J. Meas. Sci. Instrum., 2024, 15(1): 23-32.
DOI: 10.62756/jmsi.1674-8042.2024003
http://jmsi.publish.founderss.cn/
jmsi@nuc.edu.cn

Instance segmentation algorithm of electronic components based on improved YOLOv5

YANG Yining, WEI Honglei*

School of Mechanical Engineering and Automation, Dalian Polytechnic University, Dalian 116000, China

***Corresponding author:** WEI Honglei (weihl@dlpu.edu.cn)

Received: October 17, 2023 **Revised:** November 15, 2023 **Accepted:** January 24, 2024

Abstract: To address the challenge of automatic recognition of electronic components on an assembly line, an improved YOLOv5 was used to implement instance segmentation of four categories of electronic components. Firstly, multi-channel histogram equalization was used for image preprocessing. Then, the YOLOv5 was improved: Segmentation head was added; Sequeeze-and-excitation net (SE-Net) channel attention module was embedded to enhance the feature extraction capability and to compress the useless information without increasing the model complexity; GhostNet was used to make the model lightweight; and BiFPN was used to enhance model feature fusion capability. Finally, experimental results showed that the *mAP* of the proposed method could reach 96.7% and the detection time of a single image was 45.5 ms. The results prove that proposed method has superior performance than that based on mask region-based conventional neural network (Mask RCNN) and initial YOLOv5, and has practical significance for automatic detection of electronic components.

Key words: instance segmentation; deep learning; YOLOv5; components detection

0 Introduction

The electronics industry has been experiencing a rapid development, and more and more electronic components have been applied to various industrial sectors. Therefore, it is vital to rapid classification of different electronic components. In the conventional approach, such classification primarily relies on manual sorting, which proves to be a labor-intensive and time-consuming process and cannot meet the requirements of efficient production[1].

Owing to the development of machine vision techniques, numerous machine vision methods have been applied to object detection and classification of electronic components. For example, Lefkaditis et al.[2] developed an electronic components automated identification system based on the outline morphology of the components. Zeng et al.[3] used multi-level thresholding and color distribution features to identity components on PCB. Tsai et al.[4] proposed a global Fourier image reconstruction method to printed circuit board (PCB) identification. However, traditional machine vision methods have low accuracy and low efficiency[5]. In addition, lighting condition exerts a significant influence on it and image feature extraction needs to be performed manually. Furthermore, the random placement of electronic components an the assemble line make it difficult to set a general matching template.

Nowadays, deep learning is developing rapidly and the convolutional neural networks are widely used for automation extracton of image features based on their superior performance[6]. Scholars have used convolutional neural networks (CNNs) for recognition and classification of various parts, such as ship spare-parts[7], car parts[8], machine parts[9], mechanical components[10], and electronic components.For example, Varna et al.[11] applied YOLOv4 into detection of minimized moving electronic components. Similarly, Farbice et al.[12] used YOLOv4 to classify electronic components on PCB. Wu et al.[13] used YOLOv5 to realize the detection of PCB components. Atik[14] used pre-trained networks AlexNet, ShuffleNet, SqueezeNet, and GoogleNet to classify the electronic components. Wu et al.[15] detected the electronic components by the generative adversarial network (GAN) to generate training samples. In addition, Kuo et al.[16] proposed a graph convolutional network for component classification and detection on a PCB.

Object detection is a dominant approach. However, object detection only uses a rectangular box to display the approximate position of the target, but the contours of components cannot be segmented and positioned accurately. In addition, components may overlap in the production line. Therefore, in our work, we will combine

object detection with instance segmentation because it has proved to be successful in numerous fields[17-21].

Although the instance segmentation has yielded substantial advancements, several challenges persist. Notably, most of the above methods used mask region-based convolutional neural network (Mask RCNN)[22], which is an extension of Fast RCNN and a two-stage object detection algorithm. Since two-stage approach needs a long training and inference time as well as a heavy workload of quantitative computation, it is difficult to meet real-time detection requirements in industrial environments.

By comparison, one-stage detection algorithms have demonstrated noteworthy performance, one of which is YOLOv5, which not only achieves outstanding recognition results, but also boasts exceptional inference speed. YOLOv5 has also been successfully applied in various scenarios, such as camellia oleifera fruit detection and locating process[23], counting of white blood cells[24], power line intelligent inspection[25], and so on. Thus, combining YOLOv5 with semantic segmentation can lead to better instance segmentation. On the basis of the above-mentioned research, an instance segmentation algorithm using improved YOLOv5 was proposed. Firstly, the image preprocessing was implemented to enhance image quality. Then, YOLOv5 was modified by adding a segmentation head to concurrently execute object detection and segmentation task. After that, channel attention mechanism squeeze-and-excitation net (SE-Net) was added, which can make reasonable use of characteristic channel information and suppress useless information. Finally, GhostNet and bi-directional feature pyramid networks (BiFPN) were used to improve the model feature fusion capability and lightweight the model. Experimental results demonstrate the proposed method has a good effect of instance segmentation. Besides, it is well suited for industrial applications and has broad application prospect.

1 Image preprocessing

The images of electronic components were rapidly collected from the production line, which were characterized by relatively low image quality, resulting in image blurring and indistinct boundaries. Preprocessing was done to enhance the images and make the image boundaries sharper in preparation for the subsequent instance segmentation. In our work, a multichannel histogram equalization was used for image preprocessing.

