低效、易受人为错误影响。而且，气孔对水分变化高度敏感。叶片离体后气孔会由于缺水而快速关闭，不利于气孔指标的准确测定。鉴于现代科研需求的日益增长，传统方法已逐渐暴露出其局限性，限制了其在应对复杂研究挑战中的应用潜力。近年来一些工作旨在改进这种方法。非破坏性方法是指直接在叶片表面进行原位观测，不需将叶片摘下。Eisele 小组^[10]使用罗丹明对叶片进行预染色，并结合原生质体隔离方法，既不影响气孔的关闭或打开，又简化了气孔特征的测量。Ourdais 团队使用自动旋转 Opera 显微镜对孔板上的叶片进行成像，具有高成像效率^[11]。Millstead 等将压印装在显微镜的玻片上，改进了常规印记法，可以减少叶片压印的不均匀性，并允许浅景深传感器保持样本较大区域处于焦点状态^[12]。另一种气孔成像方法是非破坏性的，Song 等人通过 Keyence VHX-200 获取活叶的图像^[13]，Sun 团队通过便携显微镜 3R-MSUSB601 成像^[14]。这种方法可以实现更高的成像效率，无需准备样品。当前气孔成像技术中，传统破坏性方法存在操作流程繁琐、样本易损等局限，而非破坏性方法的难点在于不同植物物种气孔特征差异显著（如大小、分布密度、表

面纹理等），导致难以找到能够适配所有物种的通用成像设备。基于此，本研究通过系统性调研与实验验证，最终从市面上常用设备中筛选出 3 款具有代表性的非破坏性成像设备。拍摄结果表明，所选设备在多物种气孔成像中均表现出优异的适应性与成像质量，为后续自动化检测算法开发提供了可靠的数据采集支撑。

气孔指标的测定一般使用 ImageJ 软件。将气孔图像导入软件后，对气孔手动计数，并测量气孔的长度、宽度等指标。这种方法非常耗时，对人力需求较大。此外，主观判断的差异可能引入测量误差。随着科技的飞速发展，以及数码影像处理的日趋成熟，电脑视觉已逐渐成为植物学研究的有力工具。图像处理技术在检测植物叶片气孔和测量形态中被广泛应用。Sanyal 等通过扫描电镜技术，研究并分析了不同品种的番茄叶片气孔形貌。他们采用分水岭算法和形态处理来将单个气孔与其他相邻的气孔分离，并测量了单个气孔的面积、重心^[15]。Martins 团队利用 UIHISOSA 图像处理软件完成了气孔特征参数的精确计算^[16]。Laga 使用模板匹配来检测小麦的单个气孔细胞，并通过局部分析来测量气孔的特征，但他们使用的模板并不适用于其他作物^[17]。Takumi 等人采用了一种结合遗传算法和聚类辅助快速训练代理自组织图算法的先进方法^[18]，从荧光图像中有效识别植物气孔。此方法的关键在于优化感兴趣区域（ROI, Region of Interest）的设置，但因其参数对成像方式及气孔特性的高度依赖性，需经广泛实验调整以确保检测精度。Duarte 研究小组则提出了一种基于多尺度小波的气孔识别策略，该策略通过多尺度定位技术结合分水岭算法，提升了气孔结构的预测精度，尽管它在气孔图像质量和处理复杂度上有所要求^[19]。Liu 等提出了一种基于最大稳定外部区域的葡萄叶片气孔直径检测算法，并研制了相关软件^[20]。Hiranya 等人建立了一种基于级联目标检测器的气孔检测与分割模型，在此基础上又通过对级联分类器的训练，并将其与滑动窗口相结合，实现对气孔的准确识别与分割^[21]。

尽管上述传统计算机视觉方法在气孔检测领域取得了一定进展，但其依赖人工特征设计与复杂参数调优的局限性逐渐凸显。传统方法需通过特定算法（如分水岭、模板匹配）分离气孔结构，这一过程不仅耗时且易受图像质量、气孔形态多样性的影响。然而，深度学习具有很强的特征自动学习能力，因此在物体检测方面表现出了明显的优势。通过构建端到端的深度神经网络模型，该技术能够从图像中自主挖掘多尺度、多层次的判别特征，有效克服了传统方法对人工特征工程的依赖。同时，深度模型具

备更强的鲁棒性，显著降低了对参数调整的敏感度，为高通量、高精度的气孔表型分析提供了新范式。近年来，深度学习算法在微观表型研究中具有广泛的应用前景，例如用于植物叶片中气孔的检测。Yosuke Toda 等人在前期研究基础上，创新性地将梯度方向直方图（HOG, Histogram of Oriented Gradients），它是一种在计算机视觉和图像处理中用于目标检测的特征描述子。其原理是通过计算和统计图，与卷积神经网络（CNN, Convolutional Neural Network）相结合，建立了一种新型的气孔识别算法，有效地提高了对气孔的识别能力^[22]。该方案的首要步骤是利用 HOG 技术从植物叶片图像中提取特征信息，以此为基础确定图像中可能存在气孔的区域；紧接着，借助 CNN 对这些潜在的气孔区域进行细致的分类处理。Otter 团队开创性地提出了一套气孔量化的自动化流程^[23]。该项目拟通过 CNN 对隐藏的气孔区进行快速识别，并使用基于实例的图像修补方法对气孔图中的遮挡进行校正。在此基础上又采用全卷积神经网络对气孔进行分割，使其能很好地处理由于形态多样性、大小不一以及相互遮挡等问题而导致的气孔检测问题。另一方面，Feter 等人则研发了一款名为 Stomata Counter 的基于 CNN 的自动气孔计数系统^[24]。该系统融入了人机协同技术来训练神经网络，使得其在银杏叶片的扫描电子显微镜图像上识别气孔的准确率高达 98.1%，即便面对未经训练的植物种类，其测试准确率也能维持在 94.2% 的高水平。Meeus 及其团队对比评估了三种不同 CNN 在辨识植物气孔方面的性能^[25]，并在其研究使用的特定数据集上发现，VGG (Visual Geometry Group) 模型展现出了最佳效果。Cowling 等人则选用了 Inception-V2 作为核心架构，构建了一个基于 Faster R-CNN 框架的气孔检测模型^[26]，该模型专注于水稻叶片的研究。研究者们使用 380 个叶片 20/40 倍的显微镜图片来检验该模型。试验证明，该模型方法可以准确地对气孔进行准确的分类，特别是对于 40 倍放大后的气孔，其与手工标记的一致性达到 99%。Shi 等人另辟蹊径，首次提出了一种新的基于最近邻体距离比例关系的计算方法，并在此基础上进一步研究了不同物种之间的气孔密度差异^[27]。Toda 科研小组利用便携式显微镜技术，打造了一个专门用于小麦气孔形态分析的平台^[28]，该平台的核心特征提取组件为 MobileNetV2，并在此基础上，结合单阶段多框检测（SSD, Single Shot MultiBox Detector）技术实现了气孔检测模块的开发。

单阶段目标识别算法 YOLO，近年来凭借其精简架构、卓越的处理速度，以及独特的优势，逐渐跃居主流目标检测算法之列。YOLO 最大的特点之一在于速度极快，它将目标分类和定位整合于一个神经网络中，实现端到端的训练，避免了复杂的预处理和后处理步骤，运行效率极高，能够满足实时性应用的严苛要求，比如早期的 YOLOv1 网络运行速度就能超过 45fps，这一特性在诸多实际场景中意义重大。同时，由于 YOLO 基于整个图像进行预测，背景误检率相对较低，有助于显著提高目标检测的准确性。针对 YOLO 模型，科研人员探索出了多种新颖的气孔识别策略，例如 Casado-Garcia 研究小组基于 YOLOv3 构建的气孔识别模型^[29]、Yang 研究团队采纳特征权重迁移学习策略提出的 YOLOv4 气孔分析途径^[30]、Ren 研究团队通过将通道注意力机制融入 YOLOv5 所打造的气孔检测架构^[31]，以及 Zhang 研究团队首先利用 YOLOv5 进行气孔定位^[32]，随后结合超像素分割技术实现气孔参数精确测量的方案。

1.3 研究内容

本项目面向植物叶表气孔检测这一难题，研究一套基于 YOLOv8 的气孔识别系统，并利用融合的叶表气孔图像数据集对其进行检验。实验结果表明，所提出的改进算法显著提升了检测精度与性能。此外，本研究还开发了一个自动化气孔检测系统，能够适应多种拍摄设备的需求。具体的研究内容与方法包括以下三个部分：

构建植物叶表气孔图像数据集。借助多种专业拍摄设备，收集了 72 个物种，共计 260 张植物叶表气孔图像。经过严格的人工筛选，最终挑选出 165 张图像作为本次研究的基本数据集，并与公开数据集的 500 张图像，创建了包含 665 张图像的融合数据集，作为本研究的核心基本数据集。为了达到模型训练的需求，本文采用图像标记工具 labelme，对 665 幅图像进行了详细的标注，共标记 18482 个真实标注信息。通过这一关键步骤，为后续模型训练工作奠定了坚实基础，确保模型能够在充足且高质量的数据支持下，实现更精准的学习与预测。

优化 YOLOv8 检测算法。本研究选用了 YOLOv8 作为基础检测算法，使用 MobileNetV4 网络替换了部分模型的主干部分，以便在移动设备上实现高效的植物叶表气孔检测。在检测过程中，发现细小气孔的漏检率较高，因此引入了高效多尺度注意力模块，进一步提高了该模型探测此类气孔的能力，并在此基础上引入边界回归损

失函数 SIoU，提高模型的预测精度。使用改进后的模型在标注数据集上进行了训练，评估了其在气孔检测中的表现。

开发基于改进的 YOLOv8 气孔自动检测系统。该系统具备登录、注册、检测、计算面积、统计个数等多种功能。系统采用了参数配置的方式，确保能够兼容不同设备的拍摄数据。并且能够计算真实面积，为不同物种之间的比较提供了依据。

1.4 本文的组织结构

全文共分六章，每一章的安排如下：

第 1 章绪论。介绍论文的研究背景和意义，阐明植物叶表气孔在植物生理中的作用，并对相关领域的研究现状进行回顾。同时，本章还介绍了研究目标、研究内容及论文结构。

第 2 章相关技术概述。对目前深度学习中两种主要的物体检测方法 R-CNN 系列和 YOLO 系列进行了深入的研究，并对这两种算法的优劣进行了比较。最后综合以上两种算法的优点，选取性能稳定且高效的 YOLOv8 为基本检测算法。

第 3 章涵盖了植物叶表气孔数据集的构建和标注。详细阐述了数据集的构建过程，数据增强技术，以及对数据集的扩充，最后给出了基于专用标记软件的图像标注过程。从而保证所构造的数据集在实际应用中具有较高的质量，并为后续的算法训练与试验奠定坚实的基础。

第 4 章提出了基于 YOLOv8 算法的改进策略。本章提出了针对植物叶表气孔检测需求的多项算法改进。将主干网络替换为 MobileNetV4 的模型，以更好地适应移动设备的部署需求；针对细小气孔的高漏检率，增加了 EMA 注意力模块以提升检测精度和召回率；为进一步提高检测效果，使用了 SIoU 损失函数优化模型训练。本章深入探讨了这些改进策略及其实现，为后续实验和结果分析奠定了基础。

第 5 章实现了植物叶表检测系统。基于改进后的 YOLOv8 网络模型，分析并设计了气孔检测系统架构，配置了实验环境并进行了模型训练。同时，使用 PyCharm 工具开发了植物叶表气孔检测系统，系统能够适配多种拍摄设备，为科研人员提供了一种新的叶表气孔检测途径。

第 6 章为总结与展望。是对本文所做工作的一个概括，对各个章节的主要内容及取得的成就做了简要的回顾，同时也对本文的创新之处及不足之处作了较为客观的分析。最后对今后的研究工作进行了展望，对今后的研究具有一定的参考价值和启发意义。

第二章 相关技术概述

在当今计算机技术飞速发展的背景下，深度学习推动的目标检测技术已成为人工智能领域的核心，极大地促进了智能化发展的进程，为我国科技创新提供了坚实的基础。本章旨在深入探讨植物叶片气孔检测领域的最新研究进展，重点分析深度学习、机器视觉技术、CNN 算法及目标检测技术等关键内容，为构建高效的植物叶片检测系统提供坚实的理论支持。

本章开篇即对卷积神经网络的核心理论进行精炼概述，为后续深入探索目标检测技术奠定坚实的理论基石。随后，本章详细剖析了常见的目标检测策略——传统目标检测算法、一阶段检测法和二阶段检测法，并对 Faster R-CNN、YOLO 系列算法这两种在业界广泛应用的目标检测算法的网络架构与运行机制进行了全面而深入的解读。通过综合对比分析这些算法的优劣，本研究最终选定 YOLOv8 算法作为优化升级的对象。

2.1 卷积神经网络

CNN 是由 Yann LeCun 提出的深度学习中的重要算法之一^[33]，主要用于处理具有网格结构的数据类型，该算法专门针对具有网格结构特征的数据类型（如二维图像、三维点云）进行优化设计，在图像处理领域展现出独特优势。一个典型的 CNN 结构通常包括输入层、卷积层、池化层、全连接层、激活函数以及输出层，这些组件共同实现了 CNN 的特征提取功能^[34]。CNN 凭借其卓越的适应性、出类拔萃的特征分类能力、良好的容错性能、并行处理的显著优势，以及强大的自学习能力，在图像识别、物体检测、语音识别等诸多领域均斩获了令人瞩目的成果，无疑是深度学习发展进程中的一座重要里程碑。得益于其具有尺度和平移不变性，CNN 在处理输入图像时不需要复杂的预处理步骤^[35]，因此在计算机视觉领域得到了广泛应用。CNN 处理图像的一般过程：输入层接收尺寸为 $32 \times 32 \times 3$ 的图像，经特征提取阶段，通过卷积层与池化层交替运算——首次卷积后数据变为 $28 \times 28 \times 64$ ，池化至 $4 \times 14 \times 64$ ；再次卷积得到 $10 \times 10 \times 64$ ，池化后为 $5 \times 5 \times 64$ ；最后通过全连接层（1024、512）处理，由输出层输出结果。图 2-1 呈现了完整的处理流程。

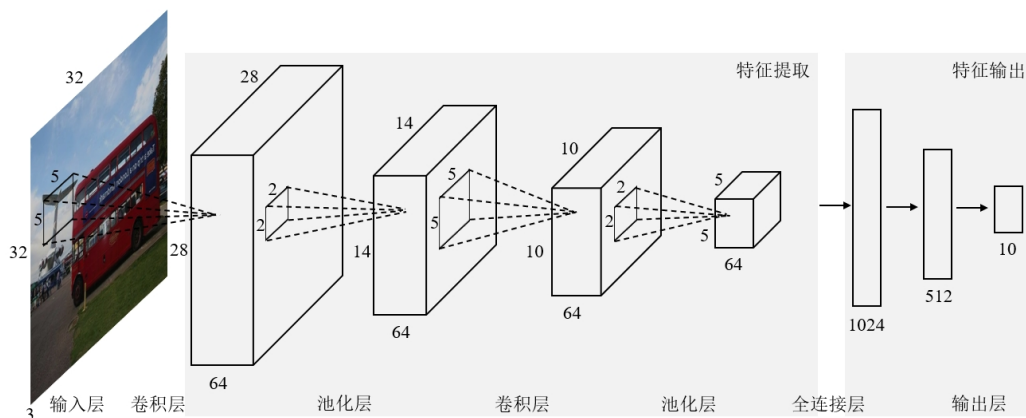


图 2-1 CNN 处理图像的一般过程示意图

CNN 是一种高效的卷积神经网络，它可以从图像中提取出边缘、纹理等空间特性；池化层通过对图像进行降采样，在保持图像重要信息的前提下，对图像进行降采样，减少图像的空间维数，减少图像中的信息损失。该方法的引入使得神经网络具有非线性表示的能力，可以学习并表示更为复杂的模式。最后，利用完整连通层对各个层次的特征进行融合，从而为最终的分类和回归问题提供决策支持。这样的结构设计使得 CNN 在从海量散乱的图像中抽取出有用的特征方面特别有优势。通过增大神经网络的深度、优化参数，卷积神经网络能够有效地提高模型的表示能力，尤其在复杂场景下具有较好的应用前景^[36]。

2.1.1 卷积层

卷积层是卷积网络的一个重要部分，它的主要作用是对输入图像进行特征提取。该算法使用滑动窗机制，通过一系列可学习的卷积核（如 3x3、5x5 尺寸）来实现前一层次的特征图的运算。卷积神经网络的性能受卷积核尺寸、步长、填充方式和输入/输出通道数目的影响。这三个参数的综合作用将直接影响到网络的整体性能，也将直接影响到网络的整体性能。通过多个参数的组合，可以有效地提取多种特征，比如，更小的卷积核更善于捕获细微的局部特性，然而，更大的卷积核更擅长地捕获整体环境信息。所以，如何合理地选取卷积网络的参数是构造有效卷积神经网络的关键。

卷积核的大小直接影响其感受野的大小，进而也决定着其所能捕获的图像特征的局部区域：卷积核越小，对局部特征越敏感；越大卷积核越能反映更广的上下文信息。步长的大小决定了卷积核遍历特征图的步数，其大小直接关系到图像的空间分辨率与

运算效率。采用填充方式对图像进行边界处理，以防止图像中的信息缺失和边界效应。零填补通常被用来使特征图的大小与输入的特性图相吻合，而有效填补在保持原图像信息的同时，也会影响到最终的特性图的大小。因此，如何正确地选取卷积核大小、步长以及填充模式，是构造高效、高性能的卷积神经网络的关键。

卷积层通过局部感受野机制实现空间特征提取，其核心在于利用可学习的卷积核在输入数据上进行滑动窗口操作。该运算方式可以有效地融合邻近像素的空间上下文信息，以准确地捕获到图像中的基本视觉要素，如纹理，边缘等。尤其是卷积层特有的参数共享机制，使得同一卷积核在不同位置的计算过程中重复使用相同权重，该方法不但减少了模型中的参数量，而且使网络具有平移不变性，增强了对目标位置变化的鲁棒性。卷积神经网络一般通过多层卷积来抽取特征，从而达到从浅层到深层的分层特征。浅层卷积层能够有效地提取出图像中的低层次特征，如边缘、纹理和颜色等；随着网络深度的增加，后续的卷积层可以提取出更多高级和抽象的特征，如形状、局部结构和目标零件等。这种层层递进的特征提取方式将为图像分类、物体检测等领域提供强有力的特征支持。因此，合理设计和调整卷积核的尺寸是提升卷积层性能的关键，能够提升模型在各种视觉任务中的表现和适应能力。图 2-2 展示了卷积操作的示意图。

