CCC Publications

Segmentation Method of Magnetic Tile Surface Defects Based on Deep Learning

Y. An, Y.N. Lu, T.R. Wu

Yu An
College of Computer Science and Technology
Jilin University,
Changchun 130012, China
anyu19@mails.jlu.edu.cn

Yinan Lu*
College of Computer Science and Technology
Jilin University,
Changchun 130012, China
*Corresponding author: luyn@jlu.edu.cn

Tieru Wu
College of Mathematics
Jilin University,
Changchun 130012, China
wutr@jlu.edu.cn

Abstract

Magnet tile is an essential part of various industrial motors, and its performance significantly affects the use of the motor. Various defects such as blowholes, breaks, cracks, frays, unevenness, etc., may appear on the surface of the magnet tile. At present, most of the defects rely on manual visual inspection. To solve the problems of slow speed and low accuracy of segmentation of different defects on the magnetic tile surface, in this paper, we propose a segmentation method of the weighted YOLACT model. The proposed model uses the resnet101 network as the backbone, obtains multi-scale features through the weighted feature pyramid network, and performs two parallel subtasks simultaneously: generating a set of prototype masks and predicting the mask coefficients of each target. The residual structure and weights are introduced in the prediction mask coefficient branch. Linearly combine the prototypes and the mask coefficients to generate masks to complete the final target segmentation. The experimental results show that the proposed method achieves 43.44/53.44 mask and box mAP on the magnetic tile surface defect dataset, and the segmentation speed reaches 24.40 fps. Compared with other segmentation methods, our method is more excellent in performance.

Keywords: YOLACT Network, Deep Learning, Target Detection, Defect Segmentation.

1   Introduction

Magnet tile is a tile-shaped permanent magnet widely used in various motors and is the core component of the engine. In the production process of magnet tiles, various surface defects will appear on the product's surface due to multiple reasons, such as raw materials and equipment. The more significant defects are frays, breaks, and unevenness, and the smaller ones have blowholes and cracks. In China, most of the defects on the surface of products still rely on manual visual inspection and visual equipment inspection. Due to the cognition difference of manual inspection personnel and the wide variety of surface defects of the magnetic tile, it is difficult for workers to fully and correctly evaluate the quality of the magnetic tile. The equipment can detect relatively large defects, but it is difficult to detect minor flaws such as blowholes and cracks. Therefore, finding an accurate and fast automatic segmentation method of magnetic tile surface defects is essential.

The early automatic segmentation of magnetic tile surface defects was mainly based on the use of traditional image processing techniques. Li et al. used texture analysis and discrete curve wave transform to process the image of the surface defect of the magnetic tile. Under the condition of strictly controlled light, the grinding texture was effectively suppressed, and the effect of splitting crack defects was better [9]. Liu et al. enhanced the defects, suppressed the background texture through the notch filter, and then used the Canny edge detection operator to segment the magnetic tile surface defects [11]. Yang et al. first used NSST to decompose the image in multi-scale and multi-direction, performed corresponding processing in the high and low-frequency domains, and then segmented the crack defects from the reconstructed image. However, this method required high contrast of fine cracks and took a long time [19].

In recent years, many scholars have successively proposed many defect segmentation methods based on machine vision represented by convolutional neural networks. The segmentation method based on deep learning gradually replaced the traditional segmentation method. Existing deep learning segmentation methods are mainly divided into two categories. One is two-stage segmentation methods represented by Faster R-CNN [14] and Mask R-CNN [4]; The other is one-stage represented by YOLACT [1] and SOLO [16]. The two-stage method divides the segmentation task into two stages. First, the regional candidate network (RPN) generates a candidate area. Then, the detection network detects the category and location of the candidate area and performs segmentation in the detection frame. The accuracy of this method is higher, but the speed is slightly slower. The one-stage method does not require the RPN stage, and the segmentation result can be obtained directly. Therefore, the segmentation speed is faster, but the accuracy is relatively low.

Deep convolutional neural networks have shown excellent performance in the field of image segmentation. There is more research work in magnetic tile defect detection, but there are fewer applications in the field of segmentation. Nowadays, there are many improved Mask R-CNN models to study defect segmentation. Wu et al. improved the feature pyramid network of the Mask R-CNN model to a combination of top-down and bottom-up, modified the anchor's size, and introduced a dropout layer in feature extraction. The improved Mask R-CNN model realized the integration of workpiece detection and segmentation [18]. Wang et al. applied Mask R-CNN to segment wheel defect images for the first time, using deep convolutional neural networks to identify and segment different wheel defects [17].

