Tech Science Press

Multi-Scale Image Segmentation Model for Fine-Grained Recognition of Zanthoxylum Rust

Fan Yang[1], Jie Xu[1,*], Haoliang Wei[1], Meng Ye[2], Mingzhu Xu[1], Qiuru Fu[1], Lingfei Ren[3] and Zhengwen Huang[4]

[1]School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, 611731, China
[2]College of Forestry, Sichuan Agricultural University, Chengdu, 611130, China
[3]Sichuan Zizhou Agricultural Science and Technology Co. Ltd., Mianyang, 621100, China
[4]Department of Electronic and Computer Engineering, Brunel University, Uxbridge, Middlesex, UB8 3PH, United Kingdom
*Corresponding Author: Jie Xu. Email: xuj@uestc.edu.cn
Received: 19 August 2021; Accepted: 09 October 2021

Abstract: Zanthoxylum bungeanum Maxim, generally called prickly ash, is widely grown in China. Zanthoxylum rust is the main disease affecting the growth and quality of Zanthoxylum. Traditional method for recognizing the degree of infection of Zanthoxylum rust mainly rely on manual experience. Due to the complex colors and shapes of rust areas, the accuracy of manual recognition is low and difficult to be quantified. In recent years, the application of artificial intelligence technology in the agricultural field has gradually increased. In this paper, based on the DeepLabV2 model, we proposed a Zanthoxylum rust image segmentation model based on the FASPP module and enhanced features of rust areas. This paper constructed a fine-grained Zanthoxylum rust image dataset. In this dataset, the Zanthoxylum rust image was segmented and labeled according to leaves, spore piles, and brown lesions. The experimental results showed that the Zanthoxylum rust image segmentation method proposed in this paper was effective. The segmentation accuracy rates of leaves, spore piles and brown lesions reached 99.66%, 85.16% and 82.47% respectively. MPA reached 91.80%, and MIoU reached 84.99%. At the same time, the proposed image segmentation model also had good efficiency, which can process 22 images per minute. This article provides an intelligent method for efficiently and accurately recognizing the degree of infection of Zanthoxylum rust.

Keywords: Zanthoxylum rust; image segmentation; deep learning

1 Introduction

Zanthoxylum bungeanum Maxim, generally called prickly ash, is a plant of the Rutaceae family, belonging to a deciduous shrub or small tree. Furthermore, Zanthoxylum plays an important role

in diet, medicine, etc. [1,2]. With the rapid development of the Zanthoxylum industry, the planting area of Zanthoxylum has been expanding. At the same time, the scope of pests and diseases of Zanthoxylum has gradually been broadening. The yield reduction rate of Zanthoxylum is more than 50%. This situation not only affects the yield of Zanthoxylum, but also the quality of Zanthoxylum. Zanthoxylum rust has the largest loss and the most widespread distribution among these pests and diseases [3,4]. Zanthoxylum may be infected with Zanthoxylum rust in various growth stages. In terms of rust prevention and control, the infected area is a key indicator for judging the degree of rust infection. At present, the degree mainly relies on manual judgment and analysis. This method has two disadvantages as follows. First, manual analysis is difficult to quantify the area and severity of disease infection, resulting in poor accuracy. Second, the high dependence of manual judgment on professional knowledge leads to poor timeliness. Therefore, we need to study intelligent recognition methods. When rust occurs, leaves produce spore piles of different shades of yellow. Because the color and boundary shape of spore piles have obvious characteristics, image segmentation technology can be used to divide the area of rust.

In the field of computer vision [5–10]. the main basis of traditional image segmentation methods for area segmentation methods is regional texture, color, shape, etc. [11–16]. In this paper, the image segmentation of Zanthoxylum rust is diverse in morphology and has a large difference in lesion size. Traditional image segmentation methods are difficult to achieve accurate results. In recent years, the rapid development of artificial intelligence technology has provided new methods for image segmentation. Image segmentation methods based on deep learning have also received a lot of attention [17]. In 2015, Long et al. [18] built a fully convolutional network (FCN) for image semantic segmentation, which completed the pixel-level image segmentation task. FCN used deconvolution technology [19] for up-sampling, recovered each pixel from abstract features, and realized the leap from image-level to pixel-level classification. However, FCN does not consider the global context information, and the efficiency is not very high. Later, Ronneberger et al. [20] proposed an image segmentation model U-Net network based on CNN (Convolutional Neural Network). The model has a small number of parameters and fast network speed. Furthermore, the model can be trained based on a small amount of data and is widely used in medical image segmentation. He et al. [21] proposed the Mask R-CNN network, which is mainly based on the Faster-RCNN network [22] to expand, adding a branch to use existing detection to perform parallel segmentation and prediction of the object [23].

In conclusion, the application of Zanthoxylum rust image segmentation technology based on deep learning can extract feature information more comprehensively and increase the accuracy and efficiency of segmentation. However, the regional identification of Zanthoxylum rust still has some problems. First, the lesion area at the early stage is small. Second, the size of the infected area varies in different periods. Third, different shapes of lesions have irregularly connected blocks. In response to these problems, this paper studies the segmentation and recognition of the Zanthoxylum rust area based on deep learning [24–29]. Our main contributions are as follows: 1) we construct a fine-grained Zanthoxylum rust image dataset. In this dataset, the Zanthoxylum rust image is segmented and labeled according to leaves, spore piles, and brown lesions; 2) we add a parallel structure to four parallel structures of the ASPP module and propose a novel FASPP module with five parallels; 3) considering the influence of leaf parts on the recognition accuracy of raised spore piles and brown diseased areas, we create the residual images for enhancing the rust regional features. Section 3 introduces the fine-grained segmentation dataset of Zanthoxylum rust. Section 4 and Section 5 proposes two image segmentation methods based on the FASPP module and enhanced rust regional features separately. The conclusion and perspectives of this paper are given in Section 6.

2 Related Work

The proposal of the deep CNN network has greatly promoted the development of computer vision [30–34]. The DeepLabV2 model is an image semantic segmentation network proposed by Chen et al. [35]. First, this structure increases the feature map size while expanding the receptive field without adding parameters [36]. Besides, DeepLabV2 uses a higher sampling density to calculate the feature map and then uses bilinear interpolation to restore the feature to the original image size. Second, DeepLabV2 refers to the idea of SPP (Spatial Pyramid Pooling) and proposes a similar structure ASPP (Atrous Spatial Pyramid Pooling) module [37,38]. Each branch of this module uses atrous convolutions with different sampling rates to extract image context information, which increases the receptive field and enables more efficient classification. Third, DeepLabV2 uses CRF (fully-connected Conditional Random Field) to enhance the model's ability to capture details, ensuring the boundary position is split accurately [39,40]. The structure of DeepLabV2 is shown in Fig. 1.
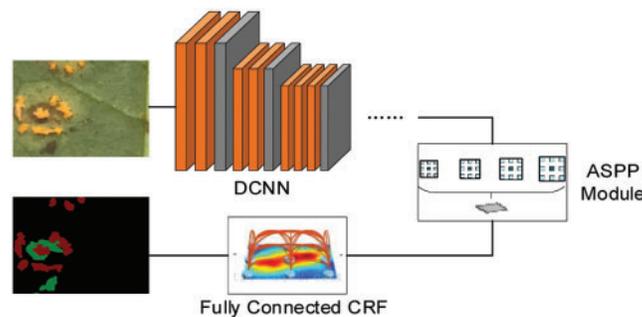
Figure 1: DeepLabV2 structure diagram

