Current Research on Convolutional Neural Network for Unmanned Driving

Xinyue Zhang

College of International, Hebei University, Tianjin, China
bjenson80807@student.napavalley.edu

Abstract. With the development of artificial intelligence and the upgrading of traditional cars, artificial neural networks have great potential in unmanned driving technology. Academics are paying increasingly close attention to computer vision research, and applications based on artificial neural networks are slowly but surely making their way into everyday life. In response to this situation, convolutional neural network models have good adaptability. Convolutional neural networks can be quite effective in identifying features and categorizing images. Technology for autonomous vehicles may recognize obstructions, road signs, and driver tiredness. Even when an image is translated, scaled, rotated, or thickened locally or globally, it may still produce the matching recognized information with high resilience and interference resistance.This article studies the following principal contents: extracting and processing color and shape segmentation regions, researching algorithms for identifying traffic signs, detecting multiple target criticalpoints based on Mask R-CNN, and comparing the performance of Mask R-CNN and Mask R-CNN combined with critical point detection.

Keywords: Artificial neural network; Unmanned driving; Identification; Scene recognition

1 Introduction

At present, the research of computer vision is attracting widespread attention from the academic community, and the application based on artificial neural networks is gradually entering people's lives. For example, computer vision is involved in new idea, unmanned driving, face recognition, intelligent image recognition, and 3D reconstruction. This paper explores the potential of artificial neural networks in driverless technology. The rapid advancement of artificial intelligence and the arrival of the 5G era have made it possible for people to have new demand for vehiclesars, stability, intelligence, comfort and other needs have promoted the upgrading of traditional cars, and the driving mode of classic cars has undergone tremendous changes, and driverless cars were born in this context [1, 2]. Driverless driving is to

B. H. Ahmad (ed.), *Proceedings of the 2023 International Conference on Data Science, Advanced Algorithm and Intelligent Computing (DAI 2023)*, Advances in Intelligent Systems Research 180,
https://doi.org/10.2991/978-94-6463-370-2_43

perceive the road conditions and the surrounding situation through sensors, and transmit them to the CPU, which judges the problem according to artificial intelligence and issues a series of instructions to the mechanical device,to control the vehicle to carry out reasonable actions. Intelligent Traffic System (ITS) is the integration of information technology, the incorporation of data communication technology, sensor technology, electronic control technology, artificial intelligence, and computer technology. It has the characteristics of timeliness, accuracy and efficiency [3, 4]. Nowadays, some computer vision applications, image recognition is not perfect. There is no direct and simple way to widely determine various unexpected situations: identify any object in any environment, such as damaged and blurry traffic signs, road obstacles. Existing technologies can and can only solve the recognition of specific things, for instance, straightforward geometric recognition, face recognition, document recognition for handwriting, or vehicle recognition. And these recognitions may require particular lighting, background, and target pose requirements in particular environments.

Segmenting the color range of traffic sign images is the result of collecting and counting the colors of a large number of traffic sign maps [1]. Through the RGB color range statistics of a large number of traffic sign images, the color range of traffic signs is obtained: red: 139<=r<=255; 0<=g<=106; 0<=b<=203; yellow: 139<=r<=255; 90<=g<=255; 0<=b<=90; blue: 0<=r<=98; 0<=g<=206; 111<=b<=255. In this range, three types of traffic signs are divided into prohibitions, warnings, and instructions [1]. Color segmentation of a real-shot traffic sign image is shown in Figure 1.

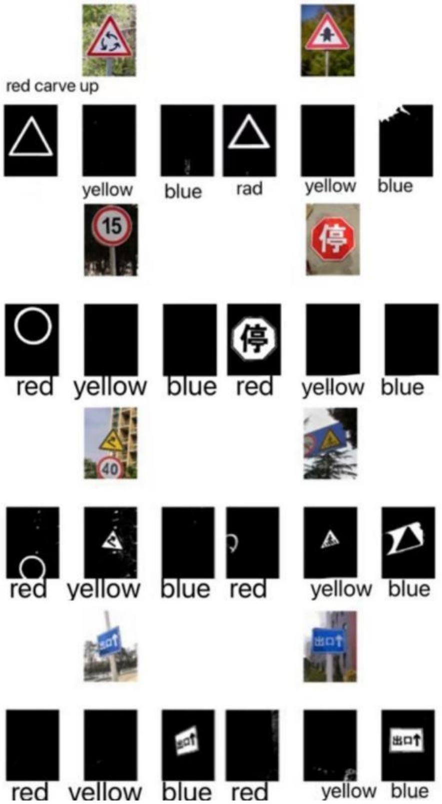

Fig. 1. Color segmentation of a real-shot traffic sign image [5]

