DOI: 10.3969/j.issn.2097-583X.2026.04.014　　文章编号：2097-583X(2026)04-0622-08

基于三重注意力机制的多视角茶叶病害检测方法

邵　罕[a,b]，蒋明毅[c]，孙艳歌[c*]，郭华平[c]

（信阳师范大学 a.茶学与食品学院； b.河南省茶树生物学重点实验室；
c.计算机与信息技术学院， 河南 信阳 464000）

摘　要：为精准检测茶叶病害，从3个不同的视角(通道-宽度、通道-高度、宽度-高度)重新解析特征图，提出了一种新型残差多尺度增强注意力模块，有效提取了病害图像的语义信息与空间特性。同时，综合上述3个视角设计了新型三重注意力特征融合增强模块，对通道和空间特征进行深度融合，显著提升了模型对叶片背景与病害的区分能力。实验结果表明：改进后的模型性能表现优异，其平均精度值(mAP)为95.2%，精度为93.1%，召回率为92.0%，优于现有的检测方法。

关键词：茶叶病害检测；目标检测；注意力机制；特征融合

中图分类号：TP391　　**文献标志码：**A　　**开放科学(资源服务)标识码(OSID)：**

Multi-view tea disease detection method based on triple attention mechanism

Shao Han[a,b], Jiang Mingyi[c], Sun Yange[c*], Guo Huaping[c]

(a. School of Tea and Food Science; b. Henan Key Laboratory of Tea Plant Biology;
c. School of Computer and Information Technology, Xinyang Normal University, Xinyang 464000, China)

Abstract: To achieve more accurate detection of tea diseases, a feature map was reexamined from three distinct perspectives: channel-width, channel-height and width-height. Based on those perspectives, a novel residual multi-scale enhanced attention module was proposed, and the semantic information and spatial characteristics of disease images were efficiently extracted. Meanwhile, a new triple attention-based feature fusion enhancement module was designed by integrating the above three perspectives. Channel and spatial features were deeply fused, which significantly improved the model's ability to distinguish between leaf backgrounds and diseases. Experimental results showed that the improved model achieved excellent performance, with a mean average precision (mAP) of 95.2%, a precision of 93.1%, and a recall of 92.0%, which was superior to existing detection methods.

Key words: tea disease detection; object detection; attention mechanism; feature fusion

0　引言

茶叶作为兼具健康价值与经济价值的重要作物，种植广泛且受到全球消费者的青睐[1]。截至2024年，中国茶叶年产量呈现稳步增长趋势，预计年增幅为4.5%～5.0%，产业总产值已突破2200亿元。然而，茶叶易受叶枯病、锈藻病、红点病等多种病害侵袭，这些病害不仅影响产量和质量，还会给茶产业带来巨大经济损失[2]。茶叶病害检测技术能够精准识别病害的类型及严重程度，为制定科学高效的防治方案提供支撑。因此，研发精准、高效的茶叶病害检测方法，对于保障茶产业的可

收稿日期：2024-11-25；**修回日期：**2025-01-07；*.**通信联系人**，E-mail：yangesun@xynu.edu.cn

基金项目：国家自然科学基金项目(62062004)；河南省留学人员科研择优资助项目

作者简介：邵　罕(1981—)，男，河南驻马店人，讲师，硕士，主要从事茶叶病虫害研究；孙艳歌(1982—)，女，河南平顶山人，副教授，博士，主要从事深度学习、机器学习与数据挖掘研究。

引用格式：邵罕，蒋明毅，孙艳歌，等.基于三重注意力机制的多视角茶叶病害检测方法[J].信阳师范大学学报(自然科学版)，2026，39(4)：622-629.

Shao Han, Jiang Mingyi, Sun Yange, et al. Multi-view tea disease detection method based on triple attention mechanism[J]. Journal of Xinyang Normal University (Natural Science Edition), 2026, 39(4): 622-629.

持续发展与经济稳定至关重要。

传统的茶叶病害检测主要依赖人工方法[3-4],但这类方法存在误诊率高、检测效率低的问题。随着深度学习的发展,其在目标检测领域的应用得到了研究人员的广泛关注与探索。深度学习驱动的目标检测方法可分为单阶段检测器与双阶段检测器。单阶段检测器通过单次回归直接预测目标边界框与类别,具备检测速度快的优势,如Wang等[5]利用YOLOv5s算法通过通道剪枝来压缩模型体积,提出了一种针对小尺寸苹果果实的精确检测方法;Sozzi等[6]评估了多种卷积神经网络模型(YOLOv3、YOLOv3-tiny、YOLOv4、YOLOv4-tiny、YOLOv5x、YOLOv5s),用于葡萄串的实时检测与计数。双阶段检测器(如Faster R-CNN[7]),首先生成目标候选区域,再进行分类与回归操作,通常具有更高的检测精度。然而,上述方法均仅依赖卷积神经网络进行特征提取,仍存在明显局限性:一是对全局上下文信息的捕获能力不足,二是在复杂场景下对细粒度特征的处理效果有限。注意力机制的引入恰好弥补了这些缺陷,通过动态调整特征权重,强化模型对关键区域的关注,从而提升了目标检测的性能与精度。

