

# The Researches Detection Method of Illegal Parking Based on Convolutional Neural Network

Xue-Hong JIANG<sup>1</sup>, Hui-Li FENG<sup>2</sup> and Shi-Yue LI<sup>3,a,\*</sup>

<sup>1</sup> Information Center of Department of Housing & Urban-Rural Development, Shijiazhuang 050011, China

<sup>2</sup> Hebei Building Technology Co., Ltd., Shijiazhuang 050011, China

<sup>3</sup> College of Artificial Intelligence and Data Science, Hebei University of Technology, Tianjin 300400, China

<sup>a</sup>lishiyue11@qq.com

\*Corresponding author

**Keywords:** convolutional neural networks, city management, image processing, machine vision

**Abstract.** Violation of parking is a legal term. Our country's law stipulates that there should be no-stop signs and markings, road sections with isolation facilities between motor vehicles and non-motor vehicle lanes and sidewalks, and crosswalks and construction sites. No parking; Railway crossings, sharp bends, narrow roads with widths less than 4 meters, bridges, steep slopes, tunnels, and sections within 50 meters of the above locations are not allowed to stop. In recent years, urban management is gradually developing towards information There are corresponding treatment mechanisms for illegal parking in urban management, but the method of detecting illegal parking based on machine vision is still under study. This paper takes a street in Hebei Province as the research object, and studies the illegal parking under the surveillance image. It also proposes to use the machine vision-based method to automatically detect the illegal parking. In the monitoring image, according to the street view image information detected by the computer, A status assessment of the vehicle placement can be achieved. The detection algorithm in this paper uses the multi-angle suggestion area to accurately locate the offending vehicles in the image and mark them on the detected vehicles. The experimental data shows that the algorithm has good adaptability to the detection of illegal parking in surveillance images, and effectively improves the inspection efficiency.

## 1. Introduction

It is common to drive illegal parking in a city. Unauthorized parking of vehicles will affect the image of the city. If the vehicles are parked indiscriminately, the order of the city will be chaotic, and the taste and image of the city will be affected. This will make the traffic environment of the city roads unable to be improved. If the parking space is occupied by the parking space, the normal traffic of other vehicles will be affected, and even the pedestrian traffic efficiency will be affected. When the traffic speed of the road decreases, it will damage everyone's traffic. right.

If a motorized vehicle is parked at will, it will lead to an increase in traffic accidents. Sudden parking of the motor vehicle will also make other drivers unable to prepare, so the collision accident and the rear-end accident will increase. The traffic safety of urban roads will be affected as a result, and it will also be a kind of damage to the lives and property of the masses.

The random parking of motor vehicles will lead to the unsettled roads, and most of the places where we park vehicles illegally are schools, or banks and restaurants, so if the roads are occupied for a long time, it will form a blockage point. Causes unnecessary blockage on the road.

Deep Learning is an evolution of the Mountain Artificial Neural Network ( ANN), which is usually composed of multiple layers of nonlinear arithmetic units. Using the lower layer output as the input of the higher layer, this structure is used to automate the analysis of massive training data and obtain abstract expressions of related features. Compared with the shallow network, the network model with

multiple hidden layers can more completely mine the data features. In recent years, the development of computer hardware has been changing with each passing day, and the deep learning model has achieved rapid development.

At present, there are many techniques applied to vehicle detection and recognition. In the literature [1], LimTR et al. used Uabor filtering and support vector machine to complete vehicle detection; Cui B et al. proposed in the literature [2] to use the deformation template to study vehicle detection. In recent years, there have also been methods based on deep learning to identify models. In the literature [3], a method based on unsupervised learning convolutional neural network is proposed. In the literature [4], Heikki Huttunen et al. use the method based on deep neural network to realize vehicle type recognition.

Convolutional neural networks [5] (CNNs) have been widely used in various fields, such as target detection [6], speech recognition [7] and so on. In the field of target detection, the development of region-based CNN (R-CNN) is very rapid. Among them, the Faster-RCNN[8] model won the 2015 COCO Testing Contest and achieved the highest target detection accuracy at PASCAL VOC2007 and PASCALVOC 2012.

