



融合通道和空间注意力的茶叶病害目标检测方法

孙艳歌, 蒋明毅, 冯岩, 郭华平, 张莉, 吴飞

引用本文:

孙艳歌, 蒋明毅, 冯岩, 郭华平, 张莉, 吴飞. 融合通道和空间注意力的茶叶病害目标检测方法[J]. 信阳师范学院学报自然科学版, 2025, 38(1): 59–65. doi: 10.3969/j.issn.2097–583X.2025.01.008

SUN Yange, JIANG Mingyi, FENG Yan, GUO Huaping, ZHANG Li, WU Fei. Tea disease detection method using fusion of channel and spatial attentions[J]. *Journal of Xinyang Normal University (Natural Science Edition)*, 2025, 38(1): 59–65. doi: 10.3969/j.issn.2097–583X.2025.01.008

在线阅读 View online: <https://doi.org/10.3969/j.issn.2097–583X.2025.01.008>

您可能感兴趣的其他文章

Articles you may be interested in

一种加权词向量的混合网络文本情感分析方法

A Mixed Network Text Sentiment Analysis Method Based on Weighted Word Vectors

信阳师范学院学报自然科学版, 2021, 34(3): 472–477. <https://doi.org/10.3969/j.issn.1003–0972.2021.03.022>

一种基于改进SSD的烟雾实时检测模型

Real-time Smoke Detection Model Based on Improved SSD

信阳师范学院学报自然科学版, 2020, 33(2): 305–311. <https://doi.org/10.3969/j.issn.1003–0972.2020.02.021>

基于语义关联性特征融合的大数据挖掘方法

Large Data Mining Method Based on Semantic Correlation Feature Fusion

信阳师范学院学报自然科学版, 2019, 32(1): 141–145. <https://doi.org/10.3969/j.issn.1003–0972.2019.01.025>

基于递归卷积神经网络的行人检测方法

Pedestrian Detection Method Based on Recurrent Convolutional Neural Networks

信阳师范学院学报自然科学版, 2021, 34(4): 655–660. <https://doi.org/10.3969/j.issn.1003–0972.2021.04.025>

基于随机森林和多特征融合的青苹果图像分割

Green Apple Image Segmentation Based on Multi-feature Fusion and Random Forest

信阳师范学院学报自然科学版, 2018, 31(4): 681–686. <https://doi.org/10.3969/j.issn.1003–0972.2018.04.032>

融合通道和空间注意力的茶叶病害目标检测方法

孙艳歌*, 蒋明毅, 冯岩, 郭华平, 张莉, 吴飞

(信阳师范大学 计算机与信息技术学院, 河南 信阳 464000)

摘要:准确、快速检测茶叶病害对茶叶产业经济意义重大。然而,病害叶片尺度变化、相互遮挡、复杂背景等问题大大降低了检测精度。提出了一种融合通道和空间注意力的茶叶病害目标检测方法,以实现准确、高效的病害检测。首先,利用多层的自注意力机制提取茶叶病害的多尺度特征,以获取茶叶图像在不同尺度下的局部细节特征;其次,引入新的通道注意力模块以提取更丰富的通道信息,同时抑制复杂背景噪声并增强模型特征表示能力。此外,设计一种新的空间注意力模块进一步提取特征的空间关系、减少冗余信息并优化计算开销。实验结果表明,融合通道和空间注意力的茶叶病害检测方法能应对叶片的尺度变化、相互遮挡和背景复杂等挑战。

关键词:茶叶病害检测;目标检测;通道注意力;空间注意力;特征融合

中图分类号:TP391.4

文献标识码:A

开放科学(资源服务)标识码(OSID):



Tea disease detection method using fusion of channel and spatial attentions

SUN Yange*, JIANG Mingyi, FENG Yan, GUO Huaping, ZHANG Li, WU Fei

(College of Computer and Information Technology, Xinyang Normal University, Xinyang 464000, China)

