LiDAR target fusion and algorithm detection based on improved YOLO

Zhuoquan Yu¹, Jingyao Liu¹, Wenyang Xu², Yiding Liu³, Chenji Lu⁴

¹ College of Electronic and Communication Engineering, Guangzhou University, Guangzhou, 510006, China
² College of Information science and Technology, Jiageng college of Xiamen University, Xiamen, 363105, China
³ College of Computer and software, Nanjing University of information Science and Technology, Nanjing, 210044, China
⁴ Peking University, Beijing, 100871, China

*Corresponding author’s e-mail: kikyou@pku.edu.cn

Abstract. Abstract. In order to achieve safe driving behavior, the most important point of automatic driving is to detect the target. At present, the judgment of obstacles is based on single sensor, so it is difficult to detect some complex road environment, and it is easy to be missed. Therefore, this paper proposes another system device combined with color camera technology in lidar. This is another detection method proposed on the basis of Yolo, which improves the detection ability of small targets such as non-motor vehicles and people. This is based on the Yolo algorithm, using images and other samples to obtain relevant useful data, and finally build the detection system model. Finally, the sensor is introduced to combine the color image and the deep image in order to improve the detection accuracy. Finally, the fusion of decision-making level is verified by test samples. The results show that the improved YOLO algorithm and decision-level fusion algorithm have higher target detection accuracy, can meet the real-time requirements, and can reduce the miss detection rate of small and weak targets such as non-motorized vehicles and pedestrians. Therefore, the method proposed in this paper has good performance and broad application prospects, while taking into account both accuracy and real-time.

1. Introduction

In order to improve road traffic safety, autonomous vehicles have become the mainstream of world traffic development in the future. Target recognition is one of the basic links to ensure safe driving of automobile, and needs the help of various sensors. In recent years, due to the excellent performance of LiDAR and color cameras in obstacle detection and modeling, the most popular sensors now include LiDAR and color cameras.

Using color cameras to capture the traffic scene where the target is located in real time, the deep learning based detection method can provide more accurate information than traditional target detection methods, so it has become a research trend. In deep learning, convolution neural network combines artificial neural network and convolution algorithm to identify various targets. It has good
impetus to some degree of distortion and distortion [1]. YOLO is a real-time target detection model based on convolution neural network. YOLO has become a benchmark in the field of target detection because of its ability to learn large amounts of data, the ability to extract point-to-point features, and the good real-time recognition effect. Li Chen et al. clustered the initial candidate boxes, recombined the feature map, expanded the number of horizontal candidate boxes, and constructed a pedestrian detector based on YOLO (YOLO-P), which reduced the pedestrian miss detection rate. However, the YOLO model is limited to the detection of static images, which makes the detection of pedestrian dynamic changes more limited. Therefore, on the basis of the original YOLO, an extracting algorithm is designed, which can reduce the loss of feature information, by combining it with the detection algorithm DPM (deformable parts model) and R-FCN (area-based full convolution network), and is used for privacy-related situation recognition in smart home environment. However, the algorithm divides the grid of recognized images into 14 by 14. Although it can extract blurred objects, its workload cannot meet the requirements of real-time, extracting the information characteristics of gray images as the input layer of YOLO model. However, the process of extracting information to form an input layer using the alternating direction multiplier method takes more time and is greatly limited in application [2].

In order to adapt to the complexity and variability of traffic environment, some studies use color cameras and LiDAR to detect targets on autonomous vehicles simultaneously, and then provide sufficient environmental information for vehicles by fusion method. Researchers use convolution neural network method to design three detectors to extract obstacle information by combining dense depth map and dense reflection map from three-dimensional LiDAR output with color image from camera output. Researchers propose a multi-sensor fusion framework for autonomous driving in visual traffic environment, which combines the sensor information of LiDAR to achieve efficient autonomous positioning and obstacle perception through geometric and semantic constraints. However, the process and algorithm of multi-sensor fusion are too complex to meet the real-time requirements. In addition, the literature does not take into account the existence of dark targets such as pedestrians and non-motorized vehicles. Based on the above analysis, a real-time target detection system with multiple sensors (color camera and LiDAR) and multimodal (color image and LiDAR depth image) is presented. First, color and depth images of obstacles are obtained using color cameras and LiDAR, respectively, and input into the improved YOLO detection model framework. Then, after convolution and pooling, the detection bounding boxes for each pattern are output. Finally, at the decision level, the two detection boundary boxes are fused to get accurate detection targets [3].

2. Overview of system methods

2.1. LiDAR and color cameras

The sensors used in this paper include color cameras and lidar, as shown in Figure 1.
LiDAR is a 64-line three-dimensional radar detection system software. It sends the inspection data signal (laser) to the overall target, and then compares the received data signal (overall target radar echo) with the sent data signal on the overall target reflecting surface. LiDAR is installed in the middle of the top of the car, and can check the natural environment information content according to the high-speed scanner. LiDAR can send 64 lasers on the top of the head. This laser is divided into four groups, each with 16 laser emitters, and the head rotation angle of view is omni-directional, with a testable interval of 120 M. 64-line LiDAR has 64 fixed laser transmitters, and obtains the surrounding information content of each at a fixed pitch angle, and obtains a series of three-dimensional output coordinate points. Then 64 points are identified, and the distance between each point in the scene and the LiDAR is used as the sharpness value to obtain the deep image. The colorful camera is installed directly under the top LiDAR, and the adjustment position is based on the horizontal and vertical orientation of the camera image. The axis of the straight management center and the ra value of the horizontal and vertical orthogonal and the plan view produced by the laser projector make the camera angle of view and yaw angle about 0, and the pitch angle about 0. Digital images can be obtained immediately, but the LiDAR and camera output images must be paired in time and space so that the information content between them is the same [4].

