

Application of Neural Network in Image Detection of Illegal Billboards

Xue-Hong JIANG¹, Hui-Li FENG² and Yi-Jie DONG^{3,a,*}

¹ Information Center of Department of Housing & Urban-Rural Development, Shijiazhuang 050011, China

² Hebei Building Technology Co., Ltd., Shijiazhuang 050011, China

³ College of Artificial Intelligence and Data Science, Hebei University of Technology, Tianjin 300400, China

^alishiyue11@qq.com

*Corresponding author

Keywords: illegal billboards, image detection, convolutional neural networks

Abstract. In recent years, urban management is gradually developing towards informationization. There are corresponding processing mechanisms for illegal billboards and illegal parking in urban management, but the method of detecting illegal billboards based on machine vision is still under study. At present, the automatic detection method of illegal billboards based on machine vision is generally carried out under ideal conditions, and the experiment is carried out under a simple background. This paper takes a street in Hebei Province as the research object, and studies the illegal billboard area under natural conditions, and proposes to use the machine vision-based method to automatically detect the illegal billboards. In the surveillance image, according to the street scene detected by the computer. Image information enables status evaluation of billboard placement. The detection algorithm of this paper uses the multi-angle suggestion area to accurately locate the illegal billboards in the image, and mark the detected billboards. Compared with the direct detection of the billboard in the image, the interference of the background factor on the target area is removed, and the false detection rate is effectively reduced. In billboard detection, this paper proposes an improved region suggestion algorithm to generate higher quality candidate regions in an image for adapting to billboards of various poses under natural conditions. Experimental data shows that the algorithm has good adaptability to the detection of billboards in surveillance images.

1. Introduction

All kinds of illegal outdoor billboards that have been approved for setting up, occupying urban space, influencing the cityscape and the hidden dangers around the main roads. The illegal billboards are not only disorderly, but also affect the beauty of the city's cityscape. The billboards that are in disrepair will be corroded and weathered, endangering the safety of the citizens; the billboards on both sides of the road are also easy to block the line of sight, causing frequent traffic accidents. Therefore, it is necessary to find and handle illegal billboards in a timely manner. However, at present, the investigation and management of illegal billboards by urban administrative departments is mainly through inspections and reports by the masses, and the efficiency is difficult to maximize. With the development of science and technology and the improvement of life, artificial intelligence has gradually entered human life. Applications based on machine vision algorithms have been popularized in life, such as vehicle detection, pedestrian detection, object tracking, target recognition and other practical applications. The traditional manual operation method greatly improves the production efficiency. Therefore, in the field of urban management, how to use the machine vision-based target detection algorithm to accurately locate the illegal billboards and then transmit the detection information to the computer, which can greatly improve the efficiency of the illegal billboards, but The illegal billboards are complicated to distribute in the image, and the detection task is difficult. It is necessary to adopt an appropriate algorithm for accurate detection.

This paper takes a street in Hebei Province as the research object, and studies the illegal billboard area under natural conditions, and proposes to use the machine vision-based method to automatically detect the illegal billboards. In the surveillance image, according to the street scene detected by the computer. Image information enables status evaluation of billboard placement. The detection algorithm of this paper uses the multi-angle suggestion area to accurately locate the illegal billboards in the image, and mark the detected billboards. Experimental data shows that the algorithm has good adaptability to the detection of billboards in surveillance images.

2. Method

Target detection is to draw a range of objects of interest in the image, then select the target frame with a rectangle and attach a category label to it. The automatic detection of illegal billboards based on machine vision technology is of great significance for the follow-up work such as statistics and inventory of billboards in urban management and the handling of illegal billboards. Improper handling can also cause false detections. There are few research methods for billboard detection. The target detection method is generally aimed at vehicles, pedestrians and other targets. The contours of this type of target are simple and unobstructed, and the background of the image is simple and easy to detect. However, billboards have the characteristics of small area, irregularity and uneven distribution, and the traditional detection methods have poor adaptability.