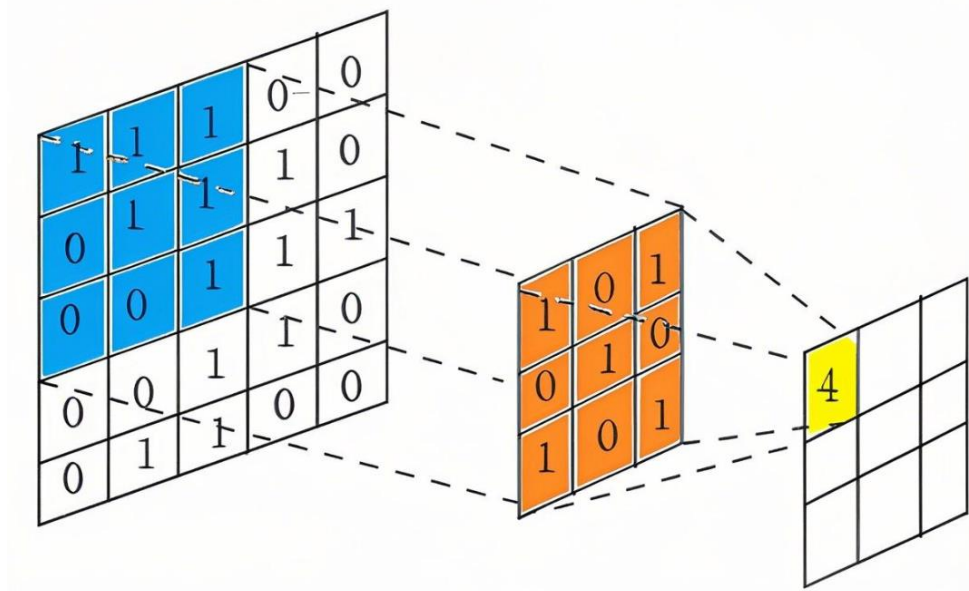


图 2-2 卷积操作示意图

2.1.2 池化层

池化层，又称下采样层，是 CNN 中至关重要的组件，通常置于连续的卷积层之间，构成了 CNN 的第二大关键部分。其核心功能在于，在卷积操作之后，通过缩减特征图的空间尺寸，来降低计算的复杂程度。具体而言，池化过程是借助移动池化窗口，并对每个窗口内的特征值进行汇总，以此实现空间维度的压缩。

在卷积神经网络中，最大池化和平均池化是两种广泛应用的方法。最大池化运算以池化窗口内的最大值作为输出；平均池化则以窗口内各像素值的平均值作为输出。这两种方法不仅能够有效地减少特征数量、参数量以及参数矩阵的维度，降低计算量，还能减少模型中的参数个数，进而避免过拟合现象的发生。同时，它们可以有效提取特征的显著值或平均值，增强特征的稳定性与抗噪性，确保图像的关键信息得以保留，不影响图像的判别与识别。

池化层的主要参数包括池化窗口尺寸（例如 2×2 、 3×3 ）和步长。这些参数的选择会直接影响特征图的大小以及信息丢失的程度。在网络架构中，池化层的输出通常会与下一个卷积层或其他池化层相连，以进行更深层次的特征提取和降维操作。合理配置池化层参数，能够有效地控制模型的复杂度，加速模型的学习过程。图 2-3 和图 2-4 分别呈现了平均池化和最大池化的示意图像。

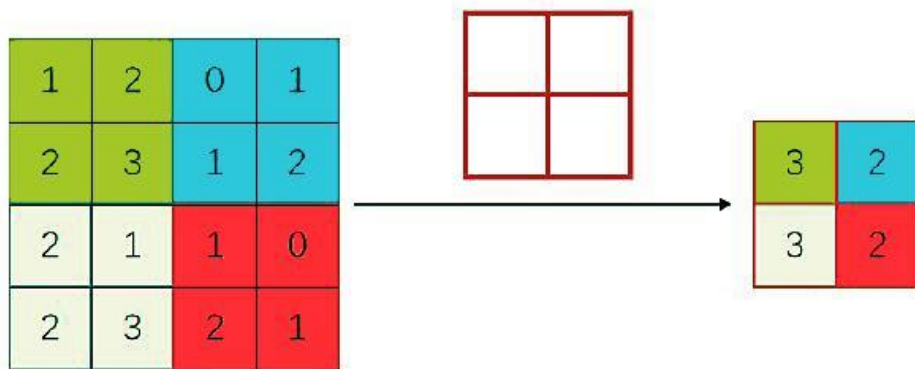


图 2-3 平均池化示意图

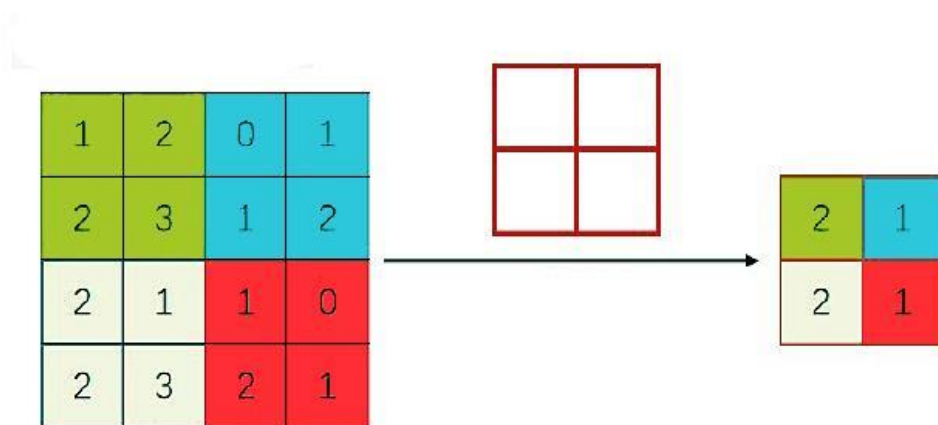


图 2-4 最大池化示意图

2.1.3 全连接层

全连接层是深度神经网络中常见的结构之一。其显著特点在于，每个输入神经元与输出神经元之间都有连接，形成了全连接的网络结构（图 2-5）。全连接层以前一层所抽取的特征矢量为输入，先做一次线性转换，将各层的权值乘以各层的权值，再加上一个偏差项。在此基础上又利用 ReLU, Sigmoid, Softmax 等非线性激活函数对其进行处理，以提高模型的表达能力。在卷积网络中，全连接层是一个典型的卷积网络，它对前一层所得到的所有特征进行融合，最后得到一个分类（如图类别）或者一个回归（如边界框坐标）。

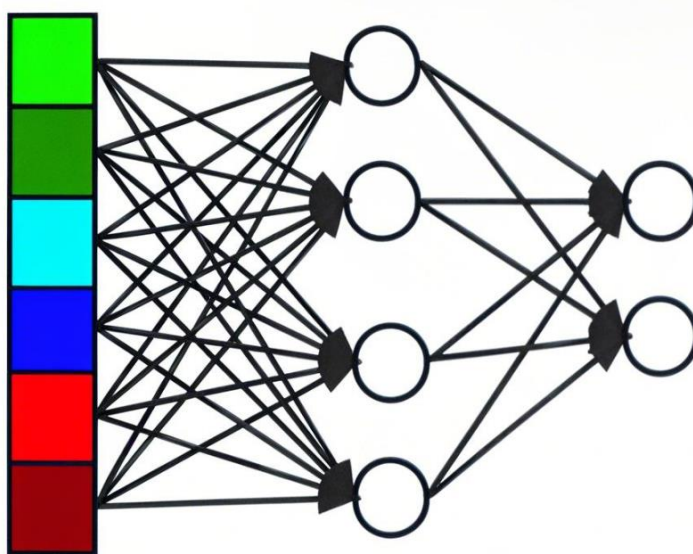


图 2-5 全连接层网络结构示意图

全连接层的主要功能是特征变换，它能够将低层次的特征通过组合与抽象转化为适应最终任务的高层次特征表示。卷积层专注于提取图像中的局部特征，而全连接层把所有的局部特征都集成起来，然后把它们转换成最后的分类或者回归结果。因此，全连接层在神经网络中起着至关重要的作用，决定了模型的最终输出。

2.1.4 激活函数

将激活函数引入到神经网络中，克服了单纯依靠线性转换所造成的局限性，可以高效地完成复杂的模式识别任务，这一转换通常在执行卷积操作之后实施。激活函数的选择对于向后续神经元传递的信息类型具有决定性影响。通过增强网络模型的学习效能和表达能力，激活函数助力神经网络高效地从庞大数据集中提炼复杂模式。

激活函数的选择直接决定了网络训练的效果和性能。非线性激活函数赋予神经网络更强的表达能力，而不同类型的激活函数对梯度的传播方式有所不同，进而影响训练速度和网络的收敛性。在卷积神经网络中，经常使用 Sigmoid 激活函数，Tanh 激活函数，ReLU 激活函数等，下面详细介绍这几个函数。

Sigmoid 激活函数是一种在深层神经网络中被广泛应用的活化函数，它被应用于 logistic 回归分析中，通过非线性转换把输入信号转换为 $(0,1)$ 区间，从而达到二分类的目的^[37]。尽管有新的激活函数出现，但 Sigmoid 在逻辑回归和二分类任务中仍具有独特优势。数学公式如 2-1 所示。函数图像见图 2-6。

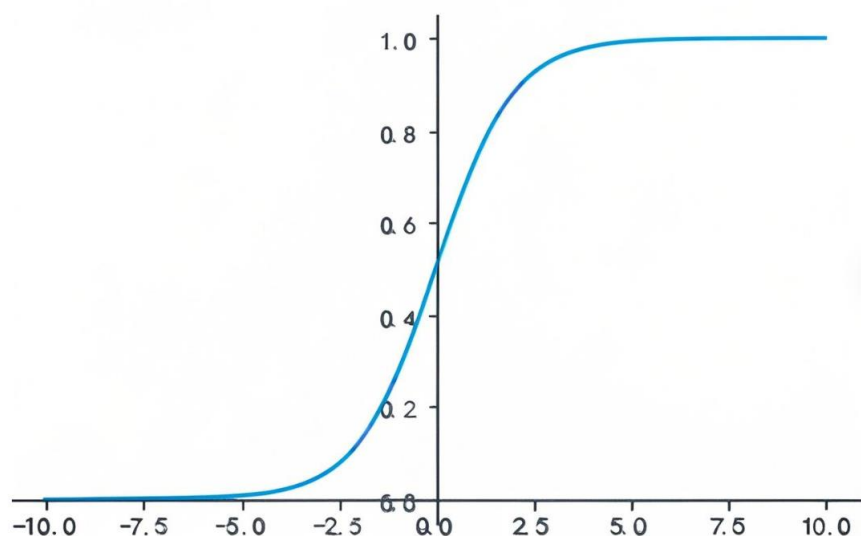


图 2-6 Sigmoid 激活函数图像

$$f(x) = \frac{1}{1 + e^{-x}} \quad (2-1)$$

与 Sigmoid 函数类似，Tanh 激活函数能够将神经^[38]网络的输入映射到一个特定的输出范围，从而在二分类任务中发挥关键作用。它不仅能够有效地捕捉输入数据的非线性特征，还能在一定程度上缓解梯度消失的问题，这对于深度神经网络的训练至关重要。因此，在二分类任务中，Tanh 激活函数同样是一个值得考虑的选择。数学公式如 2-2 所示。函数图像见图 2-7。

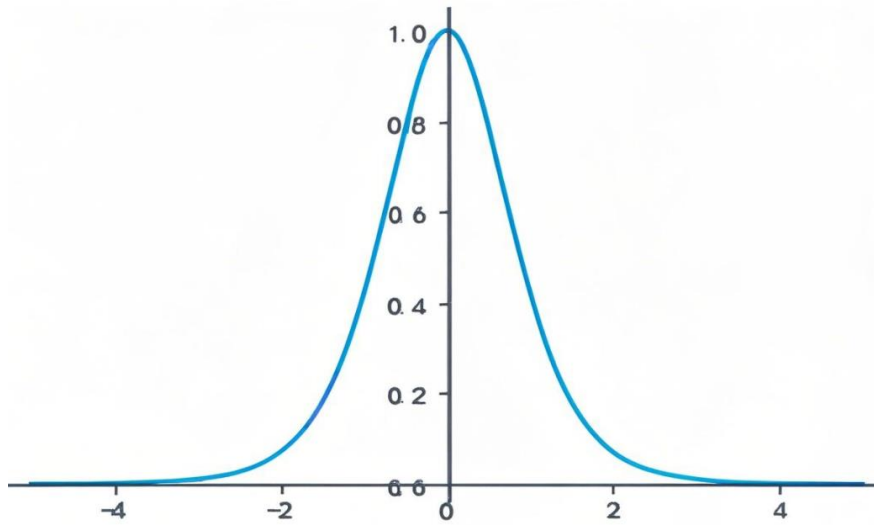


图 2-7 Tanh 激活函数图像

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2-2)$$

在卷积神经网络的回归任务中，修正线性单元 ReLU 激活函数应用广泛^[39]。它能对输入进行非线性变换，这不仅有效增强了模型的表达能力，还简化了计算过程。同时，ReLU 激活函数在一定程度上缓解了梯度消失问题，并且能够促进模型对特征的学习与选择，进而提升回归预测的精度。其数学公式如公式 2-3 所示，对应的函数图像见图 2-8。

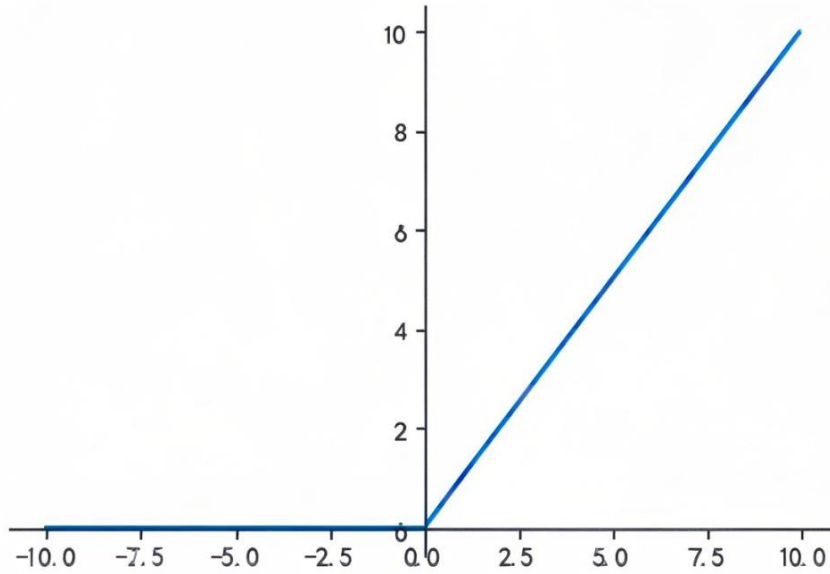


图 2-8 ReLU 激活函数图像

$$f(x) = \text{Max}(x, 0) \quad (2-3)$$

更多先进的激活函数，如 Leaky ReLU、PreLU 和 Softmax，都是在传统 Sigmoid、Tanh 和 ReLU 函数的基础上进行改良的结果。为了解决 ReLU 可能导致的神经元“死亡”问题（训练过程中某些神经元始终未被激活），Leaky ReLU 和 PreLU 提出了有效的解决方案：Leaky ReLU 通过在负输入区域引入一个小的非零斜率来缓解这一问题^[40]；而 PreLU 则进一步发展，在该斜率上引入了可学习的参数，从而有效降低了神经元死亡现象的发生^[41]。ELU 函数进一步优化了 ReLU，在负值区间采用指数衰减，使得输出接近于零均值，这一特性有助于加快神经网络的收敛速度。至于 Softmax 函数，通常适用于多分类问题的输出层，它能够将网络的原始输出转换为概率分布形式，并确保所有输出值的总和严格等于 1^[42]。

2.2 目标检测算法概述

目标检测算法的主要任务是识别图像或视频中的物体，并确定其精确位置，这项技术已广泛应用于自动驾驶、视频监控等多个领域^[43]。以前的目标检测技术多依靠人工设计的特征（如 Haar、HOG 等）及分类器，其识别效果受到特征工程的影响^[44]。随着目标检测的应用场景越来越复杂，数据量越来越大，传统的检测方法越来越显示

出它的不足，很难满足高精度、高效率的检测要求。深度学习的出现给目标检测方式带来了一场革命，以深度学习为基础的目标检测方法因其强大的特征学习能力而快速成为当前研究的热点，并获得了明显的性能提升。通过监督、半监督或无监督学习方式，深度学习算法能够从海量数据中自动提取潜在特征和模式，逐步成为目标检测技术的新兴发展方向。与传统方法依赖人工设计特征不同，深度学习可以直接从原始图像中提取出更具辨识度和层次化的特征，因此在处理复杂图像及提升鲁棒性方面具有明显的优势。

基于深度学习的目标检测方法可分为两类，基于候选区域的两阶段检测方法和基于回归的单阶段检测算法。两阶段方法如 R-CNN 方法，首先生成候选区域，然后对这些区域进行分类和定位；而单阶段检测算法，如 YOLO 和 SSD，则通过端到端的卷积神经网络直接输出目标类别和位置，避免了生成候选区域的步骤，从而提供了更高的精度与速度^{[45][46]}。尤其是在实时检测任务中，YOLO 表现最为突出。近年来，基于 Transformer 的模型进一步推动了目标检测领域的进步，通过引入自注意力机制，显著提升了复杂场景下的检测性能。

尽管目标检测算法已取得显著进展，但在实际应用中仍面临诸如尺度变化、遮挡以及背景复杂度等挑战。因此，如何在提高检测精度与速度之间找到合理的平衡，依然是当前研究的热点问题。本章将详细探讨深度学习在目标检测中的应用，并简要分析为何选择 YOLOv8 作为植物叶表气孔检测的核心算法。

2.2.1 传统目标检测算法

尽管目标检测领域已迈入深度学习的全新篇章，但其基石仍深深植根于传统目标检测的智慧之中。本章将阐述传统目标检测算法的基本架构。

传统目标检测流程主要分为三步：区域候选生成，采用多尺度滑动窗口法对输入图像进行密集遍历扫描，通过穷举式搜索生成潜在 ROI；特征工程构建，依赖人工设计的尺度不变特征变换（SIFT，Scale-Invariant Feature Transform）、局部二值模式（LBP，Local Binary Pattern）以及 HOG 等手工特征将 ROI 转化为特征向量；分类决策输出，基于支持向量机（SVM，Support Vector Machine）、或可变形部件模型

