IOP Conf. Series: Materials Science and Engineering **719** (2020) 012076 doi:10.1088/1757-899X/719/1/012076

Road Segmentation based on Deep Learning with Post-Processing Probability Layer

Weibin Chen

Network center, Jimei University, 183 Yinjiang Road, Jimei district, Xiamen, China.
Email: wbchen@jmu.edu.cn

Abstract. In order to solve the problem of road image semantics segmentation, a novel semantic segmentation method based on deep learning is proposed in this paper, and it is verified that this method can optimize the segmentation effectively comaring with traditional Full Convolution Neural Network (FCN) model. Firstly, the traditional Full Convolution Neural Network model is constructed. And then according to the principle of the post-processing probability layer method proposed in this paper, the label of all road image training sets is used to compute and transform it to form a two-dimensional array which can represent the classification probability of each pixel, and it is combined with the Full Convolution Neural Network model to be used for road image semantics segmentation. Secondly, the tensorflow neural network framework is used to simulate the above two models. Finally, the experimental results show that the CNN model with the proposed post-processing probabilistic layer is able to get better results in road semantics segmentation in KITTI data sets. The pixel accuracy is improved from 88.8% to 91.3%, and the mean pixel accuracy is increased from 82.9% to 85.7%. The mean intersection over union increased from 72.5% to 77.9%.

1. Introduction

Autonomous driving technology is a hot topic in these years. Among them, accurate semantic segmentation is one of the key technologies in the process of automatic driving. For example, fuzzy c-means algorithm based on objective function is widely used[1]. In recent years, with the progress of computer hardware, the expansion of information and data, and the rapid improvement of Deep Learning technology, Convolutional Neural Networks [2-4] are proposed and applied in image classification and semantic segmentation. In order to make convolutional network better applied to image semantic segmentation, many researchers have improved and optimized it and proposed a variety of neural networks for semantic segmentation. In 2014, a fully convolutional networks (Jonathan Long et al.) [5] was proposed. By replacing the final fully connected layer with the convolution layer, the network uses deconvolution to carry out up-sampling, so that the size of the output image can be restored to the same size as the input image, thus achieving end-to-end image semantic segmentation at the pixel level. In 2015, researchers proposed SegNet neural network (Vijay Badrinarayanan et al.) [6]. In which, the symmetric structure of encoder and decoder is used, and the maximum pooling index is applied in the decoder, so that the segmentation resolution can be improved.

In order to solve the problem of image information loss caused by pooling, DeepLab neural network (liang-chieh Chen et al.) [7] was proposed. In which, a pyramid-shaped dilated pool structure is proposed by using dilated convolution and combining it with fully connected condition random field.

However, dilated convolution structure has the disadvantages of high computational cost and large footprint, so RefineNet neural network (Guosheng Lin et al.) [8] was proposed. The method solves the disadvantages of dilated convolution structure. The segmentation precision of the network is high, but the network is complex. At present, the research focus in the field of road image semantic segmentation has gradually transformed intotreatment after joining the deep learning technology, and on this basis, new research achievement constantly results, even if the application and the promotion of these results of this study is still limited to some aspects, but with the continuous development and improvement of deep learning technology, pushes the road image semantic segmentation technology into a new era.

The networks mentioned above are proposed to solve the problem of image semantic segmentation, but these networks have a wide range of applications and are not designed specifically for a certain kind of segmentation problem. Therefore, in order to make the neural network model is more suitable for solving the problem of road detection. This paper proposes a post-processing probability layer of the new model, and the model is used for road-image semantic segmentation. Comparatived with traditional fully convolution neural network algorithms, the proposed model has a better ability for road-image semantic segmentation.

2. The Design of Post-Processing Probability Layer

The roads in data set which is used to train the model are almost in the middle and lower of the image. Therefore, this paper proposes an algorithm of post-processing probability layer after CNN based on this characteristic of road image.

The method of post-processing probability layer is shown as follows.

First, transform RGB image to binary image in training data sets. The value in the result image is 1 or 0, as shown in Figure 1. A value of 1 means that the pixel is classified as a road, while a value of 0 means that the pixel is classified as a non-road (background).