Considering the amount of calculation and storage space, the ASPP module is used in the DeepLabV2 network structure for better segmentation at multiple scales. This measure refers to the idea of the SPP module in SPPNet [41]. DeepLabV2 implements the ASPP module, a variant of the SPP module. The ASPP module is similar to the spatial pyramid structure [42]. It uses multiple sampling rates of convolutions to extract features in parallel and then merges the features. The ASPP module parallels atrous convolutions of different sampling rates to sample on a given input, which is equivalent to capturing image context at multiple scales. This method can increase the receptive field and perform segmentation tasks more efficiently. This paper proposed the FASPP (Five-branch Atrous Spatial Pyramid Pooling) module based on the ASPP module, the details will be introduced in Section 4.

3 Dataset

Zanthoxylum rust is manifested in light yellow, orange, or gray raised spore piles and brownish lesions. In the early stage of rust disease, small yellow spots and spore piles appear on the leaf surface. The shape of the spots is close to a circle. As the degree of rust deepens, lesions will expand, and the color will change to brownish. To better judge the extent and stage of rust and to more accurately calculate the area size of brownish lesions and yellowish convex spore piles, the dataset marked yellow convex spore pile areas and brownish diseased areas with different color.

Labelme software [43] is an image semantic segmentation and annotation tool. In this paper, Labelme software is used to label healthy Zanthoxylum leaves and Zanthoxylum rust leaves. Import

images that need to be labeled into Labelme software and mark the segmentation label by drawing a polygon. There are three labels included in our dataset: "Rust1", "Rust2", and "background" [44].

The part labeled "Rust1" is raised spore pile parts of the leaf. The color of raised spore piles is generally yellow or gray. Raised spore piles are shown in Fig. 2a. Marked raised spore piles are shown in Fig. 2b.
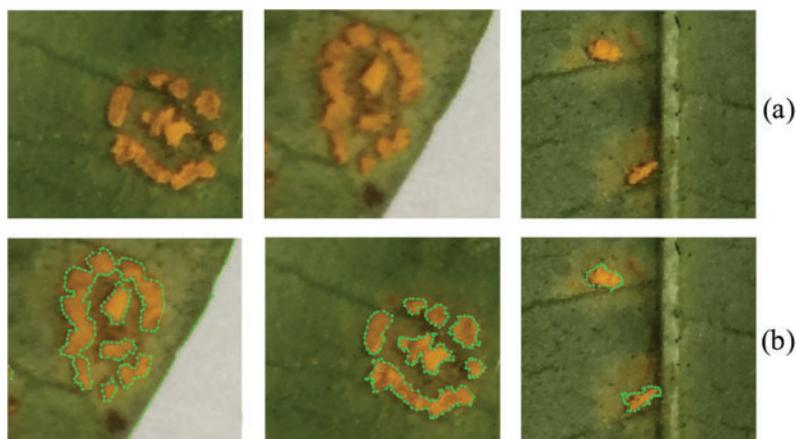
**Figure 2:** Part of original images of raised spore piles and annotated images

The part labeled "Rust2" is brownish lesions of the leaf, shown in Fig. 3a. The brown diseased leaves marked at the edge of brownish lesions are shown in Fig. 3b.
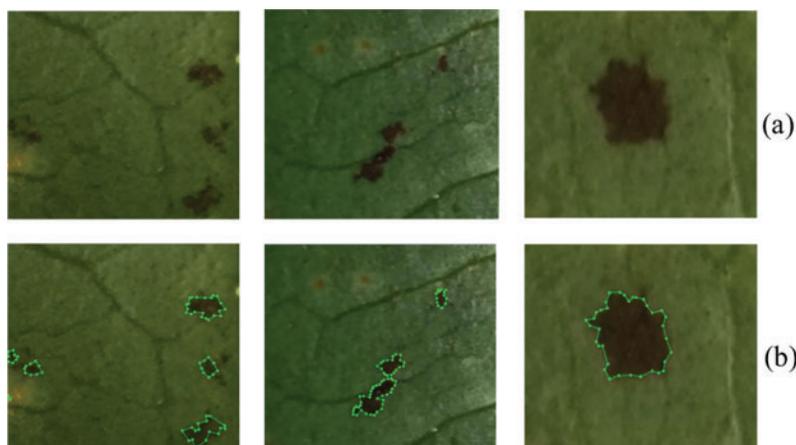
Figure 3: Original images of brown lesions and annotated images

The part labeled "background" was the background part around the leaf, and was marked in black. After marking the background part, raised spore piles and brown lesions were marked, the annotated images are shown in Fig. 4.

The label images are 8 bit images. In the images, the RGB value (0, 0, 0) corresponds to the background part "background" around the leaf, the RGB value (128, 0, 0) corresponds to the raised spot part "Rust1", the RGB value (0, 128, 0) corresponds to the brownish lesions part "Rust2", and the RGB value (128, 128, 0) corresponds to the leaf part.
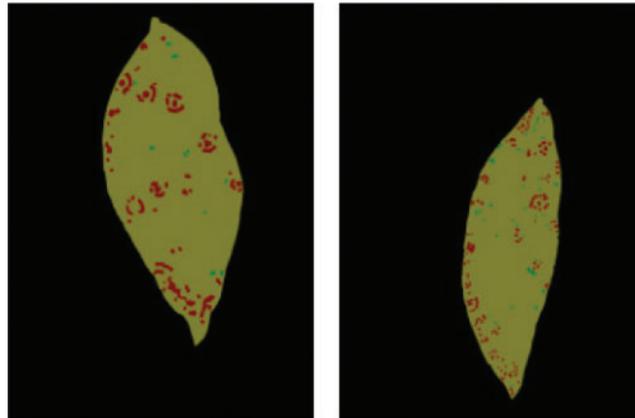
Figure 4: Images of partially completed label

This paper selected 100 leaves for labeling, including leaves with few raised spore piles and brownish lesions, leaves with many raised spore piles and brownish lesions, and leaves with different shades. This method ensured the integrity and completeness of the data. Fig. 5 shows examples of the Zanthoxylum rust dataset.

