

# Lane Line Detection System Based on Improved Yolo V3 Algorithm

Gaoqing Ji

Hebei University of Architecture

Yunchang Zheng (✉ [zyc2023@hebiace.edu.cn](mailto:zyc2023@hebiace.edu.cn))

Hebei University of Architecture

---

## Research Article

**Keywords:** Yolo v3, Lane line detection, Average detection accuracy, Real time

**Posted Date:** October 13th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-961172/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Lane line detection system based on improved Yolo V3 algorithm

Gaoqing Ji<sup>1,a</sup>

<sup>1</sup>College of Electrical Engineering, Hebei University of Architecture, Zhangjiakou 075000, Hebei, China

<sup>a</sup>jq1941@hebiace.edu.cn

Yunchang Zheng<sup>2,b,\*</sup>

<sup>2</sup>College of Electrical Engineering, Hebei University of Architecture, Zhangjiakou 075000, Hebei, China

<sup>b</sup>zyc2023@hebiace.edu.cn

**Abstract:** Aiming at the problems of low accuracy and poor real-time performance of Yolo v3 algorithm in lane detection, a lane detection system based on improved Yolo v3 algorithm is proposed. Firstly, according to the characteristics of inconsistent vertical and horizontal distribution density of lane line pictures, the lane line pictures are divided into  $s * 2S$  grids; Secondly, the detection scale is adjusted to four detection scales, which is more suitable for small target detection such as lane line; Thirdly, the convolution layer in the original Yolo v3 algorithm is adjusted from 53 layers to 49 layers to simplify the network; Finally, the parameters such as cluster center distance and loss function are improved. The experimental results show that when using the improved detection algorithm for lane line detection, the average detection accuracy map value is 92.03% and the processing speed is 48 fps. Compared with the original Yolo v3 algorithm, it is significantly improved in detection accuracy and real-time performance.

**Key words:** Yolo v3; Lane line detection; Average detection accuracy; Real time

## 1.Introduction

In recent years, with the rapid development of China's automobile industry and the continuous progress of artificial intelligence technology, by 2020, vehicles with auxiliary driving function and automatic driving function have accounted for more than 50% of the total number of vehicles in the market <sup>[1]</sup>. The development of intelligent automobile industry can not only effectively increase the safety factor of automobile driving, but also solve the problems of ecological environment damage caused by traffic congestion.

In the field of intelligent vehicles, the automatic detection of lane markings on the road is the technical basis for vehicle assisted driving and an important part of the perception module in driverless vehicles <sup>[2]</sup>. At present, the main methods of lane detection at home and abroad include: the method of extracting road features by machine vision, the method of establishing road model for detection, and the method of multi-sensor fusion detection <sup>[3]</sup>.

The method of extracting road features by machine vision mainly uses machine vision technology to classify the gray value features and color features of lane lines. After learning, the lane lines on the road can be detected. Because the gray value, color and other features in the image are often affected by external conditions such as light intensity and shadow, this method is used for lane line detection. The measurement is easily disturbed by environmental changes.

The method of establishing a road model for detection is to first establish a two-dimensional or three-dimensional image model of the road picture, and then compare the image in the established model with the road in the photo to be detected, so as to detect the lane line. The application scope of this detection method is relatively small, and it is only applicable to the roads with the characteristics of known templates. In addition, the algorithm has less computation during detection. Large, with problems such as poor real-time performance.

The method of multi-sensor fusion detection is to detect the lane line by means of high-definition camera, UAV aerial photography, GPS, radar and other fusion methods <sup>[4]</sup>. This method has high cost and is difficult to be applied in the large-scale practical lane line detection system <sup>[5]</sup>.

In view of the above characteristics, advantages and disadvantages of lane line detection technology at home and abroad, this paper proposes a lane line detection system based on improved

Yolo v3. The system is mainly composed of equal data preparation module, learning and training module and lane line detection module. After testing, compared with traditional machine learning algorithms such as r-cnn and r-fcn, this detection method has better robustness and operationTime and other aspects have improved significantly.

## 2 .Yolo V3 algorithm

### 2.1 Algorithm principle

Yolo is a real-time target detection algorithm proposed by Redmon in 2016. The detection speed of this algorithm can reach 45 frames per second, and its map value is more than twice that of other real-time detection algorithms. This algorithm is widely used in real-time object detection in cutting-edge technologies such as auto driving<sup>[6]</sup>.

Yolo v1 treats detection as a regression problem. First, adjust the size of the input image to a unified 448 \* 448 pixel, then convolute the image, and finally detect the target by the full connection layer.

