Meteorite Detection and Tracing with Deep Learning on FPGA Platform

Kuo-Kun Tseng, Jiangrui Lin, Haichuan Sun, K. L. Yung and W. H. Ip

K.-K. Tseng, J. Lin H. Sun
School of Computer Science and Technology, Harbin Institute of Technology (Shenzhen), Shenzhen, China
e-mail: kktseng@hit.edu.cn

K. L. Yung W. H. Ip
The Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hung Hom, Hong Kong

W. H. Ip
The Department of Mechanical Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada

Abstract At present, the strength of space exploration represents the strength of a country. Meteorite exploration is also part of the space field. The traditional meteorite detection and tracking technology are slow and not accurate. With the development of deep learning, computer detection technology becomes more and more accurate and efficient, which makes it possible to improve the accuracy and speed of meteorite detection. In this paper, the deep learning algorithm implemented by FPGA is applied to the meteorite detection, and the popular tracking algorithm is applied to the meteorite tracking, so that the structure of the meteorite detection and tracking system can meet the practical requirements.

Keywords Meteorite detect · Tracking · Deep learning · FPGA

1 Introduction

Meteorite detection is an important part of space detection. Meteorites vary in size, and there are space stations, satellites, and lots of space junk floating in space. How to avoid the collision between meteorites and artificial satellites and keep satellites safe by accurately identifying meteorites, detecting and tracking them, predicting the track and speed of their operation becomes a hot topic in planetary exploration.

The traditional machine learning method has some shortcomings such as low applicability and poor accuracy in image, so it is difficult to apply to practice. In recent years, with the rise of deep learning and the development of computer hardware, the field of computer vision has made explosive progress. Prior to this, because of the low accuracy and poor real-time performance of the image algorithm, the research results of machine vision can only be applied to simple scenes. With the development of science and technology, computer vision is more pressing to cope with complex scenes. With the breakthrough development of deep learning, artificial intelligence, and other fields, many excellent large data sets of computer vision have emerged, which indicates that machine vision has entered a new era. Because of its low cost, large amount of information, and so on, the development potential is self-evident.

With the progress of deep learning, we believe that meteorite detection will face a chance to improve its technology in the speed and accuracy of detection. In the paper, the related research, algorithm, and FPGA implementation have been studied as the following sections.

2 Related Research

2.1 Detection Work

Traditional object detection methods usually only aim at one object, such as popular and mature face detection, pedestrian detection, and fingerprint recognition. Feature extraction uses some human-experience-based features such as HOG [1] features or SIFT [2] features, and the classifiers are usually SVM classifiers or AdaBoost classifiers. In general, the strategy of exhaustion is used in detection. It extracts all possible target boxes, traverses these target boxes, and classifies each window.

With the development of convolutional neural networks, the advancement of hardware technology, and the rise of machine learning, people begin to study large-scale object detection. The emergence of excellent data sets is also important. Li Feifei, a professor at Stanford University, led to the establishment of a large data set, ImageNet. In addition, the common datasets are PASCAL VOC, Microsoft COCO. Some initial research work based on convolution neural network adopted sliding window method and used CNN to classify different sizes and positions of windows. There are many regions generated by exhaustive sliding windows, and these methods often take a long time to train and test, so they cannot be used in the above data sets. AlexNet's [3] success in ILSVRC 2012 has not only affected the direction of image classification, but also attracted the attention of other researchers in the field of computer vision. At that time, object detection was still carried out by traditional algorithms, but no major breakthroughs have been made in the standard test data sets for object detection such as PASCAL VOC. Therefore, Cirshick [4] applies convolution neural network to object detection, and proposes the R-CNN model. In the R-CNN model, about 2000 candidate regions are generated for each image, and for each image, all candidate regions are extracted separately, which makes the time consumed in feature extraction become the bottleneck of the total test time. The research team of Microsoft Asia Research Institute [5] applies SPP-Net to object detection, and improves the defect of R-CNN, but there are still some shortcomings. In the process of training and testing, candidate regions are proposed, image features are extracted, and classified according to the features. The three processes form a multi-stage pipeline. A direct result of this is that additional space is needed to store the extracted features for use by the classifier. So, Girshick [6], one of the designers of R-CNN, proposed a further improvement on R-CNN, called Fast R-CNN. DeepID-Net [7] is also an influential model in object detection. The model is further improved on the basis of R-CNN training process, and the model pretraining method is improved. All of these models are based on candidate regions.