Figure 5: Examples of the dataset images

Since most of the selected 100 leaves have many raised spore piles and brownish lesions, which are more complicated, directly inputting the deep neural network for training will increase training difficulty. This paper cropped 100 leaves to the size of 480×360 and the bit depth of 24 bits. After cropping, the label value and bit depth of the image change. The cropped images cannot meet the requirements of dataset and cannot be input into the neural network for training. Therefore, in this section, the cropped images were grayed processing and the raised spore piles, brown diseased spots, leaf parts, and background were relabeled. The leaf part was marked as 3, the brownish diseased part was marked as 2, the raised spore piles was marked as 1, and the background was marked as 0. After

processing the cropped pictures, 5753 pictures were obtained. The partially cropped label images are shown in Fig. 6.
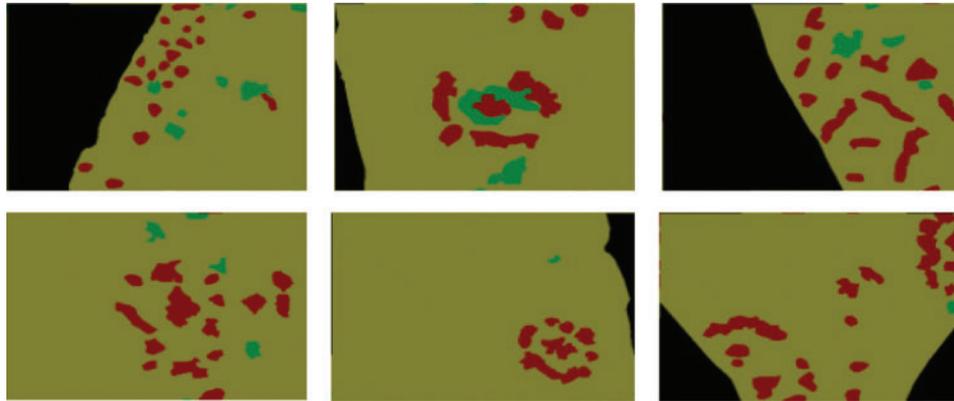

Figure 6: Partially cropped label images

## 4 Image Segmentation Method Based on FASPP Module

4.1 FASPP Module

The ASPP module in DeepLabV2 uses atrous convolutions with different sampling rates. Using convolution kernels with different sampling rates is beneficial to obtain multi-scale information of Zanthoxylum rust and reduces the loss of detailed features of rust leaves. The four parallel sampling rates of the ASPP module in DeepLabV2 are all large, but the information extracted by atrous convolutions with large sampling rates only has a good effect on the segmentation of some large objects. In order to enhance the ability of the network model to recognize small targets, this paper proposed the FASPP (Five-branch Atrous Spatial Pyramid Pooling) module based on the ASPP module.

This paper used parallel atrous convolutions to realize the variant of the ASPP module that used multiple sampling rates to extract features and then merged them. The four parallel sampling rates of the ASPP module are large, respectively [6,12,18,24]. However, the information extracted by atrous convolutions with large sampling rates only was conducive to the segmentation of some large objects, but not small objects. To enhance the ability of the network model to recognize small targets, this paper proposed the FASPP (Five-branch Atrous Spatial Pyramid Pooling) module based on the ASPP module.

In this paper, a variant of the ASPP module was implemented using parallel atrous convolutions, which used branches with multiple sampling rates to extract features and then merged them. This paper added a parallel structure to four parallel structures of the ASPP module and proposed the FASPP module with five parallels. The structure of the FASPP module is shown in Fig. 7.

From Fig. 7, this paper added a branch with a small sampling rate to the parallel structure of four larger sampling rates, fused features extracted from five branches with different sampling rates, and replaced the ASPP module to experiment.
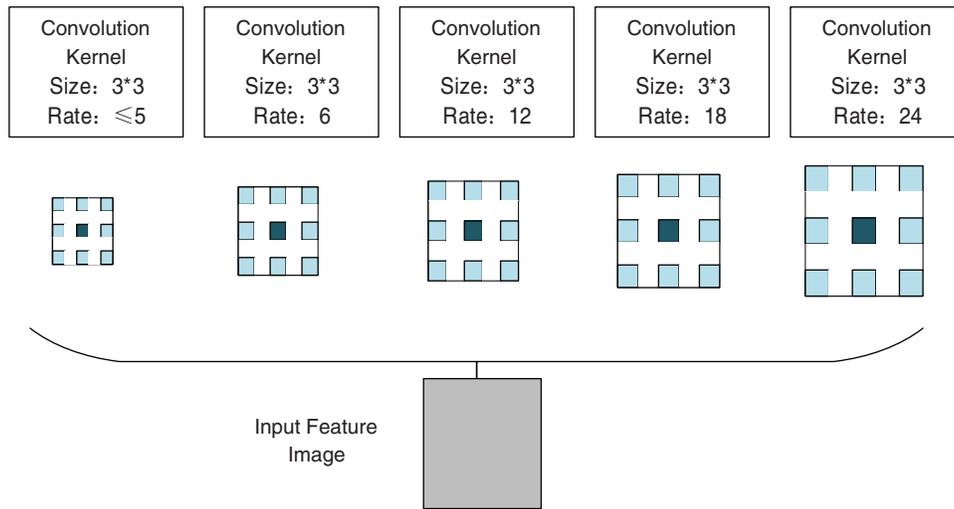
Figure 7: The structure of FASPP module

4.2 Experimental Results and Analysis

In this paper, the experimental environment was Ubuntu16.04. The computer CPU was Intel Core i7-7700. The GPU was NVIDIA GeForce GTX 1080Ti, and the deep learning framework was Caffe. The network model was trained under GPU. The learning rate was 0.001. The learning strategy was "poly". The exponent (power) was 0.9, and the maximum number of iterations (max_iter) was 80,000 times. The curve of learning rate with the number of iterations in the poly change strategy is shown in Fig. 8.
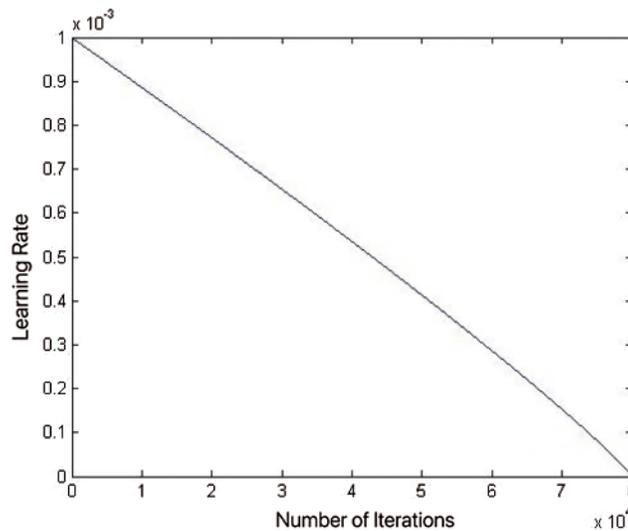
Figure 8: Poly learning rate change curve

This paper used 500 pictures with 480×360 pixel to test trained network model based on the FASPP module. In order to explore the best network segmentation performance at the sampling rate, we compared the performance of the FASPP module at different sampling rates. This paper set the

sampling rate of the new branch of FASPP to 1, 2, 3, 4, and 5 for the experiment. The experimental results are shown in Tab. 1.

