

# 杨梅树识别深度学习模型实验数据集

王东亮<sup>1\*</sup>, 罗巍<sup>2,3</sup>

1. 中国科学院地理科学与资源研究所地表格局与模拟院重点实验室, 北京 100101;

2. 北华航天工业学院, 廊坊 065000;

3. 河北省航天遥感信息处理与应用协同创新中心, 廊坊 065000

**摘要:** 杨梅树[*Myrica rubra* (Lour.) S. et Zucc]是一种常绿乔木, 树高 5–15 m, 胸径可达 60 cm, 树冠可达 5 m 以上。杨梅树广泛分布在我国江南地区, 生长于海拔 1,500 m 以下酸性红壤、山坡向阳的地理环境中。杨梅果是具有江南地理特色的水果。发展杨梅树无人机影像智能识别模型, 对提取杨梅树的位置和树冠信息, 进而实现无人机精准喷农药和估产有重要意义。作者于 2019 年 1 月 23–24 日在浙江省永嘉县大洋山森林公园采用大疆 Phantom4 无人机进行航拍, 并对杨梅树进行了多边形标记, 即采用 Mask RCNN (Region Convolutional Neural Networks) 深度学习模型对杨梅树进行自动识别, 对识别结果运用目视解译方法验证。结果表明, Mask RCNN 在杨梅树识别方面有较高精度, 总体检出率达 90.08% (其中错检率为 9.62%), 漏检率为 9.92%。杨梅树识别深度学习模型实验数据集包括: (1) 无人机拍摄的浙江省永嘉县大洋山森林公园影像 3,108 张, 影像尺寸为 5,472×3,648; (2) 杨梅树树冠样本标记数据 (284 张影像块); (3) 杨梅树深度学习模型识别结果数据 (14 张影像块)。该数据集以 .jpg、.JSON 格式存储, 有 3690 个数据文件组成, 数据量为 25.6 GB (压缩为 71 个文件, 25.5 GB)。

**关键词:** 杨梅树识别; 无人机影像; Mask RCNN 深度学习模型; 实验数据集

**DOI:** 10.3974/geodp.2019.03.10

## 1 前言

深度学习作为机器学习领域的一个分支, 可以从大数据中自动学习到一些人类难以手动提取到的特征, 因而能够取得以往传统浅层机器学习模型(如 ANN)难以达到的效果<sup>[1–2]</sup>, 被《Nature》<sup>[2–3]</sup>、《PNAS》<sup>[4]</sup>等期刊争相报道, 已成为遥感<sup>[1]</sup>、地球系统科学<sup>[2]</sup>等大数据处理领域一项重要的工具。Kellenberger *et al.*<sup>[5]</sup>使用卷积神经网络(CNNs, 一种深度学习模型), 从数千张 4 cm 分辨率无人机 RGB 影像中识别了 20 余种大型哺乳动物(实验中没有区分物种, 统一当做动物一类), 取得了较传统 EESVM 浅层机器学习<sup>[6]</sup>更高的精度(30% @ 80% 召回率 vs. 10% @ 75% 召回率)。Norouzzadeh *et al.*<sup>[7]</sup>组合使用 AlexNet、GoogLeNet、ResNet 等 9 种深度神经网络模型, 开展了基于地面红外触发相机图像中的动物的识别与分类, 并

收稿日期: 2019-07-24; 修订日期: 2019-09-16; 出版日期: 2019-09-25

基金项目: 中华人民共和国科学技术部 (2017YFC0506505, 2017YFB0503005); 中国科学院 (XDA23100200)

\*通讯作者: 王东亮 0000-0002-1377-8394, 中国科学院地理科学与资源研究所, wangdongliang@igsrr.ac.cn

数据引用方式: [1]王东亮, 罗巍. 杨梅树识别深度学习模型实验数据集[J]. 全球变化数据学报, 2019, 3(3): 290–296. DOI 10.3974/geodp.2019.03.10.

[2]王东亮, 罗巍. 杨梅树识别深度学习模型实验数据集[DB/OL]. 全球变化科学研究数据出版系统, 2019. DOI: 10.3974/geodb.2019.04.16.V1.