The algorithm divides the input image into s \* s grids. If the center point of the detected target falls into a grid, the grid is responsible for detecting the target. The parameters to be predicted in each grid include: the predicted value of B boundary boxes, the confidence conf of each boundary box, and the probability pr (class I | object) of C condition categoriesFormula 1 shows the number of parameters num to be predicted for each grid<sup>[10]</sup>.

$$\text{Num}=s*s*(B*(4+1)+C) \quad (1)$$

Each bounding box includes four predicted values: X, y, W and h. where (x, y) respectively represents the coordinate value of the center point of the bounding box, and W and H are the width and height values of the bounding box. The position and size of the bounding box can be determined through the four values of X, y, W and h. the confidence conf of each bounding box refers to whether the target is included in the bounding box and the accuracy of its prediction.

The definition of confidence conf is shown in formula 2.

$$\text{Conf}=\text{Pr}(\text{Object})\times\text{IoU}(\text{Pred},\text{Truth}) \quad (2)$$

In formula (2), IoU (PRED, Truth) refers to the matching degree between the prediction box and the real boundary box. Its definition is shown in Formula 3, where area (Boxt ∩ Boxp) represents the area where the prediction frame intersects with the real frame, area (boxt ∪ Boxp) represents the area of the union of the prediction frame and the real frame. When the prediction frame completely coincides with the real frame, IOU (PRED, truth) = 1; when the two completely do not coincide, IOU (PRED, truth) = 0.

$$\text{IoU}(\text{Pr ed},\text{Truth}) = \frac{\text{area}(\text{Boxt} \cap \text{Boxp})}{\text{area}(\text{Boxt} \cup \text{Boxp})} \quad (3)$$

Pr (object) refers to the probability that the bounding box contains the target. When the target is included, PR (object) = 1; when the target is not included, PR (object) = 0.

The probability pr (class I | object) of condition category is based on the condition that the grid contains detection targets, and each grid only predicts the probability of a specific category of targets. Its definition is shown in formula 4<sup>[12]</sup>.

$$\text{Pr}(\text{class i}|\text{object})=\text{Pr}(\text{Class i})/\text{Pr}(\text{object}) \quad (4)$$

In formula 4, Pr (class I) is the probability that the detection target is a specific class target, and PR (object) is the probability that the detection target is included in the bounding box.

Loyo V2 is a faster target real-time inspection algorithm launched at the end of 2016 based on Yolo v1. After testing, the map value of loyo V2 is 76.8% at the processing speed of 67 frames per second and 78.6% at the processing speed of 40 frames per second<sup>[13]</sup>.

Compared with Yoyo V1, the main change of YOYO V2 is to reduce the image input into the network. The original 448 \* 448 pixel input image is adjusted to 416 \* 416 pixels. In addition, the basic

network layer is also adjusted. Yoyo V2 uses darknet-19 as the classification network, and the darknet-19 network is composed of 19 volume layers and 5 maximum pool layers. After testing, it is compared with the Darknet network adopted by yoyo v1. In addition, the target classification accuracy map of darknet-19 is improved by about 3.2% [15].

Yolo V3 (you only look once V3) is a single-stage target detection algorithm proposed in early 2018. It not only maintains the operation speed of the algorithm, but also improves the prediction accuracy. It is a popular real-time target detection algorithm at present. The network structure model of Yolo V3 is shown in Figure 1 [18]

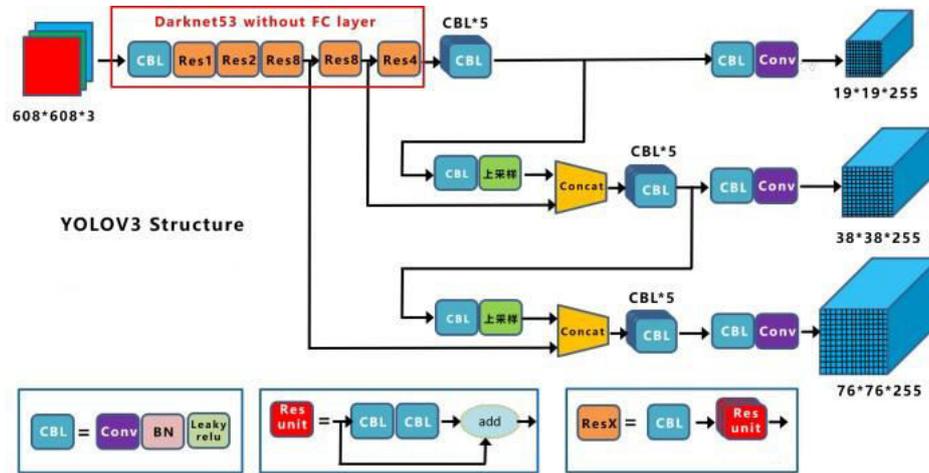


Figure 1. Network structure model of Yolo V3

Compared with the previous version of Yolo, the main differences of Yolo V3 include: using darknet-53 (106 layers in total, including 53 convolution layers) basic network model; using independent logical classifier logistic instead of the previous softmax function; using a method similar to FPN for multi-scale feature prediction. The previous version of Yolo V2 can only use 1 on a single feature map $\times 1$ . The most prominent feature of Yolo V3 is that it can detect targets in three sizes. When the feature size is  $13 \times 13$ , it is used to detect larger targets, when the feature size is  $26 \times 26$ , it is used to detect medium-sized targets, and when the feature size is  $52 \times 52$ , it is used to detect smaller targets. Therefore, it solves the function of adjusting the detection network according to the size of the target [19].

