

# 基于改进 YOLOv5 的红树林单木目标检测研究

马永康<sup>1,2</sup>, 刘华<sup>1,2\*</sup>, 凌成星<sup>1,2</sup>, 赵峰<sup>1,2</sup>, 姜怡<sup>1,2</sup>, 张雨桐<sup>1,2</sup>

<sup>1</sup>中国林业科学研究院资源信息研究所, 北京 100091;

<sup>2</sup>国家林业和草原局林业遥感与信息技术重点实验室, 北京 100091

**摘要** 针对无人机影像中红树林单木目标较小且分布密集, 对其检测时自动化程度不高、效率低等问题, 基于深度学习方法提出了一种红树林单木目标检测模型(YOLOv5-ECA), 以实现无人机影像中红树林单木快速、精确的自动识别和定位。首先利用开源软件 LabelImg 在选取的无人机影像上依次标注目标树, 构建红树林单木数据集; 其次选择 YOLOv5 为基础目标检测模型, 依据目标分布密集且尺寸较小的特点对其进行优化和改进; 使用有效通道注意力(ECA)机制对 CSPDarknet53 骨干网络进行改进, 避免降维的同时增强特征表达能力, 并在 SPP 模块引入 SoftPool 改进池化操作, 保留更多细节特征信息; 最后利用 ACON 自适应激活函数自适应地决定是否激活神经元。结果表明: 使用已构建的数据集对改进前后的网络进行训练, 在测试集上对比准确率、召回率、平均精准度的均值(mAP)@0.5 等参数, 各模型略有差异但均趋于收敛; 所提 YOLOv5-ECA 的平均检测精度较 YOLOv5 提高了 3.2 个百分点, 较 YOLOv4 提升了 5.19 个百分点, 同时训练损失也更低, 能够快速、精准且自动化地检测红树林单木目标, 较好地提升了对红树林单木的识别和定位能力。

**关键词** 红树林单木; 无人机影像; 深度学习; 目标检测; YOLOv5-ECA

中图分类号 S757.1

文献标志码 A

DOI: 10.3788/LOP202259.1828003

## Object Detection of Individual Mangrove Based on Improved YOLOv5

Ma Yongkang<sup>1,2</sup>, Liu Hua<sup>1,2\*</sup>, Ling Chengxing<sup>1,2</sup>, Zhao Feng<sup>1,2</sup>, Jiang Yi<sup>1,2</sup>, Zhang Yutong<sup>1,2</sup>

<sup>1</sup>Research Institute of Forest Resource Information Techniques, Chinese Academy of Forestry,  
Beijing 100091, China;

<sup>2</sup>Key Laboratory of Forestry Remote Sensing and Information System, National Forestry and  
Grassland Administration, Beijing 100091, China

**Abstract** In this study, an individual mangrove object detection model called YOLOv5-ECA based on deep learning is proposed to automatically identify and locate individual mangroves with high accuracy aiming at the challenges of small and dense individual mangroves in drone images, resulting in low automation and efficiency for detecting them. First, the open-source software LabelImg is used to mark the target tree on the selected drone image, which is applied to construct the individual mangrove dataset. Then, the YOLOv5 is used as the basic object detection model to maximize and enhance the target tree, and achieving this is based on the characteristics of dense distribution and small size of objects. The efficient channel attention (ECA) mechanism enhances the CSPDarknet53 backbone network to avoid dimensionality reduction while enhancing feature expression capabilities. Furthermore, the enhanced SoftPool pooling operation is introduced into the SPP module to retain more detailed feature information. Finally, the ACON adaptive activation function determines whether the neuron is activated. The results demonstrate that the constructed dataset is used to train the network before and after improvement, and the accuracy, recall, and mean average precision (mAP)@0.5 parameters are compared. The results of different models are slightly different, but they all tend to converge. The proposed YOLOv5-ECA's average detection accuracy is 3.2 percentage points higher than YOLOv5 and 5.19 percentage points higher than YOLOv4, and its training loss is also lower. The deep learning-based YOLOv5-ECA model can quickly, accurately, and automatically detect individual mangroves and significantly enhance the ability to identify and locate them.

**Key words** individual mangrove; drone image; deep learning; object detection; YOLOv5-ECA

收稿日期: 2021-07-18; 修回日期: 2021-08-07; 录用日期: 2021-08-10

基金项目: 高分湿地资源监测应用子系统(二期)(21-Y30B02-9001-19/22-2)

通信作者: liuhua@ifrit.ac.cn

# 1 引言

红树林是热带和亚热带海岸带上特有的森林群落,同时红树林生态系统也是地球上生物最为复杂、物种最为丰富的生态系统之一,具有重要的生态和经济意义<sup>[1]</sup>,在改善生态状况、维护生物多样性和沿海地区生态安全等方面发挥着重要作用<sup>[2]</sup>。红树林资源调查和动态监测是红树林科学保护与管理的基础和前提,也是加强海洋生态系统的恢复机理与技术研究的重要步骤<sup>[3]</sup>。但是人为改造等原因,我国红树林面积正在日益减小,为了更好地保护红树林生态系统,扩大红树林面积,近年来多个红树林自然保护区根据规划开展了红树林人工造林工作。随着对红树林保护意识的加强,在原有红树林中补植和人工造林成为恢复红树林资源的主要措施。因此如何实现对红树林单木快速、精确、自动化的识别和定位在新造红树林的成效评估中具有重要意义。

红树林分布地理位置特殊,多在潮间带浅滩,红树林幼苗若未成活则因为根系无附着力被水流冲走,因此存在的红树林单木幼苗即为成活的林木,在对其进行精确检测的基础上结合该区域造林规划可得成活率。同时在此特殊的环境中,低空无人机遥感系统因具有数据采集灵活、成本低且可快速获取小范围超高分辨率影像的特点<sup>[4]</sup>,非常适用于红树林区域的影像采集。基于此,将无人机应用于红树林幼苗的获取,是有效提高红树林幼苗监测精度的途径之一<sup>[5]</sup>。随着林业信息技术的发展,无人机已结合人工智能、图像处理等方法在林业病虫害监测、森林植被分类以及树木检测等领域广泛应用<sup>[6]</sup>。冯家莉等<sup>[7]</sup>采用面向对象的最近邻分类方法对红树林树种类型进行分类研究,并对比综述了无人机遥感和常规航空航天遥感技术对红树林资源调查监测的优缺点,得出无人机遥感系统非常适用于红树林资源调查的结论。Navarro等<sup>[8]</sup>使用窗口滤波算法和控制分水岭分割算法成功估算出红树林的地上生物量,为管理人员提供可重复、低成本的 UAV-SfM 调查方法。杨礼<sup>[9]</sup>用无人机高分辨率影像数据,针对单一树种的林木区域,建立了一种基于模板匹配的单株树识别方法,该方法采用相关系数测度方法,结合单株树灰度分布特征,进行单株树识别。胡馨月等<sup>[10]</sup>利用无人机数据生成的正射影像作为研究对象,提出一种 MeanShift 算法和分水岭分割算法相结合的林木株数提取方法。但是以上方法提取过程均非常复杂和繁琐,同时由于红树林幼苗分布较为密集且目标较小,对红树林单木进行逐一识别和定位较为困难。加之无人机分辨率较高,用 CPU 批处理也需要大量的时间,部分方法还需要结合人眼观察和后续的工作等<sup>[11]</sup>。如何结合无人机影像快速、精确地检测红树林单木成为新造红树林成活率核查工作中亟待解决的问题。

