doi:10.6041/j.issn.1000-1298.2025.01.036

主干信息共享与多感受野特征自适应融合的作物叶片等级和病害识别方法

罗　洋[1]　何自芬[2]　张印辉[2]　陈光晨[2]

(1.红塔烟草(集团)有限责任公司昭通卷烟厂, 昭通 657000; 2.昆明理工大学机电工程学院, 昆明 650500)

摘要: 作物叶片等级和病害的快速准确识别对开发农业智能设备以促进农产品精细化管理有着重要意义。针对作物叶片等级和病害识别准确率低、成本高等问题,提出主干信息共享与多感受野特征自适应融合的作物叶片等级和病害识别算法(Crop leaf grade and disease recognition network, CLGDRNet)。首先,CLGDRNet 采用 CSPNet、GhostNet、ShuffleNet 构建特征提取主干网络,同时将 CSPNet、GhostNet、ShuffleNet 所提取的特征信息进行共享以达到信息互补的目的;其次,设计多感受野特征自适应融合模块(Multi-receptive field feature adaptive fusion module, MRFA),将不同感受野特征图进行自适应加权融合,在增强模型局部感受野的同时突出有效通道信息;最后,提出一种深层梯度跨空间学习高效多尺度注意力模块(Efficient multi-scale attention mechanism with deep gradient cross-space learning, EMAD),将 EMAD 嵌入模型的颈部以获取深层梯度信息和目标坐标信息并跨空间融合不同尺度的上下文信息,使模型能够对深层特征图产生更精确的像素级关注。实验结果表明,CLGDRNet 在初烤烟叶分级数据集(Tobacco leaf grading dataset, TLGD)上识别精度 mAP@0.5 和 mAP@0.5:0.95 分别达到 85.0%、76.1%,在苹果叶病害数据集(Apple leaf disease dataset, ALDD)上识别精度 mAP@0.5 和 mAP@0.5:0.95 分别达到 97.6%、74.2%。相较于多种先进目标检测算法,CLGDRNet 具有更高的识别精度,可为高精度作物叶片等级和病害识别提供关键技术支撑。

关键词: 作物叶片等级; 作物叶片病害; 目标检测; 信息共享; 多感受野特征融合

中图分类号: TP391; S24　　**文献标识码**: A　　**文章编号**: 1000-1298(2025)01-0377-11　　**OSID**:

Crop Leaf Grade and Disease Recognition Method Based on Backbone Information Sharing and Multi-receptive Field Feature Adaptive Fusion

LUO Yang[1]　HE Zifen[2]　ZHANG Yinhui[2]　CHEN Guangchen[2]

(1. *Zhaotong Cigarette Factory of Hongta Tobacco (Group) Co., Ltd., Zhaotong* 657000, *China*

2. *Faculty of Mechanical and Electrical Engineering, Kunming University of Science and Technology, Kunming* 650500, *China*)

Abstract: Rapid and accurate recognition of crop leaf grade and disease is integral to the advancement of intelligent equipment for promoting refined management of agricultural products. In response to the problems of low accuracy and high cost of crop leaf grade and disease recognition, a crop leaf grade and disease recognition network (CLGDRNet) was proposed based on backbone information sharing and multi-receptive field feature adaptive fusion. Firstly, CSPNet, GhostNet and ShuffleNet were utilized to build a feature extraction backbone, and the feature information extracted by CSPNet, GhostNet and ShuffleNet was shared to achieve the purpose of information complementarity. Secondly, a multi-receptive field feature adaptive fusion module (MRFA) was designed, and the different receptive field feature maps were adaptively weighted and fused to highlight the effective channel information while enhancing the local receptive fields. Finally, an efficient multi-scale attention mechanism with deep gradient cross-space learning (EMAD) was proposed, the EMAD was embedded in the neck to obtain the deep gradient information and the target coordinate information, in addition, the context information of different scales was fused across the space, which could generate more accurate pixel-level attention to the deep feature

收稿日期: 2024-01-18　修回日期: 2024-02-15

基金项目: 国家自然科学基金项目(62171206、62061022)和中国烟草总公司云南省公司烟叶智能分级项目(HZ2021K0462A)

作者简介: 罗洋(1996—),男,助理工程师,主要从事图像处理和机器视觉研究,E-mail: luoy1990229542@163.com

通信作者: 张印辉(1977—),男,教授,博士,主要从事机器视觉和机器智能研究,E-mail: zhangyinhui@kust.edu.cn

map. The experimental results showed that the recognition accuracy of mAP@0.5 and mAP@0.5:0.95 for tobacco leaf grading dataset (TLGD) achieved 85.0% and 76.1%, respectively, and 97.6% and 74.2% for apple leaf disease dataset (ALDD), respectively. Compared with a variety of advanced target detection algorithms, CLGDRNet achieved higher recognition accuracy and faster recognition speed, which could provide key technical support for high-precision fine recognition of crop leaves.

Key words: crop leaf grade; crop leaf disease; object detection; information sharing; multi-receptive field feature fusion

0 引言

随着农业自动化技术的不断发展,基于机器学习的自动化视觉系统已经被提出并应用于农业生产中,如作物生长监测[1-2]、作物产量预测[3-5]、作物病害检测[6-7]、作物分类[8]、田间杂草检测[9]等。作物叶片可以反映出作物种类、生长状况和健康状况等信息,对于具有经济价值的作物叶片,需要根据叶片的生长位置、颜色、形状、成熟度等特征进行等级划分以保证后期加工质量,而作物叶片病害的准确识别对提高作物产量以及病害预防有着重要意义。因此,作物叶片等级和病害的精确识别对作物叶片精细化管理至关重要。

传统视觉方法根据叶片颜色、纹理、成熟度以及几何形状等特征进行组合并设计合理阈值以实现作物叶片等级和病害识别[10-12]。传统视觉方法无法提取目标深层语义特征信息且需要人工选择特征和调整大量参数,限制了识别精度和效率。

对于结构特征差异明显的不同等级作物叶片和病害类别之间的识别任务,传统视觉方法能够取得较为理想的识别效果,但在农业生产中,根据作物叶片质量,同种作物同一部位的叶片(例如初烤烟叶、茶叶)需要继续进行等级区分,同一部位的叶片之间几何特征差异较小,传统视觉方法无法胜任此任务,利用深度学习的方式对相似度较高但等级不同的叶片进行识别,可有效获取不同等级叶片的深层语义信息,找出叶片间的内在联系,从而提升识别精度[13-16]。

在深度学习算法中,YOLO 系列[17-21]目标检测模型凭借速度快且精度高的优点被广泛应用于工业缺陷检测[22]、交通车辆检测[23]、行人检测[24]等方面。目前,YOLO 系列模型已经更新到 YOLO v8,将 YOLO v8 用于作物叶片等级和病害识别任务时,由于特征提取主干网络结构简单,随着网络深度的增加,部分对作物叶片识别有益的深层特征信息随之丢失;模型局部感受野较小且单一,缺少上下文感知;此外,模型缺乏对目标局部特征信息的关注,导致其无法有效区分相似度较高的不同等级作物叶片和结构特征差异较小的病害类别。

