北京林业大学学报
JOURNAL OF BEIJING FORESTRY UNIVERSITY

基于Mask R-CNN和迁移学习的无人机遥感影像杉木单木树冠提取

谢运鸿 孙钊 丁志丹 罗蜜 李芸 孙玉军

UAV remote sensing image extraction of single tree crown of Chinese fir based on Mask R-CNN and transfer learning

Xie Yunhong, Sun Zhao, Ding Zhidan, Luo Mi, Li Yun, Sun Yujun

引用本文:
谢运鸿, 孙钊, 丁志丹, 罗蜜, 李芸, 孙玉军. 基于Mask R-CNN和迁移学习的无人机遥感影像杉木单木树冠提取[J]. 北京林业大学学报, 2024, 46(3):153-166. doi: 10.12171/j.1000-1522.20210343
Xie Yunhong, Sun Zhao, Ding Zhidan, Luo Mi, Li Yun, Sun Yujun. UAV remote sensing image extraction of single tree crown of Chinese fir based on Mask R-CNN and transfer learning[J]. Journal of Beijing Forestry University, 2024, 46(3):153-166. doi: 10.12171/j.1000-1522.20210343

在线阅读 View online: https://doi.org/10.12171/j.1000-1522.20210343

您可能感兴趣的其他文章

Articles you may be interested in

基于无人机影像的高郁闭度杉木纯林树冠参数提取

Extraction of tree crown parameters from high-density pure Chinese fir plantations based on UAV images

北京林业大学学报, 2020, 42(10): 20-26. https://doi.org/10.12171/j.1000-1522.20190386

利用无人机高分辨率影像进行树木高度提取

Tree height extraction using high-resolution imagery acquired from an unmanned aerial vehicle (UAV)

北京林业大学学报, 2017, 39(8): 17-23. https://doi.org/10.13332/j.1000-1522.20160428

机载与星载高分遥感影像单木树冠分割方法和适宜性对比

Comparison of individual tree crown extraction method and suitability of airborne and spaceborne high-resolution remote sensing images

北京林业大学学报, 2019, 41(11): 66-75. https://doi.org/10.13332/j.1000-1522.20180446

基于无人机遥感技术的废弃采石场立地条件类型划分

Site condition classification based on remote sensing technology of unmanned aerial vehicle in abandoned quarry

北京林业大学学报, 2018, 40(9): 90-97. https://doi.org/10.13332/j.1000-1522.20180051

面向钻蛀振动实时侦听的深度学习模型压缩

Deep learning model compression for real-time listening of boring vibration

北京林业大学学报, 2021, 43(6): 92-100. https://doi.org/10.12171/j.1000-1522.20200100

基于深度学习的5种树皮纹理图像识别研究

Research on image recognition of five bark texture images based on deep learning

北京林业大学学报, 2019, 41(4): 146-154. https://doi.org/10.13332/j.1000-1522.20180242

DOI:10.12171/j.1000−1522.20210343

基于 Mask R-CNN 和迁移学习的无人机遥感影像杉木单木树冠提取

谢运鸿　孙　钊　丁志丹　罗　蜜　李　芸　孙玉军
（北京林业大学森林资源和环境管理国家林业和草原局重点实验室，北京 100083）

摘要:【目的】利用无人机遥感影像对树冠进行自动化提取，获取高精度树冠信息。**【方法】**该研究提出一种基于 Mask R-CNN 和迁移学习的无人机影像单木树冠提取方法。首先，选用在 Faster R-CNN 基础上改进优化的 Mask R-CNN 实例分割模型，特征提取网络在 ResNet50 残差网络和 ResNet101 残差网络二者间选取最优。其次，引入迁移学习与 Mask R-CNN 一起训练，联合迁移学习的导向作用降低训练时间，提高训练精度。**【结果】**Mask R-CNN 模型的总体精度为 93.59%，用户精度为 65.46%，F1 分数为 76.05%，平均精度均值为 0.31；载入迁移学习后的 Mask R-CNN 模型在同等训练条件下比原模型的用户精度提升 29.53%，F1 分数提升 19.63%，平均精度均值提升 0.21；分别以 ResNet50 和 ResNet101 为特征提取网络的 Mask R-CNN 模型中，ResNet50 + Mask R-CNN 模型的总体精度、用户精度、F1 分数、平均精度均值各为 96.94%、95.57%、96.17%、0.54，ResNet101 + Mask R-CNN 模型的总体精度、用户精度、F1 分数、平均精度均值各为 96.20%、94.41%、95.19%、0.49；其中载入迁移学习的 ResNet50 + Mask R-CNN 模型在预测东西冠幅、南北冠幅、树冠面积与样方郁闭度的预测决定系数分别为 0.87、0.84、0.93 和 0.83。**【结论】**本研究提出的基于 Mask R-CNN 和迁移学习的方法得到了较为精准的树冠参数结果，为无人机遥感影像评估树木资源提供了一种快速高效的解决方案。

关键词：无人机；遥感影像；深度学习；Mask R-CNN；迁移学习；树冠提取
中图分类号：S757.2　　**文献标志码：**A　　**文章编号：**1000−1522(2024)03−0153−14
引文格式：谢运鸿，孙钊，丁志丹，等. 基于 Mask R-CNN 和迁移学习的无人机遥感影像杉木单木树冠提取 [J]. 北京林业大学学报，2024，46(3)：153−166. Xie Yunhong, Sun Zhao, Ding Zhidan, et al. UAV remote sensing image extraction of single tree crown of Chinese fir based on Mask R-CNN and transfer learning[J]. Journal of Beijing Forestry University, 2024, 46(3): 153−166.

UAV remote sensing image extraction of single tree crown of Chinese fir based on Mask R-CNN and transfer learning

Xie Yunhong　Sun Zhao　Ding Zhidan　Luo Mi　Li Yun　Sun Yujun
(National Forestry and Grassland Administration Key Laboratory of Forest Resources & Environmental Management, Beijing Forestry University, Beijing 100083, China)

Abstract: [**Objective**] The UAV image automatically extracts canopy for precise information. [**Method**] In this study, single tree crown extraction method from UAV images based on Mask R-CNN and transfer learning was proposed. Firstly, the optimized Mask R-CNN instance segmentation model based on faster R-CNN was selected, and the optimal feature extraction network was chosen between ResNet50 and ResNet101. Secondly, transfer learning was introduced to train with Mask R-CNN together, combined with the guiding role of transfer learning to reduce training time and improve training accuracy. [**Result**] The results showed that the overall accuracy of the Mask R-CNN model was 93.59%, the user accuracy, F1 score and mean average precision were 65.46%, 76.05% and 0.31, respectively. After adding transfer learning, the

收稿日期: 2021−09−03　**修回日期:** 2023−10−20
基金项目: 国家自然科学基金项目(31870620)，林业科学技术推广项目([2019]06)。
第一作者: 谢运鸿。主要研究方向：森林资源监测、林业遥感。Email: xyh1261233@gmail.com　地址：100083 北京市海淀区清华东路 35 号北京林业大学林学院。
责任作者: 孙玉军，博士，博士生导师。主要研究方向：森林资源监测、林业遥感。Email: sunyj@bjfu.edu.cn　地址：同上。
本刊网址: http://j.bjfu.edu.cn; http://journal.bjfu.edu.cn

user accuracy of Mask R-CNN model was increased by 29.53%, the F1 score was increased by 19.63%, and the average accuracy was increased by 0.21. In the Mask R-CNN model with ResNet50 and ResNet101 as feature extraction networks, the average values of overall accuracy, user accuracy, F1 score and mean average precision of the ResNet50 + Mask R-CNN model were 96.94%, 95.57%, 96.17% and 0.54, respectively. The average values of overall accuracy, user accuracy, F1 score and mean average precision of the ResNet101 + Mask R-CNN model were 96.20%, 94.41%, 95.19% and 0.49, respectively. The R^2 of ResNet50 + Mask R-CNN model loaded with transfer learning in predicting east-west crown width, north-south crown width, crown area and quadrat canopy density were 0.87, 0.84, 0.93 and 0.83, respectively. [**Conclusion**] The method based on Mask R-CNN and transfer learning proposed in this study obtains more accurate results of tree crown parameters, which provides a fast and efficient solution for tree resource assessment in UAV images.

Key words: UAV; remote sensing image; deep learning; Mask R-CNN; transfer learning; crown extraction

树冠作为树木光合作用的重要场所，是其生长的重要能量来源之一，学者往往通过研究树冠来获取树木的生长状况、生物量和木材的材性等信息[1]。利用遥感影像进行树冠提取能有效提高森林资源调查监测效率，但是早期遥感影像分辨率低，难以获得单木树冠信息，分辨树冠的形状和边界边缘非常困难[2]。随着遥感技术的进步及影像分辨率的提升，提取单木树冠信息的技术手段变得更为先进科学。相较于卫星影像，无人机获取的高空间分辨率影像能提供丰富的纹理信息，为森林资源调查提供了一种新思路[3]。无人机遥感具有高空间分辨率、高时间分辨率以及影像获取便捷等优点，已广泛应用于林业资源调查中[4]。

基于数字图像技术实现高分辨率遥感影像单木分割及树冠参数提取是当前研究热点，有效解决了树冠参数获取困难且工作量大的问题。当前基于高分辨率影像进行树冠参数提取方法主要有：(1)利用树冠独特的影像光谱和纹理信息与其他地物区分开并直接提取树冠的方法[5]；(2)基于几何光学特征和影像特征的算法(局部射线法、区域增长法、分水岭分割法和谷地跟踪法等)获取树冠的边缘[6−7]；(3)利用影像多尺度区域变化原理的面向对象多尺度分割方法提取树冠[8−10]。在这些研究方法中(1)由于算法本身的局限性不具有普适性，其中方法(2)由于影像纹理特征的复杂性及不确定性导致方法的稳定性下降，3 类方法在生产实践中均存在一定的短板，在单木树冠参数的自动提取方面有一定的局限性。

