Nature Environment and Pollution Technology
An International Quarterly Scientific Journal

p-ISSN: 0972-6268
(Print copies up to 2016)
e-ISSN: 2395-3454

Vol. 21 | No. 4 | pp. 1609-1617 | 2022

Original Research Paper

https://doi.org/10.46488/NEPT.2022.v21i04.013

Open Access Journal

Optimal Models for Plant Disease and Pest Detection Using UAV Image

Dashuang Liang, Wenping Liu† and Yugang Zhao
School of Information Science and Technology, Beijing Forestry University, Beijing, 100083, China
†Corresponding author: Wenping Liu; wendyl@vip.163.com

Nat. Env. & Poll. Tech.
Website: www.neptjournal.com

Received: 25-12-2021
Revised: 09-02-2022
Accepted: 12-02-2022

Key Words:
Anchor-free detector
Deep learning
Unmanned aerial vehicle
Automated machine learning
Pest detection

ABSTRACT

The use of deep learning methods to detect plant diseases and pests based on UAV images is an important application of remote sensing technology in modern forestry. This paper uses a CenterNet-based object detection method to construct models for plant disease and pest detection. The accuracy of the models is influenced by parameter alpha, which is used to control the affine transformation in the preprocessing of CenterNet. First, different alphas are sampled for training and testing. Next, the least square method is used to fit the curve between alpha and accuracy measured by mAP (mean average precision). Finally, the equation of the curve is fitted as mAP = -0.22 * alpha^2 + 0.32 * alpha + 0.42. In comparison, an automated machine learning (AutoML) method is also conducted to automatically search for the best model. The experiments are done with 5,281 images as the training dataset, 1,319 images as the verification dataset, and 3,842 images as the test dataset. The results show that the best alpha value obtained by the least square method is 0.733, and the accuracy of the corresponding model is 0.536 in mAP@[.5, .95]. In contrast, the accuracy of the AutoML method model is higher with the model accuracy of 0.545 in mAP@[.5, .95]. However, the training time and training resource consumption of the AutoML method are about 3 times that of the least square method. Therefore, in practice, a trade-off should be made according to the accuracy requirements, resource consumption, and task urgency.

INTRODUCTION

Plant diseases and insect pests seriously threaten the growth of forests and can be great impediments to forest health and forestry production (Zhang et al. 2010). Traditionally, the monitoring methods of forestry pests and diseases mainly consisted of field surveys. These manual scoring and counting through field surveys are expensive, and the monitoring methods have time lags and strong subjectivity. Moreover, in areas with dangerous terrain and restricted access, the surveyors are unable to discover plant diseases and insect pests in time (Chiu 1993). Therefore, it is necessary to detect plant diseases and insect pests more accurately and quickly, which will help to develop early treatment technologies and greatly reduce economic losses at the same time (Fuentes et al. 2017).

In the early 1970s, with the launch of the first remote sensing satellite, remote sensing images were used to monitor forestry diseases and pests (Gao et al. 2006, Lehmann et al. 2015). However, it is difficult to popularize and widely apply because of its inaccurate positioning accuracy, high cost, weather influence, and relatively long imaging cycle (Wu 2013). Fortunately, with the rapid development of unmanned aerial vehicle (UAV) technology and continuous improvement in its performance, UAV remote sensing has many advantages such as low cost, high precision, simple operation, and flexibility (Tang 2014). At present, people have also begun to explore the application of UAV remote sensing image in monitoring forest diseases and insect pests, combined with traditional computer vision methods and image analysis technology. The application integrates a global positioning system and geographic information system to detect the distribution of pests and achieves good results (Tetila et al. 2020, Yuan & Hu 2016).

In addition, Hinton et al. (2016) proposed the concept of deep learning. Krizhevsky et al. (2012) first applied the convolutional neural network (CNN) to the ImageNet large-scale visual recognition challenge (ILSVRC). In the ILSVRC -2012 challenge, the trained deep CNN won first place in tasks of image classification and object detection, and the error rate was far lower than the other programs. Since then, deep learning has been rapidly applied to different research fields and has achieved great success in many fields, including image classification (Huang et al. 2016), object detection (Girshick 2015, Redmon et al. 2016), image segmentation (Chen et al. 2018, Lee & Park 2020) and so on.

In recent years, deep CNN are applied to the detection of plant disease severity and has been proven to be a good method. For example, Liu et al. (2018) proposed a new CNN

model based on Alexnet to identify four common apple leaf diseases. The overall accuracy of the model was 97.62%, and the parameters were reduced by 51,206,928 compared with the standard Alex net model. Xie et al. (2020) proposed a faster DR-IACNN model by introducing the concept-v1 module, concept-resnet-v2 module, and SE module. The results showed that the average detection accuracy of the detection model for the grape leaf disease dataset was 81.1% and the detection speed was 15.01 FPS. Wang et al. (2019) proposed a corn leaf disease segmentation method based on an improved fully convolution network (FCN). The method mainly included an encoding network and corresponding decoding network, as well as a pixel-level classifier behind the decoding network. This method had good segmentation performance and could accurately segment the diseased area of corn leaves.

