



Electrolytic capacitor surface defect detection based on deep convolution neural network

Haijian Wang ^{a,b}, Han Mo ^b, Shilin Lu ^b, Xuemei Zhao ^{c,*}

^a Yinhe Biomedical Investment Co.Ltd, Beihai Guangxi, China

^b School of Mechanical and Electrical Engineering, Guilin University of Electronic Technology, Guilin Guangxi, China

^c School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin, Guangxi, China



ARTICLE INFO

Keywords:

Electrolytic capacitors
Surface defect detection
MobileNetV2
Focal Loss function
Transfer learning

ABSTRACT

The existing methods for detecting surface defects in electrolytic capacitors are typically based on conventional machine vision, with limited feature extraction capabilities, poor versatility, slow detection speed, and the inability to achieve accurate and real-time defect detection. In this study, a real-time object detection algorithm based on an improved single shot multibox detector (SSD) is proposed to achieve omnidirectional surface defect detection of electrolytic capacitors. First, an electrolytic capacitor surface image acquisition device was established to capture omnidirectional surface images of the capacitors, and an electrolytic capacitor surface defect dataset was created. Next, the visual geometry group (VGG)-16 network structure was replaced with the MobileNetv2 network structure, effectively reducing the model's parameter count and improving inference speed. Moreover, the Multibox Loss function was replaced with the Focal Loss function to increase the model's attention to difficult-to-classify samples and improve model accuracy. Additionally, a transfer learning network model was designed to apply the model to electrolytic capacitors of different colors using small sample learning. Finally, the performance of the improved network model was tested on a dataset of electrolytic capacitor surface defects. The experimental results demonstrate that the parameters quantity of improved model is 3.50 M, the mAP value reaches 92.67 %, which is improved by 2.54 %, and the Macro-F1 value reaches 92.15 %, which is 11.32 % higher than that before improvement. Thus, the proposed improved SSD model provides a theoretical basis and technical prerequisites for automated and intelligent surface defect detection in electrolytic capacitors.

1. Introduction

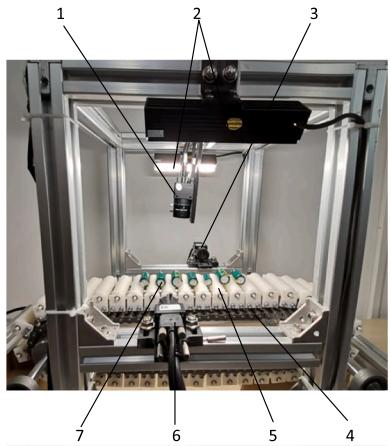
Electrolytic capacitors, benefit from the function of filtering and coupling, are widely used in the field of electronic manufacturing. However, due to processing technology or production equipment, internal defects or external defects often occurs during the production process of electrolytic capacitors. Generally, external defects have a variety of types such as pin burn, pin broken, scratch, damage and deformation. Once the defective electrolytic capacitors are applied to electronic products, it will cause product failure or even explosion, which may seriously threaten the personal safety (Shuai and Xiaohai, 2018). Thus, the quality of electrolytic capacitors should be strictly guaranteed in the manufacturing process. However, at present, manual detection is still the main surface defect detection method of electrolytic capacitors, which consumes lots of time and manpower. Moreover, manual detection is easily influenced by worker's subjectivity, ads to

misjudgment, and further significantly reduce the testing quality of electrolytic capacitors (Huimin et al., 2021; Dzhunusbekov and Orabayev, 2020).

The development of deep learning provides a fresh approach for surface defect detection of products. In 2015, Joseph Redmon team proposed the You-only-look-once (Yolo) detection algorithm (Redmon et al., 2016). As a typical one-stage detection algorithm, Yolo is fundamentally different from the two-stage detection algorithm in that the training and detection process are carried out simultaneously in the same network. Therefore, the processing speed of Yolo detection algorithm is faster, but the accuracy may not be as high as that of region with convolutional neural network feature (R-CNN) series (Redmon and Farhadi, 2017). To improve the detection accuracy, Yolo V2, Yolo V3, Yolo V4 and other Yolo detection algorithms were proposed (Redmon and Farhadi, 1804; Bochkovskiy et al., 2004). Firstly, Yolo V2 effectively improves the detection accuracy by adding Batch Normalization layer

* Corresponding author.

E-mail address: qingseyuji2010@163.com (X. Zhao).



1, 3 and 6-Camera 2-Light Source 4-Chain 5-Roller 7-Measured Electrolytic Capacitors

Fig. 1. Image acquisition system for electrolytic capacitors. 1, 3 and 6-Camera 2-Light Source 4-Chain 5-Roller 7-Measured Electrolytic Capacitors.

behind each convolution layer, using high-resolution pictures and cluster analysis of boundary boxes. Next, Yolo V3 uses Darknet-52 as the backbone network, the network structure is all convolution layer, and it adopts up-sample fusion method similar to feature pyramid network (FPN), which can effectively improve the accuracy of small target detection. Moreover, combined with spatial pyramid pooling and path aggregation network, Yolo V4 adopts a new backbone network, which is capable to make the network model have higher performance. Compared with above algorithms, the single shot multibox detector (SSD) convolutional neural network algorithm was proposed by Liu W et al. in 2016 (Liu et al., 2016), which made several improvements, such as using multi-scale feature maps for detection, setting up a default box with multiple aspect ratios, data enhancement and replacing the fully connected layer with a convolutional layer for training. As a result, the detection speed and accuracy were greatly improved, the mean Average Precision (mAP) in the visual object classes (VOC) 2007 dataset reached 74.3 %, but there are still major shortcomings for small target detection.

This paper aims to achieve high-precision detection of surface defects in electrolytic capacitors, and an experimental platform was built to collect defect images of electrolytic capacitors. Based on the collected images, a convolutional neural network was constructed, and relevant indicators such as model parameters, detection time, and control accuracy were comprehensively considered to achieve synthesized defect detection of electrolytic capacitors of different sizes and types. Moreover, aiming at the problem that the surface defects of electrolytic capacitors are diversified and the characteristics of some defects are not obvious, the neural network model was then further optimized to improve the detection accuracy, then comparing and analyzing the advantages and disadvantages of different detection methods. Finally, in response to the challenge of quickly achieving high-precision detection of different electrolytic capacitors, a transfer network model was constructed to complete the transfer learning process with a small sample size.

2. Construction of dataset

2.1. Experimental platform

In the acquisition process of electrolytic capacitor defect images, multi-angle and all-round image acquisition are required, including side (360°), top and bottom (side with pins), in which three industrial cameras are used to take charge of the above respectively. For the side of the electrolytic capacitors, a roller structure is applied to the drive system to enable it to rotate during the drive process, allowing a complete

Table 1

Information on acquisition equipment and parameters.

