Scholar Publishing Group
International Journal of Neural Network
https://doi.org/10.38007/NN.2021.020203
ISSN 2790-2692 Vol. 2, Issue 2: 17-24

Intelligent Vehicle Lane Recognition Based on Neural Network

Wenchen Liu and Guozhen Xu*

Zaozhuang University, Zaozhuang, China

**corresponding author*

Keywords: Neural Network, Intelligent Vehicle, Lane Recognition, Fully Convolutional Neural Network

Abstract: The input of external road information is the root of decision-making and control of automatic or auxiliary driving system, and ensuring the quality of input information is the premise of normal operation of the system. This paper mainly studies the application of intelligent vehicle lane recognition based on neural network. Based on the open source traffic scene dataset and the relevant theoretical basis of convolutional neural network, this paper carried out research on lane detection methods, constructed a variety of traffic scene detection models, optimized the model structure, and finally performed well on the experimental test set. Aiming at the problem of lane detection, this paper develops lane detection based on full convolutional network. By reclassifying and labeling the unclear lane lines in the sample images, a data set of lane line detection was constructed. The full convolutional network model is built and the network structure is adjusted to improve the accuracy of detection.

1. Introduction

Lane is an important part of the traffic sign system, which contains a large amount of intuitive and useful information. It uses specific colors, shapes and words to convey the directions of this section of traffic road, guide the driving direction of vehicles, and remind intelligent cars or drivers to make correct response and operation to the traffic state on the road [1]. With the further development and improvement of visual equipment, visual processing and other technologies, lane detection and recognition is not only an important function of automatic driving system, but also a basic technology to support the construction and construction of ADAS system functions, such as lane keeping function and lane departure warning function [2]. Therefore, lane detection and identification research is an essential and important demand for the development of social economy and science and technology. However, nowadays, lane detection and recognition is basically based on machine learning or deep learning, which requires a large amount of computation. It not only requires high computer configuration, but also causes interference from various external factors such as weather and light, such as lane line and background, when the contrast is low. As a result, lane recognition requires a large number of operations to eliminate and reduce interference factors in the preprocessing stage such as lane detection and positioning, which often takes a large amount of time and the algorithm is complex [3, 4].

In recent years, major foreign brand vehicle companies are racing to launch intelligent concept cars with intelligent driving, and frequently conduct road tests. Joint research and development institutions have developed models for Mercedes Vamors -p intelligent vehicles, the smart car carrying a visual navigation system based on double focus and the focal length of different camera division of labor cooperation, tracking their respective goals, short using a high-resolution camera to realize to detect obstacles, wide Angle camera used to detect the road wide range information [5]. Honda has installed a camcord-based lane detection and recognition system for its vehicles, as well as a vehicle ranging and holding system. When the distance between the vehicle and the vehicle in front is too close, the automatic driving system will realize automatic deceleration function [6]. Waymo, an autonomous driving company under the search giant Alphabet, lets a blind friend sit in a self-developed autonomous car and conduct road tests in the actual lane. The smart car is equipped with a traffic sign recognition system based on camera [7]. Google is also committed to the research and development of vision processor (VPU) and Advanced driving assistance system (ADAS) and lane recognition is one of the main functions of the processor and driving assistance system [8].

The traditional lane line detection algorithm has poor robustness and low accuracy for lane line detection in special scenes, and sometimes it even cannot be detected. Therefore, in this context, it is of great significance to study lane line detection based on convolutional neural network.

2. Lane Recognition Based on Convolutional Neural Network

2.1. Application of Convolutional Neural Network in Computer Vision

(1) Image recognition

Image recognition and classification is a basic computer vision task, mainly through various graphic processing and feature description of the input image, so as to determine its category. Since images are easily affected by interference factors such as illumination, shadow and deformation, traditional methods for image understanding and recognition based on manual features cannot effectively classify large-scale image data [9, 10]. As a kind of data-driven algorithm, convolutional neural network (CONVOLUtional neural network) was first applied to image recognition task, and showed strong feature learning ability that traditional methods could not match [11].

The classical CNN model LENET-5, which is used to realize character recognition and classification, mainly includes a feature extraction layer composed of multiple convolution-activation-pooling layers stacked together and a fully connected layer responsible for integrating information and output classification probability [12]. In addition to the LENet-5 model, network frameworks such as AlexNet, VGGNet, GoogleNet and ResNet also show excellent performance in the direction of image recognition and are widely used as the foundation networks for other deep CNN models [13]. In object detection and image segmentation, classification task refers to the classification and recognition of the objects of interest or each pixel in the image.