随着目标检测和物体识别技术的发展,深度学习开始广泛应用于目标检测,人脸识别和语音识别等是最常见的应用<sup>[12]</sup>。针对特征提取能力不足、预测框处理后检测速度降低等问题,李娜等<sup>[13]</sup>提出了一种多尺度特征融合的 Anchor-Free 轻量化检测算法。对于现有目标检测算法在复杂巡检场景下巡检缺陷时检测精度较低的问题,赵杰伦等<sup>[14]</sup>提出一种基于尺度不变特征金字塔的缺陷检测方法。但以上方法均只对相应的特定场景进行算法改进。近年来,深度学习也逐渐应用于林业行业,逐步实现对林业更加精准、快速、智能的监测。周焱等<sup>[11]</sup>针对无人机森林图像中树木尺度小、生长密集以及分布不规律等问题,提出一种基于深度学习的小目标受灾树木检测方法。但是该 SSD 算法舍弃了含有丰富信息的底层特征,对小目标检测的鲁棒性较低。遥感图像具有目标尺度差异大且类间相似度高特点,使得处理自然图像的目标检测算法直接应用于遥感图像时仍面临着一些挑战。针对上述挑战,姚艳清等<sup>[15]</sup>提出一种多分辨率特征融合的遥感图像目标检测方法。YOLO 目标检测模型经历了数次更新迭代,先后解决了多目标检测、小目标检测、漏检修复、多尺度预测等多个维度的问题。同时 YOLO 系列模型拥有较好的全局感受野,利用网格划分、锚框匹配和多语义融合检测机制,能够有效改善检测过程感受野受限的问题,从而实现对目标的快速检测。基于目标检测可自动学习特征、速度快、效率高等优势,将该方法更好地应用于红树林单木检测中以实现红树林幼苗状况的智能监测是本研究的主要目的。

为快速、精准且自动化地检测红树林单木目标,本文将单阶段目标检测的深度学习模型 YOLOv5 引入红树林单木的检测研究中,并对其进行改进,融合有效通道注意力(ECA)机制并将池化操作替换为 SoftPool,形成改进的 YOLOv5-ECA 模型,提升对红树林单木的检测能力。

## 2 数据与研究方法

### 2.1 研究区概况

广东湛江红树林国家级自然保护区地处广东省西南部的湛江市,中国最南端,东经 109°40′~110°35′,北纬 20°14′~21°35′,由沿雷州半岛海岸线带状间断性分布的多个保护小区组成,是中国红树林面积最大、分布最集中的自然保护区。它属于森林与湿地类型自然保护区,主要保护对象为热带红树林湿地生态系统及其生物多样性。保护区位于北热带向南亚热带的过渡区域,因此受季风气候和海洋气候影响较大。在保护区内,目前按照一定的规则在大量空余的坑塘人工种植了各类红树林幼苗,本研究将以此为对象展开目标检测研究。

### 2.2 数据来源

鉴于湛江红树林自然保护区新造红树林特殊的地



理位置和自身较为矮小的特点,在2020年9月,选择一天中退潮的时间段,无人机以120 m航高,通过飞行任务区进行多次断点续飞数据采集,在所选大约3000 m<sup>2</sup>的研究区内共获得2348张无人机图像,理论空间分辨率为0.05 m,可以清晰地分辨出单株红树林幼苗,从而构建本研究所需数据集。该无人机由深圳市大疆创新科技有限公司生产,搭载CCD传感器,型号为大疆精灵4RTK,设计轻巧、操作方便、性能良好。获取的部分无人机影像如图1所示。

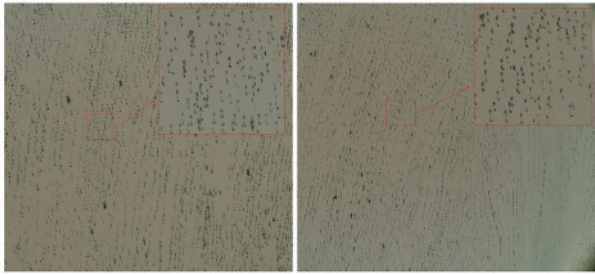


图1 部分红树林单木无人机影像  
Fig. 1 Partial drone images of individual mangrove

### 2.3 YOLOv5目标检测算法原理

YOLO系列算法是由Redmon等<sup>[16]</sup>在2016年开发出的一种单阶段目标检测算法,该算法采用单一回归的思想,更易学习待检测目标的泛化特征,是一种全新的目标检测架构。YOLO模型的网络结构形式为从端到端,在生成候选框的同时完成分类并预测存在的对象及它们的位置,使得时间成本和计算资源都大大降低。YOLO算法从最初的YOLOv1,经过不断的创

新和完善,发展到现在最新的YOLOv5。YOLOv5算法依据不同的网络深度和宽度,依次有YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x四个不同版本。同时YOLOv5在前一版YOLOv4的基础上增加了不少新的改进思路,通过训练、测试速度和精度的改善,模型性能有了很大的提升。

一个完整的目标检测算法,通常包括输入端(Input)、基准网络(Backbone)、特征融合部分(Neck)、检测头(Head)<sup>[17]</sup>。YOLOv5的网络结构如图2所示。YOLOv5的输入端包含了训练数据集图像的预处理阶段,即将输入图像缩放到网络自身所需要的输入尺寸,并依次对输入图片进行归一化等操作,为图片的输入做好准备。在输入端沿用YOLOv4的Mosaic数据增强,用来提升模型的训练速度和网络精度。同时依据不同的数据集,新加了自适应锚框计算程序,该程序可以自适应地对图片进行不同的缩放。在骨干网络部分,相比YOLOv4,YOLOv5新增了Focus结构,用来提取通用的特征等,并依此构造了两种CSP结构。Focus结构最主要的功能是对图片进行切片操作,Focus结构及切片示意图如图3所示。在骨干网络和检测头中间的特征融合部分,YOLOv5添加了FPN+PAN的组合结构,利用它进一步增强了对特征的多样性和鲁棒性。检测头作为输出层用来完成目标检测后结果的最终输出,通常包含一个分类分支和一个回归分支。综合整个YOLO算法系列的优缺点,最终选用GIOU\_Loss为训练时的损失函数,输出层的锚框机制与YOLOv4保持一致。

