ISSN: 2414-1895 DOI: 10.6919/ICJE.202211_8(11).0033

Deep Learning-based Chip Identification of Discarded Circuit Boards

Zhipeng Wang, Jin Chen, Liying Yang, Zhongxiang Wang

School of Shanghai Polytechnic University, Shanghai, 201209, China

Abstract

At present, China's semiconductor industry is developing rapidly, the annual scrap of electronic products is growing rapidly, resulting in a large number of environmental pollution and waste of resources, it is urgent to develop efficient disposal and utilization technologies and methods, in order to recycle the metal materials of the circuit board, especially those electronic components rich in gold, and other rare precious metals, if they can be dismantled separately, centralized purification treatment, will effectively improve the disposal efficiency and also reduce the amount of chemical reagents in the later purification and reduce environmental pollution. However, due to the many types of electronic components on the circuit board, the large target scale range, the complex background and other characteristics, the traditional recognition and positioning algorithm in the automatic identification and positioning of electronic components appear in the process of large amount of calculation, low accuracy, poor algorithm portability and other issues. Aiming at this problem, this paper proposes a method for automatic identification and positioning of electronic components of waste circuit boards based on deep learning. Through the construction of a deep learning network, the detection and recognition model of electronic components is constructed, and the automatic identification and positioning of electronic components of waste circuit boards is realized in combination with the camera calibration method, so that the recognition and positioning effect has been significantly improved.

Keywords

Waste Circuit Boards; Chips; Deep Learning; Identification and Positioning.

1. Introduction

In 2006, Hinton[1]et al. first proposed the academic concept of deep belief network (DBN), which is a new branch of machine learning, and the goal of DNN is to simulate human brain function, be able to effectively identify images, words, sounds and other information, and complete the analysis and learning of information. Let deep learning be further developed in theory and technology.

In 2012, Hinton[2]et al. proposed the deep structure of Alex Net and used the ReLU activation function for the first time to solve the gradient disappearance problem and use GPUs to improve the operation speed, which promoted the development of the field of image recognition.

Around 2015, there were many good target detection network models based on convolutional neural networks and candidate region algorithms, and there were new breakthroughs in the field of object detection research. For example, He[3]et al. proposed the Res Net algorithm, and the error rate on TOP-5 is only 3.75%, which alleviates the gradient disappearance problem. After that, He[4]et al. optimized algorithms such as Fast-RCNN, which can also complete tasks such as object detection and instance segmentation on the basis of recognizing a large amount of information, so that image recognition technology develops rapidly. Since then, new convolutional neural network models and

ISSN: 2414-1895 DOI: 10.6919/ICJE.202211_8(11).0033

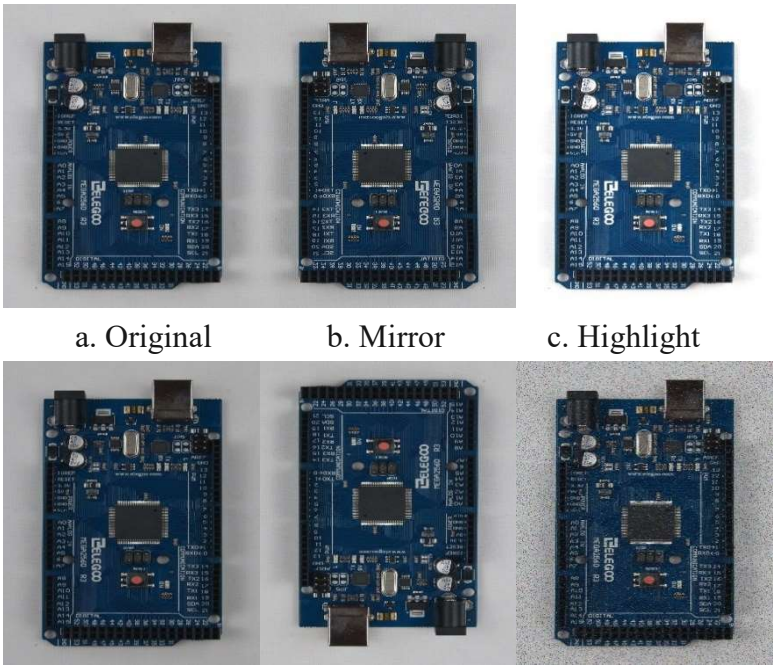
algorithms have been produced, such as VGG-Net[5], Google designed and developed GoogLe-Net[6] and so on.

