№ 3 (3), 2021 ISSN: 2682 – 9517 (print) ISSN: 2683 – 0930 (online)
INSTITUTE OF INFORMATICS AND INNOVATIVE TECHNOLOGIES

COMPUTER VISION TECHNIQUES IN INTELLIGENT TRANSPORTATION SYSTEMS

Cao Phuong Thao **Nguyen Thi Hong Hoa** **Dinh Cong Tung** **Nguyen Huy Cuu**

University of Transport and Communications, Hanoi, Vietnam

thaocp@utc.edu.vn; hoanth@utc.edu.vn; tungdc@utc.edu.vn; cuunh@utc.edu.vn

Abstract. **Vehicle's detection and classification is an important tool in intelligent transportation systems. Recently, surveillance cameras with image processing and computer vision techniques are widely applied due to their ease of installation, upgrading and maintenance, allowing direct detection and classification of vehicles. However, these approaches were unable to adequately distinguish vehicles from each other when those vehicles look similar and involve in complex transportation conditions. This paper presents a method to detect vehicles going in the wrong lane using deep learning and computer vision method, which can accurately identify vehicles and their direction. In this method, Convolutional Neural Networks and Kalman filter is used to detect and track vehicles. Then the direction of vehicles can be detected by those coordinates. Our experiments show that the proposed mechanism achieves high accuracy even with complex transportation constraints.**

Keywords: computer vision, convolution neural network, ITS, vehicles detection.

I. INTRODUCTION

Camera surveillance with computer vision techniques provide a flexible way of monitoring the vehicles in the road, especially monitor the complex transportation. In intelligent transportation systems (ITS) [1, 2, 3], detecting moving vehicles from camera video stream is a fundamental task. The task of identifying moving vehicles is normally performed in two steps: vehicle detection and vehicle tracking. In the two mentioned steps, vehicles under monitoring are detected and then tracked by the surveillance system. More specifically, given a video stream recorded by surveillance systems, detection and tracking algorithms will identify target vehicles in consecutive time. The output of these two algorithms will be send to the transport management center for further analysis such as vehicles speed detection, vehicles breaking traffic rule and traffic monitoring [4, 5].

There are number of object detection approaches being used in surveillance systems where vehicle detecting is one of essential case. In [6], the histogram of gradients (HOG) is one of method to detect vehicles in a region of interest by using sliding windows moving in the frame to calculate features. Haar features can be used to determine whether the given individual image sample contains vehicle or nonvehicle [7]. Recently, deep learning-based feature extraction and classification methods for real-time vehicle classification have been proposed [8]. After vehicles are detected, the tracking method can be applied to predict the trajectory of vehicles.

The most common method for object tracking is using Kalman filters, which are recursive estimators for states of dynamic systems [9, 10]. To increase the accuracy, mean-shift was combined with Kalman filter to predict the search regions [11]. If the system does not fit into linear model, particle filter is an important method to track the object [12]. It combines gray and contour feature particles using fusion algorithm to balance the weights according to the present scene. Motion direction and assignment can be used to track the vehicles in their lanes and calculate the speed of the vehicles [13]. Image segmentation and pattern analysis techniques are also applied in the system to detect and track the moving vehicles at day and nighttime [14] by recognize headlight and taillight of vehicles. Using cameras and the pattern recognition techniques, the traffic flow can be measured under various environments conditions by detecting vehicles.

The current techniques for detection and tracking in transportation surveillance systems are still facing challenges that are not completely solved. Under complex transportation conditions, especially where multiple vehicles run concurrently without orders, the detecting and tracking algorithms cannot accurately and efficiently track vehicles. This paper proposes a method that simultaneously detect and tracks vehicles in a sequence of video frames using convolution neural network and multiple Kalman filters. The method detects moving vehicles in each frame and associates these vehicles corresponding to those in successive frames. Particularly, Kalman filters are used to predict vehicle positions and use predicted positions for associations. Then the vehicles go in wrong way will be detected. Experimental results show that the proposed algorithm is able to perform multiple vehicles detecting in wrong way simultaneously with high level of robustness and efficiency.

The rest of this paper is organized as follow: Section 2 presents a framework of the vehicle detection and tracking. Section 3 demonstrates the accuracy and robustness of the proposed method. Finally, Section 4 states the conclusions and future works.