深度学习是近些年来快速发展的图像处理和数据分析技术，基于深度学习的分割算法在复杂背景下目标物分割中均取得了较好的效果[11−13]，深度学习分割算法主要有 3 类：语义分割[14]、实例分割[15]和全景分割[16]。目前深度学习学者主要沿着两个方向来研究实例分割[17]。第一个方向是先获取掩码建议再实行分割的检测方法，Dai 等[18]通过多任务网络级联（multi-task network cascades，MNCs)与卷积神经网络(convolutional neural networks，CNN)实现物体掩码建议和掩码分割；第二个方向是对语义分割结果进行约束和像素点聚类的分割方法，获得每个实例对象的范围，Bai 等[19]在 CNN 的基础上预测变换的分水岭能量图分割不同的实例物体，根据物体内部各像素点的能量值会随其离物体中心的距离而变化的原理区分不同的实例物体。

这些深度学习模型，在实际应用中，普遍存在数据量不足的情况，网络的运算非常复杂，在训练寻找权重的过程中需要花费大量的时间[20]。迁移学习[21]是一种基于源数据和目标数据、源任务和目标任务之间相似性的机器学习方法，利用在源数据领域中学习到的知识来解决目标数据领域中的任务，能够有效减少网络的训练时间，在同等训练条件下载入迁移学习能获得更优越的模型。李博等[22]为识别园艺作物叶部病害，分别使用 ResNet18、ResNet50 和 ResNet152 这 3 种残差网络作为特征提取网络对数据集进行迁移学习训练，训练模型对叶部病害的平均识别准确率都在 95% 以上。查志华等[23]采用迁移学习 Faster R-CNN 模型实现葡萄果穗的检测，分别采用 ResNet50、GoogleNet、VGG16、VGG19、AlexNet 5 种网络作为特征提取器，其中 VGG16 迁移学习模型检测平均精度可达 99.07%。近些年来，学者使用 Mask R-CNN 与迁移学习结合的方法在树冠提取领域中已取得较好的成果，Nuri 等[24]采用 Mask R-CNN 与迁移学习对多光谱影像单木株数提取的 F1 分数为 0.91；Anastasiia 等[25]利用 MaskR-CNN 与迁移学习对橄榄树(*Canarium album*)树冠的提取，单木株数提取的 F1 分数为 0.95；Hao 等[26]基于此方法试验了杉木(*Cuninghamia lanceolate*)人工幼龄林单木树冠的提取，将冠层高度模型(crown height model，CHM)

与 RGB 影像结合，单木株数提取 F1 分数为 0.85，对树高的预测决定系数为 0.87。

综上，本研究以无人机遥感影像中的杉木树冠为对象，构建分割树冠的 Mask R-CNN 结合迁移学习的模型，通过评价指标，探究 Mask R-CNN 和迁移学习在无人机林业遥感上单木树冠分割的可行性与有效性，提高单木树冠的分割精度。通过本研究中的分割模型对不同株数密度、不同郁闭度杉木林分的单木树冠提取试验，解决传统方法无法精准、自动化分割树冠的弊端，探索出一条适合杉木纯林单木树冠信息获取的新思路，为森林资源调查与监测提供技术支持。

1　研究区概况与研究方法

1.1　研究区概况与数据来源

研究区位于江西省赣州市大余县集体林区（25°19′N，114°08′E），具体如图 1a 和图 1b 所示。研究区属于典型的亚热带季风气候，降雨充沛。年平均气温为 19.6 ℃，年降雨量为 1 500 mm，极端最高温度为 41.2 ℃，极端最低温度为−10 ℃。研究的主要对象为杉木，共选取了 5 块具有代表性的区域，总面积为 10.7 hm^2，平均高为 9.5 m，平均胸径为 11.8 cm，平均郁闭度为 0.45，均为同龄林林分，生长状态良好。

数据来源于大疆精灵 4Pro 无人机（型号：Phantom4 Pro，品牌：DJI 大疆，国家：中国）拍摄的影像，于 2018 年 8 月采集，影像空间分辨率为 8 cm，相对起飞点飞行高度为 300 m。采用 DJI GO4 软件可以进行飞行姿态和角度的设置，使其保证航摄影像采集的同一性，因为该无人机采用的相机为非量测型全画幅相机，镜头畸变大，虽然可以根据相机畸变参数进行像片的畸变纠正，但是纠正过程会产生纠正误差，越靠近边缘，纠正误差越大，为提高精度，有效方法是加大像片重叠度，因此无人机航向重叠率和旁向重叠率均设置为 80%。为了减少图像的地理畸变，准备了 30 份红色油性漆纸板，以图像控制点应选择在具有清晰影像的表面特征上为原则进行野外布设像控点。后期采用瑞士 Pix4d 公司的无人机数据处理软件 Pix4dmapper 进行图像处理，将无人机拍摄的有效航片进行镶嵌并合成为数字正射影像，获得影像总覆盖面积为 105 hm^2，如图 1c。

以森林资源小班调查的矢量文件和地面实时调查数据为基础获取训练样本和验证样本。为了提高

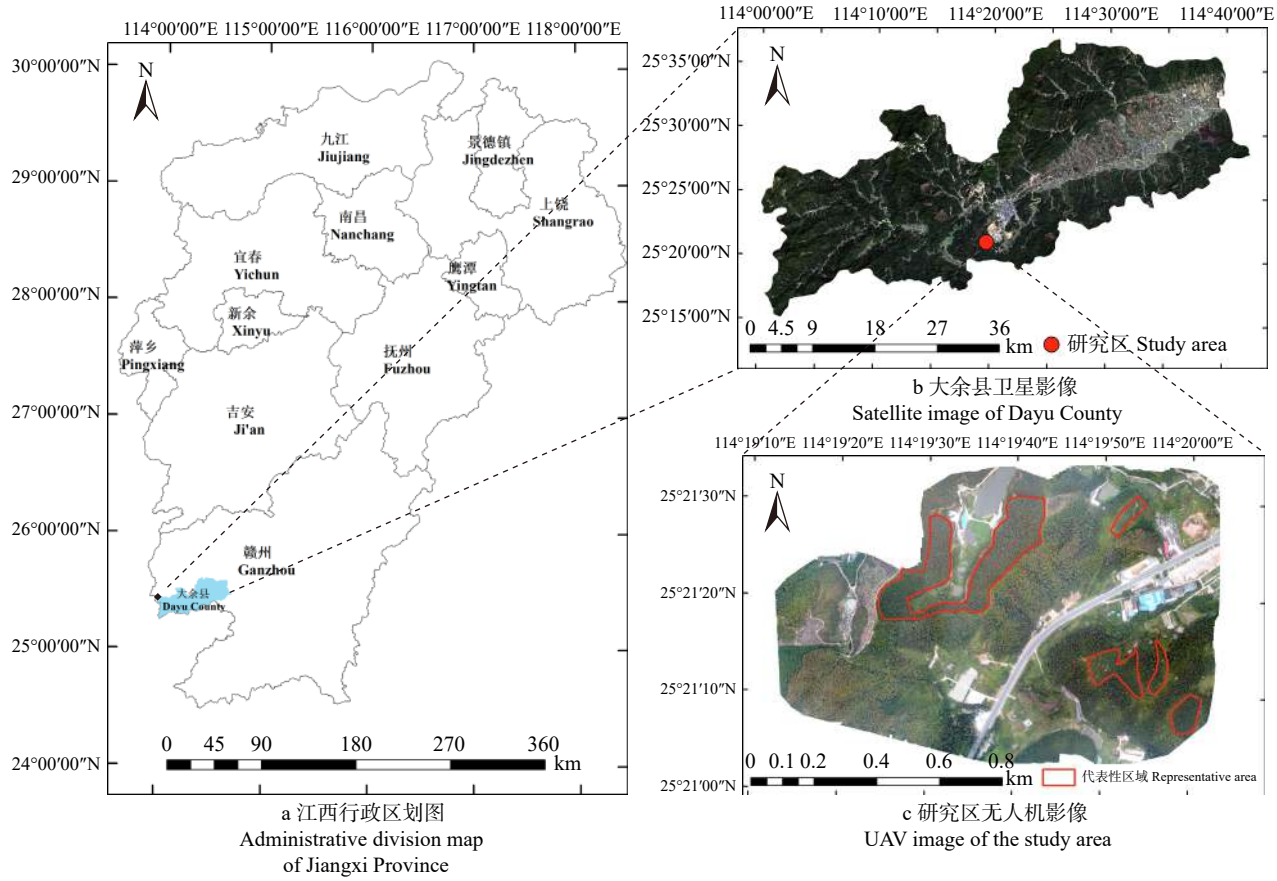


图 a 来自江西省自然资源厅−江西省县级以上行政区划图。Figure a is from the Department of Natural Resources of Jiangxi Province – the administrative division map of Jiangxi Province above the county level.

