

Title: Detection and recognition applications in unmanned retail (clothing) with computer vision technology

Industrial Applications □Intelligent Manufacturing □Intelligent Driving ■Intelligent Life □Smart Medicine □Smart City

[Overall background]

At present, the online flow is almost divided, the "new retail" concept has become the red fried chicken. Off the line, it once again becomes "the place of war". In other words, under the line, it is still the absolute big head of the retail industry.

However, the main cost of offline retailing is much higher than the yearly rentals. Statistics show that in 2016, the total rent in China increased by 7%, and the average salary of employees increased by 8%. Unmanned retail just can save human cost. To eliminate the human maintenance cost of unmanned retail equipment, the form of user self-service and payment saves the most important labor cost.

In November 11, 2016, the office of the State Council issued the opinion on promoting the transformation of the innovation of retail entities (State Office [2016] No. 78), which made clear the guiding ideology and basic principles to promote the transformation of the innovative retail innovation in China. At the same time, we made specific plans in adjusting the business structure, innovating the development mode, promoting cross-border integration, optimizing the development environment and strengthening policy support. The issue of "opinion" expect to promote the integration of online and offline. It emphasizes on: "To establish the standard and competition rules to adapt to the integration development, guide the entity retail enterprises to improve the information level gradually, and integrate the advantages of the offline logistics, service, experience and other advantages with the online business, capital flow and information flow, and expand the whole channel layout of intelligent and networked."

[Business background]

Where will no retail sale go in the future? Industry experts believe that computer vision will become the mainstream operation of unmanned retailing based on cost and promotion. Of course, the premise is that computer vision is enough mature in technical view, otherwise the cost problem is also difficult to solve.

For example, in the view of Jiang Guofei, the vice president of ant gold clothing and the head of the technical laboratory, the computer vision will become the mainstream technology direction of the unmanned retail in future, because the sensor will increase the cost of operation and the transformation of the supply chain. However, the computer vision will be able to be expanded in scale once the computer vision is mature.

In addition, the details of the goods and people captured by the computer vision, after the analysis by the algorithm model, can help the entity shop owner to better understand the guests and shops, and make a policy decision. For example, does the bursting goods should be adjusted to the position of the set; is it to be based on the guests' favorite route, and then we consider to optimize the shelves and commodity placement strategies; if the sales performance during the rush hour peak is flat, it should consider replacing the recommended goods between the peak periods; the supermarket supply chain is or not necessary to systematically optimize the other.

“If we refer to the five level technical chart of driverless technology, the fifth order of unmanned retail technology is pure computer vision technology. So far no one can come up with a mature technology plan, but technology idealists try to rivet their strength to get closer to it.” Jiang Guofei said.

Project description

[Problem description]

At present, the mainstream unmanned retail solutions, including Amazon Go, have adopted the hybrid technical solutions of "sensor" and "computer vision", but the lower cost of the pure computer vision scheme will become the mainstream of unmanned retail once the future is mature. However, at present, the unmanned retail stores have to improve the recognition rate of goods by unifying the shape specifications of the goods. This technical defect makes it difficult to popularize the unmanned retail stores. We hope to find a computer vision technology and algorithm that can be applied to unmanned retail stores as far as possible. Because there are many kinds of unmanned retail goods, the scene is complex and there are many technical difficulties involved, so we hope to gradually improve the recognition rate of retail goods by designing a single item and scene (clothes), and expect to introduce a "take then walk" retail model store in the future.

[User expectations]

The participants should train the AI model based on the training data set provided by the organizers. And in the test scene, they can correctly detect and recognize the clothing goods in the test data set in the valid time while the computing resources are determined, and detect the position of the target object in the image, and recognize the target objects.

[Expected economic effect]

In 2017, China's total retail sales reached 33 trillion and 200 billion Yuan, up 10.4% over the same period last year. At the same time, according to the "new retail business model innovation and investment opportunity depth study" issued by the prospective industry research institute, the unmanned retail stores will meet the development dividend period in the next five years. In 2020, the growth rate is expected to reach 281.3%, and the market volume will exceed 1 trillion and 800 billion Yuan by 2022.

Task requirements

[Technical path]

Based on ordinary PC computers, the performance of a single GPU on Ubuntu 16.4 system cannot be higher than that of GeForce GTX 1080 Ti.

[Technical indicators]

From the point of view of the system development, the participants will use the training set provided by the organizer to train an AI model, and finally provide a software test demo that can be run in the hardware environment described in the technical path above, and can correctly detect and recognize the images in the test set:

(1) Deviation rate between target detection position and the actual position, is to calculate the

overlap ratio of the two boxes, taking the average of all proportions, whose value range is [0,1] and can be marked as F_x ;

(2) Correct rate between the target recognition category attribute and the actual category attribute, is to calculation f_1 according to the accuracy rate and recall rate, whose value range is [0,1] and can be marked as F_y ;

(3) The processing time of a single picture should not exceed 33ms. For statistical convenience, calculate the total time for running all pictures, and divide the total number of pictures to find the average time T spent for each picture. If $T > 33\text{ms}$, the time is invalid. The value is 0. If $T \leq 33\text{ms}$, the value is $1 - (T/33)$, and the value range is [0,1], which is marked as F_z ;

(4) The formula for calculating the total score is: $\text{Score} = 0.3 * F_x + 0.5 * F_y + 0.2 * F_z$.

[Standard Submission]

The test demo is based on the form that the contestant is good at. No separate format is required. The output file format must be a CSV file containing the following elements:

- (1) Source file name, such as 201805140001.jpg;
- (2) Target position, such as "10,20,100,150", which stands for x, y, width, Height;
- (3) Target categories, such as ramper, hoodle, etc.;
- (4) Target attributes such as palm, leather, mini, etc.

[Task list]

- (1) Proposal and project documents, focusing on ideas of algorithm design;
- (2) If using neural network technology, submit training generated neural network files;
- (3) Executable file or installation package that can operate normally;
- (4) Project source code.

Reference Information

[Reference tool]

Tensorflow

References

<https://www.tensorflow.org/>

<http://www.im2maker.com/news/20161208/40b2a5606c1d4e17.html>

[Data interface]

None