Research on Structural Video Analysis Platform Based on Cloud Computing

Zuocheng Wang [1.a], Lixia Xue e[2.b] and Yanli Luo [3.c*]

[1] The 38th Research Institute of China Electronics Technology Group Corporation, HeFei, China
[2] School computer and information, HeFei University of Technology, HeFei, China
[3] AnHui Sun Create electronic CO., LTD. HeFei, China
[a] cswangzc@163.com, [b] 51003239@qq.com, [c] 442590198@qq.com
*corresponding author

Keywords: Intelligent algorithm, cloud computing, video structure, Docker

Abstract: By 2010, the world has already entered the ZB era in which the global data volume roughly doubles every two years, a day of video monitoring data of more than 1500PB. By intelligent algorithm research and cloud computing technology, video structured analysis platform based on Docker is constructed through video structure key technology research. Key information of vehicles and moving targets in surveillance video being batch extracted by system forms an information retrieval library. The structure of Docker virtual technology and cloud computing platform improves high availability of system resources and achieves high reliability and expansibility of system. In actual applications, the key information in surveillance video can be spotted quickly and accurately, which has been widely used especially in public security and transportation industry, greatly improving efficiency in such industry.

1. Introduction

By 2010, the world has already entered the ZB era. The amount of data around the world doubles every two years, and more than 1500PB generates from daily video surveillance data. Tens of thousands of the 'Skynet' cameras are used in cities.

Use of video monitoring system rests on video basic functions such as video and image preview, playback and control for a long time. With the basic functions, great potential of huge video and image resource provided by the large-scale and dense deployment of front-end video points can neither be fully excavated nor utilized. Big video data application is a direction of peace city construction in the future. In the big data era, the analysis, transformation and installation of video has become the most challenging international technical problem due to the maximum of information content, statistical amount and computing time.

The paper studies the video structure key technology and constructs structural video analysis platform, basing on structural video analysis platform of Docker, the intelligent algorithm, the Docker virtual technology and cloud computing platform support. The system extracts the video

structure information of people, vehicles and objects in the video. It can also batch extract the key information of moving targets in the video, which forms the information retrieval library。

System, using the structure of Docker virtual technology and cloud computing platform, improves high availability of system resources and achieves high reliability and expansibility. In actual applications, the key information in surveillance video can be spotted quickly and accurately, which has been widely used especially in public security and transportation industry, improving efficiency greatly in such industry.

2. Object extraction, structuring

This paper, utilizing professional GPU, realizes the structural information analysis of videos and images by analytical means, such as object detection, object recognition, feature extraction and deep learning, , which improves the picture or video content recognition from artificial way to computer extraction.

```
              ┌──────────────┐
              < start        >
              └──────────────┘
                     │
                     ▼
         ┌──────────────────────┐
         │ video, images loading │
         └──────────────────────┘
                     │
                     ▼
         ┌──────────────────────┐
         │   Target detection    │
         └──────────────────────┘
                     │
                     ▼
         ┌──────────────────────┐
         │  The target object    │
         │     recognition       │
         └──────────────────────┘
                     │
                     ▼
         ┌──────────────────────┐
         │  The target feature   │
         │      extraction       │
         └──────────────────────┘
                     │
                     ▼
         ┌──────────────────────┐
         │   Target feature      │
         │    information        │
         └──────────────────────┘
```
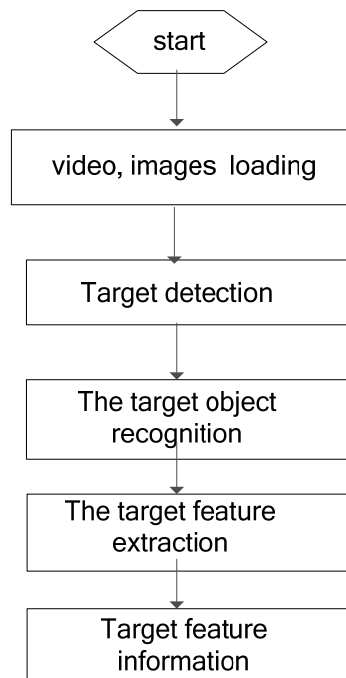
Figure 1 technical path graph

(1) Data sources: mainly including real-time video streaming, video files and images.

