



**HAL**  
open science

# Data Augmentation Method on Pine Wilt Disease Recognition

Weibin Li, Bingzhen An, Yuhui Kong

► **To cite this version:**

Weibin Li, Bingzhen An, Yuhui Kong. Data Augmentation Method on Pine Wilt Disease Recognition. 5th International Conference on Intelligence Science (ICIS), Oct 2022, Xi'an, China. pp.458-465, 10.1007/978-3-031-14903-0\_49 . hal-04666455

**HAL Id: hal-04666455**

**<https://hal.science/hal-04666455v1>**

Submitted on 1 Aug 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

# Data Augmentation Method on Pine Wilt Disease Recognition<sup>\*</sup>

Weibin Li<sup>1,2</sup>, Bingzhen An<sup>1,2</sup>, and Yuhui Kong<sup>1,2</sup>

<sup>1</sup> School of Artificial Intelligence, Xidian University, Xi'an 710075, China  
weibinli@xidian.edu.cn(W.L.); abz24423@163.com(B.A.);  
kongyuhui527@126.com(Y.K.)

<sup>2</sup> Beidou Space-Time Intelligence Research Center, Xidian University,  
Xi'an 710075, China

**Abstract.** In recent years, deep learning has made a breakthrough in image recognition. However, it often requires a large amount of label data as the sample set. In most practical applications, the neural network is prone to over-fitting or weak generalization due to the lack of annotation data. This phenomenon is especially obvious in a small-scale data set. To solve this problem, pine wilt disease data is used as an example to adopt mirroring, flipping, adding noise, rotating, scaling, and other augmentation methods to enhance the amount of the image sets. It can not only increase sample diversity but also make the network more stable for training. In this paper, the effects of different amplification methods and training samples size on the Faster R-CNN and YOLOv3 models are tested, and its results show that scaling has the greatest impact on the two models for the reason that the two models are both sensitive to the size of sample images. The accuracy of Faster R-CNN starts to decline when the number of training sets is expanded to 60 percent of the new training samples, the accuracy of YOLOv3 starts to decline when the number of training sets is expanded to 75 percent.

**Keywords:** data augmentation · image recognition · small-scale data sets · pine wilt disease.

## 1 Introduction

Recently, Pattern Recognition System based on CNN(Convolutional Neural Network) has made breakthrough advances in many tasks. Deep Learning often encounters over-fitting problems although it has shown outstanding performance in many aspects. The problems of over-fitting are mainly caused by three reasons: complex models, data noise, and limited training data. Most neural networks require a large number of parameters and sample data for training to

---

<sup>\*</sup> Supported by organizationthe Key R & D projects in Shanxi Province(No. 2021GY-102); National Natural Science Foundation of China(No. 6217020827), the Key R & D projects in Xi'an(21RGZN0012), the Key R & D projects in Xianyang(2021ZDYF-GY-0031).

make these parameters work correctly. The data set is the basis of training deep learning models. The quantity, diversity, quality, and imbalance of the training data significantly affect the robustness and generalization of the deep learning model [1]. Data sources mainly include self-established datasets and existing public datasets. Creating a data set with enough samples is often a tough and time-consuming task. Particularly, some images are hard to obtain, such as agricultural and forestry pests and diseases, or medical images. Therefore, data augmentation is an effective way to artificially increase the training data, when the number of samples is insufficient or imbalanced.

Pine is one of the important tree species in China's forest resources, accounting for 70% of the artificial forest area [2]. However, in recent years, a large number of pine resources have been attacked by the invasive pest pine wilt disease, which has seriously endangered the utilization and sustainable development of forest resources. The disease mainly affects pine tree species and has two main routes of transmission: natural transmission (*Monochamus alternatus*) and human transmission [3, 4]. Due to the difference in chlorophyll and water content between healthy pine trees and diseased pine trees, resin secretion, growth vigor, and needle color are important indicators for judging pine wilt disease.

Pine trees infected with pine wilt disease generally die especially in summer and autumn. The disease goes through three distinct stages of infection. In terms of the color of needles, dark green needles of healthy pine trees become yellow in the mildly damaged stage, yellow-brown in the moderately damaged stage, and red-brown in the severely damaged stage [5–7]. A pine tree with pine wilt disease can be identified by checking whether the pine tree has reddish brown or yellowish brown needles or the whole pine tree has withered, withered and drooped needles but does not fall off. Due to the insufficiency of available UAV images, the differences in background, illumination, and target size, as well as the differences in the number of samples at different stages of infection, lead to the imbalance in samples size. Due to the insufficiency of available UAV images, the differences in background, illumination, and target size, as well as the differences in the number of samples at different stages of infection, lead to the imbalance in samples size.