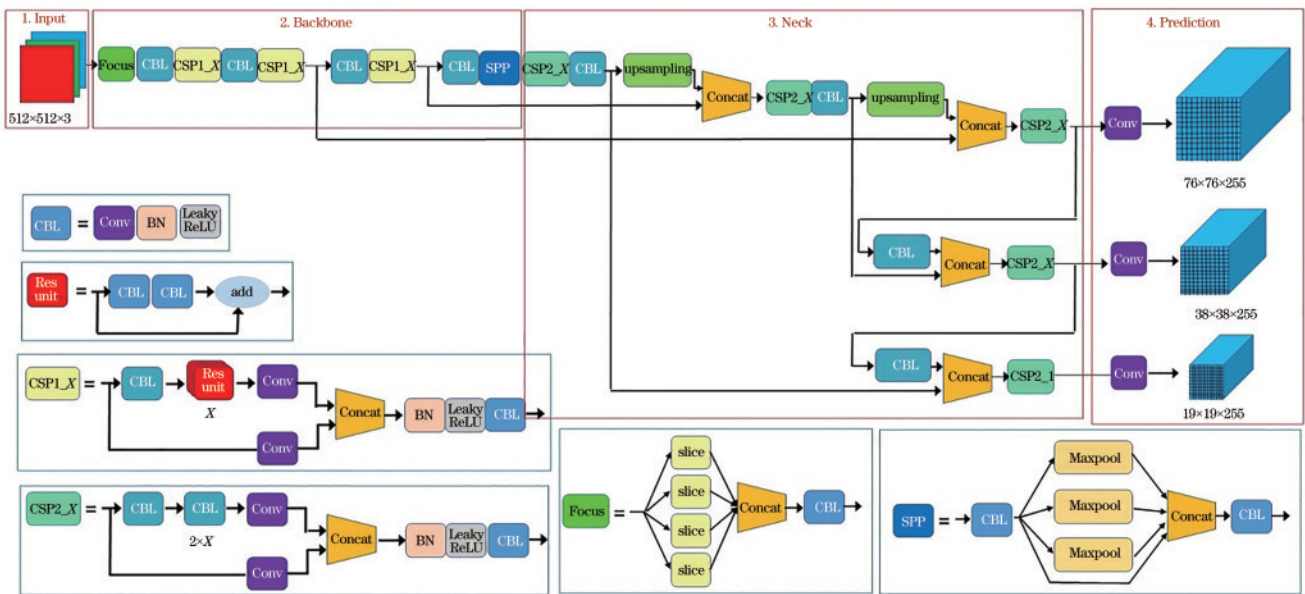


图2 YOLOv5网络结构  
Fig. 2 YOLOv5 network structure

### 2.4 模型改进与优化

YOLOv5是YOLO系列最新的算法,模型自身大

小和检测速度均已取得较大改进。但YOLOv5原模型都是在公开数据集训练而来的,针对本研究所构建

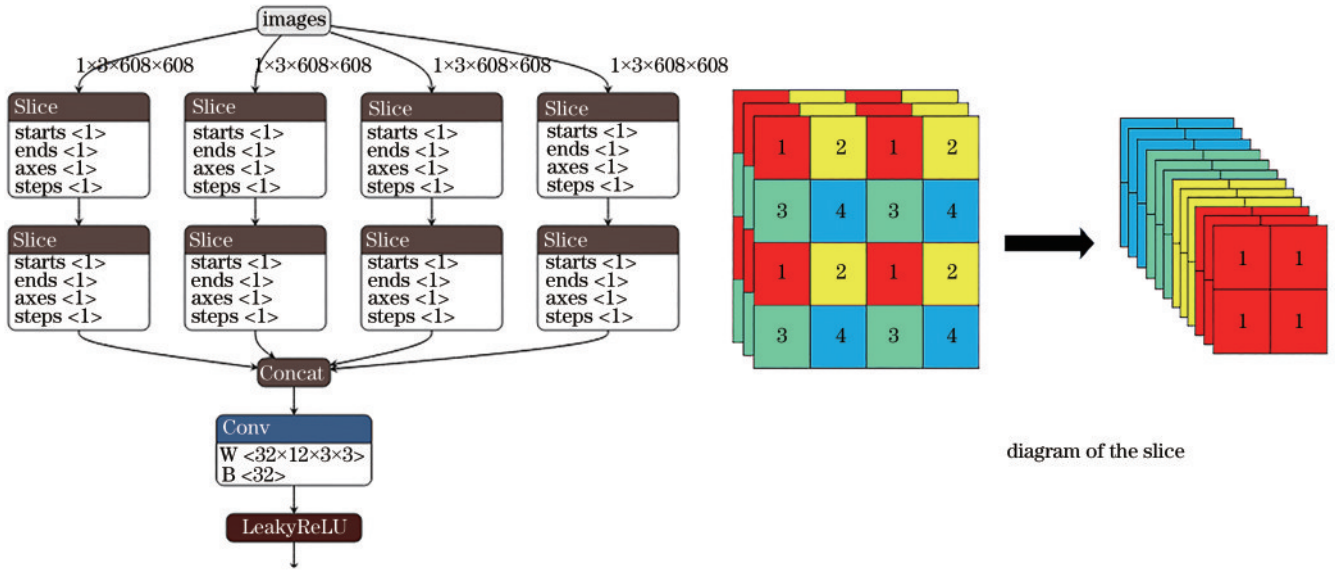


图 3 Focus 结构及切片示意图  
Fig. 3 Schematic of Focus structure and slice

的红树林幼小单木数据集的特点, 还需对该模型进行改进和优化, 以更好地识别和检测目标。

2.4.1 基于有效通道注意力机制的 YOLOv5-ECA 网络

YOLOv5 网络在提取特征时会同等对待模型的所有通道, 这在一定程度上限制了网络的检测性能。由于本次研究的对象是目标较小且分布较为密集的新造红树林幼苗, 因此在 YOLOv5 的 CSPDarknet53 特征提取网络中引入有效通道注意力机制, 改进后模型重新命名为 YOLOv5-ECA。注意力机制被证明是提升目标检测性能的重要方法, 在各类主流的检测算法中都得以广泛应用。而 ECA 最突出的特点是避免降维, 具有跨通道交互作用, 同时能降低模型的复杂性, 增强特征表达能力<sup>[18]</sup>。ECA 机制如图 4 所示。

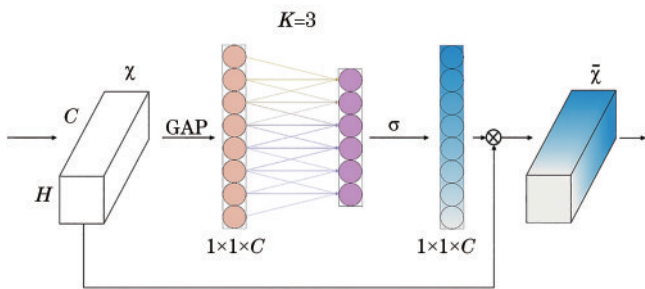


图 4 ECA 机制结构图  
Fig. 4 Structure diagram of ECA mechanism

ECA 模块通过大小为  $k$  的快速一维卷积产生通道注意, 其中核的大小完全由通道维数相关函数自适应确定。在维度保持不变的条件下输入特征图像  $\chi$ , 所有通道经全局平均池化后, ECA 模块将会利用一个可

以权重共享的一维卷积来学习特征, 并在学习特征时通过每一个通道与之  $k$  个近邻来捕获跨通道交互。 $k$  表示快速一维卷积的内核大小, 是通过跨通道信息交互作用的遮盖区域面积与通道维数  $C$  的正比关系得到的, 计算公式为

$$k = \left\lceil \frac{\log_2 C}{\gamma} + \frac{b}{\gamma} \right\rceil_{\text{odd}}, \quad (1)$$

