CNN Based Vehicle Counting with Virtual Coil in Traffic Surveillance Video

Jilong Zheng
Electrical and Communication Engineering
Beijing Institute of Technology
Email: jlzheng90@gmail.com

Yaowei Wang
Beijing Institute of Technology
Email:yaoweiwang@gmail.com

Wei Zeng
Peking University
Email: weizeng@pku.edu.cn

***Abstract*—This paper presents an efficient method of vehicle counting based on convolutional neural network (CNN) with virtual coils. Within virtual coils, foreground is obtained by background substraction. Vehicle is then detected by voting of virtual coil sub-regions. To deal with vehicle cross-lane cases, a cascade classifier combining connected component analysis (CCA) and CNN is adopted. Experiments are carried out on seven real traffic videos. The proposed approach works well on recognizing cross-lane vehicles, achieving above 90% accuracy with real-time processing speed.**

***Keywords*-vehicle counting; virtual coil; CCA; CNN**

I. Introduction

Vehicle counting, as one of key technologies of intelligent transportation system, has received considerable interest in industry and research community. Video based vehicle counting using virtual coils is widely used due to fast speed and easy implementation. However, accuracy of traditional technologies[1][2] drops dramatically during vehicles running on lane lines. To cope with the difficulty, a cascade classifier consists of connected component analysis (CCA) and convolutional neural network (CNN) is proposed. The cascade classifier can remove missing and false detection efficiently.

II. Proposed Method

Our vehicle counting algorithm includes three steps. In the first step, within virtual coils, foreground can be roughly calculated through background substraction. By voting of virtual coil sub-regions, vehicle blobs can be detected. Secondly, when foreground covers two adjacent virtual coils, the CCA is adopted to counting the connected blobs, if more than one connected blobs are detected, the CNN is adopted to recognize whether there is one vehicle running on the lane line. Finally, the counter eliminates the repeat counting if there is a cross-lane vehicle.

A. Vehicle detection using virtual coil

In traffic videos, lane lines are not parallel due to perspective phenomenon in 3D scenes. In order to fully fit lane line, trapezoidal virtual coils are adopted (see Fig. 1). A virtual coil is placed in one lane at fixed position. The virtual coil is divided into nine parts, each part is called as sub-coil and assigned with a predefined weight. While a vehicle passing through the virtual coil, each sub-coil calculates the proportion of foreground pixels. f_j represents the number of foreground pixels in the jth sub-coil; b_j represents the number of background pixels in the jth sub-coil; $g(y_j)$ represents the proportion of foreground pixels of the jth sub-coil. $g(y_j)$ is written as

$$g(y_j) = \frac{f_j}{f_j + b_j}. \quad (1)$$

Each $g(y_j)$ is multiplied by predefined weight w_j as its score, then all scores are summed and normalized, the result is defined as $Score(x)$. Each virtual coil x is scored by a function,

$$Score(x) = \Sigma_{j=1}^{9} w_j g(y_j), \quad (2)$$

where w_j is the weight of the jth sub-coil. Finally, if $Score(x)$ is above the predefined threshold, a vehicle blob is detected.

B. Cross-lane recognition based on Connected Component Analysis

When two adjacent virtual coils are both occupied by foreground, there are two possible cases: one case is two vehicles passing through two adjacent virtual coils respectively, which can be counted without error counting; the other case is one vehicle running across two lanes, may leading to repeat counting without processing(see Fig. 1). To recognize the cross-lane vehicle, a cascade classifier based on CCA and CNN is adopted. Connected component generally refers to pixels of foreground that the pixel value is the same and the positions is adjacent. CCA is used to calculate the number of connected component based on Depth-First-Search. To eliminate noise, connected components of small size are removed. When there is only one connected component, it means one vehicle is passing across two lanes.

C. Cross-lane recognition based on CNN

In traffic videos, the foreground of vehicle is usually incomplete due to inaccurate background substraction. The CCA can not accurately recognize cross-lane vehicle based on incomplete foreground. To cope with this case, a classifier based on CNN is adopted. CCA and CNN constitute a cascade classifier to recognize cross-lane vehicle.

978-1-4799-8688-0/15 $31.00 © 2015 IEEE
DOI 10.1109/BigMM.2015.56

Figure 1. Two situations that need CCA and CNN. Images in the first row show that one vehicle is running across two lanes. Images in the second row show that two vehicles are running on their own lanes.

