International Journal of New Developments in Engineering and Society
ISSN 2522-3488 Vol. 8, Issue 2: 110-114, DOI: 10.25236/IJNDES.2024.080217

Method of Dead Standing Tree Detection Based on RetinaNet Object Detection Network

Lulu Liu, Leifang Xie, Peng Zhang, Yong Wang, Han Tian

The 27th Research Institute of China Electronics Technology Group Corporation, Zhengzhou, 450047, China

Abstract: *Aiming at the recognition of dead standing trees in UAV RGB images in forest areas, this paper proposes a dead standing tree detection method based on RetinaNet target detection network. Using UAV images for RetinaNet target detection network model training, and this paper compares multiple feature extraction networks, the results show that ResNet-152-FPN is the Best, using ResNet-152-FPN-based RetinaNet for target detection on dead standing trees in forest areas, the average accuracy of dead standing tree recognition reaches 81.6%. It proves the feasibility of RetinaNet target detection network on dead standing wood recognition, and provides strong support for dead standing wood census recognition in forest areas.*

Keywords: *UAV RGB images, RetinaNet, target detection, dead standing trees*

1. Introduction

Dead standing trees, referred to as snags, are trees in a forest that have died naturally or due to other reasons but remain standing. They are an important component of forest ecosystems [1-3], and studies on their quantity and distribution play a crucial role in scientifically managing forests, understanding natural succession within forest communities, and controlling forest pests and diseases. Researchers have predominantly relied on manual field surveys to study the distribution of dead standing trees [4-7]. While this method provides detailed and accurate data with high reliability, it is inefficient and impractical for obtaining large-scale data within a short period. With the application of remote sensing technology in forestry, rapid and large-scale detection and identification of dead standing trees have become feasible.

In studies utilizing remote sensing data for dead standing tree identification, Heurich et al. [8] employed an independent window context segmentation algorithm to segment dual-color infrared images of the Bavarian Forest National Park in Germany, achieving segmentation accuracies of 88% for images from 2001 and 90% for images from 2008. Polewski et al. [9] used a deep convolutional generative adversarial network (DCGAN) to detect the shapes of dead standing trees in multispectral images, with results showing that 45% of the test data performed better than feature shape models. In contrast, this paper utilizes the RetinaNet object detection network to identify dead standing trees in forest images, aiming to achieve large-scale detection and identification of snags.

2. Study Area and Data Overview

The study area of this paper is located in Erguna City, with a total area of 26,042.9 square meters. Unmanned aerial vehicle (UAV) images were captured using the DJI Mavic Pro UAV. The DJI GS Pro ground station software was employed for flight route planning and UAV flight control. Flight parameters were set as follows: a heading overlap of 90%, a sidelap of 80%, a flight altitude of 50 meters, and a flight speed of 3 m/s. To simulate the effect of oblique photography, multiple flights were conducted with the single lens at different tilt angles. In total, 1,911 UAV images were obtained for the study area, including 384 orthoimages and 1,527 oblique images. The Metashape software was used to process the UAV image data, resulting in the generation of orthoimages for the study area.[10]

3. Research Methodology

This paper proposes a dead standing tree detection method based on the RetinaNet object detection network for UAV image recognition. RetinaNet [11] is an object detection network proposed by Facebook's

ISSN 2522-3488 Vol. 8, Issue 2: 110-114, DOI: 10.25236/IJNDES.2024.080217

AI team in 2018. Its main contribution lies in the concept of Focus Loss, which addresses the problem of "class imbalance" in object detection. Since the number of dead standing trees in this study is relatively small compared to normal tree crowns, the RetinaNet object detection network is suitable for dead standing tree identification. The method is detailed below in terms of dead standing tree dataset construction and network structure configuration.

3.1 Construction of Dead Standing Tree Dataset

After obtaining UAV aerial images of the target research area, this paper segments the UAV images into 512*512-pixel patches suitable for the RetinaNet object detection network. To avoid incomplete tree crowns caused by image edge segmentation, a 20% overlap is retained between adjacent images.

Figure 1: Dead Standing Tree Dataset Annotation

After obtaining the segmented images, the positions of dead standing trees are annotated using LabelMe software, as shown in Figure 1. The red box outlines the tree crown of the dead standing tree. A total of 101 annotated images of dead standing trees are obtained from the orthoimage of sample area T1. The images in the dataset are then subjected to random rotations, translations, stretches, and zooms to generate 1010 augmented image samples, completing the construction of the dead standing tree image dataset.

