

# Bayberry Tree Recognition Dataset Based on the Aerial Photos and Deep Learning Model

Wang, D. L.<sup>1\*</sup> Luo, W.<sup>2,3</sup>

1. Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China;
2. North China Institute of Aerospace Engineering, Langfang 065000, China;
3. Center for Intelligent Manufacturing Electronics, Institute of Microelectronics, Chinese Academy of sciences, Beijing 100029, China

**Abstract:** Bayberry (*Myrica rubra* (Lour.) S. et Zucc) is an evergreen tree with 5-15 m high; the diameter of the tree is about 60 cm at breast height and a crown of more than 5 m. Bayberry tree is widely distributed in the southern of the Yangtze River basin of China. Bayberry tree grows in acid red soil below 1500 m above sea level and sunny mountain slopes. Its fruit is a special fruit with geographical characteristics in the southern of the Yangtze River basin. The development of intelligent model for recognizing bayberry trees based on unmanned aerial systems (UAS) images is of great significance to extract the position and crown information of bayberry trees, which can then be used for precise pesticide spraying and yield estimation. The authors used DJI Phantom 4 unmanned aerial systems (UAS) to take the aerial photography in Dayangshan Forest Park, Yongjia county, Zhejiang province from January 23 to 24, 2019. Polygonal markers of Bayberry trees were then labeled before training a deep learning model, Mask RCNN (Mask Region Convolutional Neural Networks), to automatic identify the bayberry trees. The recognition results were evaluated by visual interpretation method. The experimental results show that Mask RCNN has a high accuracy for bayberry tree recognition. The overall true positive rate is 90.08%, the false positive rate is 9.62%, and the loss positive rate is 9.92%. The experimental dataset of the presented bayberry tree recognition deep learning model includes: (1) 3,080 UAS images captured in Dayangshan Forest Park, Yongjia county, Zhejiang province, with an image size of  $5,472 \times 3,648$ ; (2) Marked bayberry tree training samples, including 284 image tiles; and (3) bayberry tree recognition results by the deep learning model, including 14 image tiles. The dataset is archived in .jpg and .JSON data formats, consists of 3,690 data files with data size of 25.6 GB (compressed to 71 files with data size of 25.5 GB).

**Keywords:** bayberry tree; recognition; UAS images; mask RCNN

## 1 Introduction

As a branch of machine learning, deep learning can automatically learn some features which

---

**Received:** 24-07-2019; **Accepted:** 17-09-2019; **Published:** 25-09-2019

**Foundations:** Ministry of Science and Technology of P. R. China (2017YFC0506505, 2017YFB0503005); Chinese Academy of Sciences (XDA23100200)

**\*Corresponding Author:** Wang, D. L. 0000-0002-1377-8394, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, wangdongliang@igsnnr.ac.cn

**Data Citation:** [1] Wang, D. L., Luo, W. Bayberry tree recognition dataset based on the aerial photos and deep learning model [J]. *Journal of Global Change Data & Discovery*, 2019, 3(3): 290–296. DOI: 10.3974/geodp.2019.03.10.

[2] Wang, D. L., Luo, W. Bayberry tree recognition dataset based on the aerial photos and deep learning model [DB/OL]. Global Change Research Data Publishing & Repository, 2019. DOI: 10.3974/geodb.2019.04.16.V1.