Then, the all labels array sum averaging (normalized), get a the size of h*w matrix, the matrix is called a probability layer (as shown in Figure 2, 3), in which h, was the input label images high and width, the range of each value is 0 to 1 (including 0 and 1), the closer the value 1 indicates that the pixels is the more likely for road, the closer the value 0 indicates the pixel points is more likely for background.

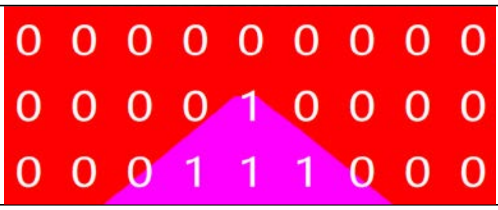


Figure 1. Schematic of the converted data label.

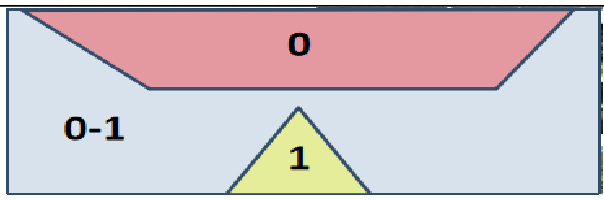


Figure 2. Probability layer concept map.

In order to enable the probability layer to act on the final semantic segmentation results, it is necessary to transform the probability layer. Subtract 0.5 from each value of the original probability layer to get the converted probability layer, as shown in Figure 4.

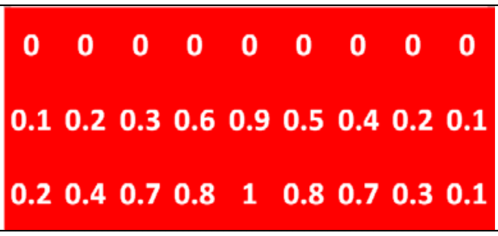


Figure 3. Original probability layer diagram.

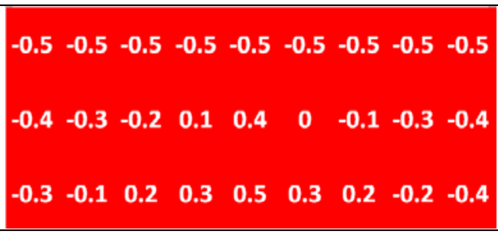


Figure 4. Probability layer schematic diagram after conversion.

IOP Conf. Series: Materials Science and Engineering **719** (2020) 012076 doi:10.1088/1757-899X/719/1/012076

Finally, when the model predicts the test set image, add the probability layer in the last layer of the model and assign an impact factor to this layer as the post-processing method of the prediction result, so as to correct some misclassified pixels.

After the road image has been trained by the neural network model, a two-dimensional array label matrix will be generated, as shown in Figure 5, with the value of 0 to 1 for each pixel. Therefore, the value of each pixel is compared with 0.5. A value greater than 0.5 indicates that the pixel is classified as a road, while a value less than 0.5 indicates that the pixel is classified as a background. As can be seen from the Figure 5, two pixel values were misclassified, and the background was predicted to be the road, and the road was predicted to be the background.

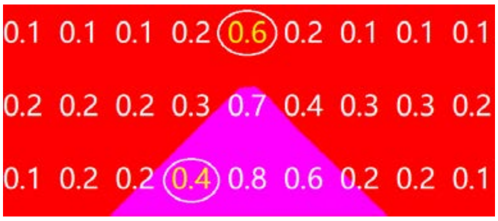


Figure 5. Prediction error label matrix

In this case, the post-processing probability layer can influence the results. By adding the predicted value of each pixel in the prediction result to the corresponding pixel value of the probability layer, as shown in Figure 6, the above prediction errors can be eliminated, thus improving the segmentation accuracy of road image semantic segmentation to a certain extent. As can be seen from the figure, two originally misclassified pixels were finally classified correctly after being processed by probability layer, and the classification results of other pixels were not changed.

