WILEY | Hindawi

Research Article

Traffic Incident Detection Based on mmWave Radar and Improvement Using Fusion with Camera

Zhimin Tao ,[1,2] Yanbing Li ,[3] Pengcheng Wang,[1] and Lianying Ji[4]

[1]*School of Transportation Science and Engineering, Beihang University, Beijing 100191, China*
[2]*Beijing Anlu International Technology Limit Co., Ltd, Beijing 100043, China*
[3]*School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China*
[4]*Muniu Linghang Technology Company, Beijing 100192, China*

Correspondence should be addressed to Zhimin Tao; taozhimin@buaa.edu.cn

Improving traffic efficiency and safety is the goal of all countries due to the increasingly congested road environment worldwide. The progress of intelligence has promoted the development of the transportation industry. As the first step to intelligence, perception technology is an important part to realize intelligent transportation. Accurate and efficient traffic management systems, such as the automatic control of traffic lights at urban intersections or highway emergency disposal, need the support of advanced environmental sensing technology. In the application of traffic perception, millimeter wave radar and camera are two important sensors. Radar has been widely used in traffic incident perception due to its all-weather working capability; however, there are problems such as inability to detect stationary targets and poor target classification performance. Camera has the advantages of accurate target angle information measurement and rich details, but there are problems of inaccurate ranging and speed measurement and performance degradation in harsh weather conditions. Considering the complementary characteristics of the two sensors in information, an improved incident detection method based on radar-camera fusion is proposed. This method combines the advantages of millimeter wave radar and camera and improves the robustness of the traffic incident detection system. The detection performance is verified in the real experiment. The results show that the detection accuracy of the proposed fusion system is better than that of a single millimeter wave radar in all scenarios, and the accuracy is improved by more than 50% in some cases.

1. Introduction

With the rapid growth of vehicle numbers worldwide [1], the urban traffic system is facing more and more challenges such as traffic congestion, low traffic efficiency, and traffic accidents [2–4]. These traffic problems will be more serious in some emerging market countries whose construction of transportation infrastructure cannot keep up with the growth of the car numbers. Therefore, the demand for efficient, intelligent transportation system becomes urgent. In recent years, with the development of information science especially machine learning technology, the intelligence degree of the transportation system is rapidly improved [5–9]. With

the introduction of intelligent technology, the transportation system can reduce road congestion and improve traffic efficiency, which has important social and economic value.

In intelligent transportation system, efficient traffic flow planning and control depends on sufficient and reliable traffic perception data [10–12]. In the congestion scenario, the dense traffic targets pose a challenge to the perception techniques. Traffic jams generally occur in areas with heavy traffic flow, such as intersections and highways. Taking the intersection as an example, the transportation department wants to know the real-time situation of traffic flow and traffic incidents in order to accurately allocate traffic light time, improve traffic

efficiency, and reduce response time in case of traffic accidents [13]. Generally, the categories of traffic incidents include speeding, retrograde vehicle (i.e., vehicle reverse traveling), illegal lane change, abnormal stop, occupation of emergency lane, lane queuing, and intrusion in forbidden area, etc. [14, 15]. The detection of these traffic incidents needs the sensor to continuously track the traffic target, obtain the change of the target state over time, and then detect the possible incidents according to the target state. This requires the sensor to have the ability to track the target within a certain range.

Traditional traffic sensing equipment such as geomagnetic coil [16] can sense the real-time traffic flow of a road cross section; however, the sensing approach of the geomagnetic coil obtains the limited amount of traffic information and has the disadvantages of high construction and maintenance cost, which makes less cost-performance ratio. Moreover, coil solution cannot be used for traffic incident detection which requires monitoring large scenarios. The new video perception technology [11, 17–19] can obtain the traffic state of a certain area on the road, which is suitable for incident detection at intersections and highways. Compared with the coil, video greatly increases the amount of traffic information available and then can detect traffic incidents. Research on cameras has focused on vehicle detection [20–23] and distance estimation [24–27], including optimization of computational complexity and robustness. Because the camera has rich image details, it has good performance for target detection. However, it is difficult to accurately estimate the specific location and speed of the vehicle for traffic camera due to the lack of depth information. In addition, video-based methods are more vulnerable to weather and have poor performance over long distances. In recent years, with the development of millimeter wave (mmWave) chip technology, the difficulty of radar design is reduced. More and more radar sensors are used in intelligent transportation systems because of their good all-weather working ability. In transportation applications, radar technology has evolved from a traditional single velocity measurement function to a tracking function for multiple traffic targets in a large area. It can track traffic targets on the road in real time and obtain accurate range information through direct measurement and accurate velocity information by using the Doppler principle, which provides the necessary information to detect traffic incidents. Despite its advantages in ranging and velocity measurement, radar still needs to face the problems of low angular resolution, unable to detect stationary targets and poor target classification ability.