2 Algorithms for Identifying Traffic Signs

Recognition Algorithm: Image enhancement can be used to more accurately identify a series of traffic signs such as prohibition signs, indication signs, and warning signs. Due to the limitations of image acquisition angle, poor visual effect, noise pollution and other conditions, it is sometimes difficult to obtain better image recognition effect, this study can highlight or remove some interference information according to specific needs, and can use image enhancement processing to solve the problem. First, you can use a grayscale transformation.

The spatial domain processing expression is: y(j, i)=T[x(j, i)], where T is the grayscale transformation function. Comparison chart before and after grayscale processing is shown in Figure 2.

Its function type can be referred to:

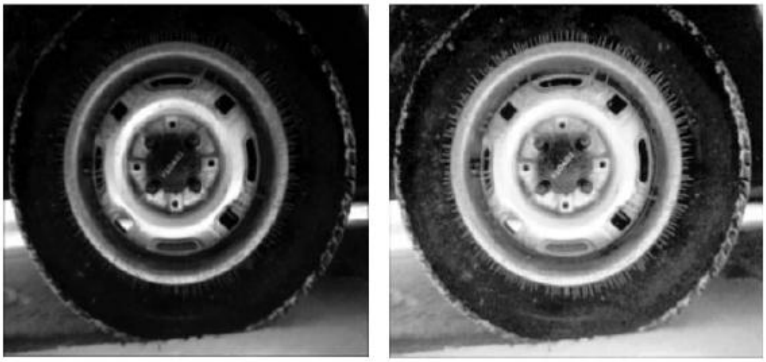

Fig. 2. Comparison chart before and after grayscale processing [6]

In this study, the image can be filtered and extracted to the instrumental part of the image.The specific area can be divided by subtraction, multiplication and other algorithms, and a particular part of the image can be obtained for the identification of different traffic signs.

3 Methods

3.1 Advantages of Convolutional Neural Networks

Convolutional Neural Networks (CNNs) belong to the category of artificial neural networks.They are a typical deep learning network architecture that was influenced by the organic visual cognitive systems. Because its network model is often composed of multiple layers, it is also known as Deep Convolutional Neural Network (DCNN), which is one of the typical models in the field of deep learning [7]. Convolutional, pooling, and fully connected layers comprise its structure; many convolutional and pooling layers are typically taken. The input data of the convolution will be locally associated with the feature map output by each convolutional layer.At the same time the local input and the corresponding connection weight matrix will be weighted and sometimes added with bias values to obtain the output value, which is equivalent to the convolution process in the field of signal processing, so it is called a convolutional neural network [8]. Deep learning algorithms represented by convolutional neural networks can automatically and accurately acquire the features of images. A typical convolutional neural network model is shown in Figure 3.

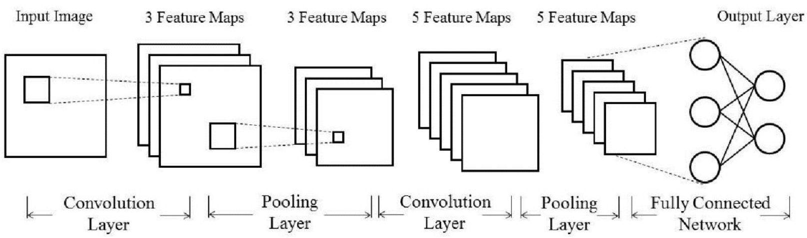

Fig. 3. A typical convolutional neural network model [9]

Convolutional neural network training process: Generally there are two steps to train CNNs. The first step is the process of continuous feature extraction of incoming image data from the shallow network to the deep web, which is generally called bold propagation. The next step is to discover the error by comparing the forward propagation results with the label(also called loss), and then reverse the mistake from the deep to the shallow network for bias derivation calculation, and then update the training step of the weight parameters, which is generally called backpropagation [10]. The training process is shown in Figure 4, and the CNN training process is shown in Figure 5.