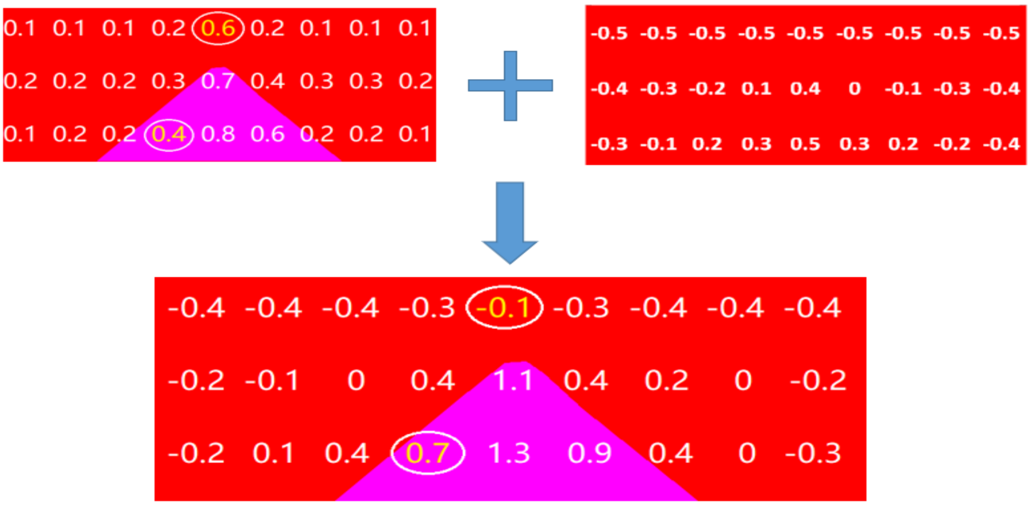


Figure 6. Combining predictional and probability layer.

Because of the particularity of road image, the road is always at a specific position in the road image. Therefore, it is possible to attach a special value to each pixel value in the image to improve the semantic segmentation result of the road image. This is the main idea of post-processing probability layer.

3. Experimental Verification and Analysis

3.1. Experimental Data

All experiments in this paper are based on road images in the data set KITTI[10] for model training. Data set KITTI was created in 2012 by Toyota's Chicago institute of technology and Karlsruhe institute of technology in Germany. It is one of the most famous data set in the field of autonomous driving. The KITTI data set mainly includes images of daily life scenes such as highways, rural areas and cities, which are mainly captured by a color camera mounted on a car. In these dataset, the road is marked purple (RGB value [255, 0,255]) and the background is marked red (RGB value [255, 0, 0]). This paper uses 289 images, and their sizes were cropped to 576*160, among which 214 images were used as the training model and 75 images were used as the result verification.

3.2. The Experimental Process

(1) Pretreatment: at the beginning of the experiment, it is necessary to preprocess the trained images and labels. First, cut the image and label; Then the label image is transformed into a vector array encoded by one-hot code to represent the classification of corresponding pixels. Finally, the training picture and the corresponding label were divided into batch of batch_size for the input of the training model.

(2)Build FCN network model

(3)Training model: the output of the neural network was normalized by softmax function, and the processed results and label results were averaged through cross entropy function to form the cost function of model training. The Adam Optimizer is then used to optimize the model in the direction of the cost function drop and to calculate the training accuracy for each iteration. In the process of model training, the learning rate is set as 0.00001.

(4)Generate probability layer: according to the probability layer principle, generate a two-dimensional array that can represent the classification probability of each pixel with all the training pictures.