Acquisition equipment	Light source system	Industrial camera	Camera lens
Parameters	Type: LED Wavelength: 460 ~ 636 Shape: Bar	Camera model: MV-HS120GM Maximum resolution: 1280 × 960 Maximum frame rate: 54fps	lens model: BT-118C1214MP5 Resolution: 5 M Focal length: 12 mm

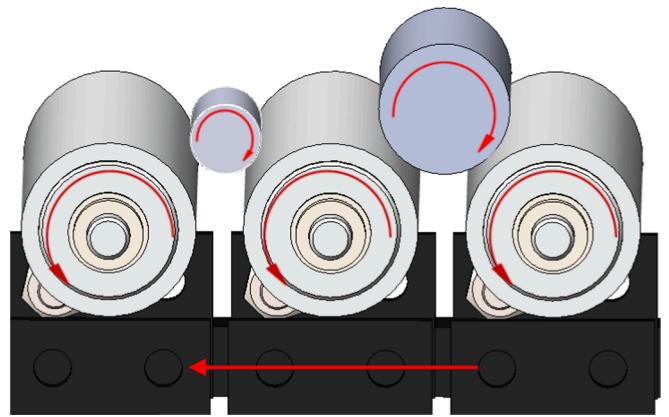


Fig. 2. Illustration of the partial structure of the drive system.

image of the side surface to be acquired. The image acquisition system is shown in Fig. 1.

The parameters of the electrolytic capacitors image acquisition system are shown in Table 1.

Fig. 2 depicts the drive system's partial construction. Each measured electrolytic capacitor is positioned between two rollers, and the rotation motor drives the horizontal movement of the electrolytic capacitors, while the sliding friction generated by the roller and the bottom board results in the counterclockwise rotation of the roller, which drives the clockwise rotation to the electrolytic capacitor. The whole sides of the electrolytic capacitor can be acquired by camera 1 in Fig. 1, while cameras 3 and 6 can be responsible for the top and bottom sides, thus realizing multi-angle and all-round dynamic acquisition of the electrolytic capacitor.

The image acquisition software adopted in this paper is MV Demo provided by Microvision, and the software interface of the image acquisition system is shown in Fig. 3.

2.2. Data sample acquisition

In this paper, six types of surface defects of electrolytic capacitors, namely, pin burn, pin broken, scratch, damage-s (side), damage-tb (top, bottom) and deformation were studied and analyzed. The position of three cameras were used to acquire images of the side, top and bottom surfaces of the electrolytic capacitors, and the six types of electrolytic capacitors surface defect images are shown in Fig. 4.

4000 images containing six types of defects were acquired after manual screening, including 668 images of pin burn defects, 668 images of pin broken defects, 664 images of scratch defects, 668 images of damage-s defects, 668 images of damage-tb defects, and 664 images of deformation defects.

2.3. Data labelling

After acquiring the surface images of electrolytic capacitors, it is required to label and generate a defect sample dataset of electrolytic

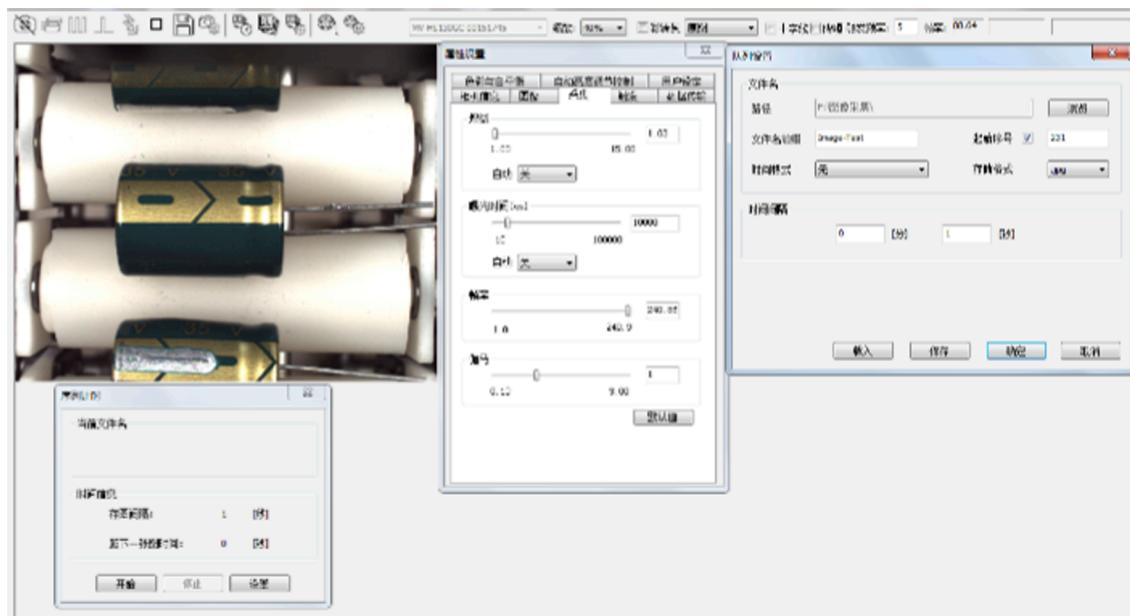


Fig. 3. Image acquisition system software interface.

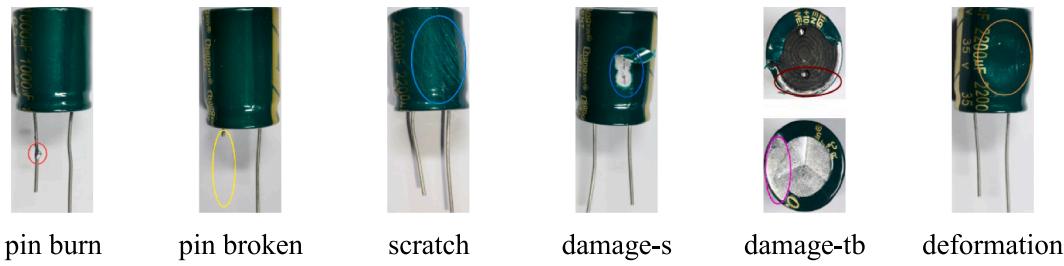


Fig. 4. Six defects in the surface of electrolytic capacitors.

