Towards Extreme Weather Detection: Dark Channel Priori Enhanced Detection with New Labelled Data

Yijun Li*

School of mathematics and statistics, Northeastern University Qinhuangdao, 066000, China

* Corresponding Author Email: lemon06pie@stu.ncwu.edu.cn

Abstract. With more and more cars entering our lives, the number of car accidents is also increasing. To solve this problem, driverless technology has entered our field of vision. An effective way to realize unmanned driving technology is the application of deep learning. At present, the practice of various large companies in this field is mainly to realize the integrated deep learning of image recognition, judgment, and response on automobiles. The model performance is satisfactory in normal condition. However, if the vehicles encounter some extreme weather, e.g., foggy, it is difficult to ensure the safety of deep learning model due to the insufficient accuracy of image recognition. To solve this severe problem, we advance three strategies. Firstly, considering high-capacity backbone will mitigate this situation, Yolov5 with stronger performance was used as the detector in the experiment. Secondly, image pre-process would be benefit to this foggy situation. Thus, we introduce dark channel priori defogging to our detection framework. Thirdly, since deep learning is data hungry, we labelled image collected in heavy fog weather condition and utilize them to train the network. Abundant experiment demonstrates the effectiveness of the proposed strategies. We gain a precision of 0.89734 after training the model 299 times and the safety can be guaranteed with the high accuracy of image recognition.

Keywords: driverless technology, deep learning, training set, yolov5.

1. Introduction

In the last 50 years, the mass production of cars has brought a lot of convenience to our lives, but at the same time, the accompanying car accidents have also, unfortunately, caused the loss of many lives. To reduce or even avoid the occurrence of traffic accidents. Driverless cars provide a feasible way.

However, to replace the original driver of the car, we need to assemble a new driver for the car. Every part of the new driver needs to be equipped, and its eyes are most suitable for deep learning.

However, it is unrealistic for such driverless vehicles to be applied to real life on a large scale. If all vehicles in a certain area can be arranged by the system in the situation of no machine failure, the probability of an accident is 0. But many experienced and confident drivers have difficulty choosing to believe in driverless driving, which leads to the probability that driverless and manned driving will coexist in future lanes. The role that driverless technology needs to play is becoming a reliable, sensitive， and experienced driver of human beings. Therefore, deep learning can endow driverless technology with such ability.

It should be noted that with the continuous optimization of cameras and deep learning training ability, image recognition based on photo pixel size and computer image recognition ability is also developing rapidly. Through the research= in the past ten years, it is found that deep learning before 2020 generally adopts smaller training sets. Most research companies directly train the recognition, judgment, and response of integrated images on vehicles. However. When the existing driverless image recognition systems encounter extreme weather, can they normally recognize pedestrians and vehicles?

I'm afraid the answer is no. To solve this problem, I decide to take three steps. First, I use Yolov5 with stronger performance as the detector in the experiment. Second, I use dark channel priori defogging in the experiment. Third, in the experiment, the label of the data set under heavy fog is labeled and used it to train the network.

2. Related works

In 2015, Toyota allocated 1 billion dollars for autonomous driving projects and cooperated with Stanford, MIT, and the University of Michigan. More importantly, Toyota is currently the car company with the most patents for autonomous driving, with more than 1400 patented technologies [1].

Volkswagen's driverless technology research and development began with participating in the DARPA driverless car challenge. In October 2015, Volkswagen cooperated with the University of Nevada in the United States to conduct a driverless test between Nogales, Arizona, and Mexico City, the capital of Mexico. The test distance was 2400 km, setting a record for the longest driverless test distance completed in Mexico [2-4].

On August 25th, 2016, using NASA+MIT technology, Nutonomy made the driverless taxi take the lead in Singapore to complete the trial operation of carrying passengers. The company's driverless taxi is equipped with six sets of lidar detection systems, and there are two cameras on the dashboard to scan obstacles and detect traffic light changes.To better develop the economic value of driverless cars, the company also developed cloud computing software to deploy and command driverless taxis and command driverless taxi fleets [5].

On September 13, 2016, NVIDIA launched NVIDIA DriveTM PX 2 auto, a smaller energy-saving in-depth learning computing platform, at the GPU technology conference. Some large companies (such as Qualcomm and Samsung) have also introduced their own embedded processors with GPU units. At present, Samsung processor chip mainly researches and develops automotive intelligent hardware and vehicle machine interconnection system to expand the product portfolio for vehicles and realize the perception function supported by multiple systems [6-14].