Abstract: Accurate and rapid detection of tea disease is of great economic significance to the tea industry. However, the problems such as scale variation, mutual occlusion and complex background of disease-affected leaves greatly reduce the prediction accuracy of the detector. A tea disease target detection method, integrating channel-spatial attention, was proposed to achieve accurate and efficient disease detection. Firstly, the multi-layer self-attention mechanism was used to extract multi-scale features of tea diseases to obtain local detail features of tea images at different scales. Secondly, the new channel attention module was introduced to extract richer channel information, suppress complex background noise and enhance the ability of model feature representation. In addition, a new spatial attention module was also proposed to further extract the spatial relationship of features, reduce redundant information and optimize computational overhead. The experimental results showed that the tea pest detection method with channel-spatial attention could cope with the challenges of leaf scale changes, mutual occlusion and complex background.

Key words: tea disease detection; object detection; channel attention; spatial attention; feature fusion

0 引言

茶叶病害检测是茶叶产业中一项具有挑战性的任务,具有重要的实际意义^[1]。全国产茶大省早期茶园面积已达到 2.93 万 km²,茶叶产量 261.6

万 t,茶叶总产值超过 2000 亿元。由于主要受到叶枯病等茶叶疾病的影响,茶叶产量每年减少约 20%。因此准确、快速地检测茶叶病害对于茶叶产业的经济至关重要。

然而,茶叶病害的复杂性和多样性给其检测带

收稿日期:2023-08-28;修回日期:2024-03-12;*.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);河南省自然科学基金项目(232300421167);河南省留学人员科研择优资助项目

作者简介: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:孙艳歌,蒋明毅,冯岩,等.融合通道和空间注意力的茶叶病害目标检测方法[J].信阳师范大学学报(自然科学版),2025,38(1):59-65.

SUN Yange, JIANG Mingyi, FENG Yan, et al. Tea disease detection method using fusion of channel and spatial attentions [J]. Journal of Xinyang Normal University (Natural Science Edition), 2025, 38(1): 59-65.

来了挑战。茶叶叶片具有尺度不一、相互遮挡和背景复杂等特点,这增加了检测器准确检测的难度。因此,设计准确、高效的茶叶病害检测方法对于茶叶产业来说至关重要,它有助于提高茶叶产量和质量,降低农业生产成本,提升茶叶产业的竞争力和可持续发展能力。

注意力机制通过模拟人类视觉系统在面对复杂场景时快速准确地捕捉关键信息的能力,在病害识别任务中得到成功应用。注意力机制通常分为空间注意力、通道注意力和自注意力等^[2-3]。LIU等^[4]利用空间注意力模块增强了模型的特征提取能力。WANG等^[5]提出的DeepPest采用通道注意力机制,获取每个通道的权重并与原图相乘,突出重要特征,从而抑制了无用的噪声。王泽钧等^[6]利用通道和空间注意力机制,增强了有用特征信息的权重并减弱了噪声和干扰信息的权重。ZHA等^[7]通过增加轻量级坐标注意力机制,提升了移动模型的特征提取能力。王志强等^[8]提出了一种基于坐标注意力机制的轻量级网络模型,使模型更加专注于叶片中的病害特征,提高了对不同病害种类的识别准确率。尽管上述研究工作利用注意力机制,提升了模型提取特征并消除冗余的能力,但没有全面综合多种注意力机制提取复杂背景下的多尺度特征,因此在尺度多变、复杂的病害检测场景中,检测精度并不高。综合通道和空间两种注意力有助于模型捕获多尺度特征,并消除背景噪声。

为此,针对茶叶叶片尺度不一、相互遮挡、背景复杂的特点,本文提出一个融合通道和空间注意力的茶叶病害目标检测模型来解决该问题。首先,利用Transformer提取茶叶病害的多尺度特征,提取不同尺度下的局部细节特征,并更好地捕捉图像中的全局和局部特征;然后,提出一种新的通道注意力模块(Channel Attention, CA),自适应计算特征图权重,与特征图加权相乘,以聚焦相关的通道信息,消除复杂背景噪声信息。此外,又设计了一种空间注意力模块(Spatial Attention, SA),突出复杂背景下多尺度病害目标的空间位置,以滤除复杂背景的干扰,同时优化计算开销,增加模型的准确性和鲁棒性。