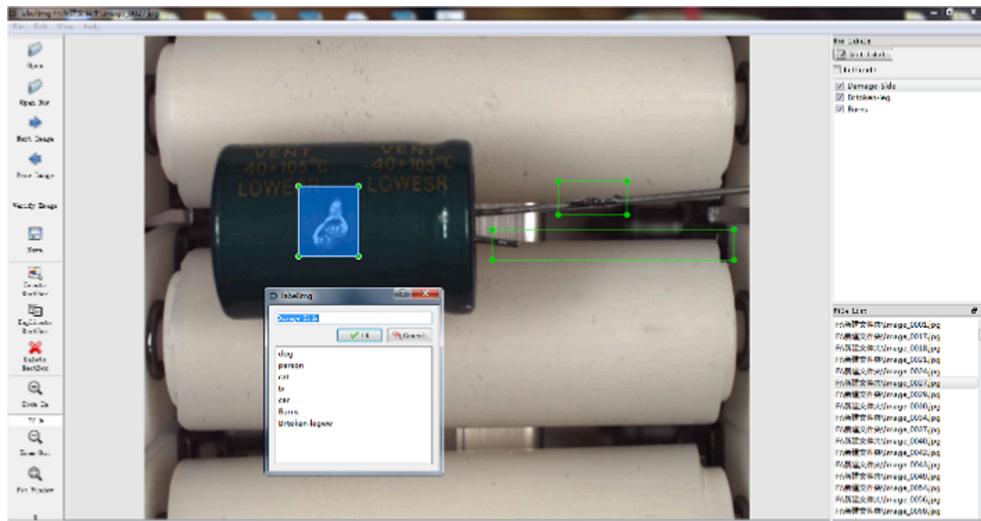


Fig. 5. Data image labelling process.

capacitors, which contains both the images and the labelled data. The labelled data is stored in a generated XML file, which contains the number, type, location and size of defects present in the image to provide the required information for model training. The data image labelling process is shown in Fig. 5.

2.4. Data expansion

Data expansion, as a method of expanding dataset, is widely adopted in current network training which can effectively augment dataset and improve the accuracy and robustness of model detection. The commonly adopted data expansion methods include random flipping, cropping,

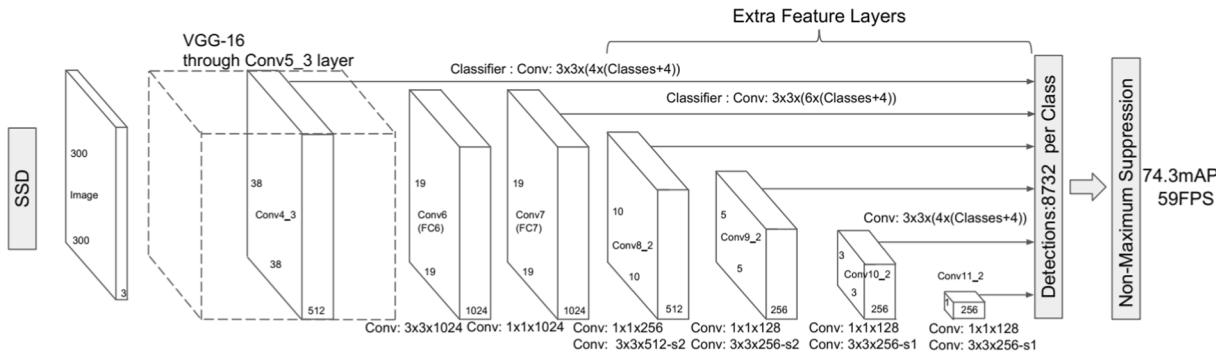


Fig. 6. Illustration of SSD network model.

scaling and adding noise, etc. Data expansion not affects the pixel characteristics of the original images, and it also diversifies the image presentation, enhances network generalization and prevents overfitting of the model. In this paper, considering the image characteristics of electrolytic capacitors and the acquisition effect, the image enlargement, image rotation, and image cropping are performed to expand the dataset. Finally, a total of 10,000 electrolytic capacitor surface images containing six types of defects were obtained by data expansion, including 1668 images of pin burn defects, 1668 images of pin broken defects, 1664 images of scratch defects, 1668 images of damage-s defects, 1668 images of damage-tb defects, and 1664 images of deformation defects. After data expansion, 80 % of the image data is used for training, 10 % for data testing and 10 % for data evaluation.

3. Algorithm improvement

3.1. SSD algorithm improvement

MobileNetV2 is a small and efficient lightweight network model, which has a smaller computational complexity. Moreover, the structure of MobileNetV2 network can significantly reduce the computational effort of the network model and improve the detection speed. Thus, compared with traditional convolutional neural networks, the improved SSD algorithm based on MobileNetV2 is constructed by replacing the visual geometry group (VGG)-16 network with a lightweight network structure.

3.1.1. SSD network model

The SSD convolutional neural network was proposed by Wei Liu et al. in 2016 and 2018 based on single-stage target detection network, which integrates the entire detection process into a standalone convolutional neural network. SSD combined elements from faster region

with convolutional neural network feature (R-CNN) and YOLO algorithms to create a novel convolutional neural network model and it simplifies training and optimization, meanwhile enhancing detection speed. The network architecture diagram is illustrated in Fig. 6 (Qi and Cai Jin Ji yuan, 2021).

From Fig. 6, SSD network model is improved based on VGG-16 network model, and the improvements are mainly reflected as follows.

- 1) Two convolutional layers were used to extract feature instead of the fully connected layers in VGG-16.
- 2) Replace the pooling layer convolution kernel size from $2 \times 2\text{-S}2$ to $3 \times 3\text{-S}1$.
- 3) Remove all dropout layers and apply the dilated convolutional algorithm.

3.1.2. Depthwise separable convolution

By replacing the traditional convolutional calculation method with depthwise separable convolution, the MobileNetV2 model reduces the parameter computation of convolutional neural networks. In depthwise separable convolution, each channel value of the convolutional kernel is set to 1, the input feature matrix channels correspond one-to-one with the convolutional kernel, and the number of convolutional kernels matches the number of channels. After depthwise separable convolution, a 1×1 pointwise convolution is added to modify the output feature matrix channels. Therefore, the convolutional layers in the MobileNetV2 model are referred to as depthwise separable convolution (Wei and Kai, 2021). Fig. 7 shows the architecture of the depthwise separable convolutional network model.

The computational complexity of convolutional neural network parameters can be effectively reduced by introducing the concept of depthwise separable convolution (Nikos, 2021), and the ratio of ordinary convolution to depthwise separable convolution P is expressed

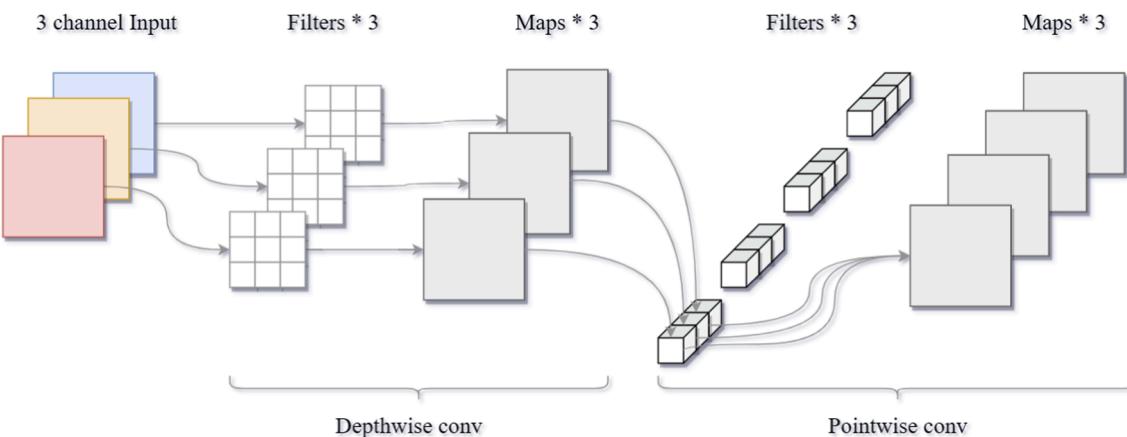


Fig. 7. Illustration of depthwise separable convolution network model.