3. Methods

The method we choose to carry out the image recognition task is yolov5 and the basic training set we select is yolov5m. After refining, again and again, the YOLO has become a really powerful tool in its fifth version, based on which we select yolov5 and its related training set.

There are two deep learning classical detection methods, which are two-stage method consisting of Faster-rCNN and mask-rCNN, and a one-stage method. YOLO, as one kind of one-stage method, has many core advantages, among which fast speed and suitability for real-time detection tasks are the most appealing. And different versions of YOLO are introduced down below, from which we choose yolov5.

In the first version, yolov1 applies tensor value interpretation to grid cells and each cell has two boxes. The model of v1 is shown in the picture below.

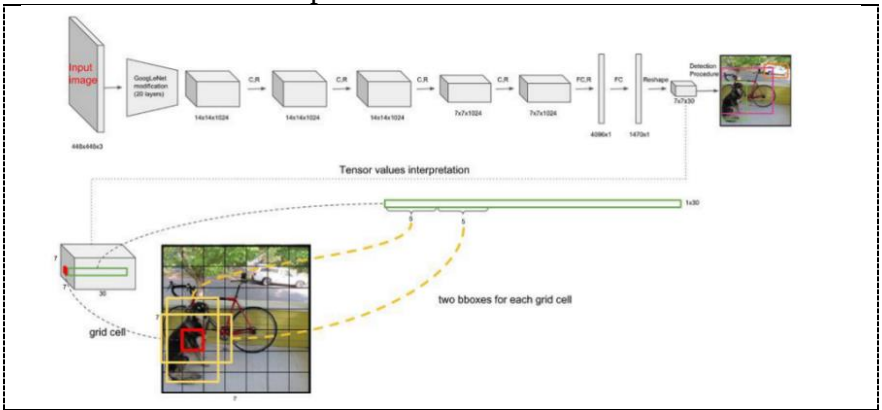


Figure 1. The model of yolov1

Then comes version 2, which had made some changes based on version 1. Dropout is dropped in version V2, and Batch Normalization has been added after convolution. yolov2 has no FC layer or 5 times down sampling. Its 1 X 1 convolution saves a lot of parameters. Besides, in version 2 users can

change the input image size after a particular iteration. And the picture down below is the network architecture of YOLO V2.

When YOLO evolves into version 3, the biggest improvement is the network structure, making it more suitable for small target detection. The features are made more detailed, and multi-continuous feature map information is integrated to predict objects of different specifications. The a priori frame is more abundant, with 3 scales, each with 3 specifications, a total of 9 types. Softmax improvement, predicting multi-label tasks.

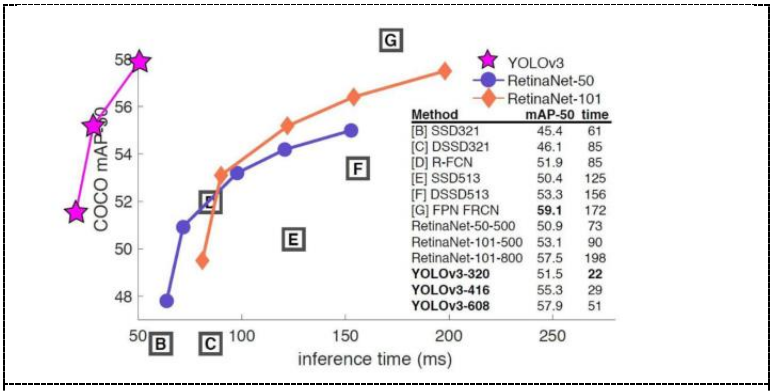


Figure 2. The mAP of yolov3

Today, almost all network architectures use the residual connection method. The idea of resnet is also used in V3, stacking more layers for feature extraction. The training effect will not be worst. And the picture down below shows the residual connection, which can help create better features.

As for yolov4, it improves from the data level and network design level. All experiments in version 4 are done on a single GPU, so experimenters don't have to worry too much about equipment.

