Research on Traffic Light Detection and Recognition Based on Computer Vision

Xiu Ru Li[1,2], FengQin[1,2], Dong Xu Liu[1,2],YingMing Wang[1],Wan Sheng Wang[1],QianYin[1],Jun Lin[1],ZhiQiang Qian[1],Zhen Hu Hu[1]
[1](Tencent Cloud Big Data College, Ma 'anshan University, China)
[2](School of Computer Science, Anhui University of Technology, China)

Supported by: 2022 Ma 'anshan University National College Students Innovation and Entrepreneurship Training Program，Project number: 202213614004X

ABSTRACT : *In recent years, the number of traffic accidents in China has exceeded 200,000, and the direct property losses from traffic accidents have reached 1.3 billion yuan. The occurrence of traffic accidents is mainly caused by the disharmony between the vehicle, the driver and the traffic environment. For example, the driver can not effectively identify the actual situation of the traffic signal light due to fatigue. On the other hand, unmanned driving technology requires robots to effectively identify the real road conditions and traffic signal lights, so intelligent detection and recognition of traffic signal lights is very important in assisting driving and unmanned driving. The detection and recognition of traffic lights are restricted by many factors, including noise generated during image acquisition and weather factors. In recent years, the rise of neural networks has driven the development of computer vision, and single-stage algorithms represented by YOLO and SSD have gradually become a research hotspot in the direction of target recognition. According to actual needs, this article describes the system based on YOLOv5 deep learning framework, design and implement a set of traffic lights detection and recognition system, the system after the training, can be effective circular traffic light color information, and the experimental effect is good, the implementation of the system for traffic lights direction of target detection and target identification has a certain reference value..*

KEYWORDS -*Traffic lights; Target recognition; YOLOv5*

I. BACKGROUND INTRODUCTION

In recent years, China's economic development is very rapid, per capita income is continuously rising, more and more families in order to convenient transportation, buy cars as a means of transportation, in recent years, car travel has become people's general behavior, become an important means of transportation in people's life. In 2021, there will be 400 million motor vehicles in China, and the increase of vehicles will increase the probability of traffic accidents at traffic junctions. According to the analysis of China's road traffic accidents in 2021, the number of traffic accidents in China in recent three years was more than 200,000, and the direct property loss amount of traffic accidents was 1.313 billion yuan. Traffic accidents caused by personnel driving vehicles are mainly caused by the inharmonious relationship between the vehicle, the driver and the traffic environment. For example, the driver can not effectively identify the actual situation of the traffic signal light due to subjective or objective factors, and violate traffic rules, which is the inducement of traffic accidents. In addition, unmanned driving technology requires robots to effectively identify the real road conditions and traffic signal lights, so intelligent detection and recognition of traffic signal lights is very important in unmanned driving and assisted driving.

1.1 PURPOSE AND SIGNIFICANCE OF THE STUDY

Image recognition, as a research hotspot in the field of image processing, has been widely studied. In this study, image recognition is divided into two main stages: mathematical modeling of object shape. Secondly, based on this model, various shapes of complex images are classified and the desired object shapes are extracted. provide reliable decision basis for city managers .

Prototype - like object is the key part of detection in various images. These targets mainly include round and rounded objects. In underwater operation, most of the inspection object parts are circular artificial objects, so circular object detection has a broad development prospect in this field. In order to identify these circular objects, we should first analyze them, thoroughly master and master the mathematical characteristics, and then study the circular object detection algorithm based on these characteristics that can be widely used with various fields. The purpose of this paper is to develop algorithms for circular object detection in complex images.

1.2 RESEARCH STATUS AND EXISTING PROBLEMS AT HOME AND ABROAD

Methods and tools for addressing object detection vary. Hough transform is the most widely used in the research. The following is the description of the content, development history and the advantages and disadvantages of Hough transform.

Hough transform is an algorithm to describe the edge state of a region, and it is also an important tool to study circular object detection. The main idea is to transform the original image space into the parameter space, so as to obtain the characteristics of the image space, and then use the obtained parameter space to select the most parameters. The good anti - noise performance is an important reason why transform is widely used in line detection. In addition, the calculation is easy and fast, and it is not sensitive to the incomplete part in the image. However, for circle detection with a parameter space larger than two dimensions, there is a large computation, and there is a drawback called point-to-face or line mapping. The time complexity and space complexity are both very high and occupy a large amount of computer memory [1].