Similar to computer vision tasks, the application of CNN in the research of plant diseases and insect pests can be roughly divided into classification networks, object detection networks, and segmentation networks according to their different tasks. The object detection network model not only gives the types of pests and diseases but also accurately finds out the location of the pests and diseases on the images. Object detection methods are further roughly divided into two categories, anchor-based object detection methods, and anchor-free object detection methods. For example, anchor-based object detection methods include the Faster RCNN series (Girshick 2015, Ren et al. 2017), YOLO series (Bochkovskiy et al. 2020, Redmon et al. 2016, 2017, 2018), SSD series (Jeong et al. 2017, Liu et al. 2016, Zhang et al. 2020), RFCN (Dai et al. 2016) and so on. All these methods rely on a set of predefined anchor boxes. Thus, to avoid the shortcomings brought by the pre-defined anchor box, people began to study the anchor-free object detection method. This method does not need to use anchors but adopts the idea of key point regression. First, some key points are defined to describe an object, and then each key point of the object is regressed. There are some classic methods, such as CornerNet (Law & Deng 2020), CenterNet (Zhou et al. 2019a), FCOS (Tian et al. 2020), and extreme (Zhou et al. 2019b), etc., all of which take object detection as a standard key point estimate problem.

In the above anchor-free object detection algorithms, CenterNet is a simple, fast, and accurate detector without any non-maximum suppression (NMS) as the post-processing method. In the training of the original CenterNet, a random affine transformation is done to each image to enrich the data, and the new object's boundary box is generated by clip operation, which is usually not correct. To make the boundary box generated by clip operation as accurate as possible, Liang et al. (2021) only used a criterion that the proportion of the remaining objects should be at least 90% of the size of the original objects. Therefore, this study mainly focuses on identifying an optimal model based on parameter alpha, which is the parameter used to control the affine transformation in the preprocessing of CenterNet, to improve the accuracy of detecting plant diseases and pests. Both the least square method and an automated machine learning (AutoML) method are proposed to find an appropriate model for detecting plant diseases and pests.

MATERIALS AND METHODS

Study Area

The study area is located in Lingyuan City, Liaoning Province, Northeast China. Lingyuan City is located in the west of Liaoning Province, bordering Hebei Province and Inner Mongolia. The main vegetation is Chinese red pine (*Pinus tabuliformis*) and scatted with some poplar. The pest dendroctonus (*Scolytidae*) has caused damage and tree death in the area.

Six sample plots that are being infected by plant diseases and pests were selected. They were named plot 1 to plot 6. Fig. 1 shows the detailed location and distribution of these plots on the map.

Data Collection

The images of the UAV were taken in the study area from August 11 to August 12 during the growing season, which can better reflect the characteristics of plant growth. UAV took off from the center of six sample plots, and a batch of images was taken in each sample plot at different heights. The model of the UAV is four rotor DJI inspire2, equipped with DJI X5 professional camera, and the resolution of each image is 5280 × 3956 pixels.

CenterNet Principles

For image I, after the full convolution network, three outputs will be generated. One is the feature map containing C (number of categories of the detection task) layers. Each feature map corresponds to one category. Each pixel on the feature map represents the score of the pixel belonging to the center of an object. Another layer represents the offset ($x/S_x - x/S_x$, $y/S_y - y/S_y$), where (x, y) is the center of an object with class C on the input image, and (S_x, S_y) is the horizon and vertical scale parameters of the input image to that of the output feature image, respectively. The number of this layer is 1, and each pixel on this layer represents the offset between the real center point position and center pixels on the feature map. The other layer represents the length and