3.2 Network Structure Configuration

In this paper, RetinaNet is adopted as the object detection network structure for dead standing tree detection. The network structure, as shown in Figure 2, consists primarily of a Residual Network (ResNet) combined with a Feature Pyramid Network (FPN) and two fully connected networks (FCN). The dataset is input into the network, and the ResNet feature extraction network is first utilized for feature extraction. Subsequently, the feature pyramid is employed to enhance the multi-scale features formed during feature extraction, thereby obtaining a collection of feature maps with richer information and multi-scale target information. Finally, two fully connected subnetworks are used to accomplish the tasks of target classification and bounding box position regression.

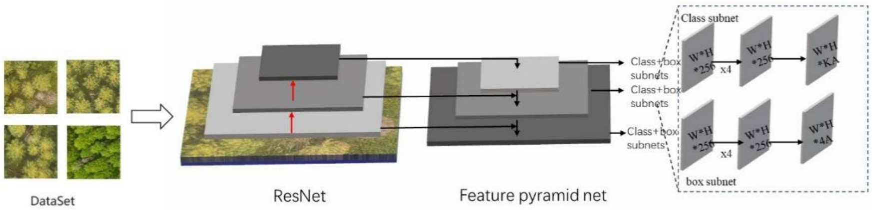

Figure 2: Network Structure

The Residual Network (ResNet) is one of the most widely used feature extraction networks in current applications. Depending on the number of layers, ResNet can be categorized into ResNet-34-FPN, ResNet-50-FPN, ResNet-101-FPN, and ResNet-152-FPN, among others. The selection of the appropriate Residual Network will be based on the actual effectiveness of target recognition; thus, this paper will subsequently choose the suitable Residual Network based on the recognition results.

ISSN 2522-3488 Vol. 8, Issue 2: 110-114, DOI: 10.25236/IJNDES.2024.080217

The major innovation of the RetinaNet network is the proposal of the Focus Loss solution. The design of Focus Loss addresses the significant imbalance between foreground and background objects, such as when the ratio between foreground and background is 1:1000. The formula for Focus Loss is as follows:

$$\mathrm{FL}(p_t) = -\alpha_t (1 - p_t)^{\gamma} \log(p_t) \tag{1}$$

In equation (1), p_t represents the probability of different target objects, while γ and α_t are hyperparameters used to adjust the loss function according to the actual situation. From equation (1), it can be observed that regardless of whether it is a foreground target object or a background region, as the target probability p_t increases, the term $(1 - p_t)^{\gamma}$ decreases. This achieves automatic balance between positive and negative samples. For samples with a larger proportion, balancing adjustment through Focus Loss can reduce their contribution to the overall loss. Conversely, for samples with a smaller proportion, Focus Loss can increase their contribution to the overall loss, thereby enhancing the sensitivity of the neural network to samples with smaller quantities and increasing the overall accuracy of the neural network.

3.3 Accuracy Evaluation

This paper evaluates the accuracy of the dead standing tree (snag) detection method based on the RetinaNet object detection network using the Average Precision (AP). AP is a measure of the average precision of object detection for a particular class. It is defined as the average precision at a set of recall levels [0, 0.1, ..., 1], where the precision value at each recall level is interpolated from the maximum precision value at the corresponding recall level. The formula is as follows:

$$AP = \frac{1}{11} \sum_{r \in \{0, 0.1, \ldots, 1\}} P_{interp}(r) \tag{2}$$

4. Experiment Validation

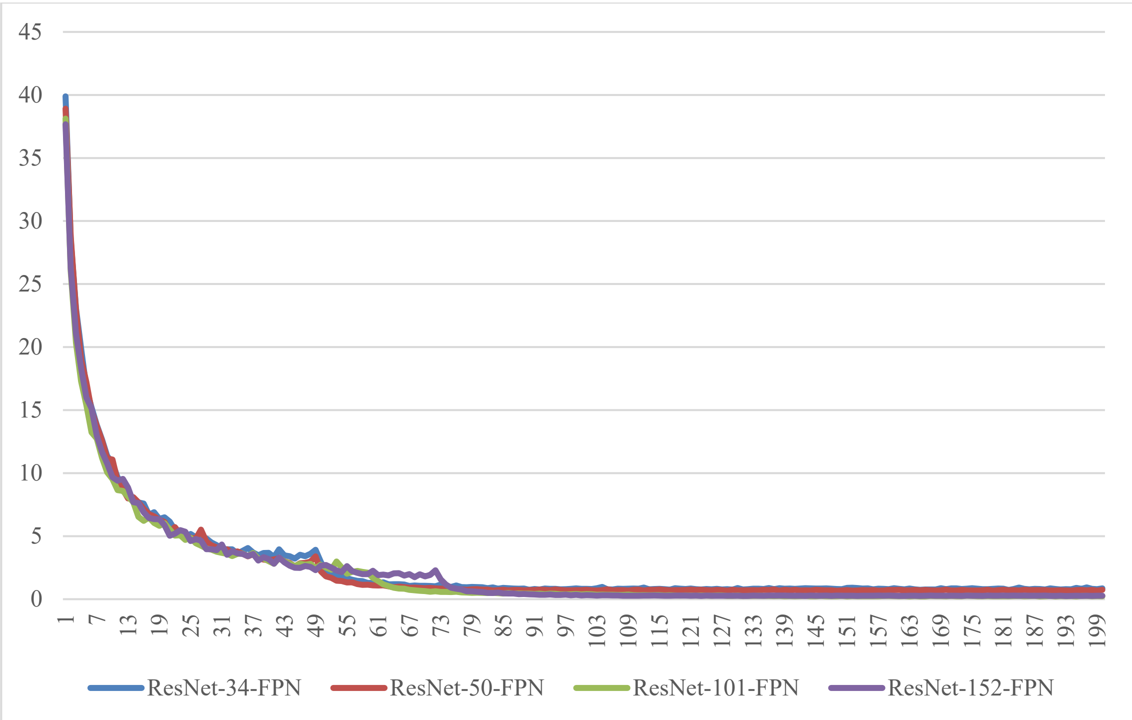

Figure 3: Loss Curves of Different Feature Extraction Networks

The hardware environment for the experiment validation consists of an Intel Xeon Gold 6132 processor running at 2.6GHz, 192GB of memory, and a P100 graphics card with 16GB of memory. The

ISSN 2522-3488 Vol. 8, Issue 2: 110-114, DOI: 10.25236/IJNDES.2024.080217

software environment comprises CentOS 7.2.1511, pytorch 1.4, and CUDA 9.0.

The dataset, which includes 1111 images containing dead standing trees, is divided into two parts. Among them, 957 images are used for training, and 154 images are used for validation. In terms of parameter settings, the neural network's batch size is set to 4 for each training iteration. The initial learning rate is set to 0.1%, the confidence threshold is set to 0.5, and the non-maximum suppression threshold is set to 0.5. The hyperparameters γ and α_t are set to 2 and 0.25, respectively.

In this study, ResNet-34-FPN, ResNet-50-FPN, ResNet-101-FPN, and ResNet-152-FPN were selected as the feature extraction networks to compare and determine the optimal one for dead standing tree recognition. After 200 training iterations, the loss curves of each feature extraction network stabilized. The loss of ResNet-34-FPN stabilized around 0.9, ResNet-50-FPN around 0.7, ResNet-101-FPN around 0.2, and ResNet-152-FPN around 0.3. From the descending trend of the loss curves, it can be observed that ResNet-101-FPN achieved the best performance, followed by ResNet-152-FPN, while ResNet-34-FPN performed the worst. The loss curves of each feature extraction network are shown in Figure 3.