In recent years, the object detection algorithm based on machine vision has been continuously innovated, and the detection accuracy has been continuously improved. Researchers at home and abroad have proposed many feasible algorithms for different types of targets, which are widely used in practical scenarios. The existing target detection algorithms are mainly built around the following three mathematical models.

2.1 Target detection algorithm based on sliding window extraction feature

The process of the target detection algorithm based on the sliding window is: decomposing an image into a certain number of sub-windows of different positions and scales, and for each window, using a classifier to determine whether the target object is included, the method is for different types of objects, Different features and classifiers are designed. Commonly used target features are color features, texture features, and direction gradient histogram features (Histogram of Oriented Gradient HOG), etc., and then such features are trained in a large number of samples, and finally It is detected by template matching. This kind of algorithm has certain anti-noise ability, and it has a good effect on the single target with clear outline in the image. However, in the case of mutual occlusion between multiple targets, the contour of a single target cannot be accurately obtained. The template does not match correctly. The variable part model algorithm proposed by Felzenszwalb[1] in 2008 (DeformablePart Model, DPM) is the most successful sliding window based detection algorithm, which is based on multi-size images. The HOG feature effectively solves the problem of different target scales and is widely used in practical scenarios.

The method needs to perform filtering scan on each position in the image, and the detection speed is slow, and the detection precision is low in the case where the multi-body is occluded under natural conditions. Therefore, under the condition of complex background and strong interference factors in the actual field of view, more advanced detection algorithms are needed to obtain higher detection accuracy.

2.2 Target detection algorithm based on candidate region

For the shortcomings of traditional detection methods, foreign researchers have proposed a novel target detection strategy to obtain candidate regions from images. The candidate area indicates the possible location of the target. The number of these areas is much smaller than the number of sliding windows, and the high recall rate can improve the efficiency when further complex classifiers are used. This method is the candidate area method. The method is suitable for the problem that the target to be detected is in a complex background, and has good robustness to disturbance factors such as

illumination intensity, image rotation, noise interference, and image blur. In [2], 12 common methods for generating candidate regions were evaluated in depth. The results show that the more accurate the candidate region is located, the greater the help for the classifier. The experimental algorithm in the paper is in the datasets of PASCAL VOC2007[3] and ImageNet[4]. The quality of the candidate regions generated by each algorithm is high and the performance is basically the same. The target category has good adaptability and certain generalization ability.

2.3 Target detection method based on convolutional neural network

Convolutional neural network as a deep learning model, which can automatically learn and extract features from data, its generalization ability is significantly better than traditional algorithms, and has been successfully applied to image recognition and speech recognition. Convolutional neural network is a multi-layer learning network consisting of input layer, hidden layer (including convolution layer and down sampling layer) and output layer. Each layer is a kind of transformation (mapping). Convolution transformation, pool transformation, batch standardization, Dropouts and other operations. Among them, the convolution transform is the core idea of convolutional neural network. The convolution operation in mathematics is used to process the data in the image, and the features in the image are extracted, which is called convolution feature. This feature reflects the deep information of the image.

Due to the introduction of convolutional neural networks, the machine vision-based target detection algorithm has a qualitative improvement in detection accuracy compared to traditional algorithms. The target detection algorithm using convolutional neural network [5] (Regions with Convolutional Neural Network, R-CNN) is a precedent in this field. For the first time, the algorithm uses the convolution feature of the image to replace HOG, Local Binary Patterns (LBP). Algorithms such as Scale Invariant Feature Transform (SIFT) combine image selection methods with Selective Search and use the candidate region as input to the convolutional neural network for each candidate region. The convolution feature extraction, including convolution, pooling, etc., obtains the output of the fixed dimension, and finally classifies according to the output vector of the previous step and returns to the target bounding box border to obtain the final output position. The proposed algorithm solves the problems in multi-objective and multi-category detection, such as mutual interference between targets and image scene changes. However, it brings high precision and takes a long time. Because each candidate region generated in the image is subjected to convolution, pooling, etc. (approximately 2000 candidate regions are generated for one image), a large number of operations cannot bring practical use. The value of Fast-Region based Convolutional Neural Network[6] (Fast-RCNN) algorithm effectively solves the problem of data operation time. This method is improved on the basis of the basic algorithm R-CNN. The network input of the algorithm is A complete image improves the efficiency of the algorithm and introduces the RoI Pooling Layer[6], which makes the last convolution layer feature pooled and produces a fixed-size convolutional feature map to match the subsequent fully connected layers. The algorithm uses the regional recommendation network [7] Region Proposal Networks, RPN, which shares the convolution feature of the full map with the detection network, so that the generation of candidate regions takes almost no time. RPN is a full-convolution network that predicts both target and target category scores at each location. By training the RPN separately, a higher quality recommendation area can be formed. High average accuracy achieved on PASCAL VOC 2007[3] and PASCAL VOC 2012[3] (mean Average Precision, mAP)