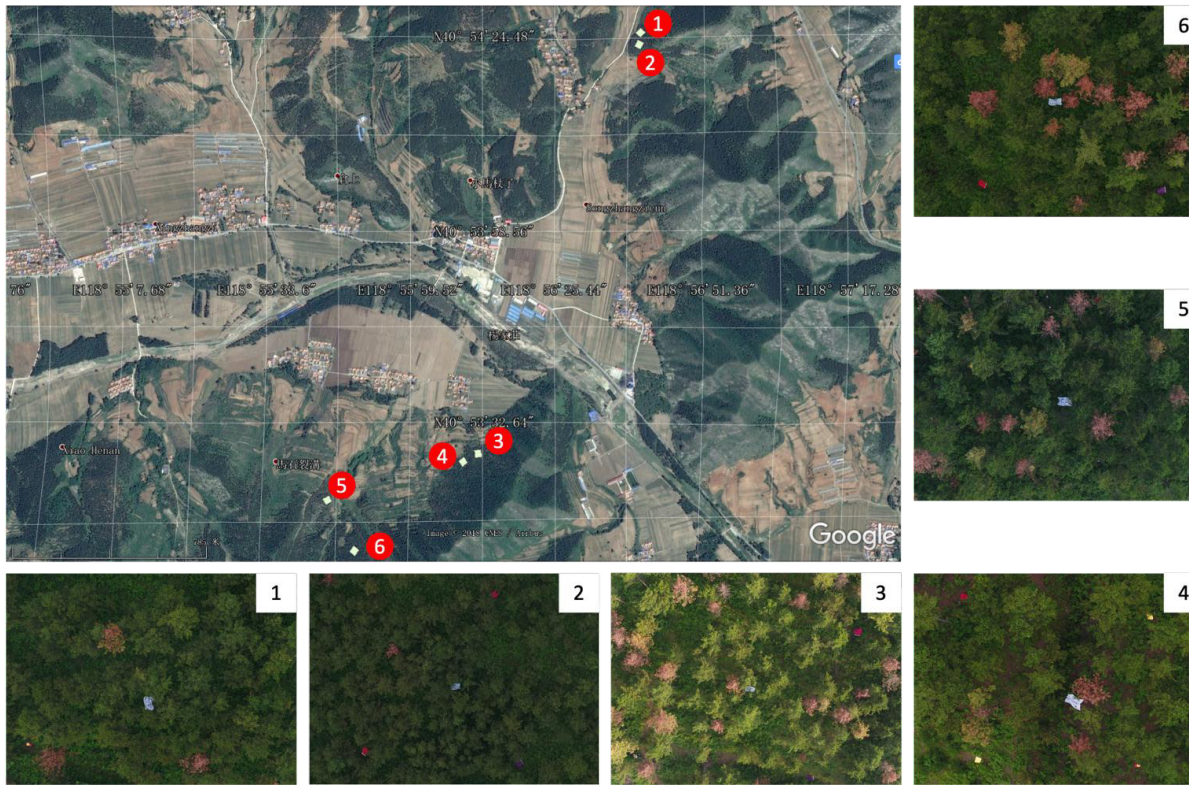

Fig. 1: The study area.

width of the boundary box of an object. Each pixel on this layer represents the length and width of the object's boundary box (box) corresponding to the point. The overall structure of CenterNet is shown in Fig. 2.

Data Enhancement by Affine Transformation

In the original CenterNet, before sending the image to the network, an affine transformation of random center translation and scaling is made to the image. As a result, some objects on the image will be removed from the new image boundary, while some objects on the image will be retained in the new image, and other objects on the image will fall on the new image boundary, as shown in Fig. 3(b). For objects that fall on the boundary, the original CenterNet only recalculates the boundary box's coordinates of the remaining objects with a simple clip operation. The new boundary box of the remaining objects calculated by the clip operation is often inaccurate, as shown in Fig. 3(b), while Fig. 3(c) shows the ground truth of the boundary box of the remaining object.

For an object that remains in the new image after affine transformation to the original image, the ratio of the area of the new boundary box (boundary box created by clip operation) to the original boundary box (boundary box in the original image) is defined as alpha. It is known that, for an object that falls on the new image's boundary after the affine transformation to an image, the larger the alpha ratio, the more accurate the new boundary box obtained by clip operation is. As the shapes of the plants' canopy are often approximate to circular, in most situations, with the growing of alpha, the remaining object' boundary box calculated by clip operation will be more accurate, as shown in Fig. 4.

Optimal Parameter Search by Least Square Fitting

Alpha is a scale factor, and its value falls within the range of [0, 1]. To quickly obtain the appropriate value of alpha, the following strategies are adopted: after affine transformation, no constraints are made for the objects that are moved out of the image or the whole objects that are still within the image scope, while the alpha value is required to be greater than 0.5 to objects that fall on the image boundary, that is, the alpha value must be limit to the range of [0.5, 1] if some objects fall on the image boundary after an affine transformation.

To explore the relationship between alpha and the detection accuracy of diseased and pest plant detectors, different alpha values are selected to train models on the same training dataset and then tested on the same test dataset. Finally, the