These object detection network model algorithms are mainly divided into two categories: one is the RCNN, Fast-RCNN, faster-RCNN represented by the RCNN, which are all object detection algorithm models based on the Region Proposal; The other type is YOLO[7], SSD[8]etc., which are all object detection algorithm models based on regression thinking.

2. Image Acquisition and Data Production

2.1 Image Acquisition

In order to further increase sample diversity, control sample size, and reduce model training time, 500 images are selected for data enhancement through blur, brightening, dim, rotation and noise, removing the unclear pictures after enhancement, and finally retaining 2000 pictures. The dataset is divided in an 8:2 ratio, with 1600 images as the training set and 400 as the test set.



a. Original b. Mirror c. Highlight

d. Dim e. Rotation f. Noise

Figure 1. Original image and data enhanced image

2.2 Dataset Production

Use image labellmg software labeling to annotate pictures, imitate the VOC2007 dataset pattern, and create 2 folders and a txt file. The JPEGImages folder stores the pictures that need to be labeled, and the Annotations folder automatically saves the xml file after the label, and predefined_classes.txt file into the class name that needs to be labeled.

3. Identification of Discarded Circuit Boards based on Yolov5

3.1 Introduction to YOLOv5 Object Detection Algorithm

YOLOv5 object detection algorithm is the fifth version of YOLO, its core idea is to take the entire map as the input of the network, in the output layer directly return to the location coordinates and categories of the target, which is characterized by high detection accuracy, fast detection speed, to meet the needs of real-time monitoring. YOLO is a single-stage object detection algorithm, and the YOLOv5 model is divided into four parts: Input, Backbone, Neck, and Prediction.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202211_8(11).0033

3.1.1 Input

Input includes Mosaic data enhancements, automatic anchor box calculations, and adaptive image scaling. Mosaic Data Enhancement randomly stitches 4 images together, enriching the background of the picture, adding a lot of small goals, and making the network more robust. The ability to calculate anchor boxes is added to the code in YOLOv5, eliminating the need to manually calculate anchor boxes in the dataset, and automatically calculating the best anchor box value in the training set at each training session. YOLOv5's adaptive image scaling is to improve the inference speed, adaptive image scaling compared to unified scaling, can reduce the amount of image redundancy information and computation, inference speed has also been improved.

3.1.2 Backbone

Backbone modules are mainly composed of Conv structure, CSP and SPPF structure. The Conv structure includes Conv2d two-dimensional convolution, BatchNormal, SiLU activation functions. The Conv structure is responsible for feature extraction on the feature map and changing the number of feature map channels. The CSP structure mainly reduces the amount of neural network computation by reducing gradient information duplication in network optimization. The CSP structure of YOLOv5 has two different structures, one with residual structure and one without residual structure. The main role of SPPF is to enhance the expression ability of feature maps.

3.1.3 Neck

The Neck part adopts the FPN + PAN structure, the FPN is from the top down, the feature layer is downsampled, the high-level strong semantic features are passed down, so that the features at all scales have rich semantic information. After the multi-layer network in the middle of the FPN, the target information at the bottom has been very blurred, and the PAN has added a bottom-up route to upsample the feature layer to compensate for and strengthen the positioning information.

3.1.4 Prediction

Prediction uses two-dimensional ordinary convolution, which uses CIoU_loss to calculate bit losses, and BCEWithLogits_loss to calculate confidence losses and classification losses. Calculation of CIoU_Loss and BCEWithLogits_Loss.

