ISSN: 2414-1895 DOI: 10.6919/ICJE.202111_7(11).0063

Express packages detection based on yolov3

Shujie Liu[1, a], Weizhong Zhang[2, b,*]

[1]College of Computer Science and Technology, QingDao University, Qingdao 266071, China;

[2]College of Computer Science and Technology, Qingdao University, Qingdao 266071, China.

[a]lotove.liu@hotmail.com

*Corresponding author Email: bzhangwz_01@aliyun.com

Abstract

Aiming at the low precision and slow speed of express packages detection under traditional image processing conditions, a deep convolutional neural network detection method based on YOLOV3-tiny is proposed. This method uses our manually labeled express parcel image data set, uses the k-means algorithm to cluster the anchor boxes of the data set, and uses YOLOV3-tiny deep convolutional neural network with data enhancement strategy to perform training and testing. Experimental results show that the average accuracy of this method model on the test set reaches 94.73%, and the real-time detection frame rate reaches 96 frames per second. The accuracy and real-time of detection have improved the existing methods to a certain extent, and have certain reference value for the development and improvement of express packages detection and separation.

Keywords

Object Detection, YOLOv3-tiny, packages detection.

1. Introduction

With the rapid development of the e-commerce industry, online shopping has become one of the main ways for people to shop daily, which has led to an increase in the number of packages in the express logistics industry and promoted the rapid development of the entire express logistics industry. The main characteristics of the express delivery industry are high labor intensity and high demand for personnel. However, the existing human resources can no longer meet the needs of the rapid development of the express industry. This situation is particularly evident in the annual e-commerce promotion. The couriers sorted violently, damaged the packages, and piled up a large number of packages in the warehouse, which reduced efficiency and made it difficult to sort, transfer and deliver some packages in time[1]

At present, the field of express packages detection is mainly based on traditional image processing methods. By processing the acquired depth image and judging by the change of the depth value, the processing of adjacent and irregular packaging is not timely, reducing the accuracy and increasing the hardware requirements. With the in-depth development of deep learning, another object detection algorithm based on convolutional neural network has emerged. According to different design ideas, the existing object detection algorithms can be divided into two categories: two-step object detection algorithms for applications such as Faster-RCNN and one-step object detection algorithms for applications such as SSD and YOLO.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202111_7(11).0063

2. YOLO Algorithm

2.1 YOLO Overview.

The YOLO object detection algorithm proposed by Redmon et al.[2] is a classic one-stage neural network model. Although it is slightly inferior to the two-stage model in accuracy, the most significant feature of the YOLO algorithm is its fast detection speed, which meets the real-time requirements of the packages detection field. On the basis of absorbing SSD[3] and FastRCNN[4] and other algorithms, YOLO has been updated in three versions, and the latest yolov3 has significantly improved its accuracy and real-time performance.

2.2 YOLOv3 network structure.

The YOLOv3 algorithm adopts the basic network of Darknet-53 and draws on the method of residual network. Residual links are set up between certain layers. The network structure of Darknet-53 is shown in Figure 1. It includes 53 convolutional layers and a large number of 3×3, 1×1 samples.

	Type	Filters	Size	Output
	Convolutional	32	3 × 3	256 × 256
	Convolutional	64	3 × 3 / 2	128 × 128
1×	Convolutional	32	1 × 1	
	Convolutional	64	3 × 3	
	Residual			128 × 128
	Convolutional	128	3 × 3 / 2	64 × 64
2×	Convolutional	64	1 × 1	
	Convolutional	128	3 × 3	
	Residual			64 × 64
	Convolutional	256	3 × 3 / 2	32 × 32
8×	Convolutional	128	1 × 1	
	Convolutional	256	3 × 3	
	Residual			32 × 32
	Convolutional	512	3 × 3 / 2	16 × 16
8×	Convolutional	256	1 × 1	
	Convolutional	512	3 × 3	
	Residual			16 × 16
	Convolutional	1024	3 × 3 / 2	8 × 8
4×	Convolutional	512	1 × 1	
	Convolutional	1024	3 × 3	
	Residual			8 × 8
	Avgpool		Global	
	Connected		1000	
	Softmax			