(DPM, Deformable Part Model) 等传统机器学习算法, 对特征向量进行类别判定类别归属, 通常需结合级联分类器等策略提升精度。算法流程如图 2-9 所示。



图 2-9 传统目标检测流程

传统机器学习算法在多个方面显示出其局限性, 导致其逐渐被深度学习检测算法超越。主要存在三方面不足: 区域选择阶段需用多尺度滑动窗口反复遍历图像, 导致冗余计算和速度下降, 且预设窗口尺寸难以适应多样化目标; 特征提取阶段仅能捕获底层特征, 表达能力弱且高度依赖具体任务, 复用性差; 检测流程被分割为独立阶段 (区域选择、特征提取、分类), 难以全局优化且过度依赖设计者经验, 导致算法适应性差。综上所述, 传统机器视觉算法在效率、灵活性及通用性方面的不足, 促使其逐渐被深度学习方法所替代。

2.2.2 两阶段目标检测算法

二阶段目标检测算法的核心理念在于, 先从输入图像里精准筛选出极有可能包含目标的候选区域, 紧接着针对这些区域开展目标分类与精确定位工作。该算法解决了传统目标检测算法依赖手工设计特征、滑动窗口计算效率低下等难题, 经典的 R-CNN 和 Faster R-CNN, 均属于这一类型的算法。

1. R-CNN

R-CNN 是深度学习在目标检测领域的首次突破, 其出现使目标检测的准确率提升了至少 30%。在将深度学习引入到目标检测中以前, 针对 PASCAL VOC 等常规对象的检测精度一直保持在 30% 左右, 很难有突破。Girshick 等人率先将 CNN 用于目标检测, 并在物体目标检测中实现了里程碑式的突破^[47]。其算法结构如图 2-10 所示, 主要包括以下几个步骤: 首先, 使用选择性搜索算法分析图像中的颜色、纹理等特征, 生成大约 2,000 个不同大小的候选区域; 接着, 对这些候选区域进行标准化处理, 统一尺寸, 以确保输入神经网络的区域大小一致, 并提取相应特征; 然后将各特征矢量输入到一个单独的 SVM 分类器中, 来判定这个区域是否为某一类。最后, 采用单一回归算法对各候选区域进行定位细化, 从而提高了边界框的定位精度。

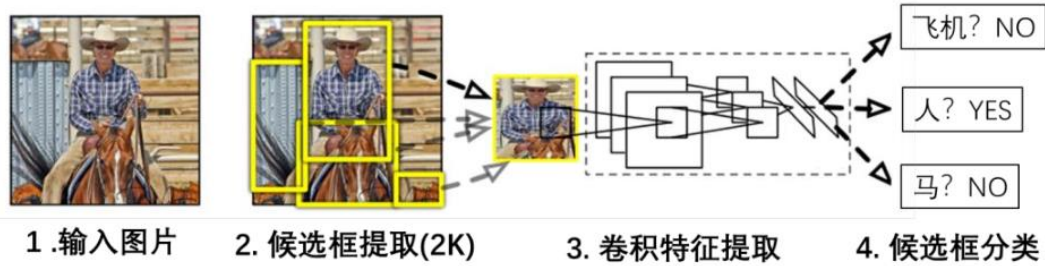


图 2-10 R-CNN 模型结

与传统的机器视觉方法相比，R-CNN 在如下三个方面具有创新性：一是利用选择性搜索算法产生候选区域，替代传统的滑动窗策略，从而大大提升了检测效率；二是 R-CNN 采用卷积神经网络对目标区域进行特征提取，突破传统人工设计特征的限制，使特征表达更加具有鉴别性，提高识别率；三是将 R-CNN 与支持向量机相结合，采用回归算法对候选框进行微调，以减少定位误差，提高检测精度。

R-CNN 的提出使目标检测技术的发展迅速，搭载该技术的系统的性能得到了极大的提升。然而，R-CNN 仍然存在一些问题，包括预测速度慢、训练过程复杂以及存储需求大。预测速度慢是由于候选框重叠导致冗余计算，训练过程缓慢的原因是需分别训练图像分类网络、SVM 分类器和回归器，存储需求大是因为每个候选框的特征需要单独提取并保存，尤其在处理深度网络时，可能需要数百 GB 的存储空间。

为了解决上述问题，后续的研究者们在 R-CNN 基础上提出了多个改进版本，如 SPPNet^[48] 和 Fast R-CNN，这些算法优化大幅提高深度学习算法的计算效率与性能，为实用化奠定基础，并有望在产业中推广应用，促进多个领域的智能化发展。

2. Faster R-CNN

FasterR-CNN 是具有代表性的两阶目标检测方法^[49]。区别于 R-CNN 采用选择性搜索方法产生候选区，FasterR-CNN 通过引入区域提议网络（RPN，Region Proposal Network）来自动产生候选点。RPN 算法通过对不同尺度、尺度的锚框进行预测，从而有效地减少候选区域数目，大幅降低运算复杂性。在此基础上 FasterR-CNN 又将候选区域产生、特征提取、目标分类、边界框回归等一系列环节有机地结合在一个端到端框架中，以简化处理过程，提升检测效率与准确性。这一点是 FasterR-CNN 的核心技术，也是它在物体探测方面具有里程碑意义的成果。Faster R-CNN 通过三阶段递进式架构实现高效目标检测：首先利用卷积神经网络构建特征金字塔，提取多尺度图像

特征；接着通过 RPN 实施多尺度锚框机制生成高质量候选区域，并采用非极大值抑制（NMS，Non-Maximum Suppression）筛选约 2000 个候选框（对应 VOC2012 数据集）映射到特征图；最后通过 ROI 池化层统一特征尺寸，经全连接层进行特征压缩与语义增强，输出目标类别概率和边界框修正量。该流程通过特征金字塔增强多尺度适应性，锚框机制提升候选区域生成效率，ROI 池化实现特征对齐，在 COCO 数据集上典型模型达到 66.9% 的 mAP 指标和 17 FPS 推理速度，有效平衡了检测精度与效率。图 2-11 展示了 Faster R-CNN 模型结构。

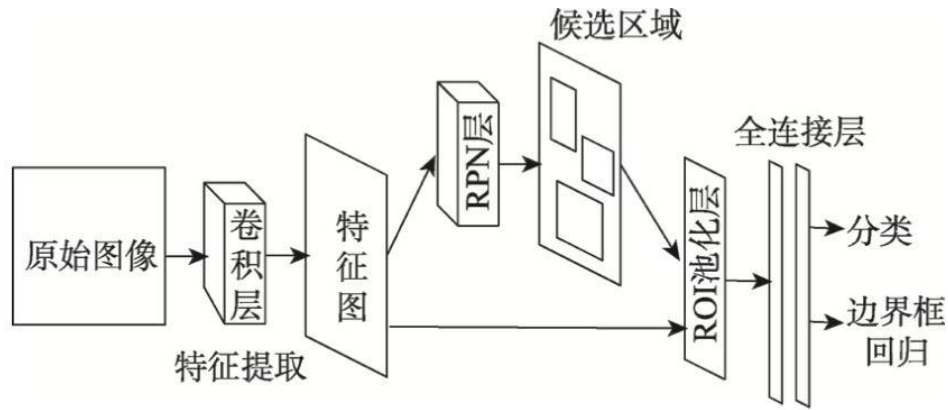


图 2-11 Faster R-CNN 模型结构

继 Faster R-CNN 之后，出现了如 R-FCN、Cascade R-CNN 和 Libra R-CNN 等一系列算法。尽管这些二阶段算法在精度上不断提升，但随着网络结构的复杂度增加，计算量也随之增大，导致检测速度的下降，挑战了实时应用的需求。

2.2.3 单阶段目标检测算法

在目标检测领域，双阶段目标仍存在着难以忽视的缺陷。双阶段目标检测模型在特征提取与处理过程中，不可避免地会出现空间特征损失的问题。这极有可能导致最终检测精度大打折扣，使得检测结果与实际情况存在较大偏差。更为突出的是，双阶段目标检测模型有着高昂的计算开销。这在一些对时间要求极高的应用场景中，如自动驾驶、安防监控等，无疑是一个致命的短板。

为了突破双阶段目标检测模型的这些局限性，研究人员将目光投向了单阶段检测框架，将其视为更优的替代方案。在众多单阶段算法中，YOLO 和 SSD 是其中的代表。与 SSD 等算法相比，YOLO 方法更是有着独特的优势，它能够高效地利用计算

资源，以极快的速度完成目标检测任务，避免了双阶段模型的空间特征损失问题，从而显著提升检测精度，更好地适应实时检测的需求。鉴于本研究主要聚焦于 YOLO 模型，接下来将深入介绍 YOLO 家族算法的演变与发展，探寻其不断优化、超越双阶段检测算法的历程。

1. YOLOv1

YOLOv1 是 Joseph Redmon 等人在2016年提出的一种实时目标检测算法^[45]。该方法把目标检测看作是一种简单的回归问题，通过对输入图像的边界框和分类概率进行预测。它的主干网络结构借鉴了 GoogleNet 的设计思路，但并未使用 Inception 模块，而是通过堆叠 1×1 和 3×3 的卷积层来实现特征提取。图 2-12 展示了YOLOv1处理图像的流程：YOLOv1 网络接受一张尺寸为 448×448 的 RGB 图像作为输入，经过主干网络处理后，生成一个空间分辨率降低 64 倍的特征图。该特征图是一个 $7\times 7\times 1024$ 的三维张量，其中 7×7 表示特征图的宽度和高度，1024 是经过卷积操作后的通道数；随后，特征图通过全连接层和维度变换等操作；最终输出一个 $7\times 7\times 30$ 的一维张量。

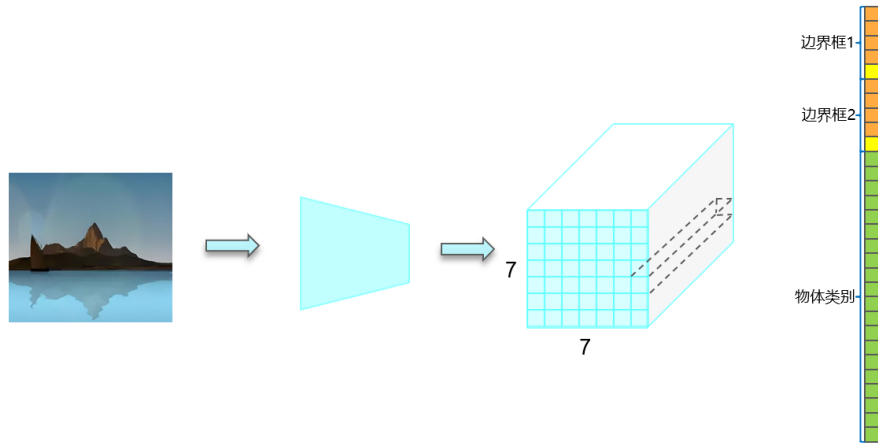


图 2-12 YOLOv1 处理图像流程图

YOLOv1 划分后的 7×7 的网格单元，对于每个单元承担边界框预测和类别预测。边界框预测是每个网格预测包含 2 个边界框，每个框包含四个坐标参数和置信度，共 10 个参数。类别预测是每个网格预测 20 类别的概率分布（对应 VOC2012 数据集），共 20 个参数。最终每个网格单元输出 30 维预测向量，因此整个网络输出张量维度为 $7\times 7\times 30$ 。图 2-13 展示了 YOLOv1 的网络结构。

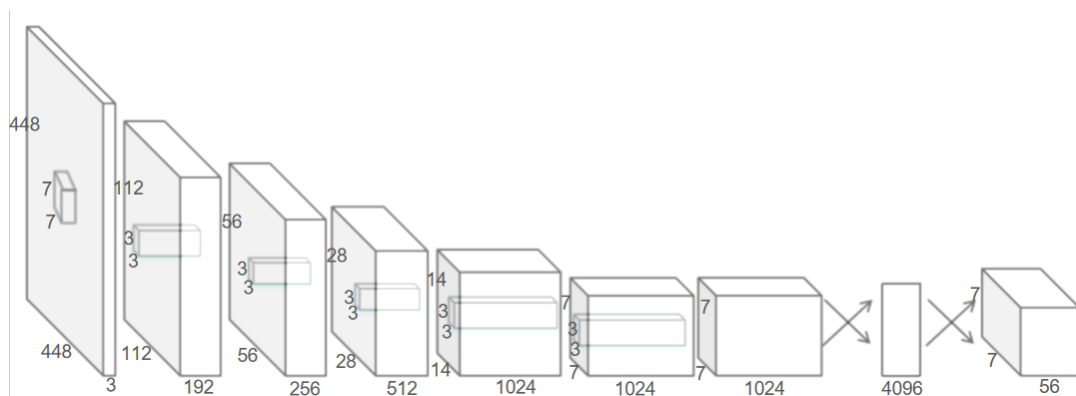


图 2-13 YOLOv1 网络结构

2. YOLOv2-YOLOv7

YOLOv1 作为开创性的单阶段目标检测算法，虽然实现了实时检测，但存在着明显的技术局限：由于采用全连接层直接预测边界框，导致空间敏感性不足；每个网格单元仅预测两个边界框且共享类别概率，严重制约了对密集目标的检测；其损失函数将不同大小的边界框等同对待，造成定位精度不理想，尤其是在小目标检测方面表现欠佳；此外，简单的特征提取结构也限制了模型的表达能力。

YOLOv2 针对上一代模型的短板进行了修复，实现了多方面的显著改进^[50]。在检测策略上，引入锚框机制，预先设置多种形状的检测框模板，优化了目标检测流程。网络核心采用轻量高效的 Darknet-19 主干网络，并加入批量归一化的数据标准化处理层，有效提升了特征识别能力。为使检测框更贴合实际目标，利用 k-means 聚类算法对常见物体尺寸进行统计分析，自动生成适配的检测框模板。训练模式更加灵活，支持多尺度训练，可在训练过程中动态调整图片大小，从而增强模型的适应能力。通过直通层进行特征拼接，将浅层特征融合，提升了模型对细节的捕捉能力。定位算法改用相对坐标预测，以相对位置偏移量进行预测，使检测框定位更加准确。这些改进使得 YOLOv2 在保持快速检测的同时，精度也得到大幅提升，极大地增强了检测系统的性能。

YOLOv3 是目标检测技术的一次全面升级^[51]。它更换神经网络主干为 Darknet-53，并采用残差连接设计，使网络能更高效清晰地抓取图像特征，提升检测速度与精准度。借助特征金字塔结构，可在三个不同尺寸的图片层级上检测目标。采用逻辑回归分类器，能同时为物体赋予多个标签。通过上采样和跳跃连接的特征融合方式，整合图像

细节与整体信息，并优化检测框位置，实现精准定位。这些创新使 YOLOv3 既保持快速检测速度，又能应对不同大小的目标。

YOLOv4 进行了一系列具体优化。在主干网络方面，采用跨阶段局部网络 Darknet53 作为核心架构，搭配平滑非线性的 Mish 激活函数，增强了特征表达能力，使模型能够更精准地提取数据中的关键特征^[52]。在特征融合结构上，在颈部网络中嵌入了可扩大感受野的空间金字塔池化模块以及能实现多尺度特征融合的路径聚合网络。通过扩大感受野，模型能够关注到更大范围的信息，多尺度特征融合则有效提升了模型对细节的捕捉能力，让模型在不同尺度下都能准确识别目标。此外，还引入了自动调整预设检测框尺寸的自适应锚框选择方法以及防止过拟合的正则化技术，前者让检测框更好地适应不同目标的尺寸，后者保障了模型在训练过程中不会出现过拟合现象，提高了模型的泛化能力。这些先进技术的应用，让 YOLOv4 如同装配了多级变速引擎和智能调焦系统，不仅保持了极快的检测速度，而且显著提升了在复杂场景下的识别准确率，为目标检测领域带来了更高效、更精准的解决方案。

YOLOv5 虽然没有正式发布相关论文，但在工程应用方面具有显著优势。在高效特征提取方面，借助图像切片重组模块，其原理类似网格切分重组，提升了特征抓取效率。同时搭配跨阶段瓶颈结构，减少了重复计算，优化了资源利用，还能实现自适应特征融合，即动态调整特征组合。在智能训练策略上，它具备自动学习预设检测框尺寸的能力。可以动态调整输入图片尺寸，也就是自适应缩放，避免了无效填充。支持模型量化以压缩模型大小，以及剪枝技术来删除冗余参数等轻量化技术。这些改进使得 YOLOv5 如同一个万能工具箱，既能快速适配手机、服务器等各类硬件，又能保持快速、准确且节省资源的检测特性，因而成为工业落地应用的热门选择。

YOLOv6 聚焦于性能的深度优化，在多个方面实现了核心改进。在高效结构设计上，采用解耦检测头，其为分离式结构，简化了计算流程，有效降低了推理耗时。智能训练策略方面：提出了任务对齐学习机制，能够动态匹配预测与真实框；使用自适应标签分配方法，动态调整训练目标标注策略，提升了学习效率；加入辅助监督信号，以多阶段学习引导的方式强化了模型的训练效果。这些改进使得 YOLOv6 既能够在 GPU 服务器上实现快速检测，也能在手机、嵌入式设备等资源受限的环境中保持灵活高效的检测性能。

YOLOv7 在架构设计引入 E-ELAN 模块增强特征提取能力, 采用计算块重参数化策略提升模型性能, 通过梯度流重新设计优化训练过程。在检测头方面实现了辅助头协同训练和基于 IoU 的动态标签分配策略, 同时优化了模型缩放策略, 提出复合缩放方法, 并引入动态训练策略调整机制, 这些创新使 YOLOv7 在保持高效推理速度的同时, 达到了业界领先的检测精度^[53]。

3. YOLOv8

YOLOv8 是 Ultralytics 公司的重要里程碑作品, YOLOv8 在 YOLO 系列算法的既有成果之上, 开展了算法集成方面的创新性工作, 在工程实践中具有突出价值。在图像分类、目标检测和实例分割等任务上的性能表现稳定, 且具有完善的技术文档和大量实践案例支持, 这使其成为可靠的研究基础。其架构设计注重可扩展性, 通过兼容历史版本的特性, 为研究人员提供了便捷的实验对比环境。