At present, there are not many researches on the use of the Faster-RCNN model for vehicle identification. In [9], Shoaib Azam et al. used Faster-RCNN to perform vehicle pose detection from four different orientations (front, back, left and right); Tianyu Tang et al. in [10] The vehicle data set obtained by aerial photography is based on the vehicle detection research based on Faster-RCNN. In real life, whether it is road monitoring, highway tolls or parking lots, the source of information for vehicle images is obtained by a single camera. Therefore, the vehicle image information will only be sourced in a certain direction, and the processing of multi-directional vehicle picture information is generally not applicable to real life.

This paper takes a certain street in Hebei Province as the research object, and applies the Faster-RCNN to the vehicles for road monitoring, and obtains more vehicle information from only one direction to realize the identification of the monitored vehicles. The research on the illegal parking area under video surveillance, and proposed to use the machine vision-based method to automatically detect the illegal parking. In the monitoring image, according to the street view image information detected by the computer, the state of the vehicle placement can be evaluated. The detection algorithm in this paper uses the multi-angle suggestion area to accurately locate the illegal parking in the image and mark it on the detected vehicle. The experimental data shows that the algorithm has good adaptability to the detection of illegal parking in surveillance images.

## **2. Method**

Compared with the common multi-layer perceptual mechanism model, the convolutional neural network has more rich architecture and more network layers. The convolution network specifically includes a convolution layer, an activation layer, a pooling layer, a fully connected layer, and the like.

Among them, the convolution layer is in the central area of the network, and plays a role in improving the level of network nonlinearization. The behavior of convolving the image (matrix) and the filter matrix (a set of weight matrices obtained by training samples) (the elements are sequentially multiplied and then summed) is called a convolution operation. The flow chart of the convolution layer is in the specific application. First, the  $l_m-1$  layer contains four channel features. As the convolution kernel, it is responsible for grabbing features. After the convolution starts, the convolution kernel will be in  $l_m$ - Convolution on layer 1 and output a feature plane, input to the  $l_m$  layer.

The activation layer has the function of nonlinear mapping. The parameters related to the convolution kernel are generally obtained through sample training, so it will produce a value lower than zero. This situation is different from the excitability of neurons in the simulated human brain thinking process. The simulation is implemented in the form of a build activation function.

Common activation functions in convolutional neural networks are other variants of ReLU functions such as Sigmoid, tanh, ReLU, and Leaky ReLU.

The above formulas are the most widely used activation functions. At the beginning of the neural network, Sigmoid is generally used as an activation function. However, the disadvantage is that if the input value is too large or too small, the gradient will disappear, and the output average of Sigmoid will be Not zero, and it takes a long time to perform index calculations. For the tanh function, even if its mean remains at 0, it faces the same problem as Sigmoid. The ReLU activation function shows the advantages of rapid decline and simple calculation, and this will cause the activation function to disappear. The functions such as Leaky ReLU and P ReLU effectively avoid the above problems.

The role of the pooling layer is to shrink the feature map and determine the strong activation point.

The pooling function is usually used to describe the partial pooling of the feature graph, which is usually refined into two types: mean pooling and maximal pooling. The former generally sums all the pixels in the  $n \times n$  pixel block, and then calculates the mean value, which is used as the corresponding activation value of the next layer; the latter generally takes the highest value of the pixel in the  $n \times n$  pixel block as the corresponding activation value of the next layer.

There is also a global average pooling layer called Avg Pooling. The number of parameters based on the full connection layer is very large and easy to over-fitting, whereas Avg Pooling implements an average pooling for the final feature map, that is to say each feature map corresponds to a feature output, and this output represents the characteristics of the relevant category. Avg Pooling not only minimizes network parameters, but also calculates the sum of spatial information, so it shows high stability. In addition, there is a unique type of pooling layer that is globally averaged, mainly used to replace the fully connected layer, implement feature mapping, and obtain better detection accuracy and fewer parameters than FC.

