Obstacle Detection

This document describes the data format and relevant evaluation criteria of the obstacle detection task in Baidu Apollo program.

1.Introduction

It provides 20,000 frames of image data, which can be divided into 10,000 frames of training sets and 10,000 frames of testing sets. 200 frames are extracted as the sample images, half of which are extracted from the testing sets and half of which are extracted from the training sets. The data are collected in some road in Beijing. The data are 1080P colorized images. Professional annotators mark five kinds of obstacles: vehicles, pedestrians, cyclists, tricycle and other unmovable obstacles on the road. There are 176,779 vehicles, 17317 tri-cycle and cyclists, 35738 pedestrian and 4633 traffic cone.

3. Acquisition Equipment

Two cameras for the data collection of obstacles are mounted on the car roof, one of which uses a lens with a focal length of 6mm and the other uses a lens with a focal length of 12mm (the open data are collected by the former one)

Camera model	LI-USB30-AR023ZWDR
Resolution @ frame rate	1920 x 1080 @ 30 fps
Number of A/D bit depth	12 bits
Sensor module	ON Semiconductor AR023Z 1080p HD
Sensor exposure mode	Line exposure
Sensor size	1/2.7”
Pixel size	3.0 um
Color output	Colorzied
Focal length of lens	6 mm, 12 mm
Data interface	USB 3.0

4. Format Description

4.1 Format of Training Sets

The training set is organized according to the following directory structure:

train_data/ // it contains the top folders of the training data

├── images numbered from 0 to 9999 // it contains image files in the training data, which are numbered from 0 to 9999

│ ├── 00000.jpg

│ ├── :

│ └── 09999.jpg

├── labels // it contains the annotation files corresponding to the images, which are numbered from 0 to 9999

│ ├── 00000.txt

│ ├── :

│ └── 09999.txt

└── list // list files of images and annotations

4.1.1 images directory

There are images of training data in images directory, which is jpeg format and numbered from 0 to 9999

4.1.2 Formats of annotation files in the labels

Each annotation file corresponds to a frame of image, where each line corresponds to a obstacle. The column is defined as shown in Table 2, where the coordinate origin is the upper-left vertex of the image. The columns are separated by a space character.

Table 2 Formats of annotation files

Column No.	Definition
1	Types of obstacle.
5~8	Bounding box of obstacle: xmin、ymin、xmax、ymax
2~4，9~15	Reserve fields

4.1.3 Formats of list files

Each line is the relative path of a pair of images and the annotation files. There is a space character between two columns. For example:

images/00000.jpg

labels/00000.txt

images/00001.jpg

labels/00001.txt

:

:

images/09999.jpg labels/09999.txt

4.2 Format of Testing Sets

The format of the testing sets is the same as that of the training sets, but there are no labels folders. The organizational form is as follows:

test_data/ // it contains the top folders of the testing data

├── images // it contains image files in the testing data, which are numbered from 0 to 9999

│ ├── 00000.jpg

│ ├── :

│ └── 09999.jpg

└── list // list files of images Each line of the list file is the relative path of a frame of image. For example:

images/00000.jpg

:

images/09999.jpg

3.3 Format of Prediction Results

Users need to save the test result according to the agreed output format. The platform will perform evaluation according to the period specified by users. All results are stored in the result files in the order of the images in the list file of the testing set. In the result files, each line corresponds to a obstacle. The column is defined as shown in Table 3, where the coordinate origin is the upper-left vertex of the image. The columns are separated by a space character.

Table 3 Format of the test results

Column No.	Definition
1	Types of obstacle.
5~8	Bounding box of obstacle: xmin、ymin、xmax、ymax
2~4，9~15	Reserve fields
16	Confidence of the detection

3.4 Evaluation Metrics

Intersection-over-union (IoU) threshold is 0.5, that is, the detection box is considered to be correct when the IoU of the detection box to the ground truth box is greater than 0.5. If there are several detection boxes whose IoUs of the ground truth box are

greater than 0.5, the detection box with the highest confidence is selected to be the correct result. The algorithm is evaluated through the precision and recall rate as well as the average precision (AP). Among them, AP is the area under the precision-recall rate curve.

5 . Document of User Interfaces

5.1 Training Interfaces

Users need to implement their own training scripts, run_train.sh. Among them, the training script needs to satisfy:

1. The location of the training data is specified in the script. The data preprocessing or format conversion should be conducted as needed.

2. The training logs need to be stored in the logs/train.log. The training loss should be output in the format of 'Iteration %d, loss = %f' for the platform to parsing and drawing. The platform will call the training script as follows. ./run_train.sh

5.2 Prediction Interfaces

Users need to implement their own prediction script, run_predict.sh. The prediction script is called by the platform based on the period specified by users. The prediction script needs to satisfy:

1. The first parameter is the address of the model file.

2. The second parameter is the path of the input data folder. The data format refers to section 3.2.

3. The third parameter is the address of the prediction result file. The output format refers to section 3.3.

4. The prediction log is stored in the logs/predict.log.

The platform will call the prediction script as follows ./run_predict.sh $modelfile $image_root_dir $predict_result_file