1 相关工作

在计算机视觉任务中,目标检测方法变得越来越重要。许多方法被用于解决病害问题,本文将这些方法分为两大类:(1)基于卷积神经网络的方法,

(2)基于注意力机制的方法。

1.1 基于卷积神经网络的方法

传统的病害检测方法通常使用决策树、支持向量机(SVM)、K-means聚类等,需要针对不同类型的病害人工设计特征提取的方法。然而,传统的图像识别方法中,手动特征工程是一项复杂的任务。因此,卷积神经网络(Convolutional Neural Networks, CNN)凭借其强大的特征提取能力,广泛应用于植物病害目标检测。ZHOU等^[9]提出了一种基于Faster R-CNN和FCM-KM融合的水稻病害检测算法。TIAN等^[10]设计了一个基于YOLOv3^[11]的系统,可以实时检测果园中3个不同阶段的苹果。在病害目标检测中表现较好的基于卷积神经网络的模型还有基于YOLOv4^[12]的改进方法等。然而这些方法没有特别关注检测对象的通道特征和位置关系,从而可能引入一些冗余和背景噪声。

1.2 基于注意力机制的方法

近年来,注意力机制在病害检测任务中得到成功应用。WANG等^[13]引入通道注意力机制,为单个通道分配权重,并与原始疾病图像相乘,在强调显著特征的同时抑制不相关的背景噪声。ZHANG等^[14]提出了一种基于多通道自动定向循环机制的番茄叶片病害检测方法。ZENG等^[15]引入了一种自聚焦卷积神经网络,该网络融合了自注意力机制,用于从作物病斑中提取相关特征以进行病害识别。GAO等^[16]引入了一种高效的作物病害识别模型,在提高植物病害识别准确率的同时最小化提取无关特征。尽管越来越多的研究利用注意力机制进行病害检测,但很少应用在茶叶病害检测中。而且,它们大多只关注一种注意力机制,没有全面结合多种注意力机制以达到更好的效果。

2 融合通道和空间注意力的目标检测方法

为了解决茶叶病害目标检测中尺度多变、相互遮挡、背景复杂等问题,本文提出一种融合通道和空间注意力的茶叶病害目标检测模型。它由3个部分组成:骨干网、颈部和头部,网络总结构如图1所示。在骨干网中,多层的自注意力机制提取病害图像的多尺度特征。颈部的综合通道-空间注意力模块自适应计算权重并与特征图加权相乘,从而消除背景噪声并突出病害特征。头部由3个并行的卷积分支组成,分别检测不同大小的茶叶病害特征。

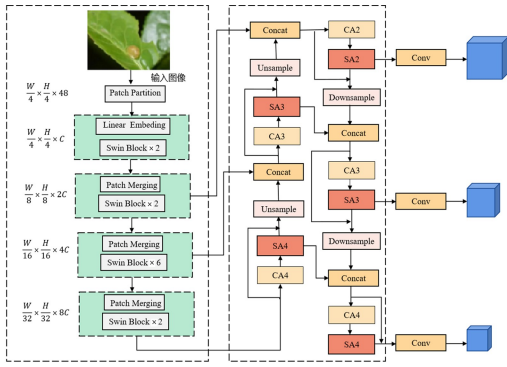


图 1 融合通道和空间注意力机制的茶叶病害目标检测结构图

Fig. 1 The fusion of channel and spatial attentions for tea disease detection network

2.1 骨干网

自注意力机制能够通过计算 Query、Key 和 Value 之间的归一化距离, 来捕获像素之间的长距离依赖关系, 以表达每个注意力分数与所有其他注意力分数之间的全局关系。然而, 全局注意力机制计算成本高, 而局部注意力机制限制了不同 token 之间的相互作用, 需要更多的计算块来实现全局注意力。基于自注意力机制的方法 Swin Transformer^[17] 引入跨窗口注意力机制, 能够大大减少全局自注意力的计算成本。其层次化结构使得模型能够在不同层级上捕获多尺度特征, 从而更好地解决茶叶病害类型多样, 尺度变化大等问题。因此, 选择 Swin Transformer 作为骨干网以有效提取茶叶病害的多尺度特征, 提高模型的准确性。