When the transform has obvious shortcomings, it cannot meet the requirements of practical application, so the basic principle of creating random transform in the future is to randomly sample several points in the complex image and determine the circular object detected by them. However, even if the complexity of the existing random transformation in time and space is reduced due to the presence of noise and complex image background interference, random sampling may lead to a large amount of invalid accumulation.

In order to obtain more effective and accurate test results, experts and scholars improved the adaptability again. First of all, it is necessary to select the element information of the sampling point to determine whether to accumulate, so as to reduce the processing time of invalid information. Second, the selected flood actinization number is reduced, so the probability of error sampling can be greatly reduced [2].

II. OBJECT DETECTION METHOD

Traditional object recognition and detection methods are generally divided into three steps: firstly, some regions are selected as candidate regions on a given image, then the required features are extracted from these regions, and finally the trained classifier is used for classification. This is shown in Fig.1. We describe each of the three stages below.

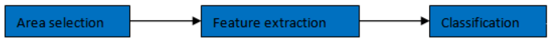

Fig.1Traditional object recognition and detection consists of three steps

(1) Region selection: This step is to locate the position of the target. Since the target may appear in any position of the image, the size and aspect ratio of the target are not fixed, so the method of sliding window is first applied to traverse the entire image, and different sizes and aspect ratios need to be set, so sliding Windows of different sizes are used. This enumeration method, which includes all possible directions of the target, also has prominent disadvantages. (Disadvantages of sliding window: Limited time complexity of the problem, because the object detection by different sizes, so the length and width of sliding window ratio is usually set several fixed, therefore, when multiple classification objects within the scope of the aspect ratio change there are big differences between the cases, even through the sliding window scanning, also unable to get ideal segmentation area and need a lot of calculation, Operation speed is slow, need a better classifier).

(2) Feature extraction: Due to different kinds of objects, different lights, different environments and other reasons, this determines a robustness (robustness refers to robust and strong, it is also the ability of the system to survive in abnormal and dangerous situations). Is not that simple. However, the advantages and disadvantages

of these characteristics are directly related to the accuracy of the analysis. The commonly used features in this stage include SIFT (scale-invariant feature conversion) HOG (orientation gradient histogram algorithm, used for object detection features) and so on.

(3) Classifier: SVM is commonly used in practice (Support vector machine, also known as vector network, is a supervised learning model and related learning algorithm that uses classification and regression analysis to analyze data).

It can be seen that there are two major problems in traditional object detection: first, there are redundant Windows and time-consuming. Second, it is not very strong and robust to the change of diversity.

Object detection is a relatively mature field in artificial intelligence, among which algorithms are generally divided into two categories:

The first category: One-stage series algorithm, or deep learning object detection method based on regression algorithm, is represented by YOLO,SSD, etc. It directly classifies based on anchor (in YOLO algorithm, there are anchor boxes with initial length and width for different data sets) and adjusts bounding boxes. Unified as a regression problem.

The second category: Two-stage series algorithm (divided into object category and object location problem or classification problem and regression problem). The two algorithms have their own characteristics and shortcomings. The first one is fast but not accurate, and the second one is high but slow.

2.1 DEEP LEARNING OBJECT DETECTION ALGORITHM BASED ON REGION PROPOSAL

In view of the problem of sliding window,region proposal gives a very good solution methods. Region proposal is to find out in advance where the target in the diagram might exist. As the Region proposal uses a large amount of texture, boundary, color and other information in the picture, it can ensure high recall when fewer Windows are selected (thousands or even millions of Windows). Thus, the problem of long time and large complexity of the first problem in traditional object detection is greatly reduced.

2.1.1R-CNN

After the target region is determined, the next thing to do is to extract the features of the target from the target region and classify it. In image classification, convolutional neural network plays a leading role in image classification. Microsoft's latest ResNet (which can make the model trained by deep network better than shallow network, and can make the model trained by deep network. The proposed new network structure) and Google's Inception V4 model (Inception network is a CNN classifier designed with excellent local topology) can get better performance. Therefore, it is also a good method to conduct image analysis on candidate regions through CNN after target detection. The R-CNN framework is designed, which makes a great breakthrough in conventional object detection and opens a new beginning of object detection based on deep learning.