In addition to the candidate-region-based approach, YOLO introduced by Redmon [8] and others adopted the idea of regression. Then, the SSD algorithm proposed by Liu et al. [9] is faster than the previous fastest YOLO algorithm in speed, and can be compared with Faster R-CNN in detection accuracy.

2.2 Tracking Work

As a key technology of intelligent video surveillance, target tracking has important application value in many fields and has been a hot research topic in related academic fields. After decades of exploration and development, researchers at home and abroad have realized the change and optimization of tracking algorithm. According to different classification criteria, visual target tracking has the following categories [10]:

1. According to whether the tracking target belongs to rigid object, it can be divided into rigid body tracking and non-rigid tracking.
2. According to the number of cameras used to collect information in the tracking scene, it can be divided into monocular vision tracking and multi-vision vision tracking.
3. According to the different representations of the target, it can be divided into region-based, contour-based, and feature-based tracking.

3 Algorithms

3.1 Object Detection

In order to ensure the accuracy of meteorite detection algorithm, the popular depth-learning algorithm is used as the benchmark algorithm, and the algorithm is optimized and improved on the basis of this algorithm.

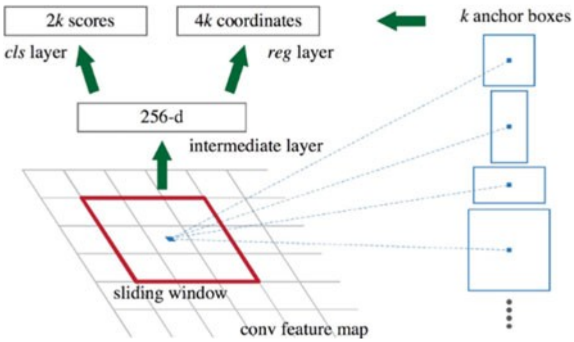

Fig. 1 The whole graph is obtained by CNN and the characteristic graph is obtained. After the convolution of kernel3*3* 2563 * 3 * 256, it is predicted that K anchor boxes are objects at each point, and the position of anchor boxes is fine-tuned. Extracts the object frame and classifies it in the same way as Fast R-CNN. Shares a CNN network with the categories [6]

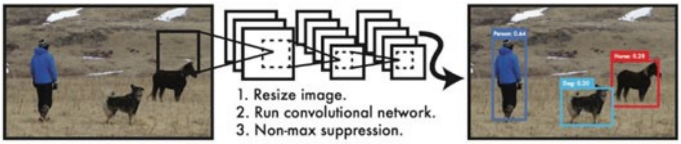

Fig. 2 Partitioned the image into an S * SS * S cell. Each cell outputs B rectangular boxes (redundant designs) containing the location information of the box (x, y, w, h) and Confidence of the object. Each cell then outputs C conditional P (Class|Object) P (Class|Object). The final output layer should have S * S × (B * 5 + C) S × S × (B 5 + C) unit [8]

R-CNN:
In the early days, window scanning was used for object recognition. R-CNN [4] removes window scanning and uses clustering method to segment and group images to get hierarchical groups of multiple candidate frames (Fig. 1).

Faster R-CNN:
Extracting candidate box is running on CPU, which is inefficient. Faster R-CNN uses CNN to predict candidate boxes. Its innovations are shown in Fig. 2.

YOLO:
Fast R-CNN needs 20 K anchor boxes to determine whether the object is, and then object recognition. It has two steps. YOLO combines the selection and recognition of object boxes, and outputs them one step at a time.

