TAD16K: An Enhanced Benchmark for Autonomous Driving

Yuming Li*, Jue Wang, Tengfei Xing, Tianlu Liu, Chengjun Li, Kuifeng Su
Autonomous Driving Lab, Tencent, Beijing, China
Email: lymhust@gmail.com

ABSTRASCT

Although promising results have been achieved in the areas of object detection and classification, few works have provided an end-to-end solution to the perception problems in the autonomous driving field. In this paper, we make two contributions. Firstly, we fully enhanced our previously released TT100K benchmark and provide 16,817 elaborately labeled Tencent Street View panoramas. This newly created benchmark, we call it Tencent Autonomous Driving 16K (TAD16K), not only contains previously labeled traffic-signs (221 types), but also creates annotations for three new objects, which are traffic lights (6 types), vehicles and pedestrians. Secondly, we provide the evaluation results of two state-of-the-art object detection algorithms (SSD and DetectNet) on our benchmark, which can be used as the baseline for future comparison purpose. Finally, we also demonstrate that the network trained on our benchmark can be directly deployed for practical application. The TAD16K, relevant additions and the source codes are publicly available.

1. TAD16K BENCHMARK

TAD16K is an updated version of our previously released traffic-sign benchmark TT100K. The TT100K is a large traffic-sign benchmark. The images in it come from 100,000 Tencent Street View panoramas. which choses 10 regions from 5 different cities in China (including both downtown regions and suburbs for each city) and downloaded from the Tencent Data Center. Presently, Tencent Street Views cover about 300 Chinese cities and the road networks linking them. The original panoramas were captured by 6 SLR cameras and then stitched together. Image processing techniques such as exposure adjustment were also used. Images were captured both from vehicles and shoulder-mounted equipment, at intervals of about 10 meters. These capture and processing methods of panoramas in Tencent Street View guarantee the high definition and resolution of our TT100K and TAD16K. Besides, panoramas in Tencent Street Views are captured in different seasons and weather, which further ensures the variety of our datasets.