取得与志愿者相近的精度（识别图像中是否存在动物的精度达 96.6%）。Madec *et al.*<sup>[8]</sup>发展了一种基于 Faster RCNN 的麦穗识别算法,取得了较高的识别精度,相对均方根误差 rRMSE 为 5.3%,甚至优于首轮人机交互标记精度<sup>[9]</sup>。和传统机器学习方法相比,深度学习精度更高,移植性更强,但训练得到的模型是不可以解释的<sup>[1]</sup>,是一种黑匣解决方案,因此在遥感领域的使用存在一定争议,在实际应用中需要标记大量样本,并根据实际情况调整模型,以便取得更高精度和效率<sup>[3,5]</sup>。

随着 Faster RCNN<sup>[8]</sup>和 Mask RCNN<sup>[10]</sup>等模型的提出,深度学习模型的精度和速度都有了较大提升,在目标识别等领域已接近或超越了人类<sup>[7]</sup>。本研究基于大量高分辨率无人机航拍影像,利用 TensorFlow 平台,采用 Mask RCNN 深度学习算法对杨梅树进行了识别和标记。将测试结果与目视解译的结果进行比较,发现 Mask RCNN 能够较准确地识别杨梅树。

2 数据集元数据简介

杨梅树识别深度学习模型实验数据集<sup>[11]</sup>的名称、作者、地理区域、时间分辨率、空间分辨率、数据年代、数据集组成、数据出版与共享服务平台、数据共享政策等信息见表 1。

表 1 杨梅树识别深度学习模型实验数据集元数据简表

条 目	描 述
数据集名称	杨梅树识别深度学习模型实验数据集
数据集短名	BayberryTreeRecogData
作者信息	王东亮 0000-0002-1377-8394, 中国科学院地理科学与资源研究所, wangdongliang@igsrr.ac.cn 罗巍 0000-0003-2226-8414, 中国科学院微电子研究所, luowei2@ime.ac.cn
地理区域	浙江省永嘉县大洋山森林公园。具体范围: 28°17'N-28°19'N, 120°26'E-120°28'E
数据年代	2019 年 1 月 23-24 日, 每天 10:00-15:00
时间分辨率	仅覆盖 1 次                      空间分辨率    3 cm
数据格式	.jpg、.JSON
数据量	25.6 GB
数据集组成	无人机原始影像、杨梅树标记数据、杨梅树识别深度学习模型实验结果数据。
基金项目	中华人民共和国科学技术部 ( 2017YFC0506505, 2017YFB0503005 ); 中国科学院 ( XDA23100200 );
出版与共享服务平台	全球变化科学研究数据出版系统 <a href="http://www.geodoi.ac.cn">http://www.geodoi.ac.cn</a>
地址	北京市朝阳区大屯路甲 11 号 100101, 中国科学院地理科学与资源研究所
数据共享政策	全球变化科学研究数据出版系统的“数据”包括元数据 ( 中英文 )、实体数据 ( 中英文 ) 和通过《全球变化数据学报》( 中英文 ) 发表的数据论文。其共享政策如下: ( 1 ) “数据”以最便利的方式通过互联网系统免费向全社会开放, 用户免费浏览、免费下载; ( 2 ) 最终用户使用“数据”需要按照引用格式在参考文献或适当的位置标注数据来源; ( 3 ) 增值服务用户或以任何形式散发和传播 ( 包括通过计算机服务器 ) “数据”的用户需要与《全球变化数据学报》( 中英文 ) 编辑部签署书面协议, 获得许可; ( 4 ) 摘取“数据”中的部分记录创作新数据的作者需要遵循 10% 引用原则, 即从本数据集中摘取的数据记录少于新数据集总记录量的 10%, 同时需要对摘取的数据记录标注数据来源 <sup>[12]</sup>
数据和论文检索系统	DOI, DCI, CSCD, WDS/ISC, GEOSS, China GEOSS

3 数据研发方法

3.1 研究区概况

大洋山森林公园位于浙江省永嘉县，距离温州市区 48.8 km，规划面积 661.93 hm<sup>2</sup>。海拔 330–893.1 m，平均坡度 30°左右。年平均气温约 17℃，无霜期约 270 天。年均降雨量 1,500–1,900 mm。公园内有针叶林、杨梅园、铁皮石斛种植基地，森林覆盖率约 87.4%。