(2) Object detection

Object detection mainly realizes the function of identifying the object of interest from the background in the input image, and accurately predicting its position and category, including object classification and border regression. Target detection models based on CNN can be roughly divided into two categories: two-stage detector and single-stage detector [14].

The two-stage detector is a kind of target detection algorithm based on candidate regions. Firstly, the candidate regions are generated by the algorithm, and then the candidate regions are finely located and classified. R-cnn algorithm is the opening work of the two-stage method. Firstly, the SS(Selective Search) algorithm is used to extract 2000 candidate regions from the input image, and then the convolutional neural network is used to extract the image features of candidate boxes. Finally, the border classification and regression are carried out by SVM classifier and regressor. Get the final test result. CNN in R-CNN is mainly responsible for extracting candidate region features. Compared with traditional feature descriptors, CNN has more abstract representation ability and can improve the detection accuracy of the target, but it has the disadvantages of high training difficulty, large storage space occupation and high time cost [15]. Based on R-CNN, Fast R-CNN and Faster-RCNN have been proposed successively, which greatly improve the performance of object detection algorithms. Fast R-CNN mainly introduces the RoI pooling structure layer into the convolutional neural network. It only needs one forward convolution process to obtain the feature information of all candidate boxes, and then predicts the category and position of the target through the fully connected layer, which effectively improves the detection speed of the overall algorithm. Faster R-CNN directly uses convolutional neural network to realize regional proposal and target classification detection [16].

Single-stage detector is a kind of target detection algorithm based on feature regression, represented by YOLO and SSD series frameworks. The core idea is to use convolutional neural network to achieve end-to-end target classification and bounding box regression without generating candidate regions. The YOLO framework evenly divides the input image into several grids. If the center of the sample box falls to a grid, the grid is responsible for predicting the category and position of the target [17]. On the basis of YOLO, YOLO9000 and YOLOv3 introduce optimization strategies such as multi-scale prediction, Anchor box and clustering, which further greatly improve the detection accuracy of the network and ensure high real-time performance. The target detection network uses convolutional layer instead of fully connected layer to integrate features and output prediction results, so that the network can process images with different input sizes [18]. Since each grid in the output feature graph is responsible for predictive classification and bounding box regression, 1×1 convolution can be used to obtain the depth features of the corresponding grid position for computing task output.

2.2. Lane Line Detection Based on Full Convolutional Network

Compared with the classical convolutional neural network model, the biggest feature of fully convolutional network is that it does not set the fully connected layer, but replaces the fully connected layer with the convolution kernel, with the size of 1×1. This has the advantage of reducing the network's requirement for image input size. In addition, the final output result of the full convolutional network is slightly different from that of the convolutional neural network, which is specifically shown as follows: After feature extraction, the full convolutional network will up-sample it by deconvolution method to make the final output image consistent with the dimension of the input image, which is conducive to improving the accuracy of target segmentation.

Full convolution network model will be divided into two parts, the first part is the convolution pooling layer, the forward propagation, this section with the classical convolution neural network model is similar, but at the end of the convolution layer, original all connection layer has been replaced by convolution, the convolution kernel size is 1 x 1, convolution kernel number and the original connection layer all have the same number of; The main function of the second part is to up-sample the results extracted from the network, which aims to adjust the final output image to the same scale as the input image. Upsampling is usually done by deconvolution. Image deconvolution is also called transpose convolution, which is the inverse of the convolution operation.

In order to better extract the deep features of lane lines, this paper designed five pooling layers in the convolution pooling layer in the first part. After five layers of pooling operation, the input sample will get a feature map with the original size of one thirty-second. At this time, we need to perform upsampling operation. However, it is found through experiments that if the up-sampling operation of 32 times is directly applied to the obtained feature map, the feature map reduced to the input sample size will be too rough, and effective lane features cannot be extracted. Based on this situation, we consider a multi-layer structure of deconvolution fusion mode in the lane detection network model.

The main purpose of the deconvolution fusion mode of multi-layer structure is to obtain more features of the convolution layer from the upsampling results. Even so, the sampling operation is not level, the more the better, this is because, there are some lower level of convolution, to extract the characteristics of them are not perfect, also do not have the appropriate semantic characteristics, it is likely the result of the experiment error of classification, if only to increase the layer and its blindly into the sampling characteristics on the diagram, Will have an impact on the final experimental results.

Based on the above ideas, we first up-sampled the results of the seventh layer by a factor of 4 when constructing the lane detection model, and then up-sampled the output feature map of the pooled fourth layer by a factor of 2. The results of these two parts and the pooling results of the third layer are fused to form a total fused feature map. Finally, the image is upsampled by a factor of 8 to restore the size of the input image.