The detection steps of the three feature maps with different sizes in Yolo V3 are 32, 16 and 8 respectively. For the pictures with  $416 \times 416$  pixels input, the first detection layer is located in the 82nd layer of darknet-53 network. Because its detection step is 32, the feature map of this layer is a picture with a resolution of  $13 \times 13$ . The second detection layer is located in the 94th layer of darknet-53 network. Because its detection step is 16, the characteristics of this layer are different. The characteristic image is a picture with a resolution of  $26 \times 26$ . The third detection layer is located at layer 106 of darknet-53 network. Because its detection step is 8, the characteristic image of this layer is a picture with a resolution of  $52 \times 52$  [22].

In Yolo V3, the number of bounding boxes  $B = 3$  and the number of target categories  $C = 80$ . According to Formula 1, the number of results of the three detection layers is  $13 \times 13 \times (3 \times (5 + 80)) = 43095$ ,  $26 \times 26 \times (3 \times (5 + 80)) = 86190$ ,  $52 \times 52 \times (3 \times (5 + 80)) = 172380$ . It can be seen that in order to improve the accuracy of target detection, Yolo V3 increases the number of predictions, resulting in a decrease in the processing speed of the algorithm.

## 2.2 Algorithm related parameters

Loss function is an important parameter in Yolo V3 algorithm, which is used to indicate the inconsistency between the predicted value and the real value in the model. The training model in the algorithm aims to reduce the value of loss function. The smaller its value, the higher the robustness of the model. Commonly used loss functions include mean square difference loss function, cross entropy loss function, etc.

The definition of loss function  $L$  in Yolo V3 algorithm is shown in formula (5), which is composed of three parts: bounding box coordinate prediction error  $L_{\text{coord}}$ , bounding box confidence error  $L_{\text{conf}}$  and target classification prediction error  $L_{\text{class}}$ .

$$L = L_{\text{coord}} + L_{\text{conf}} + L_{\text{class}} \quad (5)$$

The bounding box coordinate prediction error  $l_{coord}$  is used to represent the error between the predicted value and the real value of the bounding box position. The definition is shown in formula 6.  $\lambda$  indicates the weight coefficient of coordinate error,  $S^2$  indicates that there are  $S^2$  grids in the picture, and  $B$  indicates the number of prediction bounding boxes of each grid. It indicates the possibility that the  $j$  prediction box of the  $I$  grid contains targets. If targets are included,  $I_{ij}^{obj}=1$  otherwise  $I_{ij}^{obj}=0$ .

$(x_i, y_i, w_i, h_i)$ . The four values represent the abscissa, ordinate, width and height values of the center point of the real bounding box in the  $i$ th grid. The four values represent the abscissa, ordinate, width and height values of the center point of the predicted bounding box in the  $i$ th grid.coord<sup>[25]</sup>

$$L_{coord} = \lambda_{coord} \sum_{i=0}^{s^2} \sum_{j=0}^B I_{ij}^{obj} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 + (w_i - \hat{w}_i)^2 + (h_i - \hat{h}_i)^2]$$

(6)

The definition of confidence error  $l_{conf}$  is shown in formula 7.  $\lambda_{noobj}$  represents the weight penalty coefficient, and respectively represent the real and predicted confidence. The meaning of other parameters in the formula is the same as that in formula (6).

$$L_{conf} = \sum_{i=0}^{s^2} \sum_{j=0}^B I_{ij}^{obj} [(C_i - \hat{C}_i)^2] + \lambda_{noobj} \sum_{i=0}^{s^2} \sum_{j=0}^B I_{ij}^{noobj} [(C_i - \hat{C}_i)^2] \quad (7)$$

The definition of target classification prediction error  $l_{class}$  is shown in formula 8. There are  $C$  types of targets to be detected, taking  $C = 1, 2, 3 \dots C$ . and respectively represent the real and prediction probabilities of targets of type  $C$  in the  $i$ th grid, and the meanings of other parameters are the same as those in formula 6<sup>[26]</sup>.

$$L_{class} = \sum_{i=0}^{s^2} I_{ij}^{obj} \sum_{c \in class} I_{ij}^{obj} [P_i(c) - \hat{P}_i(C)]^2 \quad (8)$$