3.2 算法原理

Mask RCNN 是一种概念简单和框架灵活的目标检测分割算法，不仅能实现目标的准确检测，而且能够对每个对象实例产生一个高质量的分割掩膜（Segmentation Mask）。Mask RCNN 是在对原有的 Faster RCNN 进行扩展的基础上，融合了 FCN（Fully Convolutional Networks）和 FPN（Feature Pyramid Networks）思想，并添加一个分支在对一张图片所有的目标进行正确检测的同时还要对每个示例的掩膜进行预测。Mask RCNN 分三个阶段：第一阶段，由 Mask RCNN 主干网络（即 ResNet 和 FPN）提取出无人机图像的特征图；第二阶段，使用区域建议网络（Region Proposal Network, RPN）生成目标的建议框，并对建议框筛选得到感兴趣区域（Regions of Interest, ROIs）；第三阶段，对每个 ROIs 的预测类别、位置和对应的对象进行掩膜，生成目标对象掩膜多边形。

3.3 技术路线

对测区进行分块和飞行路线的设计，获取无人机影像；然后，使用 Mask RCNN 进行杨梅树训练、识别与验证。

第一步，将无人机影像进行分块，并对杨梅树标记。由于 Mask RCNN 对计算机显存或内存有较高要求，本研究使用的电脑显存为 6 GB，输入的影像尺寸不能太大，因此将原始影像分割为每块 1,024×682 像素或更小尺寸。在此基础上，对无人机影像杨梅树进行标记，并将样本分为训练样本（284 张影像块）和验证样本（14 张影像块）。

第二步，模型训练。基于 TensorFlow 搭建了 Mask RCNN 杨梅树识别模型，利用已标记的训练样本，对模型进行迭代训练，获得杨梅树识别深度学习模型。

第三步，模型验证。使用训练好的模型，对每张验证样本影像进行识别，获得杨梅树掩膜多边形，并与目视解译结果进行对比，计算模型精度。

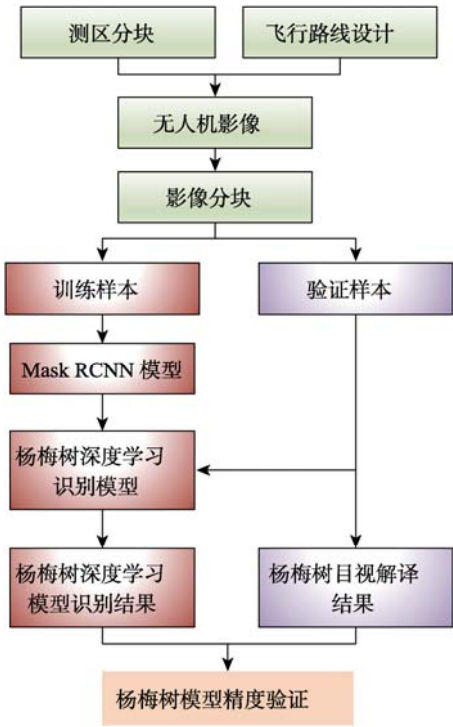


图 1 数据集研发技术路线图

4 数据结果与验证

4.1 数据集组成

数据集包括：（1）浙江省永嘉县无人机影像（表 2）；（2）无人机杨梅树影像块、样本

标记多边形及杨梅树识别结果（表 3）。

4.2 数据结果

测区划分为 4 个，飞行路线见图 2，获取影像 3,108 张，影像分辨率约 3 cm，尺寸为 5,472×3,648。训练样本包含 284 张影像块（从 18 张影像裁剪获得），在第 1、2、4 测区均有分布。用于识别结果验证的样本则包含 14 张影像块（从 2 张影像裁剪获得），只分布于第 4 测区，位置见图 2。杨梅树多边形采用 Labelme 软件标记完成，见图 3。识别结果见图 4。