式中:  $\gamma=2; b=1; \lceil \cdot \rceil_{\text{odd}}$  表示最邻近奇数。

一般的注意力机制模块在计算时会选择高层特征图来进行全局的池化, 通过降维的方式将每个通道的二维特征最终压缩成一个实数。经过此计算过程, 本可以认为把整个特征压缩以此获得全局较为完整的感受野, 但是在进行降维压缩时, 会丢失部分的空间特征信息, 同时也无法体现通道维数和空间维数之间的相互依附关系。虽然较为成熟的 Convolutional Block Attention Module (CBAM) 是同时基于空间和通道的, 但是它们的计算过程是相互独立的, 计算时间和计算量都大幅增加<sup>[19]</sup>。相较而言 ECA 的重要意义在于: 一是不进行降维, 会整合所有的通道和空间维数信息, 不丢失信息; 二是通过跨通道信息交互作用的遮盖区域面积与通道维数  $C$  的正比关系得到自适应的  $k$  值, 减小了计算量。ECA 机制降低模型复杂性的同时, 增强特征的表达能, 进而提高模型精度。

2.4.2 融合 SoftPool 池化结构的 YOLOv5-ECA

池化操作是现有的卷积神经网络 (CNN) 中最基础的步骤, 在减小特征图大小和模型的计算成本的同时, 能够在一定程度减缓模型训练时的过拟合现象, 增大后续卷积运算的感受野, 降低卷积层检测目标信息时过度敏感的风险。已有的池化方法大多是基于平均

池化和最大池化的,平均池化仅对邻域内特征点求平均值,能很好地保留背景,但容易使图片模糊。最大池化对邻域内特征点取最大,能很好地保留纹理特征,也正因为保留了池化内核中的局部最大值,更适用于突出特征图中响应较为强烈的部分,容易忽略部分细节特征。针对本研究中目标较小且像素过低的新造红树

林幼苗,利用最大值池化和平均池化进行特征映射时易损失重要检测信息。因此在 SPP 模块中引入 SoftPool 改进池化操作,保留更多细节特征信息,SoftPool 结构如图 5 所示,融合在 YOLOv5 中的位置如图 6 所示。

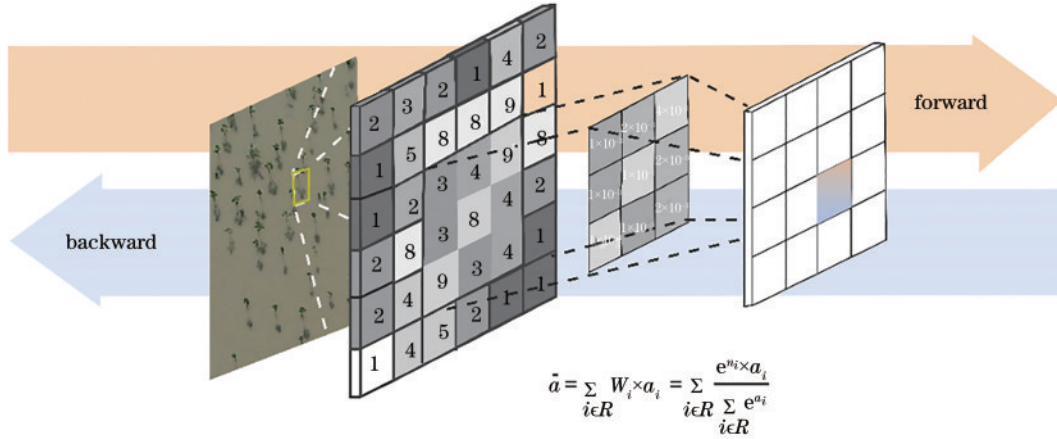


图 5 SoftPool 结构图

Fig. 5 SoftPool structure diagram

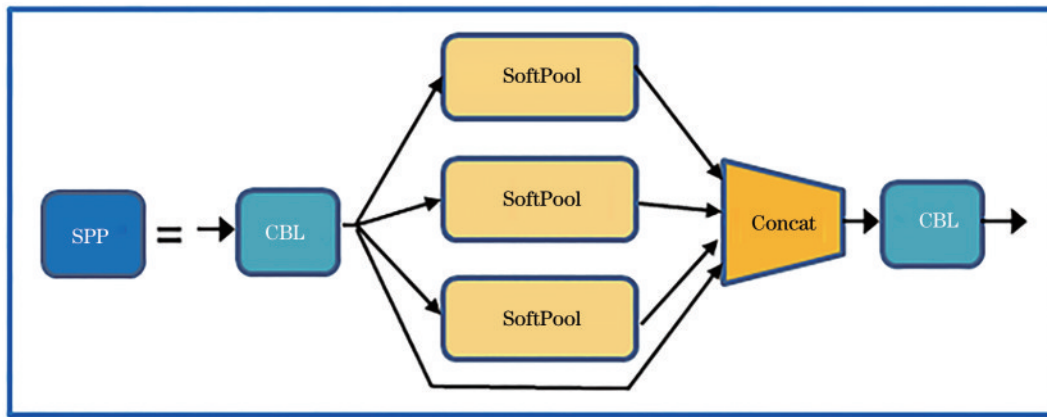


图 6 SoftPool 在 YOLOv5-ECA 中的位置

Fig. 6 Position of SoftPool in YOLOv5-ECA

与其他池化不同,SoftPool 使用 Softmax 进行加权池化,能够保持特征的表达性且操作是可微的<sup>[20]</sup>。每一次反向传播的梯度都能够得到更新,SoftPool 可综合利用池化内核每一个激活因子,仅增加极少内存,增大相近特征信息区分度的同时,保留了整体感受野的特征信息,提高算法的准确率。

SoftPool 的核心思想在于 Softmax 的利用,根据非线性特征值计算区域  $R$  的特征值权重:

$$W_i = \frac{e^{a_i}}{\sum_{i \in R} e^{a_i}}, \quad (2)$$

式中:  $W_i$  为权重;  $a_i$  为激活值。权重能够保证重要特征的传递,保证区域  $R$  内的特征值在反向传递时至少会有预设的最小梯度。在得到权重后,通过加权区

域  $R$  内的特征值,得到输出:

$$\tilde{a} = \sum_{i \in R} W_i \times a_i, \quad (3)$$

式中:  $\tilde{a}$  为 SoftPool 的输出值。SoftPool 实现了对池化内核所有激活因子的加权求和。

### 2.5 评价指标

目标检测中通常用平均精准度 (AP) 和平均精准度的均值 (mAP) 来评价模型的检测效果和性能<sup>[21]</sup>, AP 为召回率 (R) 和精确率 (P) 曲线下的面积,本研究目标为单类目标,AP 等同于 mAP。面积交并比 (IoU) 表示计算模型预测目标的矩形区域与验证集中目标标定的矩形区域的面积交并比,衡量模型的位置预测能力。精确率表示模型检测出正确的目标数占总目标数的比例,体现出模型在目标检测时的准确度。召回率



表示模型已检测出的目标数量占总目标数量的比例,体现了模型识别的查全能力。

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}}, \quad (4)$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}}, \quad (5)$$

$$P_A = \int_0^1 P(R) dR, \quad (6)$$

式中:  $N_{TP}$  为检测出正确的正样本个数,即预测框与标注框类别相同且 IoU 大于 0.5;  $N_{FP}$  为检测出错误的正样本个数;  $N_{FN}$  为检测出错误的负样本个数。由于精确率与召回率受置信度的影响,若只采用精确率与召回率来评价模型性能,会存在一定的不科学性和局限性,因此在实验中引入 AP 作为评价指标来评价模型的识别性能,它是目前评价主流目标检测算法性能的最重要指标之一。