3.2 Comparison

Comparing the performance of traditional methods, region-proposal-based algorithm, and regression-based algorithm on PASCAL VOC data sets, the results are

Table 1 Contrast of different detection algorithms in PASCAL VOC

Algorithm	Training set	mAP	FPS
30HzDPM	2007	26.1	30
YOLO	2007 + 2012	63.4	45
Fast R-CNN	2007 + 2012	70.0	0.5
Faster R-CNN	2007 + 2012	73.2	7

Fig. 3 Sample of a circulant matrix

5	6	4	4	5	6	6	4	5
8	9	7	7	8	9	9	7	8
2	3	1	1	2	3	3	1	2
2	3	1	1	2	3	3	1	2
5	6	4	4	5	6	6	4	5
8	9	7	7	8	9	9	7	8
8	9	7	7	8	9	9	7	8
2	3	1	1	2	3	3	1	2
5	6	4	4	5	6	6	4	5

shown in Table 1. Experiments show that the accuracy of deep learning method is much higher than that of traditional computer vision method, and Yolo can meet the real-time requirement with relatively high accuracy (Fig. 3).

3.3 Object Tracking

This paper uses the popular tracking algorithm KCF [11] algorithm, which is based on correlation filter tracking algorithm, the effect is very good, high speed, and ideas and implementation are very simple. Among them, the method of fast computation using circulant matrix is worth learning.

The innovation of the algorithm is that the circulant matrix is used to represent the image block, which greatly increases the speed of operation.

During training, a series of displacement sampling around the current position can be represented by a two-dimensional block circulant matrix X, and the ij block represents the result of moving down the i row and right j column of the original image. Similarly, when testing, a series of displacement sampling near the result of the preceding frame can also be represented by X. Such X can quickly accomplish many linear operations by Fourier transform.

Because the tracking algorithm requires fast real time, how to track the target quickly and accurately is a very important problem. In the KCF algorithm, besides using the cyclic matrix for fast calculation, linear regression training and kernel function method are also used to speed up the algorithm. The experimental results are shown below.

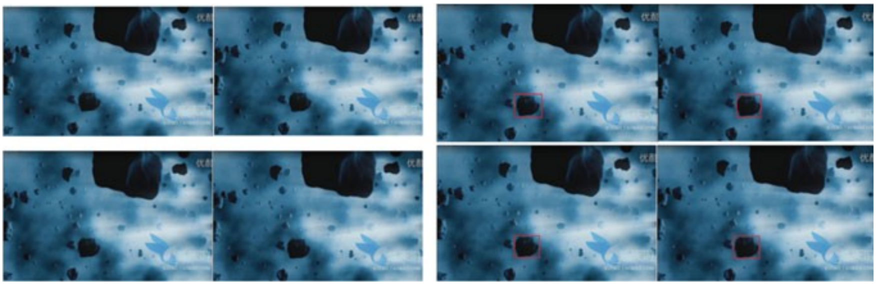

Fig. 4 Original map (left) and impression drawing (right) (http://www.youku.com/)

In the actual detection, the tracking speed reaches about 40 frames per second, which fully meets the requirements of real time. At the same time, the tracking target is not lost under the occlusion condition (Fig. 4).

4 Hardware Architecture Used FPGA

In order to save cost and meet the real-time requirement of target detection, this project intends to use FPGA instead of GPU and CPU as computing hardware resources. The advantage is that FPGA is cheaper than GPU and the computation speed is much better than that of CPU.

4.1 Object Detect with PYNQ-Z1 FPGA

The Low Power Target Detection System Design Challenge, sponsored by DAC 2018, the Top Conference on Electronic Automation Design, ended June 28, 2008 in San Francisco, California. The competition aims to design a high-precision and energy-efficient object detection system for UAVs to meet the needs of real complex scenes. In group FPGA, third teams iSmart2 use the following architecture to detect object (Fig. 5).