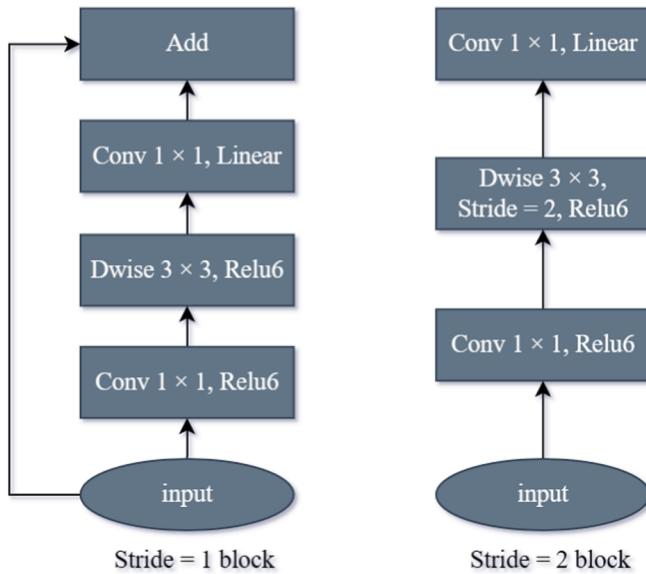


Fig. 8. Illustration of the residual structure of MobileNetV2.

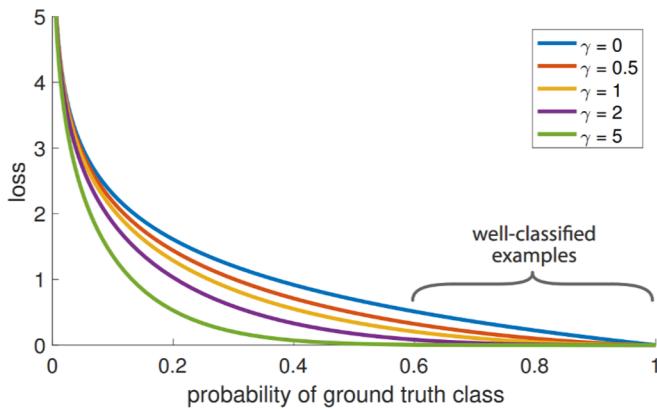


Fig. 9. Illustration of loss curves for different values of γ .

using Eq. (1) as follows:

$$P = \frac{D_K^2 \cdot M \cdot D_F^2 + M \cdot N \cdot D_F^2}{D_K^2 \cdot M \cdot N \cdot D_F^2} = \frac{1}{N} + \frac{1}{D_K^2} \quad (1)$$

where D_K denotes the size of the convolution kernel, M denotes the

channel of the input feature matrix, D_F denotes the width and height of the input feature matrix, N denotes the channel of the output feature matrix.

From Eq. (1), obviously, in terms of computational complexity, depthwise separable convolution is significantly lower than traditional convolutions, particularly when training complex network models. Thus, the depthwise separable convolutional network structure has greater advantages.

3.1.3. Residual structure

The residual structure, adopted in the MobileNetV2 network structure, is classified into two different strides, namely, stride = 1 and stride = 2, as shown in Fig. 8. When stride = 1, using ReLu6 as the activation function, the size was enlarged through 1×1 convolutional completion first. Next, a 3×3 depthwise separable convolution operation was carried out. Moreover, a 1×1 convolution was proceeded for dimension reduction. When Stride = 2, downsampling was performed by depthwise separable convolution instead of pooling layers, which prevented gradient explosion, increasing feature diversity and speeding up training (Weiwei et al., 2021).

3.2. Multibox Loss function improvement

By replacing the original Multibox loss function with the Focal Loss function, the proportion of negative samples were reduced. The network model focus on samples that are difficult to classify during training, thereby improving detection accuracy. Thus, in this paper, Focal Loss function was used to detect the defect of electrolytic capacitors. The Focal Loss function is expressed using Eq. (2) as follows:

$$FL(p_t) = -\alpha_t(1-p_t)^\gamma \log(p_t) \quad (2)$$

where α_t denotes the weight of samples, and is used to balance the number of positive and negative samples; p_t denotes the prediction probability of the model on the sample; γ is a tunable parameter, controls the weight difference of difficult and easy samples, with a value size ranging from [0, 5], Fig. 9 shows the loss curves for different values of γ , obviously, when the value of γ is set to 5, the loss value decreases rapidly, which ensures stable convergence of the model (Zhi et al., 2020).

4. Transfer learning

Transfer learning, fine-tunes the SSD convolutional neural network model by using a small amount of target training samples, thereby applying pre-trained model parameters to a new network model structure (Guangyu et al., 2020). In this paper, the constructed convolutional