Swin Transformer 由 4 个不同阶段的模块组成。如图 1 所示, 给定输入图像大小 $H \times W \times C$ 被处理以获得 token, 它被用作后续层次结构的输入。然后, 为了生成分层特征表示, 基于 Swin Transformer 的分层网络使用具有固定核大小和步长的卷积对每个阶段的输出进行下采样, 将 H 和 W 减少一半, 同时将通道数量增加一倍。阶段 i 的输出特征图有 $(H/2i + 1) \times (W/2i + 1)$ 个 tokens。其中, Swin Transformer 模块通过归一化和 MLP 提高模型的稳定性和表达能力, 并且引入了窗口多头注意力机制 (W-MSA) 和移动窗口注意力机制 (SW-MSA), 将图像划分为若干窗口。每个窗口内的像素之间进行自注意力计算, 从而减少自注意力的计算复杂度, 提高了模型的计算效率, 如图 2 所示。

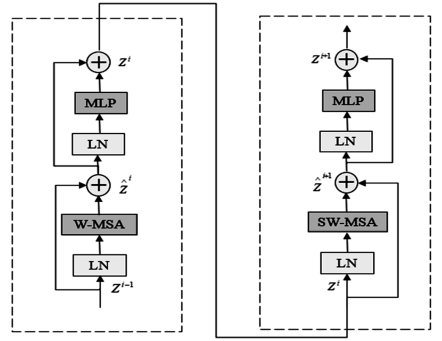


图 2 Swin Transformer 模块
Fig. 2 Swin transformer block

Swin Transformer 模块可以在不同尺度上捕获图像中的相关信息, 有利于提取多尺度特征。Swin Transformer 模块的输入输出公式如下:

$$\begin{cases} \tilde{z}^i = F_{W-MSA}(F_{LN}(z^{i-1})) + z^{i-1}, \\ z^i = F_{MLP}(F_{LN}(\tilde{z}^i)) + \tilde{z}^i, \\ \tilde{z}^{i+1} = F_{SW-MSA}(F_{LN}(z^i)) + z^i, \\ z^{i+1} = F_{MLP}(F_{LN}(\tilde{z}^{i+1})) + \tilde{z}^{i+1}. \end{cases} \quad (1)$$

2.2 颈部

颈部通过自顶向下和自底向上的反复特征融合来提取和融合茶叶病害特征图, 去除叶片遮挡和复杂背景带来的影响。

2.2.1 通道注意力模块

在颈部设计 CA 模块对通道特征进行提取, 从而突出病害中重要特征信息, 同时滤除多尺度特征图中的背景噪声, 如图 3 所示。

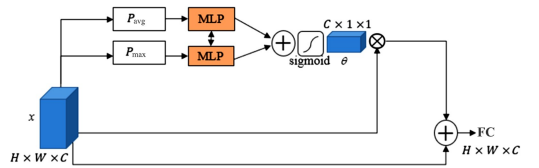


图 3 通道注意力模块

Fig. 3 Channel attention module

给定输入 x 大小为 $H \times W$ 并且通道数为 c 。首先, 同时使用平均池化 $P_{avg}(x)$ 、最大池化 $P_{max}(x)$ 获取每个通道的全局信息, 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, 利用两个连接的多层感知机 (MLP) 提取通道注意力, 并加强不同通道信息之间的关注度。最终, 两个分支分别获得的结果相加, 并经过 sigmoid 函数得到注意力系数 θ , 再与 x 进行逐元素相乘再相加的操作后, 得到 CA 模块输出的特征图 FC。公式如下:

$$FC = x \otimes \theta \oplus x, \quad (2)$$

式中: \otimes 表示逐元素相乘, \oplus 表示逐元素相加。

2.2.2 空间注意力模块

在颈部设计一种具有空间缩减注意力层(Spatial Decrease Attention, SDA)的 SA 模块来优化计算开销。如图 4 所示, SDA 接收 Q 、 K 、 V 作为输入, 并输出一个精细的特征。