In terms of the respective characteristics of camera and radar, it can be found that their perception information is complementary. Camera has absolute advantages in angle resolution and target detail information, and radar has advantages in depth information, harsh environment, and long-distance perception. In order to make up for the limitations of a single sensor, the fusion algorithm of radar and camera sensors has been widely studied [28–34]. For example, a simple coordinate transformation calibration method of radar and camera is proposed in [28], which marks the beginning of the research on radar-camera fusion. Gao et al. [29] shows the position information of each sensor on the grid cell of the aerial view and obtains the vehicle position by superimposing the information from radar and camera, but the vehicle position may appear on multiple adjacent grids in real application due to measurement error, so it will produce uncertainty. Feng et al. [30] introduce Kalman filter to improve this issue. In addition, some studies try to improve the radar accuracy through the symmetry of the rear [31], but these methods are easily affected by occlusion or different viewing angles. In general, most sensor fusion research focuses on improving the detection performance through cross verification or reducing the computational load; the weakness of each single sensor has not been fundamentally solved.

In this paper an incident detection approach is proposed via the fusion of traffic environment perception data introduced by mmWave radar and camera. The incident detection performance of radar sensors is analysed, and a method based on the fusion of radar and camera is introduced. The fusion of the sensors can improve the accuracy of incident detection in real traffic scenarios. Experiments with measured data show the effectiveness of the proposed method.

The rest of this paper is organized as follows. The signal model and target tracking flow of frequency modulated continuous wave (FMCW) mmWave radar is introduced in Section 2. Traffic incident detection method based on radar sensing data is introduced in Section 3, followed by the fusion implementation of radar-camera in Section 4. Experimental results in terms of real scenario data and corresponding discussion of performance are given in Section 5. Finally, some conclusions are drawn in Section 6.

2. FMCW Radar Signal Model and Target Tracking Flow

In transportation perception, mmWave radar is usually used for sensing the surrounding environment by emitting linear frequency modulation (LFM) wave which is defined as

$$s_t(t) = A_t \sin\left[2\pi\left(f_c + \frac{1}{2}kt\right)t\right], \tag{1}$$

where t is the time, A_t is the transmitted signal amplitude measured in units of volt, f_c is the carrier frequency in Hertz, and k is the frequency modulation slope measured in units of Hertz per second. When the transmitted signal encounters the target, it is scattered by the target to generate echo signal which is received by the radar. Then the received signal will be dechirped; i.e., the received signal is mixed with a reference signal, to obtain the baseband signal which is defined as

$$s_r(t) = A_r \sin\left[2\pi\left(k\tau t_k + f_c\tau - \frac{1}{2}k\tau\tau\right)\right]$$

$$= A_r \sin\left[2\pi\left(f_b t_k + f_c\frac{2(R_0 - vt_m)}{c} - \frac{1}{2}k\tau\tau\right)\right]$$

$$= A_r \sin\left[2\pi\left(f_b t_k - \frac{2v}{\lambda}t_m + \frac{2R_0}{\lambda} - \frac{1}{2}k\tau\tau\right)\right] \quad (2)$$

$$= A_r \sin\left[2\pi\left(f_b t_k - f_d t_m + \frac{2R_0}{\lambda} + \varphi\right)\right],$$

where A_r is the amplitude of the echo signal measured in units of volt, f_b is the beat frequency in Hertz between the transmitted signal and the echo, R_0 is the target distant from radar, v is the target radial velocity, c is the speed of light, λ is the wavelength of the transmitted signal which equals $\lambda = f_c/c$, and f_d is the Doppler frequency. It should be noticed that t_k is fast time which relates to the sampling time inside the pulse and t_m is the slow time which relates to the pulse repetition interval (PRI), and τ is the propagation time of the signal, caused by the target distance. According to (2), the range and velocity of the target can be obtained by performing 2-dimensional fast Fourier transform (FFT) processing on the echo signal [33]. In the detection stage, the radar equation can be used to calculate the signal-to-noise ratio (SNR) of the target of interest, i.e., a specific radar cross section (RCS) target such as vehicle or pedestrians. When the SNR of the target echo in the received signal exceeds the detection threshold, it is confirmed that the target exists. Usually, the detection threshold is determined by the constant false alarm rate (CFAR) method to ensure that the false alarm is at a certain level [35].