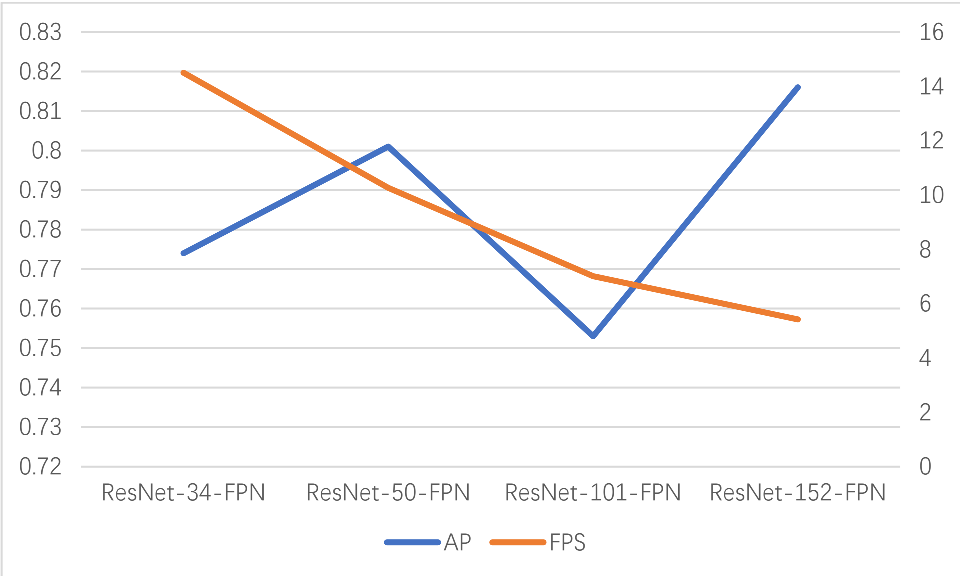

Figure 4: Comparison of Different Feature Extraction Networks in RetinaNet

The results of dead standing tree detection using different feature extraction networks on the validation dataset are shown in Figure 4. In terms of average precision (AP), ResNet-34-FPN, ResNet-50-FPN, ResNet-101-FPN, and ResNet-152-FPN achieved 77.4%, 80.1%, 75.3%, and 81.6%, respectively. The detection speeds of the four feature extraction networks for dead standing tree detection were 14.5, 10.26, 7.01, and 5.42 images per second, respectively. The ResNet-152-FPN, with the longest processing time, also achieved a detection speed of 5.4 images per second, which basically meets the requirement for real-time detection. Therefore, ResNet-152-FPN is more suitable as the feature extraction network for the study area data. The detection results of dead standing tree images are shown in Figure 5.

Figure 5: Dead Standing Tree Recognition Results (Original Images on Top, Detection Results Below)

ISSN 2522-3488 Vol. 8, Issue 2: 110-114, DOI: 10.25236/IJNDES.2024.080217

5. Conclusion

For the identification of dead standing trees in visible light images captured by drones, this paper proposes a dead standing tree detection method based on the RetinaNet object detection network and compares multiple feature extraction networks. Through experiments, it was found that the ResNet-152-FPN feature extraction network performed the best for the study area, achieving an average precision (AP) of 81.6% with only 101 dead standing tree images. This indicates a relatively high level of recognition accuracy, effectively supporting the identification and large-scale survey of dead standing trees.

References

[1] Oettel J, Lapin K, Kindermann G, et al. Patterns and drivers of deadwood volume and composition in different forest types of the Austrian natural forest reserves[J]. Forest Ecology and Management, 2020, 463(05): 118016.

[2] Franklin J, Shugart H, Harmon M. Tree Death as an Ecological Process[J]. Bioscience, 1987, 37(8): 550-556.

[3] Lu Zhijun, Liu Fuling, Wu Hao, et al. Species composition, size class, and distribution pattern of dead standing trees in the evergreen and deciduous broad-leaved mixed forest of Badagongshan Mountain[J]. Biodiversity Science, 2015, 23(02): 167-173.

[4] Guo Yili, Wang Bin, Xiang Wusheng, et al. Spatial pattern and habitat association analysis of dead standing trees in the seasonal rainforest of southwestern Guangxi[J]. Guihaia, 2016, 36(02): 154-161.

[5] An Yun, Ding Guodong, Gao Guanglei, et al. Quantity characteristics and distribution pattern of dead standing trees in natural secondary forests in the North China Loess Plateau[J]. Bulletin of Soil and Water Conservation, 2012, 32(04): 246-250.

[6] Heurich, Marco, Nielsen, et al. Automatic Mapping of Standing Dead Trees after an Insect Outbreak Using the Window Independent Context Segmentation Method[J]. Journal of Forestry, 2014,112(6):564-571.

[7] Polewski P, Shelton J, Yao W, et al. Segmentation of single standing dead trees in high-resolution aerial imagery with generative adversarial network-based shape priors[J]. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2020,XLIII-B2-2020:717-723.

[8] Briechle S, Krzystek P, Vosselman G. Silvi-Net – A dual-CNN approach for combined classification of tree species and standing dead trees from remote sensing data[J]. International Journal of Applied Earth Observation and Geoinformation, 2021,98:102292.

[9] Lin T Y, Goyal P, Girshick R, et al. Focal Loss for Dense Object Detection[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2018,PP(99):2999-3007.

[10] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition: IEEE Conference on Computer Vision & Pattern Recognition, Las Vegas, NV, USA, 2016[C].

[11] Lin T Y, Dollár P, Girshick R, et al. Feature Pyramid Networks for Object Detection: IEEE Conference on Computer Vision & Pattern Recognition, Las Vegas, NV, USA, 2016[C].