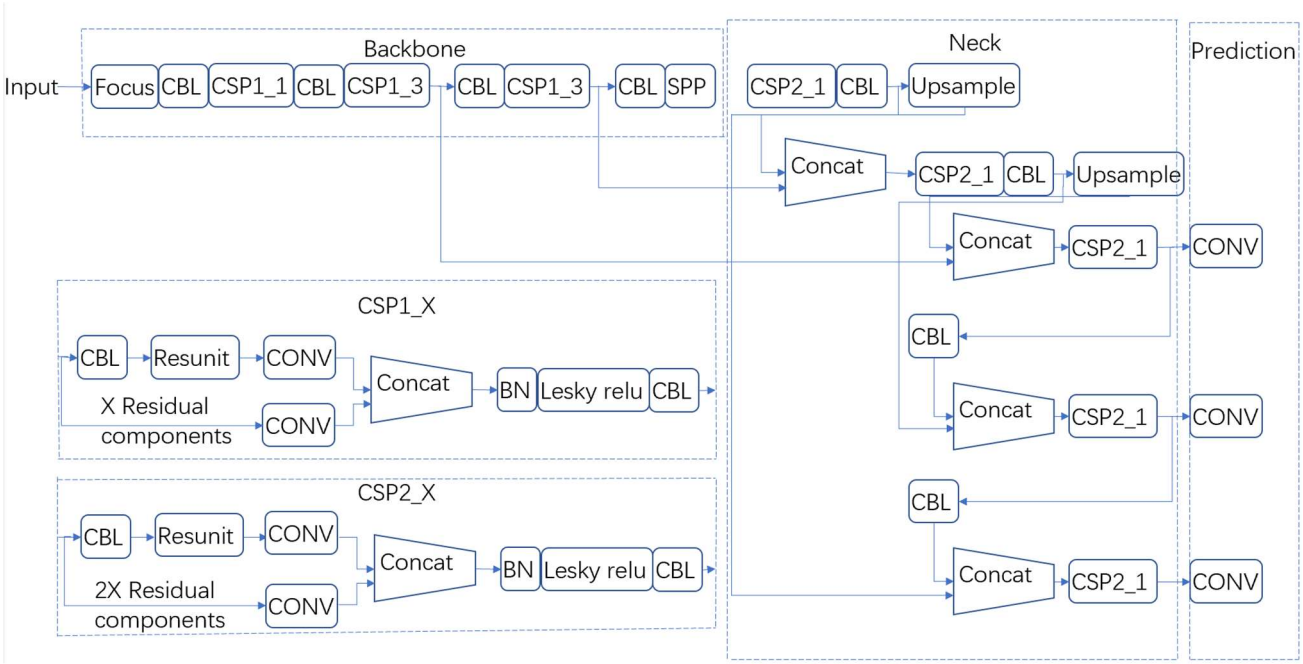


Figure 2. YOLOv5 network structure

ISSN: 2414-1895 DOI: 10.6919/ICJE.202211_8(11).0033

4. Experimental Analysis

4.1 Experimental Environment

Hardware system: abandoned circuit board SMD component removal platform based on the construction, mainly including general-purpose notebook computers, industrial cameras (JHSM500f), JHLD11808-5M lens, light source and bracket, USB transmission line, camera bracket and diffuse background paper, computer (operating system: windows 10, CPU: R7-5800H 16G memory, GPU: RTX3050, acceleration environment for CUDA11.4, python3.8, deep learning framework pytorch1.11).

4.2 Network Training

The pre-improved and post-improved algorithms employ the same hyperparameters. The learning rate is 0.01, the momentum is 0.937, the weight decay is 0.0005, the optimizer is Adam, the batch size is 90, and the number of turns is 200.