2.2. Image Calibration and Synchronization
In order to integrate information in the vehicle environment awareness system, it is necessary to complete the information calibration and synchronization.

2.2.1. Information Calibration
Installation calibration for LiDAR: The midpoint of the front bumper and windshield can be measured with a tape, and the straight line of the test vehicle's central axle can be marked with a laser projector based on these two midpoints. Then, on the center axis, at a distance of 10 meters from the rear axle of the test vehicle, a straight line is marked perpendicular to the center axis; the vertical axis of the radar center is measured with a ruler, and the vertical beam perpendicular to the ground is corrected with a laser projector to make the longitudinal axis coincide with the beam, and the lateral offset of the radar is about 0 M.

2.2.2. Information Synchronization
The target detection process based on sensor fusion is shown in Figure 2. After collecting traffic scene information, the LiDAR and color cameras output depth images and color images respectively, and input them into the improved YOLO algorithm (this algorithm is trained in multiple images collected by LiDAR and color cameras). Target detection model 1 and model 2 are constructed. Then, the final target recognition model is obtained by decision-level fusion to achieve multi-sensor information fusion.
3. Improved YOLO algorithm

In the application of the original YOLO algorithm, the following problems were found: (1) Because only two boxes can be predicted for each grid cell and only one class can be predicted, YOLO imposes strong spatial constraints on the prediction of boundary boxes. This spatial constraint limits the number of nearby objects that the model can predict. (2) In the original YOLO model, the unit division of the image is set to 7*7, which can only detect large traffic objects such as buses, cars, and trucks, but cannot meet the requirements of image unit division for small and weak targets such as non-motorized vehicles and pedestrians. When the target is close to the safety of the autonomous vehicle and the confidence of the target is low, it is easy to ignore the existence of the target and cause potential safety hazards. Based on the above shortcomings, this paper improves the original YOLO algorithm as follows: (1) To eliminate the redundant time problem caused by unexpected target recognition, the total number of categories is set to (automobile, truck, non-motorized vehicle, pedestrian, and other) according to the size and driving characteristics of common targets in traffic scenes. (2) A secondary image detection scheme is proposed for the detection of non-motorized vehicles and pedestrians. Then, the cell division of the image is maintained at 7*7 and the convolution core of the sliding window is set at 3*3.
The initial YOLO optimization algorithm cannot identify and identify the overall goal based on its characteristics, and it is very likely that some overall goals will be lost; the improved YOLO optimization algorithm can be within a certain distance based on the ambiguous characteristics of non-motorized and non-motorized lanes. Check the overall goal twice, which reduces the error diagnosis rate of the overall goal, and outputs a more comprehensive scene model to ensure the safety of the car. The entire whole process includes acquiring deep and digital images. This image is replaced by the improved YOLO optimization algorithm and the proposed management decision-making level. The improved YOLO optimization algorithm involves the secondary inspection of the image grid, so it is better than the general The whole process of identification is slow. Figure 4 shows the information to complete the calculation of the different processes of the natural environment and the optimization algorithm. You can see that the average resolution time of each frame is 81ms. Fully considering that the output power of the camera is about 10 Hz, the V-type millimeter wave radar can consider the practicality of traffic travel scenarios.
4. Conclusion
This article clearly proposes an inspection integration system that integrates LiDAR and colorful surveillance cameras. Based on the initial YOLO optimization algorithm, a second inspection plan is clearly proposed, which improves the YOLO optimization algorithm and can test non-machines. Small overall goals such as motor vehicles and non-motor vehicles. Subsequently, the management decision-making level of the sensor was introduced, combined with technology, to combine the digital image of the color camera with the deep edge detection of LiDAR to improve the accuracy of target detection. In addition, at the level of optimization algorithm, this method has excellent characteristics and broad scientific research market prospects. Because the samples specified in the article are collected in several traffic travel scenarios, the coverage area of traffic travel scenarios is relatively In the future scientific research, everyone will gradually expand the diversity of scenarios and further develop trends to improve the YOLO optimization algorithm. In the next stage of the test, environmental factors will be taken into consideration. Because the identification method based on the image is at different intervals (0-50m, 20-50m, 50-100m and> 100m) at a very large level, light The compressive strength levels are different. How to deal with light compressive strength and resolution is the key basis for target detection.

References
[1] Peng Tuyou, Wu Jie, Peng Jun. Improvement of edge detection algorithm and its application in GPR [J]. Modern Radar, 2020,42(08): 41-45.
[2] Li Chen, Deng Bingguang, Zhang Zhizhong. Improved CG detection algorithm for large-scale MIMO systems [J]. Telecommunications Technology, 2020,60(07): 839-844.
[3] Wu Rui, Zhang Anqin, Tian Xiuxia, Zhang Ting. Power data anomaly detection algorithm based on improved K-means [J]. Journal of East China Normal University (Natural Science Edition), 2020 (04): 79-87.
[4] Chen Lichao, Jiedan, Cao Jianfang, Zhang Rui. Research on real-time vehicle detection algorithm based on improved optical flow method and GMM fusion [J]. Journal of Systems: 1-9 [2020-08-24].Http://kns.cnki.net/kcms/detail/23.1538.TP.20200715.1800.008.html.