Data augmentation is an essential method to improve the training effect of CNN. There are a series of augmentation ways, including rotation, flip, scaling, translation, adding noise, and so on. Through the enhancement of training data, the network with stronger generalization ability can better adapt to the application scene. Data enhancement has been widely used in deep learning of plant pest images. For example, Deng [8] et al enhanced the diseased pine tree samples through rotation, flipping, and comparison, and the AP(Average Precision) value of the model increased from 72.5% to 80.2%. Qin [9] et al adopted multi-scale segmentation to improve the number and diversity of samples. However, few researchers pay attention to which enhancement strategy can effectively improve the accuracy of the model. Moreover, too many enhancement samples may lead to inadequate fitting. Therefore, it is necessary to select the optimal number of training samples for the enhanced data set.

This paper studies the impact of different data enhancement methods on the accuracy of the Faster R-CNN and YOLO models. By testing the influence of different training samples size on model results, the optimal training data set size is found.

## 2 Augmentations

Traditional data enhancement strategies are based on camera models and imaging principles. By using different enhancement methods, it can simulate the effects of the lens, focal length, and aperture on pictures in real scenes. For example, a fisheye lens will cause image deformation, but can obtain a wider scene in a smaller scene; Focal length will blur the background and affect the resolution of the picture; The aperture can affect the brightness and chroma of images. The fisheye effect can simulate the image shot by a fisheye lens. Scaling can simulate different focal length sizes and different shooting distances; Color transformation can simulate images under different aperture sizes and lighting; Translation and rotation can simulate different perspectives; Gaussian noise and salt-pepper noise can simulate the noise caused by camera sensors. The eight data enhancements used in this paper all rely on the UAV camera model, imaging principles, and changes in the external environment. Each augmentation technique is as follows(see Fig. 1) [10, 11]:

Flips: flip horizontally and vertically for each image in the training set.

Rotation: rotate each image in the training set randomly clockwise at the angle of from  $0^\circ$  to  $360^\circ$ ;

Scaling: the training set image is linearly enlarged or reduced. The part of the enlarged image that exceeds the original image size needs to be cropped, and the part between the reduced image boundary and the original image is filled with a black background.

Color transform: modify image brightness, contrast, and saturation.

Nonlinear scaling: the training set images are enlarged or reduced according to different aspect ratios. That is, the aspect ratio of the scaling image is different from the original image.

Translation: each image in the training set is shifted 50 pixels along the horizontal axis and filled in with a black background.

Adding noise: noise refers to the random superposition of some isolated pixels or pixel blocks on the original image which can cause strong visual effects, to disrupt the observable information of the image. Gaussian noise and salt & pepper noise are respectively added to the images of the training set, and salt & pepper noise is distributed in each image as random white or black pixels.

Fisheye effect: simulate the fisheye lens effect on the training set.