are difficult to extract manually from big data, and extract more useful feature representations with a higher accuracy than those of traditional machine learning algorithms (such as ANN)<sup>[1–2]</sup>. It has been reported by journals such as Nature<sup>[2–3]</sup>, PNAS<sup>[4]</sup>, etc. Deep learning has become an important tool in data processing fields such as remote sensing<sup>[1]</sup>, and earth system science<sup>[2]</sup>. Kellenberger *et al.*<sup>[5]</sup> trained a convolutional neural network (CNN, a recent family of deep learning algorithms) to recognize more than 20 types of large wildlife based on thousands of UAS images with a resolution of about 4cm (there was no distinction between species in the experiment), and achieved a higher accuracy than traditional EESVM shallow machine learning<sup>[6]</sup> (30%@80% recall vs. 10%@75% recall rate). Norouzzadeh *et al.*<sup>[7]</sup> combined nine deep neural network models, such as AlexNet, Google LeNet and Res-Net, to carry out the recognition and classification of animals in ground-based camera-trap images, and the accuracy is similar to that of volunteers (the accuracy of recognition is 96.6%). Madec *et al.*<sup>[8]</sup> developed a wheat ear recognition algorithm based on Faster RCNN, and the relative root mean square error (rRMSE) is 5.3%, which is even better than the first-round human-computer interaction marking accuracy<sup>[9]</sup>. Deep learning typically has a higher accuracy and better probability than traditional machine learning algorithms. However, it is a black-box solution, and extremely large training datasets and appropriate model adjustment are needed to improve the accuracy and efficiency when applying deep learning. These issues are controversial within the remote sensing community<sup>[3,5]</sup>.

With the emergence of models such as Faster RCNN<sup>[8]</sup> and Mask RCNN<sup>[10]</sup>, the precision and speed of deep learning models are close to or beyond human beings in object recognition and other fields<sup>[7]</sup>. This study presented a Mask RCNN-based model to identify and mark bayberry trees from high resolution UAS images based on TensorFlow platform. The recognized results were inspected by a visual interpretation method and showed that Mask RCNN can be used for identify bayberry tree with a high accuracy.

## 2 Metadata of Dataset

The metadata of “Bayberry tree recognition dataset based on the aerial photos and deep learning model”<sup>[11]</sup> is summarized in Table 1. It includes the dataset full name, short name, authors, geographical region, year of the dataset, temporal resolution, spatial resolution, data format, data size, data files, data publisher, and data sharing policy, etc.

**Table 1** Metadata summary of “Bayberry tree recognition dataset based on the aerial photos and deep learning model”

Items	Description
Dataset full name	Bayberry tree recognition dataset based on the aerial photos and deep learning model
Dataset short name	BayberryTreeRecogData
Authors	Wang, D. L. 0000-0002-1377-8394, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, wangdongliang@igsrr.ac.cn Luo, W. 0000-0003-2226-8414, North China Institute of Aerospace Engineering; Center for Intelligent Manufacturing Electronics, Institute of Microelectronics, Chinese Academy of sciences, luowei2@ime.ac.cn
Geographical region	Dayangshan Forest Park, Yongjia county, Zhejiang province. 28°17'N–28°19'N, 120°26'E–120°28'E
Year	January 23–24, 2019
Spatial resolution	3 cm
Data files	UAS images, marked bayberry tree training images, test images, and bayberry tree recognition results
Foundations	Ministry of Science and Technology of P. R. China (2017YFC0506505, 2017YFB0503005); Chinese Academy of Sciences (XDA23100200)
Data publisher	Global Change Research Data Publishing & Repository, <a href="http://www.geodoi.ac.cn">http://www.geodoi.ac.cn</a>
Address	No. 11A, Datun Road, Chaoyang District, Beijing 100101, China

(To be continued on the next page)

(Continued)

Items	Description
Data sharing policy	<b>Data</b> from the Global Change Research Data Publishing & Repository includes metadata, datasets (data products), and publications (in this case, in the <i>Journal of Global Change Data &amp; Discovery</i> ). <b>Data</b> sharing policy includes: (1) <b>Data</b> are openly available and can be free downloaded via the Internet; (2) End users are encouraged to use <b>Data</b> subject to citation; (3) Users, who are by definition also value-added service providers, are welcome to redistribute <b>Data</b> subject to written permission from the GCdataPR Editorial Office and the issuance of a <b>Data</b> redistribution license, and; (4) If <b>Data</b> are used to compile new datasets, the ‘ten percent principal’ should be followed such that <b>Data</b> records utilized should not surpass 10% of the new dataset contents, while sources should be clearly noted in suitable places in the new dataset <sup>[7]</sup>
Communication and searchable system	DOI, DCI, CSCD, WDS/ISC, GEOSS, China GEOSS

3 Methods

3.1 Study Area

Dayangshan Forest Park is located in Yongjia County, Zhejiang province, 48.8 km away from Wenzhou City. The park has an area of 661.93 hm<sup>2</sup>. Across the park, elevation varies from 330 m to 893.1 m. The overall terrain slope across the park is about 30°. Annual mean temperature is about 17 °C and the frost-free period is about 270 days. Annual average precipitation ranges from 1,500–1,900 mm. The park is characterized as coniferous forests, bayberry gardens and Dendrobium candidum planting bases. Forests cover 87.4% of the park’s land surface.