**Table 1:** Model performance evaluation under different sampling rate settings

New branch sampling rate setting	1 (%)	2 (%)	3 (%)	4 (%)	5 (%)	Original model (%)
Blade recognition accuracy	99.69	99.68	99.54	99.71	99.66	99.65
Recognition accuracy rate of raised spore piles	81.23	78.79	83.70	80.13	78.26	81.02
Accuracy of brown spots recognition	77.32	76.27	78.23	77.29	78.65	74.27
Mean pixel accuracy (MPA)	89.85	88.67	90.35	89.27	89.10	88.71
Mean intersection over union (MIoU)	82.20	81.29	82.73	81.97	81.54	81.14

From Tab. 1, the recognition accuracy of brownish lesions has been improved when new branches were added. But, when the sampling rate was 2, 4, and 5, the recognition accuracy of raised spore piles decreased. When the sampling rate was 3, the recognition accuracy of raised spore piles and brown lesions increased, and the effect was the best. The average pixel accuracy rate reached 90.35%, and the average intersection ratio reached 82.73%. Compared with recognition results of the original model, the recognition accuracy of raised spore piles was increased by 2.68%, the recognition accuracy of brownish lesion areas was increased by 3.96%, and MIoU was increased by 1.59%. When the sampling rate of the new branch was 3, some network tests results are shown in Fig. 9.

According to the comprehensive pixel accuracy rate, average pixel accuracy rate, average intersection ratio, and network test image results, we concluded that the boundary segmentation of the blade was accurate and the recognition accuracy was 99.54%. This result showed that blade recognition still had a high accuracy rate, which was not reduced as the FASPP module replaced the ASPP module. The FASPP model still distinguished the blade and the background well.

As shown in Fig. 10, the (b) column is the label image after color processing, and the (b) column is the model recognition results based on the FASPP module after color processing. By analyzing the raised spore pile area in these images, we concluded that the large part of raised spore piles had a higher recognition accuracy due to its obvious characteristics, and most of the smaller area also was successfully identified. Compared with experimental results of the original DeepLabV2 model, the

recognition effect on small convex spore piles improved. The recognition accuracy of convex spore piles was 83.70%, which was 2.68% higher than the original model.
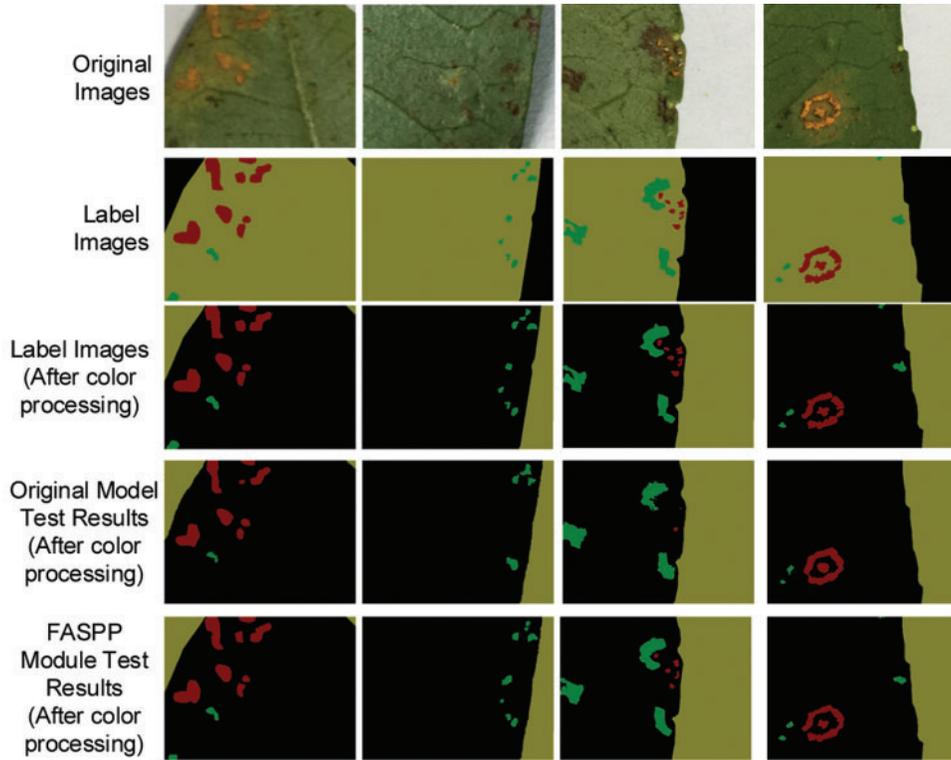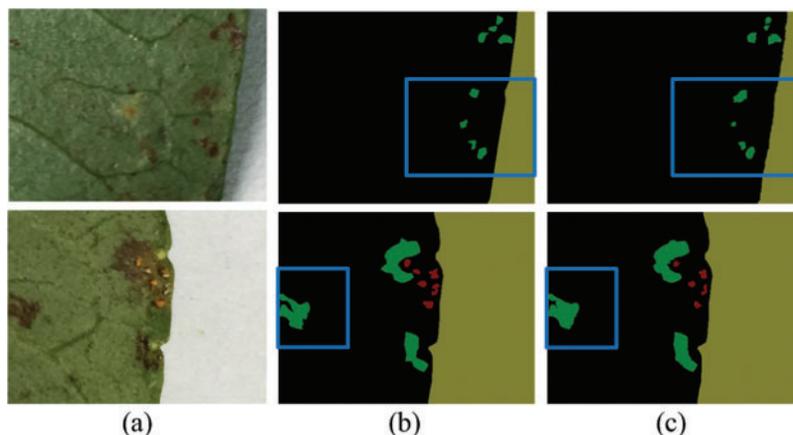

Figure 9: Test results of image segmentation model based on FASPP module
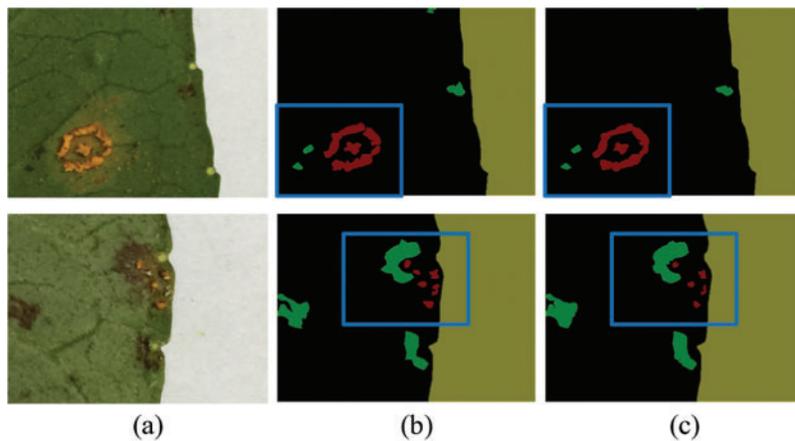
Figure 10: Comparative analysis of test results of raised spore piles

As shown in Fig. 11, the (b) column is the color-processed label image, and the (c) column is the color-processed model recognition result image based on the FASPP module. We can see that brownish lesions with a larger area can still be recognized well, and most of brownish lesions with a smaller area can be successfully identified. Compared with the original DeepLabV2 model, the recognition effect

on smaller brown lesions improved. The recognition accuracy of brown lesions was 78.23%, which increased by 3.96% compared with the original model. Since many of brownish lesions on the leaf were smaller patches, the improvement of the recognition effect of brownish lesions also greatly improved the recognition accuracy, but the accuracy was still lower than that of the convex spore piles.