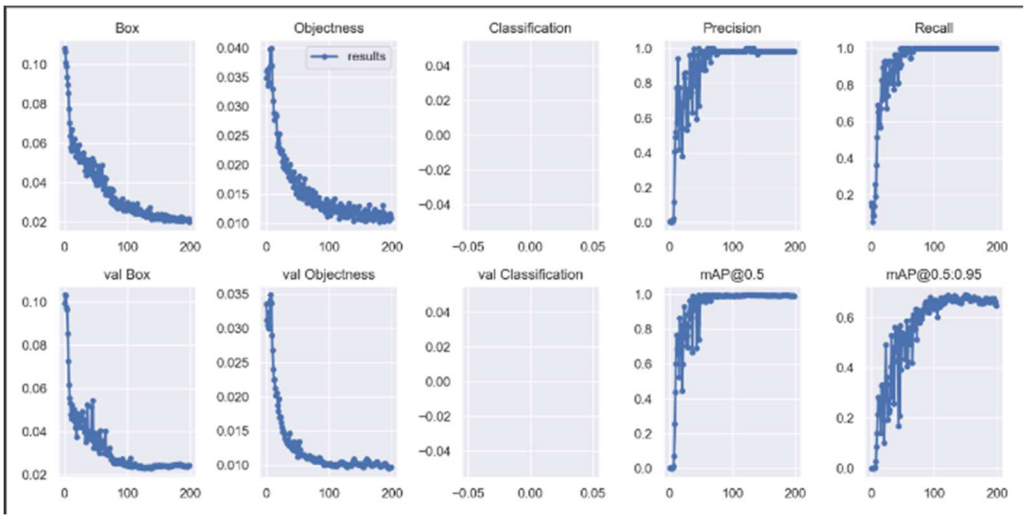


Figure 3. Training results

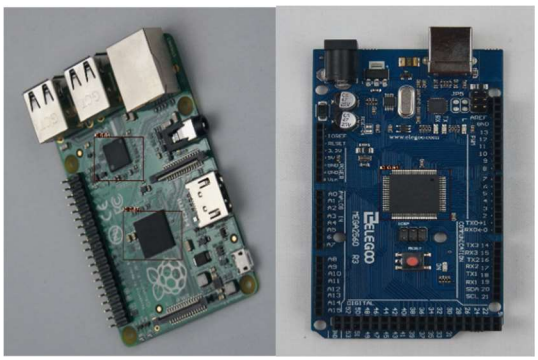


Figure 4. Predict the outcome

4.3 Algorithm comparison experiments

Table 1. Algorithm performance comparison

Model	Single image detection time/ms	Training time/h	mAP@0.5/%
Faster RCNN	60	21	89
YOLOv5	15	16	98

ISSN: 2414-1895 DOI: 10.6919/ICJE.202211_8(11).0033

5. Calibration and Positioning

5.1 Monocular Camera Calibration

The camera imaging system contains four coordinate systems: world coordinate system, camera coordinate system, image coordinate system, and pixel coordinate system.

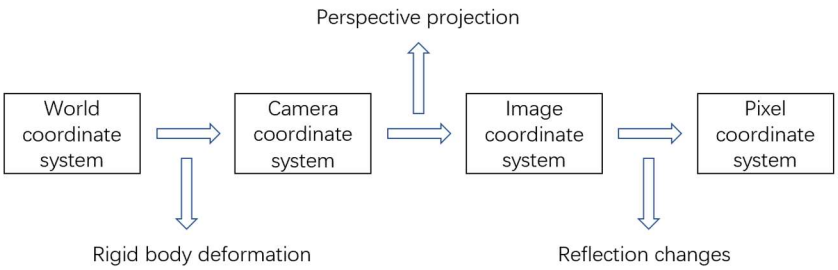


Figure 5. Coordinate system conversion

The transformation relationship between the four coordinate systems is as follows:

$$Z\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{dX} & 0 & u_0 \\ 0 & \frac{1}{dY} & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} R & T \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix} = \boldsymbol{M}_2\boldsymbol{M}_1 \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix} = \boldsymbol{M} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix} \tag{1}$$

u0 and v0 are only determined by the internal structure of the camera, M2 is the camera parameter matrix, M1 is the off-camera parameter matrix, M is the 3×4 unitarity matrix, called the camera's projection matrix.