1.1 Multichannel histogram equalization

Histogram equalization is an effective image enhancement technique[26]. By stretching the gray level with additional pixels in the image histogram and compressing the gray level with fewer pixels, the dynamic range of the difference in gray values between pixels is increased. In our work, the input image channels were separated and histogram equalization was performed in different channels. The abscissa of the histogram is the gray level of the image, the ordinate is the frequency of occurrence of each gray level. the histogram can be represented by a probability density function as

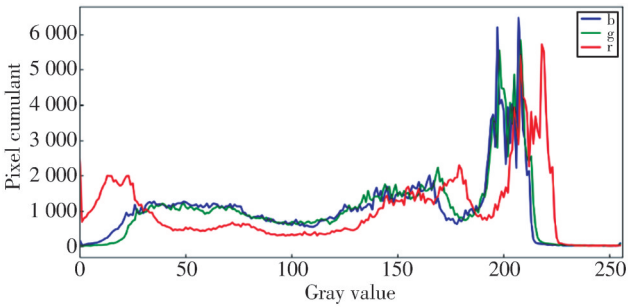
$$P_r(r_k)=\frac{n_k}{N}, \tag{1}$$

where n_k is the number of pixels on the gray level k of the image; r_k is the kth gray level; and N denotes the total number of pixels in the image.

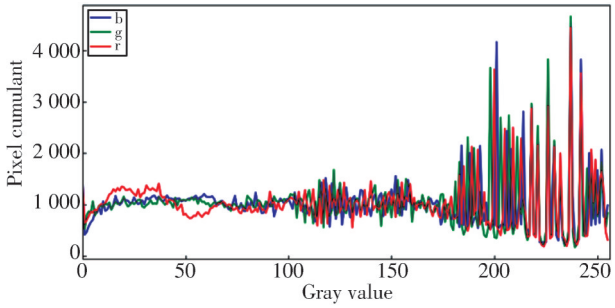
Then, the histogram equalization was performed by the cumulative distribution function of the probability of each gray level of the image, namely

$$S_k(r_k)=P_r(r_i). \tag{2}$$

Finally, the results after equalization for each channel were fused to complete the histogram equalization for multiple channels. The comparison of the histogram equalization is shown in Fig.1.



(a) Histogram before image preprocessing



(b) Histogram after image preprocessing

Fig. 1 Comparison of histograms before and after image preprocessing

Here, the red, green, and blue curves depict the grayscale histograms of three colors channels, respectively.

1.2 Image quality evaluation

To evaluate the effect of image preprocessing, the peak signal-to-noise ratio (PSNR), structural similarity (SSIM), Brenner function and sum of modulus of gray difference (SMD) function were introduced as evaluation indexes.

The equation for calculating the PSNR is expressed as

$$PSNR = \frac{G_t - G_b}{\sigma}, \tag{3}$$

where G_t is the maximum gray value of the image target area, G_b is the gray mean value of the image background area, and σ is the standard deviation of the gray value of pixels in the image background area. The SSIM is calculated by

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\mu_x^2 + \mu_y^2 + C_2)}, \tag{4}$$

$$C_1 = (k_1 l)^2, \tag{5}$$

$$C_2 = (k_2 l)^2, \tag{6}$$

where x represents the processed image; y is the original image; μ_x is the mean of x; μ_y is the mean of y; μ_x^2 is the variance of x; μ_y^2 is the variance of y; σ_{xy} is the covariance of x and y; C_1 and C_2 are the constants used to maintain stability; l is the dynamic range of pixel values; and k_1 and k_2 are set to be 0.01 here.

The SMD and Brenner function are two common blur evaluation functions, and their definitions are ignored.

The quality assessment results of the original image after preprocessing are shown in Table1.

Table 1 Image preprocessing effect

Image	PSNR	SSIM	Brenner	SMD
Original image	1	1	1	1
Preprocessed image	21.433	0.923	1.131	1.126

It can be seen that the preprocessed images exhibit minimal loss of brightness while well maintaining SSIM. Notably, PSNR has a significant enhancement, resulting in superior image quality. Furthermore, the SMD and Brenner functions demonstrate marked improvements to enhance image sharpness. Overall, this preprocessing significantly contributes to the refinement of image boundaries, ultimately yielding high-quality image.

2 Instance segmentation method based on improved YOLOv5

2.1 YOLOv5 structure

YOLOv5 model consists of four parts: input, backbone network, multi-scale feature fusion network neck and predictive classifier head. As shown in Fig.2, Firstly, the backbone network extracts image features and the neck component fuses these features. Then, the bottom-up mechanism of pass aggregation network (PANet) enriches the semantic information and compensatse for the missing information.

