Guard Alarm Device Based on Intelligent Object Recognition System

Hang Zhou[a], Wei Dong[b], and Yinghui Jiang[c]

School of Southwest Minzu University, Chengdu 610200, China

[a]z_hdhra@163.com, [b]brant.dong@163.com, [c]18780221776@163.com

Keywords: Yolo; image processing; campus; takeaway; the guard

Abstract: At present, for major takeaways and pets entering the campus, the practice of major universities is to arrange doormen to block them at the school gate, or to send permits to the takeaways to enter the campus. To reduce the workload of doormen and improve their work efficiency, this paper developed a system based on artificial intelligence that can recognize objects, which can be combined with the program installed in the camera system at the entrance of the campus. When a deliveryman enters the campus, a prompt will sound in the doorman's room. This paper is based on the image processing technology to do the following content, first, use OpenCV haar cascade way to identify the takeaway workers into the campus, and issued an alert. Secondly, when the haar cascade mode of opencv was not accurate enough, we used Yolo to label the pictures and train our own customized "takeaway worker" training set and identify the target. The results show that: this paper has achieved a specific feature of the target, we can achieve accurate identification and take the corresponding action, reduce the workload of the guard.

1. The introduction

1.1 The research background

It can be said that the discussion on whether takeout can enter the campus has never stopped. Because of various questions, it has been searched many times and aroused discussion.
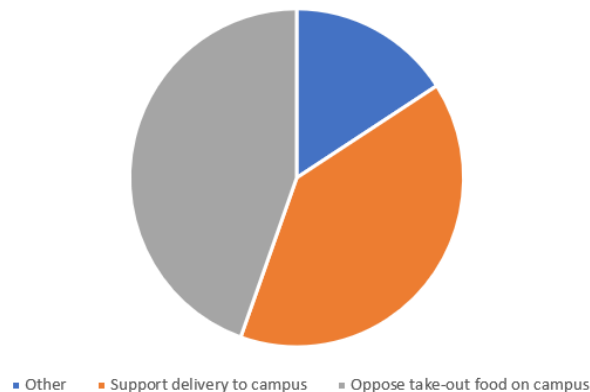

Fig 1. The statistical results

Around 2018, there was an online discussion about whether takeout could be allowed on campus, under the headline: "you have a 50% chance of being a victim of a traffic accident, a 50% chance of being a perpetrator, and a 95% chance of being a traffic violator."

In the discussion, the proportion of supporters and opponents is equal, but in another chart, most people still believe that the advantages brought by take-out in campus far outweigh the disadvantages, and there are also data showing that the number of people who order take-out is far more than that of people who do not order take-out
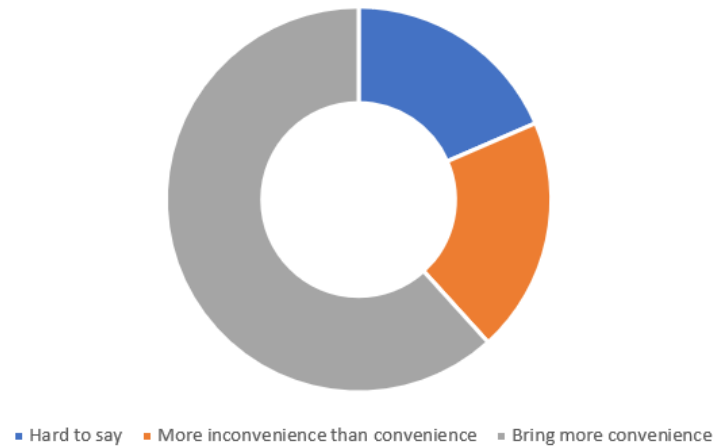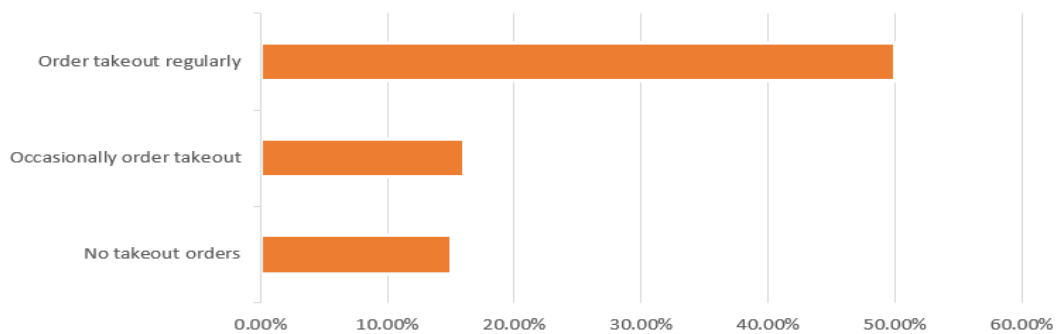
Fig 2. The statistical results