注意力机制已成功应用于目标检测领域[8],具备快速、精准捕获复杂场景下特征信息的优势。基于注意力机制的茶叶病害检测技术,大致可分为基于通道注意力的方法与基于空间注意力的方法两类。基于通道注意力机制的方法,通过捕获图像中的关键通道特征,提升检测的准确率和鲁棒性。Liu等[9]引入通道注意力机制,为单个通道分配权重并与原始病害图像相乘,在强化显著特征的同时抑制无关的背景噪声;Zhang等[10]提出基于多通道自动定向循环机制的番茄叶片病害检测方法,通过提取并重用多尺度特征,解决番茄叶片病害的相似性难题;Gao等[11]提出双分支高效通道注意力,在提升植物病害识别准确率的同时,最大限度减少无关特征的提取。基于空间注意力的叶片病害检测方法,核心在于分析并突出图像中的特定区域与位置,增强对局部结构及上下文信息的理解。Wang等[12]利用空间注意力建立不同区域间的相关性,强化关键区域在特征图中的重要性;Tang等[13]将多尺度空间注意力块融入全卷积神经网络,有效缓解空间信息丢失问题。Wang等[14]进一步提出融合相关性空间注意力的模型,通过构建不同区域间的相关性来提升关键区域的特征权重。然而现有方法仍存在明显局限性:一方面,部分方法虽能有效提取空间信息,但在通道与空间维度的特征交互方面有所欠缺;另一方面,多数方法仅聚焦单一视角的信息提取(如仅关注空间或通道特征),难以全面应对茶叶病害检测中的复杂性与多样性问题。

融合空间与通道维度的注意力机制,能够更全面地捕获目标特征、提高辨识精度,同时增强模型在复杂背景下的鲁棒性与稳定性。Sun等[15]提出改进的茶叶病害检测网络TCSA,用于自然场景下茶叶病害的自动检测与识别。该方法通过融合创新的通道与空间注意力模块,助力网络筛选高效特征,减少冗余信息干扰。Xie等[16]提出基于YOLOv8s的茶叶病害检测模型,集成可变形卷积、注意力机制及改进的空间金字塔池化模块,通过抑制无关因素干扰与实现多特征有效融合,强化模型对复杂目标及挑战性背景的处理能力。然而,现有融合通道与空间注意力权重的方法,虽然能获取丰富的特征表示,但在提取病害特征时,通道与空间维度的深层交互仍存在明显不足。

为解决上述通道与空间维度交互的局限性,提出了一种新型茶叶病害检测方法,旨在通过更深层次的特征挖掘与跨维度信息融合,全面捕获病害图像中的关键特征。具体而言,设计一种新的残差多尺度增强注意力模块(Residual Multi-scale Enhancement Attention Module,RMEA),不仅能够提取病害图像的语义信息,还能捕捉其空间特性,从而更全面地表征病害特征。同时,综合通道-宽度、通道-高度、宽度-高度等3个视角,构建基于三重注意力特征融合增强模块(Triple Attention-based Feature Fusion Enhancement Module,TAFFEM),通过深度融合通道与空间特征,进一步提升模型对叶片背景与病害的区分能力。这种多视角、多尺度的特征融合策略,显著提升了病害检测模型的性能与可靠性,有效应对茶叶病害尺度多变、自然场景背景复杂导致的检测准确率偏低等挑战。通过对比实验与消融实验,系统验证所提方法的优异性与有效性。

1　基于三重注意力机制的茶叶病害检测方法

为了解决茶叶病害目标检测中尺度多变和背景复杂等问题,下面提出一种基于三重注意力机

制的多视角茶叶病害检测模型。它由 3 个部分组成:骨干网络、颈部和头部,网络总结构如图 1 所示。骨干网络由多尺度卷积块和残差多尺度增强注意力模块组成,以捕捉多尺度特征。颈部网络采用自上而下/自下而上的方式融合骨干网络生成的多尺度特征,以帮助捕捉茶叶病害的全局信息和细节特征(分布模式、小斑点和边缘细节)。此外,提出了一种新颖的基于三重注意力的特征融合增强模块,并将其作为插件在每次特征融合操作后加入,以消除背景噪声并突出病害特征。头部由 3 个并行卷积分支组成,以检测不同大小的茶叶病害。

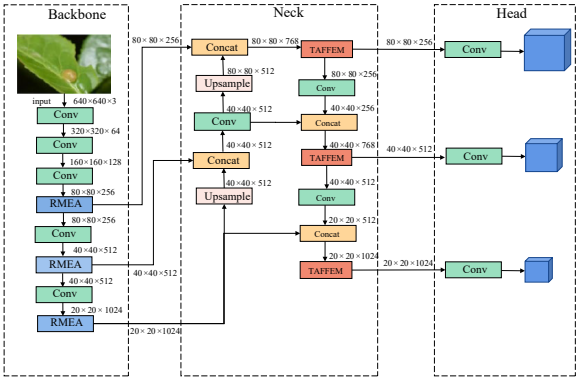


图 1　基于三重注意力机制的茶叶病害目标检测结构图

Fig. 1　Structure diagram of tea disease target detection based on triple attention mechanism