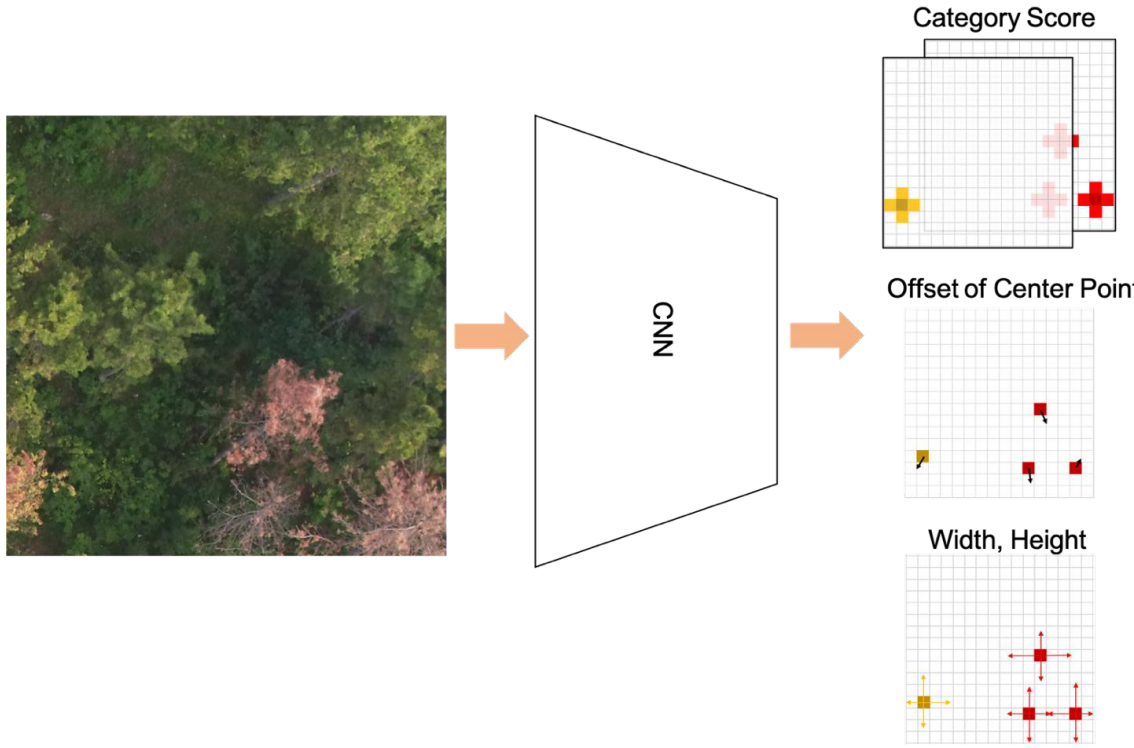

Fig. 2: Structure of Centernet network.

least square method is used to fit the curve between alpha value and object detection accuracy.

AutoML Method for Optimal Model Search

In addition to the above-mentioned method to search for an optimal parameter, another AutoML method is also used to search for an optimal model automatically for the detection of plant diseases and pests here. Following is the procedure:

Step 1. In the first epoch of training, the model trained from ImageNet is used as the initial model, and 9 different alphas values chosen from the interval of [0.5, 0.95] with a step of 0.05 are used to train the 9 models.

Step 2. In the next epoch of training, the best model among the 9 models obtained by the previous epoch is re-

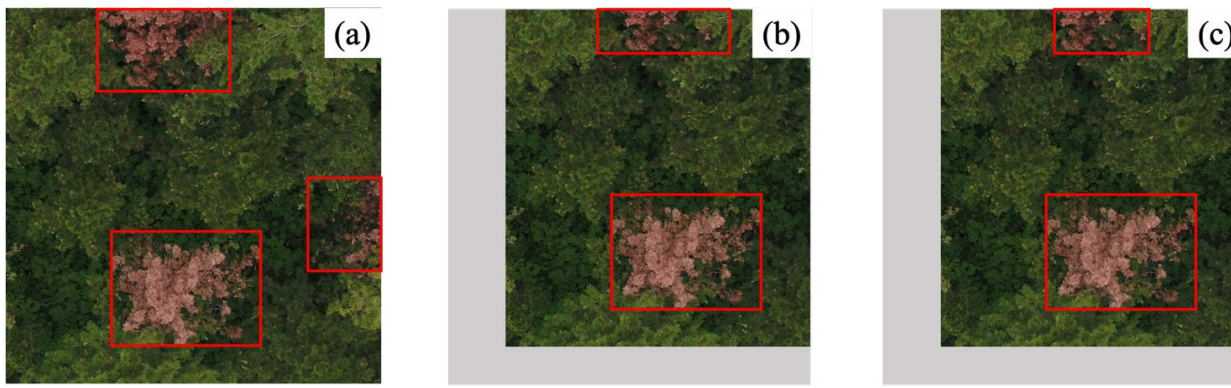

Fig. 3: (a) Original image, (b) boundary box generated by clipping after random affine transformation, (c) the ground truth of the boundary box after an affine transformation.

garded as the pre-training model, and then different alpha values within [0.5, 0.95] with the step of 0.05 are chosen again to train another 9 different models.

Step 3. Repeat Step 2 until the last epoch is completed or the training is converged.

The general schematic diagram of the whole process is shown in Fig. 5:

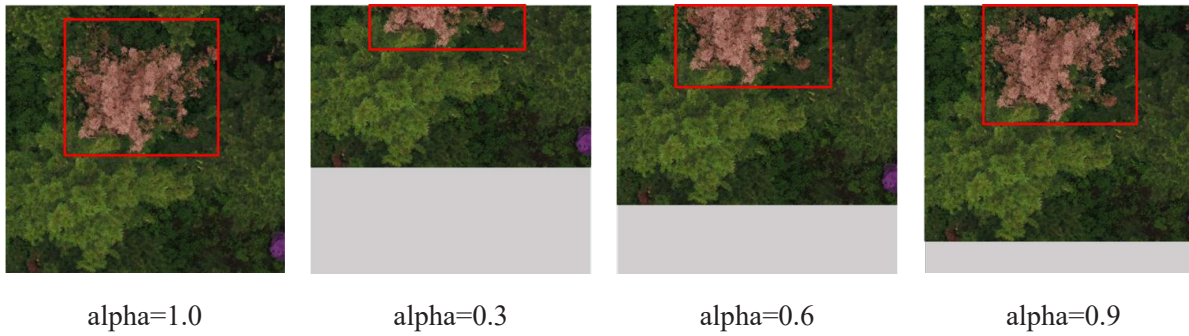

Fig. 4: Boundary box generated by clipping after random affine transformation with different alpha values.