Although Mask R-CNN has a high segmentation accuracy, it takes a long time. On the other hand, the YOLACT model of the latest single-stage real-time instance segmentation published in 2019 is relatively balanced between segmentation accuracy and speed. Therefore, it is one of the current excellent segmentation algorithms. Although the YOLACT also has problems in the test process, the target cannot be accurately located when the scene is complex, and the mask overlaps between two instances close to each other. However, the magnetic tile surface defect segmentation studied in this paper does not have this problem because each picture has only one defect. Therefore, the YOLACT model is suitable for the segmentation task of magnetic tile surface defects. However, as a one-stage method, the YOLACT algorithm has the problem of low segmentation accuracy. Therefore, we propose a weighted YOLACT (Weighted-YOLACT) method, referred to as W-YOLACT, to study the segmentation of magnetic tile surface defects. By weighting the feature pyramid and adding residual structure and weights to the prediction mask coefficient branch, we design a magnetic tile

defect segmentation algorithm that achieves higher accuracy while ensuring real-time speed.

The paper is organized as follows: The first section focuses on the background and significance of the research problems, introduces the related work, and presents the corresponding solutions we propose. The second section explains the detailed structure and innovation of our model. The third section elaborates the experimental process, results, and analysis in detail. Finally, the fourth section summarizes the paper.

2 Network Model

2.1 YOLACT Model

The YOLACT network model divides the instance segmentation task into two parallel subtasks, separating the calculations into parallel, significantly improving the calculation speed. Thus, it can achieve real-time speed while generating high-precision masks. The model in this paper is improved based on the YOLACT network. The dashed box is our improved content, as shown in Figure 1. The backbone of the model is composed of the residual network Resnet101 model [5] and the weighted feature pyramid network (described in detail in section 2.2 below), mainly for image feature extraction. The weighted feature pyramid is composed of layers P3 to P7, among which layers P3 to P5 are calculated from layers C3 to C5 corresponding to Resnet101. The P5 is down-sampled to get the P6 and P7. Input P3 into the prototype mask branch, and complete the prediction of k prototypes through the fully convolutional neural network. The other branch predicts the mask coefficients (described in section 2.3 below), which input is the feature maps of layers P3 to P7. Based on anchor target detection, a third small branch is added to predict k mask coefficients. Finally, three results of target category confidence, bounding box offset, and mask coefficient are predicted. The same operation is performed on layers P3 to P7, and these results are spliced together. Then fast non-maximum suppression algorithm (Fast NMS) is used to suppress repeated detection. Next, linearly combine the prototype mask and the corresponding mask coefficients to get the optimal segmentation map. Then, through the cropping and thresholding of each instance, the final segmentation of each instance is determined.
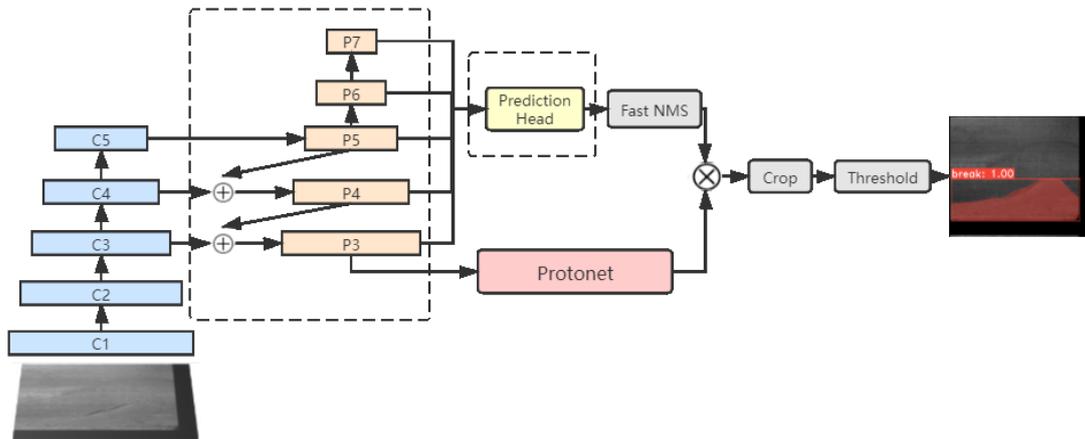