### 3 实验与结果分析

#### 3.1 基于无人机的新造红树林数据集的构建

本研究以无人机影像为主要数据源,以新造红树

林单木为检测目标。由于检测目标物较为密集,在单张影像中分布较多,在构建数据集时选取了较为清晰的分辨率为  $5472 \times 3648$  像素的原始图像,通过基于 Python 编写的分割程序,将原始影像分割为  $512 \times 512$  像素的图片 597 张,以便更好地进行目标标注,提高训练速度。为保证图像中各个检测目标的完整,使用较为常用的 LabelImg 开源软件对无人机图像中红树林单木幼苗依次进行标注,大约标注 3 万棵目标树,标注内容为新造红树林矩形包围框坐标,存储为 XML 文本文件用于 YOLOv5 模型的训练和测试。同时编写相应的程序,在 597 张裁剪后的图像上将标注目标框贴合在目标树上,核查有无漏标注的目标物,以便补充完整。然后按照 8:1:1 将已构建好的数据集划分为训练集、验证集和测试集,其中包含训练集 477 张图像,验证集 60 张图像,测试集 60 张图像。对标签数据进行目标位置及大小整体分析,得到标注框在图中的相对位置分布和目标相对大小图,如图 7 所示,目标树在图中分布较为均匀,目标的宽度大多数占图片宽度的 2%~5%,目标的高度大多数占 1%~4%。

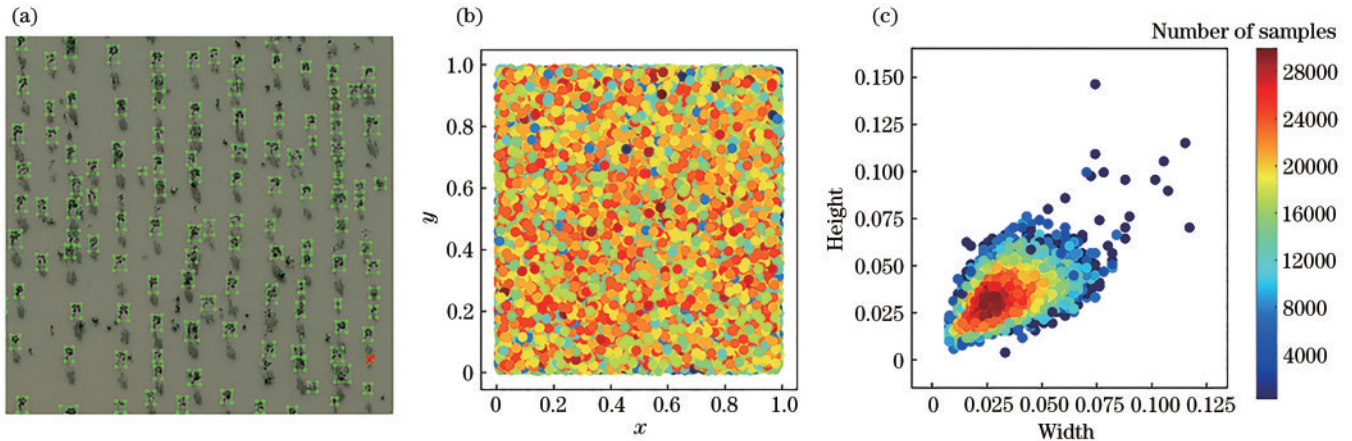


图 7 标注框概况图。(a)目标标注图;(b)归一化目标位置图;(c)归一化目标大小图

Fig. 7 Overview diagram of the label box. (a) Label of object; (b) normalized target location map; (c) normalized target size map

#### 3.2 模型训练环境

本研究的实验平台为自主配置服务器,64 位 Windows10 操作系统,处理器为 Intel(R) Xeon(R) CPU E5-2630 v3@2.40 GHz, NVIDIA Tesla K40c 显卡,显存 12G,内存 16 GB。本研究均基于 PyTorch 深度学习框架构建网络模型,开发环境为 PyTorch1.4, cuda10.1, python3.7。

使用 Adam 优化器进行训练,设置初始学习率为 0.01,权重衰减为 0.0001,动量为 0.9, batch size 设置为 16。所有实验均采用单尺度训练,图像输入大小为  $512 \times 512$  像素。依据模型自身的特点, YOLOv5 和 YOLOv5-ECA 训练的 epoch 均为 200,预训练模型为 YOLOv5x。

#### 3.3 结果分析

##### 3.3.1 训练过程损失值

在深度学习中,损失函数值可反映目标检测模型最终的预测结果和实际真实值之间的误差,用来分析和判断训练过程的优劣、模型的收敛程度以及是否过拟合等<sup>[22]</sup>。在所用的 PyTorch 框架中,损失函数可以看作是网络的某一层被放置于模型定义中,但在实际使用时更侧重于作为功能函数放到前向传播过程中<sup>[23]</sup>。

对原模型 YOLOv5 和改进后的 YOLOv5-ECA 的损失函数值进行对比分析,发现 200 epoch 过程中,两者总体都呈现下降趋势且最终趋于平缓,但 YOLOv5-ECA 的损失函数值基本上均低于 YOLOv5 原模型。YOLOv5 分别在 30 epoch 和 50 epoch 时出

现波动较大的小峰值, YOLOv5-ECA 在 20 epoch 处出现上升的突变后开始下降, 趋于平稳直至收敛。损失函数作为评价模型训练优劣的标准之一, 可以看出改进后的优化模型 YOLOv5-ECA 明显优于原模型 YOLOv5, 相同的训练轮数下 YOLOv5-ECA 有更低的损失函数值, 同时细节丢失较少, 特征学习能力更强。损失函数值随训练轮数的变化曲线如图 8 所示。

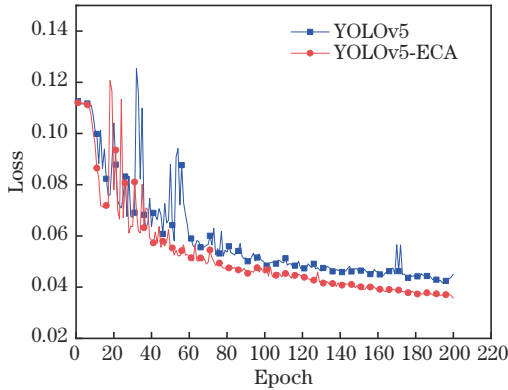


图 8 损失函数值随训练轮数的变化曲线

Fig. 8 Curve of the loss function value varying with epoch

对比模型 YOLOv4 在 400 epoch 的训练过程中, 损失函数没有特别明显的波动, 呈较为平缓的下降趋势且逐渐收敛。YOLOv4 模型本身较大, 相较 YOLOv5, 训练时间较长, 达到收敛需更多的训练轮数。YOLOv4 大致在训练轮数为 330 epoch 时损失函数值达到最低, 由于识别过程中目标只有一类, 因此 AP 等同于 mAP 值, 新造红树林单木目标检测的 AP 为 89.21%。YOLOv4 的 mAP 值如图 9 所示。