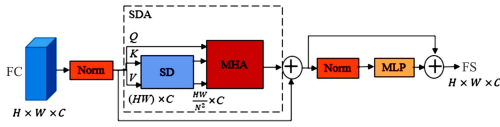


图 4 空间注意力模块

Fig. 4 Spatial attention module

SDA 降低了 K 的空间尺度和进行注意力计算前的 V , 大大减少了计算和内存的开销。其中 SDA 可表示如下:

$$SDA(Q, K, V) = \text{Con}(h_0, \dots, h_N)W^O, \quad (3)$$

$$h_i = \text{Attention}(QW_i^Q, SD(K)W_i^K, SD(V)W_i^V), \quad (4)$$

式中: N 是注意力层的头数; $\text{Con}(\cdot)$ 为将不同注意力头部的结果 h_i ($i=0, 1, \dots, N$) 连接在一起, 以实现多头注意力的计算; W^O 、 W_i^Q 、 W_i^K 、 W_i^V 分别代表在多头注意力计算中不同的权重。SD(\cdot) 是对输入序列(即 K 或 V) 进行降维, 公式如下:

$$SD(FC) = \text{Norm}(\text{Reshape}(FC, D)W^S), \quad (5)$$

式中: FC 表示输入的通道特征图; D 表示注意力层的缩减率; $\text{Reshape}(FC, D)$ 是将输入序列重塑大小的操作; W^S 为线性投影, 可将输入序列降维为 C ; $\text{Norm}(\cdot)$ 为归一化层。

空间注意力的输出是:

$$FS = \text{MLP}(\text{Norm}(SDA(\text{Norm}(FC))))), \quad (6)$$

式中: FS 为空间特征图。

最终, 通过残差连接融合后的特征图表示为:

$$FSC = FS \otimes x, \quad (7)$$

式中: FSC 为最终融合后的特征图。

2.3 损失函数

对于所提出的模型, 采用联合损失函数来训练分类、定位和置信度。

$$L = L_{\text{cls}} + L_{\text{loc}} + L_{\text{conf}}, \quad (8)$$

式中: L_{cls} 为分类损失, L_{loc} 为定位损失, L_{conf} 为置信度损失。

3 实验

3.1 实验配置

本实验在安徽农业大学所提供的茶叶病害数据集^[18]上对该方法进行评估。该数据集包含 776

张图片大小为 906×600 像素的病害图像, 根据茶叶病害常见类型分为茶藻斑病(Als)、茶饼病(Tc)、茶云纹叶枯病(Clb)、茶网饼病(Eb)、茶红锈藻病(Tr)和茶圆赤星病(Rs)等 6 大类, 如图 5 所示。

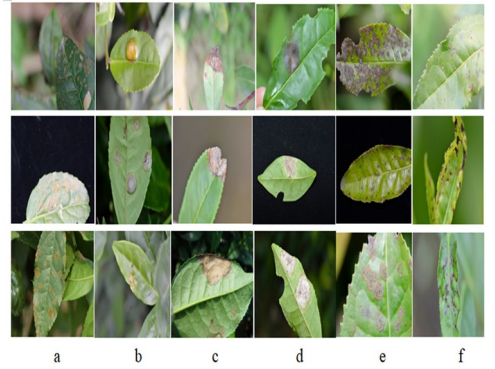


图 5 6 种常见的茶叶病害类型

a: 茶藻斑病; b: 茶饼病; c: 茶云纹叶枯病;
d: 茶网饼病; e: 茶红锈藻病; f: 茶圆赤星病

Fig. 5 Six common types of tea pests and diseases

a: Tea algae leaf spot(Als); b: Tea cake(Tc);
c: Tea cloud leaf blight(Clb);
d: Tea exobasidium blight(Eb);
e: Tea red rust(Tr); f: Tea red scab(Rs)