图 1　大余县地理位置及其 Landsant 8 卫星影像（可见光波段）和研究区无人机影像

Fig. 1　Geographic location and Landsant 8 satellite image (visible light bands) of Dayu County and UAV image of the study area

杉木覆盖面的识别精度，训练样本综合考虑了不同影像特征和光谱信息的树冠，样本尽可能多地涵盖全部杉木树冠类型，最终在图 1c 中的 5 个代表性区域截取了 58 个样方作为样本，包含了 2 480 个树冠，40 个区域用作训练，18 个区域用作验证，样本分辨率大小为 100 × 100 像素至 250 × 150 像素，面积为 0.008 6 ~ 0.032 2 hm^2。通过 VGG 图像标注（VGG image annotator，VIA）软件进行样本标签制作，样本形式包括图像文件（.jpg）和与图像对应的标签文件（.json），如图 2 所示，图 2a ~ c 为训练集部分林地图像，图 2d ~ f 为林地图像对应的标注图像。在训练集和验证集范围外，选取了 29 个实际地理面积为 0.013 5 hm^2 的样方用作模型的精度评价样本，作为测试集，测试集参数统计如表 1 所示。

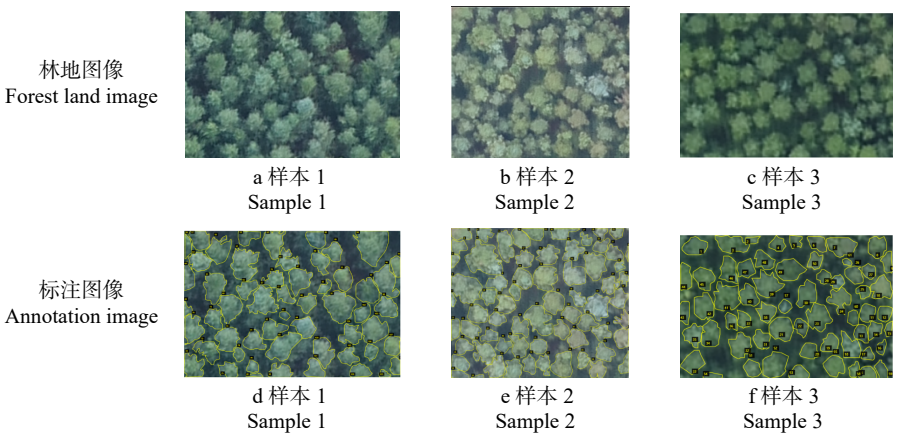


图 2 训练集示例图片

Fig. 2 Typical examples of training dataset

表 1 测试集林分参数统计

Tab. 1 Stand parameter statistics of test dataset

参数 Parameter	最小值 Min. value	最大值 Max. value	平均值 Mean	标准差 Standard deviation
东西冠幅 East-west crown width/m	0.29	2.34	1.15	0.37
南北冠幅 South-north crown width/m	0.38	2.41	1.18	0.38
树冠面积 Crown area/m^2	0.18	3.19	1.01	0.57
株数 Plant number	21.00	77.00	40.86	14.42
郁闭度 Canopy density	0.45	0.75	0.61	0.09

1.2 研究方法

1.2.1 Mask R-CNN

本研究采用 Facebook 公司（美国）在 2017 年提出的 Mask R-CNN 作为分割算法[15]，Mask R-CNN 是在 Faster R-CNN[27]的基础上优化的一种方法，模型输出分为边界框回归项、分类项和掩码项 3 个分支。由 Faster R-CNN 模型来完成边界框的回归和分类，全卷积神经网络（fully convolutional neural network，FCN）[28]来确定目标物体的掩码。首先将图片输入至特征提取模块获取特征图，分别选择 ResNet50 和 ResNet101 作为模块中的特征提取网络，ResNet 残差网络比传统的 AlexNex 网络更加简明、模块化，手动调节的超参数也更少，而且网络性能随网络深度的增加大幅提升[29]。其次两种网络得到的特征图分别利用区域提取网络（region proposal network，RPN）确定候选目标区域（region of interest，ROI），ROI 通过 ROI 对齐层映射的特征向量用作 3 个分支的预测。对多任务损失函数进行训练，通过学习不断降低损失函数的值，直至得到全局最优解，在训练过程中，所有样本的 ROI 均有一个损失函数（loss，L），其定义如式（1）所示。

$$L = L_{cls} + L_{box} + L_{mask} \quad (1)$$

式中：L_{cls} 为分类损失值；L_{box} 为边界框回归损失值；L_{mask} 为分割损失值。

（1）特征提取模块。本研究 Mask R-CNN 中的特征提取模块由特征提取网络和 FPN 两部分构成，包括自上而下和自下而上两条路径，特征提取模块同时也作为模型的主干。综合树冠特征的特殊性和其他特征提取器的适用性，采用特征表达能力较强的特征提取网络，因此选择了 ResNet50 和 ResNet101 两种残差网络。选择不同的特征提取网络，可以获得同一检测目标不同的特征属性。本试验以杉木树

冠为对象，杉木树冠在无人机影像中的形状大小、分布规律、纹理信息都有差异性，而深度学习能将目标的纹理特征进行合理学习得到很好的分类检测模型。FPN 能实现对不同尺度下的特征有效利用，使训练样本在合适的样本量下充分学习杉木树冠图像特征，能提高多尺度条件下精确快速检测小目标的能力。

（2）区域生成网络（RPN）。RPN 是基于 CNN 的一种网络，采用选择性搜索的方法得出建议区域，即通过不同尺寸的窗口在图像中进行滑动，对于每个尺寸都试图通过颜色、纹理或亮度将邻近的像素聚合在一起。RPN 的核心是锚，但是一个区域可以同时包含多个锚和多个正标签，利用建议区域与期望区域重叠部分的大小决定类别的确定，重叠部分的选取由损失函数 $L(p_i, t_i)$ 获取。

$$L(p_i,t_i)=\frac{1}{N_{\mathrm{cls}}}\sum_{i=1}^{m}L_{\mathrm{cls}}\left(p_i,p_i^*\right)+\lambda\frac{1}{N_{\mathrm{reg}}}\sum_{i=1}^{m}p_i^*L_{\mathrm{reg}}\left(t_i,t_i^*\right) \tag{2}$$

式中：i 为锚的索引；N_{cls}、N_{reg} 为分类层数和回归层数；p_i 为目标锚的预测概率；p_i^* 为期望区域的预测概率；t_i 为预测边界框的 4 个参数化坐标参数；t_i^* 为正标签锚相关的期望区域的坐标参数；L_{cls} 为目标与背景 2 个类的对数损失函数；L_{reg} 为目标与背景的回归损失函数；λ 为权重；m 为期望区域总数量。

（3）候选目标区域对齐（ROI Align）。在 Faster R-CNN 模型中，ROI 池化会导致空间位置错位，所以 Mask R-CNN 模型利用 ROI Align 去优化空间位置错位的情况。ROI 池化的区域建议是从原始图像中提取 ROI，得到特征小块在特征图像对应的位置，然后原始图像中的位置和步长乘积的比值，之后对其进行舍入取整，从而就会影响相应区域中的特征小块因舍入而未对齐。因此，ROI Align 不采用计算结果舍入取整，而是通过双线性插值确定原始图像感兴趣区域内各点的特征值，然后进行池化等运算提高精度。在本试验中，ROI Align 能消除由算法方面提取树冠造成的空间位置错位，保证每个树冠的地理位置能精准定位。

1.2.2　迁移学习

迁移学习的原理是给定一个有标注的辅助领域 D_{s} 和学习任务 T_{s}，无标注的目标领域 D_{t} 和学习任务 T_{t}，其目的是在 $D_{\mathrm{s}}\neq D_{\mathrm{t}}$ 或 $T_{\mathrm{s}}\neq T_{\mathrm{t}}$ 的条件下，通过 D_{s} 所学习到的知识完成对 D_{t} 的学习，从而降低目标领域的泛化误差[30]。而迁移学习可以解决训练数据不足的问题，新模型利用学到的参数以加快模型的训练，从而本研究将迁移学习引入到杉木树冠的分割模型中。利用免费公开的大数据集 MS COCO 上预先训练好的权重文件用作 Mask R-CNN 特征提取网络的初始导向训练文件，与 MS COCO 数据集不同的小样本正适用高层卷积神经网络，这样的网络训练不仅可以减少样本量，提高分类精度，而且可以提高运行效率。

1.2.3　Mask R-CNN 与迁移学习的单木树冠分割模型构建

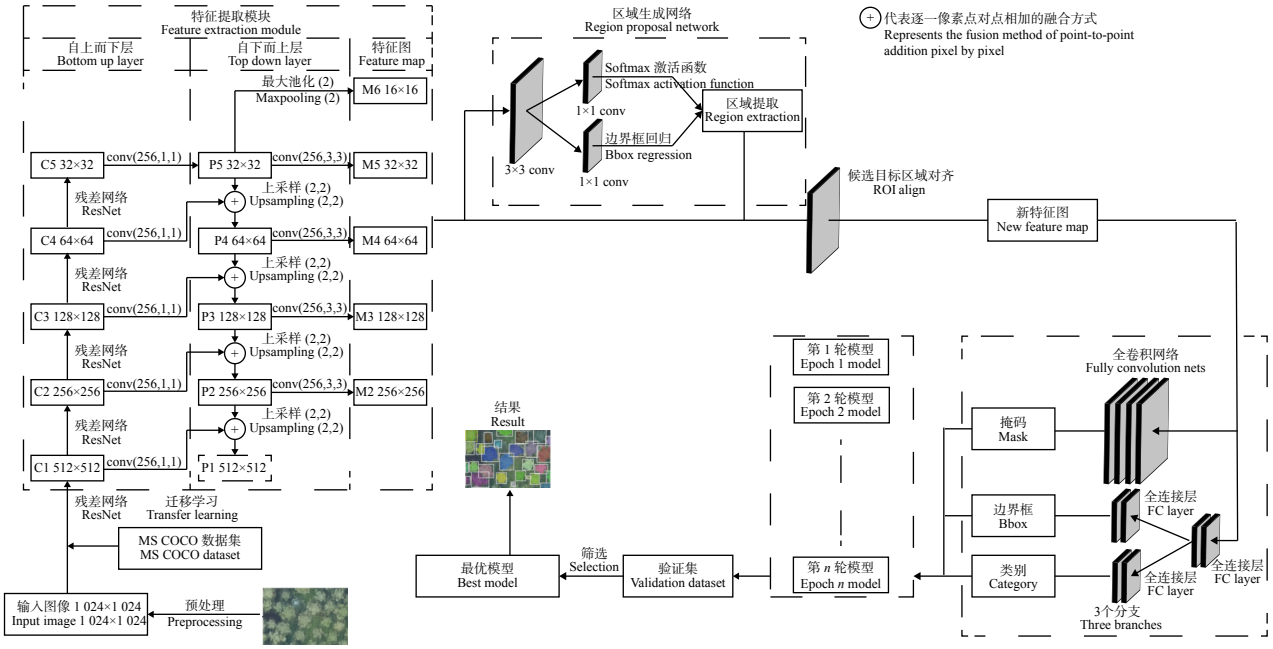
本研究中，Mask R-CNN 以 ResNet50 与特征金字塔网络（feature pyramid network，FPN）为主干的模型简称为 MRCNN50，Mask R-CNN 以 ResNet101 与 FPN 为主干的模型简称为 MRCNN101，Mask R-CNN 以 ResNet50 与 FPN 为主干并载入迁移学习的模型简称为 MRCNN50-T，Mask R-CNN 以 ResNet101 与 FPN 为主干并载入迁移学习的模型简称为 MRCNN101-T。