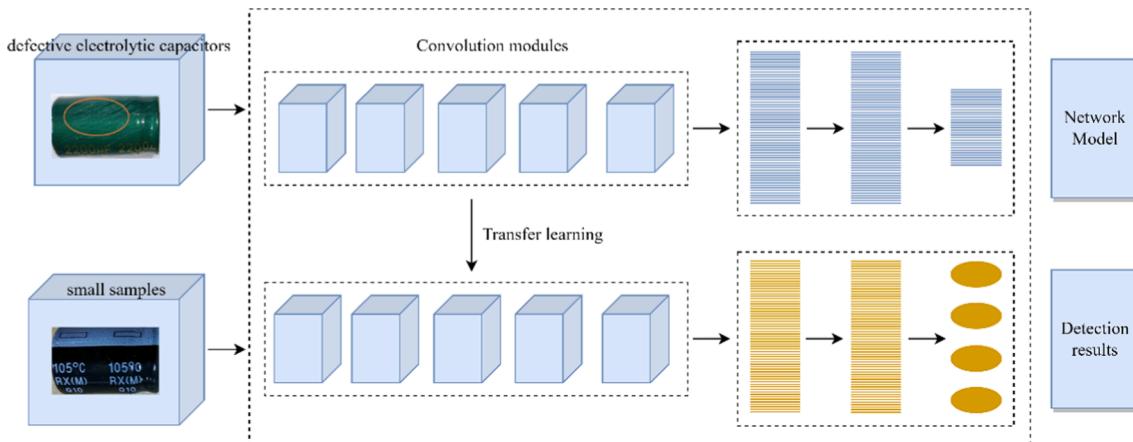
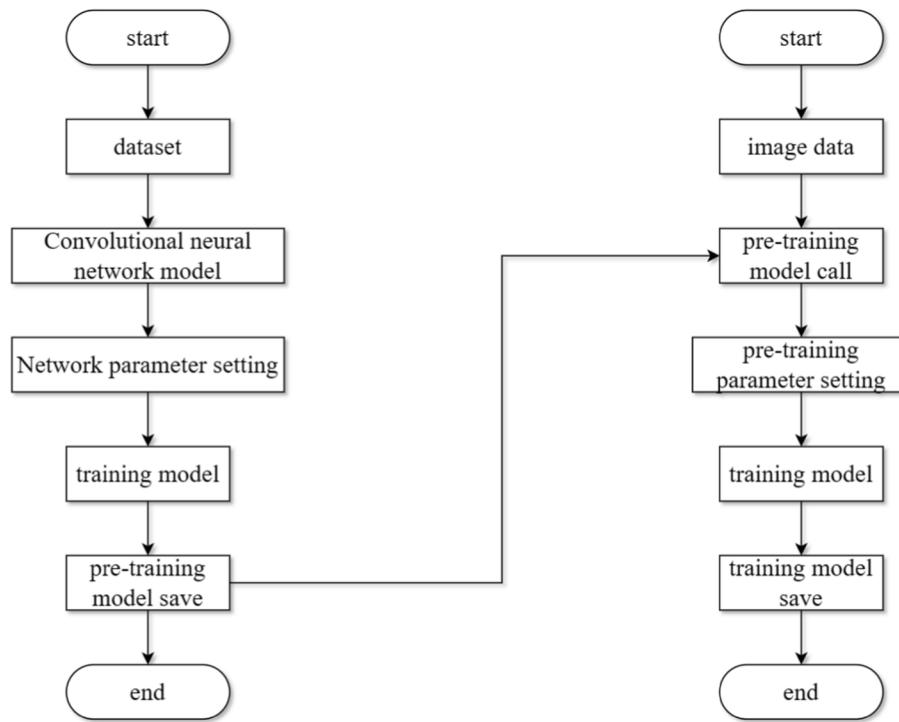
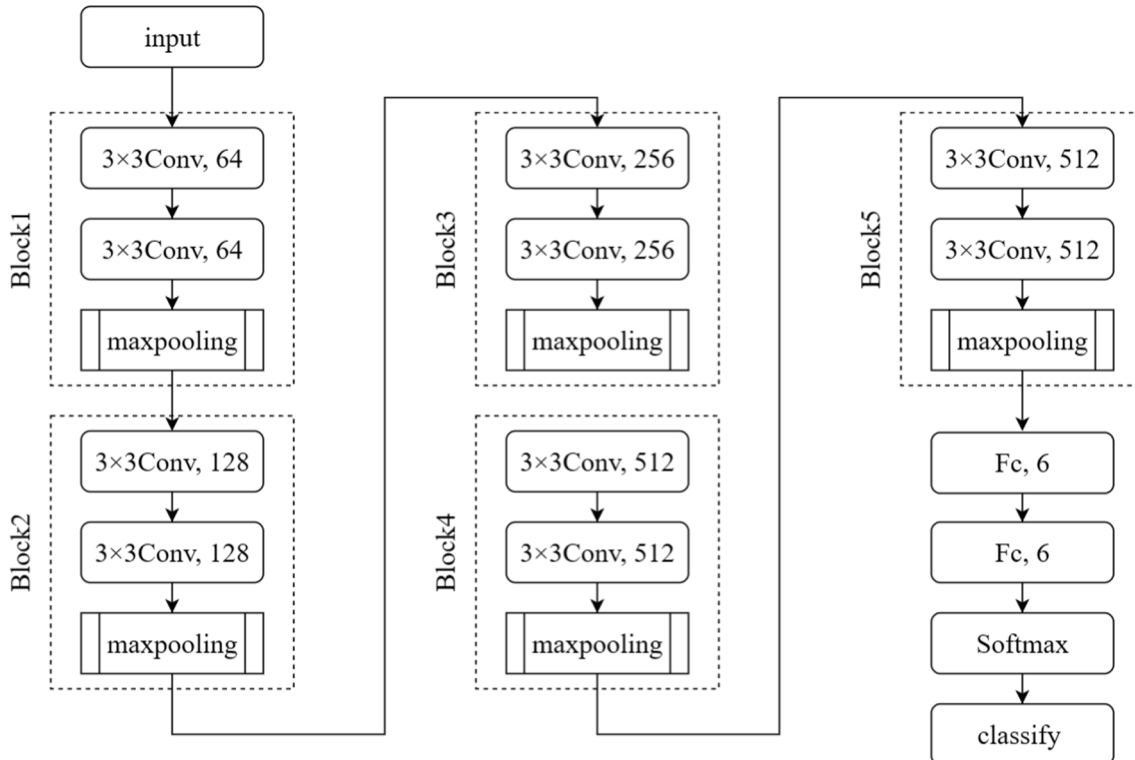


Fig. 10. Illustration of the framework for the transfer learning network model.

**Fig. 11.** Transfer learning flowchart.**Fig. 12.** VGG-16 network model in transfer learning.

neural network model is initialized with the green electrolytic capacitor dataset from Chapter 2, and the model is trained with a small amount of black target electrolytic capacitors. By adjusting the parameters of the network structure, a small-sample convolutional neural network model suitable for electrolytic capacitors was obtained, and finally realized the detection and classification of defects in the target electrolytic capacitor,

as shown in Fig. 10.

4.1. Transfer learning model training method

According to the given network structure in Fig. 10, transfer learning applies the parameters of the network model trained in the source

Table 2
Information on the experimental environment.

Name	Information
CPU	Intel(R) Xeon(R) Gold 5120
Memory	256 GB
GPU	NVIDIA Quadro GV100(32 GB, 5120 cuda core)
Operating system	Ubuntu 18.04
Development software	Cuda10.2, Pytorch1.7, Python3.5

domain to the target domain. As VGG-16 has the advantages of high accuracy and stable structure, in this paper, the VGG-16 network is used as a pre-training model for the electrolytic capacitor defect detection system and transfer learning is performed on defective electrolytic capacitors of different colors. The transfer learning is divided into two processes: pre-training and network model transferring, as shown in Fig. 11 (Jiang et al., 2021).

The left part of Fig. 11 shows the pre-training process of the network model and the right part shows the transfer process of the network model. The pre-training process is similar to the convolutional neural network training process, and the pre-trained model is output when the loss value of the network model converges stably, then the transfer learning process is completed by fine-tuning the network structure parameters through the target dataset.

4.2. Transfer learning network structure design

During the transfer learning process, it is required to pre-train VGG-16 by using the source dataset. First, transferring the parameters of the network model with well-feature extraction obtained from pre-training into the VGG-16 network of the target model as the original parameters, and then training the network model to complete the parameters fine-tuning, where the learning rate is set to 0.00001. As shown in Fig. 12, VGG-16 contains a total of 16 layers of neural network, which is divided into 5 convolution blocks, and the transfer process is completed by loading the parameters of the pre-training process into the blocks (Muhammad et al., 2021).