In CNN (convolutional neural network), the neurons in the convolutional layer are only connected to the local neuron nodes in the previous layer, while the weights of local neurons in the same layer are shared, thus reducing the total amount of training parameters. CNN generally includes: input layer, convolutional layer, excitation layer, pooling layer, fully connected layer, output layer, etc.

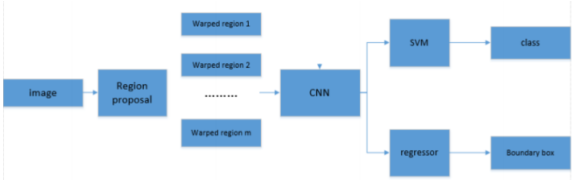

Fig.2 R - CNN frame.

In Fig.2, the frame diagram clearly shows the target detection process of R-CNN:

(1) Input the test image

(2) Using Selective Search algorithm (a region algorithm as the detection box, which is an algorithm for classification and merging based on image texture, size and shape), about 2000 region proposals are extracted from the image.

(3) Each candidate region is scaled and input to CNN to obtain a feature vector.

(4) Input the feature vectors obtained from each candidate region to SVM for analysis. To test, we first need to train the CNN model for feature extraction and SVM for classification: The CNN model for feature extraction is obtained by fine-tuning the pre-trained models LeNt5, AlexNet and VGG16 on ImageNet (a huge image library with multiple nodes), and then the CNN model is used to

extract features from the training set to train SVM.

Bounding -box regression is performed for Region proposal classified by SVM. Bounding regression is a linear regression algorithm to correct Region proposal. To ensure that the window obtained by Region proposal is consistent with the actual target window. Because the window obtained from the region proposal cannot be as accurate as that manually marked by the human, if the location of the region proposal is significantly different from that of the target, even if the classification is correct, However, since IoU(the ratio of the intersection ratio of the window between region proposal and Ground Truth, the higher the correlation degree, the higher the value) is lower than 0.5, it means that the target is still not detected.

The detection results of R-CNN on PASCAL VOC2007 (VOC dataset is a common dataset for object detection) increased directly from 34.3% of DPM HSC (which refers to the use of histogram of sparsely coded HSC to replace HOG) to 66% of mAP (AP is the precision of a category, MAP is the precision of all categories). Such a big improvement lets us see the strength of Region proposal+CNN.

Disadvantages of the R-CNN framework:

1. The training is divided into several stages with tedious steps: fine-tuning network + training SVM+ training border regressor

2. Training time and a lot of hard disk space: 5000 images produce hundreds of gigabytes of feature documents

3. Slow speed: If GPU is used, VGG16 model processing can only solve one piece every 47 seconds.

4. Image shape transformation: the candidate region needs to be scaled or intercepted for fixed size, which cannot guarantee the image deformation. Interception may cause truncation, and scaling may cause the object to be stretched.

First take a look at why the R-CNN detection speed is so slow, 47 seconds to solve one. It can be seen from the process of R-CNN that after the 2000 candidate regions were extracted from the image, he treated each region as an image and started post-processing. In fact, he had carried out the steps of feature extraction and classification for 2000 times on an image. Because we can only propose the features of the convolution layer once for the image, and then we only need to map the candidate region to the feature map of the convolution layer at the position of the original image. In this way, we only need to mention the features of the convolution layer once for an image, and then input the features of the convolution layer of each candidate region to the fully connected layer for subsequent operations. (Compared with CNN, most of the operation is consumed in convolution operation, which can save some time and resources). However, the current situation is that the scale of each region proposal is different, so it is definitely impossible to directly input the fully connected layer in this way, because the input of the fully connected layer must be a fixed size.

2.1.2SPP-NET

By mapping the attributes of these different Windows to the same dimension and treating them as fully connected inputs, we can ensure that the convolutional layer features are extracted only once from the image. Spp-net uses spatial Pyramid pooling: Each window is divided into 4*4, 2*2, 1*1 blocks, and each block is sampled, so that each window gets a feature vector of (4*4+2*2+1)*512 dimensions after passing through the SPP layer, which is used as the input of the fully connected layer for subsequent use.

After adding the SPP layer, and whatever input size is, the SPP can produce the output of a fixed size, and the ability to support multiple Windows, greatly reducing the convolution computation, enhance the accuracy in experiments, and compared with R CNN, SPP - CNN also can improve the speed of the target detection.