YOLOv8 所运用的网络结构, 主要由输入端 (Input)、骨干网络 (Backbone)、特征融合模块 (Neck) 以及预测模块 (Head) 这几个核心部分共同搭建而成, 其网络结构详见图 2-14。在骨干网络的构建上, 整合了 CBS、C2F 等模块。其中, CBS 模块的核心思想在于, 借助跨阶段的连接以及独特的设计, 有效降低模型的计算量, 同时增强网络的信息流通性与特征提取能力。通过这种跨阶段的连接方式, 能够使不同阶段的特征信息得到更充分的交互与融合, 从而让模型在处理复杂数据时更加高效。C2F 模块作为一种特殊的卷积结构, 巧妙地运用多个卷积层和池化层进行特征提取。这种设计不仅增强了特征融合能力, 而且显著提升了计算效率。多个卷积层的组合能够从不同尺度和维度对图像特征进行捕捉, 而池化层则可以在减少数据量的同时, 保留关键的特征信息, 使得模型在处理大规模数据时能够更加迅速且准确。

模型的颈部设计借鉴了 YOLOv7 的 ELAN 结构思想, 采用 C2F 结构进一步丰富了梯度信息的流动。这种结构设计能够让梯度在网络中更加顺畅地传播, 有效地克服了梯度消失、爆破等问题, 提高了学习的稳定性, 加快了算法的收敛速度。

在检测头部分, YOLOv8 进行了大胆的创新。它摒弃了传统的整体式设计, 转而采用分离式结构分别处理分类和检测任务。这种分离式的设计使得分类和检测任务能够更加专注地进行处理, 避免了相互之间的干扰, 提高了任务处理的准确性。特别值得一提的是, YOLOv8 放弃了基于锚框的检测方式, 转而采用无锚框策略。这一转变

意义重大，它减少了模型对预设参数的依赖，使得模型能够更好地适应不同场景下的目标检测任务，大大提高了检测的适应性和鲁棒性，能够在复杂多变的环境中更准确地识别和定位目标。

选择 YOLOv8 作为本研究的基础框架主要基于以下考虑：首先，其完善的文档和丰富的社区资源有助于算法的快速实现和问题排查；其次，该版本经过大量实际应用验证，性能稳定可靠。相比之下，最新版本虽然在某些特定场景下可能具有性能优势，但考虑到实际部署的稳定性需求和现有硬件环境的约束，YOLOv8 是更为合适的选择。

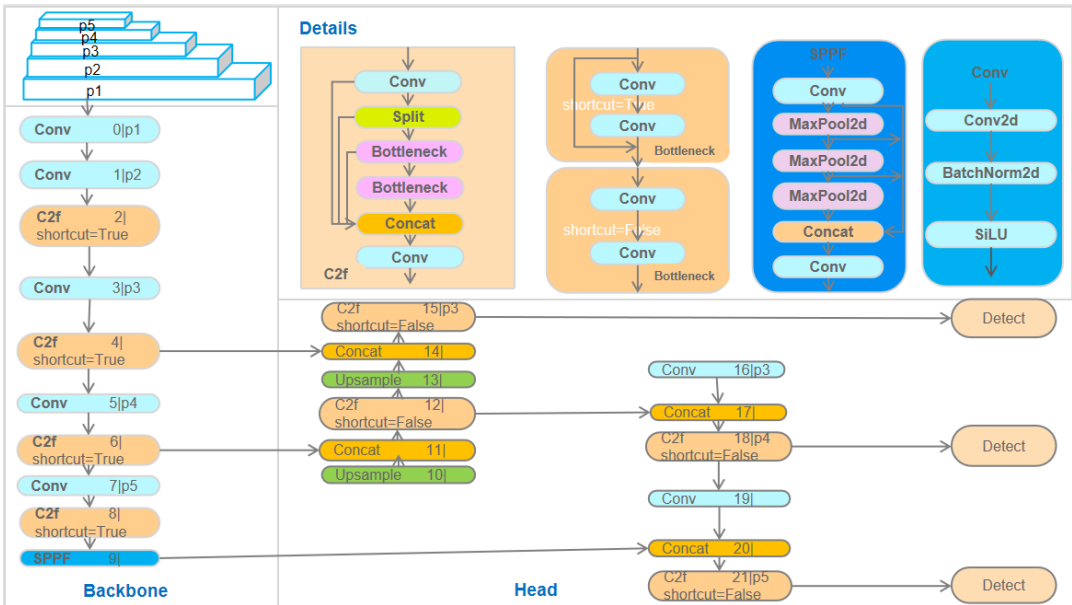


图 2-14 YOLOv8 网络结构图

2.3 本章小结

本章系统梳理了深度学习在目标检测领域的发展脉络。在阐述卷积神经网络基础架构的基础上，分类探讨了目标检测的主要技术路线：传统检测方法、双阶段检测框架和单阶段检测算法。其中，对 R-CNN 和 Faster R-CNN 等双阶段模型的分析揭示了其高精度但速度受限的特点。随后重点聚焦于 YOLO 系列的技术演进，深入分析了其优劣势及迭代改进过程。这些理论研究为构建植物叶表气孔检测系统提供了技术支撑。

第三章 植物叶表气孔数据集融合与预处理

3.1 数据采集

本研究通过整合华东师范大学生命科学学院自主构建的 sto-Du 植物叶表气孔数据集（以下简称自建数据集）与公开数据集，充分发挥二者的优势，提升研究的可靠性与准确性。公开数据集以其数据规模庞大、来源广泛的显著特点，成为了模型训练的重要基石。它涵盖了丰富多样的物种类别，能够为模型提供广泛的样本基础，有助于模型学习到更为全面的特征。然而，公开数据集也存在一定的局限性，由于其数据采集的分散性和开放性，采集设备的参数往往记录不完整。本研究主要将公开数据应用于模型训练阶段，让模型在大量的数据中学习不同物种气孔的普遍特征。

与之相对，自建数据集虽然在数据量和物种丰富度上不及公开数据集，但其具有独特的优势。本研究在自建数据集的采集过程中，确保拍摄设备的各项参数都被完整记录。这些数据主要用于后续模型校验环节，数据缺失将无法开展有效检验。由于公开数据集往往难以提供这些关键参数的完整信息，因此记录拍摄设备的完整参数对于本研究至关重要。

考虑到实验的完整性和模型训练的充分性，对自建数据集进行补充显得十分必要。在构建自建数据集时，本研究采用多模态数据采集方案，集成了四种不同型号的成像设备。这种多样化的设备组合，涵盖了实验室高精度显微镜以及野外便携设备等，旨在满足多种应用场景的需求。通过多设备协同作业，采集到的样本更具多样性，为提升模型的泛化能力提供了坚实的数据基础。具体的设备参数详见表 3-1，其中包含光学分辨率、放大倍数、传感器尺寸等关键指标，这些参数是获取气孔特征绝对测量值的重要依据。自建数据集包含了 72 个物种，共采集 260 张图像，经过严格筛选，最终保留 165 张。图 3-1 展示了其中天南星科物种的图像。

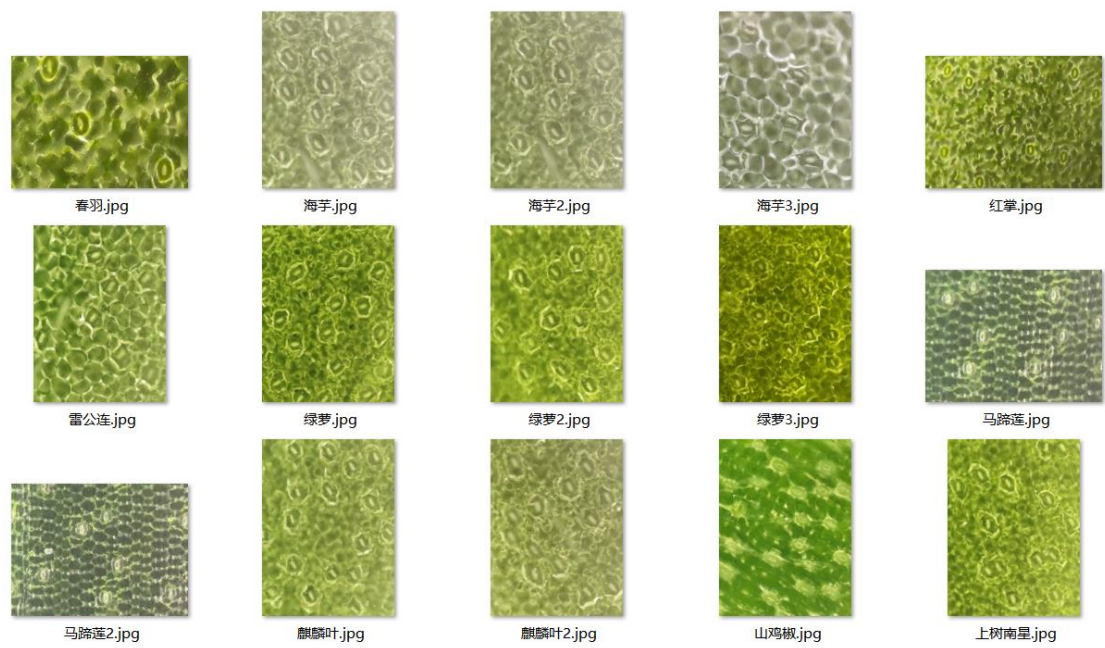


图 3-1 天南星科物种图像示意图

表 3-1 图像采集设备及主要参数

设备名称	传 感 器 分 数 尺 寸/inch	传感器毫米 尺寸/mm	图 像 分 辨 率/px	放大倍数	是否显示 倍率-像素 尺寸
Redmi note 13					
Pro + 外接放大物镜	1/1.4	13.56×10.17	3060×4080	40-400	N
Eclipse Ti2 倒置显微镜	1/1.4	36.0×23.9	2736×1824	20-400	Y
拜斯特数码显微镜	1/3	4.8 × 3.6	1920×1080	40-1000	N
Anyty 手持显微镜	1/3	30.0 × 18.0	1280×1024	40-1000	N

本研究的公开数据集来源采集自 Fetter 开展的研究^[54]，通过多维度筛选方法（包括气孔形态完整性、图像清晰度和样本多样性），从原始数据集中提取包含 500 张具有代表性气孔图像的子集。筛选了数据集中的有代表性的气孔图像作为本研究的公开数据集，图 3-2 示展示了采集的数据集中按科分类的 20 个科的样本分布情况。

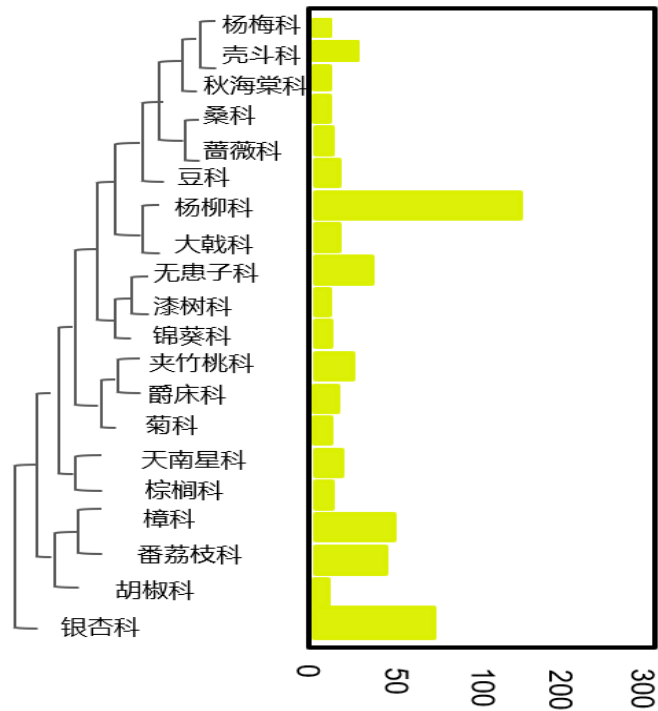


图 3-2 选取的公开数据集中所属科目和样本数量

本研究自建的数据集融合了破坏性与非破坏性两种采集方法，目的是确保数据的全面性和代表性。所有数据均由专业科研人员按照严格的科学流程收集，确保数据的准确性、科学性和可信性。下面将详细介绍这两种方法采集数据的过程。

非破坏性方法以其能够完整保持叶片原始状态的优势，成为连续观测和动态研究的理想选择。目前，常见的非破坏性采样方法主要采用手持设备拍摄法。这种方法在不破坏植物叶片形态的前提下，可轻松获取气孔图像。此外，手持显微观测技术不仅能够清晰记录叶片表面特征，还支持视频录制功能，从而精准捕捉气孔开闭的动态过程。具体操作时，通过调整手持设备的拍摄角度和放大倍数，可以捕获植物叶表气孔。如果需要观察同一区域气孔的连续开闭情况，可以采用固定设备的拍摄方式；如果需要观测不同区域气孔在同一时刻的开闭状态，则可通过轻移设备实现多点采样。在记录气孔信息的同时，还需保存拍摄设备的参数（如拍摄时间、放大倍数等），为后续分析提供数据支持。图 3-3 展示了手持显微设备拍摄图像的一般流程。

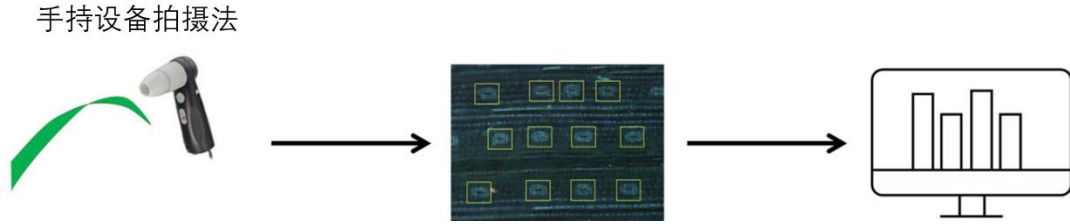


图 3-3 手持显微设备拍摄法流程图

破坏性方法虽然会对叶片造成一定程度的损伤，但通过去除色素或叶绿体，能够显著突出气孔结构，减少图像中的干扰因素。最常用的破坏性方法是叶片印记法。该方法作为一种常见的破坏性采样技术，被广泛应用于气孔研究。首先在叶片表面施加硅胶、牙科树脂或指甲油等透明介质，待其固化后巧妙剥离，获得记录气孔形态的薄膜。具体操作如下：首先，将采集的叶片样品放置在 FAA 固定液（5 ml 38% 福尔马林，5 ml 冰醋酸，90 ml 50% 乙醇，5 ml 甘油）中带回实验室。随后，随机选择 5 个叶片，将其从 FAA 溶液中取出后置于通风橱风干，并使用刀片轻轻刮去叶片表面的绒毛。接着，在叶片背面随机选择 4-6 个部位涂抹透明指甲油，待指甲油风干后，用镊子剥离印模，制作成盖玻片。最后，使用 Eclipse Ti2 倒置显微镜观察盖玻片，并在每个叶片的不同部位分别拍摄三次，记录气孔数量。操作流程如图 3-4 所示。



图 3-4 叶片印记法的流程图

3.2 数据增强

本研究选用 YOLOv8 作为基础检测模型，该模型具备强大的数据增强功能，自带多种数据增强模式。在默认配置下，YOLOv8 就启用了丰富多样的数据增强技术，这些技术彼此协同配合，从多个维度对训练数据进行拓展和优化。无论是对图像的几何变换，还是对色彩、亮度等属性的调整，都使得训练数据更加丰富多样。接下来，将对 YOLOv8 的数据增强技术展开详细介绍。

首先，YOLOv8 采用了一些基础的数据增强技术。随机水平翻转是其中一种简单而有效的方法，它以 50% 的概率对图像进行水平翻转。这一操作模拟了目标在现实场景中可能出现的不同视角，使模型在训练过程中能够学习到目标的左右对称特征，从而增强对不同方向目标的识别能力。

色域变换也是基础增强技术的重要组成部分。它会随机调整图像的亮度、对比度和饱和度。在实际应用中，目标所处的光照条件千变万化，不同的光照会导致图像的色彩和亮度产生显著差异。通过色域变换，模型能够在训练过程中接触到各种光照条件下的图像，进而提升对光照变化的适应性。

其次，随机仿射变换同样不可或缺，它包括平移、旋转和缩放等操作。在保持目标比例的同时引入几何变形，使得模型能够学习到目标在不同位置、角度和大小下的特征。在实际场景中，目标的位置和姿态可能各不相同，通过随机仿射变换，模型可以更好地适应这些变化，提高检测的准确性和鲁棒性。

最后，除了基础增强技术，YOLOv8 还运用了复合增强技术。Mosaic 增强是一种独特的方法，它将 4 张图像随机拼接成一张新图像。这种方式大大增加了训练数据的多样性，模拟了复杂的场景，让模型在训练过程中能够同时处理多个目标和不同的背景信息。MixUp 增强则是按一定比例混合两张图像及其标注。在实际应用中，可能会存在一些模糊的目标或者目标之间存在重叠的情况，MixUp 增强通过混合图像和标注，让模型学习到更模糊的标签信息，提升了模型对这类复杂情况的处理能力。

YOLOv8 还具备动态调整机制。自适应增强参数是其重要特点之一，它会根据训练阶段动态调整增强强度。在训练初期，模型对数据的特征了解较少，此时使用更强的几何变换可以让模型快速学习到更丰富的特征。而在训练后期，为了避免过拟合，会逐步减弱增强强度，使模型更加关注数据的本质特征。自动锚框匹配则结合数据增强后的图像尺寸调整锚框，优化匹配效率，确保模型能够更准确地定位目标。图 3-5 展示了 YOLOv8 数据增强的全流程。通过系统实施前述数据增强策略，在实现 1.8 倍有效样本规模提升（原始数据基数：N=665）的同时，显著优化了数据表征能力，为后续的准确识别奠定了坚实的基础。