3.2 Principle of Mask RCNN

Mask RCNN is a conceptually simple, flexible, and general framework for object instance segmentation, which can efficiently detects objects in an image while simultaneously generating a high-quality segmentation mask for each instance<sup>[10]</sup>. The method is an extension of Faster RCNN that applies fully convolutional network (FCN) and Feature Pyramid Networks (FPN), and adds a branch for predicting an object mask in parallel with the existing branch for bounding box recognition. Mask RCNN works as follows. First, a ResNet-FPN backbone is used for extracting features from UAS images. Second, a region proposal network (RPN) is used to propose candidate object bounding boxes. Region of interests (RoI) are then determined from the candidate object bounding boxes. The target category, location, and a segmentation mask are predicted for each ROIs.

3.3 Technical Flow Chart

The survey area is divided into four blocks and flight routes are designed to obtain UAS images. Training and validation of the bayberry tree recognition model are then carried out using Mask RCNN. The three steps for training and evaluate of the bayberry tree recognition model are as follows (Figure 1).

Step 1: The UAS images are split into several sub-images and bayberry trees are labeled. It was not possible to train the model with the original 5,472×3,648 pixels images because of GPU memory limitation. The original images are therefore split into multiple sub-images (tiles) to ensure that each sub-image has a size of no more than 1,024×682 pixels. The bayberry trees are then labeled to construct bayberry tree samples. The samples are then divided

into training samples (284 sub-images) and test samples (14 sub-images).

Step 2: The bayberry tree recognition model is built. Mask RCNN bayberry tree recognition model is built based on Google TensorFlow. The bayberry tree deep learning recognition model is obtained by iteratively training the model based on training samples.

Step 3: The model is evaluated. The trained model is used to recognize and mask bayberry trees for each image to obtain their segmentation masks in test samples. The accuracy of the model is calculated by comparing recognized results with visual interpretation results.

4 Results and Validation

4.1 Data Composition

The experimental dataset includes: (1) UAS images captured in Yongjia County, Zhejiang Province (Table 2); (2) Bayberry tree training sub-images, tag files, and recognition results (Table 3).

4.2 Data Products

The survey area is divided into four regions. Figure 2 shows flight lines. The campaign captured 3,108 images with a resolution of approximately 3 cm and a size of 5,472×3,648. The training images consisting of 284 sub-images (split from 18 original images) come from survey region 1, 2, and 4. The test images, consisting of 14 sub-images (split from 2 original images), are picked from survey region 4 and are used for evaluate recognition results, as shown in Figure 2. The bayberry polygons are labeled using Labelme (Figure 3). Recognition results are shown in Figure 4.

Table 2 UAS images captured in Yongjia county, Zhejiang province

Items	Description
20190123DayangshanRegion01	Images of Region01
20190123 DayangshanRegion02	Images of Region02
20190123 DayangshanRegion03	Images of Region03
20190124 DayangshanRegion04	Images of Region04
20190124DayangshanRegion04continue01	Additional Images of Region04
20190124DayangshanRegion04continue02	Additional Images of Region04

Table 3 UAS bayberry tree training sub-images, bayberry tree tag files, and bayberry tree recognition results

Items	Description
Bayberry tree training sub-images, labeled bayberry tree polygons	284 training sub-images and tag files
Bayberry tree test sub-images and recognition results	14 test sub-images and 14 bayberry tree recognition resulting images