5. Experiments

5.1. Experimental environment

In this paper, an AMAX TR40-X2-type supercomputer was used, which supports 4 pieces of double-wide GPU acceleration. Moreover, Cuda10.2, Pytorch1.7, and Python 3.5 environment configuration were

completed in Linux operating system to realize GPU acceleration, thereby accelerating the training speed. The experimental environment is shown in Table 2.

5.2. Data evaluation metrics

Four indicators were used in this experiment, namely, Recall, Precision, mAP and F1-Score, where Recall represents the ratio of the number of correctly identified targets to the total number of targets, and Precision represents the ratio of the number of correctly identified targets to the total number of identified targets. The Recall and Precision are expressed using Eq. (3) and Eq. (4), respectively (Xiaoting, 2019).

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (4)$$

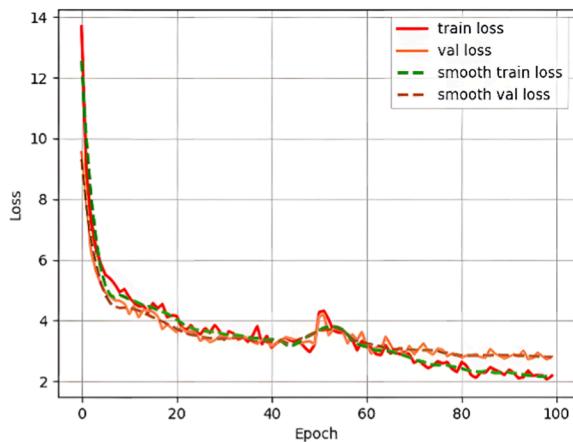
where TP denotes true positive samples with correct positive sample assignment; FP denotes negative samples assigned as positives; FN denotes positive samples assigned as negative samples; mAP denotes the average of the categories in the dataset and is commonly employed as a measure of the precision of the overall algorithm; F1-Score is commonly employed as the ultimate measure for multi-target classification. A single F1-Score can be calculated by dividing the product of Precision and Recall by the sum of Precision and Recall, as shown in Eq. (5).

$$\text{F1Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

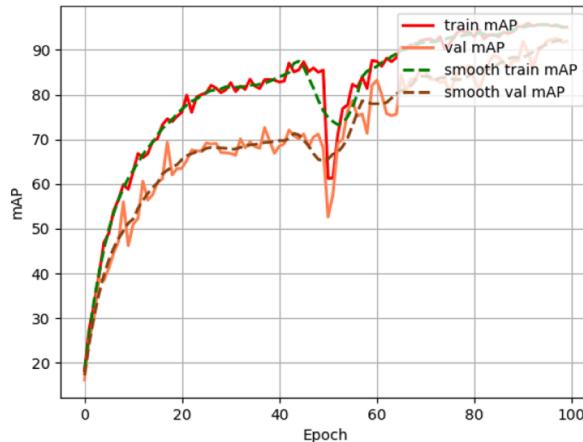
5.3. Network model training

The given model was trained with an input image size of $300 \times 300 \times 3$ and a training period of 100 epochs. Fig. 13 shows the change of loss function during the training process.

From Fig. 13(a), both the train loss and val loss reach a relatively stable state of convergence after adjusting the model parameters. Moreover, the train loss gradually converges to a stable state around 40 times in epoch, while the smooth loss also reaches a reasonable state at the end. Fig. 13(b) exhibits the mAP of the model for the training and validation set during the training process, both the train mAP and val mAP gradually increase and eventually exceed 90 %.



(a) Loss function curves



(b) mAP curves

Fig. 13. Training process curve.

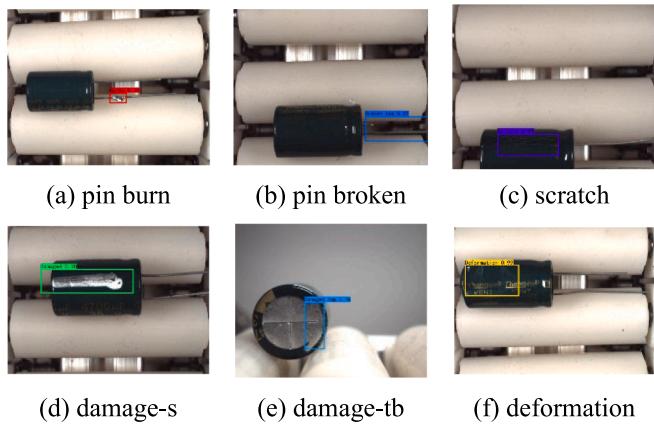


Fig. 14. Detection results.

Table 3
Analysis of indicators for identifying defects in electrolytic capacitors.

Defect Types	Indicators				
	F1-Score	Recall (%)	Precision (%)	AP (%)	mAP (%)
pin burn	0.85	81.28	88.37	85.22	89.27
pin broken	0.85	73.85	100	86.04	
scratch	0.90	82.15	100	100	
damage-s	0.89	80.46	100	99.35	
damage-tb	0.86	82.06	91.24	84.14	
deformation	0.91	84.15	100	88.16	

Table 4
Analysis of indicators after data expansion.

Defect Types	Indicators				
	F1-Score	Recall (%)	Precision (%)	AP (%)	mAP (%)
pin burn	0.89	82.89	97.02	91.17	94.42
pin broken	0.92	85.52	100	90.26	
scratch	0.96	92.18	100	96.88	
damage-s	0.91	83.42	100	95.64	
damage-tb	0.92	84.88	100	100	
deformation	0.93	86.95	100	92.56	

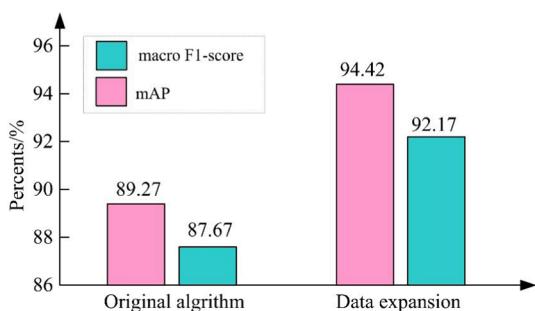


Fig. 15. Detection results before and after data expansion

5.4. Results Analysis

In this paper, six surface defect types of electrolytic capacitors, contain pin burn, pin broken, scratch, damage-s (side), damage-tb (top, bottom) and deformation, were detected. According to the detection results for each defect type, the convolutional neural network model was evaluated. The examples of detection results are illustrated in Fig. 14.

