Research on the Method of Traffic Signal Image Detection and Recognition

Jiaming Zhang*, Guanyu Chen, Yuanxin Guo
School of Information Engineering, Wuhan University of Technology, Wuhan, China

*Corresponding author: giraffe@whut.edu.cn

Abstract. Traffic lights play an important role in the distribution of conflicting traffic flow, and the correct detection and identification of traffic lights play an important role in reducing the occurrence of traffic accidents. Since the 1980s, China has begun the research and development of intelligent self-driving vehicles, which have higher requirements for all kinds of traffic sign recognition. Based on the above situation, this paper has done some research on traffic light image detection and recognition, which has far-reaching significance in the field of auxiliary driving and self-driving. The technical scheme adopted in this paper first uses the initial segmentation of RGB space and HSV color segmentation to get the selection of traffic lights, and then uses the lightweight network model to identify traffic lights and finally get the information of traffic lights. The experimental results show that the recognition accuracy and speed of this scheme can meet the requirements of the actual driving process.

1. Design background
In modern society, as a convenient means of transportation, cars have become an important part of people's daily life. With the continuous development of the city, the traffic system network is becoming more and more complex, and there are more and more traffic signs. As a kind of common traffic sign, the function of traffic signal lamp is to strengthen the road traffic management, reduce the occurrence of traffic accidents, improve the efficiency of road use, and improve the traffic condition to guide the vehicles and pedestrians to pass safely and orderly. Only by following the instructions of the traffic lights can the safety of the vehicle be ensured.

Figure 1. Traffic lights.
Therefore, we can consider using machine vision technology to identify traffic conditions. The research on visual perception technology of intelligent vehicles can be traced back to the sixties of last century, because the data computing power of computers at that time is still very limited, the research of this technology is only in its infancy. But in recent years, with the significant improvement of the performance of computer hardware, the ability of traffic light manufacturers to obtain images obtained through visual technology has also been significantly enhanced; in addition, with the popularity of visual sensors in a variety of electronic terminal equipment, the way to obtain video images has become unprecedentedly broad.

This paper hopes that through the use of image processing, on the basis of computer vision to identify traffic lights, combined with the current vehicle speed, in the auxiliary driving environment, provide safety information for drivers; provide driving information in driverless environment.

2. Technical route

For the detection and recognition of traffic lights, the specific technical route of the scheme proposed in this paper is shown in figure 2.

![Figure 2. Schematic diagram of technical route.](image)

2.1. The specific technical solutions adopted by each part are as follows:

2.1.1. Image acquisition: In order to ensure the reliability of obtaining images at night, this part intends to use ZED binocular camera. The ZED depth camera is a stereoscopic camera that provides high-definition images and accurately measures the depth of the environment. It has low level sensitivity, supports real-time depth-based visual ranging, and can capture 1080p HD video by 30FPS. Its firmware has rich API interface, which can provide enough support for the project. After testing, the camera meets the needs of car cameras.

2.1.2. Image enhancement: The image is often disturbed by many factors in the process of acquisition, transmission and storage, so it is necessary to process the image captured by the camera in advance in order to eliminate the influence of various noises and enhance the clarity of the image. This part intends to use the "smoothing" method of OpenCV to eliminate the noise in the image.

2.1.3. Image recognition: Traffic lights can be identified by setting up ROI, network model and training the network model. Using lightweight network model, it can be used in mobile and embedded end or even in ordinary PC with good accuracy. It can accurately and quickly identify the color of traffic lights and provide a guarantee for the stable implementation of the security warning system.

2.1.4. Safety warning system: When the image identifies the traffic light information of the intersection ahead, it can remind the driver in time combined with the current speed of the vehicle.

3. Technical realization plan

In this paper, the system design is analyzed from the point of view of the traffic light detection system and the traffic light recognition system. The system architecture is shown in figure 3.
3.1. Detection of traffic lights
The purpose of traffic light detection is to detect the candidate areas of traffic lights, which refers to the areas where traffic lights may exist. The system uses the size feature, shape feature, color feature and other feature information of the traffic signal to reduce the visual image area where the traffic light is located, which brings convenience for the subsequent traffic light type recognition.

3.1.1. Set up ROI (region of interest). In this application, the area where the traffic lights are located is the area of interest, and the other areas are not of interest. It can be seen from figure 4 that the traffic lights are in the upper part of the image. The setting of the region of interest can be set and trained according to the actual situation.

ROI is a kind of IVE technology, which can optimize video coding performance without loss of image quality and cut out required image parts for further processing and recognition. As shown in figure 5, the part that needs to be recognized, such as the traffic light, is cut out and then output to the subsequent recognition process, so as to avoid interference from other data in the picture and improve accuracy.
3.1.2. Morphology to background. In order to remove unwanted background factors such as buildings, trees, and the sky, a morphological background removal operation needs to be applied, which is a major preprocessing step in many computer vision-based tasks. Here, the KNN background elimination algorithm commonly used by OpenCV is used to model the background in the video to generate the mask image, and then the binary image analysis of the Mask is used to extract the area of the foreground active object, and finally the filtering operation is carried out to remove the background.