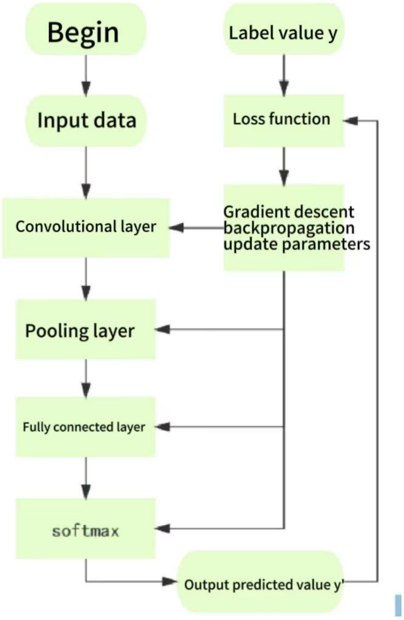

Fig. 4. The training process of the model [11]

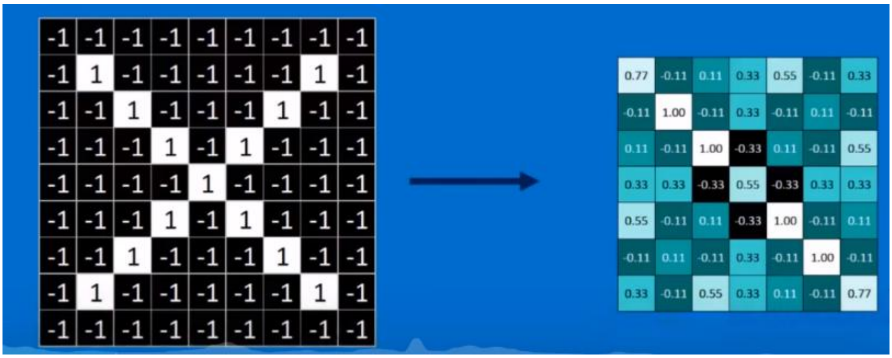

Fig. 5. CNN training process [12]

The convolution kernel is scanned on the original map, and the corresponding content is extracted and reflected on the feature image. Repeat the process of convolution, erasure, and pooling, and obtain the result after complete linking. Gradient descent is carried out, a large amount of data is given, and the error changes are observed after adjusting up and down for each feature pixel and voting weight, and the loss function is calculated to the minimum. Handwritten input number 3 uses convolutional neural networks to output training results is shown in Figure 6

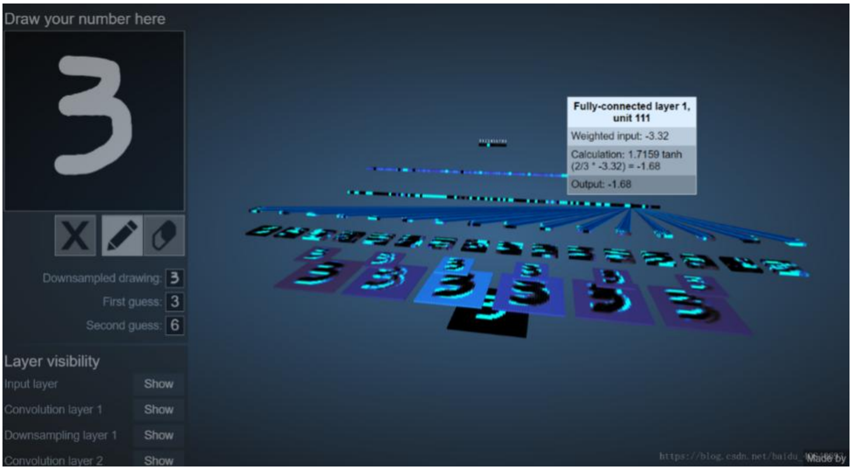

Fig. 6. Handwritten input number 3 [13]

CNNs suit tasks with complex input data and rich information, such as image recognition.