The network structure of CNN that is suitable for cross-lane recognition is designed by us. There are a total of six layers in our network structure. The resolution ratio of all the training samples are resized to 96×32. In the first convolution layer, 12 different 5×5 convolution kernels are adopted. In the second convolution layer, 24 different 3×3 convolution kernels are adopted. The downsampling layer is inserted after each convolution layer. The sixth layer, as the output layer, is fully connected to the above layers.

To construct the training dataset, five thousand training samples are collected. All training samples are collected in the real traffic videos including highway and crossroad scenes. The foreground within the region of adjacent virtual coils is extracted as a training sample. Cross-lane vehicles are chosen as positive samples, two vehicles cover two adjacent lanes respectively are chosen as negative samples. The positive and negative samples quantity are basically the same. All training samples are mirrored to fit in more camera angles.

III. Experiments

The algorithm is evaluated using seven test videos. All test videos are from the traffic surveillance video of reality scenes. The test videos consists of four surveillance videos captured from highway and three surveillance videos captured from crossroads. The length of each video is twenty minutes with resolution of 704×576 or 480×270. Camera angles of test videos relative to the road are between 30 degrees to 60 degrees. Each test video contains partial occlusion vehicles, cross-lane vehicle and incomplete foreground.

A. Evaluation Criteria

The experimental results are evaluated through two criterias: the relative accuracy P_r and the absolute accuracy P_a. P_r and P_a are defined as follows:

$$P_r = 1 - \frac{|S_r - S_m|}{S_a}. \quad (3)$$

$$P_a = 1 - \frac{S_m + S_r}{S_a}. \quad (4)$$

S_a represents the actual counting of the number of vehicle, S_a is got by manual counting; S_m represents the number of omission counting and S_r represents the number of repeating counting. To analysis the contribution of each component, P_rc and P_ac is adopted. P_rc and P_ac respectively represent P_r and P_a only using CCA without CNN.

B. Evaluation Results

Seven test videos are used to evaluate our algorithm, the results are summarized in table I. Our algorithm performs well on all test videos, the average of P_a exceeds 90% and the average of P_r exceeds 95%. By contrast, the P_a and P_r are greatly increased with CNN. The algorithm can achieve real-time counting with CPU of i5-4570, 3.20GHz. The accuracy of Crossroad2 is the lowest due to serious losses of vehicles foreground by inaccurate background substraction, leading to omission counting. The accuracy of Highway2 is the highest because of relatively stable background. Our algorithm can be well adapted to the cases of cross-lane vehicle, incomplete foreground and noise of scene.

Table I
Evaluation Results of Each Test Video

DATASET	PERFORMANCE						
	S_a	S_m	S_r	P_r	P_a	P_rc	P_ac
Highway1	600	19	23	0.992	0.93	0.933	0.87
Highway2	360	11	11	1	0.939	0.919	0.868
Highway3	570	21	23	0.996	0.923	0.964	0.891
Highway4	300	12	10	0.993	0.927	0.877	0.797
Crossroad1	390	10	19	0.977	0.926	0.895	0.844
Crossroad2	270	17	7	0.963	0.911	0.893	0.767
Crossroad3	300	8	15	0.977	0.923	0.883	0.83

IV. Conclusion

In this paper, an efficient vehicle counting approach based on CNN using virtual coils has been presented and discussed. First, within virtual coils, vehicle foreground is calculated by background substraction. Vehicle is then detected by voting of virtual coil sub-regions. To deal with cross-lane vehicle, a cascade classifier consists of CCA and CNN is adopted. The algorithm is evaluated using seven traffic surveillance video of reality scenes. Experimental results shows our algorithm can reach high accuracy: P_a exceeds 90% and P_r exceeds 95% with real-time processing speed.

References

[1] Zhu H. and Ye Jun *Adaptive vehicle detection and traffic statistics algorithm based on improved Gaussian model.* LASER & INFRARED,vol. 43, no. 11, 2007.

[2] WANG, W. and WANG, K.and HAO, Z *Video-based Statistical Algorithm for Traffic Flow.* Science Technology and Engineering,vol. 17, p. 11, 2011.