When the training and testing of the SSD neural network model were carried out, the performance of model was evaluated using F1-Score,

Table 5
Map for the four sets of experiments.

experiments	VGG-16	MobileNetV2	Multibox Loss	Focal loss	mAP/%
1	✓		✓		90.05
2	✓			✓	91.85
3		✓	✓		89.33
4		✓		✓	96.86

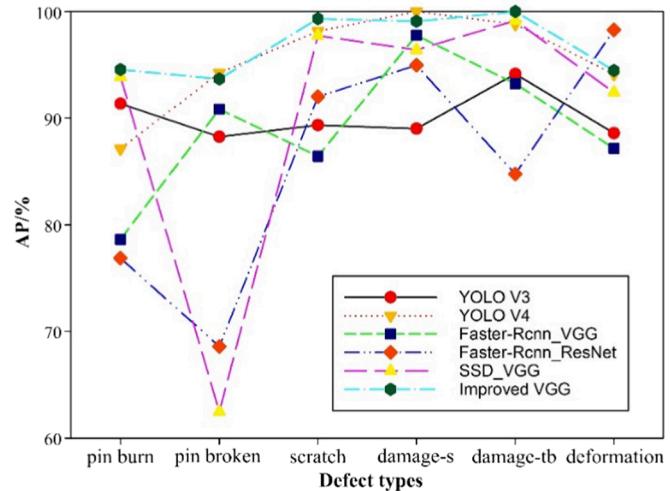


Fig. 16. AP values of six defect types based on different algorithms.

Recall, Precision, AP(Average Precision), and mAP, as shown in Table 3. The minimum Precision reaches 88.37 %, and the Recall of pin broken is slight lower than others, which is 73.85 %. Moreover, the F1-Score of each defect type is higher or equal to 0.85, and the mAP of SSD network model is 89.27 %.

5.5. Comparative analysis

5.5.1. Comparison of data expansion

Table 4 presents the results after data expansion. From the comparison between Table 3 and Table 4, both precision and recall are improved. Moreover, the F1-score after data expansion increases 0.045 than that of Table 3, and the mAP is improved by 5.15 %, as show in Fig. 15. Thus, data expansion can effectively improve the detection accuracy of the network model.

5.5.2. Comparison of algorithm improvement

(1) Comparison of mAP and AP.

The MobileNetV2 network was used to replace VGG-16 and the Focal Loss function was applied to significantly improve deflection accuracy. As shown in Table 5, four groups of experiment were carried out, combining VGG-16 and MobileNetV2 networks with two sets of functions, Multibox Loss and Focal Loss, respectively.

From Table 5, experiment 1 used the VGG-16 network model with Multibox Loss function, namely, the original SSD model, and the mAP is 90.05 %. Combined VGG-16 network model with Focal Loss function, experiment 2 improved the detection accuracy to 91.85 %. However, experiment3 adopted the MobileNetV2 network model with Multibox Loss function, the detection accuracy is merely 89.33. Finally, the accuracy of experiment 4, which used MobileNetV2 network model with Focal Loss function, reached 96.86 % and was significantly better than the other three groups.

Moreover, to verify the effectiveness of the improved SSD algorithm, the comparative experiments were carried out between the improved SSD algorithm and five mainstream algorithms. Fig. 16 shows the AP

Table 6

Comparison of algorithms before and after improvement.

Model	Training time/s	Params/M	F1-score
SSD	1680	138.36	0.88
Improved SSD	1520	3.50	0.93

values of six defect types based on different algorithms. Obviously, the AP values for the defeat of pin burn, scratch and damage-tb are maximum compared with other detection algorithms. Although the AP values of the defeat of pin broken, damage-s and deformation are slightly lower than the maximum value, the minimum value of pin broken is still quite high, reached 93.69 %. Thus, the improved SSD algorithm has high and stable detection accuracy for six defect types.

(2) Comparison of training speed, params and F1-score.

The detection speed and params are the key indicators of improved SSD algorithm performance. Table 6 shows the comparison of the training time, params and the F1-score of the network model before and after the improvement.

From Table 6, the params of improved SSD algorithm was significantly reduced, which is 39.5 times lower than SSD algorithm. Meanwhile, the training time was also shortened from 1680 s to 1520 s and the F1-score was improved from 0.88 to 0.93. Thus, the improved algorithm is significantly better than the original algorithm, both in terms of detection accuracy and detection speed.

(3) Loss function

Fig. 17 records the changes of the loss values of the original algorithm and the improved algorithm during the training process of the algorithm. From Fig. 17(a), the original algorithm has a slow decline in the loss values during the training process, and there is a fluctuation around epoch 50, which may impact on the convergence results of the model. Fig. 17(b) shows the loss function curve of improved algorithm. Obviously, the decline rate of loss function curve during the training process was improved and the decline trend was stable, indicating that the improved algorithm converges stably and is easy to train.

(4) macro-F1 comparison

The Macro-F1 can effectively reflect the accuracy of the target detection algorithm. From Fig. 18, the Macro-F1 of the YOLOv3 algorithm reaches 91.14 %, which is higher than the original SSD algorithm, YOLOv3, Faster Rcn_VGG and Faster Rcn_Resnet. Compared with YOLO v3, the Macro-F1 of improved SSD detection algorithm reaches 93.22 %, which is 2.08 % higher than YOLOv3 and 5.15 % higher than the original SSD algorithm.

5.5.3. Transfer learning results

By designing the transfer learning network model based on small samples, it is capable to identify the defects of electrolytic capacitors with different surface colors, the detection results are shown in Fig. 19.

To verify the impact of different convolutional and network layers on the detection accuracy of given network model, a comparison experiment was carried out by freezing different convolutional layers with network layers in three different training methods: (1) without freezing any network structure; (2) freezing a 5-layer network structure; and (3)

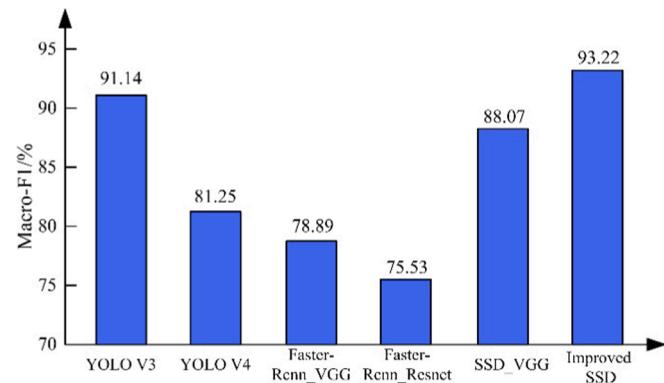


Fig. 18. Macro-F1 comparison of target detection algorithms.