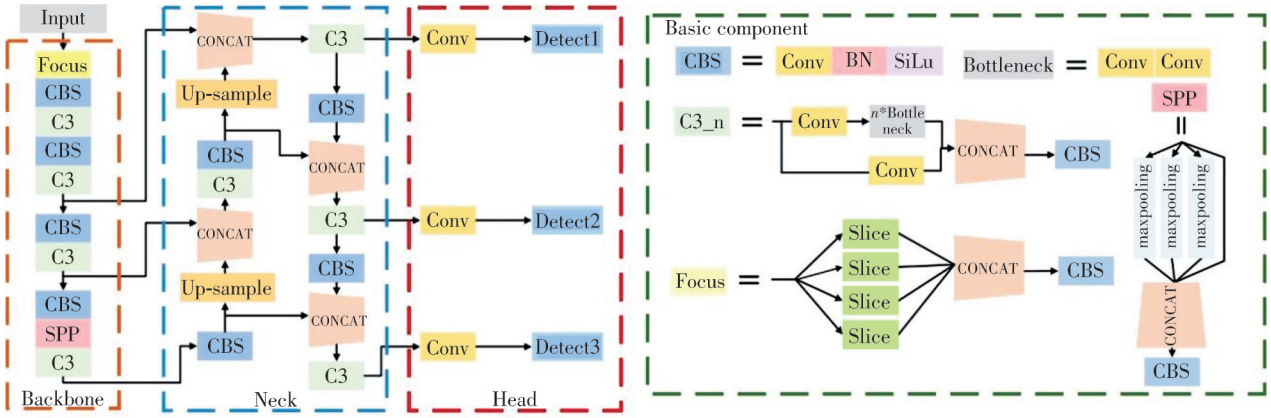


Fig. 2 YOLOv5 structure

Besides, the spatial pyramid pooling (SPP) layer increases the receptive field and enhances the nonlinear expressive power of the network by convolving the feature layers three times and maximizing the pooling. Lastly, the head part conducts detection utilizing 1×1 convolution kernel replacing the fully connected layer for prediction, and the object detection is performed on three scales.

2.2 Improved YOLOv5

On the assembly line, there are a large number of electronic components, covering a variety of sizes from small to large. This variation in the size of electronic components presents a challenge for accurate object detection. Considering that some tiny parts are easy to miss during detection and industrial production environment requires superior production efficiency, the

implementation of lightweight models for instance segmentation becomes a viable solution.

According to the characteristics of electronic components, the initial YOLOv5 model faces limitations in meeting precise detection requirements. Thus, to address these challenges, we improved the initial YOLOv5. Firstly, we added segmentation heads to implement instance segmentation when object detection. Secondly, we adopted BiFPN to facilitate multi-scale feature fusion, thereby strengthening the capability to detect small targets. Additionally, we introduced channel attention mechanism SE-Net to augment features extraction and model detection performance. Finally, GhostNet was used to lightweight the network and facilitate implementation in an industrial environment.

The improved YOLOv5 structure is shown in Table 2. The number in the second column indicates the input layer number, and −1 represents the input is from the previous layer.

Table 2 Improved YOLOv5 structure

Layer No.	Input layer No.	Module
0	−1	Conv
1	−1	Conv
2	−1	C3Ghost
3	−1	Conv
4	−1	C3Ghost
5	−1	Conv
6	−1	C3Ghost
7	−1	Conv
8	−1	C3Ghost
9	−1	SE-Net
10	−1	SPPF
11	−1	Conv
12	−1	Up-sampling
13	[1, 6]	BiFPN_Concat2
14	−1	C3
15	−1	Conv
16	−1	Up-sampling
17	[1, 4]	BiFPN_Concat2
18	−1	C3
19	−1	Conv
20	[−1, 14, 6]	BiFPN_Concat3
21	−1	C3
22	−1	Conv
23	[1, 10]	BiFPN_Concat2
24	−1	C3
25	[18, 21, 24]	Segmentation

2.2.1 Segmentation head

The initial YOLOv5 model was exclusively designed for object detection. To implement instance segmentation, a segmentation head was added in YOLOv5 to implement instance segmentation when object detection. Thus, there are four loss values in model training: classification loss, object loss, bounding box loss and segmentation loss. The total loss is the sum of four types of loss. The segmentation head takes the output of the 16th layer of YOLOv5 as input and it contains C3, SPP and up-sampling modules. The structure of segmentation head is shown in Fig.3.

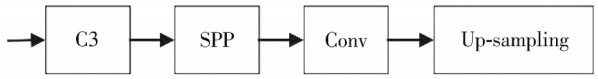


Fig. 3 Segmentation head

2.2.2 Embedded attention mechanism SE-Net

Squeeze-and-excitation net (SE-Net) [27] is a channel attention mechanism. It was embedded into YOLOv5 here. The function of SE-Net is to enhance critical features and suppress generic features. It can learn the importance of feature channels independently without changing the original spatial dimension and improve the model performance by boosting the weights of more effective feature information based on the importance of feature channels. The SE-Net mechanism consists of three parts: squeeze, excitation, and reweight operation, as shown in Fig. 4. The process of SE-Net mechanism is as follows.

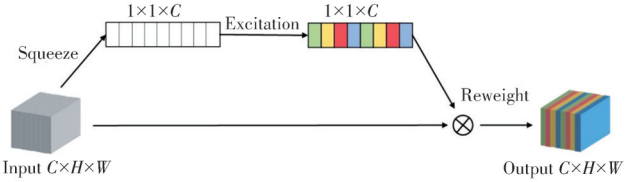


Fig. 4 SE-Net structure

1) Squeeze operation

The feature map with the size of $C \times W \times H$ is compressed into a feature map with the size of $1 \times 1 \times C$ by using average pooling, that is

$$F_{sq}(C)=\frac{1}{HW}\sum_{i=1}^{H}\sum_{j=1}^{W}C(i,j), \tag{7}$$

where H and W are the height and width of the feature map, respectively; C is the channel of the feature map; and i and j are the pixel setting variables.

2) Excitation operation

Excitation operation uses two fully connected layers to assign different weights for each channel, and the weights and channel correlations of different feature channels are represented by the parameter W as

$$F_{ex}(Z,W)=\sigma(W_2\delta(W_1,Z)), \tag{8}$$

where Z is the output of the squeeze operation, W_1 and W_2 are the channel weights, σ is the activation function, and δ is the normalization function.