(2) Target detection, identification and extraction:

This article uses the Faster - RCNN and SSD collaborative training techniques for target detection, recognition and extraction to extract the target's the structural information. The car's main information includes: shape, colour, types, license, etc.; the people's main information includes: cycling types (two rounds or three rounds), gender, age group (middle, young or old), direction, and dress colour, decoration (glasses, hats, backpacks), etc.

Faster RCNN, the best method based on the deep learning area series, is adopted for target detection. Faster-RCNN is composed of the Region Proposal Network and the detection network Fast RCNN. RPN extracts the deep features of the image by the depth convolution model. The small network used by RPN slides the window in the feature image, producing the low dimensional feature mapping. Then, the data can be enhanced by the feature mapping with multi-scale reference frame zooming. Then the enhanced feature could be recognized as the target or not through the

classification network, also the coordinates of the target can be obtained through the regional regression network.
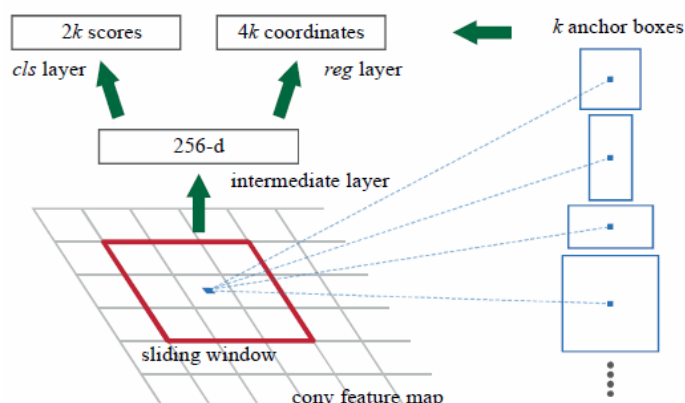

Figure2 RPN network diagram

Fast-RCNN was used to detect and identify the final target. RPN and Fast-RCNN share the convolution layer of thirteen VGG with sharing the characteristics of convolution layer in alternating training phase. Firstly, RPN is trained. Secondly, Fast RCNN is trained by proposal extracted by RPN. Thirdly, the shared convolution layer in RPN network is initialized by Faster RCNN. At last, such iterative training is executed until the end of the training. In this way, feature-shared training for convolution layer of two network can be realized.

The way of SSD (Single Shot MultiBox Detector) to obtain the target position and category is to use regression, that is, convert the target detection task into a regression problem, namely, given the input image, return the target frame and category of the location on multiple locations of the image.

SSD uses region proposal mechanism of Faster RCNN to establish a process of corresponding relation of one location and its characteristics. Use the 3*3 sliding window to extract the characteristics of each location if the feature map of certain layer(Figure b) sizes 8*8, then the feature returns to get the coordinate and category information of the target.

Multi-perspective learning hypothesis data set has two sufficient and redundant views and meets two conditions: First, each attribute set is sufficient to describe the problem, that is, it is sufficient to learn a strong learner on each attribute set if training example is enough; second, one attribute set is conditionally independent of another in a given marker.

When a task meets the fully redundant view, collaborative training can be used to generate classifiers. First, a classifier is trained on each of two views by marked examples. Then, in the co training process, each classifier picks unlabeled examples with high confidence markers (label examples with the correct mark), and the labeled example is added to labeled training set of another classifier, so that the other classifier can take advantage of these newly labeled example to update. The collaborative training process continues iteratively until all unlabeled examples are marked.

The target detection adopts the multi view cooperative training method, and the view angle is corresponding to the model. First, two models, Faster-RCNN and SSD, are trained with a small amount of labeled data. Then the unlabeled data is taken as input of the trained Faster-RCNN and SSD, and each model selects certain labels with higher score to joint each other for training according to their own score mechanism. Repeat this process until enough labeled data is generated.

(3) Information retrieval library: abstracted structural target information forms the i retrieval library.

3.　Cloud computing platform task scheduling

(1) Video structured Docker image management: the system structure uses Docker virtual technology to manage video structured services. Image management, by which each application service is produced as Docker mirror image to be uploaded to the application service warehouse, downloaded to cluster host and start the service, is the core of Docker technology.