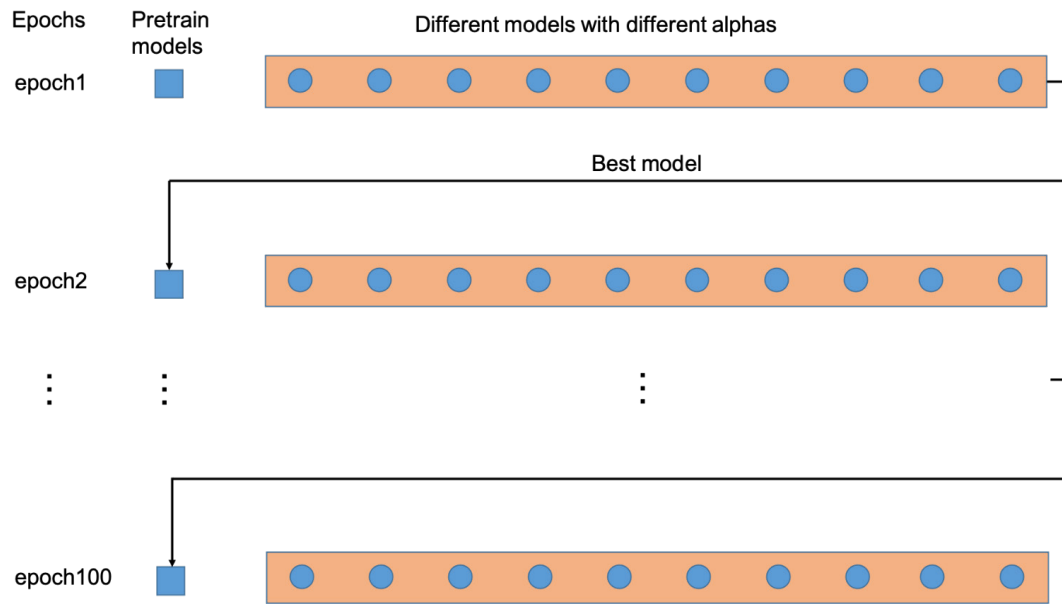

Fig. 5: Flow chart of AutoML method.

EXPERIMENTS

Data Preparation

The size of the original image obtained by the UAV is 5,280 × 3,956 pixels. Each original image will be cropped to smaller images with sizes between 1,000 and 2,000 pixels before training. This study only focuses on infected plants and dead plants, which are mainly yellow and red in the image. In the six sample plots, the images of sample plot 1, sample plot 3, sample plot 4, and sample plot 6 are divided into training and verification datasets, while the images of sample plot 2 and sample plot 5 are used as test datasets. Finally, the training and verification datasets contain 5,281 images and 1,319 images respectively, and the test dataset contains 3,842 images. In all, the training dataset contains about 2,433 infected instances and 13,539 dead instances, while the testing dataset has 938 infected instances and 3,144 dead instances, respectively. Before training, to obtain as many samples as possible, more data are created through data enhancement methods such as flipping, random color, random rotation, and random clipping.

Training

As Pytorch deep learning framework is one of the most famous and fastest deep learning frameworks, it is chosen to train the models on GPU machines in this experiment. In all these experiments, no matter how large the original image is, it will be scaled to a fixed size of 512 * 512 through affine transformation, and then sent to the backbone of the CenterNet network. The backbone network is the classic resnet-101 structure. The training is carried out from 0 and is stopped after 100 epochs. The learning strategy is Adam's method. The detailed training super parameters are shown in Table 1.

Test and Comparison

To evaluate the final test results, the official coco API (Lin & Dollar 2016) is used to calculate the accuracy of mAP, which is the average precision (AP) in multiple IOU thresholds ranging from 0.5 to 0.95 with the step of 0.05. It is simply defined as mAP@[.5, .95], which is used as the performance indicator in object detection tasks. IOU is calculated as follows:

$$\text{IOU(A, B)}=\frac{A \cap B}{A \cup B} \quad \ldots(1)$$

where A represents the boundary box of ground truth, and B represents the predicted boundary box by the network.

The Average Precision (AP) is the area under the Precision-Recall curve for the detection task. As in the COCO Challenge, the AP is computed by averaging the precision over a set of spaced recall levels from 0 to 1 with steps of 0.01.

$$AP=\frac{1}{11}\sum_{r\in\{0,0.1,\ldots,1\}} P_{\text{int}erp}(r)$$

$$P_{\text{int}erp}(r)=\max_{\tilde{r}:\tilde{r}\geq r} p(\tilde{r}) \quad \ldots(2)$$

Where $p(\tilde{r})$ is the measure precision at recall $\tilde{r}$? AP (Average Precision) is a concept of integrating precision as recall varied from 0 to 1, and mAP is defined as the average of AP for all of the object classes.

Table 1: Parameters of network training.

Argument	Value
Mini-batch size	8
Num_epochs	100
Lr_policy	Multistep
Step value	40, 80
Initial learning rate	1.25e-4
Gamma	0.1

RESULTS

Results of Least Squares Parameter Fitting Parameter Model

In the interval of [0.5, 0.95], we take 0.05 as the step to obtain different alpha values while keeping other parameters unchanged. different models are trained on the same training data and then tested on the same test set. The detailed results are shown in Table 2:

As shown in Table 2, at the lower alpha levels, the object detection accuracy increases with the increase of alpha; when alpha increases to about 0.75, the value of mAP reaches the highest. After that, the mAP decreases with the increasing value of alpha, showing a U-shaped pattern (quadratic). Therefore, based on this observation, taking the alpha parameter as the independent variable and mAP as the dependent variable, an expression of a quadratic equation is established: mAP=a*alpha^2 + b*alpha + c.