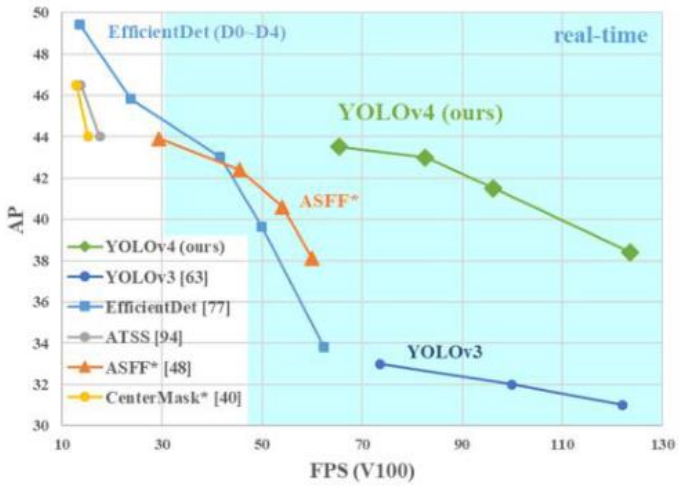


Figure 3. The model of yolov4

Finally, the yolov5 comes into our sight. To introduce version 5, I have to mention the focus module. The module first blocks, then splicing and convolution. After that are intervals to complete chunked tasks. At this time, the C of the convolution input is 12. There are not many reference experimental results, the purpose is to accelerate, and will not increase AP.

BottleneckCSP is also important in yolov5. It also includes the resnet module. Similar to the V3 version, with more CSP, the effect has been improved.

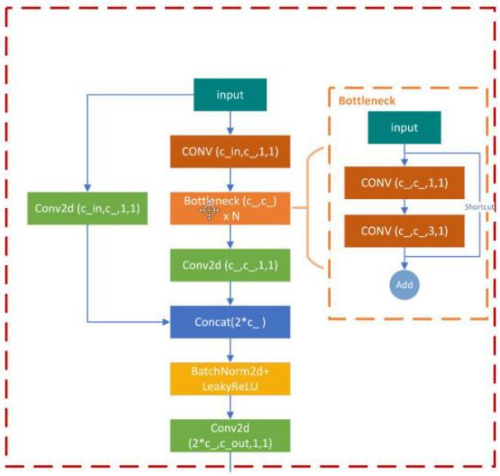


Figure 4. The BottleneckCSP of yolov5

4. Experiment

4.1. Toy experiment with poker photo

Because the image recognition training in Yolo can be strengthened again based on the already trained training set, and there are many efficient basic training sets, I chose yolov5m as the pre-training set on the GitHub official website and trained again on its basis.

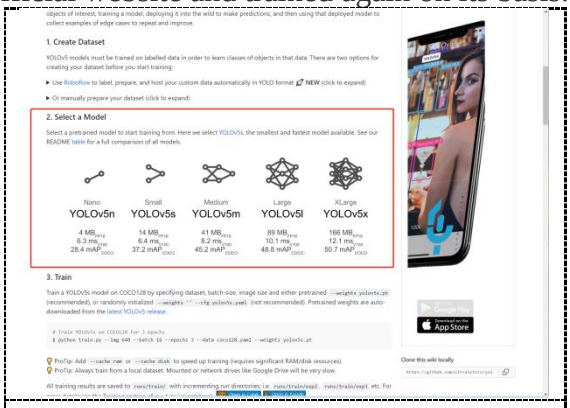


Figure 5. Pretraining set

In order to test the trainability of the training set, we used yolov5m to train 10100 poker pictures provided by Roboflow's official website, in which the weight is set to 0.4 and the number of training is 299. Identify the numbers and letters corresponding to the symbols from A to K on the playing card, and set the colors of the playing card to A, B, C, and D, as shown in the figure.

Finally, 0.975 accuracies are achieved. So we choose this training set as our basic training set.

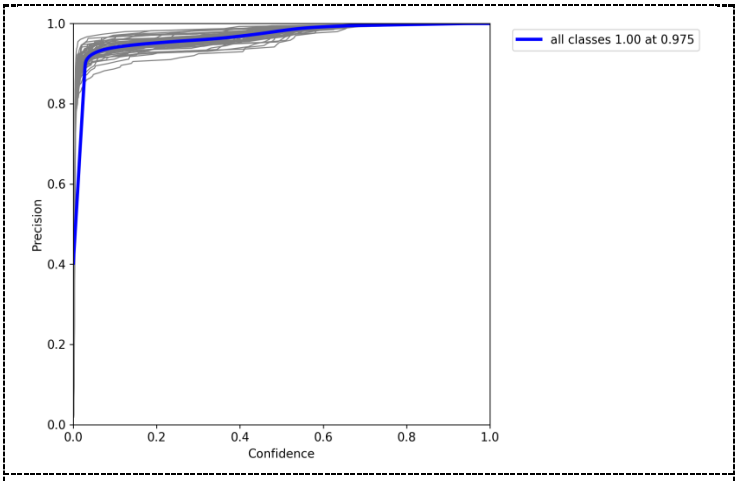


Figure 6. The precision of yolov5m in poker cards

4.2. Yolov5 with fog photo

After trying to train the recognition card, we are ready to start training road target detection in foggy weather. We collected 6813 blurry images with fog on cityscapes and used Labelme to manually label 3376 of them (Among the 6813 photos, 3376 photos are the same as the rest of the pictures, but the coefficient of fog is different, so we only mark 3376 images) and finally mark 28142 boxes. After converting the JSON data into the txt annotations required by yolov5, use these data to train on a single GPU. We use YOLOV5M as the pre-training model and select the yolov5m.YAML model for training.