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.

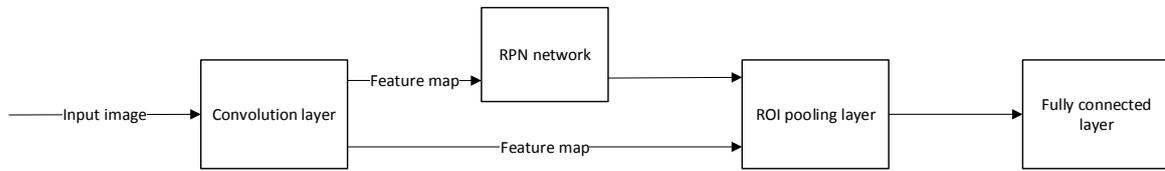


Fig. 1 Functions of Faster-RCNN.

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[8] 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 parking in the surveillance video. Usually target detection will use ImageNet classification 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.

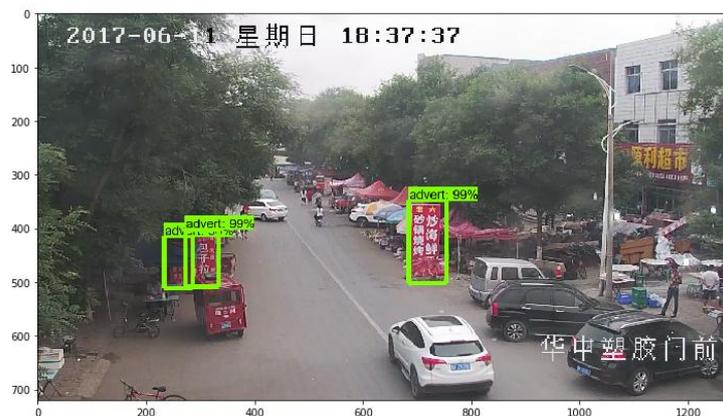


Fig.2 Monitoring video processing results.

In the streets where the illegal billboards are 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

Urban management is gradually developing towards informationization. Through image detection, detecting illegal billboards in surveillance video is a new attempt. At the same time, intelligent algorithms such as artificial neural networks have shown superior performance in image processing. Therefore, this paper studies the application of convolutional neural networks in the identification of illegal billboards.

In this paper, a Faster-RCNN detection framework based on a convolutional neural network is combined with a VGG-16 deep learning network model. The detection algorithm is improved to make the algorithm more suitable for the detection of illegal billboards in video surveillance. Management information and intelligence help.

References

- [1] Felzenszwalb P F, Mcallester D A, Ramanan D, A Discriminatively Trained, Multiscale, Deformable Part Model, CVPR (2008).
- [2] Hosang J, Dollar P, Dollar P, et al, What Makes for Effective Detection Proposals, IEEE Transactions on Pattern Analysis & Machine Intelligence. 38(4)(2016)814.
- [3] Everingham M, Eslami S M A, Gool L V, et al, The Pascal Visual Object Classes Challenge: A Retrospective, International Journal of Computer Vision. 111(1)(2015)98-136.
- [4] Deng J, Dong W, Socher R, et al, ImageNet: A large-scale hierarchical image database, IEEE Conference on Computer Vision & Pattern Recognition (2009).
- [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] Ren S, He K, Girshick R, et al, Faster R-CNN: towards real-time object detection with region proposal networks, NIPS (2015).
- [8] 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).