3.1.3. Candidate region extraction of traffic lights. RGB space is the most commonly used way to express color information. The brightness of red, green and blue primary colors is used to quantitatively represent colors, and the three colors of RGB are superimposed on each other to achieve color mixing. Different proportions of three colors will get different colors. The disadvantage is that the computation is too large and the hardware burden is heavy. Therefore, this design combines with HSV color space. HSV refers to the space where H (hue), S (saturation) and V (brightness) are used as color values to locate colors. Compared with RGB space, HSV space can be very intuitive to express the color of light and shade, hue, and bright degree, convenient for color contrast. It can segment the RGB processed image twice and improve the accuracy of color segmentation.

①Initial segmentation of RGB space
Because the traffic lights are divided into red, green and yellow, the components of different color channels are different when the traffic lights are in different states. So we can get the threshold value of three different color channels through many experiments. When the color component of one channel is greater than the color component of the other two channels to the threshold value of the channel, we initially think that the traffic signal light is the color, and then we carry out the subsequent operation.

②HSV color segmentation
Through RGB to HSV color space formula, the value of H, s, V in the image is obtained, and the HSV range of traffic lights with different colors is listed through many experiments. Then, the image preprocessed by RGB is segmented by HSV color space to segment the regions that need further recognition, so as to improve the accuracy and speed of subsequent neural network processing.

| Traffic light color | Hue       | Saturation | Value   |
|---------------------|-----------|------------|---------|
| Green light         | [105,185] | [60,255]   | [60,255]|
| Red light           | [340,365] | [80,255]   | [80,255]|
| Yellow light        | [5,35]    | [55,255]   | [120,255]|

3.2. Identification of traffic lights

3.2.1. Network model design. The network model designed by this system takes into account both the recognition accuracy and the disadvantages of the large model. By combining with the actual needs, the network model design is studied, and finally the lightweight network model is selected.

The design uses a lightweight network model of SqueezeNet, and SqueezeNet is composed of several Fire modules combined with the convolutional layer, down-sampling layer, full connection and other layers in the convolutional network. By replacing the traditional 3×3 convolution with 1×1 convolution, the number of parameters of a convolution operation is reduced by 9 times by this step. At the same time, the number of channels of 3×3 convolution is reduced: the calculation amount of a 3×3 convolution is 3×3×M×N (where M and N are respectively the number of channels of input Feature Map and output Feature Map), and M and N are reduced to reduce the number of parameters. In the end, down-sampling is placed after, and the larger Feature Map contains more information, so down-sampling is moved to the classification layer.
At the end of the network model, the global average pooling layer is used instead of the full connection layer, which greatly reduces the number of parameters.

3.2.2. Network model training. Through the network search and self-shooting to obtain three different state traffic signal image samples, and then classify the samples: set two subdirectories of squeezenet train and val to store training and test photos, set labels for different types of samples, and generate training and test data sets. Finally, the data set is used to train the neural network model, and the parameters such as learning rate and iteration times are constantly modified to achieve the best learning effect.

4. Conclusion and prospect analysis
This paper studies the detection and recognition method of traffic lights, makes full use of the color information of traffic lights combined with image processing algorithm to find the candidate area of traffic lights, and then uses convolution neural network to identify traffic lights. The detection and recognition of traffic lights is an important research direction in the field of self-driving, and it is one of the important technologies to ensure the safety of self-driving.

The detection and recognition of traffic lights provides a safety guarantee for auxiliary driving system and self-driving system, and is of great significance for the development of intelligent vehicles.

Acknowledgments
National innovation and entrepreneurship training program for college students+S202010497203

References
[1] Hinton G, Osindero S, Welling M, et al. Unsupervised Discovery of Nonlinear Structure Using Contrastive Backpropagation [J]. Cognitive Science, 2006, 30(4):725.
[2] Seide F, Li G, Chen X, et al. Feature engineering in Context-Dependent Deep Neural Networks for conversational speech transcription[C]. Automatic Speech Recognition and Understanding. IEEE, 2011:24-29.
[3] Jang C, Kim C, Kim D, et al. Multiple exposure images based traffic light recognition[C]. Intelligent Vehicles Symposium Proceedings. IEEE, 2014:1313-1318.
[4] Chiang C C, Ho M C, Liao H S, et al. Detecting and recognizing traffic lights by genetic approximate ellipse detection and spatial texture layouts [J]. International Journal of Innovative Computing Information & Control Ijicic, 2011, 7(12):6919-6934.
[5] Levinson J, Askeland J, Dolson J, et al. Traffic light mapping, localization, and state detection for autonomous vehicles[C]. IEEE International Conference on Robotics and Automation. IEEE, 2011:5784-5791.
[6] Fairfield N, Urmonson C. Traffic light mapping and detection[C]. IEEE International Conference on Robotics and Automation. IEEE, 2011:5421-5426.