（1）模型的输入与输出。样本分为训练样本与验证样本，将样本的 jpg 和 json 文件输入到模型中。为保持样本 jpg 图像长宽一致采取双线性插值及零填充方式控制图像分辨率大小为 1 024 × 1 024 像素。模型的输出格式为 jpg 格式，每个输出包括单木的分类名称、分类置信度、边界框及以树冠范围为基准的掩码，其中分类名称与分类置信度作为结果图像会因为过多显示显得杂乱无序，甚至遮掩掉树冠分割边界，所以只保留了边界框和掩码。

（2）训练平台、初始参数的选取。Mask R-CNN 中的特征提取器分别选取 ResNet50 + FPN 和 ResNet101 + FPN，独立训练 2 次。分类分为树与背景，批量为 1。试验过程中均参照 Mask R-CNN 文献[15]中原型框架和参数设置方式，但是 Mask R-CNN 模型使用默认的学习率为 0.02，经常导致权重失衡，尤其是使用小批量试验下，因此该试验选择使用较小的学习率 0.001 和梯度裁剪，可以有效提高收敛速度。其中轮数设置为 80，每一个轮数里面的迭代次数设置为 200，各个损失权重初始值设为 1，权重衰减系数为 0.000 1。本研究使用 Windows10 Pro（64 位）系统，显示核心为 NAVIDIA RTX2070（8 GB 显存），中央处理器为 i7-9750，内存为 40 G，基于 Python3.6 编程平台，采用开源的 Tensorflow1.9.0 与 Keras2.2.4 机器学习框架构建模型组件，包括样本预处理、特征提取网络、边界框与掩膜构建及多任务损失函数，并与 MS COCO 数据集协同执行迁移学习。

（3）模型的训练与验证。在模型训练阶段，通过训练样本对具有初始参数和迁移学习文件（分别为 ResNet50 和 ResNet101 残差网络的迁移学习文件）的 CNN 进行迭代训练，并根据训练结果的一系列指标对训练过程进行分析，从而不断修正和优化训练模型的相关训练参数，最终得到树冠分割模型。模

型验证阶段，将待检测样本输入分割模型得到检测结果。模型的训练与验证包含以下 4 个步骤(图 3)。步骤 1：将训练样本输入到特征提取模块中，并加载来自 MS COCO 数据集权重文件，得到相应训练样本图像的特征图，通过特征图获得多个候选 ROI。步骤 2：所有的候选 ROI 被送入 RPN 进行分类预测和边界框回归，删除部分候选 ROI，其余的 ROI 通过 ROI 对齐操作进行精准映射，对齐后的特征图用于边界框回归项、分类项和掩码项的预测。步骤 3：重复步骤 2 至 80 次，即训练 80 轮，共迭代 16 000 次，训练完所有样本得到第 1 轮到第 80 轮的权重文件。步骤 4：对部分权重文件依次利用验证样本对单株树冠分割模型进行验证，保存权重与偏差值固定模型用于后续的最优模型筛选。



特征提取模块中的 conv (256,1,1) 和 conv (256,3,3) 分别代表步长为 1 的 256 × 1 × 1 卷积核和步长为 1 的 256 × 3 × 3 卷积核；Upsampling (2,2) 代表 2 倍上采样操作；Maxpooling (2) 代表步长为 1 的 2 × 2 的滤波器的最大池化操作；C1 ~ C5、P1 ~ P5 和 M2 ~ M6 均为特征图；P1 特征图尺寸过大通常被舍弃。The conv (256,1,1) and conv (256,3,3) in the feature extraction module respectively represent 256 × 1 × 1 convolution kernel (stride = 1) and 256 × 3 × 3 convolution kernel (stride = 1); upsampling (2,2) represents 2 times upsampling operation; maxpooling (2) represents maxpooling operation with 2 × 2 filter (stride = 1); C1−C5, P1−P5 and M2−M6 are feature maps; P1 feature map is usually discarded, because its size is too large.

图 3 Mask R-CNN 与迁移学习模型结构示意图

Fig. 3 Model of Mask R-CNN with transfer learning

1.2.4 树冠参数的自动化提取

（1）东西、南北冠幅及株数的计算。根据分割结果中的边界框数组，其维度组成为(n, 4)，n 为模型提取株数，是 n 个(a, b)二维数组，该数组长 a 为(x_2-x_1)，宽 b 为(y_2-y_1)。由于所有样本内部像素空间的方位与地理空间的东南西北一致，假设东西冠幅即为边界框长的地理长度，南北冠幅即为边界框宽的地理长度，东西冠幅、南北冠幅、模型提取株数依次如式(3)～(5)所示。

$$\mathrm{EW} = k(x_2 - x_1) \tag{3}$$

$$\mathrm{SN} = k(y_2 - y_1) \tag{4}$$

$$n = \mathrm{len}(\mathrm{roi}) \tag{5}$$

式中：EW为东西冠幅，m；SN为南北冠幅，m；n 为模型提取株数；k 为单个像素边长的地理长度，m；len()为计算数组长度的函数；roi为边界框数组。

（2）各单木树冠面积及样方郁闭度的计算。根据分割结果中返回的掩膜数组，其维度组成为(a, b, n)，n 为模型提取株数，是由 False 和 True 两种元素构成的，False 代表非树冠像素，True 代表树冠像素，索引树冠像素的数量作为得到树冠面积的基础，第 i 株数的树冠面积和郁闭度的计算如式(6)～(7)所示。

$$S_i = m(\mathrm{mask}[:,:,(i-1):i] == \mathrm{True}) \times k^2 \tag{6}$$

$$w = \frac{m(\mathrm{mask} == \mathrm{True})}{n \times l} \tag{7}$$

式中：S_i为第 i 株树的树冠面积，m^2；w 为郁闭度；$m(\mathrm{mask}[:,:,(i-1):i] == \mathrm{True})$为根据 [:,:,(i−1):i] 索引的掩膜数组中等于 True 的元素数量；k 为单个像素边长的地理长度，m；$m(\mathrm{mask} == \mathrm{True})$为掩膜数组中

等于 True 的元素数量；n 为模型提取株数；l 为样方影像像素个数。

1.2.5　精度评价指标

本研究采用交并比、平均精度和平均精度均值验证实例分割模型的性能，联合总体精度、用户精度、均方根误差、F1 分数和预测决定系数对树冠的分割效果进行评价。

交并比（intersection over union，IoU）为参考树冠像素区域和预测树冠像素区域的交集与参考树冠像素区域和预测树冠像素区域的并集的比值。

IoU 从 0.5 开始，以 0.05 为步长设置 IoU 阈值，每一个 IoU 阈值对应一个精度值，所得 10 个精度的均值即为平均精度（average precision，AP）。

平均精度均值（average precision mean，mAP）为所有参考树冠区域 AP 值的平均值。AP 系列的指标不仅能衡量正确分割的数量，还可以评价目标分割失误，在目标检测任务中经常作为评价的指标，其值越高代表模型分割性能越好。

借助 VIA 图像标注软件通过人工目视判读获得参考图，参考图即正确的分割结果，将分割结果与参考图进行比较，并通过它们之间的差异来定性分割结果，当 IoU ⩾ 0.5 时记为正确分割树冠，当 IoU < 0.5 时记为错误分割树冠。单木株数的总体精度、用户精度、均方根误差和 F1 分数定义如式（8）～（11）所示。

$$A_{\mathrm{O}} = \frac{N_{\mathrm{c}}}{N_{\mathrm{d}}} \times 100\% \tag{8}$$

$$A_{\mathrm{U}} = \frac{N_{\mathrm{c}}}{N_{\mathrm{r}}} \times 100\% \tag{9}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{m=1}^{n}(N_{\mathrm{c}m} - N_{\mathrm{r}m})^2}{o}} \tag{10}$$

$$F = \frac{2A_{\mathrm{O}}A_{\mathrm{U}}}{A_{\mathrm{O}} + A_{\mathrm{U}}} \tag{11}$$

式中：A_{O}为总体精度；A_{U}为用户精度；RMSE为均方根误差；F 为 F1 分数（两种精度的加权平均值）；N_{c}为自动分割结果中被正确分割的树冠个数；N_{d}为自动分割结果中树冠个数总数；N_{r}为参考图中树冠个数总数；o 为样本数量；$N_{\mathrm{c}m}$为第 m 个样本的 N_{c} 值；$N_{\mathrm{r}m}$为第 m 个样本的 N_{r} 值。