Figure 1: YOLACT Model

The loss function used by YOLACT is the linear sum of the three losses of classification loss L_{cls}, detection frame regression loss L_{box}, and mask loss L_{mask}. The weight ratios α, β, and γ are 1, 1.5, and 6.125, respectively. The larger the weight, the greater the contribution of the mission loss. Among them, the calculation methods of classification loss and detection frame regression loss are the same as those in SSD [12], which are softmax cross-entropy and smooth-L1 loss functions [3]. The mask loss uses the $BCE(M, Mgt)$ function to calculate the pixel-by-pixel binary cross-entropy loss between the generated mask M and the ground truth mask Mgt. Thus, the joint loss function formula of YOLACT is formula (1):

$$Loss = \alpha L_{cls} + \beta L_{box} + \gamma L_{mask} \tag{1}$$

2.2 Weighted Feature Pyramid Network

Feature Pyramid Network (FPN) is a method of horizontally connecting, top-down, and extracting various scale features in images. It combines high-level semantic information-rich features with low-level precise positioning signals to strengthen the expression of the network's backbone [10]. The feature pyramid network and various variants based on it are widely used in target detection tasks. BiFPN [15] is an improvement based on PANet [13]. BiFPN balances the feature information of different scales by letting the network learn the weights of different input features by itself. If the input and output are at the same level, BiFPN will add an extra edge to integrate more features without increasing cost, which is more concise and efficient than NAS-FPN [21]. Zhao et al. proposed combining the spatial attention module with the channel attention module, which paid attention to local and global features and improved the expression of comprehensive features. The combination of dual attention modules allowed buildings to be extracted from complex backgrounds [20]. The attention mechanism inspires us to add self-learning weights to the particular feature pyramid in the YOLACT model.

Weighted feature pyramid network is referred to as WFPN. In the past, feature fusion is to add features of different scales directly. However, features of different scales contribute differently to the final output and need to be treated differently. Therefore, we introduce weights (similar to attention) based on FPN, give different weights to features of different scales, and then add them to better balance multi-scale feature information. As a result, WFPN can perform multi-scale feature fusion simply and quickly, enabling deeper feature maps to generate a more robust mask. In addition, a larger prototype mask can ensure that the final mask has a higher quality and better detection of small objects, making the features learned by the neural network richer and more conducive to segmenting target objects of different sizes.
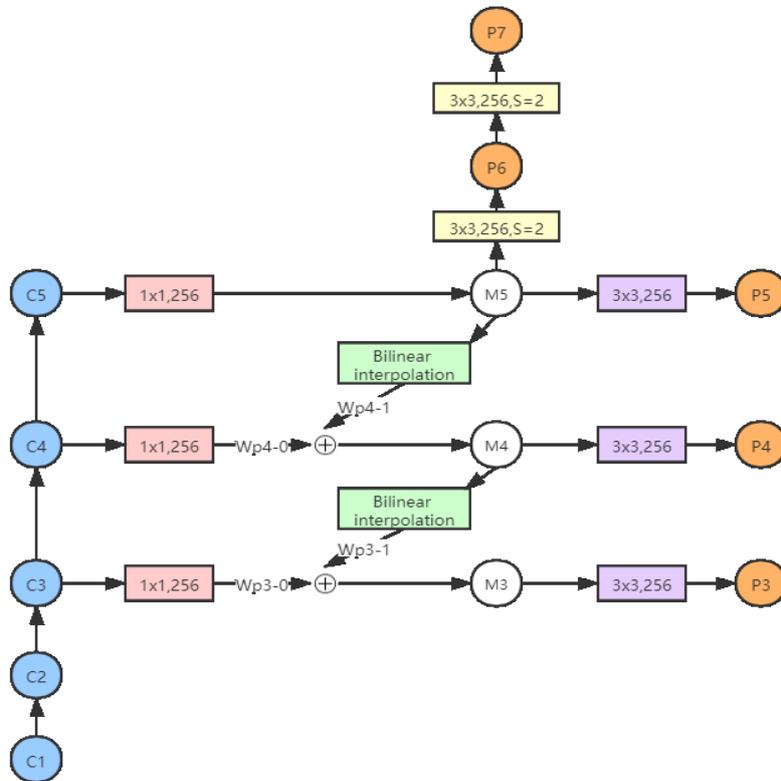

Figure 2: WFPN Structure

The structure of WFPN is shown in Figure 2, where C1-C5 represent feature maps of different scales output by the backbone network resnet101. P3-P7 represent the feature maps of different scales output by this WFPN. The pink box represents 1×1 convolution, the purpose of which is to unify all feature maps into 256 channels. The purpose of the bilinear interpolation operation is to enlarge the feature map. The purple box is a 3×3 convolution operation to prevent feature discontinuities caused by the superposition of the up-sampled image and the original image. Using the convolution operation is equivalent to the re-extraction of the features to ensure the stability of the features. M5 is obtained by convolution of C5, then perform a bilinear interpolation on M5 to double the feature map, and add the convolved C4 according to the parameter ratio to obtain the intermediate node M4. Then perform the convolution operation on M4 to obtain P4. P3 can be obtained in the same way. In addition, M5 is down-sampled by 3×3 convolution to obtain P6. P7 is obtained by 3×3 convolution on P6. Next is the parallel operation. P3 is sent to the prototype mask branch, and P3-P7 are sent to the prediction mask coefficient branch at the same time. The calculation formula is as (2):