3) Reweight operation

The weight vector generated is multiplied by the $C \times H \times W$ feature map, which is assigned a weight to get a new feature map with exactly the same size as before. S is the excitation result and expressed as

$$F_{scale}(C,S)=SC. \tag{9}$$

2.2.3 Model lightweight

Since YOLOv5 model need to deal with a large quantity of parameter counting and computation during convolutional operations, which may result in slow model training and inference and cannot meet the requirements of rapid training model and inference in industrial production environment. To make the model lightweight, the C3 module in backbone net is replaced by GhostNet[28], thus improvieng the model inference speed. GhostNet module is shown in Fig.5.

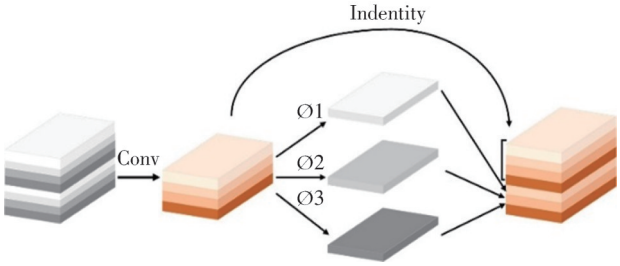


Fig. 5 GhostNet

GhostNet module splits the ordinary convolution into two parts. Firstly, 1×1 convolution is used to obtain the necessary feature concentration of input features. Then, layer-by-layer convolution is used to stack the convolution layers. Finally, the characteristic concentration obtained in the previous step is used to produce a similar feature map.

2.2.4 BiFPN feature fusion network

Since feature pyramid network (FPN) only transmits features through a top-down path, the feature information may be lost and the precision is not high. In YOLOv5, PANet is used. PANet has both top-down path and down-top path, therefore, it can fuse the top and bottom feature information. However, this method can improve the detection accuracy whereas it is prone to missing tiny targets. To address this problem, we repalced PANet with BiFPN[29] here. The structures of FPN, PANet and BiFPN are shown in Fig.6.

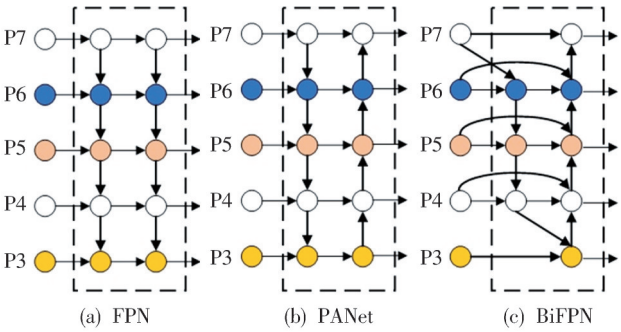


Fig. 6 Structures of FPN, PANet and BiFPN

Compared with the structure of FPN in Fig. 6(a), PANet adds an opposite path based on FPN to convey the lost location information, as shown in Fig.6(b). BiFPN is a weighted bidirectional feature pyramid network proposed by Google, as shown in Fig.6(c). It removes the nodes with only one input based on PANet to simplify the network without losing vital information. At the same time, an additional edge is added to integrate additional features when the input node and the output node are located in the same layer. More importantly, BiFPN uses a weighted feature fusion mechanism to assign a reasonable weight to each path where the data pass and then uses the learned features to continuously update the weights to obtain more valuable information.

2.3 Framework of instance segmentation method

In this study, the electronic components' image information was gathered on the production line, and then the image preprocessing based on multi-channel histogram equalization was implemented to make the boundary of the electronic components clear and easy to recognize. Next, the dataset was divided into training set, validation set and test set at a ratio of 8∶1∶1. After that, the dataset was sent to the improved YOLOv5 to train the model. Finally, we implemented the instance segmentation of electronic components. The instance segmentation results contain the object detection and semantic segmentation of electronic components. The overall framework of our method is shown in Fig.7.

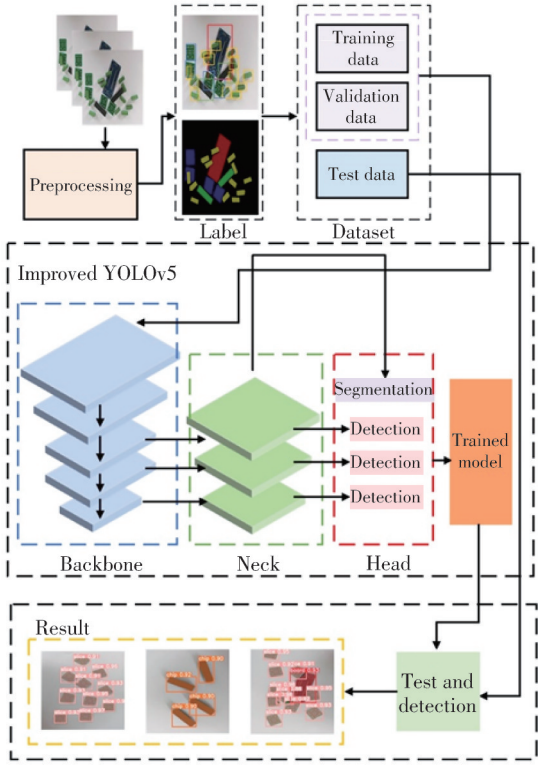
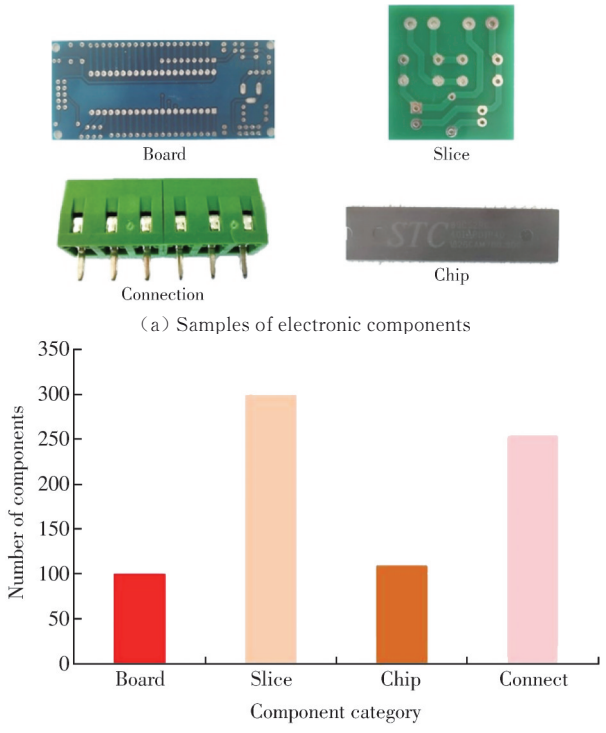