预测决定系数 R^2 表示目标变量的预测值与实际值之间的相关程度平方的数值，数值范围从 0 至 1，决定系数越接近 1 代表预测性能越好，定义如式（12）所示。

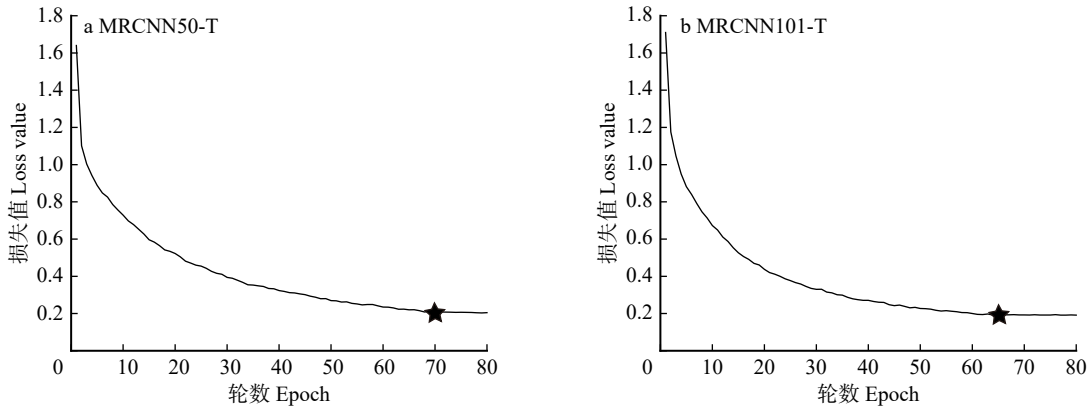
$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{12}$$

式中：n 为预测值或实际值个数；$\hat{y}_i$为第 i 个预测值；$\bar{y}$为实际值的平均值；y_i为第 i 个实际值。

2　结果与分析

2.1　模型权重选择

通过 Tensorflow 可视化工具 Tensorboard 所绘制的损失函数图（图 4），分别判断 MRCNN50-T 和 MRCNN101-T 模型的最终轮数，最终轮数对应的权重文件为最终权重文件。但是根据分割实验结果，最终权重文件所得结果往往会产生过分割现象，在单木分割实验中提取株数为关键因素，其次是每株单木分割的范围准确性。mAP 和 RMSE 指标能有效衡量模型提取株数的准确性，两个指标协同选择最优权重文件。MRCNN50-T 模型的损失值随轮数变化过程如图 4a 所示，从第一轮开始模型损失值逐步下降，下降速度逐步减小，第 70 轮损失值为 0.209 且趋于稳定（如图 4a 中星号位置所示），损失值不再明显下降，判定 MRCNN50-T 模型的最终轮数为 70。MRCNN101-T 模型的损失值随轮数变化



星号★下的轮数代表该模型的最终轮数。The epoch under the asterisk represents the final epoch of model.

图 4　MRCNN50-T 模型和 MRCNN101-T 模型损失函数图

Fig. 4　Loss function graphs of MRCNN50-T model and MRCNN101-T model

过程如图 4b 所示，MRCNN101-T 模型的最终轮数为 65。将这 2 个模型最终轮数与其前后九轮的 mAP 和 RMSE 进行比较，结果如表 2 所示，MRCNN50-T 模型的第 67 轮以 mAP 值（0.541）降序排列第 3，RMSE 值（2.344）升序排列第 2，该轮次为模型的最优轮次；MRCNN101-T 模型的第 61 轮以 mAP 值（0.541）降序排列第 2，RMSE 值（2.344）升序排列第 1，MRCNN101-T 模型的最优轮次为第 61 轮。

2.2 传统模型与深度学习模型分割结果评价

选用迭代自组织（iterative self-organization，ISO）聚类和均值漂移为传统分割模型，对比图 5 中深度学习模型的树冠分割结果，图 5b ISO 聚类[31]树冠分割结果中单木较为破碎化，相邻单木之间无法很好地分开；图 5c 均值漂移[32]树冠分割结果中各单木虽然没有和 ISO 聚类分割的破碎化，但也被划分为多个部分，树冠整体分离化；MRCNN50 和 MRCNN101 均迭代 16 000 次，如图 5d 和图 5e 所示，二者分割效果均比 ISO 聚类与均值漂移好，树冠边缘较为精准，但是存在严重的漏分割与树冠提取重叠的问题。MRCNN50-T 和 MRCNN101-T 不仅对邻近的单木之间有较好的分割效果，对图像边缘的非整体单木也有较好的分割效果，而且边缘不失真（图 5f 和图 5g）。对比图 5b ~ g，Mask R-CNN 经过高层卷积特征提取网络精确地捕捉到了单木树冠与背景、单木树冠之间的区分信息。

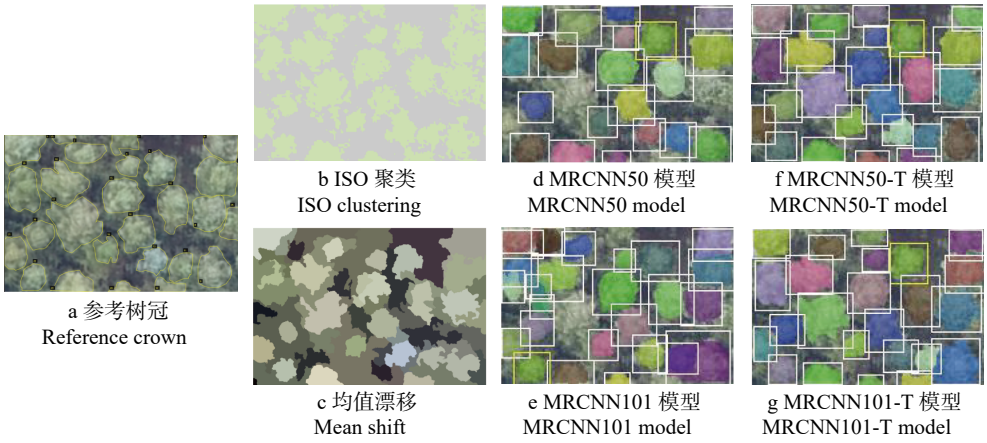
传统模型与深度学习模型的精度对比如表 3 所示，传统的 ISO 聚类和均值漂移的总体精度、用户精度、F1 分数分别为 70.02%、67.67%、68.82% 和 65.89%、75.45%、70.35%，4 个深度学习模型的总体精度、用户精度、F1 分数的平均值分别为 95.08%、80.22%、85.87%。结果表明：深度学习与 ISO 聚类和均值漂移相比较在分割效果表现更优，深度学习经过大量的卷积运算能比传统分割提取到更多的图像信息，通过更多的图像信息获得更准确的分割结果。2 个 MRCNN 模型的总体精度、用户精度、F1 分数、平均精度均值的平均值分别为 93.59%、65.46%、76.05%、0.31，对比另外 2 个 MRCNN-T 模型的总体精度、用户精度、F1 分数、平均精度均值的平均值分

表 2 MRCNN50-T 模型和 MRCNN101-T 模型权重选择参数统计

Tab. 2 Weight selection parameter statistics of MRCNN50-T model and MRCNN101-T model

MRCNN50-T			MRCNN101-T		
轮数 Epoch	平均精度均值 Average precision mean	均方根误差 RMSE	轮数 Epoch	平均精度均值 Average precision mean	均方根误差 RMSE
79	0.534	2.345	74	0.465	4.745
78	0.533	2.451	73	0.472	4.631
77	0.526	2.456	72	0.479	4.579
76	0.536	2.356	71	0.473	4.448
75	0.533	2.523	70	0.472	4.814
74	0.532	2.626	69	0.479	4.691
73	0.520	2.564	68	0.466	4.662
72	0.530	2.658	67	0.478	4.568
71	0.530	2.554	66	0.472	4.562
70	0.539	2.334	65	0.475	4.766
69	0.529	2.414	64	0.480	4.552
68	0.540	2.988	63	0.483	3.624
67*	0.541	2.344	62	0.483	5.397
66	0.542	4.156	**61***	0.490	3.096
65	0.532	2.822	60	0.484	4.552
64	0.543	3.595	59	0.490	3.591
63	0.541	2.626	58	0.481	4.000
62	0.526	3.978	57	0.482	3.648
61	0.538	2.760	56	0.491	3.737
60	0.534	3.464	55	0.489	3.619

注：加粗并带*字符的轮数代表该模型的最优权重轮数。Note: the epoch in bold with * characters represents the optimal epoch of model weight.



a 中的黄线即为每个树冠的边缘；b 中图像被分为灰色和绿色，绿色区域为树冠；c 中图像被分为多类，浅颜色区域代表树冠；d～e 和 f～g 中的不同颜色的区域代表不同单木树冠，白线代表树冠区域的边框线。The yellow line in figure a is the edge of each canopy; the image in figure b is divided into gray and green, and the green area is the canopy; the image in figure e is divided into multiple categories, and the light-colored area represents the canopy; the areas of different colors in figure c−d and figure f−g represent different single tree canopy, and the white lines represent the border lines of the canopy area.

图 5　参考树冠和 6 个模型的树冠分割结果

Fig. 5　Reference crowns and crown segmentation results of 6 models

表 3　传统模型与深度学习模型精度对比

Tab. 3　Comparison of accuracy between traditional models and deep learning models

模型 Model	总体精度 Overall accuracy/%	用户精度 User accuracy/%	F1分数 F1 score/%	mAP
ISO聚类 ISO clustering	70.02	67.67	68.82	
均值漂移 Mean shift	65.89	75.45	70.35	
MRCNN50	94.02	69.51	79.28	0.36
MRCNN101	93.15	61.40	72.82	0.25
MRCNN50-T	96.94	95.57	96.17	0.54
MRCNN101-T	96.20	94.41	95.19	0.49

注：mAP表示深度学习目标检测独有的评价指标，因此ISO聚类和均值漂移不做运算。Notes: mAP represents the unique evaluation index of deep learning target detection, so ISO clustering and mean shift are not calculated.