For the upsampling results obtained, a classifier is considered to measure the effect. We still use the most common Softmax loss function classifier. Assuming that the size of the input image is A × B, it can be considered that A × B Softmax classifiers are connected to classify each current pixel. For the detection of lane lines, the Softmax layer outputs A × B × 2 values, takes the maximum probability of each pixel as the prediction result, marks the pixels predicted as lane lines to form areas, and finally obtains the detection result of lane lines.

In this paper, the Mean Square Error (MSE) method is used as the loss function to verify the good and bad training of the model, which is the loss value between the predicted pixel value of the lane line area and the sample annotation value in the sample image. The calculation formula is:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}[y_{pred}^{(i)} - y_{true}^{(i)}] \tag{1}$$

Where, N is the total number of samples, ypred is the predicted value given by the model in a training batch, and yture is the corresponding annotation value, that is, the true value.

3. Evaluation of Detection Model Experimentss

3.1. Model Operation Environment

The detection hardware environment of the lane recognition algorithm designed in this paper is mainly composed of CPU (Central controller, Central Processing Unit) GPU (Graphics Processing Unit) RAM (random memory, random access memory). Random Access Memory (Random Access Memory) HDD (Hard Disk Drive) and other key devices. The specific parameters of these devices are shown in Table 1.

In order to better develop the software of lane recognition system on Windows platform, it is necessary to use integrated development environment to generate applications that can run on the platform. In this paper, TensorFlow and OpenCV are used to develop the lane recognition algorithm. These two software architectures adopt Python and C/C++ programming languages respectively.

Table 1. Hardware environment of lane recognition model

Device name	Equipment model
CPU	Intel Core i5-9400
GPU	NVIDIA RTX 2070
RAM	16GB DDR4
HDD	ST1000LM035

3.2. Evaluation Criteria

The recognition effect of lane recognition model is an important index to evaluate the performance. In this paper, 300 original road images are randomly selected for the experiment. Firstly, the lane recognition system is used to complete the task of lane recognition for these images. Then the number of correctly recognized and incorrectly recognized lane line coordinate points and the number of unrecognized lane line coordinate points are counted in the projection image of lane line coordinate points. Finally, the average F1 value is calculated by the F1-measure method to evaluate the recognition effect of the system. The formula of F1 value is as follows:

$$F_1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{2}$$

4. Comparative Analysis of Experimental Results

4.1. The Average of Coordinate Points of the Lane Line

Table 2. Average value of coordinate points of lane recognition algorithm

Image label	Correct identification of points (TP)	False recognition points (FP)	Number of unidentified points (FN)
001-100	16.238	1.172	0.327
101-200	16.053	1.854	0.957
201-300	19.495	0.831	0.704
Total average	17.262	1.285	0.662

Table 2 counts the coordinate points of the lane line recognized by the lane line recognition system for 300 original road images.

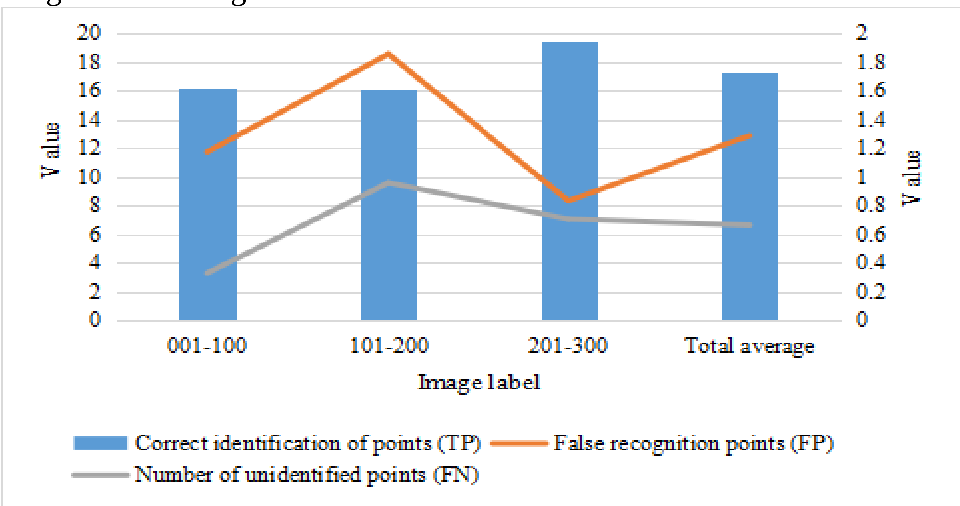

Figure 1. Diagram of the average value of coordinate points of lane line

As shown in figure 1, the use of the lane line recognition algorithm, which can identify all the image of each group respectively and then the statistics in the group image system to properly identify the coordinates of the points average average (TP) error identification, the coordinates of the points average (mean) of FP and did not identify the coordinates of the points average average (FN) Finally, the total average of each statistic was calculated. According to the experimental data, for 300 randomly selected original road images, the lane line recognition system designed in this paper can correctly identify the average number of lane line coordinate points is about 17.262, the average number of lane line coordinate points is about 1.285, and the average number of lane line coordinate points is about 0.662.

