Traffic marking recognition based on generating antagonistic neural network

Shuwei Xu¹, Shan Zhang² and Shuwei Xu¹*

¹Highway college, Chang’an University, Xi’an, Shaanxi, 710064, China
²Highway college, Chang’an University, Xi’an, Shaanxi, 710064, China

Abstract. This paper presents a method of extracting traffic lines from image images by GAN. Compared with the traditional image detection methods, the counter neural network does not need repeated sampling of Markov chain and adopts the method of backward propagation. Therefore, when detecting the image, GAN do not need to be updated with samples; it can produce better quality samples, express more clearly. Experimental results show that the method has strong generalization ability, fast recognition speed and high accuracy.

1. Introduction

Road markings are an integral part of the urban transportation system and indicate the driving specifications of the vehicle. The automatic identification of road markings plays an important role in both assisted driving and unmanned driving. Accurate identification of road markings is one of the prerequisites for achieving lane-level navigation. In recent years, the progress of deep learning has been in full swing, and deep learning has been verified to improve the accuracy of target recognition and semantic segmentation. For the traffic scene of this study, using the deep learning method based on the confrontation generation network, the longitudinal deceleration zone in the lane, the lateral deceleration zone, the mesh marking line, the zebra crossing, the diversion zone, the arrow, the horizontal solid line, the vertical solid line, Dotted lines, other symbols (including text, diamond symbols, triangle symbols) are extracted from the 10 types of markings.

1.1. Confrontational confrontation generation network

The confrontation generation network was originally proposed by IanJ. Goodfellow. Its main idea is two competitive neural network models. One takes noise as an input and generates a sample (so called a generator); another model (called a discriminator) receives samples from the generator and training data and must be able to distinguish between the two sources. These two networks are continuous games, the generator is learning to produce more and more realistic samples, and the discriminator is learning how to better distinguish between generated data and real data. These two networks are trained at the same time, and it is hoped that this competition will make the generated samples indistinguishable from the real data. It used in image modeling.

\[
\min_{G} \max_{D} \mathbb{E}_{x \sim P_{data}}[\log D(x)] + \mathbb{E}_{z \sim P_z}[\log(1 - D(G(z)))]
\]

In the above formula, \(x\) represents a real image, \(z\) represents an image that needs to be input to the G network, and \(G(z)\) represents a "false" image generated by the G network. Therefore, from the above formula, we can see that the best result we want to use GAN is:

a. The image generated by the G network has higher and higher realism, that is, the value of \(D(G(z))\) is as large as possible, and \(D(G(z))\) is increasing (approaching to 1), \(\log(1 - D(G(z)))\) will get smaller and smaller, at this point we can get \(\min_G\); b. The ability of D network to authenticate true and false is getting stronger and stronger. Similarly, the value of \(D(x)\) is as large as possible, and the value of \(\log D(x)\) is also increasing. At this time, we can get \(\max_D\).

1.2. Confrontational confrontation generation network

This experiment is based on the image translation of the conditional against the network proposed by the paper Phillip et al. The loss function is defined as:

\[
f_{GAN}(G, D) = E_{y \sim P_D}[\log D(y)] + E_{x, z}[\log(1 - D(G(x, z)))]
\]

Compared with the traditional confrontation generation network, the conditional-based confrontation generation network increases the input layer variables, because the image translation work not only cares about the generator to generate the real image, but also cares about the real image generated by the generator and the input. The images match. Therefore, the loss function is defined as:

*Corresponding author's e-mail: 756354243@qq.com

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).
\[ f_{GAN}(G, D) = E_{x,y} [\log D(x, y)] + E_{x,z} [\log (1 - D(x, G(x, z)))] \]

Considering that the image generated by the generator also has similar information to the input layer, L1 loss is added to the loss function:
\[ f_{L1}(G) = E_{x,y,z} ||y - G(x, z)||_1 \]

The total loss function is the addition of two parts:
\[ G^* = \arg \min_G \max_D L_{GAN}(G, D) + \lambda f_{L1}(G) \]

2. Experiment procedure

2.1. Data acquisition

At present, there are two main sources of road marking data, one is two-dimensional image data, and the other is three-dimensional laser point cloud data collected by on-board laser scanning equipment. The method for extracting the road markings according to these two sources also has two different ways according to the image and the three-point cloud. The method of extracting road markings on the basis of images mainly focuses on the geometric features and illumination features of road markings, such as edge detection, Hough transform, etc. However, this method is susceptible to illumination, shooting angle and distance, etc. When the image quality is poor, the reticle is blurred, or the road width changes when the scene is changed, the recognition rate of the algorithm is reduced. There are fewer methods for identifying lane markings using vehicle laser radar.

The data of this experiment is derived from the 3D laser point cloud scanned by the on-board laser scanning device, and the intensity feature image is generated by the point cloud. The reticle extraction is performed on the basis of the intensity feature image. The training data has a total of 642 intensity feature images, and each of the intensity feature images corresponds to a real road surface length of 200 meters. 3D laser point cloud image:

2.2. Production of data samples

For the traffic scene of this study, using the deep learning method based on the confrontation generation network, extracting the road markings according to these two sources also has two different ways according to the image and the three-point cloud. The method of extracting road markings on the basis of images mainly focuses on the geometric features and illumination features of road markings, such as edge detection, Hough transform, etc. However, this method is susceptible to illumination, shooting angle and distance, etc. When the image quality is poor, the reticle is blurred, or the road width changes when the scene is changed, the recognition rate of the algorithm is reduced. There are fewer methods for identifying lane markings using vehicle laser radar.