别为 96.57%、94.99%、95.68%、0.52，不载入迁移学习的 Mask R-CNN 模型精度均低于载入迁移学习的 Mask R-CNN 模型，试验结果表明同等训练条件下迁移学习在一定程度能提升模型的分割性能。MRCNN50 和 MRCNN101 载入迁移学习后总体精度、用户精度、F1 分数、平均精度均值 4 个指标分别提升了 2.92%、26.06%、16.28%、0.18 和 2.18%、33.01%、22.37%、0.24。迁移学习对模型的用户精度、F1 分数与平均精度均值较其他指标相比有大幅提升，说明使用 MS COCO 的预训练权重文件进行载入学习对 Mask R-CNN 模型减少漏检测率有显著影响，同时也降低了训练条件，使小样本实例分割成为可能。MRCNN101-T 作为更深层网络模型各评价指标均低于 MRCNN50-T，尤其是 mAP 比其低 0.05。ResNet101 深层网络运算次数更多，其参数比 ResNet50 更多更复杂，深层网络更难以优化，导致 MRCNN101 的性能比 MRCNN50 差。

2.3　单木株数计量与精度评价

本研究将 29 个不同株数密度的样方影像输入到 4 个深度学习模型进行测试，得到分割后的图片与提取株数，分割结果与人工勾画无人机遥感影像后获得的参考树冠图进行逐树冠 IoU 的计算判断后得到每个样方的正确分割株数（即所有 IoU ≥ 0.5 的树冠个数）。为验证迁移学习对整体模型的有效性，比较 MRCNN50、MRCNN50-T、MRCNN101、MRCNN101-T 这 4 个模型的单木株数计量结果，计量结果包括参考株数（人工勾画结果中的树冠个数）、提取株数（分割结果中的树冠个数）和正确分割株数。由图 6a、c 可知：MRCNN50 与 MRCNN101 的参考株数在 29 个样方中都达不到提取株数的数值，MRCNN50 模型的试验结果中参考株数与提取株数的差值范围为 [3，29]，MRCNN101 模型的试验结果中参考株数与提取株数的差值范围为 [1，41]，尤其是当参考株数大于 35 时，参考株数与提取株数

的差值呈现不断增大的趋势，差值最大为 41。结果表明 MRCNN50 与 MRCNN101 模型的稳定性随着样方株数的增加呈下降的趋势。根据测试样本和分割结果分析得知：在参考株数大于 35 的测试集中存在明显树冠相互重叠的现象，MRCNN50 和 MRCNN101 模型容易将 2～4 个重叠的树冠识别为 1 个树冠，致使样本提取株数会少于参数株数，而且随着参数株数越大的测试集，树冠重叠的数量越多，重叠比例越大，树冠之间的边缘越淡化，导致参考株数与提取株数的差值有不断增大的趋势。由图 6b、d 可知：MRCNN50-T 模型的试验结果中参考株数与提取株数的差值范围为 [0，4]，MRCNN101-T 模型的试验结果中参考株数与提取株数的差值范围为 [0，7]，2 个模型的差值范围都在试验的可接受范围内。结果表明 MRCNN50 与 MRCNN101 模型在不同株数的样方测试上具有一定的稳定性。进一步分析图 6 的结果表明：MRCNN50-T 与 MRCNN101-T 的提取株数和正确分割株数差值相比 MRCNN50 与 MRCNN101 较小，MRCNN50-T 与 MRCNN101-T 模型分割错误数量也控制在 [0，3] 的范围内。在载入迁移学习的情况下，模型不仅在株数提取性能上得到提升，还能减少分割错误的数量，并正确分离重叠状态下的各树冠。这从提取数量层面上验证了迁移学习的有效性。

2.4 单木树冠信息提取与精度评价

对 29 个样方通过系统抽样抽取 310 个树冠，分别使用 MRCNN50-T 与 MRCNN101-T 进行冠幅和树冠面积的预测，参考树冠面积、双向冠幅分别为横轴，预测树冠面积、双向冠幅分别为纵轴，进行参考数据和预测数据的散点图绘制，结果如图 7 和图 8 所示。各模型的预测值与参考值吻合较好，预测决定系数均在 0.81～0.93 之间，拟合曲线的斜率均大于 0.7。基于 MRCNN50-T 东西冠幅、南北冠幅和树冠面积的预测 R^2 分别为 0.87、0.84 和 0.93，各拟合曲线的斜率分别为 0.85、0.81 和 0.93；基于 MRCNN101-T 的预测 R^2 分别为 0.85、0.81 和 0.88，各拟合曲线的斜率分别为 0.80、0.75 和 0.89。2 个模型在预测树冠面积方面有着较好的性能，预测决定系数和拟合曲线斜率均接近 1。进一步分析图 7 和图 8 得知，MRCNN50-T 模型对东西冠幅、南北冠幅、树冠面积的拟合曲线斜率与 MRCNN101-T 模型相比，分别提升了 0.05、0.06、0.04，MRCNN50-T 模型对东西冠幅、南北冠幅、树冠面积的预测决定系数与 MRCNN101-T 模型相比，分别提升了 0.02、0.03、

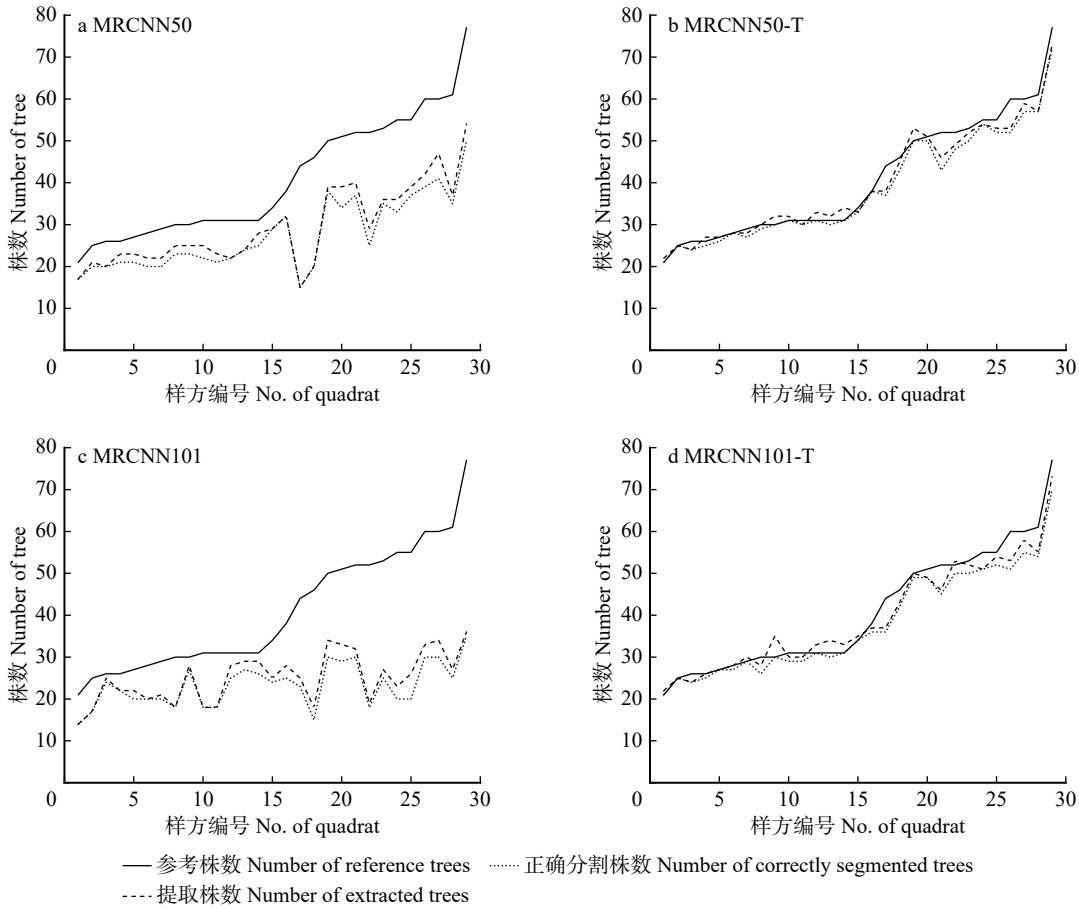


图 6 4 个深度学习模型的各株数比较

Fig. 6 Comparison of the number of trees in 4 deep learning models

0.05。试验结果表明 MRCNN50-T 在对东西冠幅、南北冠幅、树冠面积的预测性能上更有效，说明 ResNet50 特征提取网络对单木树冠个体的分割性能有着一定程度的提升作用。

2.5　样方郁闭度提取与精度评价

本试验中，树冠在林地的覆盖度作为郁闭度，将人工勾画树冠图和分割结果中的树冠覆盖度分别作为预测郁闭度和参考郁闭度，图 9 为两个模型预测郁闭度与参考郁闭度的散点图，郁闭度与树冠面积、株数密切相关，郁闭度的拟合程度可以有效代表模型的树冠提取性能。在郁闭度的提取上，MRCNN50-T 的预测决定系数和拟合曲线斜率分别为 0.83 和 0.85，MRCNN101-T 的预测决定系数和拟合曲线斜率分别为 0.72 和 0.59。试验结果表明：MRCNN50-T 的预测决定系数和拟合曲线斜率与 MRCNN101-T 的差值分别为 0.11 和 0.26，