针对上述问题,本文提出主干信息共享与多感受野特征自适应融合的作物叶片等级和病害识别算法(CLGDRNet),该算法使用 CSPNet[25]、GhostNet[26]、ShuffleNet[27]作为模型主干网络,在主干间进行特征信息共享进而保留丰富深层特征信息;设计多感受野特征自适应融合模块(MRFA)并将其嵌入网络的主干次末端,增强模型局部感受野的同时突出有效通道信息;在网络的颈部嵌入深层梯度跨空间学习高效多尺度注意力机制(EMAD),以获取深层梯度流信息并跨空间融合不同尺度的上下文信息,使模型能够对深层特征图产生更精确的像素级关注。最后建立初烤烟叶分级数据集(TLGD)和苹果叶病害数据集(ALDD)并进行消融实验,以证明本文所提方法的有效性。

1 材料和方法

1.1 数据集建立

分别建立初烤烟叶分级数据集和苹果叶病害数据集,使用两个不同农业领域的数据集进行实验以充分验证本文所提方法的有效性和模型的泛化性能。初烤烟叶图像采集自 5XYZ-9C 型智能分级设备,如图 1 所示。智能分级设备使用黑色传送带将初烤烟叶运送至装有 CA050-11U 型工业面阵相机和固定光源的暗室中进行图像采集,工业面阵相机固定于暗室顶部,其分辨率为 2 384 像素×1 528 像素,光源设备型号为 XC-BK-650-1100,固定于暗室左右两侧顶部。现行国家标准GB 2635—1992 根

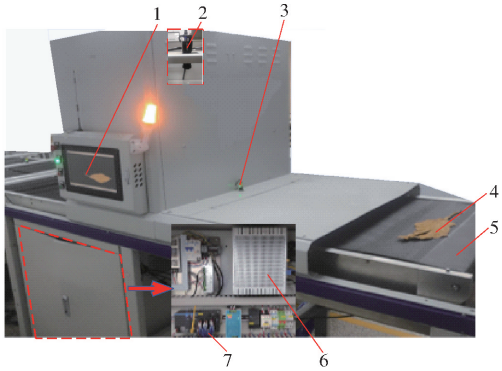
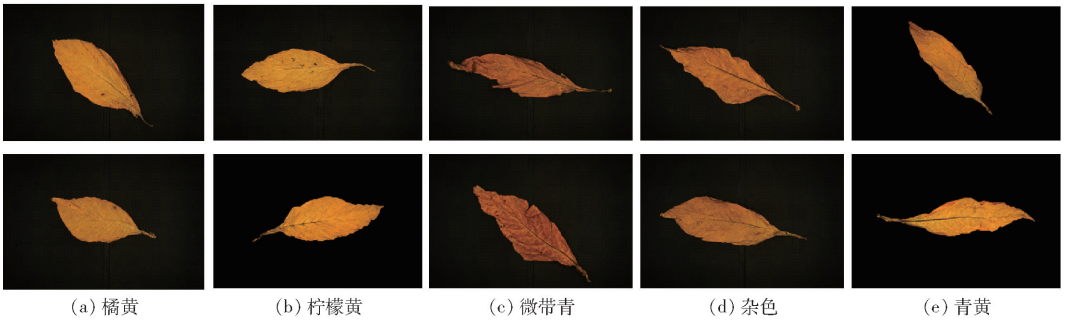


图 1 5XYZ-9C 型智能分级设备

Fig. 1 Intelligent grading equipment 5XYZ-9C

1. 显示屏 2. 工业面阵相机 3. 光电传感器 4. 初烤烟叶 5. 传送带 6. 计算机 7. PLC 控制器

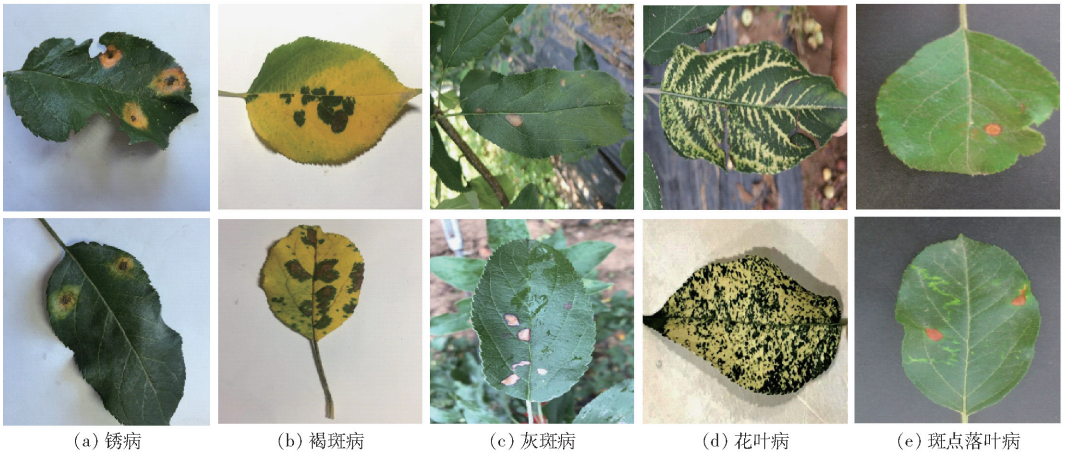
据色度、总体质量、成熟度等特征将初烤烟叶划分为橘黄(F)、柠檬黄(L)、红棕(R)、完熟(H)、杂色(K)、微带青(V)、青黄(GY)和光滑(S)共 8 个等级。由于 R、H、S 3 个等级烟叶在烟叶烤制过程中较少出现，因此仅对 F、L、V、K、GY 5 个等级的初烤烟叶进行研究。由于色度或是杂色斑点在烟叶上分布没有明显的规律且相邻等级的初烤烟叶在颜色特征上差异较小(例如橘黄烟叶和柠檬黄烟叶)，在人工分级中，分级人员只能凭借已有经验进行等级划分，这将导致分级标准不一从而限制人工分级精度和分级效率，因此使用深度学习的方法对初烤烟叶进行等级划分。利用智能分级设备获取 F、L、V、K、GY 5 个等级的初烤烟叶正面图像数据，通过数据筛选，得到高质量初烤烟叶图像 1 766 幅，5 个类别的初烤烟叶图像示例如图 2 所示。苹果叶病害数据集包含斑点落叶病(Alternaria_Boltch)、褐斑病(Brown_Spot)、花叶病(Mosaic)、灰斑病(Grey_spot)、锈病(Rust)5 种常见的苹果叶片病害图像数据共计 1 645 幅，苹果叶病害图像示例如图 3 所示。使用 LabelImg 对 TLGD 和 ALDD 中目标进行锚框标注，标注信息采用 PASCAL VOC 数据集的格式进行保存。将带有标注信息的数据集按比例 4∶1随机划分训练集和测试集，用于网络训练和网络性能测试。



(a) 橘黄　(b) 柠檬黄　(c) 微带青　(d) 杂色　(e) 青黄

图 2　TLGD 中烟叶图像

Fig. 2　Image examples of tobacco leaf in TLGD



(a) 锈病　(b) 褐斑病　(c) 灰斑病　(d) 花叶病　(e) 斑点落叶病

图 3　ALDD 中苹果叶病害图像

Fig. 3　Image examples of apple leaf disease in ALDD

1.2　CLGDRNet 架构

主干信息共享与多感受野特征自适应融合的作物叶片等级和病害识别算法(Crop leaf grade and disease recognition network，CLGDRNet)是在 YOLO v8n 基础上改进得到，其网络结构如图 4 所示。CLGDRNet 由输入、主干、颈部、输出 4 部分构成。在输入部分，模型将输入图像尺寸裁剪为 640 像素×640 像素。在主干部分，为获取更为丰富的深层特征信息，使用 CSPNet、GhostNet、ShuffleNet 作为特征提取主干网络，并在 CSPNet、GhostNet、ShuffleNet 间进行特征信息共享，同时在主干次末端嵌入多感受野特征自适应融合模块(Multi-receptive field feature adaptive fusion module，MRFA)，将具有不同感受野的特征图进行自适应加权融合，增强模型局部感受野的同时突出有效通道信息，此外，使用深层梯度跨空间学习高效多尺度注意力模块(Efficient multi-scale attention mechanism with deep gradient cross-space learning，EMAD)替换颈部的 C2f 模块，获取更多梯度流信息并融合不同尺度的上下文信息，使模型能够对深层特征图产生更精确的像素级关注。输出部分依据无锚框原理来设计检测头以达到独立处理定位、分类和回归任务