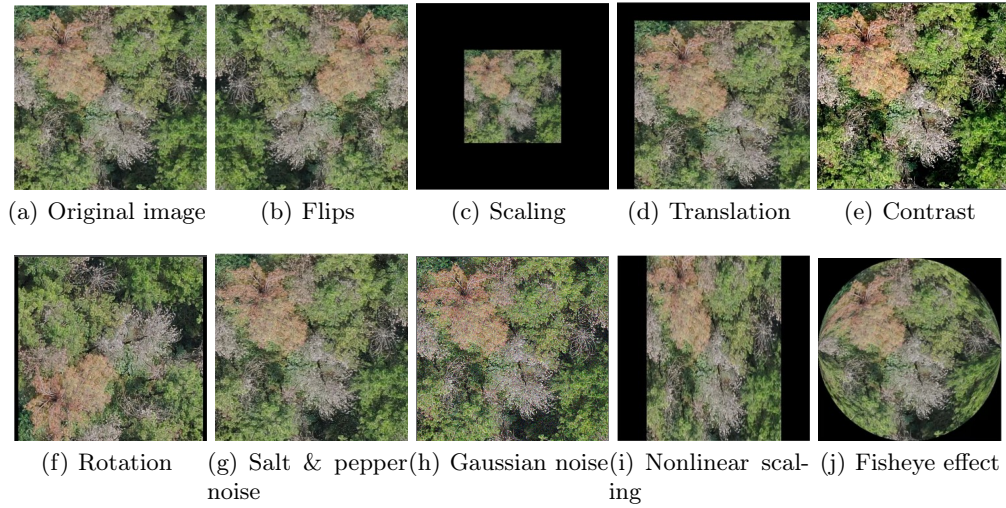


Fig. 1: Eight data augmentation methods

### 3 Materials and Methods

The UAV remote sensing image was collected by a visible-light camera mounted on Multi-rotor light UAV. The field plot is a pine wilt disease endemic area located in Xianhe Town, Hanbin District, Ankang City, Shaanxi Province, and its area is 0.0775 square kilometers. The dead trees of pine wilt disease are mainly divided into three types: yellow-brown needles moderately dead trees, red-brown needles severely dead trees, and white needles dead trees. The size of pine wilt disease images is 4000\*3000. The full-sized images are cropped to 150 sub-images about 500\*500 pixels. After cropping, sub-images are filtered to be used as the data set. We select 85 sub-images as the original data set for each augmentation technique. Among these samples, 60% are used for the training set, 20% for the validation set, and 20% for the testing set. Before training the models, we use eight augmentation strategies to generate eight new training sets. Each new training set consists of the original training sub-images and the new training sub-images enhanced by one of the above techniques. The annotation file of each sub-image is also transformed synchronously.

With the development of deep learning, object detection based on deep learning has been widely used because of its excellent performance. Faster R-CNN and YOLOv3 are the most classical algorithms in the field of object detection. As a classical two-stage model, Faster R-CNN mainly generates a series of candidate regions through the method of region proposal and then performs classification and position regression on extracted features of candidate regions, which is better in detection accuracy. YOLOv3 is a classic one-stage model based on regression, which directly extracts features from the network to predict the classification

and position of objects. And its detection speed is fast. Different target sizes and the unbalanced number of pine wilt disease samples at different stages will lead to distortion or over-fitting in the identification process. In addition, the Faster R-CNN and YOLOv3 models are easy to be ignored for small targets due to their low resolution and few features when detecting infected pine images. Therefore, more training samples are needed to learn data features and achieve higher accuracy. Data augmentation can make the model learn more invariance features through geometric transformation, thus improving the robustness of the models.

With the increase of the training data set, the detection performance of the model is improved gradually. In addition, the model may also go from over-fitting to under-fitting when increased to a certain extent. Therefore, To further analyze the impact of image data set size on the model and find the inflection point of different data-set sizes on model performance [12], we select 150 samples divided into a training set, validation set, and test set according to 6:2:2. The new data set is made with original training data and augmented data generated by the above augmentation techniques, which reach 1183 data. We randomly select 182, 364, 546, 728, 910, and 1183 data among them to form a dataset and train the Faster R-CNN and YOLOV3 models to obtain the corresponding AP curve. AP is the metric to measure the performance of object detectors like Faster R-CNN and YOLOV3. It is the average of the maximum precision at different recall values.

Before evaluating the model performance, for each augmentation method, we used the cross entropy(CE) function to compare the similarity between each image of original training set, D and corresponding image of post-augmented training set, D'. CE is a measure of the difference between two images. In this paper, the cross entropy mean value between D and D 'is used to quantify the augmentation

$$CE(p, q) = \sum p(i) \times \ln \frac{p(i)}{q(i)} + \sum q(i) \times \ln \frac{q(i)}{p(i)} \quad (1)$$

## 4 Results

The experimental operating platform is Ubuntu 20.04, Intel(R) Core(TM) i9-10900K CPU @ 3.70GHz, NVIDIA Corporation Device 2208, based on Pytorch as a deep learning framework. We use eight augmented training sets to train Faster R-CNN and YOLOV3 networks. To evaluate the impact of different data augmentation methods and different numbers of training data sets on the identification of pine wilt diseased trees, the index AP is used to evaluate model performance. The increased range is used to compare AP values before and after augmentation. The mean CE between the each augmented training set and the original data set is used to analyzed the effect of each augmentation on the training result. It can be seen from Table 1 that the scaling augmentation method has the highest AP value and CE value for the two models. a higher cross

entropy means that the image augmented by scale have the more abundant and varied images information, which enhancing the diversity of the samples.