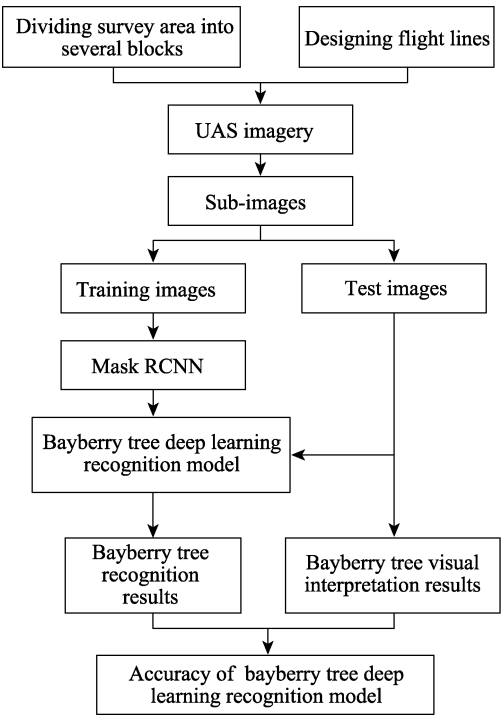
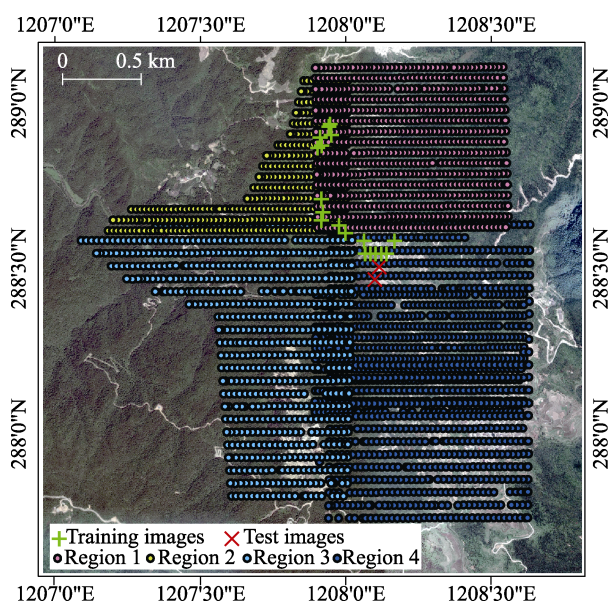
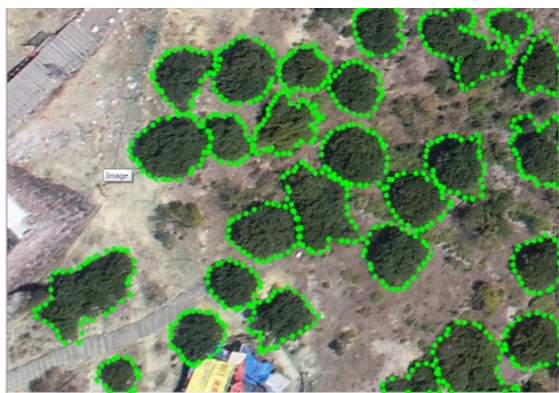


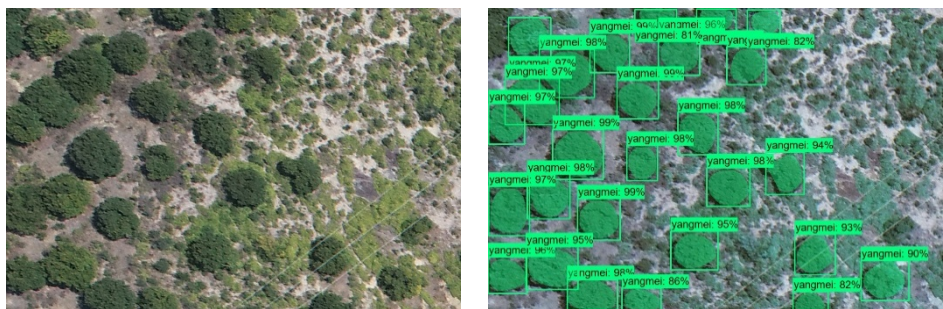
Figure 1 Technical flow chart showing generation of the dataset



**Figure 2** Unmanned aerial systems flight lines, and distribution of training and test images



**Figure 3** Example of bayberry tree polygons labeled using Labelme



**Figure 4** Bayberry tree image (left) and the recognition results (right)

### 4.3 Data Validation

The visual interpretation method is used to evaluate the accuracy of bayberry tree deep learning recognition model. The test samples includes 14 sub-images extracted from survey region 4.