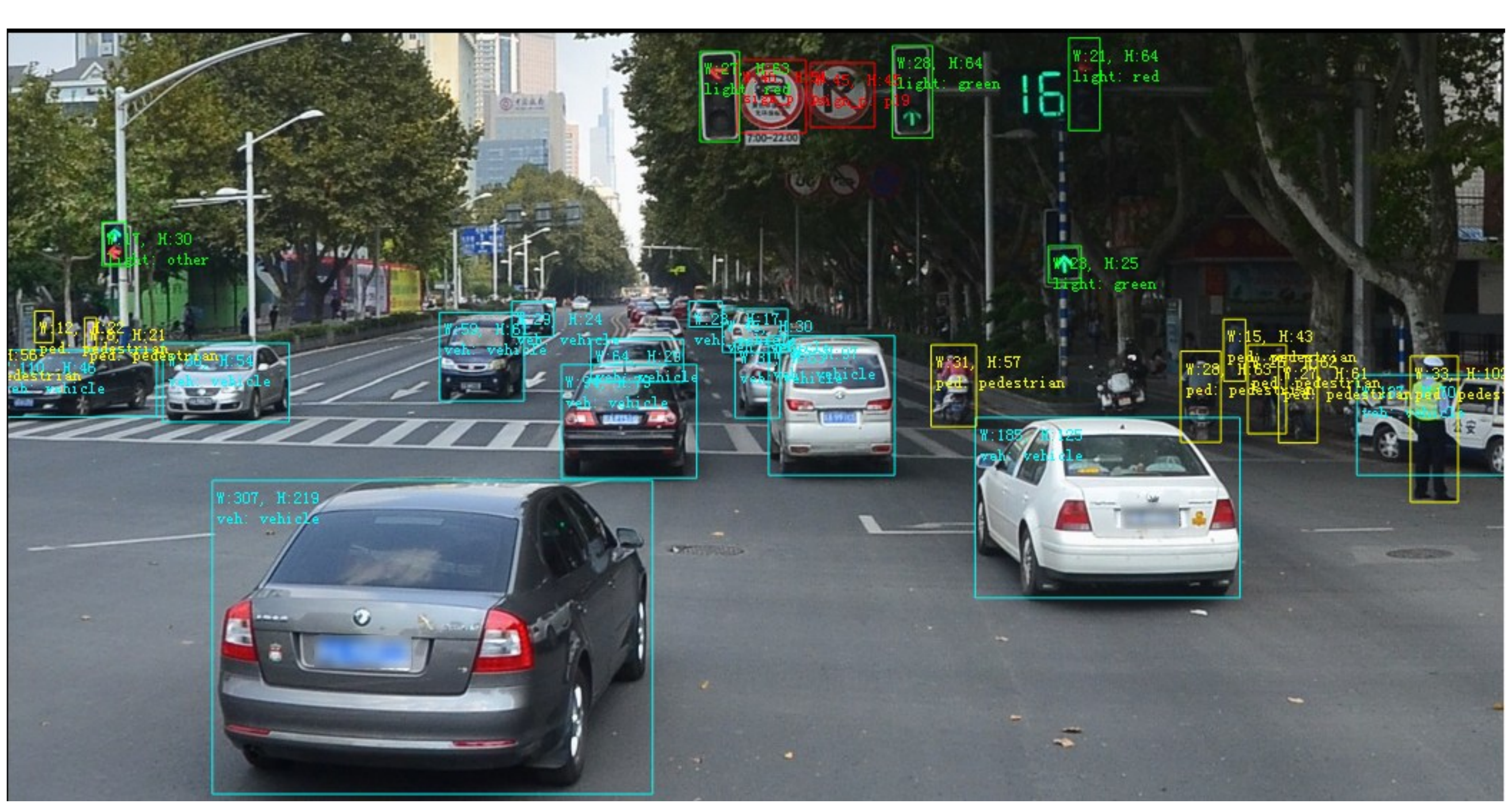

Fig. 1. Example of annotations for one image.

Table 1. Statistics of the four types of objects in

Object	(Hmin, Hmax)	(Wmin, Wmax)	Number
Sign	(7, 495)	(7, 440)	27,253
Light	(6, 512)	(4, 732)	17,369
Vehicle	(3, 1955)	(3, 2049)	114,717
Pedestrian	(5, 712)	(3, 1164)	43,046

2. BENCHMARK ALGORITHM

To facilitate the usage of TAD16K and provide a baseline for future comparison of different algorithms. We trained two state-of-the-art object detection algorithms (SSD and DetectNet) using the TAD16K benchmark and provide the evaluation results of them.