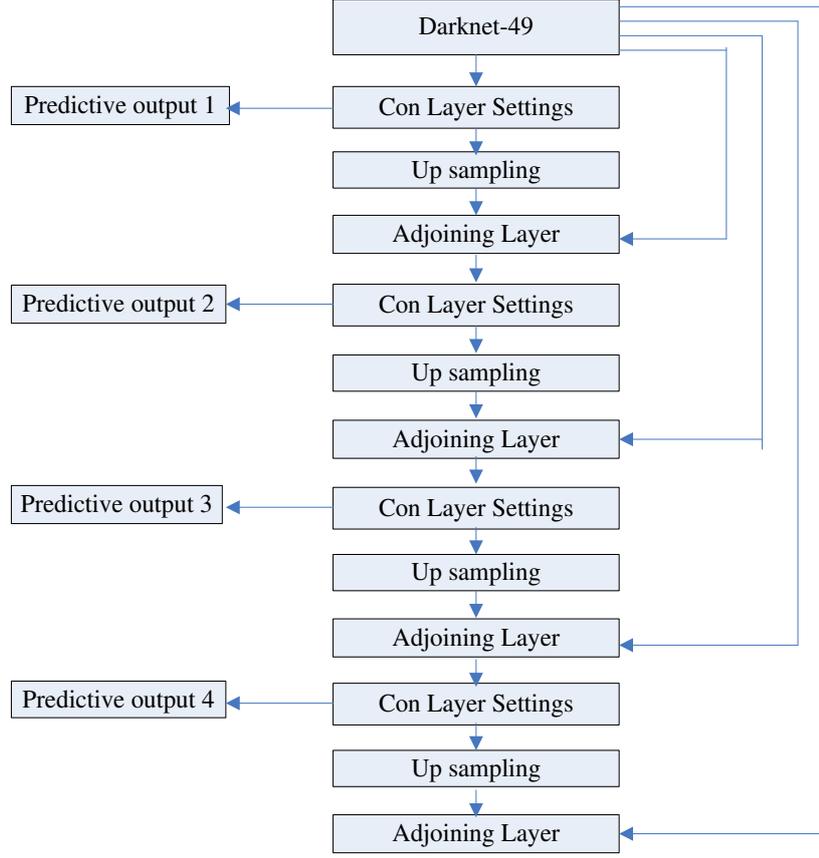
### 3. Improved Yolo V3 algorithm

#### 3.1 Improvement of network model

Firstly, in the conventional Yolo V3 algorithm, the network model is darknet-53, and the input image is divided into  $s * s$  grids. When using Yolo V3 algorithm for lane line detection, the shape feature of lane line image is that the transverse length is short and the longitudinal length is long. Therefore, in order to increase the grid detection density of image in the longitudinal direction, the algorithm is improved to divide the image into  $s * 2S$  grids. This network structure is more suitable for lane line detection system.

Secondly, aiming at the problem that the size of lane line targets is different under different road conditions, resulting in the poor detection effect of the algorithm, based on the three detection scales of yolov3 algorithm and combined with the idea of FPN algorithm, the detection scale is adjusted to four detection scales:  $13 * 13$ ,  $26 * 26$ ,  $52 * 52$ ,  $104 * 104$ . After improvement, each detection scale not only obtains the details of the bottom layer of the network, but also can be used alone. The improved detection scale is more suitable for lane line detection system than the original algorithm.

Finally, yolov3 algorithm uses darknet-53 network as the main framework. When applied to lane line detection, the problem is that the network has a large receptive field, but the spatial resolution is insufficient, which makes it easy to miss target detection when the lane line is not clear. To solve this problem, the convolution layer in yolov3 algorithm is adjusted from 53 layers to 49 layers (named darknet-49), the network structure is simplified and the loss of low-dimensional information is reduced. The network structure upsamples the input image four times to obtain four scale feature maps, which improves the detection performance. The improved network structure is shown in Figure 2.



**Figure2.** Network structure diagram of improved yolov3 algorithm

### 3.2 Improvement of algorithm parameters

#### 3.2.1 Cluster center distance parameter

Anchor parameter refers to the prediction frame with a fixed size. In the process of target detection, the size of the prediction frame directly affects the accuracy and speed of the algorithm, which is a very important parameter. In the original yolov3 algorithm, K-means clustering algorithm is used to obtain anchor parameters. In this method, different feature attributes in the distance center distance formula are treated with equal weight, and the influence of different feature attributes is not considered. In addition, if there are some noise points or isolated points in the sample, these points will have a great impact on the calculation results and cause serious deviation.

In the lane line detection system, in order to solve the problem of determining the above clustering center distance parameters, a method is proposed to replace the Euclidean distance in the original algorithm with the intersection and union ratio of the sample frame and the prediction frame. Its definition is shown in formula 9, where box is the prediction frame of the sample, CEN is the clustering center, and IOU is the area of the union of the prediction frame and the real frame (as formula 3).

$$D = \min \sum_{box=0}^n \sum_{cen=0}^k (1 - IoU) \quad (9)$$

The actual detection effect comparison between the improved method and the original yolov3 algorithm is shown in Table 1. Its average accuracy and average speed are higher than that of the original yolov3 algorithm. It shows that the improved cluster center distance is more suitable for the detection of small targets such as road routes.