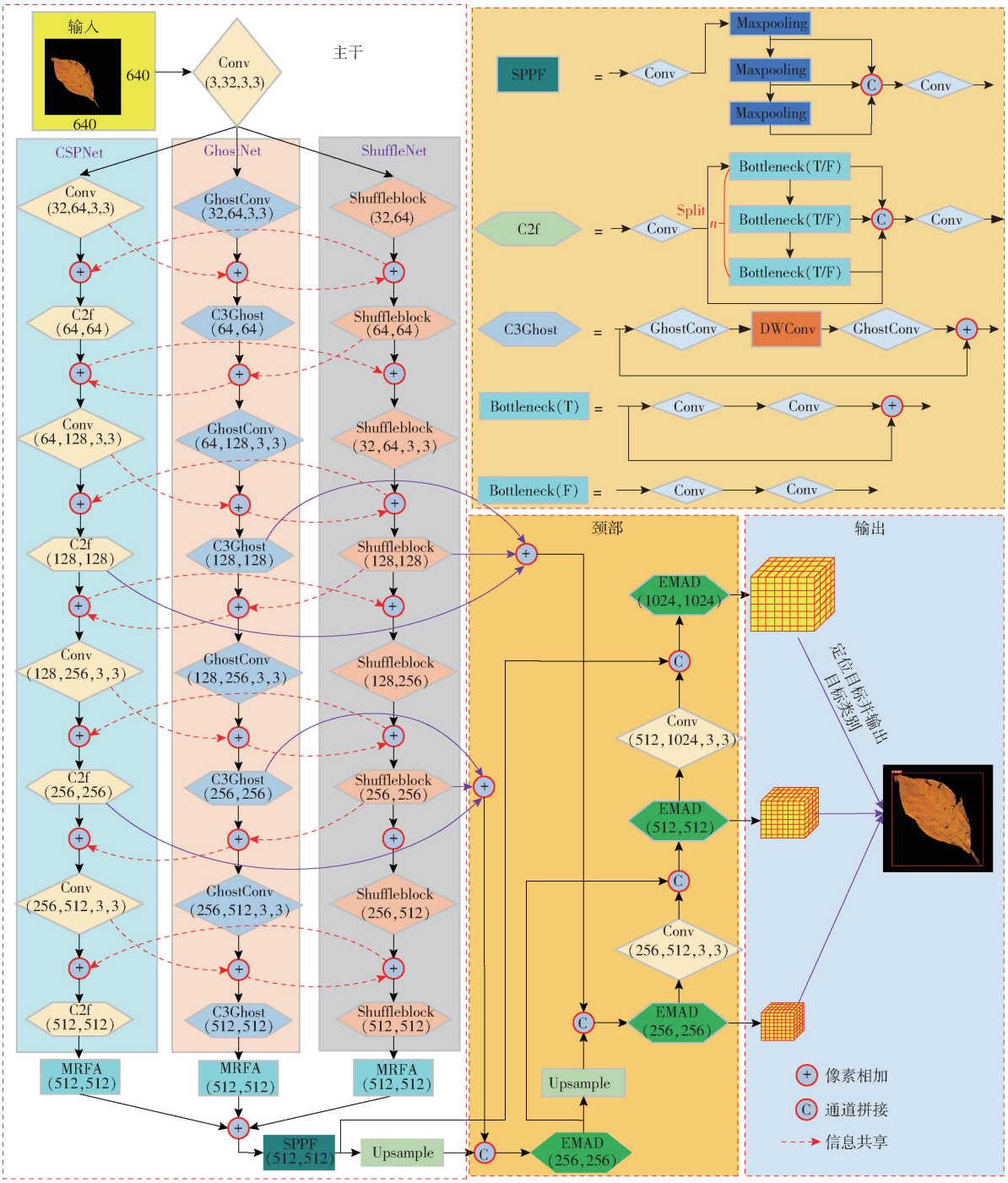


图 4 CLGDRNet 网络结构图

Fig. 4 CLGDRNet structure diagram

的目的,这种设计允许每个分支专注于各自的任务以提高模型整体准确性。

1.3 多感受野特征自适应融合模块

感受野是指卷积神经网络每层输出的特征图上的像素点在原始图像上映射区域的大小[28]。感受野尺寸的求解相当于已知输出特征图大小反向求出输入特征图的过程。计算感受野大小时,常采用自顶向底的方式计算,即先计算最终特征图在前一层上的感受野,然后逐渐传递到第 1 层。

标准卷积感受野计算公式为

$$R_i = (R_{i-1} - 1)s + h \tag{1}$$

式中 R_i——第 i 层感受野

R_{i-1}——第 $i-1$ 层感受野

s——第 i 层卷积步长

h——第 i 层卷积核尺寸

空洞卷积(Atrous convolution,AConv)感受野计算与标准卷积感受野计算相似,所不同的是,空洞卷积的卷积核尺寸 H 需要根据卷积空洞率 r 进行计算,空洞卷积的卷积核尺寸 H 计算公式为

$$H = h + (h-1)(r-1) \tag{2}$$

空洞卷积感受野计算公式为

$$R_i = (R_{i-1} - 1)s + H \tag{3}$$

多感受野特征自适应融合模块(MRFA)结构如图 5 所示。MRFA 中使用 3 层卷积结构，第 1 层和第 3 层为标准卷积，其卷积核为 1×1，用于降低通道数；为获取具有不同感受野的特征图，将 3×3、7×7、11×11 的空洞卷积并行组成第 2 层卷积结构。