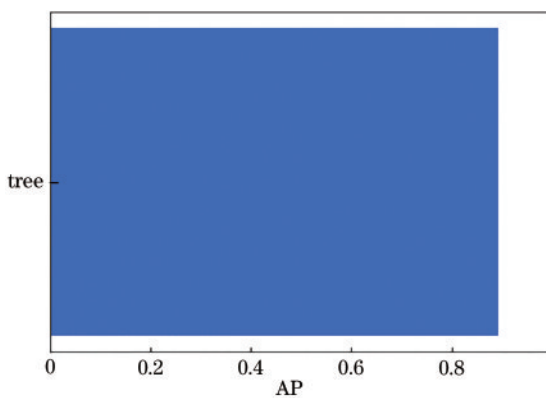


图 9 YOLOv4 的 mAP 值

Fig. 9 mAP value of YOLOv4

### 3.3.2 参数收敛结果

训练及评价参数通常体现了模型训练过程及目标检测的效果。将广义交并比损失 (Box Loss)、目标检

测损失 (Objectness Loss)、准确率 (Precision)、召回率 (Recall)、广义交并比阈值 0.5 下多类别平均精度值均值 (mAP@0.5)、变化交并比阈值下多类别平均精度值均值 (mAP@0.5:0.95) 作为主要的参数来判断模型收敛程度。Box Loss、Objectness Loss 参数值越接近 0, 表示训练效果越好; Precision、Recall、mAP@0.5、mAP@0.5:0.95 参数值越接近 1, 参数越收敛, 表示模型训练效果越好<sup>[24]</sup>。

实验结果如图 10 所示, YOLOv5 和改进后的模型 YOLOv5-ECA 的各个参数均逐渐收敛 (损失参数收敛至 0, 结果参数收敛至 1)。YOLOv5 训练过程中: Objectness Loss 在 50 epoch 附近出现较大的起伏; 结果参数 Recall、mAP@0.5、mAP@0.5:0.95 在 180 epoch 时都同时出现突降的现象, 之后继续平稳上升趋于收敛; Precision 在 0~60 epoch 会有不同程度的波动, 而后基本稳定。YOLOv5-ECA 的损失参数和结果参数均相对稳定, 突变的波动起伏较小, 平稳地趋于收敛, 结果参数均高于 YOLOv5。YOLOv5-ECA 添加了 ACON 或 ECA 结构, 以自适应学习的方式决定是否激活神经元, 可有效提高网络的检测性能, ECA 通过不降维方式使得通道和权重联系在一起, 算法检测性能更优<sup>[25]</sup>。

### 3.3.3 YOLOv5-ECA 模型检测性能

由准确率和召回率可以得出 mAP@0.5, mAP@0.5 也是衡量目标检测模型性能的主要评价指标。原模型 YOLOv5 的 mAP@0.5 为 91.2%, 改进后模型 YOLOv5-ECA 的 mAP@0.5 为 94.4%, 相较原模型 YOLOv5, 提高了 3.2 个百分点, 较 YOLOv4 提高了 5.19 个百分点, 同时训练损失也更低。YOLOv5 和 YOLOv5-ECA 的 Precision-Recall 曲线如图 11 所示。

分别对原模型 YOLOv5 和改进模型 YOLOv5-ECA 进行训练, 在所构建的无人机红树林单木数据集的测试集进行测试, 设定置信度阈值为 0.3, 使网络只检测高于置信度的幼苗目标。红树林单木目标检测效果如图 12 所示, 图 12(a) 为原模型 YOLOv5 的检测效果, 图 12(b) 为改进模型 YOLOv5-ECA 的检测效果。由图 12 可知, 因为新造红树林较小且分布密集, 在影像上不够清晰, 两种模型均有漏检的现象。综合对比发现 YOLOv5-ECA 的漏检现象较少, 一些边缘目标和模糊目标被检测到的概率更大。这是因为 SoftPool 池化操作会保留更多的细粒特征, 利于识别更小的目标木; ECA 能够降低在训练过程中的损失, 保留更多的学习特征, 提升对红树林单木幼苗的检测准确度, 较好地改善了对目标的特征提取能力, mAP@0.5 提升 3.2 个百分点。