Fig 3. The statistical results

Compared with the inconvenience brought by take-out, the benefits and convenience brought by take-out are obviously more recognized by people. At the same time, the smooth implementation of the step from food production to delivery to consumers on the platform for selling to the outside world also provides challenges, especially the security challenges caused by take-out on campus.

1.2 Research status

1.2.1 Several common practices of colleges and universities at present

For on-campus takeaways:

(1) Most universities maintain an open and cautious attitude, because when the takeaway into the campus, if there is any dispute, the school has a part of the responsibility, therefore, most schools only allow relatively large formal platforms to enter schools, at the same time, it is a one-car-and-one-license system.

(2) Few universities don't allow takeout on campus.

(3) There are no reports yet that takeaways will be allowed on campus without any requirements.

1.2.2 Problems arising

(1) the biggest problem is the safety problem, due to the university campus more frequent flow of people, not standardized delivery personnel cycling into the campus after the occurrence of safety accidents.

(2) efficiency problem, because most of the colleges and universities to the takeaway issued a certificate, the implementation of a car a licensing system, resulting in the campus delivery brother every time to delay a lot of time to verify identity, but also for students caused a certain delay in time to eat.

(3) manpower problem, because every time the guard must carry on the inspection, then needs the school gate real-time guard, some guards will put a few stools in the doorway, but in general is a kind of challenge to the guard's physical strength.

1.2.3 Combination of computer vision technology and delivery platform

At school, not a few take-out food is not strictly monitored when it enters the campus, out of the question, is often dispute continuously, such as the guard because of the busy no time to check delivery agent certificate, when takeaway little elder brother unlicensed enter the campus, had an accident, can't accurate accountability, or, when the take-away little elder brother forget to carry documents, requires the students take the initiative to go to school to receive food, it is not only time-consuming and laborious but also has a bad effect on the corresponding suppliers.

1.2.4 Combination of computer vision technology and embedded platform

(1) popular deep learning:

Computer vision has high accuracy and strong robustness in deep learning. At the same time, with the rise of popular frameworks such as TensorFlow and caffe, deep learning is gaining popularity. As a classic research direction of computer science, computer vision has great potential in the application of deep learning.

(2) combination of deep learning and embedded platform:

The deep learning algorithm deployed on the embedded system not only makes the embedded system more "intelligent" but also makes people's life more convenient.

2. Study on Yolo algorithm

2.1 The way a computer stores image

First, we found a picture of the star "Tony leung" and put him into the MCU, then we saw his picture

Fig 4. A picture of star Tony leung

But computers don't "see" the world the way we do. We zoom in and select an area

Fig 5. The enlarged image "Tony leung"

Amplified again

2	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1
2	2	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
14	0	23	0	9	10	0	0	4	7	3	1	4	6	9	12	12	8	2	
6	19	0	12	8	5	3	0	0	3	7	0	3	6	2	0	0	0	0	
0	0	15	0	0	1	14	13	0	0	0	29	2	0	5	16	8	2	7	
18	16	0	12	44	104	168	213	237	253	255	211	207	156	64	0	0	4	5	
18	0	44	176	207	240	255	255	255	250	243	251	234	244	255	208	97	15	0	
0	64	204	243	255	255	255	249	252	255	252	241	255	255	245	255	250	146	19	
82	252	255	248	250	249	246	245	250	255	255	244	255	251	240	252	255	244	201	
207	255	255	255	249	249	255	255	255	251	254	255	251	252	255	255	248	250	255	
244	255	246	255	255	255	255	255	255	255	255	254	254	254	254	254	254	254	254	
255	252	255	255	255	255	255	255	255	255	255	254	254	254	254	254	254	254	254	
255	255	246	255	255	255	255	255	255	255	255	253	253	253	253	253	253	253	253	
251	255	247	255	255	255	255	255	255	255	255	253	253	253	253	253	253	253	253	
255	239	255	255	255	255	255	255	255	255	255	253	253	253	253	253	253	253	253	
250	255	250	255	255	255	255	255	255	255	255	254	254	254	254	254	254	254	254	
246	255	242	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	
255	235	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	
255	244	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	
255	244	254	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	
255	244	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	
255	245	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	
255	247	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	
255	247	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	