3.2 Multiple Target Key point Detection based on Mask R-CNN

As the first company to develop driverless technology, Google has developed fully autonomous vehicles that can automatically start and stop [14]. In autonomous driving, the detection of critical points plays a crucial role, in the complex situations in front of the vehicle, such as pedestrians, road signs, traffic lights, arrows, obstacles, etc. To achieve good detection accuracy, can accurately and quickly distinguish them, and then make driving action decisions.

Through these critical nodes, it instantly, quickly, and accurately transmitted to the CPU of the driverless car. When you want to park the car in the parking line, as shown in the figure below, first detect the parking space road marking. Here you can use the key point selection of the road marking, and then notice, in the image, you need to find different types of critical points of the parking place, so that the car automatically drives to the parking space. Parking space critical point detection is shown in Figure 7.

Fig. 7. Parking space critical point detection [15]

Mask R-CNN is a very adaptable framework that enables the addition of numerous branches to carry out diverse tasks. Target categorization, object detection, semantic segmentation, and instance segmentation are a few instances, and recognition of human posture. Use Mask R-CNN to model the position of each key point, such as the car's tire, tail, body, front of the vehicle, etc. Each place has multiple key point types, and there is a mask, so use Mask R-CNN to predict multiple covers [14]. Figure 8 shows object detection and instance segmentation performed by Mask R-CNN.

Fig. 8. Object detection and instance segmentation performed by Mask R-CNN [14]

The mAP (mean Average Precision) measures the accuracy of object detection. For object detection, many boxes will be output, and IoU will be used to evaluate the quality of packages. IoU=area of overlap/area of union. Precision = Number of positive samples detected correctly/Total test samples Recall = Number of correct positive samples seen/Total positive models.

The specific process of calculating AP is: 1) Put the prediction boxes for all of category C that were produced by the algorithm first, ranked by confidence. 2) Set different k values. Calculate FP and TP after choosing the top k prediction boxes, so that recall is equal to 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0 respectively; calculate Precision; 3) Average the 11 Precisions, that is, get APs; AP is for a single category, mAP is to sum all categories of APs, and then take the average: mAP = sum of all categories of APs / total number of categories. The property comparison of Mask R-CNN and Mask R-CNN combined with critical point detection is shown in the following table was established in Table 1.

Table 1. Comparison of Mask R-CNN and Mask R-CNN combined with critical point detection

Algorithm	Mask R-CNN	MaskR-CNN+keypoints	Classification
mAP(%)	81.46	83.32	Car

The advantages and disadvantages of convolutional neural networks in unmanned driving are discussed. Unmanned one are affected by some unique environments, such as extreme cold, complex road conditions and other extreme environments. Unmanned cars on the road ensure that following road safety signs vehicle identification and additional information, can essentially solve the highway driving process due to fatigue and other factors caused by the danger [12]. The use of convolutional neural network can play a significant role in finding features, classing pictures, in unmanned technology, can detect road signs, obstacles, or identify whether the driver is tired even if the image has been translated, zoomed, rotated, local or overall thickening and other deformations can also output the corresponding recognition content, its robustness, anti-interference ability is powerful. Letters can be recognized through CNN is shown in Figure 9.

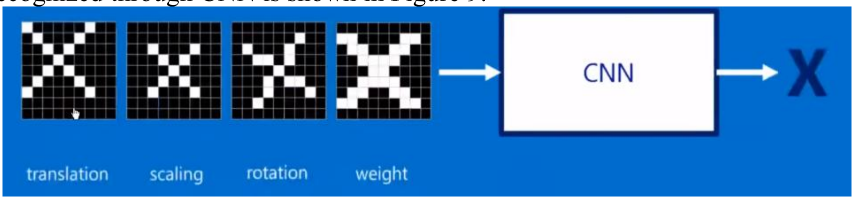

Fig. 9. Letters can be recognized through CNN [15]