Fitting the quadratic curve equation mAP = a * alpha^2 + b*alpha + c with the least square method, the result coefficients of a, b and c are -0.22, 0.32, and 0.42, respectively. Therefore, the expression of the curve is mAP = -0.22 * alpha^2 + 0.32 * alpha + 0.42, which is a bottom-up quadratic curve. In the range of [0.5, 0.95], the accuracy of mAP increases at the beginning and then decreases with the increase of alpha. When alpha reaches 0.733, the maximum value of the curve is 0.533. The overall curves are shown in Fig. 6. The red curve is the broken line connected by the original points, and the green curve is the least square fitting result of a one-dimensional quadratic equation curve.

Results of different models

For comparison, the least squares method model trained with the alpha of 0.733, which is the maximum point of the quadratic curve, together with the model created by the AutoML method and some other standard object detection models such as Faster RCNN (Ren et al. 2017), SSD (Liu et al. 2016), CenterNet (Zhou et al. 2019a) are all tested on the same test dataset. The results of all these models are as follows in Table 3.

As can be seen from the table, the accuracy of the AutoML method model and the least square method model are 0.545 and 0.536, respectively. Among the anchor-based methods (Faster RCNN, SSD, and RetinaNet), RetinaNet attained the highest accuracy, reaching 0.480 in mAP@[.5, .95]. When compared with all the models, CenterNet (AutoML) reached the best accuracy of 0.545 in mAP@[.5, .95].

It also can be seen in Fig. 7 that in the early stage of training, the convergence speed of the AutoML method model is faster than that of the least square method model, and when

it goes to about the 20th epoch, the accuracy of the AutoML method model almost equals to convergent accuracy of least square method model.

DISCUSSION

Can Affine Transformation Operation be Replaced with Resize Operation

To verify the effect of the affine transformation operation on the model's accuracy, the affine transformation operation is replaced by a simple resizing operation in the training stage. Therefore, random translation and scaling affine transformation operation are removed in the preprocess of training, and all the original images are uniformly resized to 512 * 512 pixels to train a CenterNet model. the accuracy of the experimental results can only reach 0.493 in mAP@[.5, .95]. It is lower than the original CenterNet's accuracy with 0.498 in mAP@[.5, .95]. This may be because the resizing operation uniformly scales all original images to a size of 512 * 512 pixels, on the contrary, the operation of random affine transformation in the original Centernet generates images of multiple different sizes (Fig. 8), which can enrich the richness of the training data. As a result, although resizing operation can produce a more accurate boundary box than that of the affine transformation operation, it is no better than the affine transformation operation in improving the detection model's accuracy.

Analysis of the Relationship Between Alpha and Detection Accuracy

Generally, when the affine transformation is performed on the image, some plant objects may fall on the image boundary. The larger the alpha ratio, the more accurate the newly created boundary box of the remaining plants is. However, Fig. 6 indicates that in the experiment, the accuracy of ob-

Table 2: Object detection accuracy corresponding to different alphas.

alpha	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
mAP	0.522	0.526	0.528	0.530	0.533	0.534	0.533	0.529	0.528	0.522

Table 3: Detection results comparison using different frameworks and network architectures.

Method	backbone	mAP@[.5]	mAP@[.75]	mAP@[.5, .95]
Faster RCNN	ResNet-101	0.693	0.535	0.472
SSD	ResNet-101	0.573	0.489	0.451
RetinaNet	ResNet-101-FPN	0.724	0.546	0.480
CornerNet	Hourglass-104	0.722	0.559	0.491
CenterNet	ResNet-101	0.703	0.557	0.498
CenterNet (LSM)	ResNet-101	0.726	0.584	0.536
CenterNet (AutoML)	ResNet-101	0.737	0.603	0.545

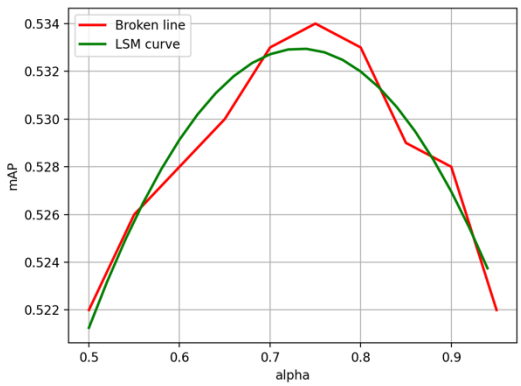

Fig. 6: Least square fitting curve and broken line connected by the original points.