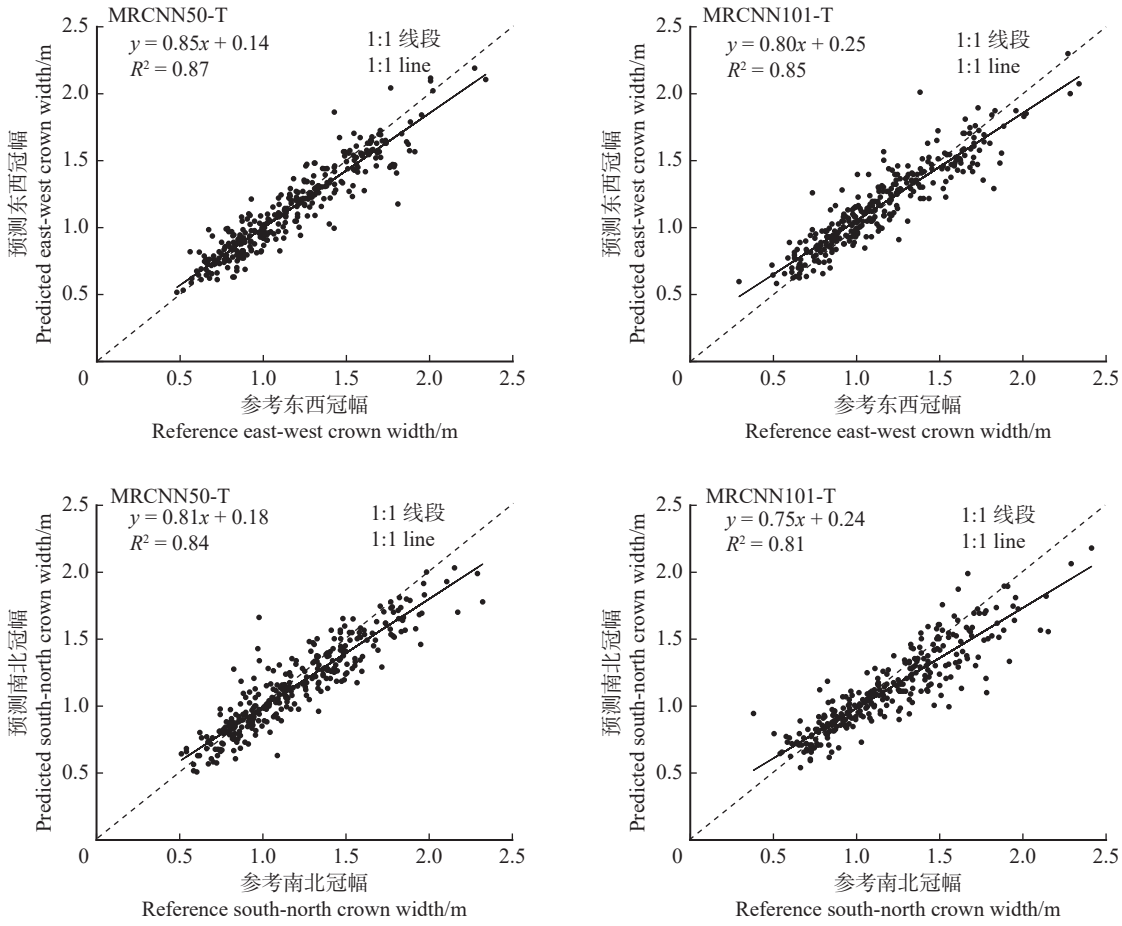


图 7　MRCNN50-T 和 MRCNN101-T 冠幅预测散点图

Fig. 7　Scatter plots of MRCNN50-T and MRCNN101-T crown width prediction

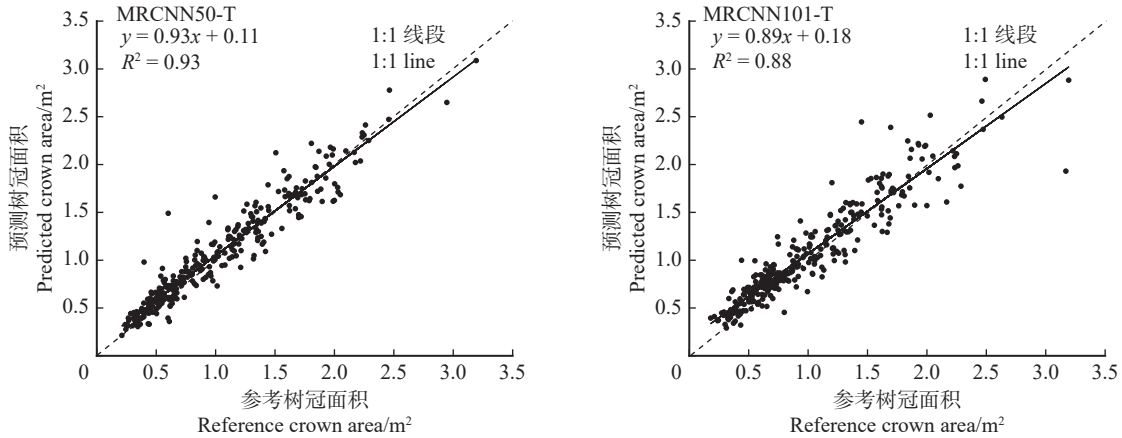


图 8　MRCNN50-T 和 MRCNN101-T 树冠面积预测散点图

Fig. 8　Scatter plots of MRCNN50-T and MRCNN101-T crown area prediction

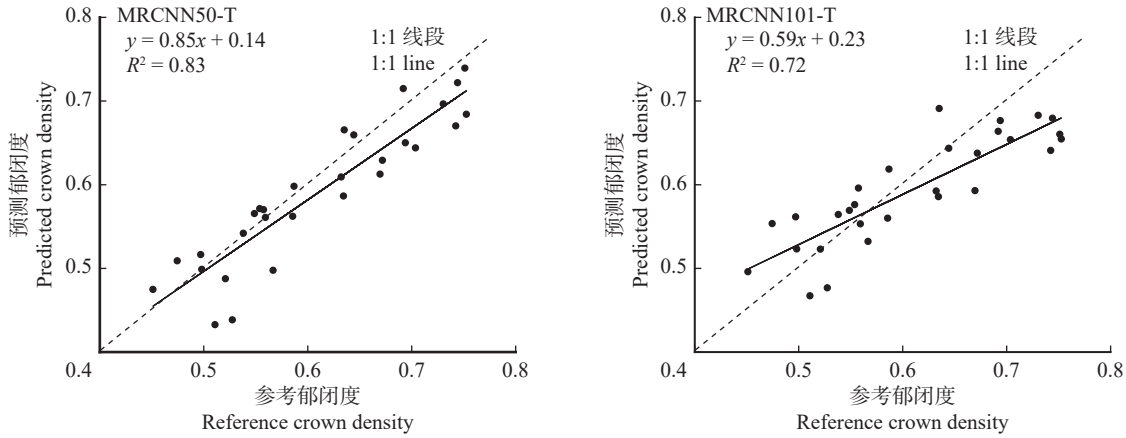


图 9 MRCNN50-T 和 MRCNN101-T 郁闭度预测散点图

Fig. 9 Scatter plots of MRCNN50-T and MRCNN101-T crown density prediction

MRCNN50-T 建立的样方郁闭度预测模型性能要优于 MRCNN101-T。

综合图 7～9 可知，在冠幅、树冠面积和郁闭度预测方面，MRCNN50-T 优于 MRCNN101-T，其拟合精度能满足生产和实践的需求。与 ISO 聚类和均值漂移分割的传统分割相比，深度学习实例分割模型的性能更为优越。主要的原因是传统分割模型不能有效识别树冠之间的重叠与边缘的环境差异。将迁移学习引入到实例分割中，进一步提升了分割精度。迁移学习将模型初始权重调整为比初始状态更优的状态，不仅从模型开始就提升了分割性能，而且节省了调整权重的训练时间。

3 讨 论

本研究试验了 Mask R-CNN 与迁移学习的无人机影像单木树冠提取，实验结果证明，不论是单木株数、单木树冠冠幅、单木树冠面积，还是样方郁闭度都能获得较高的提取准确度。Özcan 等[33]分别利用 sobel、prewitt、roberts、LoG、canny 算子和极大/小值的传统数字图像处理方法对果树的树冠边缘检测试验中，其中 sobel、prewitt、roberts、LoG、canny 各算子对树冠提取结果的 F1 分数分别为 74.93%、75.17%、70.17%、76.39%、80.51%，均低于本试验中的 MRCNN50-T 模型。对于小范围影像的单木树冠提取而言，传统数字图像处理占优，因为 Mask R-CNN 与迁移学习方法需要制作样本和训练模型，时长相比传统方法更长，但是对于大范围影像而言，Mask R-CNN 与迁移学习方法占优，因为该方法在制作样本和训练完模型后，可以实现自动化的大量提取，而且提取精度优于传统数字图像处理。

本研究利用各模型对无人机正射影像进行单木树冠提取，提取精度和模型的性能、算法的准确性、影像的质量等因素均有直接或间接的影响。研究发现不同的模型有着不同的提取精度，深度学习模型普遍比传统模型提取精度更优；影像的质量也至关重要，利用野外布设像控点、提高图像清晰度、选择合适天气拍摄、太阳高度角等因素提高影像质量。本研究只试验了同龄、生长状态良好的杉木纯林，目前国家森林资源丰富，对单木树冠提取的需求也远不止这一种林分，因此应对更复杂的天然林、异龄林等的单木树冠提取就需要更为高效、准确的算法和合适的影像。后续对复杂林分的研究中，因为认识到可见光波段信息量不足的短板，将利用数字表面模型与数字高程模型相减得到的冠层高度模型加入树冠提取模型中运算，补充可见光波段未包含的三维信息；并且将更多的纹理特征和植被指数加入深度学习特征提取模块中运算，更多的信息加入运算用于提高复杂林分树冠提取的普适性。