Fig 6. The final enlarged image " Tony leung "

Now you can see rows and rows of Numbers in the picture, and in fact, this is how computers are stored.

2.2 Opencv is used to train haar cascade classifier to identify targets:

2.2.1 Collect and make thousands of "negative" images, but make sure that the object to be detected is not captured by the crawler in the image

2.2.2 Make thousands of "positive" images, make sure they contain the object to be detected, and convert the image to vector format

opencv_createsamples -img eleme.jpg -bg neg.txt -info pos.txt -maxxangle 0.5 -maxyangle -0.5 -
opencv_createsamples -info pos.txt -num 2000 -w 20 -h 30 -vec pos.vec

Fig 7. Trained "we let me" sign

2.2.3 Test the generated haar classifier

test_haar = cv2.CascadeClassifier("test.xml")

Fig 8. Identify the results

2.3 Basic principles of Yolo algorithm:

2.3.1 Profile

Yolo is an advanced, real-time target detection algorithm.

2.3.2 Design concept

After the images are changed into the required size by the Yolo algorithm, a separate CNN model is adopted to realize target detection, and the images are sent to the CNN network. Finally, the network prediction results are processed to obtain the detected targets. Compared with the R-CNN algorithm, it is a unified framework, using only one network, not only do the classification, but also do the regression, so its speed is faster, and the training process of Yolo is also end-to-end.
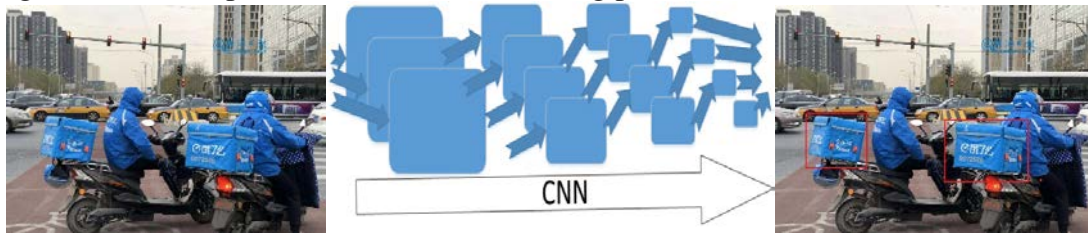

Fig 9. The process of Yolo

2.3.3 Sliding window technology

The general method is to use a sliding window target detection algorithm, which transforms the detection problem into an image classification problem. The basic principle is to use Windows of different sizes and proportions (width to height ratio) to slide on the whole picture with a certain step size, and then do image classification on the corresponding regions of these Windows, to realize the detection of the whole picture, as shown in figure 3 below. However, since we do not know the size of the target to detect, you need to set different sizes and proportions of the window to slide, and choose the appropriate step size. But this will produce a lot of sub-regions, and all must go through the classifier to make predictions, which requires a lot of calculation, so your classifier can not be too complicated, because to ensure the speed. One of the solutions is to reduce the number of subregions to be classified, which is an improvement strategy of R-CNN. It adopts the selective search method to find the Region Proposal that is most likely to contain the target. In fact, it can be regarded as a heuristic method to filter out many subregions, which will improve efficiency.
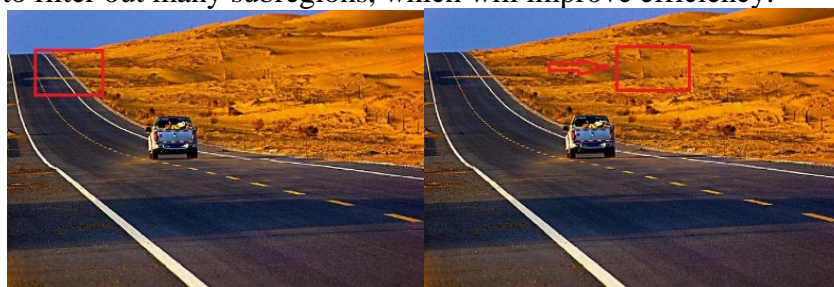