True positive rate (TPR) is the ratio of the number of bayberry trees detected out by the model to the number of visual interpretation bayberry trees ( $N_{total}$ ).

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

False positive rate (FPR) is the number of targets ( $N_{false}$ ) that are not bayberry trees but are identified as bayberry trees by the model, divided by the total number of targets ( $N_{positive}$ ) that are identified as bayberry trees by the model:

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

Loss Positive Rate (LPR) is the ratio of the number of bayberry trees are not identified as bayberry trees by the model ( $N_{loss}$ ) to the number of visual interpretation bayberry trees.

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

The recognition results and accuracy of bayberry tree recognition model based on MASK RCNN are shown in Table 4.

**Table 4** Results and accuracy of bayberry tree recognition model

Images	Number of bayberry trees	Number of true positives	Number of false positives	Number of loss positive rate	True positive rate (%)	False positive rate (%)	Loss positive Rate (%)
2019012404DJI_c_0546-1	25	25	1	0	100.00	3.85	0.00
2019012404DJI_c_0546-2	29	27	0	2	93.10	0.00	6.90
2019012404DJI_c_0546-3	0	0	10	0	—	100.00	—
2019012404DJI_c_0546-6	39	34	0	5	87.18	0.00	12.82
2019012404DJI_c_0546-7	43	41	0	2	95.35	0.00	4.65
2019012404DJI_c_0546-8	10	7	0	3	70.00	0.00	30.00
2019012404DJI_c_0546-9	46	43	0	3	93.48	0.00	6.52
2019012404DJI_c_0546-10	34	34	0	0	100.00	0.00	0.00
2019012404DJI_c_0526-13	33	33	0	0	100.00	0.00	0.00
2019012404DJI_c_0526-18	22	19	7	3	86.36	26.92	13.64
2019012404DJI_c_0526-19	46	39	0	7	84.78	0.00	15.22
2019012404DJI_c_0526-20	34	29	0	5	85.29	0.00	14.71
2019012404DJI_c_0526-21	29	25	0	4	86.21	0.00	13.79
2019012404DJI_c_0526-22	28	25	1	3	89.29	3.85	10.71
Sum/average	418	381	19	37	90.08	9.62	9.92

## 5 Discussion and Conclusion

A Mask-RCNN-based bayberry tree recognition model is built based on 284 training sub-images. The recognition accuracy is relative high. In 14 test sub-images, 418 bayberry trees are found by the visual interpretation method, and 381 bayberry trees are detected out by the deep learning recognition model. The overall true positive rate is 90.08%, false posi-

tive rate is 9.62%, and loss positive rate is 9.92% (Table 4).

Where as, there is still considerable room to improve the accuracy of recognition. This paper only uses the basic MASK RCNN deep learning framework for bayberry tree recognition. The false positive rate and loss positive rate are both approximately 10%. The use of a negative sample learning method in combination with revising the model according to characteristics of bayberry trees will be addressed in our future work to improve the accuracy and reliability of recognition. The yields of bayberry trees will also be estimated according the relationship between ground-measured yields and the crown sizes of bayberry trees calculated from model-generated crown masks.

### Author Contributions

Wang, D. L. contributed to the UAS data collection and processing. Wang, D. L. and Luo, W. made a general design for the development of datasets, labeled the samples, designed, trained, and evaluated the model, and wrote the data paper.

### Acknowledgements

The authors would like to thank Zhejiang Senminglin Technology Co., Ltd. and the staff of Yangshan Forest Park for their support in this UAS campaign.

### References

- [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: 436–444.
- [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] Wang, D. L., Luo, W. Bayberry tree recognition dataset based on the aerial photos and deep learning model [DB/OL]. Global Change Research Data Publishing & Repository, 2019. DOI: 10.3974/geodb.2019.04.16.V1.
- [12] GCdataPR Editorial Office. GCdataPR data sharing policy [OL]. DOI: 10.3974/dp.policy.2014.05 (Updated 2017).