Fig. 7 Framework of our method

3 Experiment and analysis

3.1 Dataset acquisition and experimental environment

The image information was gathered on the production line, and the image was saved as JPEG format. In the experiments, a total of 516 images were collected as the dataset. The training set, validation set and test set were divided with a ratio of 8∶1∶1. There were four categories of electronic components in the dataset: chip, electronic board, electronic slice and electric connection piece, as shown in Fig. 8(a). The quantity distribution of electronic components is shown in Fig. 8(b). Besides, before model training, each electronic component instance on the image in the dataset was labeled individually.



(a) Samples of electronic components

(b) Quantity distribution of electronic components

Fig. 8 Electronic component samples and quantity distribution

In the experiments, the computer with Windows10 system, the CPU with AMD Ryzen 74 800 Hz with 16 GB running memory, and the CUDA11.1.114 parallel computing framework with graphics card NVIDIA GTX1650 were used.

3.2 Model training

The training epoch was set to be 100, learning rate was set to be 0.01, batch size was set to be 1, and the parameter optimizer adopted SGD. Each image was resized to the size of 640×640 pixels before being fed into the model. The deep learning framework used PyTorch. The loss during training mainly results from segmentation, classification, rectangular box and object detection, as shown in Fig. 9. The total loss is equal to the sum of four parts, as shown in Fig.10.

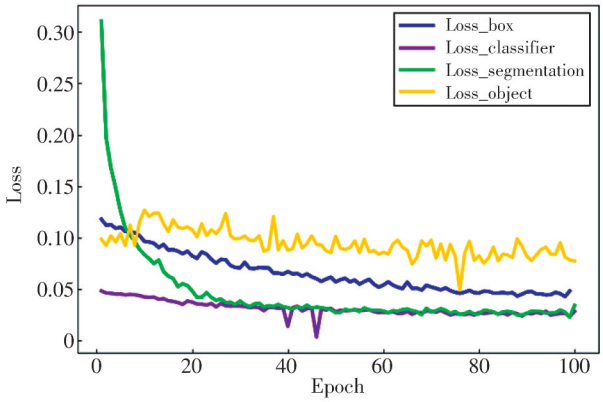


Fig. 9 Losses of four parts

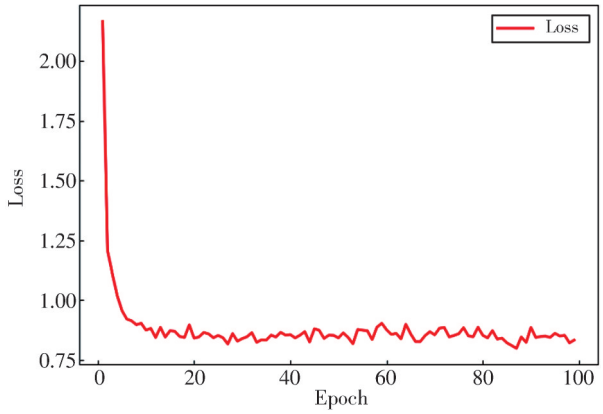


Fig. 10 Total loss

In Fig.9, it is evident that the loss of segmentation shows a significant reduction, reaching a stable state after approximately 40 training rounds, which indicates that the training effect is well and the loss value finally stabilizes at a low level. Furthermore, substantial reduction can be observed in the loss of rectangular boxes. The loss of classification is maintained at a low level at all time, indicating the well-trained performance of classification. The loss of object detection decreases with a slight fluctuation due to the distinctly discriminative features on different types of components. In Fig.10, the progressive increase in training epochs yields a gradual reduction total loss, and it finally decreases to about 0.2, which is a relatively low level. The loss decreases rapidly and steadily throughout the training process, which indicates that the model is well trained.

3.3 Comparative experiments and analysis

3.3.1 Evaluation metrics

The main evaluation metrics of the experiments are

precision (P), recall (R), mean average precision (mAP) and *Params*.

P refers to the proportion of predictions that are correct for all positive sample results and is calculated by

$$P=\frac{TP}{TP+FP}, \tag{10}$$

where TP is the number of positive samples predicted correctly, and FP is the number of negative samples with prediction errors.

R refers to the ratio of correctly predicted positive samples to all true positive samples and is calculated by

$$R=\frac{TP}{TP+FN}, \tag{11}$$

where FN is the number of positive samples with prediction errors.

AP refers to the region below the PR curve. It averages the accuracy values at different recall points and is calculated by

$$AP=\int_0^1 P(r)\mathrm{d}r, \tag{12}$$

mAP is the mean average precision and is calculated by

$$mAP=\frac{\sum_{i=1}^{N} AP_i}{N}, \tag{13}$$

where N is the number of classes, and $mAP@0.5$ represents mAP value in case of confidence 0.5.