Use Dockerfile to create a structural video image and create a new directory and Dockerfile. Then write Dockerfile and use docker build to generate mirror images. The image, containing all dependencies and running environment necessary for the algorithm application service running, uploads the created video structured image to application service warehouse. The cluster host can directly download, run and deploy the algorithm from the warehouse application to realize the dynamic distributed operation.

(2) The cloud computing platform for scheduling structural video service: cluster service scheduling and management are operated by the cloud computing platform; to create a video structured analysis task is to create an analytic task in the cluster through cloud computing platform, which assigns service model and specified configuration parameters for the services.

To start the virtual bayonet task and enter the startup parameters including the number of services to start and hosts deployed by the services is to start a new video structured analysis tasks. After the task being started, its state turned on and it is dynamically assigned to the host.

The virtual bayonet management can individually start, stop, restore, force shutdown the already-started services on the list.

4.　Information retrieval

It is mainly to achieve multi criteria query, image search, task scheduling, resource management and other functions for structured data in the video structure analysis system. The search results are displayed in a grid way, which mainly contains the searched images and the showing time of the target, with matching degree to retrieval conditions in the lower right corner. Click on the target image to open the details page.

5.　Simulation experiment

The adopted figure in the experiment in this paper shows the video of a one-way 3 lane and two-way 3 lanes. The result figure structured by the algorithm proposed in this paper for the video is shown as follows: Figure 3 is the result of the vehicle target structured by the algorithm proposed in this paper. Figure 4 is the result of people target structured by the algorithm proposed in this paper.
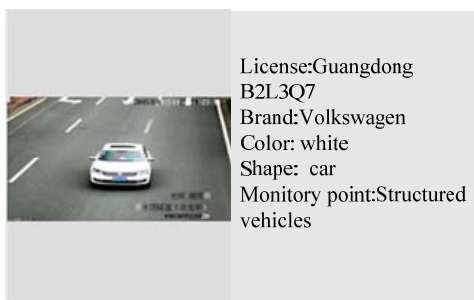
License:Guangdong B2L3Q7
Brand:Volkswagen
Color: white
Shape:　car
Monitory point:Structured vehicles

Types: Two rounds
Gender: man
Age: young
Monitory point: changjiang road

Fig. 3The Structured result of car　　　Fig. 4 The Structured result of people

License:Guangdong BP41A8
Brand:Ford Motor
Color: gray
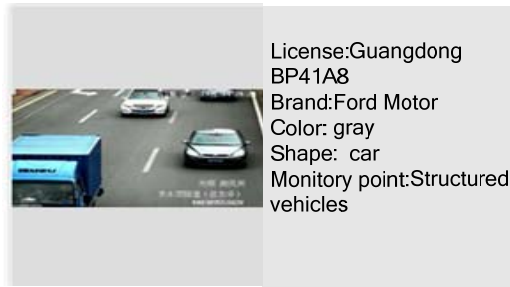Shape: car
Monitory point:Structured vehicles

Fig. 5 The Structured result of car

6. Conclusions

Based on structural video analysis system of cloud computing, the paper makes a study of structural video analysis algorithm, Docker virtual technology and cloud computing platform to solve the problems like increasing scenario application of massive video data and increased cost of data storage on petabytes level. All monitoring video, traffic bayonet, social surveillance data, etc. will be integrated into a big data center in which filed data are structurally processed, stored, managed, analyzed, calculated and utilized.

Acknowledgements

References

[1] Miao Panpan. The research of some techniques video retrieval based on content [D]. Nanjing University of Science and Technology, 2010.

[2] Lao Guofang. Applied research based on video structured surveillance system. Soochow University, 2012.

[3] X Yin, Y Shang. Research and Implementation PaaS platform based on Docker [J]. Microcomputer Applications, 2016.

[4] Liu Caiyun, Cao Jianrong, Li Hongyan. Video summarization technology based on object [J]. Computer system application, 2012, Vol.21, No.1:204-207.

[5] Zhang Jian, Xie Tianjun. Based on the Docker platform as a service architecture research [D]. Information technology and information technology, 2015 (10) : 131-134 .

[6]D Liu, L Zhao. The research and implementation of cloud computing platform based on docker [J], International Computer Conference on Wacelet Active Media Technolgy&Information Processing, 2014:475-478.