Disadvantages:

1. The training process consists of multiple stages and tedious steps.

2. Spp-net stabilizes the convolution layer when fine-tuning the network. As long as there is a new target, it is necessary to fine-tune both the fully connected layer and the convolution layer. Besides semantic information, the feature information extracted by the classification model also needs the location information of the target.

2.1.3 FAST R-CNN

Aiming at the above two problems, RBG proposes Fast R-CNN, a concise and Fast target detection framework. Compared with the R-CNN frame diagram, an ROI pooling layer and loss function are added after the convolutional layer is finished again, and multi-task loss function is adopted to directly add border regression to the CNN network for training.

The ROI pooling layer is a simplified version of SPP-NET because various size pyramid maps are used for each proposal in SPP-NET, while the ROI pooling layer only requires a 7x7 feature map to be sampled. For the VGG16 network (a network pre-trained based on a large number of real image libraries) CONV5_3 has 512 feature maps, such that all region proposals correspond to a feature vector of dimension 7*7*512, which is used as the input of the fully connected layer.

R-cnn training process is divided into three steps, and Fast R-CNN directly uses softmax (normalization index) to replace SVM classification, and uses multi-task loss function border regression to also add to the network. Therefore, the complete training is from end to end (excluding region proposal extraction process). In the process of network fine-tuning, Fast R-CNN also fine-tunes part of the convolutional layers to achieve better detection results [3].

Fast R-CNN combines the essence of R-CNN and SPP-NET, and introduces multi-task loss function at the same time, which makes the training and testing of the whole network very simple. The input of the convolutional neural network for feature extraction in Fast R-CNN is the whole image instead of each proposed region.

Still lacking: Selective Search algorithm is adopted to propose Region proposal, but most of the time needed is consumed in this process (Region proposal 2~3s, while feature classification is only 0.32s). In addition, the region proposal can not be selected by Selective search for the exact degree of end-to-end training test.

2.1.4FASTER R-CNN

In order to solve the shortcoming of long extraction time, the quality of region proposal affects the accuracy of target detection. If a method is found that provides only a few hundred Windows and has a very high recall rate, this method can not only improve the speed of object detection, but also improve the performance of object detection.

Therefore, Faster R-CNN proposed a new neural network called RPN(Region Proposal Network). The core of RPN is that convolutional neural network can be used to directly form Region Proposal, and the method adopted is essentially sliding window. The design of RPN is quite exquisite, because RPN only requires sliding on the last convolutional layer once, because the Anchor mechanism and border regression can obtain region proposals with more scale and aspect ratio.

After selecting the image, the feature map of the last layer is obtained by convolutional operation. Then, on this image, the feature map is convolved with a 3*3 sliding window to get the last convolutional layer. The last convolutional layer has 256 feature maps and is generated by RPN. After screening, a region proposal that is close to the target is obtained.

At the same time, it can make the RPN network and Fast R-CNN network realize the weight sharing of convolutional layer, during the trainingRPN and Fast R-CNN used a four-stage training method:

(1) By initializing the network parameters of the pre-trained model on ImageNet, the RPN network can be fine-tuned;

(2) In the application stage (1), RPN network is used to extract region proposal to train Fast R-CNN network;

(3) The Fast R-CNN network of (2) was used to re-initialize the RPN and fine-tune the fixed convolutional layer;

(4) Fix the convolutional layer of Fast R-CNN in (2) to fine-tune the region proposal extracted by RPN in (3).

The RPN and Fast R-CNN after weight sharing will improve the accuracy of target detection.

Through the trained RPN network and given the test image, the region proposal after edge regression can be intuitively obtained. Then, according to the category score of region proposal, the RPN network can be ranked. And select the first 300 Windows as the target detection input window of Fast R-CNN,VOC2007 test set test mAP can reach 73.2% (Selective Search + Fast R-CNN is 70%), and the speed of target detection can do 5 frames per second. It should be noted that the latest version has combined RPN network and Fast R-CNN network and merged candidate box extraction into deep network. This is the contribution of Faster R-CNN, which is a framework for end-to-end target detection using one CNN network [4].