4.3 数据结果验证

采用目视解译方法，对杨梅树深度学习模型的识别精度进行验证。验证样本包括从第 4 测区抽取的 14 张影像块。

表 2 浙江省永嘉县无人机影像

条目	描述
20190123 大洋山 01 测区	01 测区影像
20190123 大洋山 02 测区	02 测区影像
20190123 大洋山 03 测区	03 测区影像
20190124 大洋山 04 测区	04 测区影像
20190124 大洋山 04 测区补拍	04 测区补拍影像
20190124 大洋山 04 测区继续	04 测区继续影像

表 3 无人机杨梅树影像块、样本标记多边形及杨梅树识别结果

条目	描述
杨梅树训练影像块和样本标记多边形	包含 284 张影像块
杨梅树测试影像块和识别结果	包含 14 张待识别影像块、14 张影像杨梅树识别结果

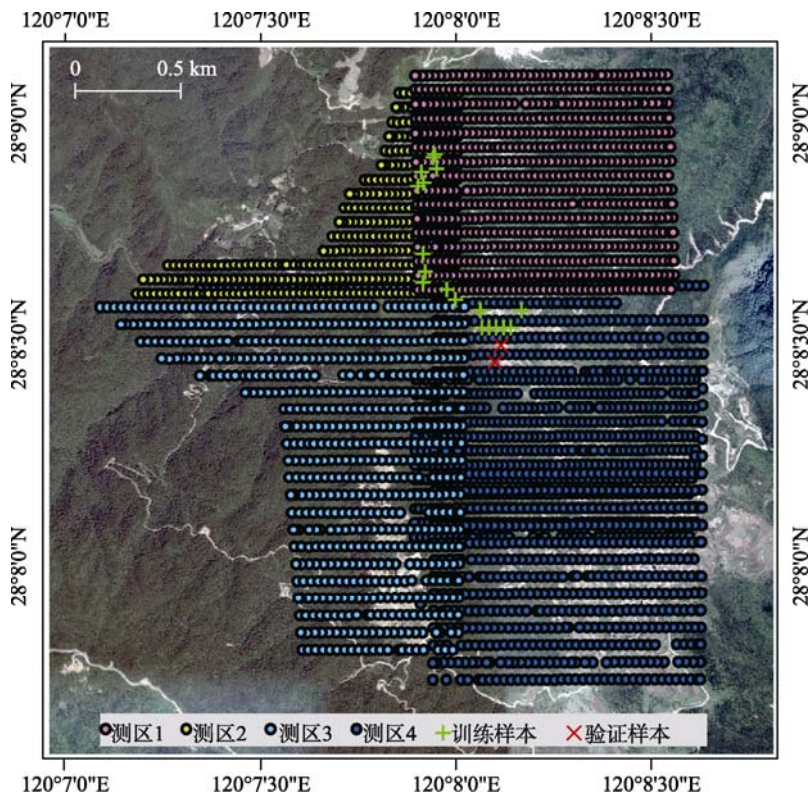


图 2 测区划分与样本分布



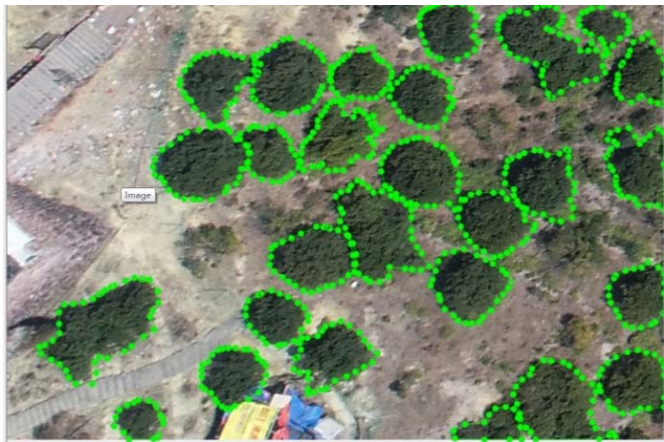


图 3 基于 Labelme 标记的杨梅树多边形

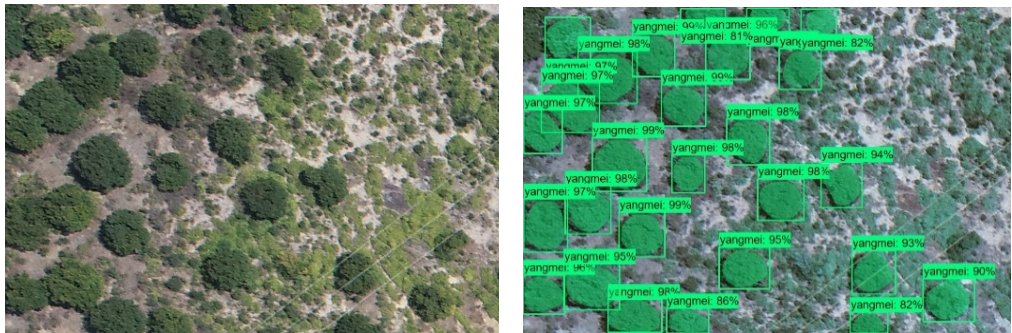


图 4 杨梅树影像（左）与深度学习模型智能识别结果（右）

检出率（True Positive Rate,  $TPR$ ）为被模型检测的正样本（ $N_{positive}$ ）数目占目视解译目标总数（ $N_{total}$ ）的比值：

$$TPR = \frac{N_{positive}}{N_{total}} \quad (1)$$

错检率（False Positive Rate,  $FPR$ ）表示一个样本集中，被分类器错误判定为目标的数目（ $N_{false}$ ）占有所有被模型判定为目标的目标的数目（ $N_{positive}$ ）的比值：

$$FPR = \frac{N_{false}}{N_{positive}} \quad (2)$$

漏检率（Loss Positive Rate,  $LPR$ ）表示一个样本集中，未被模型检测到的目标数量（ $N_{loss}$ ）占目视解译目标总数（ $N_{total}$ ）的比值：

$$LPR = \frac{N_{loss}}{N_{total}} \quad (3)$$