Fig 10. The sliding window is used for target detection

2.3.4 Yolo grid partitioning technology

While Yolo uses the generation of the grid, directly determine the confidence in the grid, generate some identified objects in the box, finally, the use of a non-maximum suppression algorithm, the repeated overlapping box removed, showing the object to be detected.

Fig 11. Grid generation algorithm

3. Realization of the whole system:
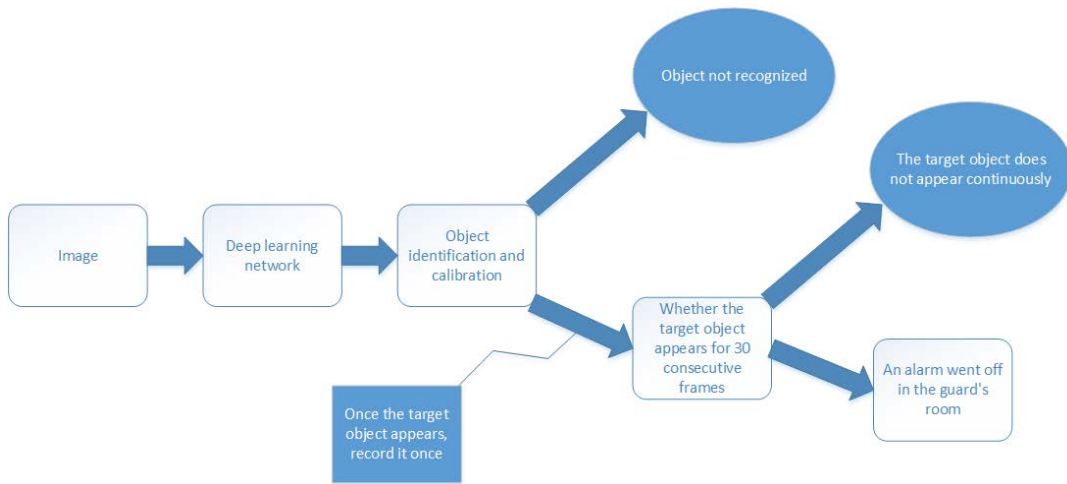
3.1 System schematic

Fig 12. System schematic

The camera collects pictures, and the system judges whether there is a takeaway. In order to eliminate the wrong interference and see whether the target object continuously appear in the field of vision if the conditions are met, an alarm is sent to remind the doorman to open the window to have a look, or to check the takeaway certificate

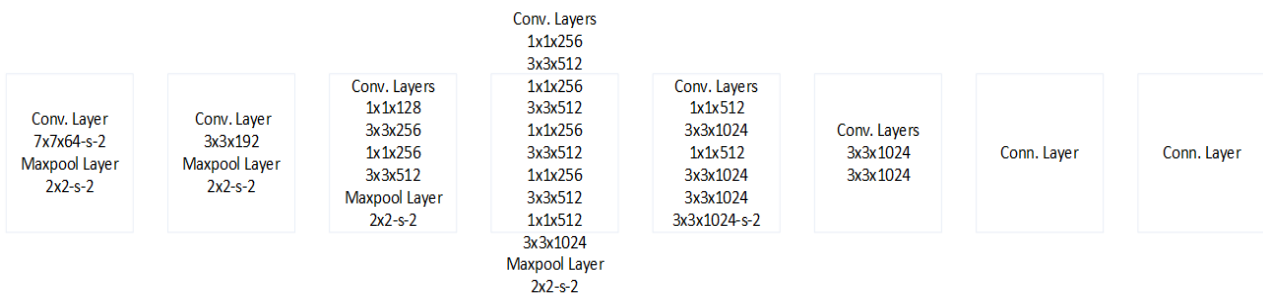### 3.2 Overview of Yolo visual algorithms

Fig 13. Schematic diagram of Yolo algorithm

It consists of 24 convolutional layers and 2 full connection layers. The convolutional layer is used to extract image features, and the full connection layer is used to predict image position and category probability values. Of course, we need to make some modifications in the end.