1.1　骨干网络

骨干网络用于提取病害的多尺度粗特征。选择 CSPdarkNet53[17] 作为骨干框架来提取茶叶病害的多尺度特征。此外,将基于倒残差移动自注意力块(IRMB)[18] 的残差多尺度增强注意力(RMEA)模块插入到 CSPdarkNet53 中,以解决 CSPdarkNet53 在整合全局信息和处理长距离依赖方面的局限性并增强其提取多尺度特征的能力,如图 2 所示。

图 2(a)显示了以 IRMB 为基础的 RMEA 整体架构。IRMB 模块与传统的残差块相比不仅能够优化计算效率,同时还能增强模型提取茶叶病害多尺度特征的能力。在 IRMB 中,1×1 卷积用于压缩和扩展通道数量,以优化计算效率;3×3 深度可分离卷积用于捕捉空间特征,而注意力机制则用于捕捉特征之间的全局依赖性。所提出的 RMEA 则是在 IRMB 的基础上加入了全新的注意力机制 EMA,从而进一步增强模型在多尺度特征

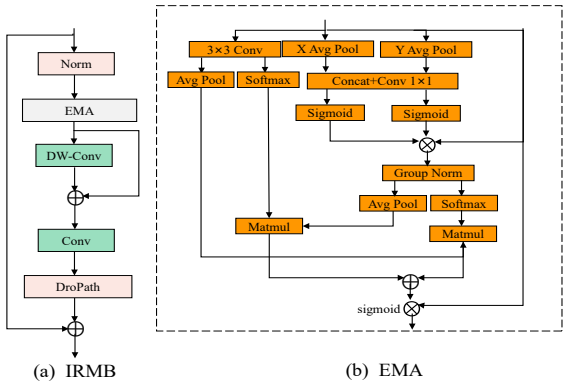


图 2　残差多尺度增强注意力模块

Fig. 2　Residual multi-scale enhancement attention module(RMEA)

提取和自适应权重计算方面的能力。如图 2(b)所示,注意力机制 EMA 模块通过通道-高度和通道-宽度两个视角进行池化操作,捕捉多尺度信息。池化后的特征首先经过 1×1 卷积融合后,先进行组归一化,再通过 sigmoid 函数激活。随后,3×3 卷积进一步提取多尺度特征。接着,自适应池化和 softmax 被用于计算注意力权重,并通过矩阵乘法将其应用于特征图,从而增强多尺度特征表达。将 RMEA 插入 CSPdarkNet53 结构后,使模型通过多尺度特征捕捉与自适应权重计算,显著提升了特征表达能力和整体性能,增强了模型提取茶叶病害多尺度特征的能力。

1.2　颈部

颈部网络通过自上而下(或自下而上)的方式反复融合多尺度茶叶病害特征图。此外,在颈部网络中使用了基于三重注意力的特征融合增强模块(TAFFEM)来增强融合特征的上下文语义,并消除复杂背景中的冗余和噪声,如图 3 所示。

1.2.1　基于三重注意力的特征融合增强模块

该特征融合增强模块基于三重注意力机制,包含三重注意力和融合增强注意力两个部分。三重注意力分别从通道-宽度、通道-高度以及宽度-高度等 3 个视角计算相关性,并分配注意力权重,从而实现通道与空间维度的有机交互。这种机制更有效地捕捉茶叶病害相关特征,同时抑制无关背景特征。此外,模块中引入了融合增强注意力机制,进一步整合通道特征与空间信息,提升特征表达能力。

三重注意力模块的具体结构如图 3 所示,该模块由 3 个并行分支构成,旨在捕捉输入张量 x 在不

同维度上的交互信息。给定输入张量$\boldsymbol{x}$(尺寸为$C\times H\times W$),首先将$\boldsymbol{x}$传递给3个分支,每个分支专门负责在特定维度上建立特征关联。

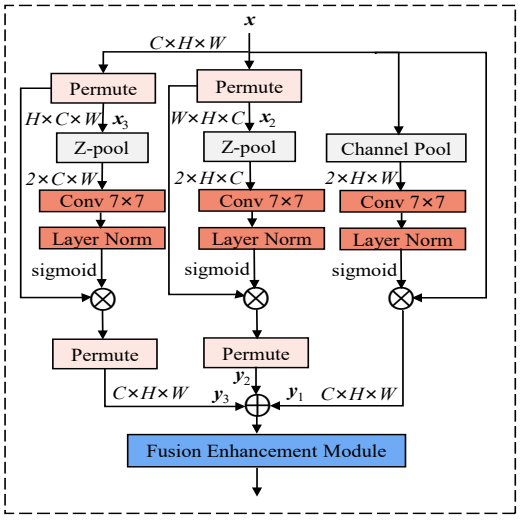


图3　基于三重注意力的特征融合增强模块

Fig. 3　Feature fusion enhancement module based on triple attention

第一个分支的运算流程如下:首先,对输入张量$\boldsymbol{x}$在通道维度C上执行高度-宽度方向池化,将其降维至$2\times H\times W$;随后,通过3个7×7卷积层、Layer Norm层、sigmoid激活函数,生成对应的注意力权重;最终,该权重与原始输入张量$\boldsymbol{x}$进行逐元素相乘,具体如式(1)所示。