**Table 1.** Comparison of detection results of cluster center parameters by different methods

Methods	Anchor	Average accuracy (%)	Average speed (frame / s)
Yolo v3	(2,8),(4,10),(5,14),(10,24),(14,33), (40,89),(100,283),(162,320)	89.32	441.78
Improved algorithm	(3,9),(6,12),(8,15),(12,28),(20,46), (49,102),(113,290),(185,348)	91.56	443.21

### 3.2.2 Loss function

The loss function is used to estimate the inconsistency between the predicted value and the real value in the model, which is the basis for the evaluation of the false detection rate by the neural network and reflects the convergence performance of the network model. The loss function in the yovv3 algorithm consists of three parts, as shown in formula 5, in which the boundary box coordinate prediction error adopts the form of mean square deviation, while the boundary box confidence error and the target classification prediction error cross Entropy loss function form.

When calculating the loss function in yolov3 algorithm, the boundary box coordinate prediction error, boundary box confidence error and target classification prediction error are activated by logistic function. The function expression is shown in formula 10 and the derivative formula of the function is shown in formula 11.

$$f(x) = \frac{1}{1+e^x} \quad (10)$$

$$f'(x) = \frac{e^x}{2e^x + e^{2x} + e^{2x}(\frac{1}{e^x})^2} \quad (11)$$

The characteristic of logistic function is that it is single and continuous, and the function value is limited, so the data will not diverge during network transmission; the characteristic of the derivative of this function is that its value is less than 1. When the input value of the network is large, the derivative value will be very small, so the error of the calculated loss function will become larger, and the convergence performance of the network model will become worse.

In order to solve the above problems, the focal loss function is proposed to replace the logistic function in the original algorithm. The expression of the focal loss function is shown in formula 12. Where  $p$  is the output value of the logistic function,  $(1-p)\gamma$  is the adjustment factor of the network system,  $\alpha$  is the weight system for target category ( $0 \leq \alpha \leq 1$ ),  $\gamma$  is the focusing coefficient ( $\gamma \geq 0$ ),  $y$  is the predicted probability value of the tag.

$$FL(p, y) = -\alpha y(1-p)^\gamma \lg p - (1-\alpha)(1-y)p^\gamma \lg(1-p) \quad (12)$$

$$L_{coord} = \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{obj} [FL(x_i, \hat{x}_i) + FL(y_i, \hat{y}_i)] + \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{obj} [FL(w_i, \hat{w}_i) + FL(h_i, \hat{h}_i)] \quad (13)$$

$$L_{conf} = \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{obj} [FL(C_i, \hat{C}_i)] + \lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{noobj} [FL(C_i, \hat{C}_i)] \quad (14)$$

$$L_{class} = \sum_{i=0}^{S^2} I_{ij}^{obj} \sum_{c \in class}^{S^2} [FL(P_i(c), \hat{P}_i(c))] \quad (15)$$

The boundary frame coordinate prediction error, boundary frame confidence error and target classification prediction error of the improved loss function are shown in formulas 13, 14 and 15 respectively. After adopting the improved adjustment factor, the reception range of small targets such as lane line is effectively increased and the detection accuracy is improved.

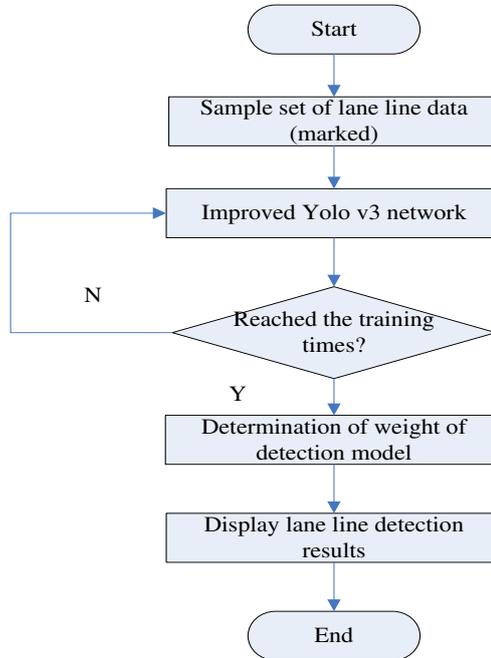
#### 4. Lane detection system design

Based on the improved yolov3 algorithm network structure, the lane detection system is mainly composed of equal data preparation module, learning and training module and lane line detection module. The implementation process is shown in Figure 3.