Besides, evaluation metrics *Params* is the number of model parameters and it used to measure the size of the model.

3.3.2 Comparative experiments of different attention mechanisms

To verify the performance and lightness of the channel attention mechanism, for SE-Net, we compared it with two common attention mechanisms, coordinate attention (CA) and convolutional block attention module (CBAM). The experiment results are shown in Table 3.

Table 3 Comparison of three attention mechanisms

Method	P/%	R/%	$mAP@0.5$/%	*Params*/M
YOLOv5+CBAM	92.5	71.3	98.5	7 813 651
YOLOv5+CA	82.9	73.6	86.2	7 806 433
YOLOv5+SE	83.4	96.4	91.4	7 473 281

It can be seen from Table 3 that based on YOLOv5, SE-Net performs better than CA, with P value of 83.4% and R value of 96.4%, which demonstrates that SE-Net has superior performance in comparison to CA and CBAM. Additionally, the mAP with SE-Net exceeds that with CA, but slightly trails behind that with CBAM whereas CBAM considerably increases the number of parameters, which is adverse to model lightweight. After comprehensive consideration, SE-Net was adpoted as the preferred choice since it has comprenhensively better performance and does not take too much computation resource.

3.3.3 Ablation experiments

To verify the effectiveness of various improvement strategies for YOLOv5, ablation experiments were conducted and the results are listed in Table 4.

Table 4 Ablation experiment

Method	P/%	R/%	$mAP@0.5$/%	*Params*/M
YOLOv5	66.8	91.3	80.5	7 406 513
YOLOv5+SE	83.4	96.4	91.4	7 473 281
YOLOv5+BiFPN	79.3	94.8	93.4	7 472 058
YOLOv5+GhostNet	66.9	99.0	95.4	6 244 025
YOLOv5+GhostNet+SE	83.6	100	95.6	6 276 793
YOLOv5+GhostNet+SE-net+BiFPN	85.7	98.7	96.7	6 506 178

It can be seen that the incorporation of SE-Net leads to a reduction in the number of total model parameters while significantly yields substantial improvements in P, R, and mAP. The integration of BiFPN within the model not only augmentes its ability to capture inherent semantic information, but also facilitates the extraction of semantic features from the same layer. Thus, BiFPN feature fusion within the network significantly enhances the performance of improved YOLOv5. In addition, the incorporation of GhostNet improves R and markedly reduces the number of model parameters so as to lightweight the model. Ultimately, these enhancements culminate in a well-behaved model performance, with a P of 85.7%, an R of 98.7% and an mAP of 96.7%. Notably, the *Params* of improved model is substantially smaller compared to that of the initial YOLOv5. Therefore, ablation experimental results show that the improvements of YOLOv5 enhance the model performance and finally achieve a balance between performance and lightweight.

3.3.4 Comparasion of different models

To verify the detection performance of the improved YOLOv5, we conducted a comparasion of the initial YOLOv5, Mask RCNN and the improved YOLOv5. The results are shown in Table 5.

Table 5 Comparison of different models

Method	P/%	R/%	$mAP@0.5$/%	Single image detection time/ms
YOLOv5	66.8	91.3	74.5	35.2
Mask RCNN	83.2	81.5	70.5	170.0
Proposed method	85.7	98.7	96.7	45.5

It can be seen that the proposed method outperforms Mask RCNN by 0.5% in P value, exhibits substantially

higher R value compared to that of the other two methods, and more importantly, the mAP value is much higher than that of the other two methods. This indicates that the proposed method can efficiently identify four electronic components. Mask RCNN is an improved instance segmentation method based on two-stage object detection method Fast RCNN, and performs well in precision. However, since it requires the generation regions of interest firstly and then arranges them, the detection time of single image is much more than that of YOLOv5 and the proposed method. In summary, the proposed method can efficiently implement instance segmentation of electronic components with fast detection speed while guaranteeing detection accuracy. Thus, the proposed method meets the requirements of industrial production.

Besides, the detection results of three methods are shown intuitively in Fig.11.

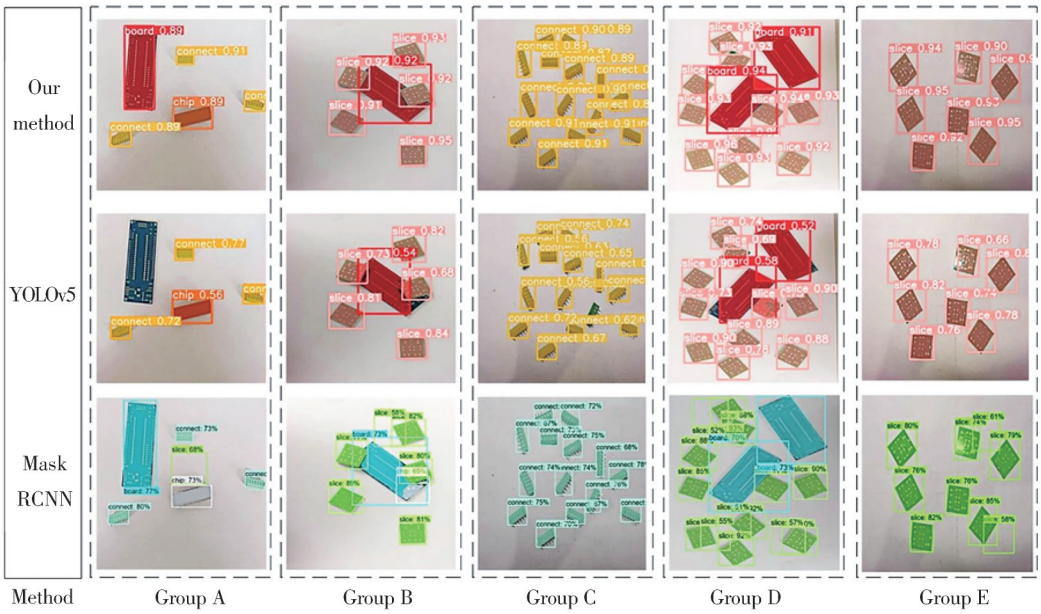


Fig. 11 Comparison of three methods for instance segmentation of electronic components