$$\boldsymbol{y}_1=\boldsymbol{x}\otimes\delta(\mathrm{LN}(\mathrm{Conv}_{7\times 7}(\mathrm{channelpool}(\boldsymbol{x})))\times 3), \tag{1}$$

式中:$\otimes$表示逐元素相乘。

第二个分支的运算流程如下:首先将输入张量$\boldsymbol{x}$绕H轴逆时针旋转90°,得到张量$\boldsymbol{x}_2$(尺寸为$W\times H\times C$);随后对$\boldsymbol{x}$在通道维度C上执行通道-高度方向的Z-pool池化操作,将宽度维度W压缩至2,具体如式(2)所示;接着通过3个7×7卷积层、Layer Norm层及sigmoid激活函数,生成注意力权重;最终将该权重与$\boldsymbol{x}_2$进行加权相乘,再沿W轴顺时针旋转90°得到输出$\boldsymbol{y}_2$,确保其与原始输入$\boldsymbol{x}$的尺寸保持一致,具体如式(3)所示。

$$\boldsymbol{x}_2=\mathrm{Permute}\left(\text{Z-pool}(\boldsymbol{x})\right), \tag{2}$$

$$\boldsymbol{y}_2=\boldsymbol{x}_2\otimes\delta\left(\mathrm{LN}\left(\left(\mathrm{Conv}_{7\times 7}(\boldsymbol{x}_2)\right)\times 3\right)\right)。 \tag{3}$$

第三个分支的运算流程如下:首先,将输入张量$\boldsymbol{x}$绕W轴逆时针旋转90°,得到张量$\boldsymbol{x}_3$(尺寸为$H\times C\times W$);随后,对$\boldsymbol{x}$在通道维度C上进行通道-宽度方向的Z-pool池化操作,将高度维度H压缩至2,如式(4)所示;接着,通过3个卷积层、Layer Norm层及sigmoid激活函数,生成注意力权重;最终,将该权重与$\boldsymbol{x}_3$进行加权相乘,再沿H轴顺时针旋转90°得到输出$\boldsymbol{y}_3$,确保其与原始输入$\boldsymbol{x}$的尺寸保持一致,具体如式(5)所示。

$$\boldsymbol{x}_3=\mathrm{Permute}\left(\text{Z-pool}(\boldsymbol{x})\right), \tag{4}$$

$$\boldsymbol{y}_3=\boldsymbol{x}_3\otimes\delta\left(\mathrm{LN}\left(\left(\mathrm{Conv}_{7\times 7}(\boldsymbol{x}_3)\right)\times 3\right)\right)。 \tag{5}$$

综合3个分支的输出结果,可得到不同维度下建立的特征关联信息,将三者相加,形成三重注意力模块的最终输出,具体如式(6)所示。该设计能够帮助模型更高效地捕捉特征的全局语义信息,实现通道与空间维度的跨维度交互,进而提升对目标区域的关注度。

$$\boldsymbol{y}=\frac{1}{3}(\boldsymbol{y}_1\oplus\boldsymbol{y}_2\oplus\boldsymbol{y}_3), \tag{6}$$

式中:$\oplus$表示逐元素相加。

融合增强注意力模块的具体结构如图4所示,由通道注意力与空间注意力协同融合构建而成。其中,通道注意力模块通过自适应学习获取各特征通道的权重系数,对不同通道的特征信息进行加权融合,从而自适应地调整每个通道在特征表示中的重要性占比。

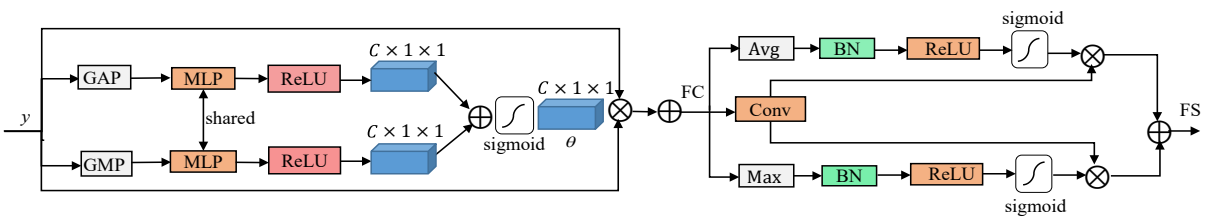


图4　融合增强模块

Fig. 4　Fusion enhancement module

如图 4 所示,给定的输入 $\boldsymbol{y}$ 的大小为 $H\times W$,通道数为 C。

首先,使用平均池化 $P_{\text{avg}}(\boldsymbol{y})$ 与最大池化 $P_{\max}(\boldsymbol{y})$ 提取各通道的全局特征信息,从而在多尺度特征图中突出茶叶病害的重要特征并滤除背景噪声。然后,利用 2 个分支的多层感知机(MLP)提取通道注意力,并加强不同通道信息之间的关注度。随后,将 2 个 MLP 分支的输出结果相加,经 sigmoid 激活函数生成注意力系数 $\boldsymbol{\theta}$。最后,将 $\boldsymbol{\theta}$ 与原始输入 $\boldsymbol{y}$ 进行逐元素相乘后再逐元素相加,得到 CA 模块输出的特征图 FC:

$$\text{FC}=\boldsymbol{y}\otimes\boldsymbol{\theta}\oplus\boldsymbol{y}。\tag{7}$$

进入空间注意力模块运算:首先,对输入的 FC 分别执行局部平均池化与局部最大池化操作,全面捕获病害特征的局部信息。接着,通过规范化、ReLU 激活函数及 sigmoid 激活函数运算,生成基于局部平均池化的空间注意力权重 W_{avg} 与基于局部最大池化的空间注意力权重 $W_{\max}$,如式(8)与(9)所示;最后,对特征 FC 进行 3×3 卷积操作,并分别与 W_{avg} 和 $W_{\max}$ 相乘,再相加得到最终的输出 FS,具体计算如式(10)所示:

$$W_{\text{avg}}=\delta\Big(\text{ReLU}\big(\text{BN}\big(\text{avg}(\text{FC})\big)\big)\Big),\tag{8}$$

$$W_{\max}=\delta\Big(\text{ReLU}\big(\text{BN}\big(\max(\text{FC})\big)\big)\Big),\tag{9}$$

$$\text{FS}=\big(W_{\text{avg}}\otimes\text{Conv}(\text{FC})\big)\oplus\big(W_{\max}\otimes\text{Conv}(\text{FC})\big)。\tag{10}$$

1.3　损失函数

模型的损失函数由分类损失与回归损失两部分构成:分类损失采用 VariFocal Loss(VFL)[19],回归损失则结合 CIoU 损失与分布焦点损失(DFL)[20]共同实现。

VFL 的主要改进在于引入不对称加权操作。该损失函数的公式如下:

$$\text{VFL}(p,q)=\begin{cases}-q\big(q\ln(p)+(1-q)\ln(1-p)\big), & q>0,\\ -\alpha p^{g}\ln(1-p), & q=0,\end{cases}\tag{11}$$

式中:p 代表预测的 IoU 感知分类得分,q 代表目标得分。

DFL 的核心思想是基于距离的特征学习,其通过交叉熵损失优化与目标标签 y 最接近的两个位置的概率分布,引导网络快速聚焦于目标位置附近的特征分布,提升对目标边界的回归精度。

$$\text{DFL}(S_i+S_{i+1})=-\big((y_{i+1}-y)\ln(S_i)+(y-y_i)\ln(S_{i+1})\big),\tag{12}$$

$$S_i=\frac{y_{i+1}-y}{y_{i+1}-y_i},\tag{13}$$

$$S_{i+1}=\frac{y-y_i}{y_{i+1}-y_i},\tag{14}$$

式中:y 代表标签,y_i 和 y_{i+1} 代表与标签 y 最接近的两个离散位置($y_i\leqslant y\leqslant y_{i+1}$)。

CIoU 损失在 IoU 的基础上进一步考虑重叠面积、中心点距离与宽高比,以解决叶片间相互遮挡的问题,并且更好地回归到边界框。

$$\text{CIoU}=\text{IoU}-\frac{\rho^2(b,b^{\text{gt}})}{s^2}-\alpha\upsilon,\tag{15}$$

式中:α 是权重系数;b 和 b^{gt} 分别代表预测框与实际框的中心点;ρ 表示欧氏距离度量,即两中心点间的欧氏距离;s 为能同时覆盖预测框与实际框的最小闭包区域的对角线长度;υ 用于表示宽高比的相似性。

2　实验

2.1　实验配置

基于安徽农业大学[21]提供的茶叶病害数据集对所提方法进行性能评估。该数据集包含 776 张分辨率为 906×600 像素的病虫害图像,涵盖茶藻斑病(Als)、茶饼病(Tc)、茶云纹叶枯病(Clb)、茶网饼病(Eb)、茶红锈藻病(Tr)及茶圆赤星病(Rs)等 6 类常见茶叶病害,如图 5 所示。