In addition, Xilinx provides an example of pynq running CNN. Pynq is equipped with CIFAR10 and LeNet5 networks, respectively. The CIFAR10 network has a test time of 3.3333906849999977 s and an accuracy of 73.2% at 500 batch size. The LeNet5 network has a test time of 1.1257555660000023 s at 600 batch size and an accuracy of 98.3%. (https://github.com/Xilinx/PYNQ-ComputerVision).

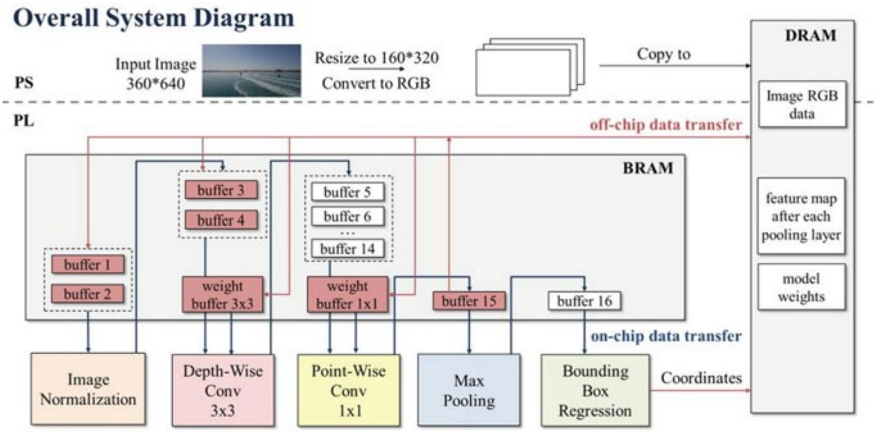

Fig. 5 The team iSmart2 adopted Mobilenet-based lightweight network design, with 12 layers. In terms of hardware implementation, the team utilizes PYNQ-Z1 and uses a module-based (IP) reuse architecture, allowing the same kind of network layer to reuse the same module to save hardware resources (https://baijiahao.baidu.com/s?id=1606045093668063557&wfr=spider&for=pc)

5 Conclusion

The deep learning architecture based on FPGA runs fast and has high accuracy. With the popular target-tracking algorithm, it can effectively solve the meteorite detection and tracking tasks. Believe that it can help the development of space industry in the future.

References

1. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2005 San Diego, CA, USA. https://doi.org/10.1109/cvpr.2005.177
2. Lowe, D.G.: Object Recognition from local scale-invariant features. IEEE **2**(3), 1150–1157 (1999)
3. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: International Conference on Neural Information Processing Systems, pp. 1097–1105. Curran Associates Inc. (2012)
4. Girshick, R., Donahue, J., Darrell, T., et al.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 580–587. IEEE Computer Society (2014)
5. He, K., Zhang, X., Ren, S., et al.: Delving deep into rectifiers: surpassing human-level performance on ImageNet classification. In: IEEE International Conference on Computer Vision, pp. 1026–1034. IEEE (2016)
6. Girshick, R.: Fast R-CNN. Comput. Sci. (2015)
7. Ouyang W., Zeng, X., Wang, X. et al.: DeepID-Net: deformable deep convolutional neural networks for object detection. IEEE Trans. Pattern Anal. Mach. Intell. (99), 1–1 (2016)

8. Redmon, J., Divvala, S., Girshick, R., et al.: You Only Look Once: Unified, Real-Time Object Detect. 779–788 (2015)
9. Liu, W., Anguelov, D., Erhan, D., et al.: SSD: Single Shot MultiBox Detect. 21–37 (2015)
10. Huang, K., Reng, W., Tieniu, T.: Survey of image object classification and detection algorithms. J. Comput. Sci. **37**(6), 1225–1240 (2014)
11. Henriques, J.F., Caseiro, R., Martins, P., Batista, J.: High-Speed tracking with Kernelized correlation filters. IEEE Trans. Pattern Anal. Mach. Intell. **37**(3), 583–596 (2015)