Table 1: Influence of traditional image data augmentation techniques on Faster R-CNN and YOLOV3 models.

Augmentation type	CE	Faster R-CNN		YOLOV3	
		AP	Increase range	AP	Increase range
Original image		0.573	0	0.380	0
Saturation	0.0005	0.576	0.003	0.391	0.011
Flip	0.0006	0.721	0.148	0.504	0.124
Gaussian noise	0.0222	0.668	0.095	0.386	0.006
Salt & pepper noise	0.0395	0.720	0.147	0.416	0.036
Rotation	0.1615	0.630	0.057	0.498	0.118
Translation	0.1947	0.648	0.075	0.426	0.046
Contrast	0.2011	0.688	0.115	0.487	0.107
Brightness	0.2244	0.786	0.213	0.469	0.089
Fisheye effect	0.2444	0.777	0.204	0.410	0.03
Nonlinear scaling	0.3663	0.798	0.225	0.423	0.043
Scaling	0.6971	0.894	0.321	0.642	0.262

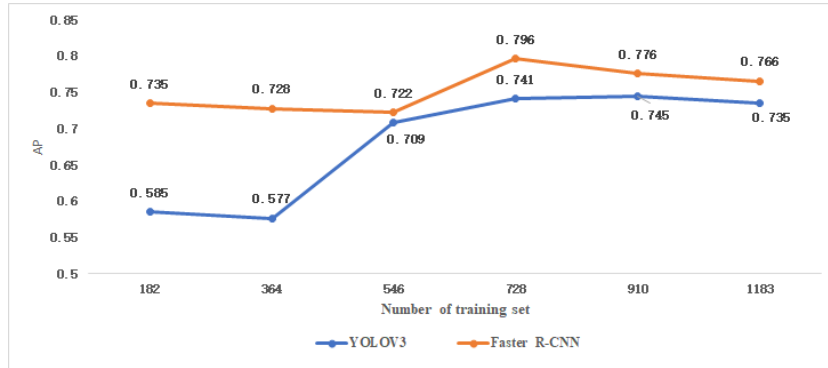


Fig. 2: Impacts of the growing number of datasets on the performance of YOLO3 and Faster R-CNN.

As can be seen from (see Fig. 2), the size of the training data set has a great influence on the detection capability of the models. With the increase in training set size, the detection performance of the model is improved gradually. However, when the data set grows to 60 percent of the new training samples, AP arrives at the maximum value and then begins to decline for the Faster R-CNN network. And for the YOLOV3 network, AP arrives at the maximum value when the data set increases to 75 percent of the new training samples.

## 5 Discussion and Conclusion

Traditional data augmentation techniques rely on rotation, translation, and other methods to fine-tune the existing data set and generate a large number of data



sets slightly different from the original data set for model training. This technique can not only increase the number of training samples but also improve the robustness and the generalization ability of the model. In this paper, the optimal augmentation techniques and the optimal number of training samples for the identification of pine wilt disease were studied.

Looking at the results from Table 1, We can see that the cross entropy is proportional to the AP value of the Faster R-CNN model, in addition to Gaussian noise, Salt & pepper noise and flip. In other word, the new training set that has cross entropy about 0.6971-0.2444 (Scaling, Nonlinear scaling, Fisheye effect, Brightness) has AP value of 0.894-0.786. On the contrary, the cross entropy in 0.2011-0.1615 (Contrast, Translation, Rotation) corresponds to AP in 0.688-0.630. we notice that the Salt & pepper noise, Gaussian noise and flip has the low cross entropy about 0.0395, 0.0222 and 0.0006, but has the relatively high AP value about 0.720, 0.668 and 0.721. The same augmentation method has different gains for different networks, so it is necessary to find suitable augmentation strategies according to different network models.

Although the data augmentation technique solves the problem of insufficient samples in the data set, poor data quality and insufficient sample diversity tend to cause overfitting, which affects the classification performance of the deep learning algorithm. In recent years, with the rapid development and wide application of deep learning, data augmentation through Generative Adversarial Networks (GAN) [13] has become a common method at the present stage. It is a generative model proposed by Goodfellow et al in 2014 and consisted of a generative network and a discriminant network. GAN can increase the number of training set samples, reduce overfitting and improve the recognition effect [14–16]. Jain et al [17]used a Conditional Generative Adversarial Network (CGAN) to generate synthetic images of tomato leaves, to improve the detection accuracy of tomato diseases. However, it needs a large amount of training data as support, and when the model takes a long time to train, it may appear instability and other problems. So the next step, we will compare the traditional data augmentation method with the GAN augmentation method, and select the optimal augmentation technology and the number of training samples.