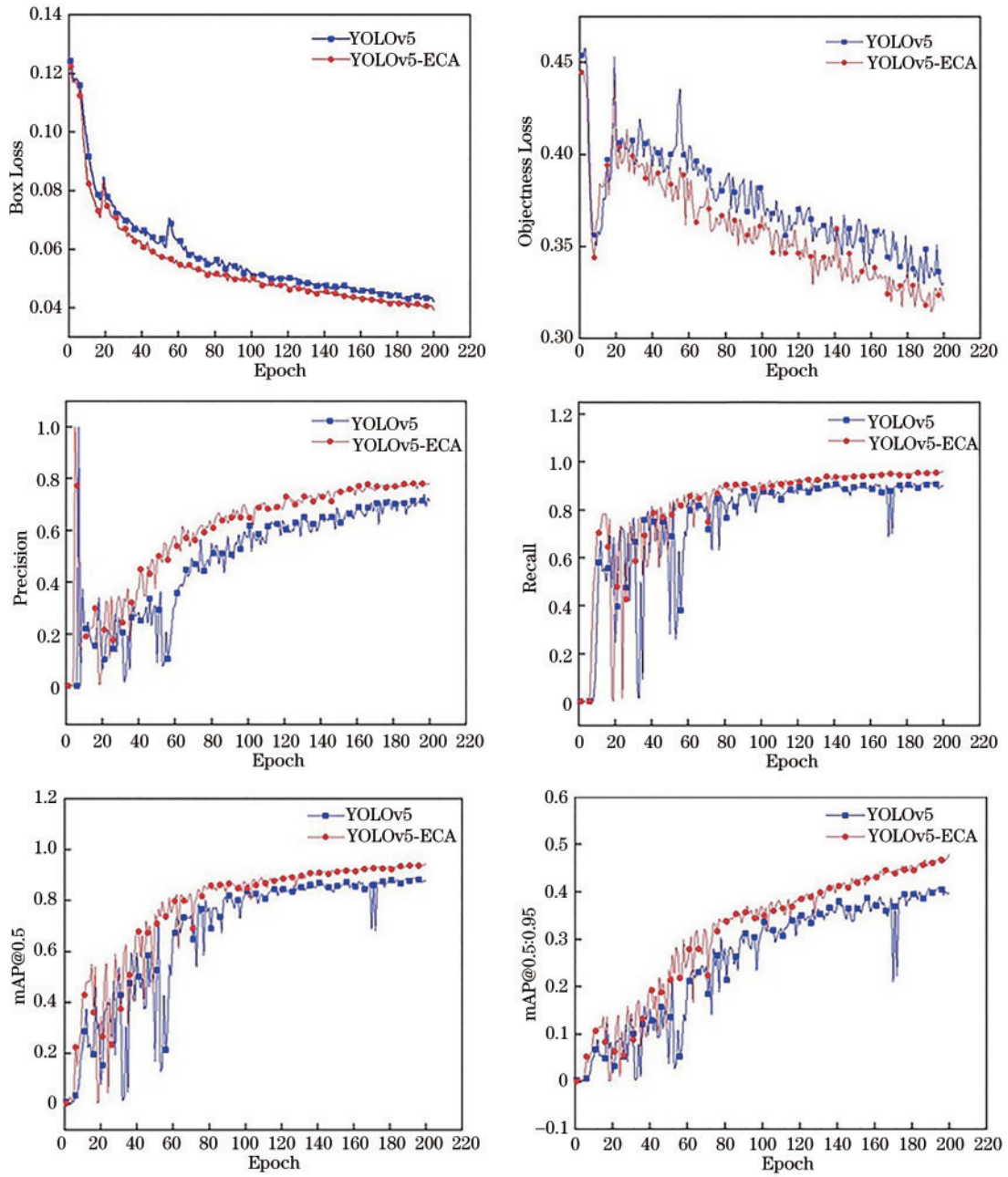


图 10 参数收敛结果对比

Fig. 10 Comparison of parameter convergence results

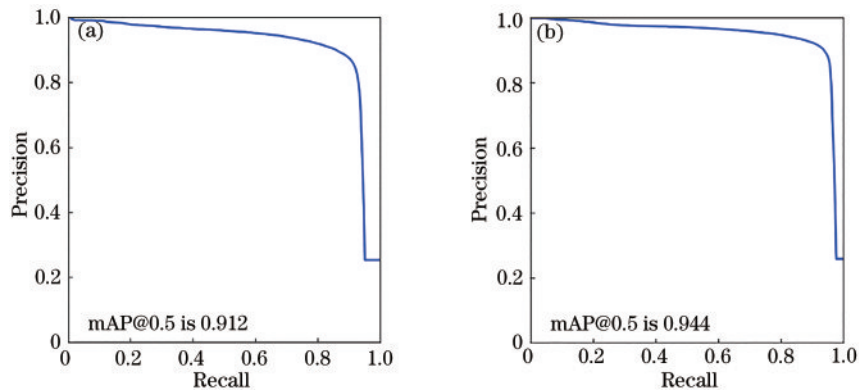


图 11 Recall-Precision 曲线。(a) YOLOv5; (b) YOLOv5-ECA

Fig. 11 Recall-Precision curves. (a) YOLOv5; (b) YOLOv5-ECA



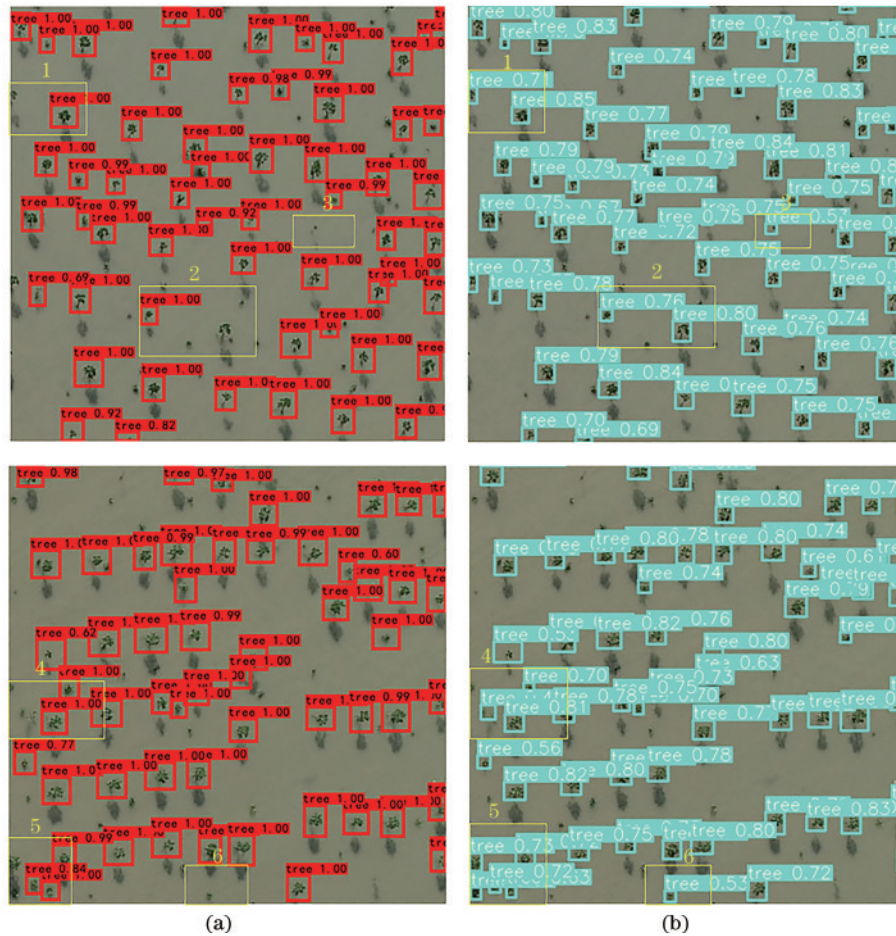


图 12 目标检测结果图。(a) YOLOv5 检测结果; (b) YOLOv5-ECA 检测结果

Fig. 12 Result of object detection. (a) Detection result of YOLOv5; (b) detection result of YOLOv5-ECA

## 4 结 论

针对现有从无人机影像识别新造红树林单木效率低、自动化程度不高等问题,以广东湛江红树林自然保护区的人工红树林单木幼苗为研究对象,提出了一种基于改进的 YOLOv5 模型识别新造红树林单木的方法 (YOLOv5-ECA)。YOLOv5-ECA 是在 YOLOv5x 的基础上进行优化和改进的,首先引入有效通道注意力机制,通过不降维和跨通道信息交互作用来降低模型的复杂性,同时增强特征的表达力;其次在 SPP 模块将池化操作替换为 SoftPool,使用 Softmax 进行加权池化,能够保持特征的表达性且操作是可微的,保留了整体感受野的特征信息,提高了算法的准确率。YOLOv5-ECA 算法经过训练后得到的 Box Loss、Objectness Loss、Precision、Recall、mAP@0.5:0.95 等参数值均趋于收敛并接近理想状态,且模型的平均查准率达 94.4%,相比原始模型提高了 3.2 个百分点,较 YOLOv4 提升了 5.19 个百分点,可快速、准确、自动地对新造红树林单木进行检测,与已有无人机树木传统检测方法相比,具有明显优势,同时退还坑塘用来人工种植红树林将成为必然趋势,因此 YOLOv5-ECA 将为提高新造红树林成活率检查精度和效率提供技术

支持。

由于红树林特殊的地理分布以及新造红树林幼苗自身尺寸较小等特点,加之在生长过程中存在幼苗被海水冲走而缺失的情况,所以在利用深度学习进行目标检测的过程中,只有成活且存在的树苗会被识别。若在有一定规则的人工种植下,未存活树苗会只剩下点位信息。因此在实现对红树林单木幼苗自动化识别的基础上,如何结合该地区造林规划统计未存活的点位信息,得到该地新造红树林的成活率,将是接下来继续研究的内容。

## 参 考 文 献

- [1] 丁冬静, 廖宝文, 管伟, 等. 东寨港红树林自然保护区滨海湿地生态系统服务价值评估[J]. 生态科学, 2016, 35(6): 182-190.  
Ding D J, Liao B W, Guan W, et al. Evaluation on service value of coastal wetland ecosystem in Dongzhai harbor mangrove nature reserve[J]. Ecological Science, 2016, 35(6): 182-190.
- [2] 高宇, 林光辉. 典型红树林生态系统藻类多样性及其在生态过程中的作用[J]. 生物多样性, 2018, 26(11): 1223-1235.  
Gao Y, Lin G H. Algal diversity and their importance in