The geometric calibration of the camera in this article adopts Zhang Zhengyou's[9] checkerboard calibration method. In the calibration process, the camera is used to shoot 15 self-made checkerboard image sets with 11x8 corner points, of which the square size of the checkerboard grid is 20mm, and the OpenCV library function is used in Python to import the self-made checkerboard calibration version image set, and finally the parameters of the camera mainly include the main point, focal length, distortion coefficient and pixel error.

```
D:\Anaconda\envs\pytorch\python.exe C:/Users/85848/Downloads/yolov5-5.0/calib_IR.py
ret: 0.4961699633095678
internal matrix:
 [[473.06210998   0.         322.24326533]
 [  0.         472.49047299 246.44975935]
 [  0.           0.           1.        ]]
distortion cofficients:
 [[-1.07820544e-01 -1.26222911e-01  9.54382415e-05  1.89862906e-03
   1.71939079e-01]]
```

Figure 6. Camera calibration parameters

5.2 Pixel Coordinates are Converted to World Coordinates

YOLOv5 recognition can get the pixel coordinates of the bounding box of the recognized object, and the following formula is used to convert the actual position coordinates of the object.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202211_8(11).0033

$$Z\begin{pmatrix}u\\v\\1\end{pmatrix}=K(R\begin{pmatrix}X_w\\Y_w\\Z_w\end{pmatrix}+T) \tag{2}$$

The formula is solved:

$$\begin{pmatrix}X_w\\Y_w\\Z_w\end{pmatrix}=R^{-1}(K^{-1}Z_c\begin{pmatrix}u\\v\\1\end{pmatrix}-T)=R^{-1}K^{-1}Z_c\begin{pmatrix}u\\v\\1\end{pmatrix}-R^{-1}T \tag{3}$$

R, K, T are all known.

Table 2. Board chip positioning data.

	The calculated coordinates of the upper left vertex (mm)	The calculated coordinates of the lower right vertex (mm)	The actual coordinates of the upper left vertex (mm)	The actual coordinates of the lower right vertex (mm)	Maximum Error (mm)
1	(50.1,-60.5)	(59.3,-55.9)	(49.5,-59.5)	(-59.5,-54.5)	1.4
2	(-48.3,2.0)	(-43.4,-11.9)	(-47.5,1.5)	(-42.5,-11.5)	0.9
3	(15.2,-15.8)	(49.9,15.6)	(14.5,-14.6)	(44.5,-14.5)	1.3

6. Conclusion

This paper simply implements the positioning theory and method of circuit board electronic chip components, also introduces the imaging principle and calibration method of the camera, and does the accuracy verification test for the calibration of the camera, and finally completes the positioning test of the high-value electronic components of the circuit board based on YOLOv5 electronic component target detection position border regression, and the maximum positioning error in the experiment is 1.4mm, which meets the positioning requirements. The use of deep learning-based methods to detect and identify and locate high-value electronic components of waste circuit boards is very meaningful, the method has stronger intelligence, through the sample training set and related training strategies to enhance the detection algorithm's ability to adapt to the environment, because its strong environmental adaptability makes the algorithm have stronger portability, more importantly, this method can be extended to more applications, such as vehicles, ship detection and positioning, counting, etc. Of course, there is still room for improvement in the research results of this paper.

References

[1] G. E. Hinton and R. R. Salakhutdinov. Reducing the Dimensionality of Data with Neural Networks[J]. Science, 2006, 313(5786) : 504-507.

[2] Alex Krizhevsky and Ilya Sutskever and Geoffrey E. Hinton. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6) : 84-90.

[3] He K, Zhang X , Ren S , et al. Deep Residual Learning for Image Recognition[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016:770-778.

[4] Ren S, He K , Girshick R , et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6):1137-1149.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202211_8(11).0033

[5] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. Computer Science, 2014:1409-1556.

[6] C. Szegedy et al., "Going deeper with convolutions," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp:1-9.

[7] Redmon J, Divvala S , Girshick R , et al. You Only Look Once: Unified, Real-Time Object Detection[C]// Computer Vision & Pattern Recognition. IEEE, 2016: 779-788.

[8] Liu W, Anguelov D , Erhan D , et al. SSD: Single Shot MultiBox Detector[J]. 2015:21-37.

[9] Zhang Z. Flexible camera calibration by viewing a plane from unknown orientations[C]// Seventh IEEE International Conference on Computer Vision. IEEE, 1999.