The fully connected layer is the final output of the convolutional layer. The fully connected layer is composed of a classifier composed of the mountain BP neural network, which is equivalent to transforming the feature map obtained by the previous convolution into a one-dimensional feature vector. If the processing of the convolutional layer, the pooling layer, and the activation function is interpreted as mapping the initial value to the hidden feature space, then the fully connected layer maps the learned feature representation to the sample mark space.

### 3. Experience

The Faster-RCNN algorithm unifies the entire inspection process into a network for end-to-end training. The algorithm replaces the candidate region generation method of Selective Search in the above with the RPN network, so that the generated candidate region shares the convolution feature with the target detection network, and the process of extracting the candidate region can also be concentrated in the GPU, compared with the Fast-RCNN. The detection efficiency of the algorithm is higher without affecting the detection accuracy. The detection structure of Faster-RCNN is as follows.

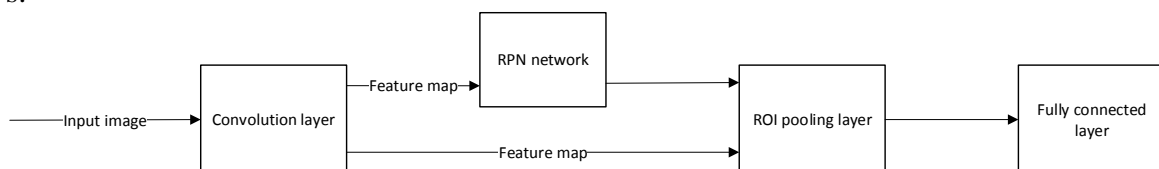


Fig.1 Functions of Faster-RCNN.

The detection process of Faster-RCNN is as follows:

- (1) Input image (224\*224 three-dimensional image);
- (2) Convolutional transformation of images. The convolutional network in this paper uses VGG-16[9] network. The network has 13 convolutional layers and 3 fully connected layers, totaling 16 layers of networks.

(3) On the last layer of the convolutional feature map, the candidate region is extracted using the RPN network.

(4) Using the ROI layer to integrate candidate regions of different sizes into a vector of fixed dimensions for classification. The input of the ROI layer is the same as that in the Fast-RCNN above. It consists of two parts. The first part is the last layer convolutional feature map (the feature size obtained by VGG16 is 7\*7); the second part is the candidate area. Window position.

(5) Classify all candidate areas and frame finishing.

After the Faster-RCNN was proposed, it achieved high detection accuracy in the PASCAL VOC 2007 database test set. The Faster-RCNN algorithm is widely used in practical scenarios. The data shows that the PASCAL VOC2007 data test set has an average precision value (mAp) of 73.2070 for 20 types of targets. For the content of this paper, the improved Faster-RCNN algorithm is used to monitor the parkings in the surveillance video. Usually target detection will use ImageNet classification [12] pre-trained models for training network initialization, which is also the standard practice of target detection methods [11]. The three convolutional neural network models obtained by ImageNet classification pre-training are used for the initialization of the shared convolutional layer in the Faster-RCNN. Most deep learning methods deal with data sets by distributing the data sets in a certain proportion to generate training sets, validation sets, and test sets.

In the experiment, the BIT-Vehicle data set randomly generated the training verification set and the test set in a ratio of 8:2; in the training verification set, 80% was used as the training data set, and the remaining 20% was used as the verification data set. Use the assigned BIT-Vehicle data set for Faster-RCNN. The experimental results are shown below.



Fig. 2 Monitoring video processing results.

In the streets where the illegal parking is dense, the monitoring system based on the convolutional neural network also has a good performance. The results of the field test are shown below.