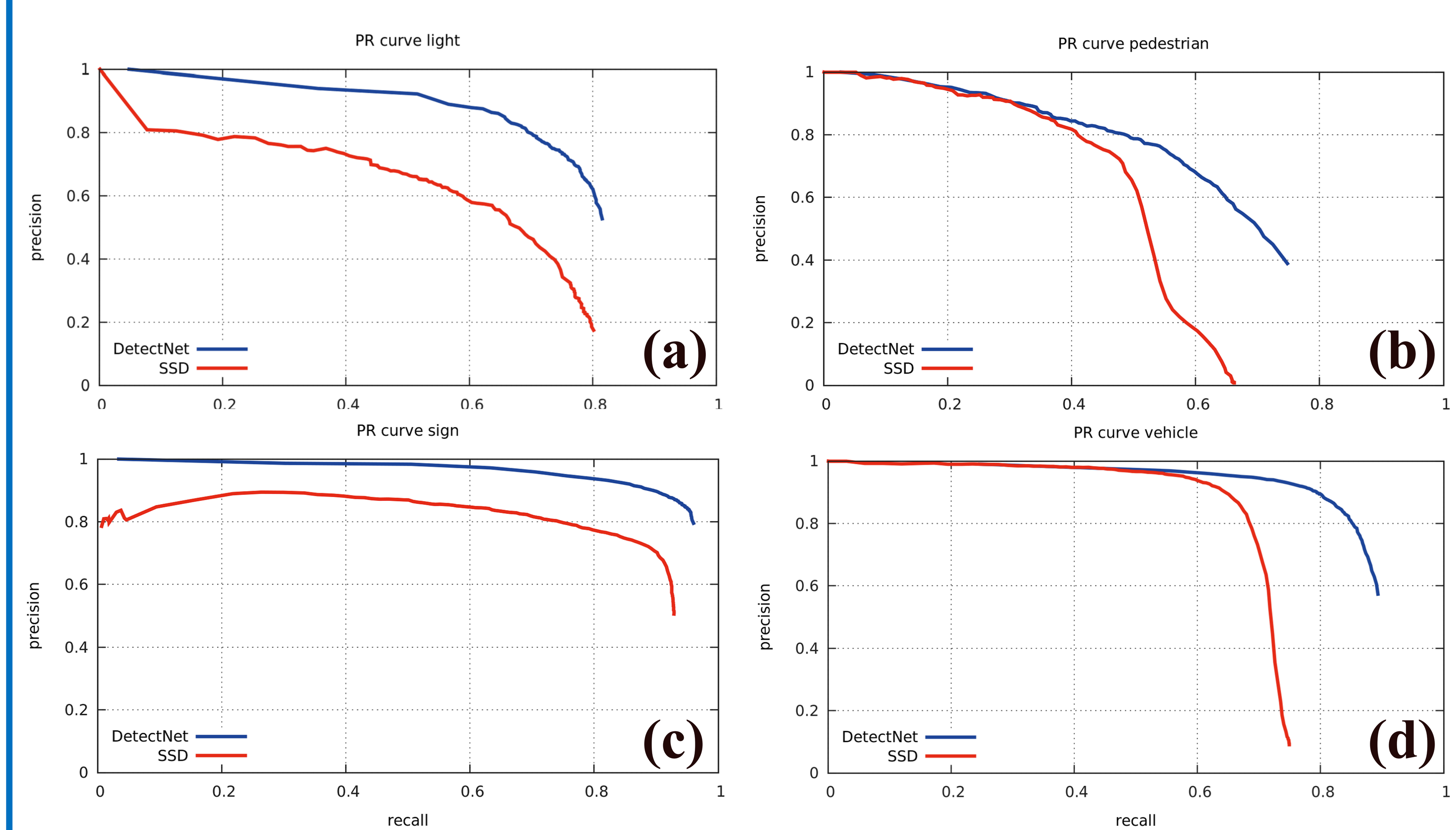

Fig. 2. PR curve of DetectNet and SSD for four types of objects in TAD16K. (a) Traffic light. (b) Pedestrian. (c) Traffic sign. (d) Vehicle.

Fig. 3. The trained models (eg. DetectNet) using TAD16K can be directly deployed for real applications. (a) PointGrey frames. (b) Basler frames. (c) Common car driving recorders. (d) Common web cameras .

Table 2. Detection results for two methods on TAD16K.

	mAP			
	Sign (T=20)	**Light (T=20)**	**Vehicle (T=50)**	**Pedestrian (T=30)**
DetectNet	0.93	0.74	0.85	0.61
SSD	0.77	0.54	0.70	0.49

3. CONCLUSIONS

We have created a new benchmark for simultaneously detecting and classifying objects concerned by self-driving cars, which are traffic signs, traffic lights, vehicles and pedestrians. Compared with previous benchmarks related to autonomous driving, images in this benchmark are more variable and have a higher resolution, and the object size in these images is in a large range. In addition, we have trained two networks (DetectNet and SSD) on this benchmark and the evaluation results of them are provided. These two methods can be used as a baseline for future research. To assist research in this field, we make this benchmark, trained models and source code public available. We hope this benchmark can act as a new challenge for the object detection and recognition community.