II. VEHICLES DETECTION AND TRACKING

The goal of this research is to track multiple vehicles in complex transportation situations. In order to achieve this purpose, this paper propose to use multiple Kalman filters to track multiple vehicles concurrently. To do this, firstly, convolution neural network is used to detect vehicles existing in a frame. According to the number of detected vehicles, a corresponding number of Kalman filters are, then, created. Finally, those filters are used to track detected vehicles in successive frames. The general framework of the method is given in figure 1.

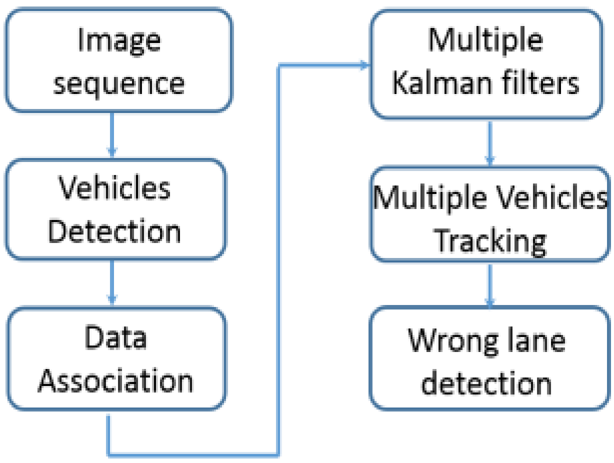

Fig. 1. The framework of the multiple tracking method

A. Convolution Neural Network for Vehicles Detection

Convolution Neural Networks (CNN) method is used to detect vehicles running in the road. A CNN comprises of convolution and pooling layers [15]. Those layers are then connected to a fully connected layers. Convolution and pooling layers extract the feature maps, which are two dimensional matrices of CNN neurons. With the input image x_i from the video sequence, the output of a convolution layer j, denoted by $y_j = b_j + \sum_i (k_{ij} \otimes x_i)$, where $\otimes$ denoted the convolution operator, b_j is a trainable bias parameter, k_{ij} is a convolution layer filter. The feature map y is calculated for any node $y(m,n)$:

$$y(m,n) = k \otimes x = \sum_{u=0}^{U} \sum_{v=0}^{V} k(u,v)x(m+u,n+v),$$

where k is the kernel of size $A*B$ and x is the input image with size $U*V$. The size of the output convolutional is $M*N$ where $M=U-A+1$ and $N=V-B+1$.

The multi-layer structure of CNN brings advantages to the task of vehicle detection. When frames are processed in convolution layers, those layers incrementally learn features from raw images and outputs of the previous layers, which are high level features such as shapes and edges. Convolution layers, thus, represent an image frame into multiple representations at each convolution layer with different levels of abstraction from low to high. This effectively helps in cancelling out noises and refining detection information. The final step of detection is done at pooling layer at which feature maps are extracted and processed so that vehicles are detected regardless of translation, rotation, scaling and other kinds of geometric transformations. As a result, CNN can provide robust detection regardless of where in road a vehicle is captured and which camera is used to capture the vehicle

B. Vehicles Tracking

We use Kalman filter to predict each vehicle in a specific point in time. Basically, a Kalman filter is used to estimate states of a linear system where states are assumed to be Gaussian random variables. Kalman filter algorithm comprises of two steps: prediction and correction. In prediction step, a state is estimated using a state equation. After that, the correction step takes current observations to adjust and update the estimated state in the prediction step. In this paper, to

№ 3 (3), 2021 ISSN: 2682 – 9517 (print) ISSN: 2683 – 0930 (online)
INSTITUTE OF INFORMATICS AND INNOVATIVE TECHNOLOGIES

track multiple vehicles simultaneously, multiple Kalman filters as number of vehicles is used [16]. Each Kalman filter is represented as below:

$$x_k = Ax_{k-1} + w_k$$

$$z_k = Hx_k + v_k$$

where $x = \left[p_x\, p_y\, v_x\, v_y\right]^T$, p_x, p_y are the center position of x-axis and y-axis, respectively. v_x, v_y are the velocity of x-axis and y-axis. Matrix A represents the transition matrix, matrix H is the measurement matrix, and T is the time interval between two adjacent fames. w_k and v_k are the Gaussian noises with the error covariance Q_k and R_k. The Kalman filter is process as follow:

- Update the state: $x_{k|k-1} = Ax_{k-1|k-1}$;
- Predict the measurement: $z_{k|k-1} = Hx_{k|k-1}$;
- Update the state error covariance: $P_{k|k-1} = AP_{k-1|k-1}A^T + Q_k$

To track multiple vehicles in complex transportation, matching between vehicles and measurements should be performed correctly. In this paper, we employ the data association method, which split and merge the vehicles [17]. Overall of the tracking method is given in figure 2.