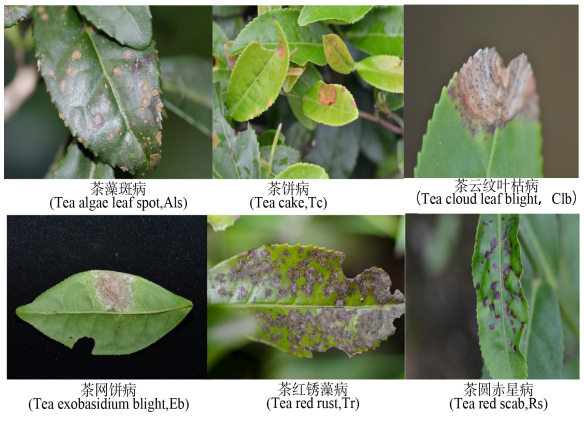


图 5　6 种常见的茶叶病害类型

Fig. 5　Six common types of tea diseases

2.1.1　实验细节

实验基于 Linux 操作系统与 NVIDIA A100 80 GB PCIe GPU 硬件平台完成,使用 Python 3.9

编程语言,依托PyTorch 1.8.2深度学习框架实现模型搭建与训练。训练参数设置为:优化器选用Adam,批次大小(batch size)设为32,初始学习率为0.005,总迭代轮次数为400。

原始数据集未提供茶叶病害目标的边界框信息,导致模型难以精确定位、区分病害区域,无法开展有效的训练与评估。为解决该问题并优化数据集,首先,通过LabelImg标注软件手动标注病害目标边界框,构建标准茶叶病害目标检测数据集,为准确完成检测任务提供数据支撑。其次,为增强模型的鲁棒性并扩充样本量,采用翻转、添加噪声等数据增强策略,数据集的图像数量从776张增至6549张。最后,按8∶1∶1的比例将增强后的数据集划分为训练集、测试集与验证集,用于保持数据分布均衡,提升实验结果的可靠性。

2.1.2　评价指标

采用精度(P)评估模型检测结果的准确性,计算公式如下:

$$P=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}。$$

召回率(R)用以评估模型识别的全面性,计算公式如下:

$$R=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}},$$

式中:TP为模型预测为正的正样本,FP为模型预测为正的负样本,FN为模型预测为负的正样本。

AP用于全面评估模型的检测精度,其公式为:

$$\mathrm{AP}=\int_0^1 P(r)\mathrm{d}r。$$

mAP为AP的均值,能综合反应模型在所有类别上的检测与识别性能,其公式为:

$$\mathrm{mAP}=\sum_{i=1}^{k}\mathrm{AP}(i)/k。$$

2.2　消融实验

实验以CSPdarkNet53为基础骨干网络,通过逐步纳入注意力机制模块构建不同模块变体,结果如表1所示。整体来看,随着子模块的逐步添加,TAFFEM的性能有所提升,并且包含所有模块的模型达到最高准确率。CSPdarkNet53+RMEA模型相较于基础模型,mAP提升0.8个百分点;CSPdarkNet53+TAFFEM模型的mAP较基础模型提升1.6个百分点。性能改进的原因可归结为:RMEA模块能够在骨干网络中从3个不同维度有效捕捉多尺度茶叶病害特征,强化关键病害信息的提取;TAFFEM模块则作用于网络颈部,通过细化颈部特征表达并融合多尺度信息,优化病害目标的特征表征。上述消融实验结果充分验证了所提TAFFEM方法在茶叶病害检测任务中的有效性。

表1　融合增强模块消融实验结果

Tab. 1　Ablation results of the fusion enhancement module

Backbone (CSPdarkNet53)	RMEA	TAFFEM	mAP/%
√	×	×	93.3
√	√	×	94.1
√	×	√	94.9
√	√	√	95.2

2.3　对比实验

为全面验证所提TAFFEM方法的性能优势,将其与SSD[22]、YOLOv5s[23]、Tea-YOLOv8s[16]、AX-RetinaNet[24]、RT-DETR[25]、TCSA[15]等6种主流检测模型,在精度、召回率及mAP方面进行比较,结果如表2及图6—图7所示。

表2呈现了7种模型的综合性能:TAFFEM方法在精度、召回率、mAP指标上均优于其他对比模型。其中,mAP指标的提升尤为突出,相较于经典模型SSD和YOLOv5s,TAFFEM的mAP分别高出6.8和6.6个百分点;即便与Tea-YOLOv8s、AX-RetinaNet、RT-DETR及TCSA方法相比,TAFFEM模型在mAP上也分别提高了1.9、1.8、1.0、0.6个百分点。

表2　不同测试模型的对比结果

Tab. 2　Performance comparison of different test models

模型	P/%	R/%	mAP/%
SSD[22]	86.5	89.1	88.4
YOLOv5s[23]	87.4	82.5	88.6
Tea-YOLOv8s[16]	92.7	89.2	93.3
AX-RetinaNet[24]	91.7	90.8	93.4
RT-DETR[25]	89.6	87.8	94.2
TCSA[15]	92.9	89.6	94.6
TAFFEM	93.1	92.0	95.2