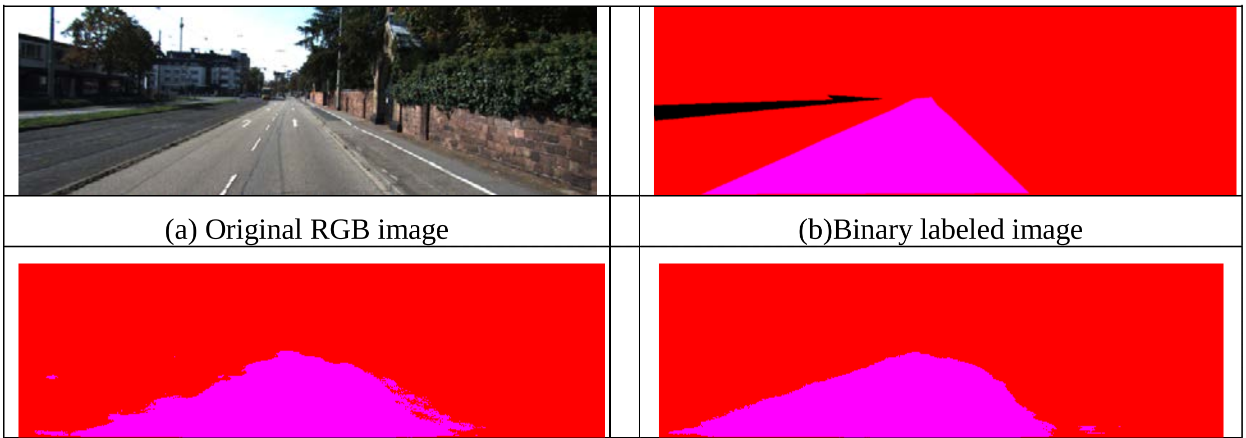
(5)Model prediction: after completion of the training model, the model parameters are saved, and then input the test pictures to already trained model, and use probability layer makes these generated prediction labels, then compared prediction labels with real labels, and find out the pixel accuracy (PA), average pixel accuracy (MPA) and the average intersection over union ratio (MIoU).

3.3. Compare Experimental Results and Analysis

In order to verify the effect of this algorithm in road image segmentation, the following work is mainly done on the data set:

(1)Fully convolutional neural network (FCN) model is used to segment road image semantically.

(2)A fully convolutional neural network model based on probability layer is used for road semantic segmentation.



(a) Original RGB image (b)Binary labeled image

(c)FCN predicting result (d) Combining result

Figure 7. The results of road segmentation.

It can be seen from the above figures that the FCN model algorithm based on probability layer proposed in this paper has a better segmentation effect on road images than the traditional FCN model algorithm. After adding probability layer, FCN model can make the segmented road outline smoother and make the road boundary more coherent and complete. At the same time, after adding the probability layer, some unreasonable discrete points can be removed to make the prediction region as connected as possible, which is also consistent with the real road scenario.

Then, the prediction results of these two models were quantitatively analyzed and their pixel accuracy (PA), average pixel accuracy (MPA) and average intersection over union ratio (MIoU) were calculated respectively. The results are shown in table 1:

The results shown in Table 1. Compared with the traditional FCN model algorithm, the proposed algorithm in this paper has improved the segmentation accuracy. Among them, pixel accuracy (PA) improved by 2.5 percentage points to over 90%.The average pixel accuracy (MPA) increased by 2.8 percentage points to over 85%.The average intersection over union ratio (MIoU) increased by 5.4 percentage points.

Table 1. Formatting sections, subsections and subsubsections.

Metric	FCN	FCN+Probability Layer
PA	88.8%	91.3%
MPA	82.9%	85.7%
MIoU	72.5%	77.9%

According to the above results, the application of probability layer can improve the semantic segmentation effect of road image. The probability layer records the special location of the road in the road image, and makes use of the characteristic of the road image to carry on the later intervention to each pixel value in the image, so as to improve the semantic segmentation result of the road image.

4. Conclusion

According to the latest research results of FCN model in the field of road image semantic segmentation algorithm in the world, considering the particularity of road image, this paper carried out the research of road image semantic segmentation method based on the idea of transformation and calculation. The main work is as follows:

(1) In view of the traditional algorithm FCN model segmentation defect with low accuracy, according to the particularity of road image: road always in a certain location in the road image, this paper proposes a new idea, each pixel of image attach a particular value, indicates the probability of classification. That can improve the road image semantic segmentation results, and improve each evaluation index through the model of semantic segmentation;

(2) The tensor flow open source neural network framework was used to train and learn the neural network framework in this paper based on the KITTI data set. Parameters and training model were determined via multiple experiments.