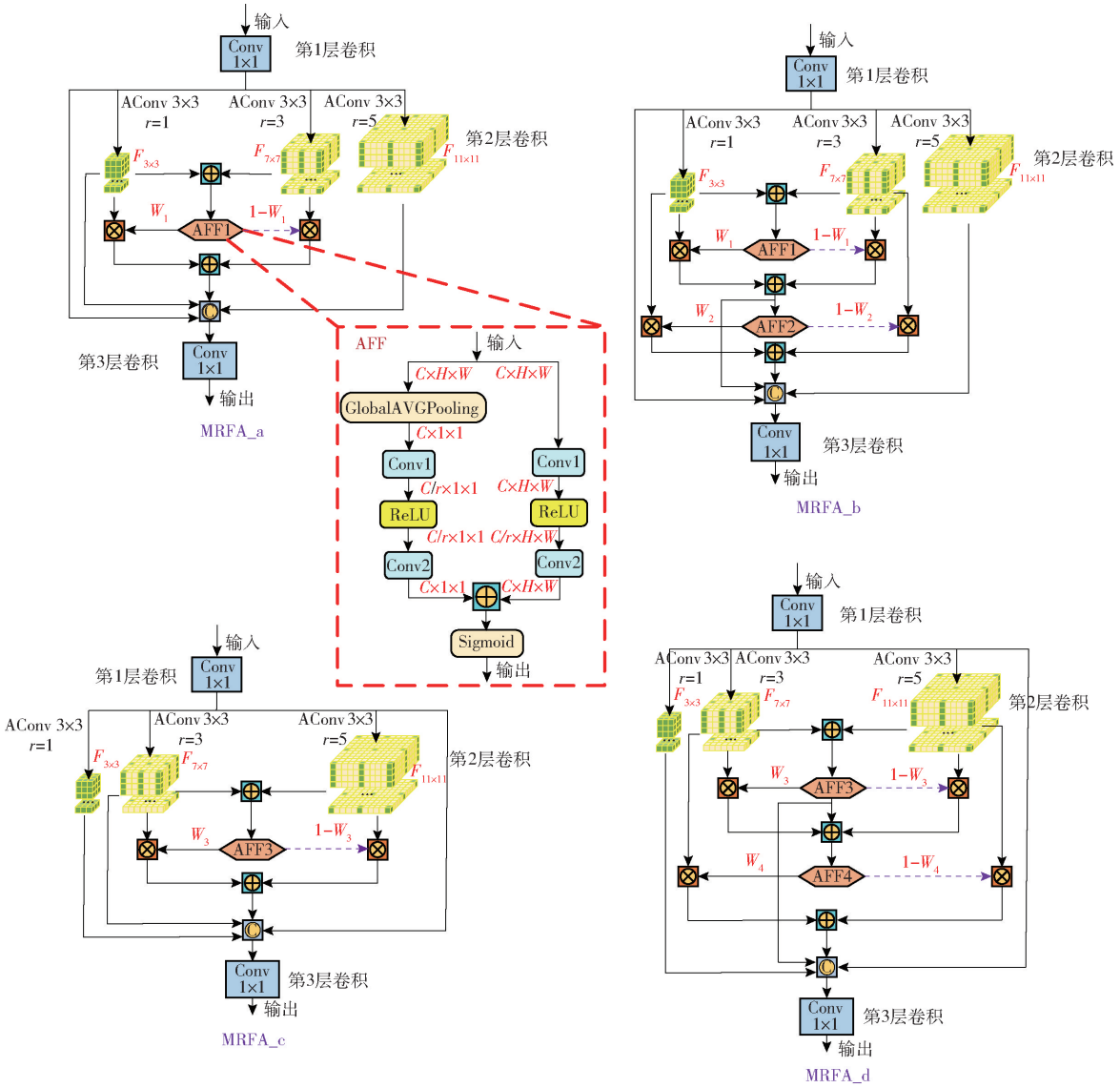


图 5　MRFA 模块

Fig. 5　MRFA module

MRFA 第 3 层卷积输出特征图感受野为 1×1，即 $R_3 = 1$，根据式(3)计算，第 2 层卷积核尺寸为 3×3、7×7、11×11 的空洞卷积输出特征图感受野分别为 3×3、7×7、11×11，即 $R_2^{3\times3} = 3$、$R_2^{7\times7} = 7$、$R_2^{11\times11} = 11$，卷积核尺寸为 11×11 的空洞卷积输出特征图感受野最大，再次利用式(3)计算，得到 MRFA 模块输出特征图的一个像素点在原始特征图的感受野为 11×11。

多感受野特征自适应融合模块使用不同空洞率的空洞卷积获取感受野为 3×3、7×7、11×11 的特征图，本文定义相邻感受野特征图中较小感受野特征图为 $F_{a\times a}$，感受野较大特征图为 $F_{b\times b}$，将 $F_{a\times a}$ 与 $F_{b\times b}$ 进行像素相加，再将像素相加结果输入注意力特征融合 AFF 机制中，获得通道注意力权重 W。上述过程表示为

$$W = \mathrm{AFF}(F_{a\times a} \oplus F_{b\times b}) \tag{4}$$

AFF 是一种能够在通道方向上筛选出重要特征信息的注意力机制，其原理如图 5 所示。在获得通道注意力权重 W 后，权重 W 与 $F_{a\times a}$ 进行通道相乘，同时使用 $1-W$ 与 $F_{b\times b}$ 进行通道相乘，再将通道相乘结果进行像素相加，即对相邻感受野特征图进行加权融合，上述过程表示为

$$Z = W \otimes F_{a\times a} \oplus (1-W) \otimes F_{b\times b} \tag{5}$$

式中 Z——相邻感受野特征图进行加权融合输出

根据感受野特征信息自适应融合机制的不同，设计 4 种 MRFA 结构，分别为图 5 所示的 MRFA_a、MRFA_b、MRFA_c、MRFA_d。定义 3×3、7×7、11×11 感受野特征图为 $F_{3\times3}$、$F_{7\times7}$、$F_{11\times11}$。MRFA_a 将 $F_{3\times3}$、$F_{7\times7}$进行自适应加权融合，随后将其与感受野为 3×3 和 11×11 以及第 1 层卷积输出的特征图进行通道拼接，最后再输入 1×1 卷积中将通道数减半。MRFA_a 特征处理过程可以表示为

$$\mathrm{Output}_{\mathrm{MRFA_a}} = \mathrm{Conv}(\mathrm{Concat}(C_{3\times3}, C_{\mathrm{AFF1}}, C_{11\times11}, C_0)) \tag{6}$$

式中 Conv——卷积核尺寸为 1×1 的卷积

Concat——通道拼接

$C_{3\times3}$、$C_{11\times11}$——$F_{3\times3}$、$F_{11\times11}$对应通道

C_{AFF1}——$F_{3\times3}$与 $F_{5\times5}$自适应加权融合特征对应通道

C_0——第 1 层卷积输出特征通道

MRFA_b 在 MRFA_a 的基础上进行二次特征加权融合，进一步增强特征图中的有效信息，MRFA_b 特征处理过程表示为

$$\mathrm{Output}_{\mathrm{MRFA_b}} = \mathrm{Conv}(\mathrm{Concat}(C_{\mathrm{AFF1}}, C_{\mathrm{AFF2}}, C_{11\times11}, C_0)) \tag{7}$$

式中 C_{AFF2}——$F_{3\times3}$与 $F_{7\times7}$二次加权融合特征对应通道

MRFA_c 与 MRFA_a 相似，所不同的是 MRFA_c 将 $F_{7\times7}$与 $F_{11\times11}$进行自适应加权融合，其特征处理过程表示为

$$\mathrm{Output}_{\mathrm{MRFA_c}} = \mathrm{Conv}(\mathrm{Concat}(C_{\mathrm{AFF3}}, C_{3\times3}, C_{7\times7}, C_0)) \tag{8}$$

式中 C_{AFF3}——$F_{7\times7}$与 $F_{11\times11}$加权融合特征对应通道

$C_{7\times7}$——$F_{7\times7}$对应通道

MRFA_d 在 MRFA_c 的基础上进行二次特征加权融合，其特征处理过程表示为

$$\mathrm{Output}_{\mathrm{MRFA_d}} = \mathrm{Conv}(\mathrm{Concat}(C_{\mathrm{AFF3}}, C_{\mathrm{AFF4}}, C_{3\times3}, C_0)) \tag{9}$$