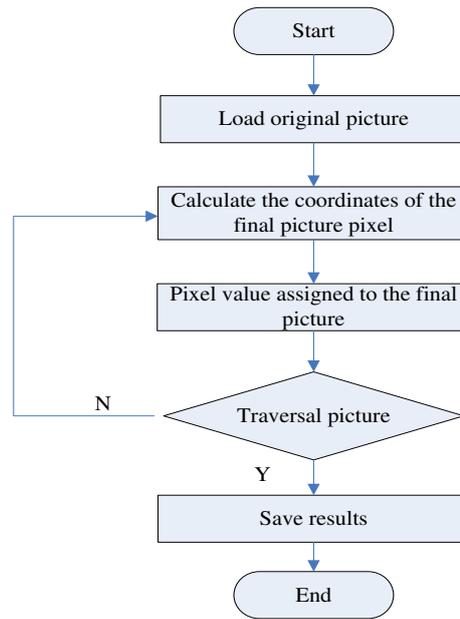
The data preparation module mainly completes the collection, marking, screening and preprocessing of lane line samples. The sample collection method is to install a high-definition camera on the car and take lane line photos on the way. Data marking mainly refers to marking lane lines in the picture, including white solid line, white dotted line, yellow solid line, yellow dotted line and other lane lines. The marking method is The main work in the screening and preprocessing stage is to select high-quality pictures from the marked pictures and preprocess them into the format required by yolo v3 neural network. The system uses nearest neighbor down sampling as the main method of picture preprocessing.

The flow chart of the nearest neighbor down sampling algorithm is shown in Figure 4. First, traverse the coordinates of each pixel in the final picture, and use the nearest neighbor interpolation method to convert it into the corresponding coordinates in the original image. For the nearest neighbor interpolation method, see formula 16, where  $t(x', y')$  is the coordinate in the final picture  $(x', y')$ ,  $s(x, y)$  is the RGB value at the original image coordinate  $(x, y)$ , and  $\text{int}$  is the rounding operation. Next, extract the pixel value of the original image coordinate, and finally assign it to the coordinates of the final image and save it.

$$t(x', y') = s(\text{int}(x + 0.5), \text{int}(y + 0.5)) \quad (16)$$



**Figure 3.** Implementation process of lane line detection system



**Figure4.** Flow chart of nearest neighbor down sampling algorithm

The preprocessed images are sent to the improved yolov3 algorithm network for lane line feature extraction. The four features of lane line are extracted at layers 67, 75, 89 and 97 of the network, and the features are sent to the Yolo layer for training. When the specified number of training times is reached, the iterative process is stopped, the final detection model is generated, and the learning and training is completed. Next, the vehicle used for testing is used. The lane line photos are sent into the trained detection model, and the system can output the lane line detection results.

## 5. Experiment and result analysis

### 5.1 Experimental environment

When the system is tested, in order to improve the computing speed and reduce the running time, the computer configuration is adopted: the CPU is Intel i7 10700f, the graphics card is NVIDIA gtx1060, the memory size is 16GB, the operating system is Ubuntu, and the interactive language is python.

### 5.2 Experimental data set

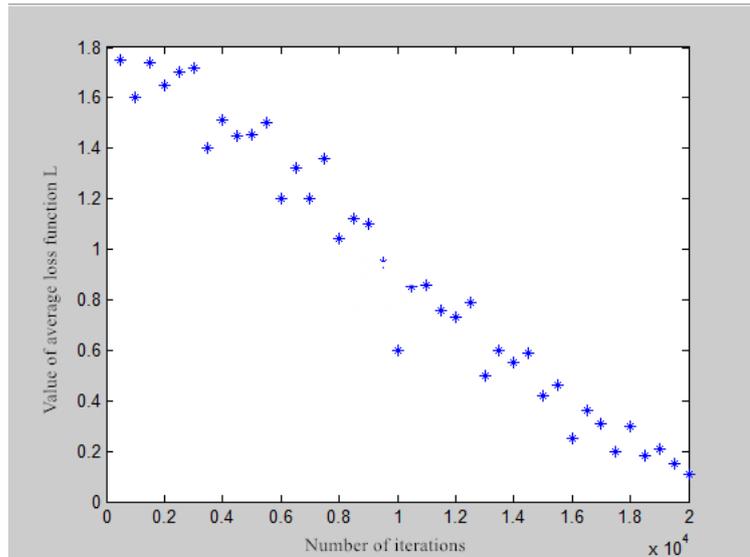
In target detection, the selection of experimental data set and feature annotation are very important, and its accuracy directly affects the effect and speed of training. During the test, a total of 500 photos containing road routes in cities were selected and grouped, including 400 learning samples and 100 test samples.

After the learning samples are preprocessed by the nearest neighbor down sampling method, the lane lines in the pictures are marked. The marking method is to mark the learning sample pictures manually by using the professional data set marking software labelme, and then save them as JSON format files.