$$P4 = Conv\left(\frac{W_{p4\_0} * Conv(C4) + W_{p4\_1} * Interpolate(M5)}{W_{p4\_0} + W_{p4\_1} + \varepsilon}\right) \qquad (2)$$

where $P4$ is the output feature of the fourth level on the top-down path. $M5$ is the intermediate feature of the fifth level on the top-down path. $C4$ is the fourth-level input feature on the bottom-up path. The outer layer $Conv$ is a 3×3 convolution operation, and the inner layer $Conv$ is a 1×1 convolution operation. $Interpolate$ is a bilinear interpolation operation. W_{p4_0} and W_{p4_1} are the parameters that need to be learned, $\varepsilon=0.0001$, in order to avoid numerical instability.

2.3 Prediction Mask Coefficient Branch

The prediction mask coefficient branch is to predict the coefficient of each mask. Based on the class and box branches of the target detector, an additional mask branch is added to predict k mask coefficients, and each mask coefficient corresponds to a prototype. This paper introduces the residual structure [6] and weights based on the original YOLACT, gives the forward neural network and the shortcut connection with different weights, and then adds them to better balance the characteristic information of the residual structure. As a result, it guarantees the depth of the network and avoids network degradation. The extracted information is also richer.
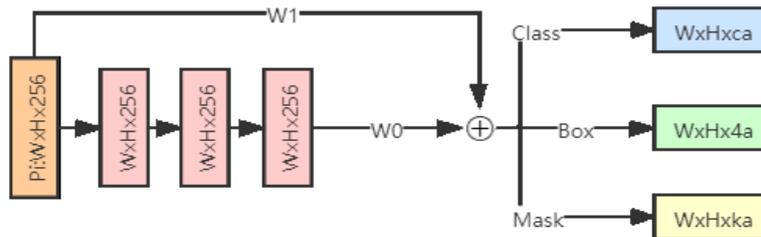

Figure 3: Prediction Head

The network structure of the prediction head is shown in Figure 3. Its input is the P3-P7 five-level feature maps, and the three branches use a shared 3×3 convolutional layer to increase the speed. The shared convolution part is realized through the residual structure and weights. The original feature map Pi and the Pi after three 3×3 convolutions are added according to the weight ratio. The result of the addition is input three branches respectively, and finally, three types of output are generated: category confidence. The mt data set used in this article has six categories (including background), namely c=6. Each pixel generates three anchors (a=3). The ratio is 1:1, 1:2, and 2:1. Its dimension is a×6; Position offset. The position is positioned by the four vertices of the bounding box. The dimension is a×4; Mask coefficient. K represents the number of prototypes, set to 32, and the mask coefficient dimension is a×32. Among them, W0 and W1 are the parameters that need to be trained. The operations of P3-P7 are the same, and the final output result is obtained by adding these results.

3 Experiment

3.1 Experimental Environment

The experiment is carried out in the environment of Intel Core i5-6500 CPU, GeForce RTX 2080 Ti graphics card, Ubuntu 18.04, and 64-bit version. The network model of this article is based on the cuda10.1+pytorch1.4.0 deep learning framework. The running environment is the integrated environment of Python 3.6 under Anaconda.

3.2 Dataset

The magnetic tile surface defect data set [7] used in this article was collected and published by a research group of the Institute of Automation, Chinese Academy of Sciences, with 784 images. There are 392 defective images and 392 marked images, and each defective image has its corresponding binary mask marked image. The defective image is in jpg format, and the marked picture is in png format. There are five kinds of defective images, among which 115 are blowholes, 85 are breaks, 57 are cracks, 32 are frays, and 103 are unevenness.

3.3 Data Preprocessing

In deep learning, most models require a sufficient number of samples because the larger the number of samples, the better the effect of the trained model and the stronger the generalization ability. However, the sample quality is often poor in the actual research process, or insufficient sample quantity. Therefore, it is necessary to perform data enhancement processing on the sample to improve the sample quality. Taking images as an example, the methods of data enhancement include data inversion, data rotation, image scaling, image cropping, and image translation.

Since the number of images in the magnetic tile defect dataset is too small to support deep learning research, the images are data enhanced. Firstly, rename all pictures and add the defect name to the picture name. Next, rotate all images by 90 degrees, 180 degrees, 270 degrees, and then flip all the original and rotated pictures horizontally, vertically, and horizontally and vertically. In the end, there are 3840 defective images in the training set and 1088 defective images in the verification set. Finally, the binary mask label information is converted into a JSON file in the format MS COCO.