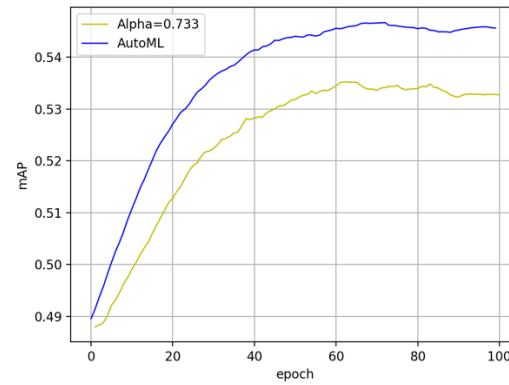

Fig. 7: Convergence curves of the least square fitting model with alpha=0.733 and AutoML model.

ject detection does not increase with the increase of alpha all the time, but actually increases at first and then begins to decrease. The reason is mainly that during the increase of alpha from 0.600 to 0.733, the new boundary box calculated by clip operation for the remaining plants becomes more accurate. As the alpha continues to increase, although the new boundary box obtained by the clip operation is more accurate, the training data's diversity begins to decrease. Nevertheless, the deep learning model needs to be driven by big data with high richness to get better accuracy. It can be seen from Fig. 6 that after the alpha increases to 0.733, the accuracy of the model begins to show a downward trend, which is mainly because the low data richness begins to play a greater impact on the model's accuracy from then on. On the whole, in the range of 0.5 to 0.95, alpha presents a univariate quadratic monotonic function to the accuracy of the model, and it makes the accuracy of the model increase at first and then begins to decrease.

Advantages and Disadvantages of Automl Method and Least Squares Method

In comparison to the least square method model, the AutoML model's accuracy is higher and it is an end-to-end training method without manual intervention, but the training time and training resources of the AutoML model are also higher. From the overall point of view, the final accuracy of the AutoML model is 0.545 in mAP@[.5, .95], which is higher than the accuracy of 0.536 in mAP@[.5, .95] selected by the least square method. However, since the variables alpha and mAP@[.5, .95] conform to a quadratic equation of one variable, only 3 different alphas and mAPs@[.5, .95] are needed to determine the equation expression. that is, the univariate quadratic equation can be determined only after 3 rounds of training and testing with different alphas. By contrast, when training the AutoML model, in each epoch 9 models need to be trained with 9 different alpha values from the parameter interval of [0.5, 0.95] with a step of 0.05. To sum up, though the AutoML method is an end-to-end method, the time cost and training resource cost of the AutoML method to find an optimal is about 3 times that of the least square method.

CONCLUSIONS

This article mainly proposed an AutoML method and a least square method for searching for an optimal model for detecting diseases and pest plants on the UAV image. The least square method finally fits an optimal alpha with the value of 0.733, and the accuracy of the corresponding model can reach 0.536 in mAP@[.5, .95]. The AutoML method uses the best model in the previous epoch as the initial model in each epoch, and the accuracy of the AutoML method is 0.545 in mAP@[.5, .95], which is higher than that of the optimal model find by the least square method. However, the training duration and training resource consumption of the AutoML model is about 3 times that of the least square method model. Therefore, in the actual application, a trade-off can be made according to the accuracy requirements, resource consumption, and task duration.

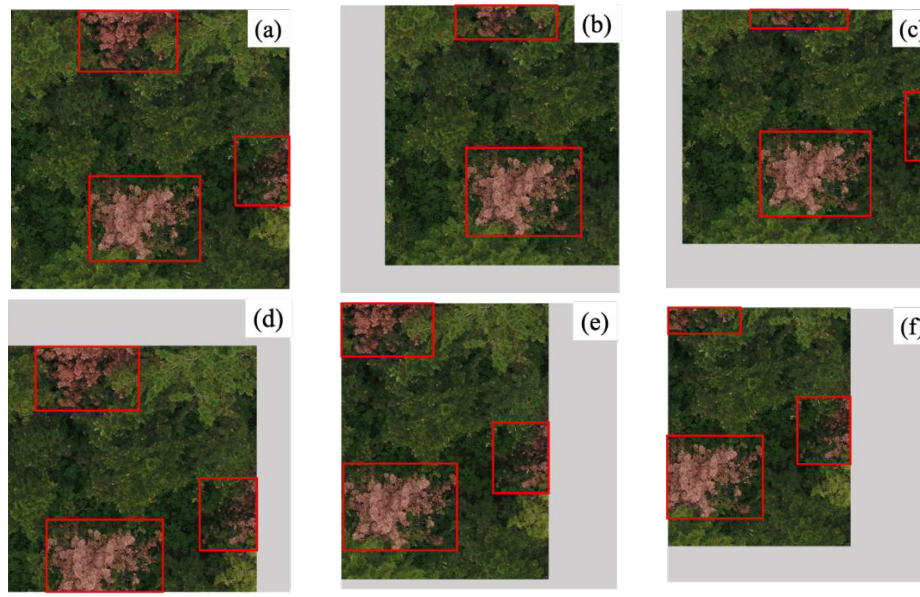

Fig. 8: (a) original image, (b), (c), (d), (e) and (f) images and boundary box labels generated by affine transformation operation.

ACKNOWLEDGMENTS

The authors are very grateful to the Lingyuan Forestry Bureau for assisting in the data collection process.

REFERENCES

Bochkovskiy, A. Wang, C. and Liao, H. 2020. Yolov4: Optimal speed and accuracy of object detection. arXiv Preprint arXiv:2004.10934.

Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K. and Yuille, A.L. 2018. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFS. IEEE Trans. Pattern Anal. Mach. Intell., 40(4): 834-848.

Chiu, S. F. 1993. Investigations on botanical insecticides in south China: An update. Bot. Pest. Integr. Pest Manag., 19: 134-147.

Dai, J.F. Li, Y., He, K.M. and Sun, J. 2016. R-FCN: Object detection via region-based fully convolutional networks. Adv. Neural Inform. Process. Sys., 41: 379-387.