It can be seen from Fig.11 that in Group A, the initial YOLOv5 with segmentation head fails to detect the boards, while the Mask RCNN has an additional detection box, which is an error recognition. Meanwhile, the confidence of the proposed method is higher than that of the other two methods. In Group B, the initial YOLOv5 with segmentation head cannot segment the overlap of boards and slices precisely, and the segmentation is not exhausive enough. Besides, the Mask RCNN method incorrectly detects the overlap of boards and slices as chips. In Group C, the initial YOLOv5 with segmentation head misses a tiny object connection while the detection effect for tiny object using the proposed method has been better improved. The confidence of our method is also higher than that of other two methods. In Group D, the Mask RCNN method cannot precisely detect overlapping slices, and the segmentation of the board using the initial YOLOv5 is incomplete and not accurate enough. In Group E, Mask RCNN also suffers from false detection and the initial YOLOv5 fails to segment the slices completely, while our method achieves better performance.

3.3.5 Comparison of similar researches on electronic components detection

To verify the advantages of the proposed method for the recognition of electronic components, the proposed method was compared with similar research on electronic components detection, and the comparison results are shown in Table 6.

Table 6 Comparison of similar researches

Refs.	Method	P/%	mAP@ 0.5/%	$Params$/ M
Ref. [30]	Binary shape analysis	84.5	—	—
Ref. [14]	CNN	98.8	—	—
Ref. [31]	Improved YOLOv3	—	89.9	61.73
Proposed method	Improved YOLOv5	85.7	96.7	6.20

In Ref. [30], binary shape analysis, a typical traditional machine vision method, was used for electronic components classification and recognition by calculation of the similarity of the shape parameters for binary images. Its P value is 84.5% lower than that of our method by 1.2%. However, it is difficult to meet the fast production environment due to low efficiency of manual feature extraction. In Ref. [14], an improved CNN was used. Although it achieves a P value of

98.8%, it can classify images only when there is one electronic component in one picture at a time, therefore, it cannot meet the practical production demands where there are a diversity of electronic components intermingled. In Ref.[31], an improved YOLOv3 was used. Its *mAP* is 89.9% lower than that of proposed method by 6.8%. However, its model parameter reaches 61.7 MB, which is nearly ten times larger than that of our method and does not favor the lightweight of the model. Besides, it only detects the electronic components and cannot segment the instance. In summary, the proposed method is able to well implement the instance segmentation of electronic components on the assembly line, meeting the detection requirements of the production and exhibiting certain advantages over similar researches.

4 Conclusions

To tackle the challenge of classifying and detecting different electronic components in a production line, we proposed on improved YOLOv5 for the purpose of instance segmentation of electronic components. The segmentation head was added for segmentation, SE-Net channel attention mechanism was embedded, GhostNet was added and BiFPN was applied to enhance network feature fusion capability. Ablation and comparison experiments demonstrate that the proposed method can efficiently and accurately detecte electronic components while balancing lightweight and performance of the model. Due to a large number types of electronic components, subsequent work will increase the types of components for training and improve the robustness of the model.

Acknowledgement

This work was supported by 2021 General Project Scientific Research Funds of Education Department of Liaoning Province (Nos. LJKZ0535, LJKZ0526)

Declaration of conflicting interests

The authors have no conflict of interests related to this publication.

References

[1] DOS SANTOS M M, SILVA FILHO A G D, DOS SANTOS W R. Deep convolutional extreme learning machines: Filters combination and error model validation. Neurocomputing, 2019, 329: 359-369.

[2] LEFKADITIS D, TSIRIGOTIS G. Intelligent optical classification system for electronic components. Elektronika ir Elektrotechnika, 2010, 98(2): 10-14.

[3] ZENG Z, MA L Z, ZHENG Z Y. Automated extraction of PCB components based on specularity using layered illumination. Journal of Intelligent Manufacturing, 2011, 22: 919-932.

[4] TSAI D M, HUANG C K. Defect detection in electronic surfaces using template-based Fourier image reconstruction. IEEE Transactions on Components, Packaging and Manufacturing Technology, 2018, 9(1): 163-172.

[5] WEI B, HU L, ZHANG Y, et al. Parts classification based on PSO-BP//IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference, June 12-14, 2020, Chongqing, China. New York: IEEE, 2020, 1: 1113-1117.

[6] HAN H, LI Y, ZHU X. Convolutional neural network learning for generic data classification. Information Sciences, 2019, 477: 448-465.

[7] YANG K, YANG T, YAO Y, et al. A transfer learning-based convolutional neural network and its novel application in ship spare-parts classification. Ocean & Coastal Management, 2021, 215: 105971.

[8] LIN C H, YU C C, CHEN H Y. Augmentation dataset of a two-dimensional neural network model for use in the car parts segmentation and car classification of three dimensions. The Journal of Supercomputing, 2022, 78(17): 18915-18958.