### 5.3 Experimental results

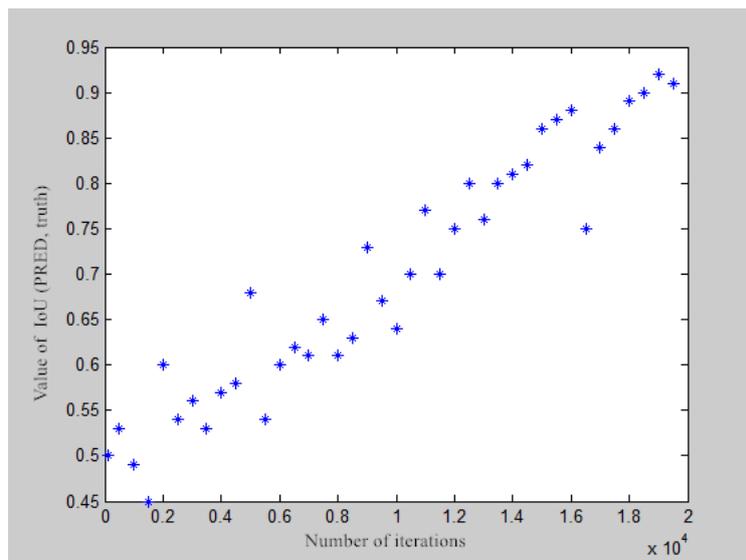
The marked lane line pictures are sent to the improved yolov3 algorithm network for training. In the training stage, the size of the pictures used is  $416 * 416$  pixels. During the training process, several important index parameters in the algorithm are dynamically recorded.

The change process of the average loss function L value is shown in Figure 5. It can be seen from Figure 5 that at the beginning of training, the loss function value is about 1.8. When the number of training increases, the convergence of the loss function value decreases, and when the number of iterations is about 20000, the loss value is 0.1 left and right to achieve the expected effect.



**Figure 5.** Variation Trend of average loss function value

The change process of IoU (PRED, truth) value of the matching degree between the prediction box and the real boundary box is shown in Figure 6. It can be seen from Figure 6 that at the beginning of training, IOU (PRED, truth) value is 0.5. When the number of training increases, IOU (pred, truth) value also increases gradually, and when the number of iterations reaches 20000, IOU (PRED, truth) value is about 0.93, and the detection accuracy has achieved the expected effect.



**Figure 6.** Change trend of matching degree value

According to the average loss function L value, IOU (PRED, truth) value and other parameters, the training times of the model are determined to be 20000. After the sample learning and training, 100 lane line photos are tested, and the system can automatically identify the lane lines in the pictures. The test results of the system under different external conditions are shown in Figure 7.



**Figure7.** System tests in different external conditions

#### 5.4 Performance comparison of different algorithms

The lane line detection system using the above improved yolov3 model has good performance in detection accuracy, map value, detection time, missed detection rate and so on. The performance parameters of the algorithm are compared with those of traditional yolov3, r-cnn and fast r-cnn, as shown in Table 2. It can be seen from table 2 that the performance of the lane line detection system of the improved yolov3 model is better than the other three algorithms in all aspects.

**Table 2.** Comparison of test performance of different algorithms

algorithm	FPS	mAP (%)	Number of iterations when $l = 0.1$	Average missed detection rate (%)
YOLO v3	8	84.87	20922	2.4
R-CNN	3	87.35	22178	4.2
Fast R-CNN	4	86.27	21896	8.1
Improved Yolo V3	12	91.03	19985	1.1

#### 6. Conclusion

Aiming at the problem that the traditional algorithm in lane detection can not give consideration to the detection accuracy and detection speed, a detection system based on the improved yolov3 algorithm is proposed in this paper. The main improvements include the following aspects:

1. According to the characteristics of inconsistent vertical and horizontal distribution density of lane line images, it is proposed to divide the images into  $s * 2S$  grids to improve the vertical detection density.

2. The detection scale is adjusted to four detection scales:  $13 * 13$ ,  $26 * 26$ ,  $52 * 52$ ,  $104 * 104$ , which is more suitable for the detection of small targets such as lane lines

3. The convolution layer in the original yooV3 algorithm is adjusted from 53 layers to 49 layers, which simplifies the network and improves the system performance.

4. The parameters such as cluster center distance and loss function are improved, which are more relevant and more suitable for lane line detection environment.

The actual test shows that the improved algorithm has good detection performance when detecting flat roads, but when the roads have large slopes, the detection is easy to be affected. Therefore, solving the problem of lane line detection in large slope scenes will be the focus of the next research.

## Acknowledgements

This work was supported by Basic Scientific Research Project in Hebei Province (No. 2021QNJS13, No. 2021QNJS06);

This work was supported by Project of Zhangjiakou science and Technology Bureau (No. 1911002B).