式中 C_{AFF4}——$F_{7\times7}$与 $F_{11\times11}$二次加权融合特征对应通道

将多感受野特征自适应融合模块嵌入 CLGDRNet 中，增加模型局部感受野的同时突出有效特征信息，进一步提升模型表征能力。

1.4 深层梯度跨空间学习高效多尺度注意力模块

在目标检测模型中，注意力机制的应用可以帮助模型更好地聚焦关键特征，削弱无关信息在特征图上的权重，从而提高识别准确率。如图 6 所示（其中，“//”表示整除，C 表示通道数，G 表示组数，H、W 表示特征图的高、宽），EMA[29] 注意力机制包括 2 条分支，分别为坐标信息嵌入分支和卷积分支。坐标信息嵌入分支可以沿着一个方向捕获长距离的依赖关系，同时可以沿着另一个方向保留准确的位置信息，能够有效地将空间坐标信息聚集到生成的注意图中，生成注意力权重信息，利用该注意力权重信息调整各通道权重，突出有效通道信息。卷积分支中存在一个卷积核为 3×3 的标准卷积，用于制造不同尺度的特征图。将坐标信息嵌入分支和卷积分支的输出进行平均池化及 Softmax 操作，并进行空间交互学习，将交互学习的结果进行像素相加，再进行 Sigmoid 操作，进而得到一个新的注意力权重，随后将该注意力权重与原输入特征信息相乘，得到 EMA 模块输出。

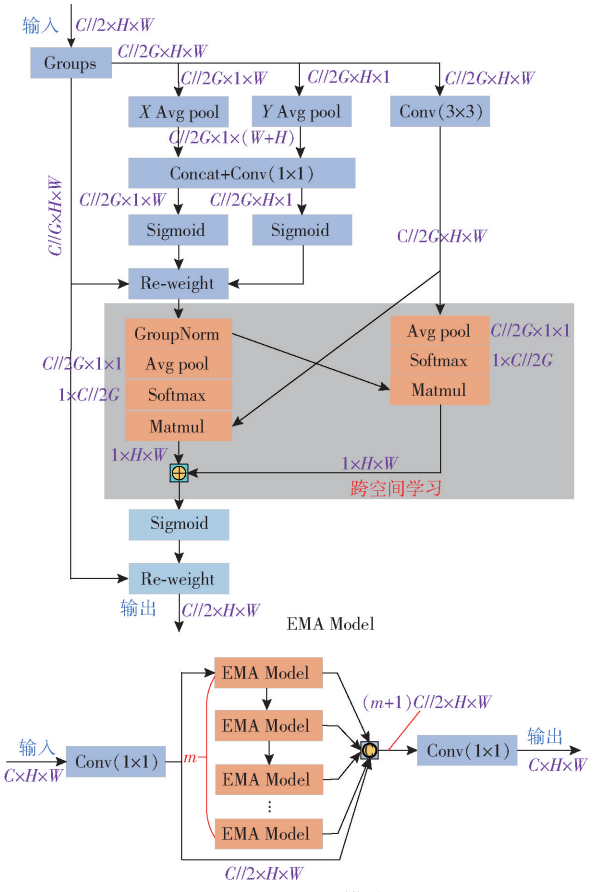


图 6 EMAD 模块

Fig. 6 EMAD module

EMAD 模块结构如图 6 所示。EMAD 模块将 m 个 EMA 注意力模块进行串联以获取更多深层梯度流信息。具体来说，对于输入特征 $X \in \mathbf{R}^{C//2\times H\times W}$，使用卷积核为 1×1 的卷积将通道数减半，随后输入 m 个串联的 EMA 注意力模块，将所有 EMA 模块及卷积核为 1×1 的卷积输出进行拼接，得到 $X \in$

$\mathbf{R}^{(m+1)C//2\times H\times W}$，最后再次使用卷积核为 1×1 的卷积将通道数量恢复为 C，上述过程可以表示为

$$\begin{cases} E_1 = EMA(X) \\ E_2 = EMA(EMA(X)) \\ \quad\vdots \\ E_m = EMA(EMA(EMA(\cdots))) \\ \text{Output}_{\text{EMAD}} = \text{Conv}(\text{Concat}(E_1, E_2, \cdots, E_m)) \end{cases} \tag{10}$$

EMAD 模块可以捕获长距离的依赖关系，并保留准确的位置信息，同时将不同尺度的特征信息进行空间交互学习，有效提高模型识别精度。

2　实验结果

2.1　实验环境及评价指标

在 Ubuntu 18.04 操作系统和 NVIDIA GeForce RTX 3060(12GB)显卡下，通过 Pytorch 深度学习框架来实现模型的搭建、训练和验证，使用 CUDA 11.1，CUDNN 8.0.4 进行网络训练。初始学习率为 0.01，批尺寸设置为 8，迭代周期设置为 200。在训练过程中，通过 Mosaic 算法对输入数据进行在线增强，采用 SGD 对训练过程进行优化。

采用平均精度均值 mAP@0.5、mAP@0.5:0.95 作为模型精度的评价指标，采用推理时间、模型参数量作为模型性能的评价指标。mAP@0.5 表示交并比(IOU)为 0.5 时平均精度均值，mAP@0.5:0.95 表示 IOU 从 0.5 到 0.95，间隔为 0.05 时平均精度均值，推理时间表示模型的单帧图像处理时间，推理时间越短，模型图像处理速度越快，模型参数量指标可直观反映出模型的参数。

2.2　CLGDRNet 识别结果与定量分析

2.2.1　主干信息共享实验

构建 3 条主干网络，并在主干网络间进行特征信息共享以获取更为丰富的深层有效特征信息，本节在 TLGD 上对单主干、双主干特征信息共享、三主干特征信息共享进行消融实验，实验结果如表 1 所示。根据表 1 可知，单一主干网络在特征提取过程中容易损失较多深层特征信息，CLGDRNet 分别采用 CSPNet、GhostNet、ShuffleNet 作为特征提取主干网络时，尽管模型在参数量、推理时间上占有优势，但识别精度较低。相较于单一特征提取主干，CLGDRNet 在使用双主干特征信息共享方式后，模型识别精度均有不同程度的提升，验证了将不同主干网络的特征信息进行共享有助于提升网络识别效果的结论。为进一步获取更为丰富的深层特征信息，CLGDRNet 将 CSPNet、GhostNet、ShuffleNet 所提取的特征信息进行共享，实验结果表明，使用三主干特征信息共享方式获得最好的识别精度，即 mAP@0.5 为 79.7%、mAP@0.5:0.95 为 71.9%。

表 1　主干信息共享实验结果

Tab.1　Experiment results of backbone information sharing

主干网络	参数量	推理时间/ms	mAP@0.5/%	mAP@0.5:0.95/%
CSPNet	3.0×10^6	2.3	76.7	68.8
GhostNet	3.5×10^6	2.0	68.0	60.8
ShuffleNet	2.0×10^6	2.1	77.6	68.4
CSPNet + GhostNet	4.9×10^6	2.5	78.8	70.9
CSPNet + ShuffleNet	4.7×10^6	2.4	78.7	69.4
GhostNet + ShuffleNet	4.8×10^6	2.5	78.6	69.3
CSPNet + GhostNet + ShuffleNet	5.0×10^6	4.0	79.7	71.9