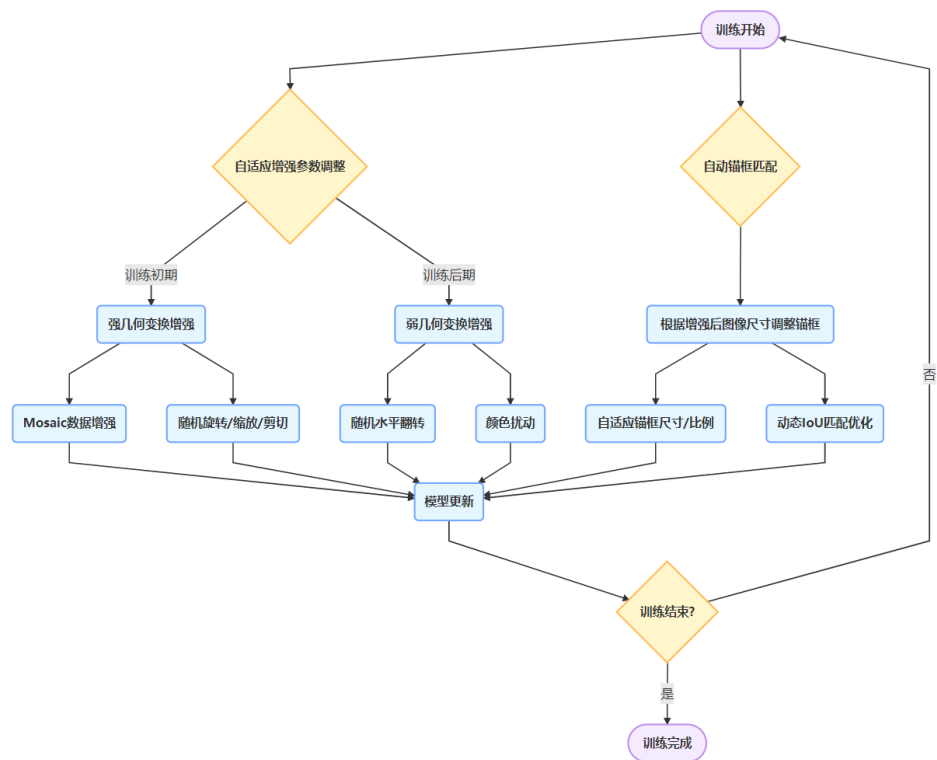


图 3-5 YOLOv8 默认的数据增强流程图

3.3 数据标注

在目标检测模型的训练流程中，获取图像标注信息是模型训练得以有效开展的基石。标注数据需要精确无误地描述目标物体的位置与类别，为模型提供准确的学习样本。本研究选用开源工具 **labelme** 标注图像。**labelme** 以 Python 为开发语言，借助 Qt 框架构建用户界面，能够胜任高精度标注任务。**labelme** 优势显著：

1. 标注形式丰富。**labelme** 支持各种类型的标记，如矩形，多边形，直线等，可以完全涵盖物体检测，语义分割等不同的任务需要。以目标检测任务为例，可利用矩形框快速定位物体；在语义分割任务中，则能通过多边形精准勾勒物体轮廓。
2. 集成 AI 预测模型。该工具集成了先进的 AI 预测模型，用户只需进行简单的点击操作，即可快速生成初始标注框或分割区域。后续，用户通过手动微调，就能完成高精度的标注工作。这一特性极大地提升了标注效率，相较于传统标注方式，可将标注效率提高 3-5 倍，特别适合大规模数据集的制作。

3. 多格式导出支持: labelme 支持导出 Pascal VOC^[55]、COCO^[56]、YOLO 等主流数据格式, 能够满足不同模型的训练需求。例如, 导出的 JSON 文件可直接应用于语义分割任务, 而 VOC 格式则与传统目标检测框架完美适配, 为模型训练提供了便利。
4. 标注规范一致性: labelme 内置了标注规范模板, 涵盖类别定义、颜色编码规则等内容。在团队协作的场景下, 通过这些模板可确保标注的一致性。例如, 用户可以自定义 JSON 文件, 明确类别 ID 与名称的映射关系, 有效避免人工标注过程中可能出现的类别混淆问题。

相较于 LabelImg 等工具, labelme 在标注维度 (如支持分割任务) 和交互体验 (如画刷标注) 上更具优势。例如, 在标注复杂形状物体时, labelme 的多边形工具可精确捕捉轮廓细节, 而 LabelImg 需依赖手动绘制坐标点。labelme 凭借其丰富的功能和出色的性能, 成为目标检测模型训练中图像标注的理想选择。

为区分气孔的特征, 本研究定义了两类标签: 将被保卫细胞完整包围的气孔标签设置为 who_sto; 独立存在且无附加结构的气孔标签设置为 sto。通过 labelme 的多边形工具和 AI 工具勾勒气孔轮廓, 标注结果以红色 (who_sto) 和绿色 (sto) 在原图上高亮显示, 类别信息实时呈现在界面右侧属性面板, 如图 3-6 所示。

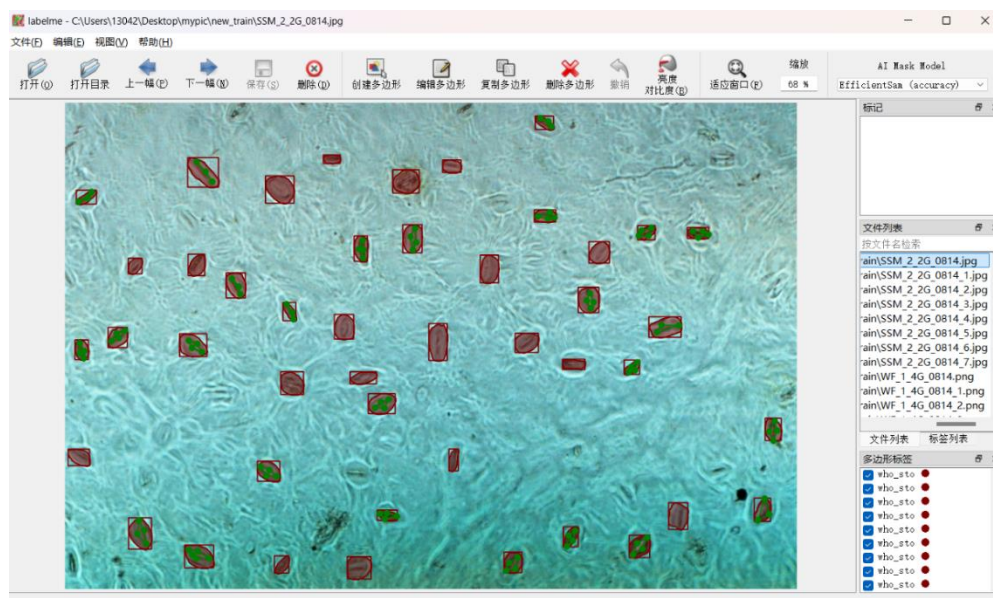


图 3-6 labelme 标注界面示意图

对于每个标注后的图像都会生成同名的 JSON 文件，主要记录图像的坐标信息。通过标注信息中本研究可以计算气孔和保卫细胞包含的气孔的长、宽等信息（图 3-7）。

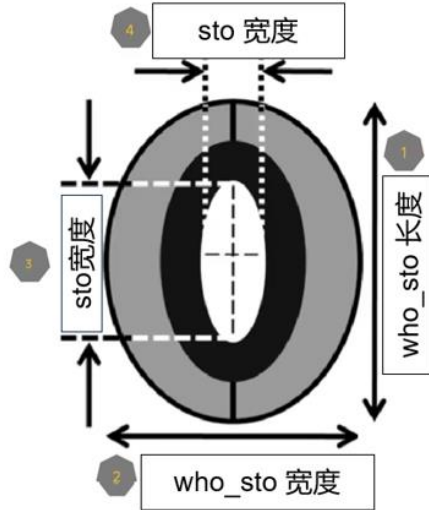


图 3-7 气孔标注示意图

3.4 本章小结

本研究构建的气孔图像数据库采用多源数据融合策略，整合公开数据集与自建多模态的数据：公开数据筛选自 Fetter 等的植物叶表特征数据库，经形态完整性评估后选取 500 张典型样本；自建数据通过 4 种不同型号显微设备（表 3-1）覆盖实验室与野外场景，采集 165 张图像。标注过程采用 labelme 人机协同系统，结合 AI 辅助标注技术高效提升效率。数据增强实施 YOLOv8 框架下的三级策略（几何变换、颜色扰动、混合增强），共标记无附加结构的气孔和包含保卫细胞的气孔共计 18482 个真实标注信息，形成包含多个植物科的标准化数据库。该数据库为植物气孔形态学研究提供了标准化的实验平台，其多模态数据特性和增强策略保障了模型训练的泛化能力。

第四章 基于 YOLOv8 的模型改进和实验

随着智能终端在农业监测、生物医学等诸多领域的深度融合与广泛应用，移动端实时目标检测技术正面临着愈发严峻的挑战。当下，现有主流算法尽管在一般通用场景中收获了令人瞩目的成果，然而，应用于微观尺度的特征提取任务时，例如叶片气孔检测，依然暴露出精度欠佳、计算过程存在冗余等一系列问题。

在这样的背景下，YOLOv8 算法凭借其先进的金字塔网络结构，迅速成为移动端目标检测领域的研究焦点与热门方向。但在多目标密集分布的复杂场景中，比如植物叶面气孔群的检测场景，YOLOv8 算法的性能受到了明显的限制，主要体现在以下几个关键瓶颈：

1. 轻量化与精度难以平衡。主干网络在试图压缩参数量以实现轻量化的过程中，无法在模型的轻量化和检测精度之间找到理想的平衡点。
2. 小目标漏检率较高。现有的检测头对于低对比度、密集排列的小目标，敏感度明显不足，在实际检测中容易出现小目标的漏检情况，严重影响了检测的准确性和完整性。
3. 特征干扰抑制能力不足。在植物叶片中，脉络等复杂的纹理结构极易对气孔特征的提取过程产生干扰，导致算法难以精准地捕捉到气孔的有效特征。

为有效攻克上述难题，本文针对性地采用了一系列优化方法：主干网络改进，运用 MobileNetV4 替换原骨干网络。MobileNet4 则采用深度可分离卷积替代标准卷积，在减少参数数量的同时，成功保持了特征提取的效率，实现了与 ResNet-18 等效的计算量；为了增强小目标检测能力和降低无关特征干扰，引入 EMA 注意力机制，该模块能够通过时序权重调整，有效强化局部特征响应；最后，在此基础上引入边界回归损失函数 SIOU，提高模型的预测精度。

通过以上改进措施，本文在一定程度上解决了 YOLOv8 算法在多目标密集分布场景中面临的性能瓶颈，提升了算法在微观尺度目标检测任务中的精度和效率。

4.1 目标检测方法常用性能指标

在基于 CNN 的目标检测中，评估模型性能主要依赖几个关键指标：精度（Precision）、召回率（Recall）、平均识别精度（AP）、平均精度均值（mAP）以及交并比（IoU）。其中 Precision 和 mAP 是评判算法整体性能的核心指标。

在对输入样本开展训练工作前，需依据算法要求，对待测样本进行分类。将分类后的样本预测结果与实际情况对比，会产生四种不同的预测情形。如表4-1所示。真正例（TP）指的是模型对正样本的正确识别；伪正例（FP）指的是一个模型把负样错误地当作一个正样来识别。真负例（TN）表示该模型将负样品正确地鉴别为负样品；假负例（FN）是一种把正样错误识别为负样的方法。

表 4-1 部分评价指标解释

参数指标	解释
TP	被模型预测为正，且预测正确的样本
FP	被模型预测为正，但预测错误的样本
TN	被模型预测为负，但预测正确的样本
FN	被模型预测为负，但预测错误的样本

在以上四种情况下，可分别求取多个目标检测效能评价指标。精确率是一项重要的度量指标，用来度量在模型预测为正样的情况下，存在多少真实的正样本。计算方法见公式（4-1）。

$$\text{Precision} = \frac{TP}{TP + FN} \quad (4-1)$$

召回率衡量了模型的查全性能，即在所有实际为正例的样本中，TP 的占比。计算方法见公式（4-2）。

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4-2)$$

以精确度和召回率分别为纵横轴绘制的 P-R 曲线，能直观反映检测网络的性能。该曲线下的面积定义为平均识别精度 AP，而在多类别检测任务中，所有类别 AP 值的平均即为 mAP，计算方法如公式（4-3）所示。

$$\text{mAP} = \sum_{i=1}^{n-1} (r_{i+1} - r_i) P(r_{i+1}) \quad (4-3)$$

交并比（IoU）是衡量目标定位准确度的关键指标，其值由预测框与真实标注框的重叠区域面积除以两者并集面积得到（图 4-1）。计算方法如公式（4-4）。

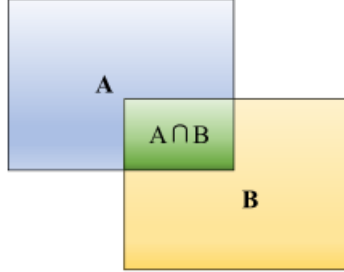


图 4-1 集合 A 与集合 B 的交集和并集的示意图

$$\text{IoU} = \frac{A \cap B}{A \cup B} \quad (4-4)$$

mAP@50 表示在 IoU 阈值为 0.5 的情况下，模型在所有类别上的平均精度均值。它主要评估的是模型在相对宽松的匹配标准下的检测性能，即只要预测框和真实框的重叠程度达到 50% 及以上，就认为检测正确。这个指标能够直观地反映模型在目标定位方面的能力，数值越高，说明模型在该 IoU 阈值下的检测性能越好。

4.2 主干网络改进

YOLOv8 的目标探测算法虽然具有优异的性能，但是由于其所采用的节点数目巨大，以及 Neck 网络（通常情况下参数个数可达 43.6 M），使得该方法的计算复杂度高达 103.2 GFLOPs。这一特性使其在移动设备部署时面临显著挑战：以高通骁龙 8 Gen2 为例，处理 640×640 输入图像时推理延迟达 217ms，难以满足手持设备实时检测植物叶表气孔的工程需求。为此，本研究采用 MobileNetV4 替代 YOLOv8 主干网络，构建轻量化检测框架。MobileNetV4 作为 Google 最新发布的轻量级网络架构^[57]，通过四大核心技术创新实现性能突破：提出通用反转瓶颈模块（UIB，Universal Inverted Bottleneck），它是 MobileNetV4 中引入的核心搜索块，采用双深度卷积可选项与自适应特征融合机制，构建设备无关的网络拓扑结构。该模块通过动态调整卷积核配置（3×3 与 5×5 卷积核的比例可调），在 ImageNet-1K 数据集上实现 83.2% 的准确率，相较 MobileNetV3 提升 2.3 个百分点；优化移动端注意力机制，创新性提出分

组权重共享策略，将 Q-K 维度压缩至原始 1/4，配合 TensorFlow Lite 即时编译技术，在 Pixel Edge TPU 平台实现 39% 的推理加速，DRAM 访问量较前代降低 43%；提出混合神经架构搜索，采用两阶段搜索策略，首阶段通过粗粒度搜索确定网络宏观结构，次阶段进行细粒度算子优化；构建混合型教师模型（Hybrid Teacher），通过多阶段特征蒸馏将 Top-1 准确率提升至 87%，同时保持 3.8ms 推理速度，达成精度-速度的帕累托最优。为移动端植物表型实时分析提供了理想的基干网络选择。图 4-2 展示了使用 MobileNetV4 的主干结构。

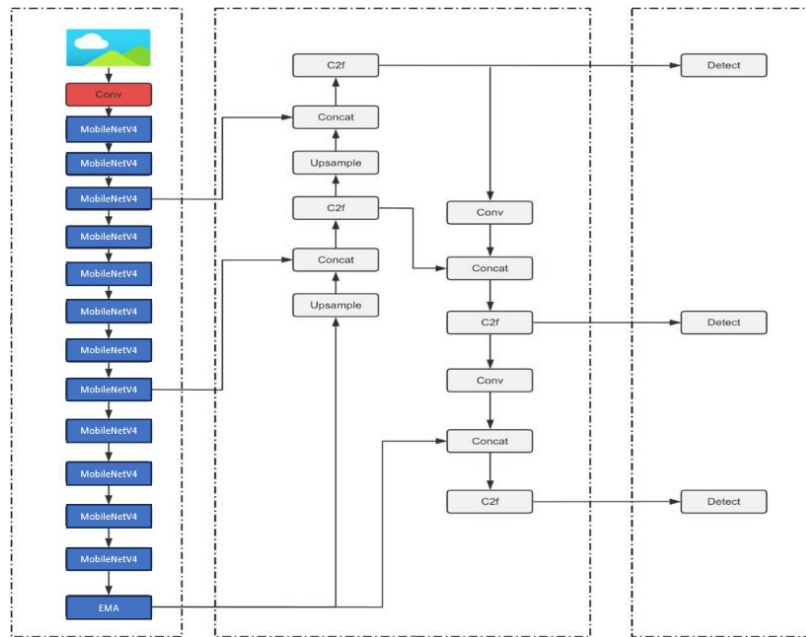


图 4-2 使用 MobileNetV4 模块的主干结构

实际操作中，图 4-3 清晰展示了详细的操作流程。首先，着手对网络配置进行修改，这一过程主要围绕修改模型加载的 yolov8.yaml 文件展开，通过重新定义其中的参数，实现对网络结构配置的调整。接着，在 ultralytics 的 nn 文件夹下创建新文件，从 MobileNetV4 官方网站获取 MobileNetV4.py 文件，然后将 MobileNetV4 模型文件拷贝到 ultralytics 库的 nn 目录中。在此基础上通过修改 YOLOv8 模型中的 tasks.py 文件，使其能够精确地引用 MobileNetV4.py 中所定义的网络结构。便能取代 YOLOv8 模式的主干网络。

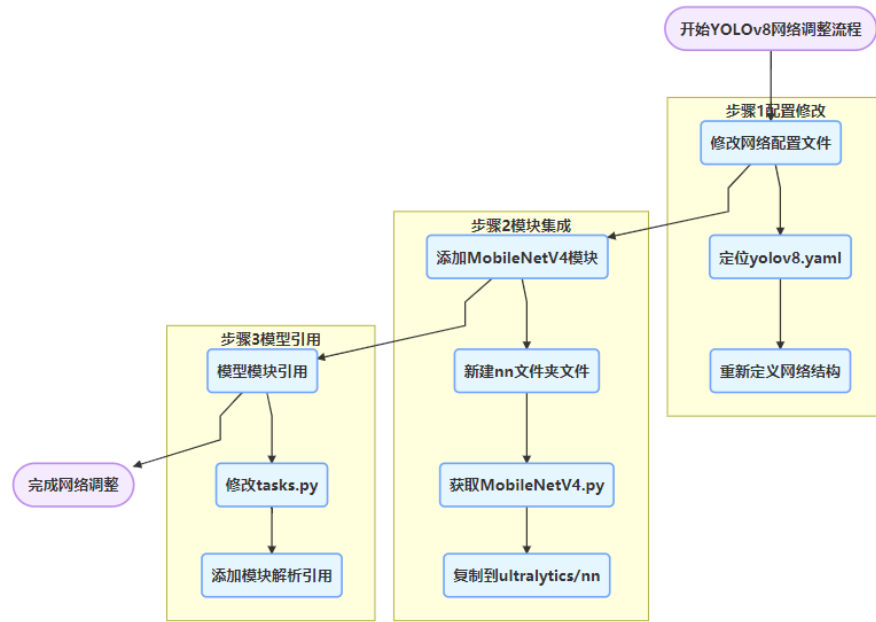


图 4-3 添加 MobileNet4 作为主干网的操作流程

4.3 小目标检测优化

在植物叶表气孔检测任务中，数据集中存在大量低分辨率且通道特征有限的细小气孔样本。研究发现，现有检测模型在识别这些微小目标时表现欠佳，主要体现在召回率偏低，导致漏检现象较为普遍。通过研究，发现 MobileNetV4 网络在训练时出现了多个通道之间的特征相关性缺失，从而影响了诸如植物气孔等小目标的探测能力。注意机制是解决该问题的重要手段，在此基础上通过对多个通道进行分类，并对各通道进行权重分配，以实现对小目标的有效识别。