3.3 Algorithm overview

First, the camera collects the pictures of people entering the campus and inputs them into the system to see if there is a take-out operator entering the gate. After eliminating the interference of identification errors, it is determined that the take-away deliveryman enters the campus if the target is always present in the image for more than 30 consecutive frames.

1) Preprocessing by opencv
(a) Make all images the same size (length: 640, width: 480)
2) Create and initialize the network
(b) Establish 28 convolution layers.
(c) The first two layers are not nonlinear, and the last layer is linear.
(d) Load and save the generated network
3) Recognize images through weight files generated from training data
4) Display results

3.3.1 Software

```
# Load the necessary libraries
import numpy as np
import cv2
import tensorflow as tf
import time
import sys
import os
# Create a network
build_networks(self)
# Train and save the created network
self.saver = tf.train.Saver()
self.saver.restore(self.sess,self.weights_file)
# Images are identified by weights generated from training data
detect_from_cvmat(self,img)
# Show the results
if "person" in class_results_set:
self.detected+=1
# cv2.imwrite(new_img_path,img_cp)
if self.filewrite_img :
if self.disp_console : print '      image file writed : ' + self.tofile_img
is_saved = cv2.imwrite(self.tofile_img,img_cp)
if is_saved == True:
  print ("Saved under:", self.tofile_img)
else:
  print ("Saving error!s")
if self.imshow :
cv2.imshow('YOLO_small detection',img_cp)
cv2.waitKey(1)
if self.filewrite_txt :
if self.disp_console : print ' txt file writed : ' + self.tofile_txt
ftxt.close()
```

4. Summary and outlook

4.1 Summary

In response to the national policy on the implementation of artificial intelligence, at the same time with China's information construction and the improvement of the international level, computer

vision and artificial intelligence will be more attention.

Based on the rapid development of computer and image processing technology, today's social demand for machine vision, developed by our team of object recognition system based on deep learning has carried on the preliminary application, around the problem of take-out food into the campus, put forward a solution, and make a set of system.

According to the computer vision technology at home and abroad, the research significance and research direction of this subject are proposed. Secondly, it analyzes the existing image processing methods and proposes a technical route suitable for the research in this paper. Finally, consult relevant literature, put forward a set of solutions.

The main work and results of this paper are as follows:

1) We studied the Yolo algorithm and made a set of systematic solutions to the problem that it is not convenient to check takeout on the campus.

2) Made a set of valuable training set by crawling to get pictures of network takeaway.

4.2 outlook

This paper can do further research on the follow-up work.

1) The Yolo algorithm can be further improved to make it faster and more accurate

2) This does not address the safety and efficiency of take-out food entering the campus only reduced the workload of the doormen. People can continue to study and propose a more efficient solution.

3) The production of training sets is tedious and the cycle is long, so it is necessary to find a faster method.

References

[1] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).

[2] Girshick, R. (2015). Fast r-cnn. Computer Science.

[3] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., & Fu, C. Y., et al. (2016). SSD: Single Shot MultiBox Detector. European Conference on Computer Vision.

[4] Hager, G. D., Dewan, M., & Stewart, C. V. (2004). Multiple kernel tracking with SSD. IEEE Computer Society Conference on Computer Vision & Pattern Recognition.

[5] Jiang Yin, & Zhang Ling. (2016). Analysis of college students' satisfaction with o2o food delivery platform -- taking "ele. me" as an example.

[6] Hong Jingyue, & Hou Yanli. (2016). Horizontal and vertical research on the impact of campus takeout o2o on college students' physical and mental health (10), 91-92.

[7] Tang Rong, & Li Zhengjun. (2016). Research on the distribution mode of campus takeaway third-party logistics. Journal of hunan university of technology (social science edition), 21(2), 118-121.