Fig. 1 Darknet-53 network structure

2.3 Multi-scale feature detection of YOLOv3 algorithm.

YOLOv3 uses the K-means clustering method to obtain the size of the a priori box, and predicts the object in three sizes of 13×13, 26×26, and 52×52, each of which predicts 3 target boxes. When detecting an image, if the initial grid is divided into $N \times N$ and C object categories need to be predicted, then the tensor for each scale will be:

$$N \times N \times [3 \times (4+1+C)] \tag{1}$$

Among them, 4 is the offset coordinate of the object frame; 1 is the confidence score. The confidence score refers to the probability that each bounding box contains an object. The confidence calculation formula is:

$$C_j^i = Pr(Object) * IOU_{truth}^{pred} \tag{2}$$

In the formula: represents the confidence of the j-th bounding box of the i-th grid cell; $Pr(Object)$ is the predicted target probability; pred is the predicted bounding box area; truth is the actual bounding box area; IOU is their intersection and union divide.

3. Packages Detection

The original image is shown in Figure 2. The cluttered packages around will cause interference to the result. To avoid this impact and to facilitate subsequent separation of packages and division of locations, the original image is cropped and only the region of interest is left.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202111_7(11).0063

Fig. 2 original image

At the same time, in order to provide more accurate data and avoid errors caused by the conveyor belt, the difference method is adopted to extract the packages. The formula is as follows:

$$I = O - B \tag{3}$$

Among them: O is the original image, B is the background image.

The test environment in this paper is the Windows10, the Python3.6 integrated environment of Anaconda, CUDA10.0, cuDNN 7.3.0, NVIDIA RTX2080TiGPU, and 3000 images are randomly selected from the image collection for model training. For model evaluation; using the Adam optimizer with a learning rate of 0.0001, the batch size is set to 16, and the various parameters are shown in Figure 3:

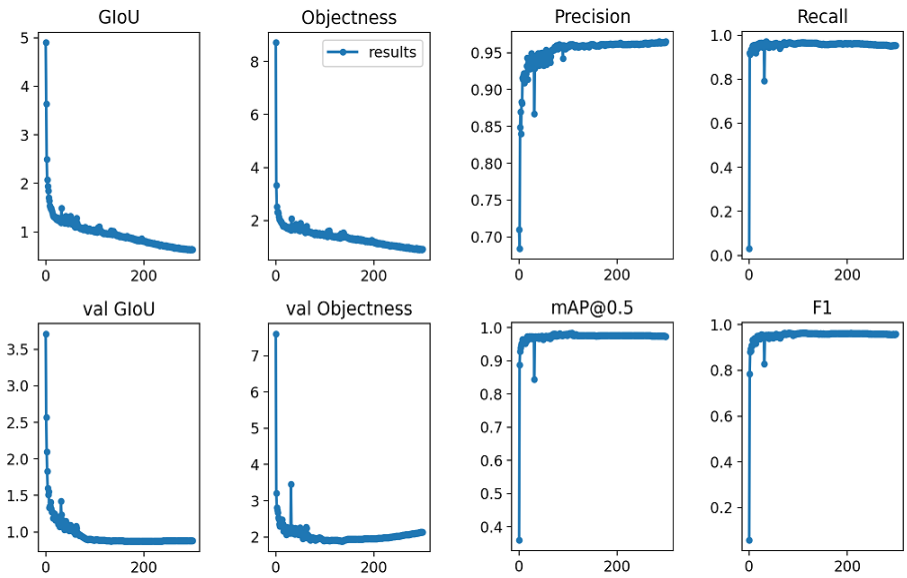

Fig. 3 yolov3 train results

Among them:

$$Precision = TP / (TP + FP) \tag{4}$$

$$Recall = TP / (TP + FN) \tag{5}$$

$$F1 = 2 * (Precision * Recall) / (Precision + Recall) \tag{6}$$

The result is shown in Figure 5:

ISSN: 2414-1895 DOI: 10.6919/ICJE.202111_7(11).0063

Fig. 5 yolov3 test results

4. Positioning enhancement

In order to obtain more accurate location information, the saliency detection in the RGB-D image[5] is performed on the extracted packages location.