3.1.1 实验细节

实验是在 PyTorch 深度学习框架下, 基于 Intel(R) Xeon(R) CPU E5-2650 v4 @ 2.20 GHz 处理器, 4 块 TITAN Xp GPU (12 GM 内存) 和 Ubuntu 操作系统的服务器上实现的。Adam 作为优化器训练网络参数, 其中初始学习率、批处理大小、迭代次数分别设置为 1×10^{-5} 、8 和 300 fps。

3.1.2 评价指标

采用平均精度(mAP)为评估本文算法的检测精度的度量, 计算公式如下:

$$P = \frac{TP}{TP + FP}, \quad (14)$$

式中: TP 为模型预测为正的样本, FP 为模型预测为负的正样本。召回率用以评估模型识别的全面性, 计算公式如下:

$$R = \frac{TN}{TN + FN}, \quad (15)$$

式中: TN 为模型预测为负的负样本, FN 为模型预测为正的负样本。

F1 分数为模型的性能提供了全面的评估。其公式为:

$$F1 = \frac{2TP}{2TP + FP + FN}. \quad (16)$$

AP 提供对模型检测精度的全面评估。其公

式为:

$$AP = \int_0^1 P(r) dr. \quad (17)$$

mAP 指标为 AP 值的均值, 可以更好地平衡 P 和 R 之间的关系, 从而反映模型的检测和识别性能。其公式为:

$$mAP = \sum_{i=1}^k AP(i) / k. \quad (18)$$

3.2 对比分析

选取 7 种模型: SSD^[19]、AX-RetinaNet^[20]、YOLOv3^[13]、YOLOv4^[14]、Centernet^[21]、M2det^[22]、EfficientNet^[23], 作为对比来验证本文提出的方法的性能。如表 1 所示, 本文方法在精度方面优于其他病害检测模型。具体而言, 与 SSD^[19]、YOLOv4^[14]、Centernet^[21]、M2det^[22] 相比, 模型精度分别提高了 9.2%、1.0%、2.0% 和 10.0%。此外, 该模型显示出更高的召回率, 表明注意力机制通过关注局部区域内部和之间的细粒度细节, 实现更高的准确性和特征提取能力。

表 1 不同测试模型性能对比 %

Tab. 1 Performance comparison of different test models

模型	P	R	$F1$	mAP
SSD ^[19]	89.5	83.7	0.865	88.4
AX-RetinaNet ^[20]	91.2	84.5	0.889	85.1
YOLOv3 ^[13]	92.4	69.5	0.789	84.8
YOLOv4 ^[14]	97.7	57.7	0.827	79.3
Centernet ^[21]	96.7	57.3	0.723	79.3
M2det ^[22]	88.7	77.5	0.827	84.1
EfficientNet ^[23]	94.5	81.2	0.873	84.3
本文方法	98.7	84.5	0.910	92.8

如图 6 所示, 该模型输出测试结果的准确率图, 包括预测框和预测分数。第 1 列图像显示, 在简单背景下, 多个叶片形状清晰且准确检测出不同尺度的病害。第 2 列显示了在叶片相互遮挡的复杂背景下茶叶病害的检测。

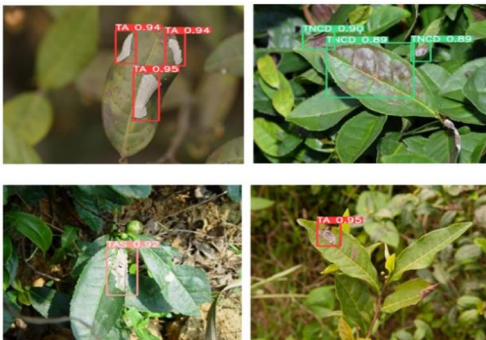


图 6 可视化检测结果

Fig. 6 Visualized detection results

根据以上分析, 表明注意力机制的引入能有效提升茶叶病害目标检测的性能。通过将 Swin Transformer 与通道注意力和空间注意力相结合, 可以在尺度多变、背景复杂的情况下, 准确检测出茶叶病害。