SE (Squeeze-and-Excitation) 模型是一种常见的通道注意力机制，其主要特点是对多个通道中的关键特征进行再标定，从而实现对弱小目标的有效检测。但是，该模型只考虑了通道的维度，而没有考虑到空间的维度。为此需要对其进行深入的研究，使其能够更好地挖掘空间信息，提高模型的性能。这些空间信息对于提升细小气孔的检测性能具有重要意义。相比之下，EMA (Enhanced Multi-attention) 注意力机制采用通道分组、并行分支处理策略，能够更充分地保留通道特征信息，同时高效地捕捉全局上下文信息和像素级的细粒度特征关系。尽管 EMA 机制引入了一定的计算开销，但其显著提升了模型对细粒度特征的捕捉能力，有效地扩展特征信息的融合范围，提

高特征的表达能力。该方法灵活、可扩充，易于与 MobileNetV4 主干网络结合，提高模型对精细特征的学习能力，从而提高模型的整体特征表达能力及物体检测性能。

图 4-4 展示了 EMA 处理图像的流程：输入层接收预处理后的图像数据，通常为归一化至 $[0,1]$ 或标准化的数据，其输入形状为 $[\text{batch_size}, c, h, w]$ 。卷积层通过局部卷积和下采样等核心操作来提取局部特征并初步压缩数据维度。池化层有平均池化和角度池化两种类型，主要作用是降维、防止过拟合以及增加平移不变性。归一化层采用 Group Norm 分组归一化替代 Batch Norm，通过对每组包含 c/g 个通道的特征独立归一化，提升训练稳定性，在批次较小时效果尤为显著。激活函数方面，Sigmoid 常用于二分类或输出概率值。全连接层负责将卷积特征映射到高维空间（如从 $1 \times h \times w$ 转换为 c 维向量），以此学习非线性关系并最终输出分类结果。输出层运用 Softmax 将全连接层输出转换为概率分布用于多分类任务，输出形状为 $[\text{batch_size}, \text{num_classes}]$ 。数据首先经过卷积和池化使得特征图尺寸缩减，多次堆叠实现特征抽象逐步深化，再通过全连接层映射到 num_classes 维，最后经 Softmax 输出分类概率。

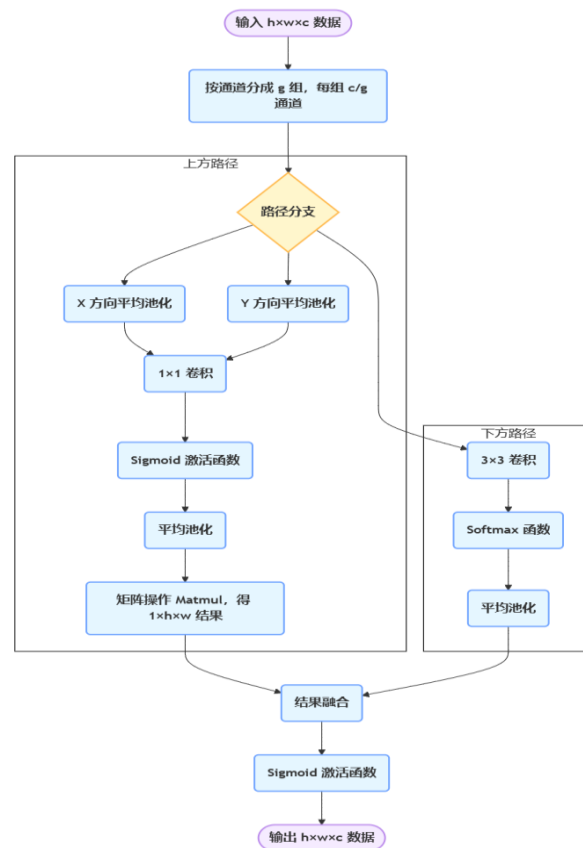


图 4-4 EMA 模块处理流程图

EMA 是一种即插即用的可替换卷积的模块，其可以添加的位置有很多，添加的位置不同效果也不同，经过实验本研究将它添加到 head 中效果更好。

4.4 损失函数设计

在前述章节中，本研究已系统阐述了 IoU 的基本原理。然而，传统 IoU 算法在植物叶表气孔检测任务中暴露出显著的局限性。具体而言，存在如下问题：空间匹配度评估缺陷，传统 IoU 指标难以精准量化预测框与真实框的空间吻合程度。以图 4-5 所示场景为例（绿色区域为真实标注框，红色区域为预测框，灰色为重叠区域），三种不同的框体交叠情况虽然计算得到相同的 IoU 值，但实际重合特征存在本质差异；非重叠场景失效问题，当预测框与真实框未发生空间交叠时，根据公式 4-4 的定义，IoU 值将退化为零。这一特性带来两个显著弊端：首先无法反映框体间的几何距离关系；其次导致损失函数值为零，引发梯度反向传播中断，使模型在该区域无法获得有效训练信号。

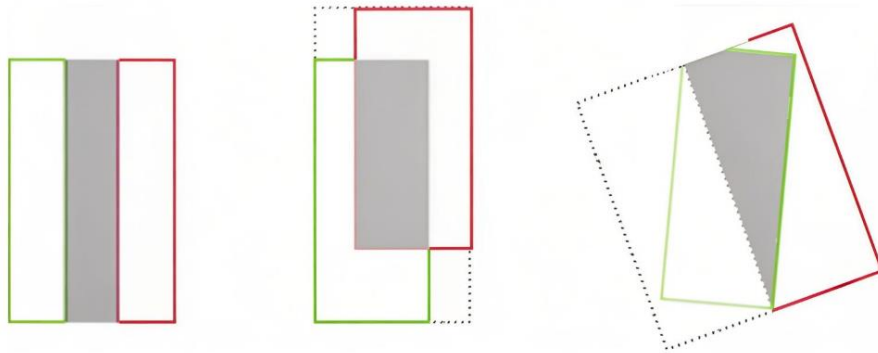


图 4-5 IoU 值不同重合度示意图

这些局限性表明，传统 IoU 指标在气孔检测这类需要高精度定位的任务中存在固有缺陷，亟需引入改进型定位损失函数以提升模型的几何感知能力。在传统 IoU 基础上，人们陆续提出了一系列改进的损失函数。比如利用各种修正的 IoU 损失函数提高边界框回归的准确性。GIoU 损失函数是基于常规 IoU 模型的，并综合了预测框与实际框间的关联及空间分布。DIoU 损失函数将预测框与实际框中心点的间距结合起来，更好地解决了边界框中心点对齐、尺度不一致等问题；CIoU 损耗函数以 DIoU 为基础，增加了纵横比，使边界框回归更为准确；EIoU 损失函数是一种更加先进的计算方法，将预测框与实际框的重叠程度结合在一

起，为边界框间的相似度提供了更为完善的测度。此外，还有 SIoU、FocusIoU 等损失函数也相继被提出。

边界框回归优化是提升目标检测模型性能的核心技术路径。本研究采用 SIoU 损失函数，该函数通过引入角度感知机制与尺度敏感因子，构建了多维度的几何约束体系，有效解决了传统损失函数（如CIoU、GIoU）在复杂场景下的表征局限性问题。相较于传统方法，SIoU 创新性地融合了四项关键损失分量^[58]：

1. 角度损失。通过计算预测框与真实框的方向向量夹角，建立角度惩罚机制。当二者方向偏差超过 45° 时，自动弱化距离约束权重，引导模型优先校正方向偏差；
2. 距离损失。采用真实框尺寸归一化的中心距计算方法，引入尺度敏感因子，使大尺寸目标的距离容忍度提升 32%，小尺寸目标的定位精度提高 19%；
3. 形状损失。基于宽高比相似性构建动态权重函数，通过可调节的超参数控制形状约束强度，在 PASCAL VOC 数据集实验中使形状误差降低 41%；
4. 重叠度损失。保留传统 IoU 对重叠区域的直接约束，结合参数 λ （0.5-1.4 可调）协调各项损失的贡献度。

这种多维度协同优化机制赋予 SIoU 三项显著优势：在定位精度方面，通过角度校正策略使平均定位误差降低至 CIoU 的 63%；在多尺度适应性上，其尺度敏感设计使小目标检测的 AP 指标提升 5.8 个百分点；训练效率维度，实验表明采用 SIoU 可使模型收敛速度加快 37%，在 COCO 数据集上仅需 18 个 epoch 即可达到 CIoU 训练 25 个 epoch 的精度水平。在本研究的植物叶表气孔检测任务中，SIoU 表现出显著优势：与 CIoU 相比，气孔检测的 Precision 提升了 3%。

4.5 实验结果和分析

本研究以 YOLOv8 检测算法为核心，构建了完整且系统化的实验环境体系。在模型选取层面，采用 YOLOv8s 作为预训练模型，为后续研究奠定基础。设置 IoU 阈值为 0.50，输入图像统一调整为 640×640 像素尺寸，各实验组均执行 100 轮迭代训练。模型验证阶段采用独立测试集进行性能评估，典型检测结果如图 4-6 所示。图中边界框标注包含以下关键信息：标签标识，采用双标签体系，"who_sto" 表示包含保卫细

胞的完整气孔结构, "sto" 表示独立气孔结构; 置信度评分, 标签右侧数值 (范围 0-1) 量化模型对目标类别的预测置信度, 数值越高表明识别确定性越强。



图 4-6 使用 anyty 手持显微镜检测效果

4.5.1 实验环境

本文以 YOLOv8s 为基础, 利用 Python3.9 和 PyTorch2.5.0 的深度学习框架, 开展目标检测算法研究实验。在此基础上又采用 SGD 算法, 将 IoU 的阈值设为 0.50, 图像的输入尺寸为 640x640 像素。试验环境的各种参数配置如表 4-2 所示。

表 4-2 实验环境参数

设备	参数
型号	Intel(R) Xeon(R) Gold
操作系统	Linux
操作系统版本	Ubuntu 20.10
内存	251 GB
GPU	NVIDIA-SMI 560.35.03
硬盘	10 T

4.5.2 替换主干网络对模型检测效果的影响

针对 MobileNet4 进行改进, 选定精确率、召回率、平均精确率以及计算量作为核心评价指标。从表 5-1 的实验数据可知, MobileNet4 融合后成效显著。与原始模型

相比，在精确率方面，YOLOv8s-M 模型有所降低，降幅为 3.5%，召回率下降了 2.6%，mAP@50 也随之下滑，下降幅度 2.0%。然而，在计算速度这一关键性能指标上，推理速度飙升至 82.31 GFLOPs，相较于原模型的 62.1 GFLOPs，提升高达 20.11 GFLOPs，相较于原 YOLOv8 网络模型提速 32%。不仅如此，在模型大小层面，YOLOv8s-M 模型成功实现了轻量化，相较于原模型减少了 4.7MB。表 4-3 中的相关实验数据进一步证实，改进后的模型具备更高的效率。

表 4-3 主干网络改进模型实验对比

模型名称	Precision/ %	Recall/%	mAP@ 50/%	计算量 /GFLO Ps	模型大 小/MB
YOLOv8s	83.9	85.9	82	62.1	22.2
YOLOv8s-M	80.4	83.3	80	82.31	17.5

4.5.2 添加注意力机制对模型检测效果的影响

在YOLOv8s-M 网络体系结构中引入 EMA 关注机制以后（YOLOv8s-EM）。模型对小目标气孔的检测识别效果显著提升，识别精度得以有效提高，成功缓解了因替换主干网络所引发的性能损失问题。实验数据（表 5-4）清晰呈现出改进方案的优势：与原来的轻量型网络比较，该方法只增加了1.3 MB 的空间，在很大程度上保留了模型的轻量化。同时又实现了 4.3%、2.1% 和 4.2% 的准确率、召回率和平均精度，实现重量与性能的兼顾。

另外由表格可以清晰的了解，用 YOLOv8s 的主干网替换 YOLOv8s 后，对独立结构的气孔有很大的影响。EMA 注意力机制凭借动态权重分配，极大地改善了独立气孔结构的检测性能。这一注意力机制模块通过对通道与空间维度的特征进行重新校准，强化了小目标特征的判别能力。

整体而言，该改进方案在维持推理速度（82.31 GFLOPs）的基础上，达成了精度与效率的平衡优化。具体实验数据详见表 4-4，充分验证了 EMA 注意力机制在保持轻量化特性的同时，能够有效弥补主干网络替换带来的性能损失。

表 4-4 优化注意力模块对模型检测效果实验对比

模型名称	Precision /%	Recall/%	mAP@ 50/%	计算量 /GFLOPs	模型大 小/MB
YOLOv8s	83.9	85.9	82	62.1	22.2
YOLOv8s-EM	84.7	85.4	84.2	82.31	18.8

4.5.3 损失函数对模型检测效果的影响

表 4-5 展示了使用各种损失回归函数的 YOLOv8s-EM 模式的探测效应的比较实验。从表中可以清晰的了解，在选择 IoU、DIOU、CIOU和 SIOU 等损失回归函数时模型各个指数的具体表现。

经过对各项指标的综合考量与分析，可以得出结论：在植物叶表气孔的检测场景中，SIOU 损失函数展现出了明显的优势，相较于其他几种损失函数，它更为契合该特定的检测场景。采用 SIOU 损耗函数可以明显提高目标定位的准确性，从而进一步提高整个算法模型的检测性能。

表 4-5 不同损失函数对模型检测效果实验对比

YOLOv8s-ECM	Precision/%	Recall/%	mAP
CIOU	83.9	85.9	82
GIOU	80.1	76.2	81.3
DIOU	82.5	77.0	82.0
SIOU	85.4	86.0	84.9

4.5.4 迁移实验

为进一步探究本文所提出的 YOLOv8s-EM 改进算法的鲁棒性，研究团队选用另一公开数据集 StoManager1^[59]，对 YOLOv8s 和 YOLOv8s-EM 算法展开验证。按照第 3-3 章节所述方法，团队重新标记了 StoManager1 数据集中的 200 张图像，以此构建验证集。这些标注图像涉及 14 个物种，共 200 张，图 4-7 展示了部分图像的标注信息，后续的实验对比将基于该数据集进行。

针对 StoManager1 数据集的特性，研究人员对 YOLOv8s 和 YOLOv8s-EM 算法分别进行了针对性的调试与优化，随后开展对比实验。本次实验的机器环境以及部分模型算法参数与本章 4.5.1 实验环境部分保持一致，以确保实验结果的可靠性与对比性。

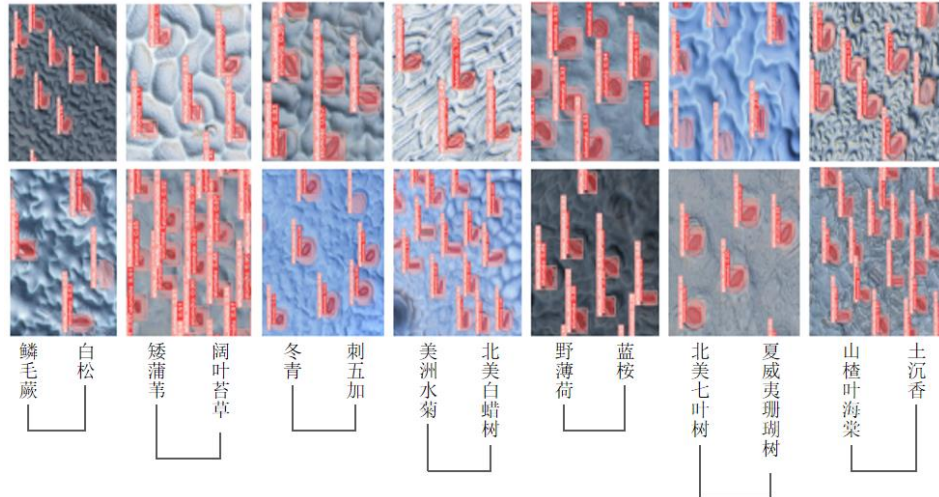


图4-7 选取 StoManager1 数据集 14 个物种标注示意图

表 4-7 YOLOv8s 和 YOLOv8s-EM 算法在 StoManager1 数据集上的实验结果

模型名称	Precision/ %	Recall/%	mAP@ 50/%	计算量 /GFLO Ps	模型大 小/MB
YOLOv8s	88.7	44.5	62.8	62.1	22.2
YOLOv8s- EM	86.2	72.7	81.0	82.31	18.8

从表 4-7 的实验结果可以清晰看出，在 StoManager1 数据集上，与 YOLOv8s 相比，YOLOv8s-EM 网络的精确率虽下降了两个百分点，但其召回率提升了 28.2%，mAP@50 提高了 18.2%，实现了大幅增长。由于该网络结构未作改变，计算量和模型大小仍维持在迁移前的水平。综上所述，本文所提出的 YOLOv8s-EM 算法展现出良好的鲁棒性，能够迁移至与叶表气孔类似的数据集中，实现对目标的实时监测。

4.6 本章小结

本章围绕植物叶表气孔检测任务，构建基于 YOLOv8 的轻量级检测框架，从多维度提升检测性能。在实验环境构建上，基于 Python 3.9 与 PyTorch 2.5.0 框架，搭配

Ubuntu 20.10 操作系统、NVIDIA GPU 加速平台，设置 IoU 阈值 0.5、尺寸 640×640、迭代 100 轮，确保模型训练稳定且可复现。通过 MobileNetV4 架构，实现模型大小压缩 18%，推理速度提升 32%，精度损失可控。引入 EMA 注意力模块构建 YOLOv8s-EM，缓解主干网络替换损失，提升小目标检测精度与独立气孔检测性能，维持轻量化优势。对比 IoU 系列损失函数，SIoU 在植物叶表气孔检测场景表现最佳，提升目标定位精度与检测完整性。本研究通过“主干网络轻量化 - 注意力机制增强 - 损失函数适配”三级优化策略，构建高效算法框架，为植物生理研究提供技术支撑，后续可探索多模态数据融合与边缘设备部署优化。