基于 MASK RCNN 的杨梅树识别结果及模型精度如表 4 所示。

表 4 杨梅树深度学习模型识别结果及精度

序号	影像编号	杨梅树 个数	正确识别 个数	错误识别 个数	漏识别 个数	检出率 (%)	错检率 (%)	漏检率 (%)
1	2019012404DJI_c_0546-1	25	25	1	0	100.00	3.85	0.00
2	2019012404DJI_c_0546-2	29	27	0	2	93.10	0.00	6.90
3	2019012404DJI_c_0546-3	0	0	10	0	-	100.00	-
4	2019012404DJI_c_0546-6	39	34	0	5	87.18	0.00	12.82
5	2019012404DJI_c_0546-7	43	41	0	2	95.35	0.00	4.65
6	2019012404DJI_c_0546-8	10	7	0	3	70.00	0.00	30.00
7	2019012404DJI_c_0546-9	46	43	0	3	93.48	0.00	6.52
8	2019012404DJI_c_0546-10	34	34	0	0	100.00	0.00	0.00
9	2019012404DJI_c_0526-13	33	33	0	0	100.00	0.00	0.00
10	2019012404DJI_c_0526-18	22	19	7	3	86.36	26.92	13.64
11	2019012404DJI_c_0526-19	46	39	0	7	84.78	0.00	15.22
12	2019012404DJI_c_0526-20	34	29	0	5	85.29	0.00	14.71
13	2019012404DJI_c_0526-21	29	25	0	4	86.21	0.00	13.79
14	2019012404DJI_c_0526-22	28	25	1	3	89.29	3.85	10.71
合计/平均		418	381	19	37	90.08	9.62	9.92

5 讨论和总结

研究利用 284 张训练影像块，训练生成了基于 MASK RCNN 构建的深度学习网络，在 14 张测试影像中，总体检测精度较高，其中目视解译发现杨梅树 418 棵，深度学习模型检出 381 棵，错检 19 棵，漏检 37 棵。总体检出率达 90.08%，错检率为 9.62%，漏检率为 9.92%（表 4）。