- ecological processes in typical mangrove ecosystems[J]. *Biodiversity Science*, 2018, 26(11): 1223-1235.
- [3] 郑德璋, 廖宝文, 郑松发, 等. 广东省红树林及其保护的重要性[J]. *广东林业科技*, 1997, 13(1): 8-14.  
Zheng D Z, Liao B W, Zheng S F, et al. The mangrove swamps and their protection significance in Guangdong Province[J]. *Forestry Science and Technology of Guangdong Province*, 1997, 13(1): 8-14.
- [4] 李德仁, 李明. 无人机遥感系统的研究进展与应用前景[J]. *武汉大学学报·信息科学版*, 2014, 39(5): 505-513, 540.  
Li D R, Li M. Research advance and application prospect of unmanned aerial vehicle remote sensing system[J]. *Geomatics and Information Science of Wuhan University*, 2014, 39(5): 505-513, 540.
- [5] Pham T D, Yokoya N, Bui D T, et al. Remote sensing approaches for monitoring mangrove species, structure, and biomass: opportunities and challenges[J]. *Remote Sensing*, 2019, 11(3): 230.
- [6] 周振超, 李贺, 黄翀, 等. 红树林遥感动态监测研究进展[J]. *地球信息科学学报*, 2018, 20(11): 1631-1643.  
Zhou Z C, Li H, Huang C, et al. Review on dynamic monitoring of mangrove forestry using remote sensing[J]. *Journal of Geo-Information Science*, 2018, 20(11): 1631-1643.
- [7] 冯家莉, 刘凯, 朱远辉, 等. 无人机遥感在红树林资源调查中的应用[J]. *热带地理*, 2015, 35(1): 35-42.  
Feng J L, Liu K, Zhu Y H, et al. Application of unmanned aerial vehicles to mangrove resources monitoring[J]. *Tropical Geography*, 2015, 35(1): 35-42.
- [8] Navarro A, Young M, Allan B, et al. The application of Unmanned Aerial Vehicles (UAVs) to estimate above-ground biomass of mangrove ecosystems[J]. *Remote Sensing of Environment*, 2020, 242: 111747.
- [9] 杨礼. 融合 UAV 遥感影像与 SFM 点云的树木识别及参数提取[D]. 焦作: 河南理工大学, 2018.  
Yang L. Tree parameters extraction using UAV remote sensing image and SFM point cloud[D]. Jiaozuo: Henan Polytechnic University, 2018.
- [10] 胡馨月, 倪海明, 戚大伟. 基于无人机影像的树木株数提取[J]. *森林工程*, 2021, 37(1): 6-12.  
Hu X Y, Ni H M, Qi D W. Tree counts extraction based on UAV imagery[J]. *Forest Engineering*, 2021, 37(1): 6-12.
- [11] 周焱, 刘文萍, 骆有庆, 等. 基于深度学习的小目标受灾树木检测方法[J]. *林业科学*, 2021, 57(3): 98-107.  
Zhou Y, Liu W P, Luo Y Q, et al. Small object detection for infected trees based on the deep learning method[J]. *Scientia Silvae Sinicae*, 2021, 57(3): 98-107.
- [12] 张荣, 李伟平, 莫同. 深度学习研究综述[J]. *信息与控制*, 2018, 47(4): 385-397, 410.  
Zhang R, Li W P, Mo T. Review of deep learning[J]. *Information and Control*, 2018, 47(4): 385-397, 410.
- [13] 李娜, 武阳阳, 刘颖, 等. 基于多尺度注意力网络的行人属性识别算法[J]. *激光与光电子学进展*, 2021, 58(4): 0410025.  
Li N, Wu Y Y, Liu Y, et al. Pedestrian attribute recognition algorithm based on multi-scale attention network[J]. *Laser & Optoelectronics Progress*, 2021, 58(4): 0410025.
- [14] 赵杰伦, 张兴忠, 董红月. 基于尺度不变特征金字塔的输电线路缺陷检测[J]. *计算机工程与应用*, 2022, 58(8): 289-296.  
Zhao J L, Zhang X Z, Dong H Y. Defect detection in transmission line based on scale-invariant feature pyramid networks[J]. *Computer Engineering and Applications*, 2022, 58(8): 289-296.
- [15] 姚艳清, 程焱, 谢星星, 等. 多分辨率特征融合的光学遥感图像目标检测[J]. *遥感学报*, 2021, 25(5): 1124-1137.  
Yao Y Q, Cheng G, Xie X X, et al. Optical remote sensing image object detection based on multi-resolution feature fusion[J]. *National Remote Sensing Bulletin*, 2021, 25(5): 1124-1137.
- [16] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 779-788.
- [17] Zou Z X, Shi Z W, Guo Y H, et al. Object detection in 20 years: a survey[EB/OL]. (2019-05-13)[2021-02-05]. <https://arxiv.org/abs/1905.05055>.
- [18] Wang Q L, Wu B G, Zhu P F, et al. ECA-net: efficient channel attention for deep convolutional neural networks [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 11531-11539.
- [19] 汪亚妮, 汪西莉. 基于注意力和特征融合的遥感图像目标检测模型[J]. *激光与光电子学进展*, 2021, 58(2): 0228003.  
Wang Y N, Wang X L. Remote sensing image target detection model based on attention and feature fusion[J]. *Laser & Optoelectronics Progress*, 2021, 58(2): 0228003.
- [20] Stergiou A, Poppe R, Kalliatakis G. Refining activation downsampling with SoftPool[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV), October 10-17, 2021, Montreal, QC, Canada. New York: IEEE Press, 2021: 10337-10346.
- [21] 候瑞环, 杨喜旺, 王智超. 一种基于 YOLOv4-TIA 的林业害虫实时检测方法[J]. *计算机工程*, 2022, 48(4): 255-261.  
Hou R H, Yang X W, Wang Z C. Defect detection in transmission line based on scale-invariant feature pyramid networks[J]. *Computer Engineering*, 2022, 48(4): 255-261.
- [22] 高云鹏. 基于深度神经网络的大田小麦麦穗检测方法研究[D]. 北京: 北京林业大学, 2019.  
Gao Y P. Study on detection method of wheat ear in field based on deep neural network[D]. Beijing: Beijing Forestry University, 2019.
- [23] 宫一男, 谭孟雨, 王震, 等. 基于深度学习的红外相机动物影像人工智能识别: 以东北虎豹国家公园为例[J]. *兽类学报*, 2019, 39(4): 458-465.

- Gong Y N, Tan M Y, Wang Z, et al. AI recognition of infrared camera image of wild animals based on deep learning: Northeast Tiger and Leopard National Park for example[J]. *Acta Theriologica Sinica*, 2019, 39(4): 458-465.
- [24] Wang R J, Zhang J, Dong W, et al. A crop pests image classification algorithm based on deep convolutional neural network[J]. *Telkomnika*, 2017, 15(3): 1239.
- [25] 张涛, 张乐. 一种基于多尺度特征融合的目标检测算法[J]. *激光与光电子学进展*, 2021, 58(2): 0215003.
- Zhang T, Zhang L. Multiscale feature fusion-based object detection algorithm[J]. *Laser & Optoelectronics Progress*, 2021, 58(2): 0215003.