4 Discussion and Analysis

This article mainly focuses on the extraction and processing of color and shape segmentation regions. By analyzing the RGB color range of many traffic sign images, the color range of traffic signs can be determined, and symptoms such as bans, warnings, and instructions can be determined based on it. The research on traffic sign recognition algorithms can highlight or remove certain interference information according to specific needs and solve the problem by using image enhancement and other processing methods. Further processing of the image through grayscale transformation can achieve ideal results. Introduced the content structure and training process of convolutional neural networks. Multiple target key point detection based on Mask R-CNN requires good detection accuracy for complex situations in front of vehicles, such as roads, and the ability to distinguish them accurately and quickly before making driving decisions. And the performance comparison of Mask R-CNN/Mask R-CNN combined with key point detection. Combined with Mask R-CNN, annotation classification training is used to model and train various key point types of vehicles such as tires, body, front and rear. It is found that detecting multiple key points of the vehicle can better detect targets, such as the car being impeded by a large number of people, but by seeing a small number of key points, a multi critical point selection prediction method is adopted, it can identify occluded targets and some small targets that are far away. The discovery of the advantages of convolutional neural networks in autonomous driving technology is highly worth exploring and utilizing.

By studying the detection of critical points on the road, the selection of multiple critical points of the vehicle class combined with the object detection of Mask R-CNN and the direct use of Mask R-CNN for instance segmentation object detection, the mAP is improved by 1.86%.More tiny and obscured targets can be identified through the recognition effect. It can better ensure the safety and recognition accuracy of cars in autonomous driving. It is proved that detecting critical points to the category can improve the recognition accuracy of Mask R-CNN. Make autonomous driving technology more scientific, standardized, and standardized. This provides research value and direction for the detection of key road points in future unmanned driving technology. The development of deep learning technology has successfully encouraged the fusion of data across many sizes and tasks, allowing the integration of data to progress from two-dimensional to three-dimensional approaches. It has exceptional practical importance for creating models for extracting road vital points.

5 Conclusion

In this paper, through the research on convolutional neural network object detection, convolutional neural network structure and operation are explained in detail. The image enhancement method of image segmentation method is also studied, and the critical point detection method is also studied, and the methods and characteristics of critical point selection are understood. It has a theoretical basis for detecting uncrewed vehicles on the road ahead of them.

Through research, this paper finds that driverless technology can be improved by studying convolutional neural network models and using computer vision to process images and improve the recognition principle and effect of existing vehicle detection algorithms for unmanned car road information detection technology, to play a role in this field more quickly and bring more convenience to people's lives. In the future, it is a trend to realize driverless driving through deep learning, and it is believed that deep learning algorithms can make driverless cars run stably and safely on the world's roads.

References

1. Han Xinhui. Unmanned driving scene recognition based on deep learning [D]. Sun Yat sen University, 2018.
2. Gong Qiang. Research on road front object detection of driverless car based on Mask R-CNN[D].Nanchang University, 2021.
3. Li Jie. Research on key point detection algorithm based on deep convolution neural network [D]. University of Chinese Academy of Sciences (Institute of Optoelectronics Technology, Chinese Academy of Sciences), 2023.
4. Qiao Weigao, Xu Xuejin. Development status and direction of driverless vehicles[J]. SAIC, 2007, 07: 40-43.
5. Han J,Yu-Qiu X U,Guan X,et al. Driverless Car Based on Logic Gate Intelligent Control[J].
   Computer&Telecommunication,2019.
6. Zhu Z, Lu J, Martin R R, et al.An Optimization Approach for Localization Refinement of Candidate Traffic Signs[J]. IEEE Transactions on Intelligent Transportation Systems, 2017: 1-11.
7. Wang Zefa, Tang Xingguo. Research on image enhancement method based on grayscale transformation[J].Science and Technology Innovation Herald, 2011, 01: 119.
8. Zhang Yinan. Research on traffic sign detection algorithm based on artificial neural network[D].Northwest Normal University, 2021.
9. Lee H S,Kang K.Simultaneous Traffic Sign Detection and Boundary Estimation Using
10. Convolutional Neural Network [J]. IEEE Transactions on Intelligent Transportation Systems, 2018, 99:1-12.
11. Zhou Feiyan, Jin Linpeng, Dong Jun. Review of Convolutional Neural Network Research[J]. Chinese Journal of Computers, 2017, 40(6): 1229-1251.
12. Yu Yijie. Research on visual vehicle detection algorithm based on improved Faster R-CNN[D].Jiangsu University, 2021.
13. Gong Qiang. Research on road front object detection of driverless car based on Mask R-CNN[D].Nanchang University, 2021.