4.2. F1 - Measure Values

Table 3. F1-measure evaluates the data

Image label	Precision	Recall	F1
001-100	92.042%	96.321%	94.862%
101-200	88.093%	93.164%	90.697%
201-300	94.435%	93.178%	95.723%
Total average	91.523%	94.221%	93.761%

Table 3 is the F1-measure evaluation data table of the lane recognition system. The data table calculates the average Precision, Recall and F1 obtained by the lane recognition system designed in this paper for each image grouping, and then calculates the total average of each statistical data.

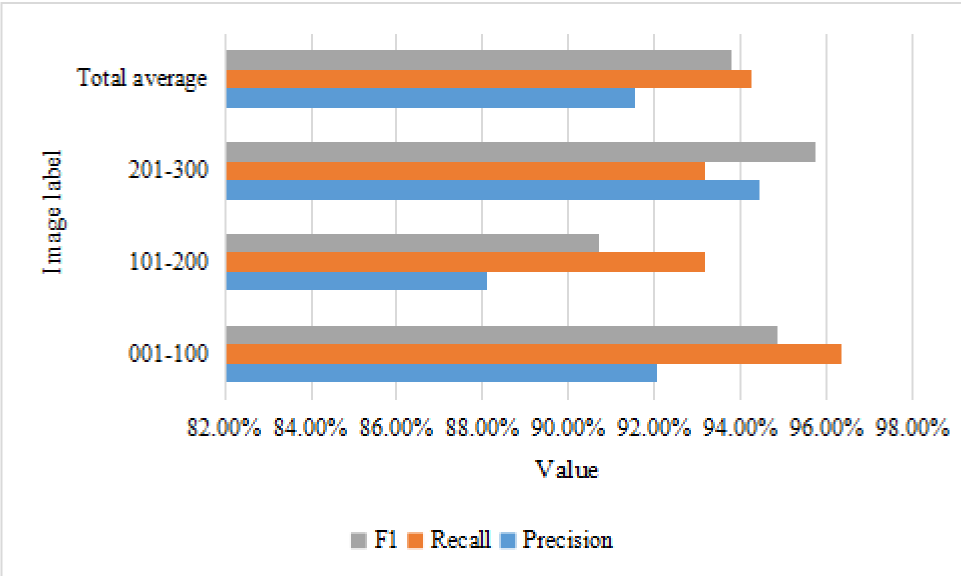

Figure 2. Precision, recall and f1 average obtained by f1-measure method

As shown in Figure 2, F1 is always between Precision and Recall, which comprehensively measures the full ability and accurate recognition ability of lane line recognition system for lane line coordinate points. Through the evaluation results of F1-measure method, the average Precision, Recall and F1 of the lane recognition system designed in this paper is 91.523%, 94.221% and 93.761%, respectively. It shows that the comprehensive performance of the system is in the advanced level among the similar systems.

5. Conclusion

In recent years, with the rapid development of intelligent vehicles and deep learning, intelligent vehicles have become the inevitable trend of the future development of automobiles. Lane detection technology is an important part of the environment perception of the driving scene of intelligent driving vehicles. Based on the research of convolutional neural network in the field of image processing, this paper builds a lane detection model based on convolutional neural network model. Limited by the research time, the lane recognition technology based on neural network designed in this paper still needs to be further studied: neural network still relies on a large amount of data as the input of offline training, and enough data will often train the network with better performance. However, the current deep neural networks still have the problems of unsatisfactory generalization performance and low training efficiency. Therefore, more intelligent modeling methods or training strategies are needed to solve these problems in the future.

Funding

This article is not supported by any foundation.

Data Availability

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Conflict of Interest

The author states that this article has no conflict of interest.

References

[1] Padmajadevi G, Aprameya K S. Offline Signature Verification & Recognition Using Angle Based Feature Extraction & Neural Network Classifier. International Journal of Electronics Engineering Research. (2017) 9(5):711-725.