3.4 Evaluation Indicators

The model in this paper uses the classic evaluation method of target detection and instance segmentation algorithms, mean Average Precision (mAP). Accuracy rate (P) indicates how accurate the model predicts a certain type of samples and the ratio of the positive samples to the total number of such samples in the prediction result. Recall rate (R) represents the ratio of the number of samples correctly predicted by a certain sample type to all the ground truth of this sample type. The AP value is the average value of the maximum precision rate under different recall rates (0-1). The AP values of all detected defects are averaged to obtain mAP, as shown in formula (3). Different thresholds are used to evaluate the overall performance of the detector. In this paper, the mAP value is calculated every 0.05 in the interval of the IOU threshold from 0.5 to 0.95.

$$mAP = \frac{1}{C} \sum_{i=1}^{C} AP_i \tag{3}$$

where C is the number of detection target classifications, and AP_i is the segmentation accuracy of a certain class.

This article uses 101-point interpolation to calculate the AP value, and the calculation formula is as (4):

$$AP = \frac{1}{101} \sum r \in \{0, 0.01, \dots, 1\} pinterp(r) \tag{4}$$

where $pinterp(r)$ represents the interpolation value corresponding to each point that is greater than or equal to the largest Precision value in the Recall interval, the formula is as (5):

$$pinterp(r) = \max_{\bar{r}:\bar{r}\geq r} p(\bar{r}) \tag{5}$$

Fps is an index used to evaluate the segmentation speed of the network model, that is, the number of pictures that can be segmented per second, and the calculation formula is as shown in formula (6):

$$fps = \frac{NumImage}{AllTime} \tag{6}$$

where $NumImage$ is the total number of pictures, and $AllTime$ is the total time of segmentation.

3.5 Training Details

When the model is trained, using YOLACT's loss function, the input image size is $550{\times}550$. We train models with batch size 8 on a single 2080 Ti GPU using resnet101_reducedfc model weights pre-trained on ImageNet. We train with SGD for 300k iterations starting at an initial learning rate of 0.001, dividing by 10 at iterations 30k, 80k, 150k, 220k iterations, and using a weight change gradient decay of 0.0005, a momentum of 0.9.

3.6 Experimental Results

The three-loss function changes of the classification loss Class Loss, the detection frame regression loss B-Box Loss, and the mask loss Mask Loss in the training process of W-YOLACT are shown in Figure 4. The linear sum of the three losses is the overall training loss. As shown in Figure 5, it tends to converge after 100,000 iterations and is basically unchanged.
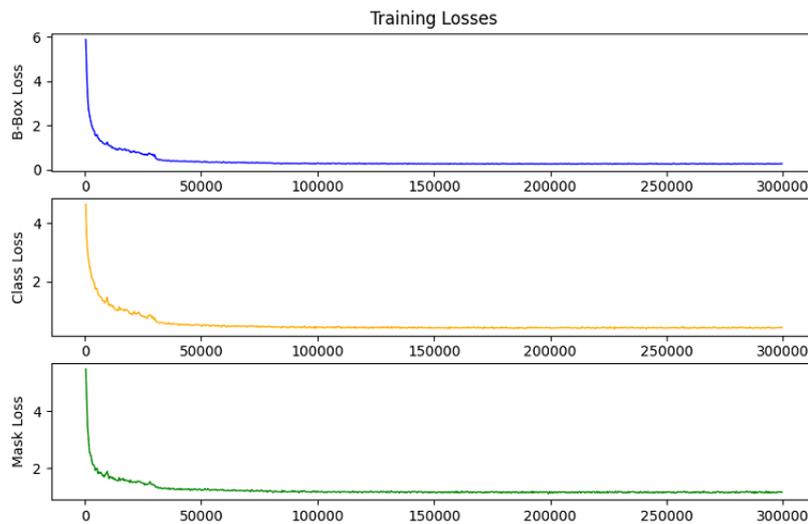