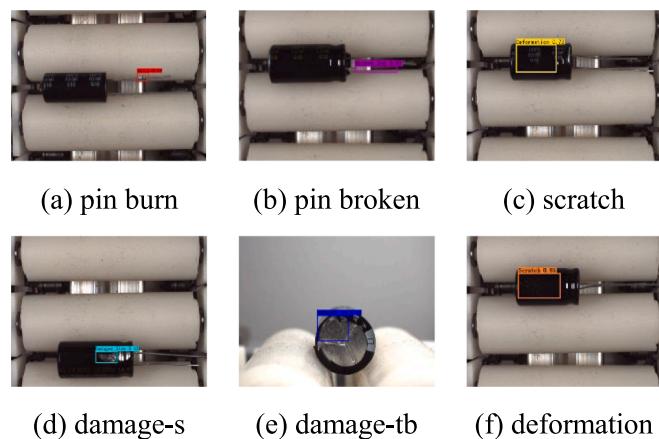
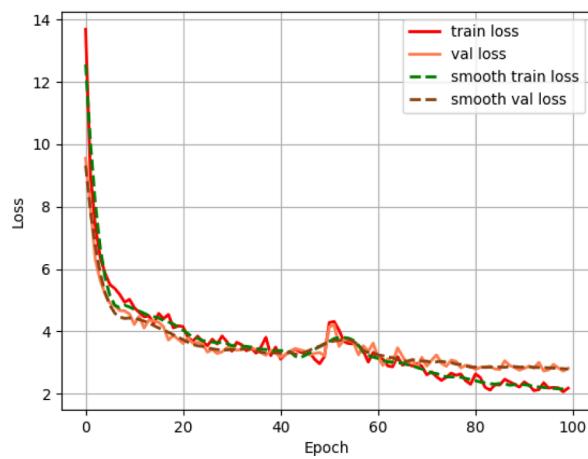
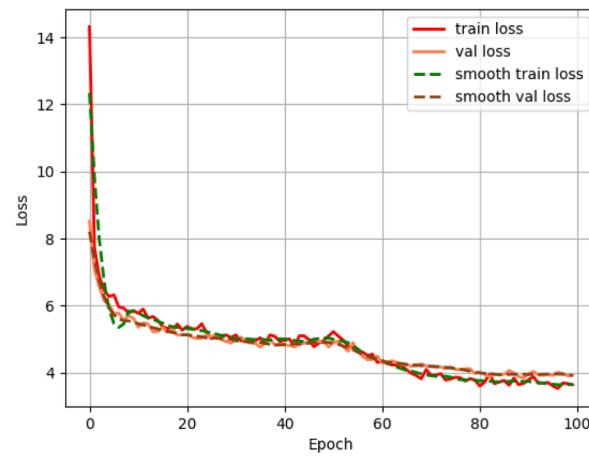


Fig. 19. Transfer learning detection results.



(a) Original algorithm



(b) Improved algorithm

Fig. 17. Loss function curve.

Table 7
Comparison experiment results.

Experiments	mAP/%	Macro-F1/%	Training time/s
1	87.63	78.36	1423
2	88.35	82.57	1221
3	86.46	81.32	1068

freezing a 10-network structure (Zhu Zhaotong et al., 2020). The experimental results are shown in Table 7.

From Table 7, the detection results of the network structure by freezing different layers are slightly different. Normally, the higher the number of frozen layers is, the higher the recognition accuracy and the faster the training speed will be. However, when frozen to 10 layers, the recognition accuracy began to decrease, which indicated that the convergence state has been reached before freezing 10 layers. Thus, the best result is achieved when the network structure is frozen at 5 layers, and a satisfactory result can be achieved.

6. Conclusion

The results of the present study indicate that the proposed improved SSD algorithm can be effectively used to accurately detect the surface defect of electrolytic capacitors. The conclusions of the proposed method are summarized as follows.

- 1) An electromechanical system that can achieve omnidirectional dynamic detection of the electrolytic capacitor surface was designed, and a dataset of six defect types (pin burn, pin broken, scratch, damage-s, damage-tb, and deformation) to train the network model was established;
- 2) Improving the original SSD algorithm by replacing VGG-16 with the MobileNetV2 network model, which significantly reduced the number of parameters in the network model and training time, and applying the Focal Loss function to the target detection algorithm to improve the detection accuracy of the algorithm;
- 3) Training and optimizing the parameters of a deep convolutional neural network-based model to construct a small-sample deep convolutional neural network-based electrolytic capacitor recognition model for small-sample transfer learning.

Declaration of competing interest

The authors declare that they have no known competing financial

interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Bochkovskiy A, Wang C Y, Liao H Y M. Yolov4: Optimal speed and accuracy of object detection[J]. arXiv preprint arXiv: 2004. 10934, 2020.
- Dzhunusbekov, E.J., Orazbayev, S.A., 2020. Electrolytic capacitor life time calculation under varying operating conditions[J]. Journal of Vibroengineering 22 (3), 721–734.
- Guangyu, S., Jun, C., Yizhuo, Z., 2020. Near infrared wood defects detection based on transfer learning[H]. Electric Machines and Control 24 (10), 159–166.
- Jiang, H.a., Liangcai, Z., Li Gongfa Zhaojie, J.u., et al., 2021. Learning for a Robot: Deep Reinforcement Learning, Imitation Learning, Transfer Learning[J]. Sensors 21 (4), 1278-1278.
- Liu, W., Anguelov, D., Erhan, D., et al., 2016. SSD: Single shot multibox detector[C]// European conference on computer vision. Springer, Cham, pp. 21–37.
- Mao Huimin, Wang Jianzhong, Jiang dong, et al. Analysis of causes and effects of white spot on anode foil of aluminum electrolytic capacitor[J]. China Plant Engineering, 2021(15): 138-139.
- Muhammad, I., Oualid, D., DeokJin, L., 2021. Transfer Learning Based Semantic Segmentation for 3D Object Detection from Point Cloud[J]. Sensors 21 (12), 3964-3964.
- Nikos, P., 2021. Measurement of Fish Morphological Features through Image Processing and Deep Learning Techniques[J]. Appl. Sci. 11 (10), 4416-4416.
- Qi, Y., Cai Jin Ji yuan, Q.I.U., et al., 2021. Image Recognition of Blade Surface Integrity Based on VGG-16 Convolutional Neural Network [J]. Aviation Precision Manufacturing Technology 57 (05), 4–8.
- Redmon J, Farhadi A. Yolov3: An incremental improvement[J]. arXiv preprint arXiv: 1804. 02767, 2018.
- Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]. Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 7263-7271.
- Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 779-788.
- Shuai, T., Xiaohai, H.u., 2018. Research on the Reliability of Tantalum Electrolytic Capacitor for Equipments[J]. Electronic Product Reliability and Environmental Testing 36 (01), 34–38.
- Wei, L.i., Kai, L., 2021. Confidence-Aware Object Detection Based on MobileNetv2 for Autonomous Driving[J]. Sensors 21 (7), 2380-2380.
- Weiwei, Z., Huimin, M.a., Xiaohong, L.i., et al., 2021. Imperfect Wheat Grain Recognition Combined with an Attention Mechanism and Residual Network[J]. Appl. Sci. 11 (11), 5139-5139.
- Xiaoting, G., 2019. Research on Rail Surface Defect Detection and Data Transmission Technology[D]. Shijiazhuang Tiedao University.
- Zhi, Q., Austin, B., Glass Lucas, M., et al., 2020. FLANNEL: Focal Loss Based Neural Network Ensemble for COVID-19 Detection. [J]. Journal of the American Medical Informatics Association: JAMIA 28 (3), 444-452.
- Zhu Zhaotong, Fu Xuezhi, Hu Youfeng. A Sonar Image Recognition Method Based on Convolutional Neural Network Trained through Transfer Learning, Journal of Unmanned Undersea Systems,2020,28(01):89-96.