Faster R-CNN combines the long-separated region proposal with the traditional CNN division method, and realizes the target analysis through end-to-end network. Therefore, FasteR R-CNN improves both in efficiency and accuracy. However,FasteR R-

CNN still cannot achieve the real target detection, because the region proposal has been obtained in advance and the calculation of each proposal type is quite heavy. But in contrast, the emergence of target monitoring methods such as YOLO makes real-time performance possible.

In general, along the way from R-CNN, SPP-NET, Fast R-CNN and Faster R-CNN, the process of object detection based on deep learning has become more and more streamlined, with higher accuracy and Faster speed. It should be said that the R-CNN series target detection method based on Region proposal is the most important branch of current target detection.

2.2 DEEP LEARNING OBJECT DETECTION ALGORITHM BASED ON REGRESSION METHOD

The algorithm of Faster R-CNN is the most commonly used target detection method at present. However, in the process of processing, Faster R-CNN cannot achieve high real-time speed. Methods such as YOLO have become increasingly important, using the idea of regression, in which targets in that area are directly regress across multiple locations in the image.

2.2.1 YOLO

YOLO(You Only Look Once) series algorithm is the most used target detection algorithm in the current series of algorithms. Its biggest feature is its fast detection speed and increasing accuracy, so it is called the most popular target detection algorithm [5].

Flow chart of YOLO object detection in Fig.3:

(1) Given an input image, first divide the image into a 7*7 grid.

(2) For each grid, we expect two borders.

(3) Based on the previous step, 7*7*2 target Windows can be predicted, and the target Windows with lower probability can be removed according to the threshold. Finally, the NMS (network management system) can be used to remove the redundant Windows.

We can see that this process is very concise. It does not require direct regression after finding a target, so that location and category can be judged.

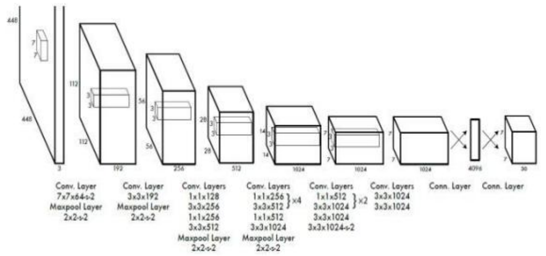

Fig.3 Structure of YOLO network

It is an important process to regression the position and category information of the target on the grid of different positions. Figure 2.3 is the network structure diagram of YOLO. After the convolutional layer, a fully connected layer of 4096 dimensions is connected, and then it is fully connected to a tensor of 7*7*30 dimensions.

In fact this 7*7 is the number of grids divided, which means that for each grid to predict a target, two possible positions are predicted along with the confidence and category of the target at this position, i.e. two targets are predicted per grid, each with information about the centroid and aspect, one being the confidence of the target, and the number of categories 20 (20 categories on the VOC), which is a total of 30 dimensional vectors. This method regresses the information needed for target detection on the grid (border information plus categories) based on the 4096 dimensional image features of the previous edge.

Disadvantages of YOLO: Without the Region proposal mechanism, only 7*7 grids can be used, and regression will be inaccurate, resulting in low detection accuracy of YOLO.

2.2.2SSD

SSD is a classical one-stage algorithm, which can effectively overcome the problems of poor detection effect and long detection time for small targets by Faster R-CNN. First of all, SSD is the same as YOLO in the way of obtaining the features and types of target areas, both of which adopt feature regression. However, YOLO expects specific locations to adopt the features of the whole map, while SSD expects specific locations to adopt the features around these locations. So how do you determine a desired feature and location? Use the anchor mechanism of Faster R-CNN. If the size of the feature map of a certain layer is 8*8, then a 3*3 sliding window is used to extract the features of each position, and then this feature regression can obtain the coordinate information and category information of the target [6].

SSD integrates the regression idea in YOLO with the anchor mechanism in Faster R-CNN. By regressed the multi-scale regional characteristics of different regions in the whole figure, it not only ensures the fast characteristics of YOLO, but also ensures the results of window estimation and Faster R-CNN Is generally more accurate.

III. IMPROVE OBJECT DETECTION METHODS

R-cnn series standard test architecture and YOLO Target test architecture provide us with two infrastructures to achieve target test. In addition, the research and development team based on this structure from other areas to provide a number of improved target detection techniques.