Figure 4: Change Graph of Each Loss

We use the Mask R-CNN model commonly used in the segmentation field to do the comparative experiment in this article. Figure 6 is the comparison result of defect segmentation. The first line is the original image, the second line is the ground truth of the image segmentation mask, the third line is the segmentation effect of the Mask R-CNN method, and the fourth line is the segmentation effect of the YOLACT method. The fifth line is the segmentation effect of our method. The five columns on the graph are named five defect names. It can be seen from Figure 6 that our method can segment the defects better. Compared with the Mask R-CNN method with the highest accuracy in the image segmentation field, the classification accuracy of crack and uneven defects is slightly worse, but the classification accuracy of the other three defects is equal. The positioning and segmentation effects on break, crack, even, and fray defects are better than Mask R-CNN.
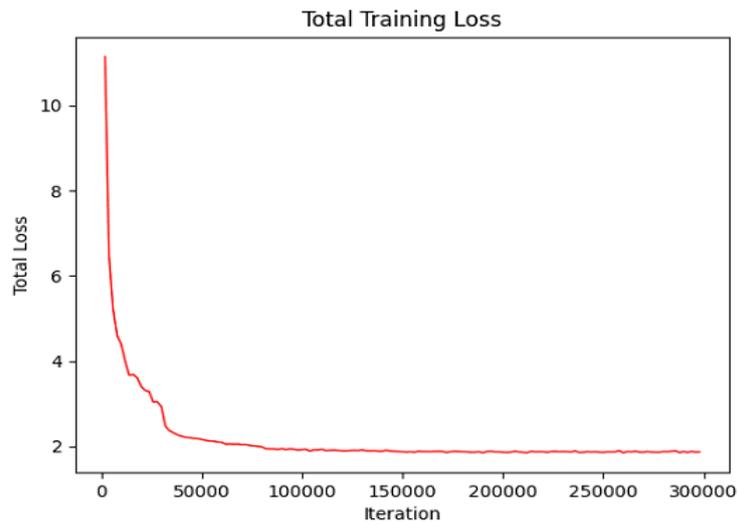
Figure 5: Change Graph of Total Training Loss

It can be seen from the break defect that the right side of the original YOLACT method positioning frame exceeds the actual defect position. The crack's head and tail defects are insufficient in positioning and segmentation, and the entire crack is not completely covered. Throughout the five defects, the classification accuracy of the method proposed in this paper is higher than that of YOLACT, and the positioning frame and segmentation mask are also better than the YOLACT method.
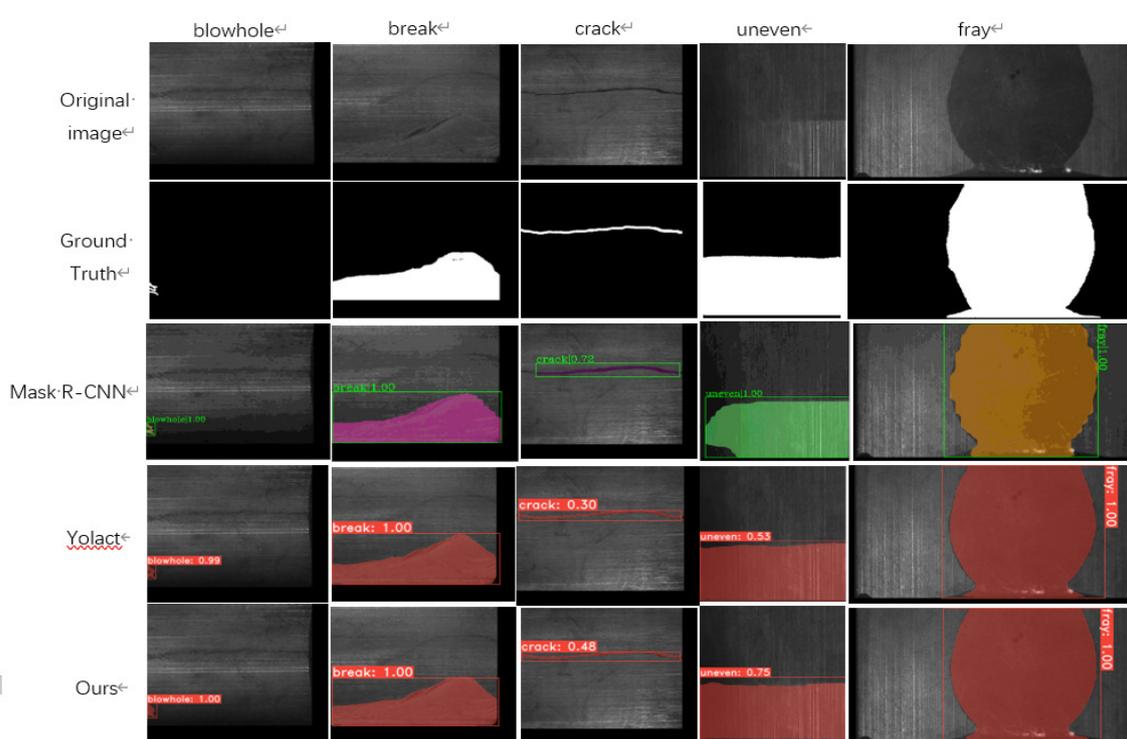