4 结 论

本研究采用无人机遥感技术与 Mask R-CNN 和迁移学习相结合的方式进行单木树冠参数的提取，得出以下结论。

（1）深度学习实例分割的分割性能优于 ISO 聚类和均值漂移分割，4 个深度学习模型的总体精度、用户精度、F1 分数的平均值分别为 95.08%、80.22%、85.87%。

（2）迁移学习的载入提高了 Mask R-CNN 的分割精度，其用户精度和 F1 分数分别提升 29.53% 和 19.53%，mAP 提升了 0.21。

（3）MRCNN50-T 在预测各树冠参数和郁闭度方面都优于 MRCNN101-T，其东西冠幅预测 R^2 为 0.87，南北冠幅预测 R^2 为 0.84，树冠面积预测 R^2 为 0.93，样方郁闭度预测 R^2 为 0.83。

本研究利用无人机平台获取的三波段彩色影像，建立集成 Mask R-CNN 深度学习和迁移学习的

模型自动化提取树冠参数，获得较高的预测精度。研究结果克服了传统分割方法精度低、目视解译工作量繁重等弊端，促进了深度学习在树冠提取研究的应用，为大范围林业调查、推动精准林业的发展提供借鉴经验。今后可将实例分割方法应用到混交林及不同地理环境的林分中，提高林分参数自动化提取精度，促进无人机林业遥感的进一步发展。

参 考 文 献

[1] 覃先林, 李增元, 易浩若. 高空间分辨率卫星遥感影像树冠信息提取方法研究 [J]. 遥感技术与应用, 2005(2): 228−232.
Qin X L, Li Z Y, Yi H R. Extraction method of tree crown using high-resolution satellite image[J]. Remote Sensing Technology and Application, 2005(2): 228−232.

[2] 孙振峰, 张晓丽, 李霓雯. 机载与星载高分遥感影像单木树冠分割方法和适宜性对比 [J]. 北京林业大学学报, 2019, 41(11): 66−75.
Sun Z F, Zhang X L, Li N W. Comparison of individual tree crown extraction method and suitability of airborne and spaceborne high-resolution remote sensing images[J]. Journal of Beijing Forestry University, 2019, 41(11): 66−75.

[3] Santos A A, Marcato J J, Araújo M S, et al. Assessment of CNN-based methods for individual tree detection on images captured by RGB cameras attached to UAVs[J]. Sensors, 2019, 19(16): 3595.

[4] 于东海, 冯仲科. 基于无人机倾斜航空影像的树冠体积测算方法 [J]. 农业工程学报, 2019, 35(1): 90−97.
Yu D H, Feng Z K. Tree crown volume measurement method based on oblique aerial images of UAV[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2019, 35(1): 90−97.

[5] 汪小钦, 王苗苗, 王绍强, 等. 基于可见光波段无人机遥感的植被信息提取 [J]. 农业工程学报, 2015, 31(5): 152−159.
Wang X Q, Wang M M, Wang S Q, et al. Extraction of vegetation information from visible unmanned aerial vehicle images[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2015, 31(5): 152−159.

[6] Huang H, Li X, Chen C. Individual tree crown detection and delineation from very-high-resolution UAV images based on bias field and marker-controlled watershed segmentation algorithms[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11(7): 2253−2262.

[7] 李丹, 张俊杰, 赵梦溪. 基于 FCM 和分水岭算法的无人机影像中林分因子提取 [J]. 林业科学, 2019, 55(5): 180−187.
Li D, Zhang J J, Zhao M X. Extraction of stand factors in UAV image based on FCM and watershed algorithm[J]. Scientia Silvae Sinicae, 2019, 55(5): 180−187.

[8] 张凝, 冯跃文, 张晓丽, 等. 结合航空影像纹理和光谱特征的单木冠幅提取 [J]. 北京林业大学学报, 2015, 37(3): 13−19.
Zhang N, Feng Y W, Zhang X L, et al. Extracting individual tree crown by combining spectral and texture features from aerial images[J]. Journal of Beijing Forestry University, 2015, 37(3): 13−19.

[9] 孙钊, 潘磊, 孙玉军. 基于无人机影像的高郁闭度杉木纯林树冠参数提取 [J]. 北京林业大学学报, 2020, 42(10): 20−26.
Sun Z, Pan L, Sun Y J. Extraction of tree crown parameters from high-density pure Chinese fir plantations based on UAV images[J]. Journal of Beijing Forestry University, 2020, 42(10): 20−26.

[10] 徐永胜. 基于无人机影像的林木冠幅提取与树种识别研究 [D]. 哈尔滨: 东北林业大学, 2021.
Xu Y S. Research on tree crown extraction and tree species recognition bases on UAV images[D]. Harbin: Northeast Forestry University, 2021.

[11] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation[C]//International Conference on Medical image computing and computer-assisted intervention. Munich: Technische Universität München, 2015.

[12] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Boston: Computer Vision, 2015.

[13] Li W J, Fu H h, Yu L, et al. Deep learning based oil palm tree detection and counting for high-resolution remote sensing images[J]. Remote Sensing, 2017, 9(1): 22.

[14] 田萱, 王亮, 丁琪. 基于深度学习的图像语义分割方法综述 [J]. 软件学报, 2019, 30(2): 440−468.
Tian X, Wang L, Ding Q. Review of image semantic segmentation based on deep learning[J]. Journal of Software, 2019, 30(2): 440−468.

[15] He K, Gkioxari G, Dollár P, et al. Mask R-CNN[C]//Proceedings of the IEEE international conference on computer vision. Venice: Computer Vision, 2017.

[16] Li Q, Arnab A, Torr P H S. Weakly-and semi-supervised panoptic segmentation[C]//Proceedings of the European conference on computer vision (ECCV). Munich: Computer Vision, 2018.

[17] 高云, 郭继亮, 黎煊, 等. 基于深度学习的群猪图像实例分割方法 [J]. 农业机械学报, 2019, 50(4): 179−187.
Gao Y, Guo J L, Li X, et al. Instance-level segmentation method for group pig images based on deep learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(4): 179−187.

[18] Dai J, He K, Sun J. Instance-aware semantic segmentation via multi-task network cascades[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Las Vegas: Computer Vision, 2016.

[19] Bai M, Urtasun R. Deep watershed transform for instance segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Hawaii: Computer Vision, 2017.

[20] Tan C, Sun F, Kong T, et al. A survey on deep transfer learning[C]//International conference on artificial neural networks. Springer: European Neural Network Society, 2018.

[21] Gu S, Ding L. A complex-valued vgg network based deep learning algorithm for image recognition[C]//2018 Ninth

International Conference on Intelligent Control and Information Processing (ICICIP). Wanzhou: International Conference on Intelligent Control and Information Processing (ICICIP), 2018.

[22] 李博, 江朝晖, 谢军, 等. 基于迁移学习的园艺作物叶部病害识别及应用 [J]. 中国农学通报, 2021, 37(7): 138−143.
Li B, Jiang C H, Xie J, et al. Leaf disease recognition of horticultural crops based on transfer learning and its application[J]. Chinese Agricultural Science Bulletin, 2021, 37(7): 138−143.

[23] 查志华, 周文静, 吴杰. 基于迁移学习 Faster R-CNN 模型田间红提葡萄果穗的识别 [J]. 石河子大学学报 (自然科学版), 2021, 39(1): 26−31.
Zha Z H, Zhou W J, Wu J. Identification of red globe grape cluster in grapery with Faster R-CNN model based on transfer learning[J]. Journal of Shihezi University (Natural Science Edition), 2021, 39(1): 26−31.

[24] Ocer N E, Kaplan G, Erdem F, et al. Tree extraction from multi-scale UAV images using Mask R-CNN with FPN[J]. Remote Sensing Letters, 2020, 11(9): 847−856.

[25] Safonova A, Guirado E, Maglinets Y, et al. Olive tree biovolume from uav multi-resolution image segmentation with mask r-cnn[J]. Sensors, 2021, 21(5): 1617.

[26] Hao Z, Lin L, Post C J, et al. Automated tree-crown and height detection in a young forest plantation using mask region-based convolutional neural network (Mask R-CNN)[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2021, 178: 112−123.

[27] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. Advances in Neural Information Processing Systems, 2015, 28: 91−99.

[28] Sun W, Wang R. Fully convolutional networks for semantic segmentation of very high resolution remotely sensed images combined with DSM[J]. IEEE Geoscience and Remote Sensing Letters, 2018, 15(3): 474−478.

[29] Sharma N, Jain V, Mishra A. An analysis of convolutional neural networks for image classification[J]. Procedia Computer Science, 2018, 132: 377−384.

[30] Torrey L, Shavlik J. Transfer learning[M]//Handbook of research on machine learning applications and trends: algorithms, methods, and techniques. Hershey: IGI Global, 2010.

[31] 梁泽, 黄姣, 韦飞黎, 等. 基于夜光遥感影像与百度 POI 数据的中国城市群空间范围识别方法 [J]. 地理研究, 2020, 39(1): 92−102.
Liang Z, Huang J, Wei F L, et al. Identifying the spatial range of urban agglomerations in China based on night light remote sensing and POI data[J]. Geographical Research, 2020, 39(1): 92−102.

[32] 伍艳莲, 赵力, 姜海燕, 等. 基于改进均值漂移算法的绿色作物图像分割方法 [J]. 农业工程学报, 2014, 30(24): 161−167.
Wu Y L, Zhao L, Jiang H Y, et al. Image segmentation method for green crops using improved mean shift[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2014, 30(24): 161−167.

[33] Özcan A H, Hisar D, Sayar Y, et al. Tree crown detection and delineation in satellite images using probabilistic voting[J]. Remote Sensing Letters, 2017, 8(8): 761−770.

（责任编辑 范 娟
责任编委 张会儒）