(1) Hard to distinguish sample mining: R-CNN used the idea of hard to distinguish sample mining when training SVM classifier, but the last two methods could not use this method because they used different strategies, which were end-to-end. Instead, they could only set the ratio of additive and subtractable samples and select them randomly. Abstract The Training Region-based Object Detectors with Online Hard Example Mining(Oral) of CVPR2016 embed Hard sample Mining in SGD Detectors. In the process of training, Fast R-CNN automatically selects appropriate region proposals according to the loss of region proposals as positive and negative examples for training to solve the problem of false positives. The experimental results show that the Fast R-CNN algorithm using OHEM (online Hard mining algorithm) mechanism can improve the mAP of VOC2007 and VOC2012 by about 4%.

(2) Multi-layer feature fusion: Both Fast R-CNN and Faster R-CNN achieve target measurement based on the characteristics of the last convolution computing layer, but due to the characteristics of the last convolution computing layer at the highest level, many details are lost, so the localization cannot be very accurate [7]. HyperNet and other methods use the multi-layer feature fusion of CNN for target detection, which has the following three advantages:

First, the features of deep layer, middle layer and shallow layer are fused together to complement each other's advantages. It not only uses the text data of the characteristics of the top layer, but also fully considers the detailed texture data of the characteristics of the lower layer, so as to make the target detection and positioning more accurate.

Second: Feature map resolution is 1/4, feature details are rich, conducive to the detection of small objects.

Third: the features are calculated before the region generation and subsequent prediction, without any redundant calculation.

(2) Use context information: When the obtained feature information of region proposal is used as target detection, the detection efficiency is usually better if the context message of region proposal is integrated.

IV. DESIGN AND IMPLEMENTATION

4.1 SYSTEM DESIGN

To complete the computer vision-based traffic light detection and recognition, Anaconda and Pycharm software, YOLOV5 project were used.

4.2 REALIZATION OF FUNCTIONS

Architecture diagramshow in Fig.5.

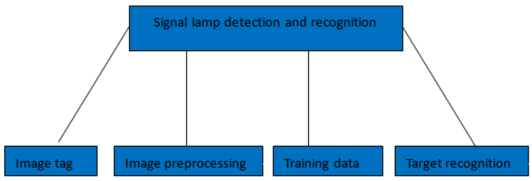

Fig.4Architecture diagram

4.2.1 YOLOV5 NETWORK STRUCTURE

The current phase of YOLOv5 target detection is the latest research result of YOLO series. It keeps the advantage of faster detection rate of the previous algorithm and makes the network construction more convenient [8][9].

According to the depth and width of the network structure, YOLO divides it into four different versions from small to large: YOLOv5s, YOLOv5m, YOLOv5x, YOLOv5l model. The structure of YOLO consists of input terminal, trunk network, neck, and output terminal. Its main functions are as follows:

(1) Input terminal

Use Mosaic data to enhance the image, in accordance with the provisions, the size of the width of the input, and then through the 4 pictures, optional compression, shear at random, or splicing, it randomly divided method of small target test effect is very good, greatly enriched the testing data set, random scale, increased the small goal, let the network robustness and better robustness.

(2) Trunk network

The convolutional neural network of image features is formed by aggregating different image small targets, in which Focus and CSP structures are used. The Focus structure is responsible for slicing pictures before they enter the main network. CSP structure is to divide the feature map of the base layer into two parts, and then merge them. It can not only reduce the amount of computation, but also improve the accuracy, and reduce the calculation bottleneck and memory cost.

(3) Neck

YOLOv5 now uses the FPN+PAN structure and the CSP2 structure of the backbone network to strengthen the ability of the fusion of network features.

(4) Output end

YOLOv5 uses Giou-loss as the bounding-box Loss function and uses weighted non-maximum rejection when carrying out non-maximum rejection, which has a good detection effect on some overlapped objects in the detection image without increasing computing resources.

4.2.2 Identification program flow chart

This experiment first needs to input the image, and need to judge the type of signal lamp, such as judging the vertical signal lamp or the horizontal signal lamp, then preprocessing the image, marking the picture, conducting the basic training and testing of the model, and finally identifying the signal lamp. The flow chart is shown in Fig.5.