The quality of the image recognition algorithm is related to the quality and quantity of the dataset. However, due to the limitation of time cost and financial cost, the image data set obtained is insufficient, its quality is poor, and the sample distribution is unbalanced. This makes the task of image recognition difficult. To solve the above problems and reduce the over-fitting probability, this paper adopts image data augmentation technology to increase the number of training samples of the model. The main work of this paper is to introduce the existing traditional image data augmentation techniques and find the optimal augmentation method and the optimal data set size suitable for model training, which provides research ideas for researchers to optimize the dataset and improve the accuracy of the model by using the corresponding data augmentation method.

## References

1. Feng, X., Shen, Y., Wang, D.: Overview of the development of image-based data enhancement methods. *Computer Science and Application* **11**, 370 (2021)
2. Fang, X.: Research on the application practice of unmanned aerial vehicle monitoring pine wilt disease. *Economic and Technical Cooperation Information* (5), 1 (2021)
3. Jiang, M., Huang, B., Yu, X., Zheng, W., Jin, Y., Liao, M., Ni, J.: The distribution, harm, and control countermeasures of pine wilt disease. *Zhejiang Forestry Science and Technology* **38**(6), 9 (2018)
4. Hunt, D.: Pine wilt disease: a worldwide threat to forest ecosystems. *Nematology* **11**(2), 315–316 (2009)
5. Zhang, S., Tan, J., Tang, X., Wang, Y., Huang, J., Song, Q., Min, J.: Study on spectral characteristics and estimation model of masson pine damaged by pine wood nematode. *Spectroscopy and Spectral Analysis* **39**(3), 8 (2019)
6. Kim, S.R., Lee, W.K., Lim, C.H., Kim, M., Kafatos, M.C., Lee, S.H., Lee, S.S.: Hyperspectral analysis of pine wilt disease to determine an optimal detection index. *Forests* **9**(3), 115 (2018)
7. Vollenweider, P., Günthardt-Goerg, M.S.: Diagnosis of abiotic and biotic stress factors using the visible symptoms in foliage. *Environmental Pollution* **140**(3), 562–571 (2006)
8. Deng, X., Tong, Z., Lan, Y., Huang, Z.: Detection and location of dead trees with pine wilt disease based on deep learning and uav remote sensing. *AgriEngineering* **2**(2), 294–307 (2020)
9. Qin, J., Wang, B., Wu, Y., Lu, Q., Zhu, H.: Identifying pine wood nematode disease using uav images and deep learning algorithms. *Remote Sensing* **13**(2), 162 (2021)
10. Hussain, Z., Gimenez, F., Yi, D., Rubin, D.: Differential data augmentation techniques for medical imaging classification tasks. In: *AMIA annual symposium proceedings*. vol. 2017, p. 979. American Medical Informatics Association (2017)
11. Shorten, C., Khoshgoftaar, T.M.: A survey on image data augmentation for deep learning. *Journal of big data* **6**(1), 1–48 (2019)
12. Luo, Z., Yu, H., Zhang, Y.: Pine cone detection using boundary equilibrium generative adversarial networks and improved yolov3 model. *Sensors* **20**(16), 4430 (2020)
13. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. *Advances in neural information processing systems* **27** (2014)
14. Zhao, Y., Chen, Z., Gao, X., Song, W., Xiong, Q., Hu, J., Zhang, Z.: Plant disease detection using generated leaves based on doublegan. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* (2021)
15. Cao, K., Wu, F., Qian, X., Yang, Z.: Gan-based uav aerial image reconstruction. *Electronic technology* **32**(8), 5 (2019)
16. Hu, G., Zhu, Y., Wan, M., Bao, W., Zhang, Y., Liang, D., Yin, C.: Detection of diseased pine trees in unmanned aerial vehicle images by using deep convolutional neural networks. *Geocarto International* pp. 1–20 (2020)
17. Jain, S., Gour, M.: Tomato plant disease detection using transfer learning with c-gan synthetic images. *Computers and Electronics in Agriculture* **187**(2021) (2021)