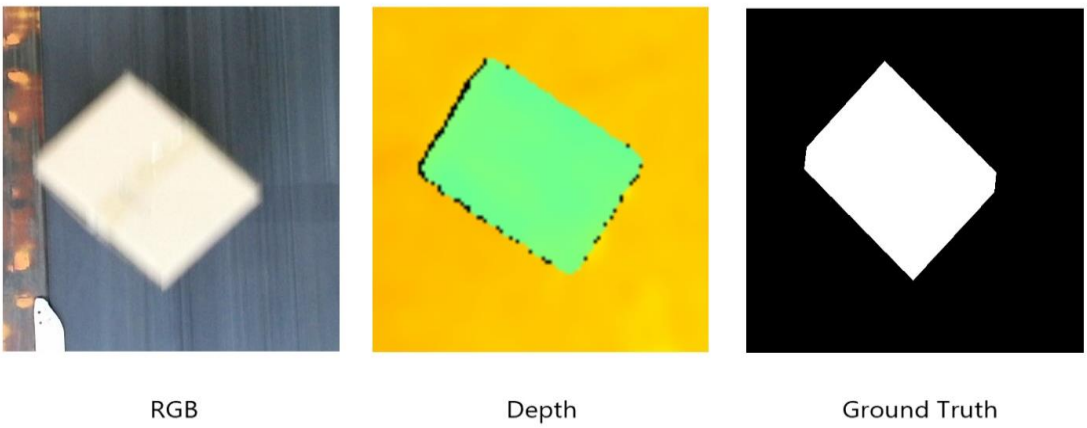

Fig. 6 pictures we used for training

First use yolo to process the image, crop the results obtained to minimize the impact of other factors, and then use a novel data-level recombination strategy to fuse RGB with D (depth) before deep feature extraction[6] for saliency detection to obtain the actual boundary.

5. Summary

In response to the detection requirements of express packages, this paper proposes a detection method based on deep learning YOLOv3 algorithm. This method locates the packages under the original image, extracts the position of the packages with high positioning accuracy, prepares for the subsequent separation of the packages, and improves the intelligence and efficiency of packages detection. Experiments show that the method proposed in this article has high accuracy and practicability for the detection of express packages.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant 61602269

References

[1] DiFeng Hu, XiangYang Zhou. Automatic three-dimensional sorting system based on Internet and database technology. Logistics Technology. Vol. 40 (2017) No. 5, p. 53-55.

[2] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, p. 779-788

ISSN: 2414-1895 DOI: 10.6919/ICJE.202111_7(11).0063

[3] Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks, IEEE transactions on pattern analysis and machine intelligence, Vol.39(2016), No.6,p. 1137-1149.

[4] Joseph Redmon, Ali Farhadi; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, p. 7263-7271.

[5] Chenglizhao Chen, Jipeng Wei, Chong Peng, Weizhong Zhang, Hong Qin, Improved Saliency Detection in RGB-D Images Using Two-Phase Depth Estimation and Selective Deep Fusion,IEEE Transactions on Image Processing, vol. 29(2020), p: 4296-4307.

[6] Xuehao Wang, Shuai Li,Chenglizhao Chen*, Yuming Fang, Aimin Hao, Hong Qin,Data-Level Recombination and Lightweight Fusion Scheme for RGB-D Salient Object Detection,IEEE Transactions on Image Processing, vol. 30(2020), p: 458-471.