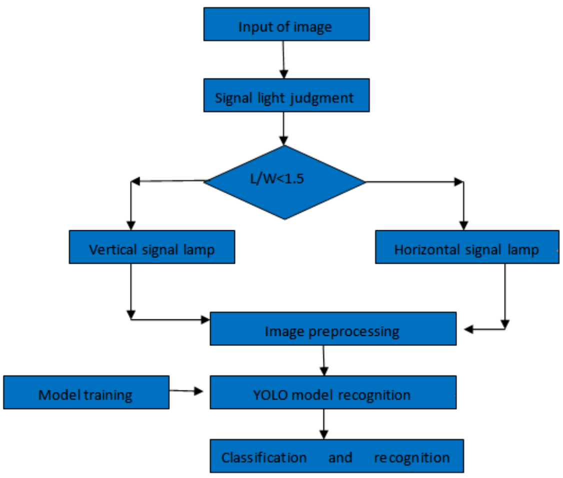

Fig. 5 Flowchart of identification program

V. Experimental results and analysis

Image red, green and yellow recognition, in the preprocessing, HSV color segmentation to classify categories. HSV three-color channels respectively represent hue H, saturation S and luminance V. The three components are independent of each other and meet the requirements of signal lamp detection and recognition.

The red light picture in the training set is imported, and the detection category is red light. The result of running the program is shown in Fig.6.

Fig. 6 Identification program red light detection

By importing the green light image from the training set, the category can be detected as green. The result of running the program is shown in Fig.7.

Fig. 7 Identification program green light detection

Select the yellow light picture from the training set and it can be recognized as yellow. The result of running the program is shown in Fig.8.

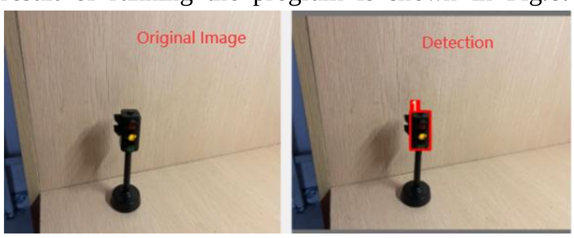

Fig. 8 Identification program yellow light detection

From the above pictures and results, it can be clearly seen that the color of the signal lamp can identify the color of the traffic light, which is in line with the expected effect and objective fact.

VI. Conclusion

Through the analysis of signal light detection and recognition, in the early stage of the experiment is not eager to complete the experiment, but to see the related literature, is to guide teachers to link requirements for the design, the implementation of the signal recognition, however, the selected images is simpler, images of a single, relatively less number, can only simple to identify the color of the light, don't know if I can in the complex traffic environment, to detect the traffic lights or do not know to realize the accuracy of the YOLOv5 algorithm is used to realize the recognition and detection of traffic lights. In the training data and model training, the effectiveness of the method can be obtained. In the next step, the recognition speed and accuracy will be improved, and the number of samples will be increased.

REFERENCES

[1] Umair Muhammad et al. Efficient Video-based Vehicle Queue Length Estimation using Computer Vision and Deep Learning for an Urban Traffic Scenario[J]. Processes, 2021, 9(10) : 1786-1786.

[2] Yamagata Koichi et al. Computer Vision System for Expressing Texture Using Sound-Symbolic Words [J]. Frontiers in Psychology, 2021, 12 : 654779-654779.

[3] Rodolfo Quispe, HelioPedrini Improved person re-identification based on saliency and semantic parsing with deep neural network models[J] Image and Vision Computing, 2019, 92(C)

[4] A. Kumar, H. Shwe, K.Juan, J. Chong Location-Based Routing Protocols for Wireless Sensor Networks: A Survey, *Scientific Research Publishing*, 2017.

[5] M. Dave, D Dalal, Simulation & Performance Evaluation of Routing Protocols in Wireless Sensor Network, *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 2, no. 3, 2013.

[6] Dong-Xu Liu, Feng Qin, Ying-Ming Wang, Xiu-Ru Li, Qian Yin, Dan-Dan Zuo, Yi-Ran Zheng, Design and Implementation of Graduate Student Integrated Information Management System — Degree Management, *Journal of Software* vol. 17, no. 3, pp. 121-129, 2022.