(3) This paper applies the method to the real road image recognition experiment, and verifies the effect and optimization performance of post-processing probability layer model by analyzing the results.

Through comparison and analysis with experimental results of various deep learning algorithms, it is shown that compared with traditional fully convolutional neural network algorithm, the algorithm proposed in this paper can better segment road images, with higher accuracy and better reliability, and can be used for semantic segmentation of road images. The final experimental results show that the post-processing probability layer can improve the pixel accuracy from 88.8% to 91.3%, the average pixel accuracy from 82.9% to 85.7%, and the average intersection ratio from 72.5% to 77.9%.

Because of the existing open source road image data set stock rarely, this algorithm could not carried out in a large number of road image segmentation experiments, the subsequent will continue to carry out the research work on establishing road image sample libraries and label data set, further verify the algorithm effect that is applied to a large number of different scenarios of road image segmentation. At the same time, other mathematical changes can be attempted in the steps of generating the probability layer, such as nonlinear changes such as tanh function and sigmoid function. The difference between linear change and nonlinear change and their advantages and disadvantages will be explored to make them suitable for road images in various situations.

5. References

[1] Bezdek, James C. Pattern Recognition with Fuzzy Objective Function Algorithms. Advanced Applications in Pattern Recognition, 1981, 22(1171):203-239.
[2] Shi J, Malik J. Normalized cuts and image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(8):888-905.
[3] Rother C, Kolmogorov V, Blake A. Grabcut: Interactive foreground extraction using iterated graph cuts. ACM, 2004, 23(3):309-314.
[4] Krizhevsky A, Sutskever I, Hinton G. ImageNet Classification with Deep Convolutional Neural Networks. Advances in neural information processing systems, 2012, 25(2).
[5] Long J, Shelhamer E, Darrell T. Fully Convolutional Networks for Semantic Segmentation. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 39(4):640-651.
[6] Badrinarayanan V, Kendall A, Cipolla R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE transactions on pattern analysis and machine intelligence, 2017, 39(12):2481-2495.
[7] Chen L C, Papandreou G, Kokkinos I, et al. Semantic image segmentation with deep convolutional nets and fully connected crfs. arXiv preprint arXiv: 1412.7062, 2014.
[8] Lin G, Milan A, Shen C, et al. Refinenet: Multi-path refinement networks for high-resolution semantic segmentation. Proceedings of the IEEE conference on computer vision and pattern recognition. 2017:1925-1934.
[9] Garcia-Garcia A, Orts-Escolano S, Oprea S, et al. A review on deep learning techniques applied to semantic segmentation. arXiv preprint arXiv: 1704. 06857, 2017.
[10] Geiger A, Lenz P , Stiller C , et al. Vision meets robotics: The KITTI dataset. The International Journal of Robotics Research, 2013, 32(11):1231-1237.
[11] Panhai Zheng, Ling Guo, Libing Ding. Research and realization of convolutional neural network based on Tensor Flow. Electronic technology and software engineering, 2018(18):20-22.
[12] Lihua Wu. Google. Application of Tensor flow framework in automatic vehicle recognition function. Automobile practical technology, 2018(20):38-40.
[13] Feng Xiao, Ting Rui, Tongwei Ren, Dong Wang. Fully convolution semantic segmentation and object detection network. Chinese journal of image and graphics, 2019, 24(03):474-482.
[14] Yanfang Xing, Hongxiu Duan, Guangwei He. Application of Tensor Flow in image recognition system[J/OL].Computer technology and development, 2019(05):1-5
[15] Zongsheng Wu, Weiping Fu, Gaining Han. Road scene understanding based on deep convolutional neural network. Computer engineering and applications, 2017, 53(22):8-15.
[16] Juting Dai, Xinyi Tang, Peng Liu, Baotai Shao. Scene semantic segmentation network based on color - depth image and deep learning. Science, technology and engineering, 2018, 18(20):286-291.