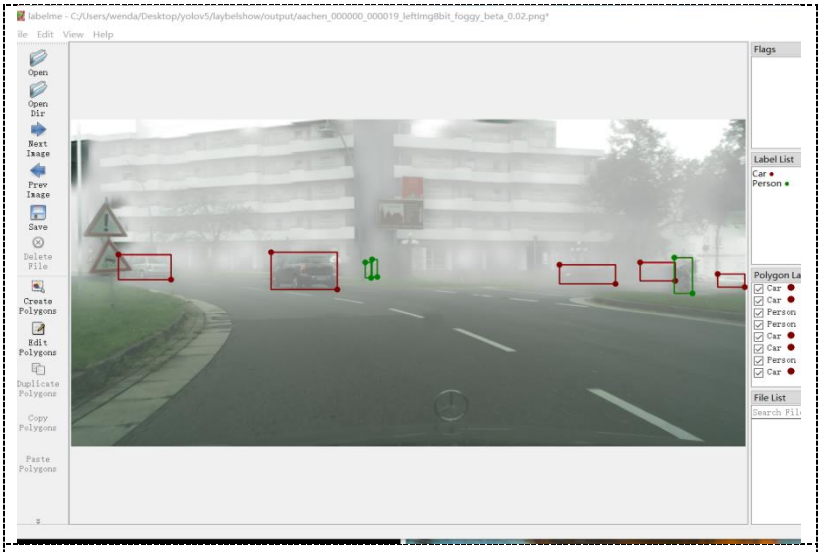


Figure 7. Labeling the pictures of foggy weather

After we labeled all the foggy pictures, we train the model in the same way as we train the poker cards.

4.3. Training results

As a result, on the second training, our neural network has already obtained excellent metrics.

After 299 training sessions: metrics/precision reached 0.89734, metrics/recall reached 0.73708 and metrics/mAP (0.5) reached 0.82891.

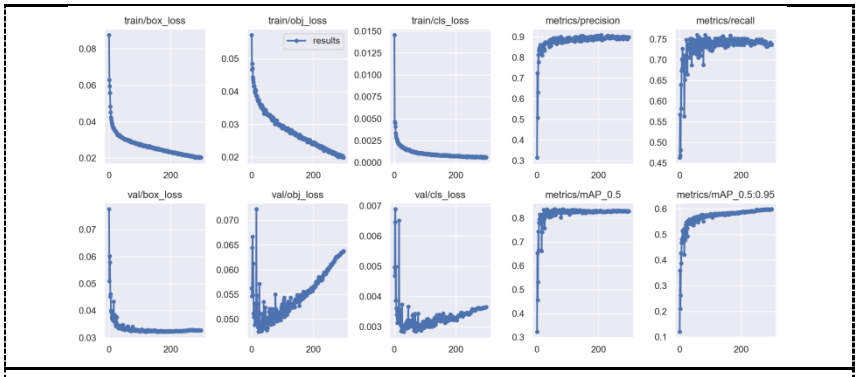


Figure 8. The result of the training of pictures of foggy weather

Table 1. The result after each particular time of training

Number	metrics/precision	*Metrics/recall*	*Metrics/mAP(0.5)*
Train 1	0.31374	0.46311	0.32275
Train 2	0.72441	0.56732	0.65319
Train 100	0.89398	0.72385	0.82425
Train 299	0.89734	0.73708	0.82891

As we can see in the pictures down below, both the cars and the pedestrians are recognized clearly.