[7] Haiyue Fang, Xiaogang Wang, Zheyuan Cai et al. Learning semantic abstraction of shape via 3D region of interest[J] Graphical Models, 2019, 105(C)

[8] Ma YuPin Improved Interaction of BIM Models for Historic Buildings with a Game Engine Platform[J] *Applied Sciences*, 2022, 12(3)

[9] Fischer D.A., Goel K., Andrews R. et al. Towards interactive event log forensics: Detecting and quantifying timestamp imperfections[J] *Information Systems*, 2022, 109

[10] XieWuping, XueJinyun, Jiang Dongming et al. An iteration-based interactive analysis method to design dynamic service-oriented systems:AN ITERATION-BASED INTERACTIVE ANALYSIS AND DESIGN METHOD[J] *Software: Practice and Experience*, 2018, 48(2)

[11] Vasanthi Raghupathy, Osamah Ibrahim Khalaf, Carlos Andrés Tavera Romero et al. Interactive Middleware Services for Heterogeneous Systems[J] *COMPUTER SYSTEMS SCIENCE AND ENGINEERING*, 2022, 41(3)

[12] Willig James H, Krawitz Marc, Panjamapirom Anantachai et al. Closing the feedback loop: an interactive voice response system to provide follow-up and feedback in primary care settings.[J] *Journal of medical systems*, 2013, 37(2)

[13] Pengyi Zheng, Yuan Rao NetDDS: A Real-time Interactive Platform Based on the Publish-subscribe Mechanism[J] *Recent Advances in Electrical & Electronic Engineering*, 2022, 15(2)

[14] Hempel Thorsten, Al-Hamadi Ayoub An online semantic mapping system for extending and enhancing visual SLAM[J] Engineering Applications of Artificial Intelligence, 2022, 111

[15] Dong-Xu Liu, Feng Qin, Wan-Sheng Wang, Xiu-Ru Li, Lu-Lu Zhang, Lin Jiang, Ying-Ming Wang, Dan-Dan Zuo, Yi-Ran Zheng, Design and Implementation of Logistics Model 2019-NCOV Epidemic Development Simulation System in the Context of Big Data, *Journal of Software* vol. 17, no. 5, pp. 182-191, 2022.

[16] Huan Linxi, Zheng Xianwei, Gong Jianya GeoRec: Geometry-enhanced semantic 3D reconstruction of RGB-D indoor scenes[J] ISPRS Journal of Photogrammetry and Remote Sensing, 2022, 186

[17] Dong-Xu Liu, Research on In-Memory Database and Technology Selection，*International Review on Computers and Software(I.RE.CO.S.)*, Vol. 17, N. 2，2022

[18] XueDabin, Hsu Li-Ta, Wu Cheng-Lung et al. Cooperative surveillance systems and digital-technology enabler for a real-time standard terminal arrival schedule displacement[J] Advanced Engineering Informatics, 2021, 50

[19] Alabi B.N.T., Saeed T.U., Amekudzi-Kennedy A. et al. Evaluation criteria to support cleaner construction and repair of airport runways: A review of the state of practice and recommendations for future practice[J] Journal of Cleaner Production, 2021, 312

[20] Ioanna E. Manataki, Konstantinos G. Zografos A generic system dynamics based tool for airport terminal performance analysis[J] Transportation Research Part C, 2009, 17(4)

[21] White Greg Stochastic strength rating of flexible airport pavements using construction data[J] International Journal of Pavement Engineering, 2020, 21(4)

[22] da Rocha Phelipe Medeiros, Costa Helder Gomes, da Silva Glauco Barbosa Gaps, Trends and Challenges in Assessing Quality of Service at Airport Terminals: A Systematic Review and Bibliometric Analysis[J] Sustainability, 2022, 14(7)

[23] Xianliang Gu, JingchaoXie, Zhiwen Luo et al. Analysis to energy consumption characteristics and influencing factors of terminal building based on airport operating data[J] Sustainable Energy Technologies and Assessments, 2021, 44

[24] Park Kyudong, Han Sung H., Lee Hojin et al. Shared steering control: How strong and how prompt should the intervention be for a better driving experience?[J] International Journal of Industrial Ergonomics, 2021, 86

[25] Bhardwaj Akshay, Slavin Daniel, Walsh John et al. Estimation and decomposition of rack force for driving on uneven roads[J] Control Engineering Practice, 2021, 114

[26] CahitBartuYazıcı, Emir Kutluay, Yavuz SamimÜnlüsoy Steering optimization for multiaxle vehicles with multiaxle steering[J] Journal of Mechanical Science and Technology, 2021(prepublish)