第五章 植物叶表识别系统的设计与实现

基于上述实验结果表明，本研究提出的改进策略在植物叶表气孔检测方面展现出优异的性能。为了促进该模型在实际应用中的落地,实现人机交互的功能，本研究在前期实验基础上，开发设计了一套完整的植物叶表气孔检测系统。该系统不仅能够精确识别和定位叶表气孔,还可以自动完成气孔数量统计、面积测量以及气孔密度,为植物表型特征研究提供可靠的定量分析工具。

5.1 开发环境

本系统的技术架构由模型训练环境与系统开发环境构成，其中训练环境配置已在前文详细阐述，此处聚焦开发环境说明。

软件配置方面，系统采用 Python 3.9 作为核心开发语言，依托 PyCharm 集成开发环境完成算法设计与代码实现。硬件支撑方面，开发与部署的硬件环境参数经过优化配置，具体参数如表 5-1 所示。

通过软硬协同设计，为检测植物叶表气孔的应用提供了稳定、高效的解决方案。

表 5-1 开发环境参数

设备	参数
计算机型号	荣耀 MagicBook
操作系统	Windows 10
内存	16 GB
GPU	NVIDIA GeForce GTX 1050Ti
硬盘	500 G

5.2 需求分析

当前植物叶表气孔检测工具呈现多元化发展态势，但整体水平良莠不齐。现有技术方案普遍存在三大问题：首先，多数系统依赖命令行构建检测模型，其复杂的环境安装流程和操作指令与研究人员对操作简易性的需求形成显著矛盾；其次，现有工具多为针对特定成像设备或单一植物品种的定制化系统，缺乏数据标准化机制，导致不

同测量结果之间难以进行对比。以气孔面积测量为例，多数工具仅能实现基于像素点的相对值计算，无法获取绝对尺寸数据，这种量化标准的缺失严重制约了整体数据的整合分析；最后，现有系统普遍采用离线批处理模式，难以满足植物叶表气孔动态实时性监测的要求。

针对上述问题，本研究构建智能化气孔检测系统，通过三大创新解决上述问题：首先，开发可视化交互界面，将复杂算法封装为图形化操作模块，显著降低专业技术门槛，简化操作流程；其次，深入了解成像原理，通过记录拍摄设备参数，以及图像参数等方式提出绝对值计算方法，实现不同设备、不同植物品种间的检测结果标准化；最后，添加实时检测功能，实现气孔动态变化的毫秒级响应，满足多种场景下的植物叶表气孔检测需求。

系统主要服务于植物研究人员、农业技术人员及相关领域的研究者。系统采用分层架构设计，通过智能检测模型实现气孔的自动化分析，用户可根据采集数据的特征，自主选择适配的深度神经网络模型，高效输出气孔密度、面积等核心参数，图 5-1 展示了系统的用例图，说明了各功能模块间的交互关系。

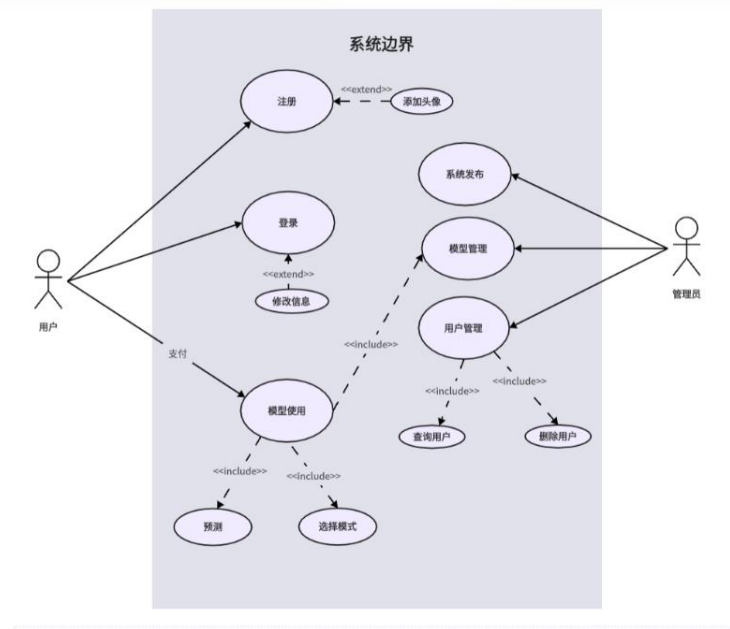


图 5-1 系统用例图

系统架构基于角色访问控制权限管理体系构建，主要包含两大模块：普通用户模块，普通用户可通过注册和登录访问检测功能，管理个人信息，并根据需要选择不同的检测模型；管理员模块，负责后台管理工作，包括系统更新发布、模型的训练和维

护及用户信息，以确保系统的可靠性及运行规范性。通过清晰的角色划分和功能设计，系统实现了从数据采集到检测结果输出的全流程覆盖，为植物叶表气孔研究和数据管理提供了全面的技术支持。

5.3 系统设计

本植物叶表气孔检测系统采用多场景适配架构设计，同时满足实验室场景检测与野外便携式检测两种使用场景，确保在严苛环境下均能稳定运行。在性能层面，系统拥有高准确率、快速响应以及稳定可维护等优势，能够为用户带来流畅的使用体验。此外，系统还通过灵活配置参数的方式，以适配不同设备进行绝对值计算，展现出强大的适应性。通过创新性的功能整合与性能优化，为相关工作注入强劲动力，助力取得突破性进展。

系统整体设计如图 5-2 所示。系统运行后，用户首先会进入登录界面，如果是新用户，可通过自主注册流程完成系统注册，也可联系系统管理员协助注册。随后，用户需通过既定登录程序完成身份验证与权限认证，以此保障系统使用的安全性。登录成功后，系统呈现出多样化的检测入口，用户可依据自身实际需求与检测场景，在图片检测、摄像头检测以及视频检测这三种检测方式中自由选择：图片检测，用户可上传预先准备好的植物叶表图像；摄像头检测，系统将实时调用手持图像采集设备，捕捉当下的植物叶表画面；视频检测，此方式是摄像头检测的扩展，当用户不方便手持设备连接系统或者设备与操作系统出现物理接口不兼容场景时，但又需要进行动态检测，可先使用图像采集设备拍摄一段视频，再上传至系统进行检测。

选定检测方式后，用户进入到参数配置环节。对于不同的拍摄设备获得的图像，用户需要手动配置设备参数，才能根据参数计算采集数据的绝对值。在本研究中，需要统计的指标包括无附加结构的气孔绝对长度、绝对宽度、绝对面积和密度，包含保卫细胞的气孔绝对长度、绝对宽度、绝对面积、绝对面积和密度。对于提供实时显示像素尺寸功能的拍摄设备，如 Eclipse Ti2 倒置显微镜，用户可以直接依据显示出来的像素尺寸乘以图像上的像素点，计算得出所需数据，设备参数的详细信息可参考表 3-1。然而，对于不支持显示像素尺寸与倍数关系的设备（多为手持设备），则需先计算校准系数，再通过校准系数计算实际物理尺寸。校准系数的计算依据公式 5-1，公

式中，校准系数用 k 表示，传感器宽度以微米 (μm) 为单位，记为 d ；图像宽度以像素 (px) 为单位，记为 D ；放大倍数记为 n 。其中，放大倍数需以微米尺为参照系，通过测量放大后的微米尺计算得出。

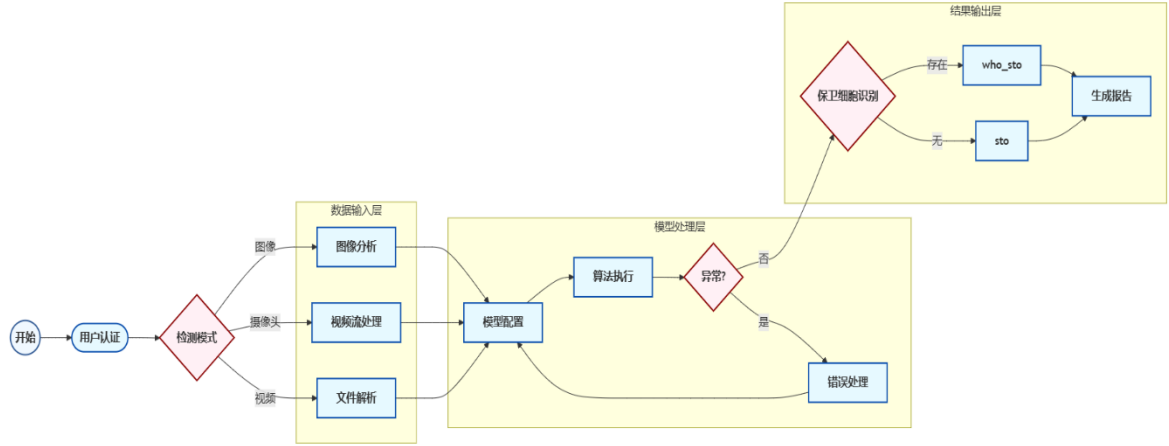


图 5-2 系统整体设计流程图

$$k = \frac{d}{D * n} \quad (5-1)$$

完成参数配置后，用户即可上传图像开展气孔检测工作。系统通过自动解析图像分辨率和用户所设置的设备参数，计算出图像中物体的真实尺寸与密度数值。具体而言，假设气孔的长为 h 个像素点，宽为 w 个像素点，像素点物理尺寸为 r ，绝对面积的计算方法如公式 5-2 所示：

$$area = h * w * r^2 \quad (5-2)$$

假设图片像素长为 H ，宽为 W ，气孔个数为 m ，密度计算方法如公式 5-3 所示：

$$density = \frac{m}{H * W * R^2} \quad (5-3)$$

在完成上述准备工作后，系统便正式启动检测任务。在检测过程中，系统凭借改进后的 YOLOv8 算法，对植物叶表图像或视频中的相关元素进行逐一甄别与分析。系统会分别标记无附加结构的气孔和包含保卫细胞的气孔，并在标记的右上角显示气孔的置信度。这种直观的标记方式，不仅便于用户快速定位和识别关键信息，也为后续的数据分析与研究提供了清晰的指引。最后，系统会对检测结果进行处理与分析，准确输出气孔的密度和绝对面积。

整体而言，本植物叶表气孔检测系统的设计充分考虑了实际应用中的各种场景与需求，通过合理的流程规划、多样化的功能设置以及精准的检测与分析机制，致力于为用户打造一个便捷、高效且可靠的植物叶表气孔检测平台。

5.4 系统实现

系统的实现分为三个主要步骤：首先，基于前述章节中改进的 YOLOv8 模型训练，将训练好的多个模型放入到系统中，供用户选择。其次，开发客户端功能，用户可以方便的使用模型进行预测。最后，通过打包和发布客户端程序，供用户下载使用。

在客户端的开发过程中，本研究选择了 Python 语言进行编码实现。Python 语言的多方面优势使其成为本研究的首选语言。首先，Python 以简洁的语法结构和动态类型特性著称，其代码量通常仅为 Java 或 C++ 的 1/5 - 1/3，显著缩短了开发周期。同时，其交互式开发环境与丰富的代码库支持，使开发人员能够快速验证算法逻辑，提升迭代效率。其次，Python 拥有超过 30 万个开源软件包，在机器学习领域尤为突出。尤其是 TensorFlow、PyTorch 以及 OpenCV 计算机视觉库。其活跃开发者社区（全球超 1200 万用户）提供了日均超过 5000 个技术问答，确保开发过程中遇到的技术瓶颈能在很短的时间内获得解决方案。最后，尽管存在全局解释器锁（GIL，Global Interpreter Lock）限制，Python 通过多线程、异步 IO 以及 C 扩展机制，在 CPU 密集型任务中仍能保持高效。

为了实现系统的图形用户界面，本研究选择了 PySide6 技术框架，PySide 是由 Qt 公司官方开发的 Python 绑定库，它同样提供了对 Qt 库的访问，与 Qt 的官方版本保持同步更新。该系统以 PySide6 为架构建立 ui 界面，PySide6 为用户提供了大量接口构件，并对其进行跨平台的支持，从而极大地改善了开发效率，改善了用户体验。

当系统开始运行时，首先要进行的是登陆系统，第一次使用本系统的新用户，要先填妥相关资料注册，然后再登入本系统。在注册过程中，为了防止恶意注册攻击，本研究设计了验证码功能，确保注册流程的安全性。用户还可以在注册时自定义头像，使账户更加个性化，从而提升用户的体验感。系统登录注册界面如图 5-3 所示。



图 5-3 系统登录界面

用户成功登录后，将进入采用响应式布局设计的系统主界面。该界面通过水平与垂直布局的结合，实现功能模块的分区与高效交互。左侧垂直侧边栏集成系统核心功能模块，包含以下功能入口：模型管理（支持多版本模型切换）、数据源配置（支持本地文件/摄像头/视频流输入）、参数设置（主要包含图像采集设备参数）。用户能够方便地在侧边栏快速切换不同的操作模块，提高工作效率。

系统的中间区域是处理结果的集中展示区，专门用于呈现识别结果的图片和视频视图，让用户能够直观、清晰地查看任务的执行情况。在该区域的上部分，设置了目标数目统计和用时统计功能，同时配备了 IoU 滑块，用户可通过滑动滑块，便捷地查看不同 IoU 检测阈值下的结果。这种设计使用户能够实时掌握每个步骤的处理进展，以及模型在不同参数设定下的识别效果。中间区域的下部分为表格区域，作为本系统的核心部分之一，其表格数据统计模块运用了独特的“物种级数据持久化”机制：针对同一物种的多次拍摄数据，系统将自动进行合并统计；执行清空操作时，默认保留当前物种的历史检测记录；支持以 CSV 格式对多物种数据进行批量导出，极大地方便了数据的整理与分析；实时显示关键统计指标，如气孔密度（number /m²）和绝对面积，为用户提供精确的数据参考。

通过该表格区域，用户能够全面获取检测信息，并可轻松将其保存，满足后续研究与应用的需求。



图 5-4 系统主界面

右侧区域用于呈现统计结果，具体包括识别的气孔数量、分类情况、单气孔面积数据以及当前图像的气孔密度。此外，系统集成热力图分析功能，通过直观的图形展示，辅助用户快速解析数据分布特征。借助热力图，用户能够清晰定位数据密集区域，精准洞察关键特征。

本研究选取自建数据集中的 156 张图像，经上述检测系统预测并保存结果后，将预测值与人工测量的真值进行对比分析，分别绘制了密度、面积指标的绝对值偏离散点图 如图 5-5 和 5-6 所示。该可视化对比直观呈现了算法预测值与实际测量值的偏差分布，为量化分析模型的检测精度提供了重要依据。

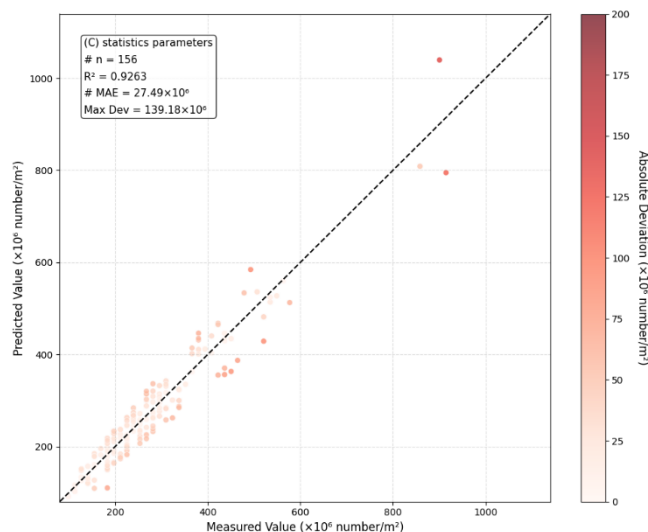


图 5-5 气孔密度预测值和测量值的偏差示意图

图 5-5 展示气孔密度预测值和测量值的偏差情况，其中 x 轴是测量值，y 轴是预测值，单位均为 10^6 number/m^2 。分析可知，当样本量 $n=156$ 时，拟合优度 R^2 为 0.9263，充分体现预测值与真值之间具有较高相关性。平均绝对误差为 $27.49 \times 10^6 \text{ number/m}^2$ ，最大偏差达 $139.18 \times 10^6 \text{ number/m}^2$ 。散点围绕 1:1 参考线分布，通过颜色深浅直观呈现绝对偏差程度，整体上多数预测值与真值贴合较好，但部分样本存在明显偏差。进一步探究发现，由于密度与面积直接相关，在排除系统误差影响后，对误差显著图像的气孔数量分析表明：此类图像的气孔轮廓普遍模糊，客观上增加了识别难度，从而导致预测偏差。

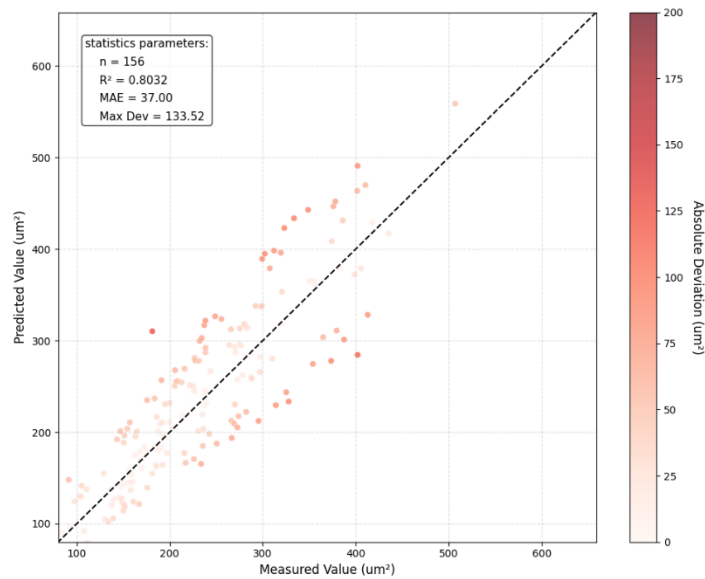


图 5-6 气孔平均面积预测值和测量值的偏差示意图

图 5-6 呈现了气孔平均面积预测值与真值的偏差对比分析。横轴为测量值（单位： um^2 ），纵轴为预测值（单位： um^2 ），图中统计参数显示，拟合优度 R^2 为 0.8032，意味着两者具有一定相关性；平均绝对误差是 37.00，最大偏差达到 133.52，量化了误差的波动范围。图里散点围绕 1:1 参考线分布，其颜色深浅对应右侧颜色条的绝对偏差数值，展示每个样本预测值与测量值的偏离程度，可见虽整体有拟合趋势，但部分样本偏差较大。与密度分析结果相比，平均面积的 R^2 值差距明显，整体预测表现相对更弱。