[9] NING F, SHI Y, CAI M, et al. Various realization methods of machine-part classification based on deep learning. Journal of Intelligent Manufacturing, 2020, 31: 2019-2032.

[10] ZHANG M. Part classification prediction based on convolutional neural network. Mobile Information Systems, 2022: 5767818.

[11] VARNA D, ABROMAVICIUS V. A system for a real-time electronic component detection and classification on a conveyor belt. Applied Sciences, 2022, 12(11): 5608.

[12] FABRICE N, LEE J J. SMD detection and classification using YOLO network based on robust data preprocessing and augmentation techniques. Journal of Multimedia Information System, 2021, 8(4): 211-220.

[13] WU J H, YAN X Y, GE L S. Fast PCB defects detection method based on improved YOLOv5. Journal of Measurement Science and Instrumentation, 2023, 14(3): 340-349.

[14] ATIK I. Classification of electronic components based on convolutional neural network architecture. Energies, 2022, 15(7): 2347.

[15] WU H, LYU Q, YANG J, et al. Electronic component detection based on image sample generation. Soldering & Surface Mount Technology, 2022, 34(1): 1-7.

[16] KUO C W, ASHMORE J D, HUGGINS D, et al. Data-efficient graph embedding learning for PCB component detection//2019 IEEE Winter Conference on Applications

of Computer Vision, January 7, 2019, Waikoloa, USA. New York: IEEE, 2019: 551-560.
[17] LIU X, ZHAO D, JIA W, et al. Cucumber fruits detection in greenhouses based on instance segmentation. IEEE Access, 2019, 7: 139635-139642.
[18] OJHA A, SAHU S P, DEWANGAN D K. Vehicle detection through instance segmentation using mask R-CNN for intelligent vehicle system//2021 5th International Conference on Intelligent Computing and Control systems, May 6, 2021, Madurai India. New York: IEEE, 2021: 954-959.
[19] ZUO L, HE P, ZHANG C, et al. A robust approach to reading recognition of pointer meters based on improved Mask RCNN. Neurocomputing, 2020, 388: 90-101.
[20] KHAN M A, AKRAM T, ZHANG Y D, et al. Attributes based skin lesion detection and recognition: A Mask RCNN and transfer learning-based deep learning framework. Pattern Recognition Letters, 2021, 143: 58-66.
[21] HOU S, DONG B, WANG H, et al. Inspection of surface defects on stay cables using a robot and transfer learning. Automation in Construction, 2020, 119: 103382.
[22] HE K, GKIOXARI G, DOLLAR P, et al. Mask R-CNN//The IEEE International Conference on Computer Vision, October 22, 2017, Venice, Italy. New York: IEEE, 2017: 2961-2969.
[23] CHEN S, ZOU X, ZHOU X, et al. Study on fusion clustering and improved YOLOv5 algorithm based on multiple occlusion of camellia oleifera fruit. Computers and Electronics in Agriculture, 2023, 206: 107706.
[24] ESCOBAR F I F, ALIPO-ON J R T, NOVIA J L U, et al. Automated counting of white blood cells in thin blood smear images. Computers and Electrical Engineering, 2023, 108: 108710.
[25] LIU M, LI Z, LI Y, et al. A fast and accurate method of power line intelligent inspection based on edge computing. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-12.
[26] CHOI W, HUH H, TAMA B A, et al. A neural network model for material degradation detection and diagnosis using microscopic images. IEEE Access, 2019, 7: 92151-92160.
[27] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 42(8): 2011-2023.
[28] HAN K, WANG Y H, TIAN Q, et al. GhostNet: More features from cheap operations//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 14, 2020, Seattle, USA. New York: IEEE, 2020: 1577-1586.
[29] TAN M X, PANG R M, LE Q V. EfficientDet: scalable and efficient object detection//2020 IEEE/ CVF Conference on Computer Vision and Pattern Recognition, June 14, 2020, Seattle, USA. New York: IEEE, 2020: 10778-10787.
[30] KAMIL M, KRZYSZTOF O. Application of multi-descriptor binary shape analysis for classification of electronic parts. Journal of Universal Computer Science, 2020, 26(4): 479-495.
[31] JING L, WEIYE L, YING Q C, et al. Research on object detection of PCB assembly scene based on effective receptive field anchor allocation. Computational Intelligence and Neuroscience, 2022: 536711.

基于改进YOLOv5的电子元器件实例分割算法

杨祎宁，魏鸿磊*

大连工业大学 机械工程与自动化学院，辽宁 大连 116000

摘　要：为解决电子元器件组装中不同类型电子元器件难以自动区分的问题，利用改进的YOLOv5网络对电子元器件进行了实例分割，实现了不同元器件的自动识别分类。首先，使用三通道直方图均衡化对图像进行预处理。其次，在不增加模型复杂度的前提下，使用SE-Net通道注意力模块增强网络的特征提取能力，压缩无用信息；利用GhostNet实现网络轻量化；采用BiFPN增强网络特征融合能力。实验得出，采用改进的YOLOv5方法对电子元器件实例分割，其*mAP*为96.7%，单图的检测时间平均为45.5 ms. 试样结果表明，该实例分割方法优于同类方法，对提高电子元件的自动化检测水平具有实用意义。

关键词：实例分割；深度学习；YOLOv5；元器件识别

引用格式：YANG Yining, WEI Honglei. Instance segmentation algorithm of electronic components based on improved YOLOv5. Journal of Measurement Science and Instrumentation, 2024, 15(1): 23-32.