Figure 6: Comparison Results of Defect Segmentation

The evaluation results of the improved methods of the YOLACT algorithm in this paper are shown in Table 1. We calculate the mAP value every 0.05 on the interval of the IOU threshold value 0.50-0.95. Take the average of all the results as the final result (all). Each improved method is added to the YOLACT algorithm separately, and the mAP of the mask and box is improved. WFPN enhances feature extraction ability, improves the segmentation effect, and increases the mAP of the mask by 0.26. The addition of the Residual structure makes the information extracted by the network richer,

Table 1: Improved algorithm accuracy comparison table

		all	.50	.55	.60	.65	.70	.75	.80	.85	.90	.95	fps
mask /mAP (%)	YOLACT	42.18	76.86	70.42	61.85	55.03	49.16	42.60	30.09	18.11	11.53	6.19	25.72
	+WFPN	42.44	77.56	69.06	62.32	55.88	49.27	43.23	29.77	19.03	12.13	6.13	25.87
	+Residual	42.46	78.68	71.31	63.59	55.33	48.85	40.70	27.86	18.95	13.02	6.30	24.61
	+WFPN +Residual (Ours)	43.44	78.20	72.92	65.89	56.56	49.41	43.28	30.32	18.56	12.47	6.79	24.40
box /mAP (%)	YOLACT	52.71	88.93	84.27	78.63	73.38	64.37	52.71	40.49	25.35	15.08	3.85	25.72
	+WFPN	52.23	86.74	84.18	77.95	72.35	61.46	50.96	40.85	27.97	15.67	4.21	25.87
	+Residual	53.50	88.21	85.37	80.00	74.19	64.23	51.89	41.07	28.15	18.71	3.21	24.61
	+WFPN +Residual (Ours)	53.44	88.48	85.43	79.44	72.49	64.25	52.17	41.30	28.44	17.28	5.17	24.40

Table 2: Various types of defect recognition accuracy table

	mask/AP(%)					box/AP(%)				
IOU=0.50	blowhole	break	crack	fray	uneven	blowhole	break	crack	fray	uneven
YOLACT	89.47	68.40	60.20	84.55	81.10	97.30	88.53	79.46	93.65	85.53
Ours	90.89	72.14	58.03	86.91	82.63	97.15	88.64	74.52	94.59	87.25

which increases the mAP of the mask by 0.28. WFPN and Residual structure are jointly added to the YOLACT model to make the mAP of the mask reach 43.44, which is an apparent increase of 1.26, and the mAP of the box reaches 53.44, which has a good improvement effect. Segmentation speed running on a single 2080Ti GPU was still above 24 fps. Thus, under the premise of ensuring the speed, we have achieved the goal of improving accuracy.

Table 2 shows the average accuracy of various types of defects when IOU=0.50. Among them, blowhole, break, fray, and uneven defects segmentation effects have been improved and higher accuracy. However, the average accuracy rate of crack defects segmentation is low and has declined. It can also be seen in Figure 6 that the crack classification accuracy rate is low, and the detection effect is poor. The model needs to be improved to solve this shortcoming. On the whole, the improved model has higher segmentation and detection accuracy than the original YOLACT model.

Comparing Mask R-CNN, CenterMask [8], YOLACT, YOLACT++ [2] methods, and our method, the results are shown in Table 3. Compared with Mask R-CNN and CenterMask, our method has faster segmentation speed and higher mAP. Although YOLACT++ has a higher mAP than YOLACT's mask and box, the fps drops more, so we select the YOLACT model as the basic network. Compared with the YOLACT model, our method improves the mask mAP of defect segmentation by 1.26 and the box mAP by 0.73. The speed drop is not obvious, and it is also improved compared with YOLACT++.

Table 3: Comparison of segmentation performance of different algorithms

	mask/mAP(%)			box/mAP(%)			fps
	all	.50	.75	all	.50	.75	
Mask R-CNN	36.10	68.20	31.20	49.50	77.30	55.40	12.90
CenterMask	36.15	72.86	32.55	48.56	81.30	49.53	19.23
YOLACT	42.18	76.86	42.60	52.71	88.93	52.71	25.72
YOLACT++	43.28	79.68	41.69	53.72	88.92	52.79	22.92
Ours	43.44	78.20	43.28	53.44	88.48	52.17	24.40

4 Conclusion

In this paper, the problem of segmentation of the surface defects of the magnetic tile is deeply studied, and the deep learning method is introduced into the segmentation of the surface defects of the magnetic tile. Based on the YOLACT model, a weighted feature pyramid network is proposed, and the residual structure and weight are introduced. Based on the above improvements, we propose an efficient segmentation model for magnetic tile surface defects (W-YOLACT). By comparing the results of ablation experiments, the detection accuracy of the W-YOLACT model has been improved to a certain extent, and the model's performance in the segmentation task of magnetic tile surface defects is improved under the requirements of hardware resource constraints and maintaining speed. Compared with the other four segmentation models, the W-YOLACT model has a faster segmentation speed and good segmentation accuracy. Therefore, it provides a reliable basis for future applications in industrial production and has high research significance. However, for crack defects, the method in this paper has the problem of low recognition accuracy. In the future, we will focus on improving the algorithm to satisfy the comprehensive defect segmentation further.