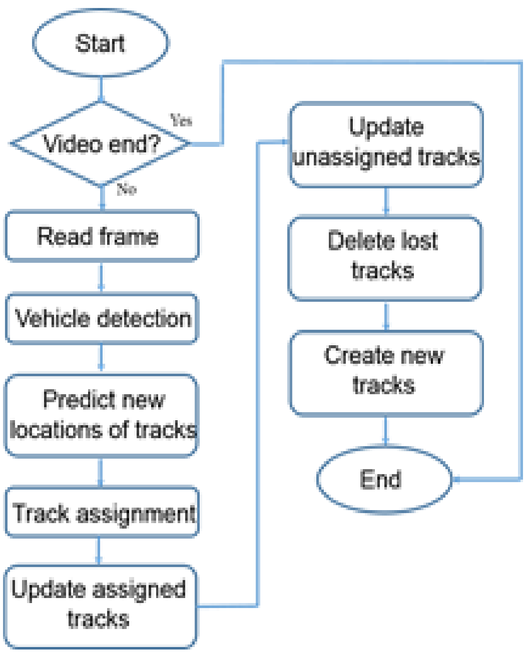

Fig. 2. The flow chart of vehicles tracking method

III. Experimental Results

Wrong way vehicles detection method adapted from [18]. First, vehicles are detected using convolution neural network. Then each vehicle was assigned an identification and track using Kalman filter. To check the direction of vehicles, the difference of the pixel in range of frames will be calculated. For example, the initial center of bounding box of vehicle is C_0 with coordinate (x_0, y_0). After several frames, the center of vehicle is $C_n(x_n, y_n)$. Here n is frame number after initial frame.

The difference of position will be $DeltaC = x_n - x_0$.

If delta$x$ is negative and xn is smaller than x1, the vehicle is moving from right to left, otherwise, the vehicle is moving from left to right. By regulation the rule, we can identify the vehicles go right or wrong way. The results of vehicles detection are shown in figure 3 and 4.

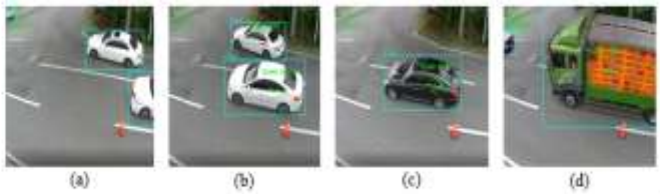

Figure 3. Right way detection

Figure 3 (a) shows the multiple cars detected, then these cars will be tracked by multiple Kalman filters. From trajectories, the center of bounding box of each car is calculated and the direction of the cars have identified. We regulation the rule with the right way from right to left. It means two cars were identified as go in to right way (figure 3 (b)). Figure 3 (c) and 3 (d) shows another case with difference vehicles in right way also detected.

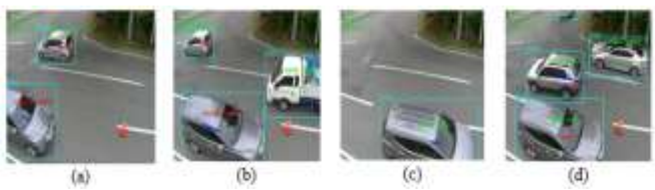

Figure 4. Wrong way detection

Figure 4 (a) and (b) shows the wrong way detected when car is moving from left to right. However, because we use the difference pixel position to detect the direction, if the car go backward (reverse gear), the detection is miss classified as in figure 4 (c) and 4 (d).