Figure 11: Comparative analysis of test results of brown spots

The addition of the FASPP module improved the model's ability to recognize small targets of Zanthoxylum rust. Compared with the original DeepLabV2 model, the image segmentation model based on the FASPP module had a better recognition effect on raised spore piles and brown lesions, and improved the model's segmentation accuracy of rust areas.

5 Image Segmentation Method with Enhanced Rust Regional Features

### 5.1 Model with Enhanced Regional Characteristics

Considering the influence of leaf parts on the recognition accuracy of raised spore piles and brown diseased areas, this paper created the residual images. The residual images are obtained by subtracting the leaf surface image from the original image, and only the raised spore pile part, the brown lesion part, and the background image are left. The purpose of making the residual images as a new branch input into the network is to enhance the rust regional features and reduce the influence of the leaf surface on the rust regional recognition.

In practical applications, making residual images by marking is difficult, this paper used the method of pixel processing to process 100 leaves. The method removed pixels with an R-value of 78 to 183, a G-value of 84 to 191, and a B-value of 45 to 118. The image after pixel processing and the image after label are shown in Fig. 12.

In this paper, 100 blades were selected, corresponding to 100 residual images. Also, owing to most of residual images contain more raised spore piles and brownish lesions than original images, and more complicated. Besides, directly inputting the neural network for training will increase training difficulty. This paper cut 100 leaves, same as the original dataset. The cropped parts of the residual image dataset are shown in Fig. 13. The size of the cropped images is 480×360, and the bit depth is 24 bits. After cropping the pictures, 5753 pictures are obtained.

**Figure 12:** Original image and residual image
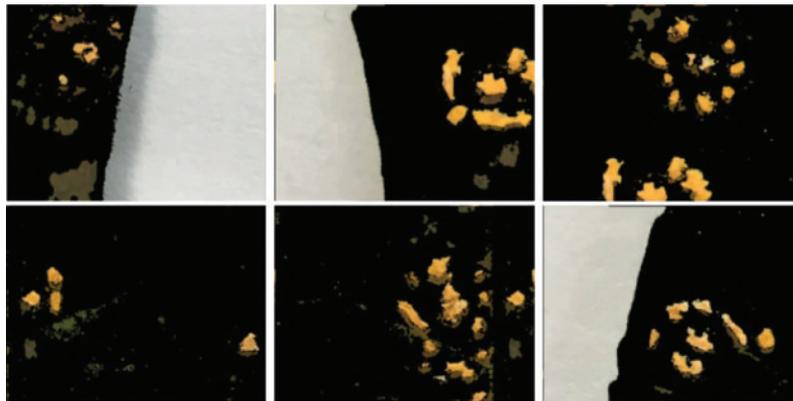
Figure 13: Cropped partial residual images

In the images of leaves infected with rust, the manifestations of rust are more complex, usually in small dots, long strips, and irregularly connected blocks. The regional feature information of images can help the recognition model to distinguish the rust-infected area effectively, so as to perform fine-grained segmentation. Hence, enhancing the regional feature information of rust disease spots can effectively improve the accuracy of Zanthoxylum rust identification. To strengthen the feature information of leaf rust, this paper inputs residual images into the image segmentation model for training. Considering the effect of leaves on rust regional segmentation result, this paper processed the original dataset on the pixel value, and made the residual image dataset. The residual image dataset was inputted as a new branch into the network. As shown in Fig. 14, we use the VGG-16 network structure as the main branch of the image segmentation model with enhanced rust regional features. The other branches of the model adopt the first 6 convolutional layers of the VGG-16 network structure and merge them. The main branch inputs the original image dataset, and another branch inputs the residual image dataset. The FASPP module is used in the network.

5.2 Experimental Results and Analysis

On the trained image segmentation model with enhanced regional features, this experiment used 500 images with a size of 480×360 for testing. The results of some network tests are shown in Fig. 15. The experimental results are shown in Tab. 2. Compared with the model based on the FASPP module,

the recognition accuracy of raised spore piles increased by 1.16%, and the recognition accuracy of brown lesions increased by 3.24%, and the MPA increased by 1.25%, and the MIoU increased 1.63%, indicating that the addition of the residual images branch improved the accuracy of the segmentation and the performance of the model.
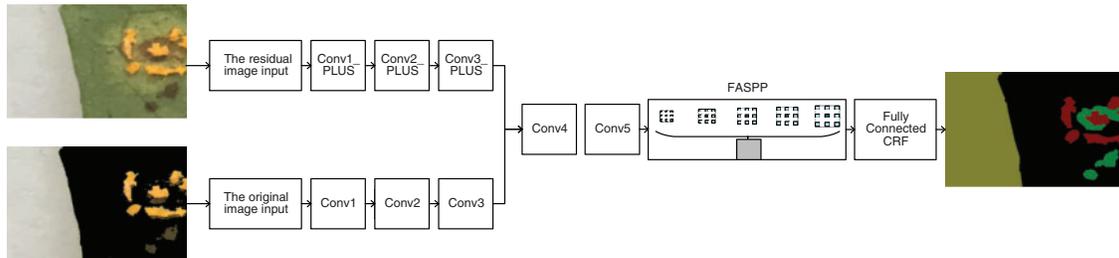
Figure 14: The model structure diagram with enhanced rust regional features
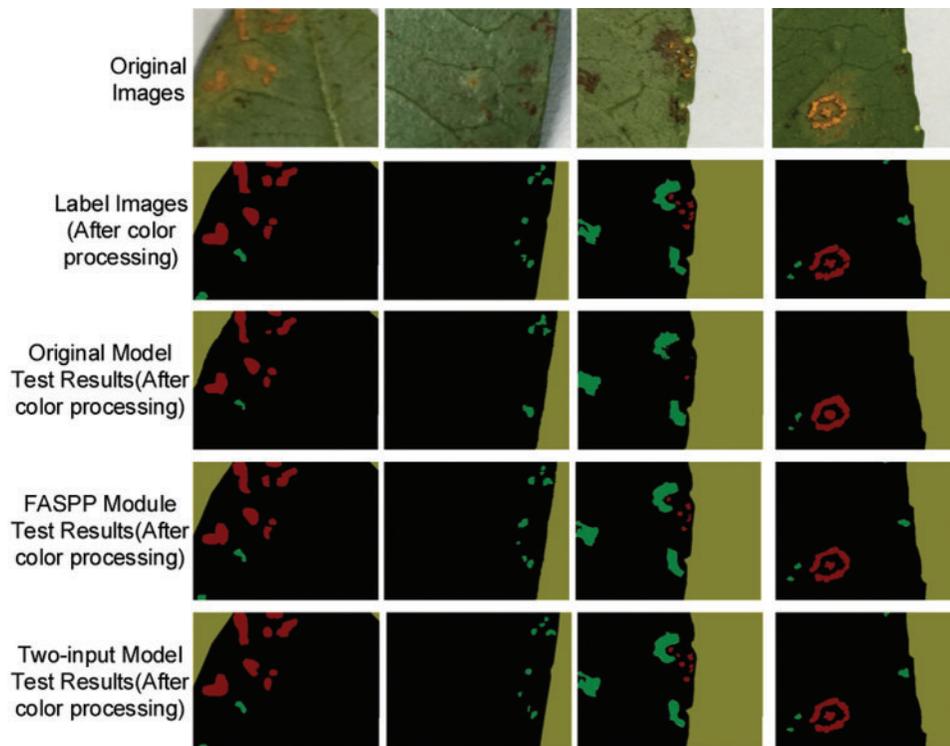
Figure 15: Model test results with enhanced rust regional features