14 张影像中，存在错检现象的影像有 4 张（第 1、3、10、14 张），其中第 3 张和第 10 张影像分别错检 10 棵和 7 棵杨梅树。主要原因为：第 3 张影像中存在大面积水域，以及与杨梅树相近的松杉树，却不存在杨梅树，导致模型错误地将松杉树识别成杨梅树；第 10 张影像中同时存在杨梅树和松杉树，但松杉树与杨梅树的颜色和纹理等都较为接近，难以区分。存在漏识别现象的影像有 10 张（第 2、4、5、6、7、10、11、12、13、14 张），每张影像漏检杨梅树 2-7 棵，主要原因是杨梅树的色彩、形状、纹理与正常值存在较大差异。如由于人工修剪等原因，杨梅树通常呈现为圆形，但有的杨梅树没有修剪，形状不规则，与松杉接近；部分杨梅树叶发黄，与正常绿色差异较大。此外，由于本研究设定只输出得分大于 0.8 的杨梅树（得分小于 0.8 的不被识别为杨梅树），部分杨梅树虽然被识别出来了，但得分小于 0.8，在显示时被自动剔除，造成漏检，需根据真实杨梅树识别的得分情况对输出分数阈值进行调整，降低因阈值设置不合理造成的漏检。

考虑到本文仅使用基本的 MASK RCNN 深度学习框架进行杨梅树识别，错检率和漏检率均近 10%，精度有待提高。今后将针对杨梅树的树冠像素个数等特点，并采用负样本学

习等方法,对模型进行改进,降低模型的错检率和漏检率,提升模型的精度和可靠性。此外,今后还将利用模型勾勒的树冠掩膜多边形,计算其树冠大小,并结合部分地面测得的杨梅树产量,开展杨梅树的估产研究。

**作者分工:**王东亮等人在大洋山森林公园工作人员协助下,采集和处理了无人机影像;王东亮、罗巍对数据集的开发做了总体设计,标记样本数据集,对模型进行设计、测试和验证,撰写了数据论文。

**致谢:**本次无人机航拍得到浙江生命林科技有限公司和大洋山森林公园工作人员的大力支持,在此表示衷心的感谢!

## 参考文献

- [1] Zhu, X. X., Tuia, D., Mou, L., *et al.* Deep learning in remote sensing: a comprehensive review and list of resources [J]. *IEEE Geoscience & Remote Sensing Magazine*, 2018, 5(4): 8–36.
- [2] Reichstein, M., Camps-Valls, G., Stevens B., *et al.* Deep learning and process understanding for data-driven Earth system science [J]. *Nature*, 2019, 566(7743): 195–204.
- [3] Lecun, Y., Bengio, Y., Hinton, G. Deep learning [J]. *Nature*, 2015(521): 436.
- [4] Waldrop, M. M. News Feature: what are the limits of deep learning? [J]. *Proceedings of the National Academy of Sciences*, 2019, 116(4): 1074–1077.
- [5] Kellenberger, B., Marcos, D., Tuia, D. Detecting mammals in uav images: best practices to address a substantially imbalanced dataset with deep learning [J]. *Remote Sensing of Environment*, 2018(216): 139–153.
- [6] Rey, N., Volpi, M., Joost, S., *et al.* Detecting animals in African Savanna with UAVs and the crowds [J]. *Remote Sensing of Environment*, 2017(200): 341–351.
- [7] Norouzzadeh, M. S., Nguyen, A., Kosmala, M., *et al.* Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning [J]. *Proceedings of the National Academy of Sciences*, 2018, 115(25): E5716–E5725.
- [8] Ren, S., He, K., Girshick, R., *et al.* Faster R-CNN: towards Real-Time Object Detection with Region Proposal Networks [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017, 39(6): 1137–1149.
- [9] Madec, S., Jin, X., Lu, H., *et al.* Ear density estimation from high resolution RGB imagery using deep learning technique [J]. *Agricultural and Forest Meteorology*, 2019(264): 225–234.
- [10] He, K., Gkioxari, G., Dollár, P., *et al.* Mask R-CNN [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017, 1(99): 2961–2969.
- [11] 王东亮, 罗巍. 杨梅树识别深度学习模型实验数据集[DB/OL]. 全球变化科学研究数据出版系统, 2019. DOI: 10.3974/geodb.2019.04.16.V1.
- [12] 全球变化科学研究数据出版系统. 全球变化科学研究数据共享政策 [OL]. DOI: 10.3974/dp.policy.2014.05 (2017 年更新) .