TAFFEM与其他6种模型的mAP曲线变化情况如图6所示。从图6可见,TAFFEM的mAP曲线呈稳定上升趋势,并在150次迭代后超过其他模型。

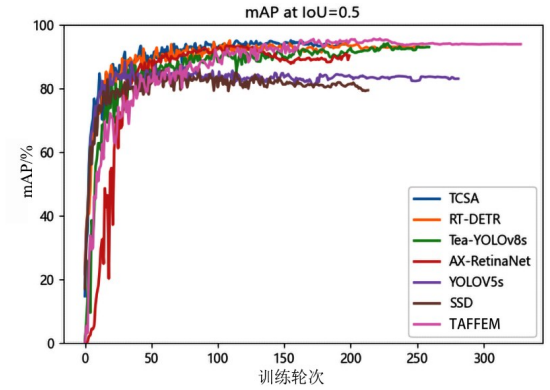


图6　7种不同模型的mAP对比

Fig. 6　Comparison of the mAP variations of seven different models

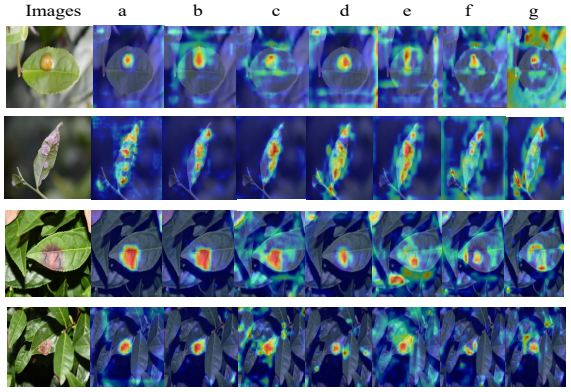


图7　网络热力图可视化识别结果

Fig. 7　Recognition results of network heatmap visualization

注　a:TAFFEM; b:TCSA; c:RT-DETR; d: Tea-YOLOv8s; e: AX-RetinaNet; f: YOLOv5s; g:SSD。

图7展示了TAFFEM、TCSA、Tea-YOLOv8s、AX-RetinaNet、RT-DETR、YOLOv5s、SSD的注意力热力图,显示了识别网络对图像不同区域的关注程度。在这些热力图中,颜色的变化表示识别网络对图像不同区域的注意力程度,其中红色区域表示较高的注意力,蓝色区域表示较低的注意力。图7中的原始病害图像从上到下分别代表了单一病害、多尺度病害、叶片遮挡和复杂背景下的病害。在这些场景中,该方法从通道-宽度、通道-高度、宽度-高度的视角重新提取特征,相较于其他方法,该机制能够更精确地覆盖目标病害区域。

2.4　TAFFEM在不同病害上的表现

使用所提供的数据集评估了TAFFEM在不同病害上的表现,如表3和图8所示。

表3　TAFFEM在不同病害上的检测结果

Tab. 3　Detection results of different diseases in TAFFEM

病害	P/%	R/%	mAP/%
Als	88.3	92.0	90.7
Tc	98.0	95.8	97.0
Clb	92.0	97.0	96.8
Eb	95.2	98.4	98.5
Tr	92.2	94.4	98.8
Rs	86.4	82.8	88.4

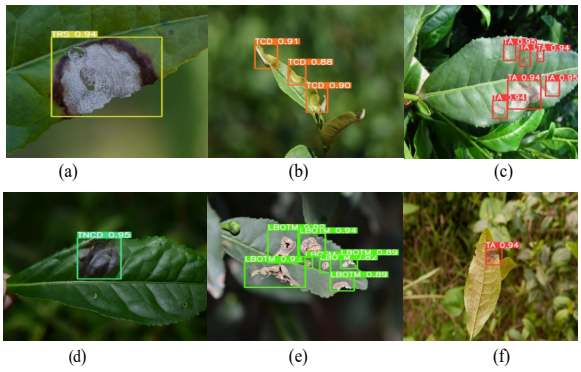


图8　可视化识别结果

Fig. 8　Visual recognition results

表3展示了TAFFEM对不同病害的准确率(P)、召回率(R)和mAP值,而图8则展示了这些病害的可视化结果。TAFFEM在Eb与Tr两种病害的mAP达到最高的98.5%和98.8%。此外,在检测多种茶叶病害(如Als、Tc、Rs和Clb)时也表现出良好的效果,表明该模型在多种条件下具备良好的鲁棒性和适应性。

3　结束语

提出的基于三重注意力机制的茶叶病害检测方法,从通道-高度、通道-宽度、宽度-高度的视角进行多尺度特征提取,全面捕捉茶叶病害的关键信息,有效提升算法在叶片相互遮挡、病害特征多变等复杂场景下的检测性能。该方法实现了95.2%的mAP、93.1%的精度、92.0%的召回率,优于现有的检测方法。

未来研究将聚焦于模型在实际应用中的实时检测性能,特别是如何在时间与计算资源受限的条件下保持高效、精准的病害检测。本研究可为茶叶病害的早期检测提供可靠的技术支持,助力提升茶叶产量与品质。

参考文献:

[1] Fang Jian, Sureda A, Silva A S, et al. Trends of tea in cardiovascular health and disease: a critical review[J]. Trends in Food Science & Technology, 2019, 88: 385-396.