[2] Noubigh Z, Mezghani A, Kherallah M. Densely Connected Layer to Improve Vggnet-Based CRNN for Arabic Handwriting Text Line Recognition. International Journal of Hybrid Intelligent Systems. (2021) (3):1-15. https://doi.org/10.3233/HIS-210009

[3] Tlemsani R, Benbakreti S, Benyettou A. On-Line Arabic Characters Recognition using Enhanced Time Delay Neural Networks. International Journal on Communications Antenna and Propagation. (2017) Vol 7(4):254-264. https://doi.org/10.15866/irecap.v7i4.13204

[4] Bonilla Cardona D A, Nedjah N, Mourelle L M. Online Phoneme Recognition using Multi-layer Perceptron Networks Combined with Recurrent Non-linear Autoregressive Neural Networks with Exogenous Inputs. Neurocomputing. (2017) 265(nov.22):78-90. https://doi.org/10.1016/j.neucom.2016.09.140

[5] Al-Jubouri M, Abusaimeh H. Offline Arabic Handwritten Isolated Character Recognition System Using Support Vector Machine And Neural Network. Journal of Theoretical and Applied Information Technology. (2017) 95(10):2315-2322.

[6] Mohammed D A, Mezher A, Hadi H S. Off-Line Handwritten Character Recognition Using An Integrated DBSCAN-ANN Scheme. Indonesian Journal of Electrical Engineering and Computer Science. (2019) 14(3):1443-1451. https://doi.org/10.11591/ijeecs.v14.i3.pp1443-1451

[7] Matija R, Offei A, Wang Q. Object Recognition in Aerial Images Using Convolutional Neural Networks. Journal of Imaging. (2017) 3(2):21. https://doi.org/10.3390/jimaging3020021
[8] Alkhateeb J H, Turani A A, Abuhamdah A, et al. An Effective Deep Learning Approach for Improving Off-Line Arabic Handwritten Character Recognition. International Journal of Computer Systems & Software Engineering. (2021) 6(2):104-112.
[9] Sharma A, Jayagopi D B. towards Efficient Unconstrained Handwriting Recognition Using Dilated Temporal Convolution Network. Expert Systems with Applications. (2021) 164(2):114004. https://doi.org/10.1016/j.eswa.2020.114004
[10] Fannas L Y, Sasi A B. Off-Line-Signature-Recognition-Based-On-Angle-Features-and-GRNN-Neural-Networks. International Journal of Parallel Programming. (2019) 7(10, 2013):229-233.
[11] Rashid N, Alim S A, Hashim N, et al. Receiver Operating Characteristics Measure for the Recognition of Stuttering Dysfluencies using Line Spectral Frequencies. IIUM Engineering Journal. (2017) 18(1):193-200. https://doi.org/10.31436/iiumej.v18i1.578
[12] Togootogtokh E, Shih T K, Kumara W, et al. 3D Finger Tracking and Recognition Image Processing for Real-Time Music Playing with Depth Sensors. Multimedia Tools and Applications. (2017) 77(8):1-16. https://doi.org/10.1007/s11042-017-4784-9
[13] Lima J, JVMD Oliveira, Oliveira R. Distributed Artificial Intelligence for Recognition and Position Validation of Printed Circuit Board. European Academic Research. (2019) 7(9):4673.
[14] Aly W, Alya S, Almotairi S. User-Independent American Sign Language Alphabet Recognition Based on Depth Image and PCANet Features. IEEE Access. (2019) (99):1-1. https://doi.org/10.1109/ACCESS.2019.2938829
[15] Khashman A, Oyedotun O K, Olaniyi E O, et al. In-Line Grading System for Mango Fruits Using GLCM Feature Extraction and Soft-Computing Techniques. International Journal of Applied Pattern Recognition. (2019) 6(1):58. https://doi.org/10.1504/IJAPR.2019.10026052
[16] Antwi-Bekoe E, Zhan Q, Xie X, et al. Insulator Recognition and Fault Detection Using Deep Learning Approach. Journal of Physics: Conference Series. (2020) 1454(1):012011 (9pp). https://doi.org/10.1088/1742-6596/1454/1/012011
[17] Tuan L N, Nguyen K C, Nguyen C T, et al. Recognition of Anomalously Deformed Kana Sequences in Japanese Historical Documents. IEICE Transactions on Information and Systems. (2019) E102. D (8):1554-1564. https://doi.org/10.1587/transinf.2018EDP7361
[18] Shinohara N, Hashimoto M. Automatic Determination of Reference Pixels Using Middle Layer of Deep Neural Network for Template Matching. Journal of the Japan Society for Precision Engineering. (2018) 84(12):1079-1084. https://doi.org/10.2493/jjspe.84.1079