3.3 不同病害检测效果比较

本文方法与 SSD^[19]、AX-RetinaNet^[20]、YOLOv3^[13]、YOLOv4^[14]、Centernet^[21]、M2det^[22] 和 EfficientNet^[23] 在相同数据集上进行比较, 结果如表 2 所示。

表 2 不同病害检测平均精度对比结果 %

Tab. 2 Comparison of different disease detection mAP

模型	Als	Tc	Clb	Eb	Tr	Rs
SSD ^[19]	89.5	83.7	86.5	85.4	75.8	90.4
AX-RetinaNet ^[20]	90.5	84.2	85.4	88.1	80.5	89.4
YOLOv3 ^[13]	90.4	70.5	81.8	86.4	77.8	85.4
YOLOv4 ^[14]	91.5	71.7	82.7	85.3	70.5	88.7
Centernet ^[21]	92.4	84.3	72.3	80.2	84.5	91.2
M2det ^[22]	88.7	77.5	82.7	78.5	77.4	89.0
EfficientNet ^[23]	90.1	81.2	87.3	80.4	80.6	88.9
本文方法	94.7	93.0	93.4	89.7	89.4	94.5

对比结果表明: 所提出的方法对于分布广、尺度变化大、背景干扰复杂的病害图像, 如茶藻斑病 (Als)、茶圆红星病 (Rs) 等具有较好的检测效果, mAP 分别为 94.7% 和 94.5%。这可能由于该模型利用跨窗口注意力机制有效地扩大了注意力区域, 并且综合注意力机制提取多尺度特征中的通道信息和空间信息。

3.4 消融实验

为了验证改进模块的有效性, 对模型进行了消融实验, 结果如表 3 所示。原始模型 Swin Transformer 单独检测各种疾病的准确率并不高。本文方法准确率明显高于原始模型 Swin Transformer, 最高平均准确率达到 92.8%。SA 模块单独加入颈部结构后在大规模可变茶叶病害 Tc 上取得了 7.7% 的性能提升, 这可能归因于 SD 注意力机制在保证识别精度的同时大大减少计算量, 优化了模型对大规模疾病的检测性能。最后, 融合 CA 和 SA 使模型的整体 mAP 提升了 7.4%, 与单独添加 CA 模块和 SA 模块相比, 分别提升了 4.3% 和 2.4%。该模型显著提高茶藻斑病和茶圆红星病两种病害

类型的数据,对分布广泛、背景复杂的病害图像具有较高的鲁棒性。

表 3 不同模型消融实验结果 %

Tab. 3 Ablation experiment results of different models

方法	Als	Tc	Clb	Eb	Tr	Rs	mAP
Swin	89.4	85.0	88.9	74.5	75.6	89.9	85.4
Swin+CA	92.7	89.0	91.2	87.8	88.3	94.5	88.5
Swin+SA	93.5	92.7	92.0	85.4	87.0	91.3	90.4
本文方法	94.7	93.0	93.4	89.7	89.4	94.5	92.8

4 结论

提出了一种基于通道和空间注意力融合的茶叶病害目标检测方法。通道注意力用于提取最相关的通道特征,从而降低冗余信息和背景噪声。空间注意力目的是捕获空间上的重要区域,优化计算开销增强模型的鲁棒性和泛化能力。这两种注意力机制的融合使该模型适应多种类型、不规则的茶叶病害。实验结果表明,与以往的网络模型相比,所提出的融合通道和空间注意力的检测方法提高了模型提取复杂背景下茶叶病害图像的能力。

参考文献:

- [1] 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: 100-353.
- [2] 王文冠, 沈建冰, 贾云得. 视觉注意力检测综述[J]. *软件学报*, 2019, 30(2): 416-439.
WANG Wenguan, SHEN Jianbing, JIA Yunde. Review of visual attention detection[J]. *Journal of Software*, 2019, 30(2): 416-439.
- [3] GUO Menghao, XU Tianxing, LIU Jiangjiang, et al. Attention mechanisms in computer vision: A survey[J]. *Computational Visual Media*, 2022, 8(3): 331-368.
- [4] LIU L, WANG R, XIE C, 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.
- [5] WANG Fangyuan, WANG Rujing, XIE Chengjun, et al. Fusing multi-scale context-aware information representation for automatic in-field pest detection and recognition[J]. *Computers and Electronics in Agriculture*, 2020, 169: 105222.
- [6] 王泽钧, 马凤英, 张瑜, 等. 基于注意力机制和多尺度轻量型网络的农作物病害识别[J]. *农业工程学报*, 2022, 38(S1): 176-183.
WANG Zejun, MA Fengying, ZHANG Yu, et al. Crop disease recognition using attention mechanism and multi-scale lightweight network[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2022, 38(S1): 176-183.
- [7] ZHA Mingfeng, QIAN Wenbin, YI Wenlong, et al. A lightweight Yolov4-based forestry pest detection method using coordinate attention and feature fusion[J]. *Entropy*, 2021, 23(12): 1587.
- [8] 王志强, 于雪莹, 杨晓婧, 等. 基于 WGAN 和 MCA-MobileNet 的番茄叶片病害识别[J]. *农业机械学报*, 2023, 54(5): 244-252.
WANG Zhiqiang, YU Xueying, YANG Xiaojing, et al. Tomato leaf diseases recognition based on WGA and MCA-mobilenet[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023, 54(5): 244-252.
- [9] ZHOU Guoxiong, ZHANG Wenzhuo, CHEN Aibin, et al. Rapid detection of rice disease based on FCM-KM and faster R-CNN fusion[J]. *IEEE Access*, 2019, 7: 143190-143206.
- [10] TIAN Yunong, YANG Guodong, WANG Zhe, et al. Apple detection during different growth stages in orchards using the improved YOLO-V3 model[J]. *Computers and Electronics in Agriculture*, 2019, 157: 417-426.
- [11] REDMON J, FARHADI A. Yolov3: An incremental improvement [J]. *arXiv preprint*, 2018, arXiv:1804.02767:1-6.
- [12] BOHCENKO A, WANG C Y, LIAO H Y M, et al. Yolov4: Optimal speed and accuracy of object detection[J]. *arXiv preprint*, 2020, arXiv:2004.10934: 1-17.
- [13] WANG X, CAO W. Bit-plane and correlation spatial attention modules for plant disease classification[J]. *IEEE Access*, 2023, 11: 93852-93863.
- [14] 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.

- [15] ZENG Weihui, LI Miao. Crop leaf disease recognition based on self-attention convolutional neural network[J]. Computers and Electronics in Agriculture, 2020, 172: 105341.
- [16] 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.
- [17] LIU Ze, LIN Yutong, CAO Yue, et al. Swin transformer: Hierarchical vision transformer using shifted windows [C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada: IEEE, 2021: 9992-10002.
- [18] THOLKAPIYAN M, ARUNA DEVI B, BHATT D, et al. Performance analysis of rice plant diseases identification and classification methodology[J]. Wireless Personal Communications, 2023, 130: 1317-1341.
- [19] LIU Wei, ANGUELOV D, ERHAN D, et al. Ssd: Single shot multibox detector[C]//14th European Conference, Amsterdam, The Netherlands: Springer, 2016: 21-37.
- [20] 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.
- [21] ZHOU X, KOLTUN V, KRÄHENBÜHL P. Tracking objects as points[C]// ECCV 2020. Lecture Notes in Computer Science, Cham: Springer, 2020:474-564.
- [22] ZHAO Qijie, SHENG Tao, WANG Yongtao, et al. M2det: A single-shot object detector based on multi-level feature pyramid network [C]//Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference and Ninth AAAI Symposium on Educational Advances in Artificial Intelligence, Honolulu, Hawaii, USA: AAAI Press, 2019: 9259-9266.
- [23] TAN Mingxing, PANG Ruoming, LE Q V. Efficientdet: Scalable and efficient object detection[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA: IEEE, 2020: 10778-10787.

责任编辑:郭红建