From Fig. 15 and Tab. 2, while the background and leaf parts were quite different, the boundary segmentation of the leaf was still relatively accurate. We concluded that the blade recognition accuracy rate was 99.66%. The results meant that the blade recognition maintained a high accuracy rate, and the model had a high accuracy in segmenting the blade and the background.

As shown in Fig. 16, the (a) column is the label images, the (c) column is the recognition result images based on the FASPP module, and the (d) column is the recognition result images of the enhanced regional features model. Comparing the recognition results of raised spore piles, we

concluded that the larger and smaller raised spore piles were well recognized by the dual-input model. Besides, compared with the FASPP module, the edge part was more accurate, and the recognition effect improved. The recognition accuracy rate was 84.86%, which was 1.16% higher than the FASPP module.

Table 2: Model performance comparison evaluation

Network model	Leaf blade (%)	Raised spore piles (%)	Brown spots (%)	MPA (%)	MIoU (%)
Model based on FASPP module	99.54	83.70	78.23	90.35	82.73
Model with enhanced rust regional features	99.66	84.86	81.57	91.50	84.36
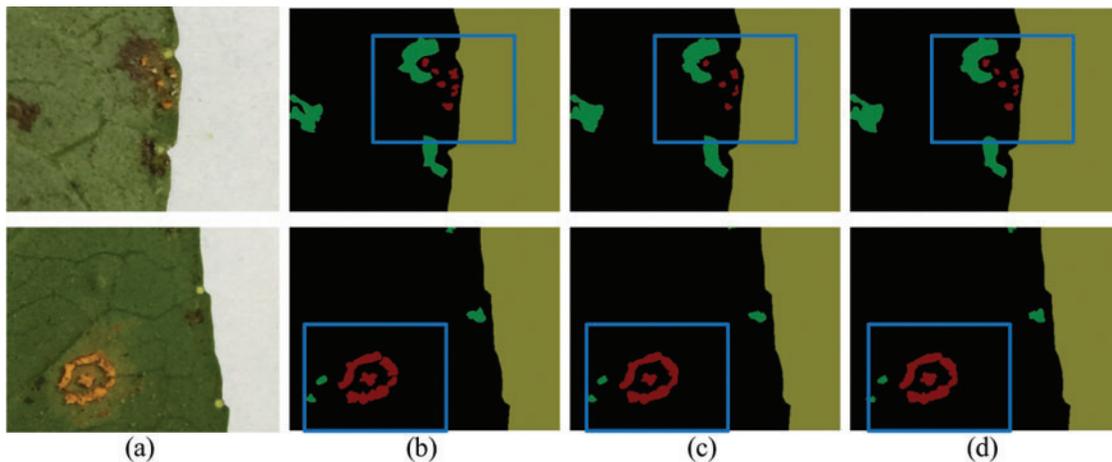
Figure 16: Contrastive analysis of identification and test results of raised spore piles

As shown in Fig. 17, the (b) column is the label images, and the (c) column is the recognition result images based on the FASPP module, and the fourth column is the recognition result images of the dual-input model. By comparing the recognition results of brownish lesion images, we concluded that the larger and smaller brownish lesions were well recognized in the recognition results of the dual input model. Besides, compared with the model based on the FASPP module, the edge of the brown lesion was more accurate, and the recognition effect was better. The recognition accuracy of the brown lesions was 81.57%, which was 3.24% higher than the model based on the FASPP module. However, the recognition accuracy of brownish lesions was still lower than that of raised spore piles. The main reason was there were some tiny brown lesions in the leaves, and this part of the lesions was too small. This part of the dataset was not labeled, so the recognition effect of the brown lesion areas will be affected during training.
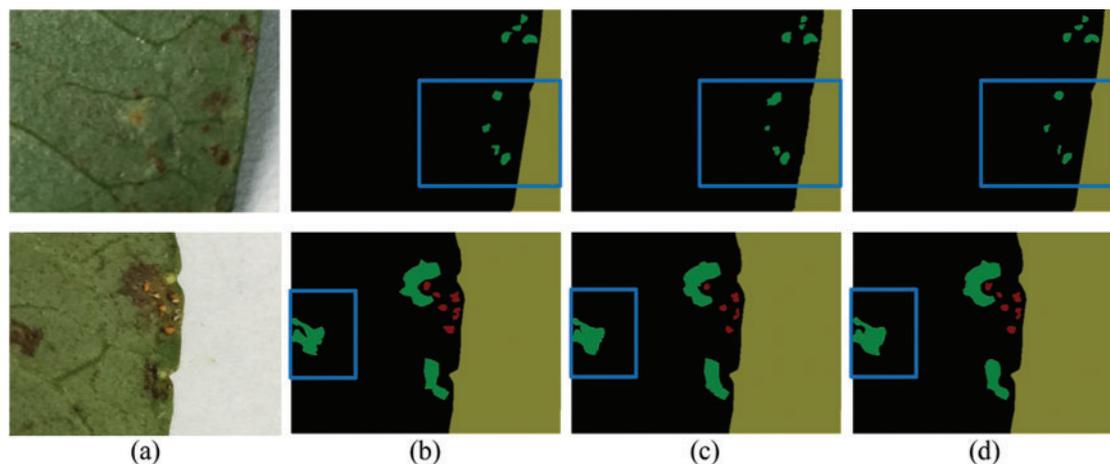
Figure 17: Contrastive analysis of identification and test results of brown spots

5.3 Improved Efficiency of the Model with Enhanced Rust Regional Features

In practical application, we hope that the model can have better efficiency to run on portable devices. In this section, we improved the model based on the dual-input branch of the model with enhanced regional features, reduced the time for the network to test an image without the reduce in accuracy, and improved the test efficiency of the network. The main branch of the network model still uses the VGG-16 network structure, and the other branch used the first several convolutional layers of the VGG-16 network structure.