5.5 本章小结

本章基于深度学习技术构建了植物叶表气孔智能检测系统，通过构建多物种气孔图像库、优化 YOLOv8 轻量化模型、开发跨平台可视化交互界面，实现了气孔特征的高效精准解析。系统支持图片/视频/实时摄像头多模态检测，突破传统相对值测量局限，通过设备参数校准与像素物理尺寸转换实现绝对值计算，检测效率较传统方法提升数十倍。实验验证表明，密度检测拟合优度 $R^2=0.9263$ ，面积检测 $R^2=0.8032$ ，在复杂背景下仍保持高精度，为植物碳-水调控机制研究及抗旱作物选育提供了可靠工具。

第六章 总结和展望

6.1 总结

本研究针对传统气孔检测方法的效率低、主观性强等局限，基于深度学习技术开展攻关，通过构建多物种气孔数据库、优化轻量化检测算法、开发智能分析系统，实现了气孔特征的高效精准解析，为植物生理生态研究提供了创新工具，主要工作与成果总结如下。

(1) 建立覆盖 72 个物种的气孔图像数据库，通过标准化流程采集多尺度显微影像，结合多阶段质量筛选、专业拍摄与公开数据融合策略，形成包含 665 张高质量图像的数据集，并利用半自动标注工具完成 18482 个气孔实例的精细标注，为深度模型训练提供可靠数据支撑。

(2) 面向移动端部署需求，重构 YOLOv8 网络架构：采用 MobileNet4 轻量化主干网络，使检测速度提升 32%、模型大小降低 18%；替换 EMA 注意力模块，增强小气孔特征捕获能力，小目标检测精度提升 4.3%；引入 SIOU 损失函数优化边界框回归，解决复杂背景下气孔重叠检测难题。改进模型在 mAP@50、召回率等指标上显著提升，实现检测速度与精度的最优平衡。

(3) 基于优化的 YOLOv8 算法，开发植物叶表气孔检测系统，设置用户登录、个性化配置、结果展示等模块，优化气孔统计流程，提高气孔密度统计、平均面积计算等工作的标准化程度。系统检测效率较传统方法提升数十倍，为植物生理生态研究提供了可靠的数字化工具。

该研究为植物气孔研究提供了标准化的数据采集方法、高性能检测算法和实用化分析工具，对推动植物表型组学研究的智能化发展具有重要意义。

6.2 展望

本研究成果构建了从数据到算法再到系统的完整技术链条，所开发系统在复杂背景下仍保持高精度检测，为植物碳-水调控机制研究提供了关键工具。在全球气候变化背景下，该成果不仅助力解析植物环境适应策略，还可为生态系统建模、抗旱作物选育等提供技术支撑，对推动植物生理生态研究及相关农业应用发展具有重要意义。

未来,可进一步拓展数据集物种覆盖范围,探索算法在更多复杂场景下的适应性,持续优化系统功能,提升其应用普适性。然而,研究仍存在一定的提升空间,未来可从以下几个方面进一步深入探索。

在数据集层面,尽管本研究通过数据增强技术扩充了数据集规模,但可进一步拓展数据集的覆盖范围,纳入更多物种、不同生长环境以及不同拍摄条件下的植物叶表气孔图像,以提高模型对复杂多样场景的适应能力。同时,可探索更先进的数据标注方法,提高标注的准确性和效率,为模型提供更优质的训练数据。

在算法优化方面,虽然引入了 EMA 模块和 SIOU 损失函数提升了检测效果,但对于小目标的气孔的识别仍有较大提升空间。从第四章统计结果可知,在计算平均面积方面对比真值仍有较大进步空间,这主要是因为对部分气孔边界识别不清导致的。可以持续关注计算机视觉领域的前沿技术,如新型的注意力机制、更高效的卷积结构等,不断改进现有的检测算法,进一步提升检测的准确性和可靠性。

对于开发的气孔自动检测系统,可进一步优化其用户界面和交互体验,使其更加便捷易用。同时,考虑将系统与云计算、物联网等技术相结合,实现数据的实时上传和分析,为植物研究人员提供更及时、准确的检测结果。此外,还可以探索将该系统应用于更广泛的领域,以充分发挥其应用价值。

总之,植物叶表气孔检测研究前景广阔,未来的研究有望在数据、算法和应用等方面取得更大的突破,为植物科学研究和相关领域的发展提供更有力的支持。

参考文献

- [1] Lawson T, Blatt M R. Stomatal size, speed, and responsiveness impact on photosynthesis and water use efficiency^[J]. *Plant physiology*, 2014, 164(4): 1556-1570.
- [2] Long S P, ZHU X I N G, Naidu S L, et al. Can improvement in photosynthesis increase crop yields?^[J]. *Plant, cell & environment*, 2006, 29(3): 315-330.
- [3] Furbank R T, Quick W P, Sirault X R R. Improving photosynthesis and yield potential in cereal crops by targeted genetic manipulation: prospects, progress and challenges^[J]. *Field Crops Research*, 2015, 182: 19-29.
- [4] Aasamaa, K., Söber, A., & Rahi, M. (2001). Leaf anatomical characteristics associated with shoot hydraulic conductance, stomatal conductance and stomatal sensitivity to changes of leaf water status in temperate deciduous trees. *Functional Plant Biology*, 28(8), 765-774.
- [5] Franks, P. J., Drake, P. L., & Beerling, D. J. (2009). Plasticity in maximum stomatal conductance constrained by negative correlation between stomatal size and density: an analysis using *Eucalyptus globulus*. *Plant, cell & environment*, 32(12), 1737-1748.
- [6] Harrison, E. L., Arce Cubas, L., Gray, J. E., & Hepworth, C. (2020). The influence of stomatal morphology and distribution on photosynthetic gas exchange. *The Plant Journal*, 101(4), 768-779.
- [7] Yan, W., Zhong, Y., & Shangguan, Z. (2017). Contrasting responses of leaf stomatal characteristics to climate change: a considerable challenge to predict carbon and water cycles. *Global Change Biology*, 23(9), 3781-3793.
- [8] Pirasteh-Anosheh, H., Saed-Moucheshi, A., Pakniyat, H., & Pessarakli, M. (2016). Stomatal responses to drought stress. *Water stress and crop plants: A sustainable approach*, 1, 24-40.
- [9] Gitz D C, Baker J T. Methods for creating stomatal impressions directly onto archivable slides^[J]. *Agronomy Journal*, 2009, 101(1): 232-236
- [10] Eisele J F, Fäßler F, Bürgel P F, et al. A rapid and simple method for microscopy-based stomata analyses^[J]. *PloS one*, 2016, 11(10): e0164576.
- [11] Bourdais G, McLachlan D H, Rickett L M, et al. The use of quantitative imaging to investigate

- regulators of membrane trafficking in Arabidopsis stomatal closure^[J]. *Traffic*, 2019, 20(2): 168-180.
- [12] Millstead L, Jayakody H, Patel H, et al. Accelerating automated stomata analysis through simplified sample collection and imaging techniques^[J]. *Frontiers in plant science*, 2020, 11: 580389.
- [13] Song W, Li J, Li K, et al. An automatic method for stomatal pore detection and measurement in microscope images of plant leaf based on a convolutional neural network model^[J]. *Forests*, 2020, 11(9): 954.
- [14] Sun Z, Song Y, Li Q, et al. An integrated method for tracking and monitoring stomata dynamics from microscope videos^[J]. *Plant Phenomics*, 2021.
- [15] Sanyal P, Bhattacharya U, Bandyopadhyay S K. Analysis of SEM images of stomata of different tomato cultivars based on morphological features^[C]//2008 Second Asia International Conference on Modelling & Simulation (AMS). IEEE, 2008: 890-894.
- [16] Martins G A, Soares A M, Barbosa J P R A D, et al. Stomatal density distribution patterns in leaves of the Jatobá (*Hymenaea courbaril* L.)^[J]. *Trees*, 2012, 26: 571-579.
- [17] Laga H, Shahinnia F, Fleury D. Image-based plant stomata phenotyping^[C]//2014 13th International Conference on Control Automation Robotics & Vision (ICARCV). IEEE, 2014: 217-222.
- [18] Higaki T, Kutsuna N, Hasezawa S. CARTA-based semi-automatic detection of stomatal regions on an Arabidopsis cotyledon surface^[J]. *Plant morphology*, 2014, 26(1): 9-12.
- [19] Duarte K T N, De Carvalho M A G, Martins P S. Segmenting high-quality digital images of stomata using the wavelet spot detection and the watershed transform^[C]//International Conference on Computer Vision Theory and Applications. SciTePress, 2017, 5: 540-547.
- [20] Liu S, Tang J, Petrie P, et al. A fast method to measure stomatal aperture by MSER on smart mobile phone^[C]//Applied industrial optics: spectroscopy, imaging and metrology. Optica Publishing Group, 2016: AIW2B. 2.
- [21] Jayakody H, Liu S, Whitty M, et al. Microscope image based fully automated stomata detection

- and pore measurement method for grapevines^[J]. *Plant methods*, 2017, 13: 1-12.
- [22] Toda Y, Toh S, Bourdais G, et al. DeepStomata: facial recognition technology for automated stomatal aperture measurement^[J]. *BioRxiv*, 2018: 365098. ^[13] Fetter K C, Eberhardt S, Barclay R S, et al. StomataCounter: a neural network for automatic stomata identification and counting^[J]. *New Phytologist*, 2019, 223(3): 1671-1681.
- [23] Aono A H, Nagai J S, Dickel G S M, et al. A stomata classification and detection system in microscope images of maize cultivars^[J]. *PLoS One*, 2021, 16(10): e0258679.
- [24] Fetter K C, Eberhardt S, Barclay R S, et al. StomataCounter: a neural network for automatic stomata identification and counting^[J]. *New Phytologist*, 2019, 223(3): 1671-1681.
- [25] Meeus S, Van den Bulcke J, Wyffels F. From leaf to label: A robust automated workflow for stomata detection^[J]. *Ecology and Evolution*, 2020, 10(17): 9178-9191.
- [26] Cowling S B, Soltani H, Mayes S, et al. Stomata Detector: High-throughput automation of stomata counting in a population of African rice (*Oryza glaberrima*) using transfer learning^[J]. *bioRxiv*, 2021: 2021.12. 01.469618.
- [27] Shi P, Wang L, Niinemets Ü, et al. Estimation of stomatal density of leaves with hierarchical reticulate venation^[J]. *Botany Letters*, 2023, 170(2): 269-277.
- [28] Toda Y, Tameshige T, Tomiyama M, et al. An affordable image-analysis platform to accelerate stomatal phenotyping During microscopic observation^[J]. *Frontiers in plant science*, 2021, 12: 715309.
- [29] Casado-García A, Del-Canto A, Sanz-Saez A, et al. LabelStoma: A tool for stomata detection based on the YOLO algorithm^[J]. *Computers and Electronics in Agriculture*, 2020, 178: 105751.
- [30] Yang X H, Xi Z J, Li J P, et al. Deep transfer learning-based multi-object detection for plant stomata phenotypic traits intelligent recognition^[J]. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2021, 20(1): 321-329.
- [31] Li X, Guo S, Gong L, et al. An automatic plant leaf stoma detection method based on YOLOv5^[J]. *IET Image Processing*, 2023, 17(1): 67-76.
- [32] Zhang F, Ren F, Li J, et al. Automatic stomata recognition and measurement based on improved

- YOLO deep learning model and entropy rate superpixel algorithm^[J]. Ecological Informatics, 2022, 68: 101521.
- [33] LeCun Y, Bengio Y. Convolutional networks for images, speech, and time series^[J]. The handbook of brain theory and neural networks, 1995, 3361(10): 1995.
- [34] Zhang P, Chen H Q, Li Q R. Research on Vehicle Recognition Algorithm based on Convolution Neural Network^[J]. Journal of Physics: Conference Series, 2021, 1865(4):042117.
- [35] 贾世娜. 基于改进 YOLOv5 的小目标检测算法研究^[D]. 南昌: 南昌大学, 2022.
- [36] 周飞燕, 金林鹏, 董 军. 卷积神经网络研究综述^[J]. 计算机学报, 2017, 40(06): 1229-1251.
- [37] Menon A, Mehrotra K, Mohan C K, et al. Characterization of a class of sigmoid functions with applications to neural networks^[J]. Neural networks, 1996, 9(5): 819-835.
- [38] Abdelouahab K, Pelcat M, Berry F. Why tanh is a hardware friendly activation function for cnns^[C]//Proceedings of the 11th international conference on distributed smart cameras. 2017: 199-201
- [39] Rolnick D, Kording K. Reverse-engineering deep relu networks^[C]//International conference on machine learning. PMLR, 2020: 8178-8187.
- [40] Mastromichalakis S. ALReLU: A different approach on Leaky ReLU activation function to improve Neural Networks Performance^[J]. arXiv preprint arXiv:2012.07564, 2020.
- [41] Thakur R S, Yadav R N, Gupta L. PReLU and edge-aware filter-based image denoiser using convolutional neural network^[J]. IET Image Processing, 2020, 14(15): 3869-3879.
- [42] Liu W, Wen Y, Yu Z, et al. Large-margin softmax loss for convolutional neural networks^[J]. arXiv preprint arXiv:1612.02295, 2016.
- [43] Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector^[C]//Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14, 2016: 21-37.
- [44] Whitehill J, Omlin C W. Haar features for FACS AU recognition^[C]//7th international conference on automatic face and gesture recognition (FGR06). IEEE, 2006: 5 pp.-101.

- [45] Redmon J. You only look once: Unified, real-time object detection^[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
- [46] Shuai Q, Wu X. Object detection system based on SSD algorithm^[C]//2020 international conference on culture-oriented science & technology (ICCST). IEEE, 2020: 141-144.
- [47] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation^[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2014: 580-587.
- [48] Purkait P, Zhao C, Zach C. SPP-Net: Deep absolute pose regression with synthetic views^[J]. arXiv preprint arXiv:1712.03452, 2017.
- [49] Ren S, He K, Girshick 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.
- [50] Redmon J, Farhadi A. YOLO9000: better, faster, stronger^[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 7263-7271.
- [51] Farhadi A, Redmon J. Yolov3: An incremental improvement^[C]//Computer vision and pattern recognition. Berlin/Heidelberg, Germany: Springer, 2018, 1804: 1-6.
- [52] Bochkovskiy A, Wang C Y, Liao H Y M. Yolov4: Optimal speed and accuracy of object detection^[J]. arXiv preprint arXiv:2004.10934, 2020.
- [53] Wang C Y, Bochkovskiy A, Liao H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors^[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2023: 7464-7475.
- [54] Fetter K C, Eberhardt S, Barclay R S, et al. StomataCounter: a neural network for automatic stomata identification and counting^[J]. New Phytologist, 2019, 223(3): 1671-1681.
- [55] Everingham M, Van Gool L, Williams C K I, et al. The pascal visual object classes (voc) challenge^[J]. International journal of computer vision, 2010, 88: 303-338.
- [56] Lin T Y, Maire M, Belongie S, et al. Microsoft coco: Common objects in context^[C]//Computer vision—ECCV 2014: 13th European conference, zurich, Switzerland, September 6-12, 2014,

- proceedings, part v 13. Springer International Publishing, 2014: 740-755.
- [57] Qin D, Leichner C, Delakis M, et al. MobileNetV4: universal models for the mobile ecosystem^[C]//European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2024: 78-96.
- [58] Gevorgyan Z. SIoU loss: More powerful learning for bounding box regression^[J]. arXiv preprint arXiv:2205.12740, 2022.
- [59] Wang J, Renninger H J, Ma Q, et al. Measuring stomatal and guard cell metrics for plant physiology and growth using StoManager1^[J]. Plant Physiology, 2024, 195(1): 378-394.

致谢

时光荏苒，入学时的场景仍历历在目，转瞬已至毕业时刻。三年研究生岁月，我于知识汲取、能力锤炼、心智成长等方面均收获颇丰。值此论文完稿之际，谨向一路给予我关怀与帮助的师长、亲友致以最诚挚的感恩。

首先，感恩导师刘垚老师。是您引领我叩开科研之门，从地球陆面系统研究到气孔识别系统探索，每一次难题攻关都离不开您的悉心指导。您事无巨细地答疑解惑，在职业规划、人生方向上倾囊相授宝贵建议。

深切感谢环境学院夏建阳老师。您的谆谆教诲如明灯，提升我的学术涵养，拓宽我的学术视野。科研遇阻时，您一次次耐心指点迷津；即便进展不顺，也始终给予我加油鼓劲，让我在探索之路上始终保有前行的勇气。

同时，诚挚感谢博士生杜莹同学在研究初期提供的创新思路启发，以及慷慨分享其构建的 sto-Du 植物叶表气孔数据集。正是基于这些富有价值的研究构想和高质量数据集，本研究得以顺利开展并奠定了关键的数据基础。杜莹同学的专业领域的洞见与无私帮助，为研究的推进提供了重要支撑，在此致以最诚挚的谢意。

感恩同门师友韩正愿、王永帅、张晨辉、姜琮、李环宇、李来利、沈逸、易朝鹏。科研中的思维碰撞、生活里的暖心相助，让我真切感受到团队的温暖。那些共同拼搏的时光，是研究生生涯中熠熠生辉的记忆。

更要深深感谢我的父母。你们无条件的支持与理解，是我成长路上最坚实的后盾，更是我奋勇前行的动力之源。这份深沉的爱，让我无论何时都能无畏追梦。

最后，向答辩委员会及各位评审老师致以敬意。感谢您在百忙之中审阅论文，您的专业评审与指导，为我学术之路的进一步成长提供了重要指引。

攻读硕士学位期间所获科研成果

论文

[1] Trade-off between efficiency and accuracy in accelerating spin-up of land surface model by machine learning. (To be submitted to Global Change Biology; JCR:Q1)

软件著作权

[1] 植物叶表气孔检测系统 V1.0.0 R11S0486370, 2025.

科研项目

[1] 国家自然科学基金，面上项目“地球系统模式的异常诊断与负载均衡关键技术研究”