2.2.2　模块消融实验

以 CSPNet + GhostNet + ShuffleNet 为 Baseline，在 TLGD 上验证模块 MRFA_d、EMAD 的识别性能，模块消融实验结果如表 2 所示。由表 2 可知，模型在采用 MRFA_d 模块后，模型识别精度 mAP@0.5 为 82.8%、mAP@0.5:0.95 为 74.3%，验证了 MRFA_d 模块增强模型局部感受野的同时突出有效通道信息进而提高作物叶片识别精度的结论；将 Baseline 中颈部的 C2f 模块替换为 EMAD 模块后，相较于 Baseline，mAP@0.5 和 mAP@0.5:0.95 分别上升 1.6、0.9 个百分点，证明了 EMAD 模块可以获取梯度流信息并跨空间融合不同尺度的上下文信息来增强模型的识别性能；在 Baseline 中同时嵌入 MRFA_d 和 EMAD 模块后，尽管模型的推理时间减慢 1.8 ms，但模型取得最好的识别精度，即 mAP@0.5 为85.0%、mAP@0.5:0.95 为 76.1%，相

表 2　模块消融实验结果

Tab.2　Module ablation experiment results

模块	MRFA	EMAD	参数量	推理时间/ms	mAP@0.5/%	mAP@0.5:0.95/%
Baseline	×	×	5.0×10^6	4.0	79.7	71.9
Baseline + MRFA_d	√	×	5.3×10^6	4.7	82.8	74.3
Baseline + EMAD	×	√	3.1×10^6	5.1	81.3	72.8
Baseline + MRFA_d + EMAD	√	√	5.0×10^6	5.8	85.0	76.1

注：√表示使用该模块，×表示不使用该模块。

较于 Baseline,mAP@0.5 上升 5.3 个百分点,mAP@0.5:0.95 上升 4.2 个百分点。

2.2.3 多感受野特征自适应融合模块实验

在 CSPNet、GhostNet、ShuffleNet 三主干特征信息共享框架下,在 TLGD 上验证多感受野特征自适应融合模块在作物叶片识别任务中的性能。设计了 4 种多感受野自适应特征融合结构 MRFA_a、MRFA_b、MRFA_c、MRFA_d,并将其嵌入网络的次末端,4 种多感受野特征自适应融合结构性能对比如表 3 所示。多感受野特征自适应融合模块使用具有不同空洞率的卷积生成具有不同感受野的特征图进行自适应加权融合,由表 3 可知,模型嵌入 4 种 MRFA 模块后,尽管模型参数量有所上升,推理时间减慢至 4.7 ms,但识别精度均有不同程度的上升。在同一 MRFA 结构下,过大的空洞率导致模型损失了较多的像素特征从而限制了模型识别精度,实验结果表明,卷积空洞率为 1、3、5 时模型取得更好的识别效果。在卷积空洞率为 1、3、5 的前提下,MRFA_a 将 $r=1$ 和 $r=3$ 的空洞卷积所生成特征图进行自适应加权融合,突出有效特征信息,mAP@0.5 和 mAP@0.5:0.95 分别上升 1.6、0.9 个百分点;MRFA_b 将 $r=1$ 和 $r=3$ 的空洞卷积所生成特征图进行二次自适应加权融合,进一步突出有效特征信息,实验结果表明,模型在使用 MRFA_b 后,识别精度进一步提升,mAP@0.5 达到 81.9%,mAP@0.5:0.95 达到 73.5%;MRFA_c 将 $r=3$ 和 $r=5$ 的空洞卷积所生成特征图进行自适应加权融合;MRFA_d 将 $r=3$ 和 $r=5$ 的空洞卷积所生成特征图进行二次自适应加权融合,取得最好的识别精度,mAP@0.5 为 83.4%、mAP@0.5:0.95 为 74.6%。

表 3 4 种 MRFA 性能对比实验结果

Tab.3 Performance comparison experiment results of four MRFA structures

MRFA	空洞率 r			参数量	推理时间/ms	mAP@0.5/%	mAP@0.5:0.95/%
	AConv1	AConv2	AConv3				
MRFA_a	1	3	5	5.3×10^6	4.7	81.3	72.8
	3	5	7	5.3×10^6	4.7	81.2	72.5
	5	7	9	5.3×10^6	4.7	80.5	72.3
MRFA_b	1	3	5	5.3×10^6	4.7	81.9	73.5
	3	5	7	5.3×10^6	4.7	79.9	71.5
	5	7	9	5.3×10^6	4.7	80.2	71.7
MRFA_c	1	3	5	5.3×10^6	4.7	81.6	73.0
	3	5	7	5.3×10^6	4.7	80.9	72.7
	5	7	9	5.3×10^6	4.7	80.2	71.7
MRFA_d	1	3	5	5.3×10^6	4.7	83.4	74.6
	3	5	7	5.3×10^6	4.7	82.1	73.2
	5	7	9	5.3×10^6	4.7	80.6	71.4

2.2.4 深层梯度跨空间学习高效多尺度注意力模块优化实验

将网络颈部的 C2f 模块替换为所设计的深层梯度跨空间学习高效多尺度注意力模块(EMAD),在 EMAD 中,EMA 模块数量的不同,模型在获取目标坐标信息和跨空间融合不同尺度的上下文信息的同时可获得不同深度的特征信息,这会对模型识别性能产生一定程度的影响,在 Baseline + MRFA_d 框架下,在 TLGD 上对 EMAD 中 EMA 模块数量进行优化实验,根据 EMA 模块数量设置 4 组对比实验,实验结果如表 4 所示。由表 4 可知,使用 EMAD 替换模型颈部的 C2f 模块后,模型识别精度均有不同程度的上升,验证了 EMAD 可以捕获长距离的依赖关系,并保留准确的位置信息,同时将不同尺度的特征信息进行空间交互学习,从而有效提高模型识别精度。随着 EMA 模块数量增加,模型参数量逐渐上升,推理时间逐渐减慢。根据实验结果,EMAD 在选用 2 个 EMA 模块时,模型取得最好的识别精度,即 mAP@0.5 为 85.0%、mAP@0.5:0.95 为 76.1%。

表 4 EMAD 优化实验结果

Tab.4 EMAD optimization experiment results

EMA 模块数量	参数量	推理时间/ms	mAP@0.5/%	mAP@0.5:0.95/%
1	4.9×10^6	5.2	84.2	75.1
2	5.0×10^6	5.8	85.0	76.1
3	5.1×10^6	6.6	83.6	75.3
4	5.2×10^6	7.1	84.0	74.9

2.2.5 不同网络对比

为客观评价算法识别性能,选用 YOLO v5n、YOLO v5s、YOLO v5m、YOLO v7 - tiny、YOLO v7、YOLO v8n、YOLO v8s、YOLO v8m 共 8 种识别算法,在实验设备不变条件下与本文方法进行对比实验,

此外，为进一步验证本文模型泛化能力，在 TLGD 和 ALDD 上进行实验，实验结果如表 5、6 所示，图 7 为 YOLO v8n、YOLO v8s、YOLO v8m 以及 CLGDRNet 的识别结果。