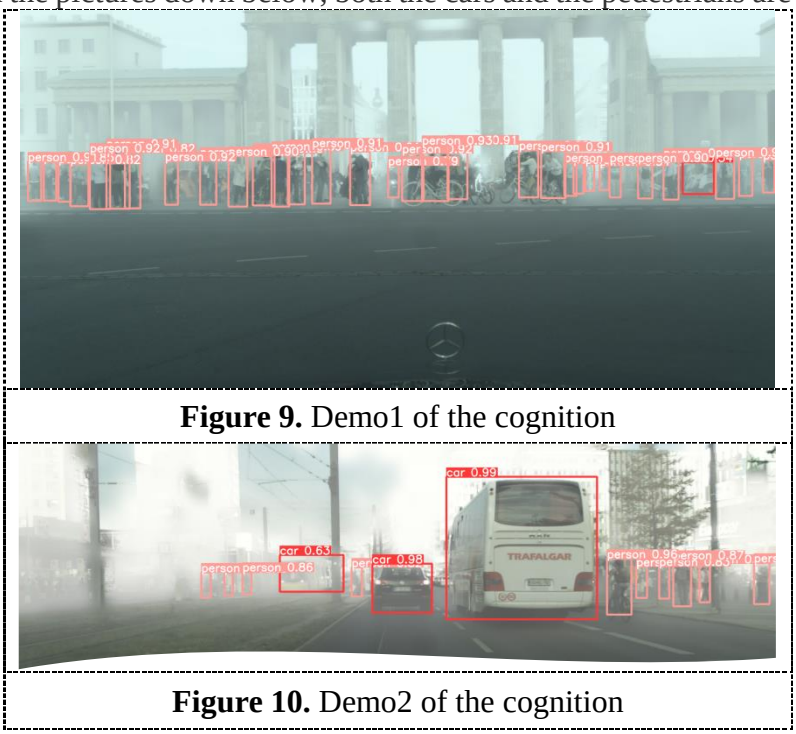


Figure 9. Demo1 of the cognition

Figure 10. Demo2 of the cognition

5. Conclusion

In general, to enhance the ability of unmanned image recognition in extreme weather, at first, Yolov5 with stronger performance was used as the detector in the experiment. Then I used dark channel priori defogging in the experiment. After that, I labeled the label of the data set under heavy fog and used it to train the network. Finally, after I examined the efficiency of yolov5m, I used it as the pre-training set to train the coded fog photos 299 times and high accuracy is achieved. We find that yolov5m is suitable for object recognition in foggy weather, and we start to wonder whether it

can be applied to other situations, for example, rainy weather, and has a great effect. I believe that after we find sufficient pictures of other extreme weather, we will begin to train.

References

[1] BRETT J A. (2016). Thinking local about self-driving cars: a local framework for autonomous vehicle development in the united states [D]. In Washington DC: University of Washing- ton.

[2] CORNICK M, KOECHLING J, STANLEY B, etal. (2016). Localizing ground penetrating RADAR: a step toward robust autonomous ground vehicle localization [J]. In Journal of field robotics.

[3] MACHT J. (2000). Automatic for the People [J]. In Tennis.

[4] CHEN Gang, ZHANG Weigong. (2015). Design of prototype simulation system for driving performance of electromagnetic unmanned robot applied to automotive test [J]. In Industrial robot: an international journal.

[5] DUAN Jianmin, SHI Hui, ZHAN Yuchen. (2016). Integrated navigation system for unmanned intelligent vehicle based on vision[J]. In Application of electronic technique.

[6] BADRINARAYANAN V, HANDA A, CIPOLLA R. (2015). SegNet: a deep convolutional encoder-decoder architecture for robust semantic pixel-wise labeling[J]. In Computer Science.

[7] KENDALL A, CIPOLLA R. (2015). Modelling uncertainty in deep learning for camera relocalization[C]. In Proceedings of 2016 IEEE International Conference on Robotics and Automation.

[8] JAYCIL Z. VARGHESE. (2015). Overview of autonomous vehicle sensors and systems[C]. In Proceedings of the 2015 International Conference on Operations Excellence and Service Engineering Orlando.

[9] HUANG Wuling, WEN Ding, GENGJ, etal. (2014). Task-Specific performance evaluation of UGVs: case studies at theIVFC [J]. In IEEE transactions on intelligent transportation systems.

[10] CHEN Xiang, QIAO Yu. (2015). Road segmentation via iterative deep analysis[C]. In Proceedings of 2015 IEEE International Conference on Robotics and Biomimetics.

[11] MONTEMERLOM, BECKERJ, BHAT S, etal. (2008). Junior: the Stanford entry in the urban challenge [J]. In Journal of field robotics.

[12] ŽBONTAR J, LECUN Y. (2016). Stereo matching by training a convolutional neural network to compare image patches[J]. In The journal of machine learning research.

[13] POMERLEAU D A. (1989). Alvinn: an autonomous land vehicle in a neural network [R]. In Pittsburgh: Carnegie-Mellon University.

[14] Liu, J., Guo, X., Li, B., & Yuan, Y. (2021, September). COINet: Adaptive Segmentation with Co-Interactive Network for Autonomous Driving. In 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 4800-4806). IEEE.