Fuentes, Y., Kim, S.C. and Park, D. S. 2017. A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. Sensors, 17(9): 2022.

Gao, Y., Liu, D., Zhang, F. and Yang, X. 2006. The application development of satellite remote sensing technology in the assessment of forest damage. Chin. Agric. Sci. Bull., 22(2): 113-117.

Girshick, R. 2015. Fast R-CNN. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 1440-1448.

Huang, G., Liu, Z., Laurens, V.D.M. and Weinberger, K.Q. 2017. Densely Connected Convolutional Networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, pp. 2261-2269. https://doi.org/10.1109/CVPR.2017.243.

Jeong, J., Park, H. and Kwak, N. 2017. Enhancement of SSD by concatenating feature maps for object detection. arXiv:1705.09587.

Krizhevsky, A., Sutskever, I. and Hinton, G. E. 2012. Imagenet classification with deep convolutional neural networks. Adv. Neural Inform. Process, Syst., 25: 1097-1105.

Law, H. and Deng, J. 2020. Cornernet: Detecting objects as paired key points. Int. J. Comp. Vision, 128(3): 642-656.

Lee, Y. and Park, J. 2020. Centermask: Real-time anchor-free instance segmentation. Proceedings of the IEEE/CVF Conference On Computer Vision And Pattern Recognition, Seatle, WA, pp. 13906-13915.

Lehmann, J.R.K. Nieberding, F., Prinz, T. and Knoth, C. 2015. Analysis of unmanned aerial system-based cir images in forestry: A new perspective to monitor pest infestation levels. Forests, 6(3): 594-612.

Liang, D., Liu, W., Zhao, L., Zong, S. and Luo, Y. 2021. An improved convolutional neural network for plant disease detection using unmanned aerial vehicle images. Nature Environ. Pollut. Technol., 10: 386. https://doi.org/10.35940/ijrte.F1110.038620.

Lin, T.Y. and Dollar, P. 2016. Ms coco API. https://github.com/pdollar/coco.

Liu, B., Zhang, Y., He, D. and Li, Y. 2018. Identification of apple leaf diseases based on deep convolutional neural networks. Symmetry, 10(1): 11. https://doi.org/10.3390/sym10010011.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y. and Berg, A.C. 2016. SSD: Single shot multibox detector. Europ. Conf. Comput. Vision., 19: 21-37.

Redmon, J., Divvala, S. Girshick, R. and Farhadi, A. 2016. You only look once: Unified, real-time object detection. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779-788. https://doi.org/10.1109/CVPR.2016.91.

Redmon, J. and Farhadi, A. 2017. YOLO9000: Better, faster, stronger. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779-788. https://doi.org/10.1109/CVPR.2016.91.

Redmon, J. and Farhadi, A. 2018. Yolov3: An incremental improvement. arXiv:1804.02767.

Ren, S., He, K., Girshick, R. and Sun, J. 2017. Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Trans. Pattern Anal. Mach. Intell., 39(6): 1137-1149.

Tang, Y. 2014. Research on the vegetation identification method based on UAV image acquisition, Chengdu University of Technology, Chengdu, Sichuan.

Tetila, E.C., Machado, B.B., Astolfi, G., de Souza Belete, N A. Amorim, W. P. Roel and A .R. Pistori, H. 2020. Detection and classification of soybean pests using deep learning with UAV images. Computers and Electronics in Agriculture, 179, 105836. https://doi.org/10.1016/j.compag.2020.105836.

Tian, Z., Shen, C., Chen, H. and He, T. 2020. FCOS: Fully convolutional one-stage object detection, 2019 IEEE/CVF International Conference on Computer Vision (ICCV). arXiv:1904.01355.

Wang, Z. Shi, Y. and Li, Y. 2019. Segmentation of corn leaf diseases based on improved fully convolutional neural network. Computer Engineering and Applications. 55(22): 127-132. (Abstract in English)

Wu, Q. 2013. Research on bursaphelenchus xylophilus area detection based on remote sensing image. Anhui University, Hefei, Anhui.

Xie, X., Ma, Y., Liu, B., He, J. and Wang, H. 2020. A deep-learning-based real-time detector for grape leaf diseases using improved convolutional neural networks. Front. Plant Sci., 11: 751. https://doi.org/10.3389/fpls.2020.00751.

Yuan, Y. and Hu, X. 2016. Random forest and objected-based classification for forest pest extraction from UAV aerial imagery. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLI-B1: 1093-1098.

Zhang, S., Wen, L., Lei, Z. and Li, S.Z. 2020. Refinedet++: Single-shot refinement neural network for object detection. IEEE Trans. Circ. Syst. Video Technol:z, 31(2): 674-687.

Zhang, T. Zhang, X. Liu, H. and Pei, X. 2010. Application of remote sensing technology in monitoring forest diseases and pests. Journal of Anhui Agricultural Sciences, 38(21): 11604-11607.

Zhou, X. Wang, D. and Krhenbühl, P. 2019a. Objects as points. arXiv:1904.07850.

Zhou, X. Zhuo, J. and Krhenbühl, P. 2019b. Bottom-up object detection by grouping extreme and center points. arXiv:1901.08043.