表 5　不同模型在 TLGD 上性能对比

Tab. 5　Performance comparison of different models on TLGD

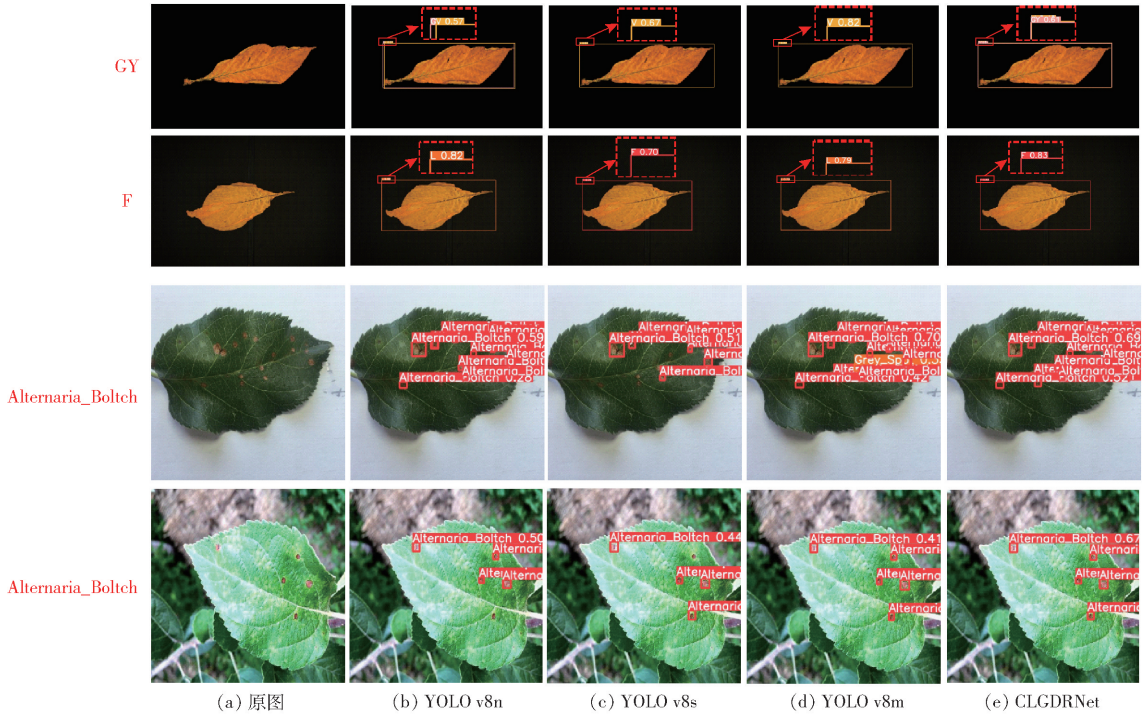
模型	参数量	推理时间/ms	mAP@0.5/%	mAP@0.5:0.95/%
YOLO v5n	1.8×10^6	2.1	71.3	62.9
YOLO v5s	7.0×10^6	4.0	72.7	64.4
YOLO v5m	2.1×10^7	9.2	73.9	65.2
YOLO v7 - tiny	6.0×10^6	2.0	63.1	56.7
YOLO v7	3.7×10^7	8.6	73.1	66.5
YOLO v8n	3.0×10^6	3.5	76.7	68.8
YOLO v8s	1.1×10^7	5.4	77.8	70.6
YOLO v8m	2.6×10^6	14.2	78.3	69.6
CLGDRNet	5.0×10^6	5.8	85.0	76.1

在 TLGD 上进行实验，在识别精度上，CLGDRNet 的 mAP@0.5 为 85.0%、mAP@0.5:0.95 为 76.1%，较其他模型有着较大优势。在模型参数量上，由于模型结构简单，YOLO v5n 模型参数量仅为 1.8×10^6，本文模型使用了主干信息共享结构，其模型参数量有所增大。在推理时间上，本文模型 CLGDRNet 推理时间为 5.8 ms，相较于其他模型，尽管推理时间减慢，但仍具有较高的实时性。

表 6　不同模型在 ALDD 上性能对比

Tab. 6　Performance comparison of different models on ALDD

模型	参数量	推理时间/ms	mAP@0.5/%	mAP@0.5:0.95/%
YOLO v5n	1.8×10^6	2.7	90.3	55.0
YOLO v5s	7.0×10^6	5.0	90.7	56.7
YOLO v5m	2.1×10^7	12.1	93.3	62.7
YOLO v7 - tiny	6.0×10^6	3.9	94.0	64.6
YOLO v7	3.7×10^7	12.8	96.0	70.5
YOLO v8n	3.0×10^6	3.5	94.9	68.9
YOLO v8s	1.1×10^7	7.2	95.6	67.4
YOLO v8m	2.6×10^6	17.8	95.4	69.8
CLGDRNet	5.0×10^6	8.8	97.6	74.2



(a) 原图　(b) YOLO v8n　(c) YOLO v8s　(d) YOLO v8m　(e) CLGDRNet

图 7　识别结果可视化

Fig. 7　Visualization recognition results

在 ALDD 上进行实验，CLGDRNet 取得了最好的识别精度，mAP@0.5 为 97.6%、mAP@0.5:0.95 为 74.2%。相较于 YOLO v5n、YOLO v5s、YOLO v5m，CLGDRNet 在识别精度上有着较大的优势；相较于 YOLO v7 - tiny、YOLO v7，CLGDRNet 的 mAP@0.5 上升 3.6、1.6 个百分点，mAP@0.5:0.95 上升 9.6、3.7 个百分点。此外，CLGDRNet 参数量仅为 YOLO v7 的 13.7%。相较于 YOLO v8n、YOLO v8s、YOLO v8m，本文模型 mAP@0.5 分别上升 2.7、2.0、2.2 个百分点，mAP@0.5:0.95 分别上升 5.3、6.8、4.4 个百分点。

由图 7 可知，在 TLGD 和 ALDD 上，YOLO v8n、YOLO v8s、YOLO v8m 均有不同程度的错检或漏检，CLGDRNet 取得了更好的识别效果。

综上所述,CLGDRNet 可对作物叶片等级及叶片病害做出精确的识别,可为高精度作物叶片等级及病害识别提供技术支撑。

3 结论

(1)针对相似度较高但类别不同的作物叶片及病害难以区分问题,本文提出了主干信息共享与多感受野特征自适应融合的作物叶片等级和病害识别算法(CLGDRNet)。CLGDRNet 将 CSPNet、GhostNet、ShuffleNet 所提取的特征进行共享以达到特征信息互补的目的,进而保留更多对目标识别有益的深层特征信息;通过多感受野特征自适应融合模块将不同感受野的特征图进行自适应加权融合,增强模型局部感受野的同时突出有效通道信息;使用深层梯度跨空间学习的高效多尺度注意力模块替换网络颈部的 C2f 模块,获取目标位置信息和深层梯度流信息的同时跨空间融合不同尺度的特征信息,进一步提高模型识别精度。

(2)实验结果表明,CLGDRNet 对相似度较高的 5 个等级初烤烟叶识别精度 mAP@0.5 达到 85.0%、mAP@0.5:0.95 达到 76.1%,对常见的 5 个苹果叶病害类别识别精度 mAP@0.5 达到 97.6%、mAP@0.5:0.95 达到 74.2%,均优于所对比的 YOLO v5n、YOLO v5s、YOLO v5m、YOLO v7-tiny、YOLO v7、YOLO v8n、YOLO v8s、YOLO v8m,此外,CLGDRNet 对初烤烟叶和苹果叶病害的推理时间分别为 5.8、8.8 ms,具有较高的实时识别性能。

参 考 文 献

[1] KHAN M U A, KHAN M A, ABBASI U A, et al. Kiwi plant growth monitoring with soil and climatic conditions in the semi-arid region of Pakistan[J]. Environmental Sciences Proceedings, 2023, 23(1): 36.