[2] Hu Gensheng, Yang Xiaowei, Zhang Yan, et al. Identification of tea leaf diseases by using an improved deep convolutional neural network[J]. Sustainable Computing: Informatics and Systems, 2019, 24: 100353.

[3] 龙樟, 姜倩, 王健, 等. 茶叶嫩芽视觉识别与采摘点定位方法研究[J]. 传感器与微系统, 2022, 41(2): 39-41, 45.

[4] Zhang Lei, Zou Lang, Wu Chuanyu, et al. Locating famous tea's picking point based on Shi-tomasi algorithm[J]. Computers, Materials & Continua, 2021, 69(1): 1109-1122.

[5] Wang Dandan, He Dongjian. Channel pruned YOLOv5s-based deep learning approach for rapid and accurate apple fruitlet detection before fruit thinning[J]. Biosystems Engineering, 2021, 210: 271-281.

[6] Sozzi M, Cantalamessa S, Cogato A, et al. Automatic bunch detection in white grape varieties using YOLOv3, YOLOv4, and YOLOv5 deep learning algorithms[J]. Agronomy, 2022, 12(2): 319.

[7] Ren Shaoqing, He Kaiming, 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, 2017, 39(6): 1137-1149.

[8] Dai Guowei, Fan Jingchao. An industrial-grade solution for crop disease image detection tasks[J]. Frontiers in Plant Science, 2022, 13: 921057.

[9] Liu Liu, Wang Rujing, Xie Chengjun, et al. PestNet: an end-to-end deep learning approach for large-scale multi-class pest detection and classification[J]. IEEE Access, 2019, 7: 45301-45312.

[10] Zhang Yukai, Huang Shuangjie, Zhou Guoxiong, et al. Identification of tomato leaf diseases based on multi-channel automatic orientation recurrent attention network[J]. Computers and Electronics in Agriculture, 2023, 205: 107605.

[11] Gao Ronghua, Wang Rong, Feng Lu, et al. Dual-branch, efficient, channel attention-based crop disease identification [J]. Computers and Electronics in Agriculture, 2021, 190: 106410.

[12] Wang Xiaotian, Cao Weiqun. Bit-plane and correlation spatial attention modules for plant disease classification[J]. IEEE Access, 2023, 11: 93852-93863.

[13] Tang Ziyi, Zhang Ruimao, Peng Zhanglin, et al. Multi-stage spatio-temporal aggregation transformer for video person re-identification[J]. IEEE Transactions on Multimedia, 2023, 25: 7917-7929.

[14] Wang Fangyuan, Wang Rujing, Xie Chengjun, et al. Fusing multiscale context-aware information representation for automatic in-field pest detection and recognition[J]. Computers and Electronics in Agriculture, 2020, 169: 105222.

[15] Sun Yange, Jiang Mingyi, Guo Huaping, et al. Multiscale tea disease detection with channel-spatial attention[J]. Sustainability, 2024, 16(16): 6859.

[16] Xie Shuang, Sun Hongwei. Tea-YOLOv8s: a tea bud detection model based on deep learning and computer vision[J]. Sensors, 2023, 23(14): 6576.

[17] Wang C Y, Mark Liao H Y, Wu Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops(CVPRW). June 14-19, 2020, Seattle, WA, USA. Piscataway, NJ:IEEE, 2020: 1571-1580.

[18] Zhang Jiangning, Li Xiangtai, Li Jian, et al. Rethinking mobile block for efficient attention-based models[C]//2023 IEEE/CVF International Conference on Computer Vision(ICCV). October 1-6, 2023, Paris, France. Piscataway, NJ:IEEE, 2023: 1389-1400.

[19] Zhang Haoyang, Wang Ying, Dayoub F, et al. VarifocalNet: an IoU-aware dense object detector[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 20-25, 2021, Nashville, TN, USA. Piscataway, NJ:IEEE, 2021: 8510-8519.

[20] Li Xiang, Lü Chengqi, Wang Wenhai, et al. Generalized focal loss: towards efficient representation learning for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(3): 3139-3153.

[21] Sun Yunyun, Jiang Zhaohui, Zhang Liping, et al. SLIC_SVM based leaf diseases saliency map extraction of tea plant [J]. Computers and Electronics in Agriculture, 2019,157: 102-109.

[22] Liu Wei, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[C]//Leibe B, Matas J, Sebe N, et al. Computer Vision-ECCV 2016. Cham: Springer International Publishing, 2016: 21-37.

[23] Jocher G. YOLOv5 by Ultralytics[EB/OL]. [2025-01-07]. https://github.com/ultralytics/yolov5.

[24] Bao Wenxia, Fan Tao, Hu Gensheng, et al. Detection and identification of tea leaf diseases based on AX-RetinaNet [J]. Scientific Reports, 2022, 12: 2183.

[25] Zhao Yian, Lü Wenyu, Xu Shangliang, et al. DETRs beat YOLOs on real-time object detection[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 16-22, 2024, Seattle, WA, USA. Piscataway, NJ:IEEE, 2024: 16965-16974.

责任编校:郭红建　申晓伟