In Tab. 3, Conv $(1+2+3)$ indicates that the new branch structure is the first seven convolutional layers of the VGG-16 network structure. Among them, Conv $(1+2)$ +Conv3_ $(1+2)$ represents the first six layers, and Conv $(1+2)$ + Conv3_1 represents the first five layers, and Conv $(1+2)$ represents the first four layers, and Conv1+Conv2_1 represents the first three layers, and Conv1 represents the first two layers, and Conv1_1 represents the first one layer.

Table 3: Model performance evaluations under different network structure settings

Network structure setting	Leaf blade (%)	Raised spore piles (%)	Brown spots (%)	MPA (%)	MIoU (%)	Testing time (s)
Conv $(1+2+3)$	99.66	84.86	81.57	91.50	84.36	2.843
Conv $(1+2)$ +Conv3_ $(1+2)$	99.69	84.81	81.96	91.57	84.26	2.804
Conv $(1+2)$ +Conv3_1	99.66	84.78	81.35	91.20	84.08	2.744
Conv $(1+2)$	99.61	85.16	82.47	91.80	84.99	2.692
Conv1+Conv2_1	99.71	81.96	81.47	90.77	83.83	2.622
Conv1	99.70	84.28	78.38	90.57	83.97	2.574
Conv1_1	99.64	84.10	77.93	90.39	83.54	2.519

From the experimental results, Conv $(1 + 2)$ which composed of the first four convolutional layers of the VGG-16 network structure got the best performance. At this time, MPA was 91.8%, and MioU was 85.99%, and the test time was 2.692 s. The model can process 22 pictures per minute, which was faster than the model before.

6 Conclusion

Focusing on the image segmentation technology based on deep learning, we carried out algorithm research to deal with the difficulty to accurately calculate the rust area of Zanthoxylum rust. The corresponding solutions and model architecture were proposed. Because most of Zanthoxylum rust spots are small, this paper proposed an image segmentation model based on the FASPP module from fully extracting the multi-scale information of the image. We added a branch based on the ASPP module and set the atrous rate value of this branch to be smaller. Compared with the results of the original model, the recognition effect of raised spore piles and brown lesions improved. We proved that this module was helpful for the improvement of the small target recognition and model performance. To further improve the segmentation effect, this paper proposed an image segmentation model with enhanced rust regional features. Considering the influences of leaves on the segmentation effect of raised spore piles and brownish lesions, this paper processed the pixel value of the original dataset, and made a residual images dataset, and took the residual image dataset as the input of another branch. The function of the residual image branch is mainly to reduce noise and increase the network model's attention to the area of the rust. The two branches of the original image and the residual images were separately entered into the network and then merged. The experimental results showed that the addition of the residual images branch made the segmentation accuracy of raised spore piles and brown lesions improved. The segmentation accuracy rates of leaves, spore piles and brown lesions reached 99.66%, 85.16% and 82.47% respectively. MPA reached 91.80%, and MIoU reached 84.99%.

Although the model proposed in this paper achieved a certain improvement in the image segmentation effect of Zanthoxylum rust, there are still some problems that need to be completed in the follow-up work. In this paper, the image segmentation model based on the FASPP module improved the segmentation effect on small targets. However, the void ratio of other branches of FASPP is not the most suitable parameter for the segmentation of Zanthoxylum rust due to the particularity of Zanthoxylum rust. Therefore, in the follow-up research work, the dilation rate of other branches of FASPP should be adjusted and tested for training. Find a set of most suitable parameter values, so that the model can better recognize raised spore piles and brownish lesions.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1]   F. Alam, K. M. Din, R. Rasheed, A. Sadiq and A. Khan, "Phytochemical investigation, anti-inflammatory, antipyretic and antinociceptive activities of Zanthoxylum armatum DC extracts-in vivo and in vitro experiments," *Heliyon*, vol. 6, no. 11, pp. e05571, 2020.

[2] H. Ling, L. Yong-an and L. Wen-guang, "Research on segmentation features in visual detection of rust on bridge pier," in *Chinese Control and Decision Conf.*, Hefei, China, pp. 2084–2088, 2020.

[3] A. P. Marcos, N. L. Silva Rodovalho and A. R. Backes, "Coffee leaf rust detection using convolutional neural network," in *XV Workshop de Visão Computacional*, São Bernardo do Campo, Brazil, pp. 38–42, 2019.

[4] A. Yadav and M. K. Dutta, "An automated image processing method for segmentation and quantification of rust disease in maize leaves," in *Int. Conf. on Computational Intelligence & Communication Technology*, Ghaziabad, India, pp. 1–5, 2018.

[5] P. Poply and J. A. Arul Jothi, "Refined image segmentation for calorie estimation of multiple-dish food items," in *Int. Conf. on Computing, Communication, and Intelligent Systems*, Greater Noida, GN, India, pp. 682–687, 2021.

[6] P. R. Kumar, A. Sarkar, S. N. Mohanty and P. P. Kumar, "Segmentation of white blood cells using image segmentation algorithms," in *Int. Conf. on Computing, Communication and Security*, Patna, India, pp. 1–4, 2020.

[7] S. B. Nemade and S. P. Sonavane, "Image segmentation using convolutional neural network for image annotation," in *Int. Conf. on Communication and Electronics Systems*, Coimbatore, India, pp. 838–843, 2019.

[8] A. Feng, Z. Gao, X. Song, K. Ke, T. Xu *et al.,* "Modeling multi-targets sentiment classification via graph convolutional networks and auxiliary relation," *Computers, Materials & Continua*, vol. 64, no. 2, pp. 909–923, 2020.

[9] A. Gumaei, M. Al-Rakhami and H. AlSalman, "Dl-har: Deep learning-based human activity recognition framework for edge computing," *Computers, Materials & Continua*, vol. 65, no. 2, pp. 1033–1057, 2020.

[10] A. Qayyum, I. Ahmad, M. Iftikhar and M. Mazher, "Object detection and fuzzy-based classification using uav data," *Intelligent Automation & Soft Computing*, vol. 26, no. 4, pp. 693–702, 2020.

[11] C. Guada, D. Gómez, J. T. Rodríguez, J. Yáñez and J. Montero, "Classifying image analysis techniques from their output," *International Journal of Computational Intelligence Systems*, vol. 9, no. 1, pp. 43–68, 2016.