[2] 王鹏新,田惠仁,张悦,等. 基于深度学习的作物长势监测和产量估测研究进展[J]. 农业机械学报,2022,53(2):1-14.
WANG Pengxin, TIAN Huiren, ZHANG Yue, et al. Crop growth monitoring and yield estimation based on deep learning: state of the art and beyond[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022,53(2):1-14. (in Chinese)

[3] 于珍珍,邹华芬,于德水,等. 融合田间水热因子的甘蔗产量 GA-BP 预测模型[J]. 农业机械学报,2022,53(10):277-283.
YU Zhenzhen, ZOU Huafen, YU Deshui, et al. Sugarcane yield GA-BP prediction model incorporating field water and heat factors[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022,53(10):277-283. (in Chinese)

[4] 罗琦,茹晓雅,姜元,等. 基于机器学习与气象灾害指标的苹果相对气象产量预测[J]. 农业机械学报,2023,54(9):352-364.
LUO Qi, RU Xiaoya, JIANG Yuan, et al. Prediction of apple relative meteorological yields based on machine learning and meteorological disaster indices[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023,54(9):352-364. (in Chinese)

[5] 严海军,卓越,李茂娜,等. 基于机器学习和无人机多光谱遥感的苜蓿产量预测[J]. 农业工程学报,2022,38(11):64-71.
YAN Haijun, ZHUO Yue, LI Maona, et al. Alfalfa yield prediction using machine learning and UAV multispectral remote sensing[J]. Transactions of the CSAE, 2022,38(11):64-71. (in Chinese)

[6] 张家瑜,朱锐,邱威,等. 基于选择性注意力神经网络的木薯叶病害检测算法[J]. 农业机械学报,2024,55(5):254-262,272.
ZHANG Jiayu, ZHU Rui, QIU Wei, et al. Cassava leaf disease detection algorithm based on selective attention neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024,55(5):254-262,272. (in Chinese)

[7] PANCHAL A V, PATEL S C, BAGYALAKSHMI K, et al. Image-based plant diseases detection using deep learning[J]. Materials Today: Proceedings, 2023, 80(1): 3500-3506.

[8] REEDHA R, DERICQUEBOURG E, CANALS R, et al. Transformer neural network for weed and crop classification of high resolution UAV images[J]. Remote Sensing, 2022, 14(3): 592.

[9] 亢洁,刘港,王勃,等. 基于动态剪枝神经网络的杂草检测算法研究[J]. 农业机械学报,2023,54(4):268-275.
KANG Jie, LIU Gang, WANG Qing, et al. Weed detection algorithm based on dynamic pruning neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023,54(4):268-275. (in Chinese)

[10] CHAUDHARY P, CHAUDHARI A K, CHEERAN A N, et al. Color transform based approach for disease spot detection on plant leaf[J]. International Journal of Computer Science and Telecommunications, 2012, 3(6): 65-70.

[11] ARIVAZHAGAN S, SHEBIAH R N, ANANTHI S, et al. Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features[J]. Agricultural Engineering International: CIGR Journal, 2013, 15(1): 211-217.

[12] CHAKI J, PAREKH R, BHATTACHARYA S. Plant leaf recognition using texture and shape features with neural classifiers[J]. Pattern Recognition Letters, 2015, 58(6): 61-68.

[13] 黄海松,陈星燃,韩正功,等. 基于多尺度注意力机制和知识蒸馏的茶叶嫩芽分级方法[J]. 农业机械学报,2022,53(9):399-407,458.
HUANG Haisong, CHEN Xingran, HAN Zhenggong, et al. Tea buds grading method based on multiscale attention mechanism and knowledge distillation[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022,53(9):399-407,458. (in Chinese)

[14] 何自芬,罗洋,张印辉,等. 多感受野特征自适应融合及动态损失调整的初烤烟叶等级检测[J]. 光学精密工程,2024,32(2):301 - 316.
HE Zifen, LUO Yang, Zhang Yinhui, et al. Flue-cured tobacco leaf grade detection through multi-receptive field features fusing adaptively and dynamic loss adjustment[J]. Optics and Precision Engineering, 2024,32(2):301 - 316. (in Chinese)

[15] 刘文波,叶涛,李颀. 基于改进 SOLO v2 的番茄叶部病害检测方法[J]. 农业机械学报,2021,52(8):213 - 220.
LIU Wenbo, YE Tao, LI Qi. Tomato leaf disease detection method based on improved SOLO v2[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021,52(8):213 - 220. (in Chinese)

[16] 刘斌,徐皓玮,李承泽,等. 基于快照集成卷积神经网络的苹果叶部病害程度识别[J]. 农业机械学报,2022,53(6):286 - 294.
LIU Bin, XU Haowei, LI Chengze, et al. Apple leaf disease identification method based on snapshot ensemble CNN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022,53(6):286 - 294. (in Chinese)

[17] HUANG Y Q, ZHENG J C, SUN S D, et al. Optimized YOLO v3 algorithm and its application in traffic flow detections[J]. Applied Sciences, 2020, 10(9): 3079.

[18] YU J, ZHANG W. Face mask wearing detection algorithm based on improved YOLO - v4[J]. Sensors, 2021, 21(9): 3263.

[19] ZHANG Y, GUO Z, WU J, et al. Real-time vehicle detection based on improved YOLO v5[J]. Sustainability, 2022, 14(19): 12274.

[20] LI S, TAO T, ZHANG Y, et al. YOLO v7 - CS: a YOLO v7-based model for lightweight bayberry target detection count[J]. Agronomy, 2023, 13(12): 2952.

[21] LI Y, FANG Q, HUANG H, et al. A modified YOLOv8 detection network for UAV aerial image recognition[J]. Drones, 2023, 7(5): 304.

[22] 王蕾,郭文平,陈欣慰,等. 融合多尺度特征的蜗杆表面缺陷检测[J]. 光学精密工程,2024,32(11):1746 - 1758.
WANG Lei, GUO Wenping, CHEN Xinwei, et al. Worm surface defect detection with fusion of multi-scale features[J]. Optics and Precision Engineering,2024,32(11):1746 - 1758. (in Chinese)

[23] 周飞,郭杜杜,王洋,等. 基于改进 YOLOv8 的交通监控车辆检测算法[J]. 计算机工程与应用,2024,60(6):110 - 120.
ZHOU Fei, GUO Dudu, WANG Yang, et al. Vehicle detection algorithm based on improved YOLOv8 in traffic surveillance [J]. Computer Engineering and Applications,2024,60(6):110 - 120. (in Chinese)

[24] GAO F, CAI C, JIA R, et al. Improved YOLOX for pedestrian detection in crowded scenes[J]. Journal of Real-Time Image Processing, 2023, 20(2): 24.

[25] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020: 390 - 391.

[26] HAN K, WANG Y, TIAN Q, et al. Ghostnet: more features from cheap operations[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 1580 - 1589.

[27] ZHANG X, ZHOU X, LIN M, et al. Shufflenet: an extremely efficient convolutional neural network for mobile devices[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 6848 - 6856.

[28] 张印辉,杨宏宽,刘强,等. 基于模型压缩与感受野增强的下茧实时检测[J]. 纺织学报,2021,42(11):29 - 38.
ZHANG Yinhui, YANG Hongkuan, LIU Qiang, et al. Real-time detection of inferior cocoons through model compression and receptive field enhancement[J]. Journal of Textile Reaearch,2021,42(11):29 - 38. (in Chinese)

[29] OUYANG D, HE S, ZHANG G, et al. Efficient multi-scale attention module with cross-spatial learning[C]//ICASSP 2023—2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2023: 1 - 5.