## Reference

- [1] HG Zhu: An Efficient Lane Line Detection Method Based on Computer Vision[J], 2020, 1802-032006
- [2] C Y Fu, W Liu, A Ranga. DSSD: Deconvolutional single shot detector[J] arXiv preprint arXiv: 1701. 06659, 2017
- [3] Lee S, Kim J, Kweon I S. VPGNet: vanishing point guide network for lane and road marking detection and recognition. In: Proceedings of the International Conference on Computer Vision. Venice, Italy: 2017. 2380–7504
- [4] Gupta A and Choudhary A. 2018. A framework for camera-based realtime lane and road surface marking detection and recognition. IEEE Transactions on Intelligent Vehicles, 3 (4): 476-485 [DOI: 10. 1109 /TIV. 2018. 2873902]
- [5] Zhang X, Yang W, Tang X L and Liu J. 2018b. A fast learning method for accurate and robust lane detection using two-stage feature extraction with YOLO v3. Sensors, 18 (12): #4308 [DOI: 10. 3390 /s18124308]
- [6] Dai Xiaogang. Research on lane detection and classification based on deep learning [D]. Sichuan: Southwest Jiaotong University, 2018: 45-46
- [7] Li Mao-Hui, Wu Chuan-Ping, Bao Yan, Fang Zhuo-Qun. On the application principle of YOLO algorithm in machine vision. Journal of Computer Applications, 2018, 5(41): 174–176
- [8] Jiang ting. Development status and future prospect of unmanned sensor system [J]. China Equipment Engineering, 2018, 2018 (21): 180-181
- [9] Shi Linjun, Yu Su. Lane line detection method based on Hough transform under multiple constraints [J]. Computer measurement and control, 2018, 26 (9): 9-12
- [10] Deng G, Wu Y F. Double Lane Line Edge Detection Method based on Constraint Conditions Hough Transform[C]. IEEE 2018 17th International Symposium on Distributed Computing and Applications for Business Engineering and Science, Jiangsu, China, 2018: 4
- [11] Redmon J, Divvala S, Girshick R, et al. You Only Look Once: Unified, Real-time Object Detection[C]. IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 779-788.
- [12] Redmon J, Farhadi A. YOLO9000: Better, Faster, Stronger[C]. Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017: 6517-6525.
- [13] Redmon J, Farhadi A. YOLOv3: An Incremental Improvement[C]. Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018: 1-4.
- [14] Sun Ting, Qi Yingchun, Geng Guohua. Moving target detection algorithm based on inter frame difference and background difference [J]. Journal of Jilin University (Engineering Edition), 2016, 46 (04): 1325-1329

- [15] Xu Jiuqiang, Jiang Pingping, Zhu Hongbo, Zuo Wei. Improvement of vibe algorithm for moving target detection [J]. Journal of Northeast University (NATURAL SCIENCE EDITION), 2015,36 (09): 1227-1231
- [16] Zheng Yuping. Detection and extraction of moving targets in complex background based on embedded platform [D]. Shanghai Normal University, 2020
- [17]Redmon J, Farhadi A.YOLO9000: better, faster, stronger[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition, 2017: 7263-7271.
- [18]Science - Applied Sciences; Shenyang Institute of Automation Researchers Provide New Data on Applied Sciences (The Application of Improved YOLO V3 in Multi-Scale Target Detection)[J]. Science Letter,2019
- [19]Tian Juan-Xiu, Liu Guo-Cai, Gu Shan-Shan, Ju ZhongJian, Liu Jin-Guang, Gu Dong-Dong. Deep learning in medical image analysis and its challenges. Acta Automatica Sinica, 2018, 44(03): 401–424
- [20] Liang Leying. Research on lane detection algorithm based on deep learning [D]. Beijing: Beijing Jiaotong University, 2018: 26-30
- [21] Zhang zemiao, Huo Huan, Zhao Fengyu. Overview of target detection algorithms based on deep convolution neural network [J]. Small microcomputer system, 2019, 40 (9): 1825-1831
- [22]Olson C F. Constrained Hough Transforms for Curve Detection[J]. Computer Vision & Image Understanding, 2016, 73(3): 329-345.
- [23]Lee S, Kim J, Kweon I S. VPGNet: Vanishing Point Guide Network for Lane and Road Marking Detection and Recognition[C]. International Conference on Computer Vision, Venice, Italy, 2017: 2380-2384.
- [24] 4 Wei Yong Ming, Quan Ji Cheng, Hou Yu Qing Yang. Aerial image location of unmanned aerial vehicle based on Yolo v2. Laser and Optoelectronics Progress, 2017, 54 (11): 101 – 110 (Wei Ming, Quan Jicheng, Hou Yu Qingyang. Research on unmanned aerial image location based on Yolo v2. Progress in laser and optoelectronics, 2017, 54 (11): 101 – 110)
- [25]S Ren, K He, R Girshick, et al. Faster r-cnn: towards realtime object detection with region proposal networks [J] .IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39( 6) : 1137—1149
- [26]Redmon J, Farhadi A. YOLOv3: an incremental improvement. In: Proceedings of the Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: 2018. 1–4



**Gaoqing Ji** was born in Zhangjiakou, Hebei,P.R. China, in 1987. He received the master's degree from the University of Electronic Science and technology. Now, he works in Hebei University of Architecture. His research interests include artificial intelligence control and machine vision

E-mail: [jgq1941@hebiace.edu.cn](mailto:jgq1941@hebiace.edu.cn)



**Yunchang Zheng** was born in Zhangjiakou, Hebei,P.R. China, in 1990. He received the master's degree from the University of Electronic Science and technology. Now, he works in Hebei University of Architecture. His research interests include Signal and information processing

E-mail: [zyc2023@hebiace.edu.cn](mailto:zyc2023@hebiace.edu.cn)