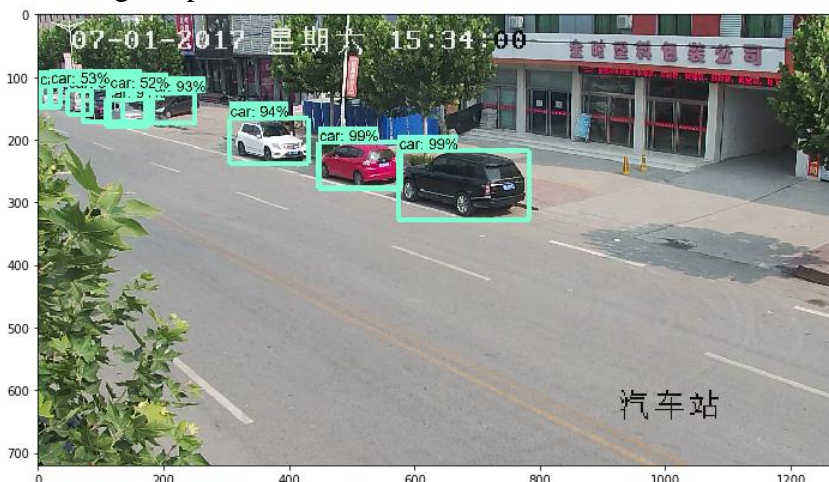


Fig.3 More complicated processing results.

#### **4. Conclusion**

At present, the continuous development of the modern technology industry, production in various industries is gradually replacing the manual method with automated equipment. The automation industry has saved production costs and improved human life. In the highly automated production industry, where machine vision is currently dominant, artificial intelligence is widely used in various fields. Urban management is gradually developing in the direction of informatization. There are corresponding processing mechanisms for illegal parking in urban management, but the method of detecting illegal parking based on machine vision is still under study.

This paper introduces the concept of convolutional neural network, common network model and three target detection algorithms based on convolutional neural network, namely R-CNN, Fast-RCNN, Faster-RCNN. For the research content of this paper, the Faster-RCNN detection framework is combined with the VGG-16 deep learning network model. In the initial parking detection process, the detection algorithm is improved, and the algorithm is more suitable for parking detection in monitoring video.

#### **References**

- [1] Thiang R L, Guntoro A T, Car Recognition using Gabor Filter Feature Extraction, Conference on Circuits & Systems. IEEE(2002).
- [2] Cui B, Xue T, Yang K, Vehicle Recognition Based on Support Vector Machine, International Symposium on Intelligent Information Technology Application Workshops. IEEE(2008).
- [3] Zhen D, Pei M, Yang H, et al, Vehicle Type Classification Using Unsupervised Convolutional Neural Network, International Conference on Pattern Recognition(2014)1-10.
- [4] Huttunen H, Yancheshmeh F S, Chen K, Car Type Recognition with Deep Neural Networks (2016).
- [5] Girshick R, Donahue J, Darrell T, et al, Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation, CVPR (2014).
- [6] Girshick R, Fast R-CNN, Computer Science (2015).
- [7] He K, Zhang X, Ren S, et al. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence. 37 (9) (2014) 1904-16.
- [8] Ren S, He K, Girshick R, et al, Faster R-CNN: towards real-time object detection with region proposal networks, NIPS (2015).
- [9] Liu B, Zhang X, Gao Z, et al, Weld Defect Images Classification with VGG16-Based Neural Network, International Forum on Digital TV and Wireless Multimedia Communications(2017).
- [10] Tang T, Zhou S, Deng Z, et al, Vehicle Detection in Aerial Images Based on Region Convolutional Neural Networks and Hard Negative Example Mining, Sensors. 17(2) (2017)336.
- [11] Long J, Shelhamer E, Darrell T, Fully Convolutional Networks for Semantic Segmentation, IEEE Transactions on Pattern Analysis & Machine Intelligence. 39 (4) (2014)640-651.
- [12] Russakovsky O, Deng J, Su H, et al, ImageNet Large Scale Visual Recognition Challenge, International Journal of Computer Vision. 115(3)(2014)211-252.