CONCLUSIONS

In this paper, we presented a method for multiple vehicles detection and tracking based on convolution neural network and Kalman filter. For each vehicle, a Kalman filter was established and it uses bounding box as feature. The Kalman filter estimates states based on the state equation and corrects using the current observations to update the vehicle states. From trajectories of vehicle, the direction of vehicle has detected and used for wrong way detection. Results of this paper show that this method can be applied in transport management system for transport violation detection.

ACKNOWLEDGEMENTS

This research was supported by University of Transport and Communications under grant number T2021-CN-001.

REFERENCES

[1] B. E. Flinchbaugh, and T. J. Olson, “Autonomous video surveillance”, Proceedings of the SPIE, vol. 2962, pp. 144–151, 1997.

[2] H. Moon, R. Chellapa, A. Rosenfeld, “Performance analysis of a simple vehicle detection algorithm”, Image and Vision Computing 20. 1–13, 2003.

[3] S. M. Baljit and K. Satish, “A Review of Computer Vision System for The Vehicle Identification and Classification from Online and Offline Videos”, Signal & Image Processing: An International Journal (SIPIJ) Vol.6 (5), 2015.

[4] C. Wang, C. Thorpe and A. Suppe, “Lidar-Based Detection and Tracking of Moving Objects from a Ground Vehicle at High Speeds”, Proceeding of the IEEE Intelligent Vehicles Symposium, 2003.

[5] C. Setchell, and E. Dagless, “Vision-based Road Traffic Monitoring Senso”, IEE Proc. Vision, Image Signal Processing, Vol. 148 (1), pp. 78-84, 2001.

[6] T. Gandhi and M.M. Trivedi. Video based surround vehicle detection, classification and logging from moving platforms: Issues and approaches. IEEE Intelligent Vehicles Symposium, 2007, Istanbul, Turkey, June 13–15, 2007. IEEE, New York City, NY.

[7] G.Y. Song, K.Y. Lee, and J.W. Lee. Vehicle detection by edge-based candidate generation and appearance-based classification. IEEE Intelligent Vehicles Symposium, Eindhoven, the Netherlands, June 3–6, 2008. IEEE, New York City, NY.

[8] Z. Dong, Y. Wu, M. Pei, and Y. Jia, “Vehicle type classification using a semi supervised convolutional neural network,” IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 4, pp. 2247–2256, 2015.

[9] G. S. Manku, P. Jain, A. Aggarwal, and L. Kumar, “Object tracking using affine structure for point correspondence”, IEEE conference of Computer Vision and Pattern Recognition, 704-709, 1997.

[10] L. Xin, W. Kejun, W. Wei, and L. Yang, “A Multiple Object Tracking Method Using Kalman Filter”, IEEE International Conference on Information and Automation, pp. 1862-1866, 2010.

[11] D. Comaniciu, and V. Ramesh, “Mean shift and optimal prediction for efficient object tracking”, IEEE Int. Conf. Image Processing, vol. 3, pp. 70–73, Vancouver, Canada, 2000.

[12] D. P. Bai, B. B. Lee, “Based-on Particle Filter For Vehicle Detection and Tracking In Digital Video”, Proceedings of the Seventh International Conference on Machine Learning and Cybernetics, Kunming, pp. 12-15, 2008.

[13] H. S. Lai, and H. C. Yung, “Vehicle-Type Identification Through Automated Virtual Loop Assignment and Block-Based Direction-Biased Motion Estimation”, IEEE Transactions on Intelligent Transportation System, Vol. 1(2), 2000.

[14] K. Dhanya, M. Manimekalai, B. Asmin, and G. Vani, “Tracking and Identification of Multiple Vehicles”, 2014.

[15] S. Christian, T. Alexander, E. Dumitru, “Deep Neural Networks for Object Detection”, Advances in Neural Information Processing Systems 26, pp. 1-9, 2013.

[16] J. M. Jeong, T. S. Yoon, and J. B. Park, “Kalman Filter Based Multiple Objects Detection-Tracking Algorithm Robust to Occlusion”, SICE Annual Conference, Japan, pp. 941-947, 2014.

[17] S. Zheng, “Tracking Multiple Objects In Surveillance Cameras”, Technical Report, 2010.

[18] Saidrasul Usmankhujaev, Shokhrukh Baydadaev, Kwon Jang Woo, "Real-Time, Deep Learning Based Wrong Direction Detection", Applied Sciences, Vol. 10(7), 2543, 2020.