[12] M. C. Keuken, B. R. Isaacs, R. Trampel, W. van der Zwaag and B. U. Forstmann, "Visualizing the human subcortex using ultra-high field magnetic resonance imaging," *Brain Topography*, vol. 31, no. 1, pp. 1–33, 2018.

[13] S. Y. Wan and W. E. Higgins, "Symmetric region growing," *IEEE Transactions on Image Processing*, vol. 12, no. 9, pp. 1007–1015, 2003.

[14] A. P. Moore, S. J. D. Prince, J. Warrell, U. Mohammed and G. Jones, "Superpixel lattices," in *IEEE Conf. on Computer Vision and Pattern Recognition*, Anchorage, AK, USA, pp. 1–8, 2008.

[15] R. Pan and G. Taubin, "Automatic segmentation of point clouds from multi-view reconstruction using graph-cut," *Visual Computer*, vol. 32, no. 5, pp. 601–609, 2016.

[16] E. P. Ijjina and C. K. Mohan, "View and illumination invariant object classification based on 3D color histogram using convolutional neural networks," in *Asian Conf. on Computer Vision*, Singapore, Singapore, pp. 316–327, 2014.

[17] S. Alheejawi, M. Mandal, R. Berendt and N. Jha, "Automated melanoma staging in lymph node biopsy image using deep learning," in *Canadian Conf. of Electrical and Computer Engineering*, Edmonton, AB, Canada, pp. 1–4, 2019.

[18] J. Long, E. Shelhamer and T. Darrell, "Fully convolutional networks for semantic segmentation," in *IEEE Conf. on Computer Vision and Pattern Recognition*, Boston, MA, USA, pp. 3431–3440, 2015.

[19] H. Noh, S. Hong and B. Han, "Learning deconvolution network for semantic segmentation," in *IEEE Int. Conf. on Computer Vision*, Santiago, Chile, pp. 1520–1528, 2015.

[20] O. Ronneberger, P. Fischer and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention*, Munich, Germany, pp. 234–241, 2015.

[21] K. He, G. Gkioxari, P. Dollár and R. Girshick, "Mask R-CNN," in *IEEE Int. Conf. on Computer Vision*, Venice, Italy, pp. 2980–2988, 2017.

[22] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.

[23] H. Zhao, J. Shi, X. Qi, X. Wang and J. Jia, "Pyramid scene parsing network," in *IEEE Conf. on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, pp. 6230–6239, 2017.

[24] M. Suganthi and J. G. R. Sathiaseelan, "An exploratory of hybrid techniques on deep learning for image classification," in *Int. Conf. on Computer, Communication and Signal Processing*, Chennai, India, pp. 1–4, 2020.

[25] C. Iorga and V. Neagoe, "A deep CNN approach with transfer learning for image recognition," in *Int. Conf. on Electronics, Computers and Artificial Intelligence*, Pitesti, Romania, pp. 1–6, 2019.

[26] G. Madhupriya, N. M. Guru, S. Praveen and B. Nivetha, "Brain tumor segmentation with deep learning technique," in *Int. Conf. on Trends in Electronics and Informatics*, Tirunelveli, India, pp. 758–763, 2019.

[27] H. Li, C. Pan, Z. Chen, A. Wulamu and A. Yang, "Ore image segmentation method based on u-net and watershed," *Computers, Materials & Continua*, vol. 65, no. 1, pp. 563–578, 2020.

[28] H. Li, W. Zeng, G. Xiao and H. Wang, "The instance-aware automatic image colorization based on deep convolutional neural network," *Intelligent Automation & Soft Computing*, vol. 26, no. 4, pp. 841–846, 2020.

[29] H. Zheng and D. Shi, "A multi-agent system for environmental monitoring using boolean networks and reinforcement learning," *Journal of Cyber Security*, vol. 2, no. 2, pp. 85–96, 2020.

[30] Y. Liu, Y. Qi and Y. Wang, "Object tracking based on deep CNN features together with color features and sparse representation," in *IEEE Symposium Series on Computational Intelligence*, Xiamen, China, pp. 1565–1568, 2019.

[31] J. Xu, W. Wang, H. Y. Wang and J. H. Guo, "Multi-model ensemble with rich spatial information for object detection," *Pattern Recognition*, vol. 99, pp. 107098, 2020.

[32] J. Xu, R. Song, H. L. Wei, J. H. Guo, Y. F. Zhou *et al.,* "A fast human action recognition network based on spatio-temporal features," *Neurocomputing*, vol. 441, pp. 350–358, 2020.

[33] Z. T. Li, W. L. Li, F. Y. Lin, Y. Sun, M. Yang *et al.,* "Hybrid malware detection approach with feedback-directed machine learning," *Science China (Information Sciences)*, vol. 63, no. 3, pp. 139103, 2020.

[34] Z. T. Li, J. Zhang, K. H. Zhang and Z. Y. Li, "Visual tracking with weighted adaptive local sparse appearance model via spatio-temporal context learning," *IEEE Transactions on Image Processing*, vol. 27, no. 9, pp. 4478–4489, 2018.

[35] L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy and A. L. Yuille, "DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 834–848, 2018.

[36] Z. Xu, J. Wang and L. Wang, "Infrared image semantic segmentation based on improved deeplab and residual network," in *Int. Conf. on Modelling, Identification and Control*, Guiyang, China, pp. 1–9, 2018.

[37] Q. Wang, X. He, Y. Lu and S. Lyu, "DeepText: Detecting text from the wild with multi-ASPP-assembled deeplab," in *Int. Conf. on Document Analysis and Recognition*, Sydney, NSW, Australia, pp. 208–213, 2019.

[38] Y. Wu, X. Shen, F. Bu and J. Tian, "Ultrasound image segmentation method for thyroid nodules using ASPP fusion features," *IEEE Access*, vol. 8, pp. 172457–172466, 2020.

[39] L. Chen, G. Papandreou, I. Kokkinos, K. Murphy and A. L. Yuille, "DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 834–848, 2018.

[40] I. Batatia, "A deep learning method with CRF for instance segmentation of metal-organic frameworks in scanning electron microscopy images," in *European Signal Processing Conf.*, Amsterdam, Netherlands, pp. 625–629, 2020.

[41] H. Zhou, Y. Wang and M. Ye, "A method of CNN traffic classification based on sppnet," in *Int. Conf. on Computational Intelligence and Security*, Hangzhou, China, pp. 390–394, 2018.

[42] Y. Wu, X. Shen, F. Bu and J. Tian, "Ultrasound image segmentation method for thyroid nodules using ASPP fusion features," *IEEE Access*, vol. 8, pp. 172457–172466, 2020.

[43] A. Torralba, B. C. Russell and J. Yuen, "LabelMe: Online image annotation and applications," in *Proc. of the IEEE*, vol. 98, no. 8, pp. 1467–1484, 2010.

[44] J. Xu, X. D. Jian, L. Wang, Y. F. Shen, K. F. Yuan *et al.,* "Family-based big medical-level data acquisition system," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2321–2329, 2019.