Author contributions

The authors contributed equally to this work.

Conflict of interest

The authors declare no conflict of interest.

References

[1] Bolya, D.; Zhou, C., Xiao, F. (2019). Yolact: Real-time instance segmentation, *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 9157-9166, 2019.

[2] Bolya, D.; Zhou, C., Xiao, F. (2020). YOLACT++: Better Real-time Instance Segmentation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 99, 1-1, 2020.

[3] Girshick, R. (2015). Fast R-CNN, *2015 IEEE International Conference on Computer Vision (ICCV)*, 1440-1448, 2015.

[4] He, K.; Gkioxari, G.; P Dollár. (2017). Mask R-CNN, *Proceedings of the IEEE international conference on computer vision*, 2961-2969, 2017.

[5] He, K.; Zhang, X.; Ren, S.; Sun, J. (2016). Deep Residual Learning for Image Recognition, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770-778, 2016.

[6] He, K.; Zhang, X.; Ren, S. (2016). Identity Mappings in Deep Residual Networks, *Springer, Cham*, 630-645, 2016.

[7] Huang, Y.; Qiu, C.; Guo, Y.;Wang, X.;Yuan, K. (2018). Surface Defect Saliency of Magnetic Tile, *2018 IEEE 14th International Conference on Automation Science and Engineering (CASE)*, 612-617, 2018.

[8] Lee, Y.; Park, J.; (2020). CenterMask: Real-Time Anchor-Free Instance Segmentation, *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 13906-13915, 2020.

[9] Li, X. Q.; Jiang, H. H.; Yin, G. F. (2014). Detection of surface crack defects on ferrite magnetic tile, *NDT and E International*, 62, 6-13, 2014.

[10] Lin, T. Y.; Dollar, P.; Girshick, R. (2017). Feature Pyramid Networks for Object Detection, *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 936-944, 2017.

[11] Liu, G. P.;Hu, H. X.;Hu, R. H. (2015). Magnetic tile surface defect extraction based on texture features, *Modern Manufacturing Engineering*, 07, 119-123, 2015.

[12] Liu, W.; Anguelov, D.; Erhan, D. (2016). SSD: Single Shot MultiBox Detector, *Springer, Cham*, 21-37, 2016.

[13] Liu, S.; Qi, L.; Qin, H.; Shi, J.;Jia, J. (2018). Path Aggregation Network for Instance Segmentation, *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 8759-8768, 2018.

[14] Ren, S.; He, K.; Girshick, R. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks, *Advances in neural information processing systems*, 28, 91-99, 2015.

[15] Tan, M.; Pang, R.;Le, Q. V. (2020). EfficientDet: Scalable and Efficient Object Detection, *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 10778-10787, 2020.

[16] Wang, X.; Kong, T.; Shen, C. (2020). SOLO: Segmenting objects by locations, *European Conference on Computer Vision. Springer, Cham*, 649-665, 2020.

[17] Wang, T. R.; Wang, M. Q.; Zhang, J. S. (2021). Wheel defect segmentation technology based on Mask R-CNN, *Foreign Electronic Measurement Technology*, 40(02), 1-5, 2021.

[18] Wu, X. Y.(2020). Detection and Segmentation of Industrial CT Image Defects by Mask R-CNN[D], *Lanzhou Jiaotong University*, 2020.

[19] Yang, C. L.;Yin, M.;Jiang, H. H.(2017). Magnetic tile surface crack detection based on non-subsampled Shearlet transform, *Transactions of the Chinese Society of Agricultural Machinery*, 48(03), 405-412, 2017.

[20] Zhao D.D.; Zhao H.S.; Guan R.C.; Yang C. (2021). Efficient Building Extraction for High Spatial Resolution Images Based on Dual Attention Network, *International Journal of Computers Communications & Control*, 16(4), 4245, 2021.

[21] Zoph, B.; Vasudevan, V.;Shlens, J.;Le, Q. V.(2018). Learning Transferable Architectures for Scalable Image Recognition *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 8697-8710, 2018.

C | O | P | E

Member since 2012
JM08090

This journal is a member of, and subscribes to the principles of,
the Committee on Publication Ethics (COPE).
https://publicationethics.org/members/international-journal-computers-communications-and-control