The data of this experiment is derived from the 3D laser point cloud scanned by the on-board laser scanning device, and the intensity feature image is generated by the point cloud. The reticle extraction is performed on the basis of the intensity feature image. The training data has a total of 642 intensity feature images, and each of the intensity feature images corresponds to a real road surface length of 200 meters. 3D laser point cloud image:

the longitudinal deceleration zone in the lane, the lateral deceleration zone, the mesh marking line, the zebra crossing, the diversion zone, the arrow, the horizontal solid line, the vertical solid line, Dotted lines, other symbols (including text, diamond symbols, triangle symbols) are marked with 10 types of markings.
2.3. Data preprocessing and data enhancement

Data preprocessing: Since each image of the intensity feature image generated from the three-dimensional point cloud conversion is different, this experiment performs framing processing on each whole intensity feature image, and the image size after segmentation is 1024*1024. An overlapping area of 200 pixels is reserved between adjacent images, and the actual length of the corresponding road surface is 4 meters, and the number of images after the width is 8008.

Data enhancement: Because a small amount of data is not easy to converge during training, it is easy to produce over-fitting, so the experiment carried out data enhancement: each image was rotated at different angles, increasing the number of training images.

2.4. Parameter settings and experimental results

The loadSize is set to 1076, the finesize is set to 1024, and the number of iterations is 100. This experiment compares the choice of the generator and whether the test results of adding L1Loss in the generation network and the loss function are selected.

|                | Longitudinal speed bump | Entrance and exit mesh line | sidewalk guide belt | various arrows | lateral speed belt | solid line dotted line | symbol | All       |
|----------------|-------------------------|----------------------------|---------------------|----------------|-------------------|-----------------------|--------|-----------|
| Training accuracy | Accuracy rate           | 0.9648                     | 0.8152              | 0.9369         | 0.9525            | 0.9736                | 0.9536 | 0.9341    | 0.8758    | 0.5337    | 0.9024    |
| Recall rate      |                         | 0.8324                     | 0.3119              | 0.8264         | 0.9057            | 0.8267                | 0.8934 | 0.9083    | 0.8467    | 0.8224    | 0.8825    |

Table 1 Use training accuracy against generated networks /%.

|                | Longitudinal speed bump | Entrance and exit mesh line | sidewalk guide belt | various arrows | lateral speed belt | solid line dotted line | symbol | All       |
|----------------|-------------------------|----------------------------|---------------------|----------------|-------------------|-----------------------|--------|-----------|
| Test accuracy  | Accuracy rate           | 0.9756                     | 0.9624              | 0.8857         | 0.9424            | 0.8783                | 0.9376 | 0.9042    | 0.8542    | 0.9083    | 0.7496    | 0.8714    |
| Recall rate    |                         | 0.7937                     | 0.2668              | 0.7973         | 0.8676            | 0.7924                | 0.8686 | 0.9437    | 0.8924    | 0.7977    | 0.8668    | 0.8644    |

Table 2 Use test accuracy against generated networks /%.

The experimental results against the generated network will be made into a more intuitive line chart, as shown below.

![Use training accuracy against generated networks](https://doi.org/10.1051/e3sconf/201913604076)
Test Results:

(a) Solid and dashed lines
(b) Various arrows
(c) Longitudinal speed bump
(d) Sidewalk
(e) Mesh line

3. Conclusion

In this experiment, the deep learning method against the generation network is used to extract the features of the reticle: the longitudinal deceleration zone, the lateral deceleration zone, the mesh reticle, the zebra crossing, the diversion zone, the arrow, the horizontal solid line, the vertical solid line, the dotted line, Other symbols (including text, diamond symbols, triangle symbols) were
extracted from the 10 types of markings. Good classification results have been obtained. The experimental results show that the method used in this experiment to generate network extraction reticle is better than the traditional method and can be applied to actual production. Points of style

References

1. Yu, M., Yu, B., (2000) Detection of indicator marks based on color images. J. Computer Engineering and Applications.,165-168.
2. Zhang, B., (2000) Research on traffic sign recognition method based on fuzzy neural network. J. Journal of Highway and Transportation Technology and Technology.
3. Song, W.,Fu, M.,Yang, Y., (2015) An efficient traffic sign recognition technology for driverless cars. J. ROBOT., 102-111.
4. Li, L.,Ma, G. , (2009) Research on traffic sign detection and classification algorithm in natural scenes. J. Journal of Harbin Institute of Technology., 29-33.
5. CYGANNEK, B., (2007) Soft system for road signs detection. J. Advances in Soft Computing., 316-326 CYGANNEK, B., (2007) Soft system for road signs detection. J. Advances in Soft Computing., 316-326.
6. Zhang, Z, (2000) A flexible new technique for camera calibration. J. Pattern Analysis and Machine Intelligence., 133-142.
7. Ministry of Transport (2015) Code for setting up urban road traffic signs and markings. Ministry of Transport, BeiJing.
8. Liang, W., Cai, L., Guo, W. (2010) New method for traffic sign recognition. In: The 15th National Conference on Image and Graphics. GuangZhou.