When the radar has multiple antennas, supposing the echo signal returned from the direction θ as shown in Figure 1, considering antenna 0 as a reference, the signal received by the nth antenna is

$$s_{rn}(t) = A_r \sin\left[2\pi\left(f_b t_k - f_d t_m + \frac{2R_0 + (n-1)d\sin\theta}{\lambda} + \varphi\right)\right], \quad (3)$$

where $n = 1, \ldots, N$ is the antenna index in the array. The echo signal received by the antenna array is

$$s_{ra}(t) = \begin{bmatrix} s_{r1} \\ \vdots \\ s_{rN} \end{bmatrix}. \quad (4)$$

Then, the third FFT will be applied to the received signal along with the array, and the angle information of the target is obtained. After obtaining the target's range, velocity, and angle information, Kalman filter is used to track the target [36]. The processing flow is shown in Figure 2 and the following steps are performed:

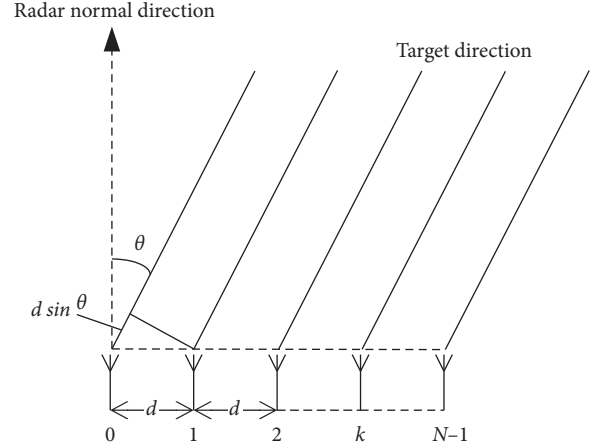(1) Cluster the detection points to merge multiple scattered points of the same target
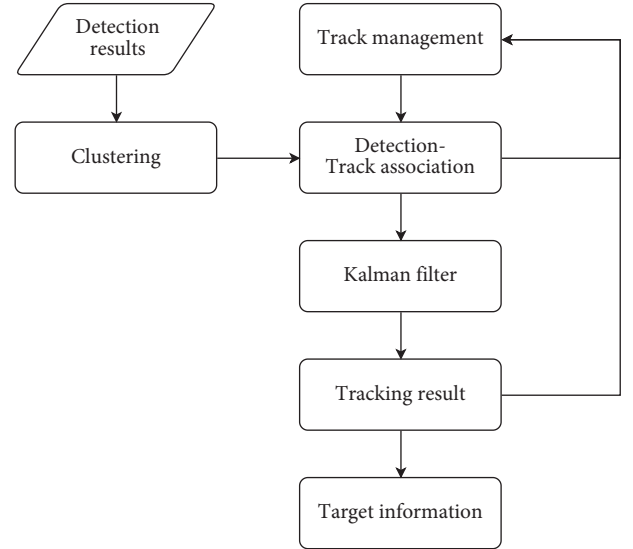

FIGURE 1: Angle measuring diagram of antenna array.

FIGURE 2: FMCW radar signal processing flow.

(2) Associate detections and tracking trajectory to make sure the detection and trajectory from the same target are matched

(3) Use Kalman filter to update the status of the trajectory and the target information including filtered range, angle, velocity, and target length are obtained

A typical protocol of radar output information is shown in Table 1. In traffic applications, the data reporting period of radar is about 100 ms, which ensures the real-time perception of traffic scenario.

3. Millimeter Wave Radar Traffic Incident Detection Method

After obtaining the tracking result of the traffic targets, the radar output information can be used for traffic incident detection. The incident detection can be divided into the following categories according to the implement mechanism and the type of used information.

TABLE 1: Radar output data format.

Data type	Description
ID	Target indication number
Velocity	Radial velocity of target
Range	Radial distance between target and radar
Azimuth	The angle between the target bearing and the radar normal
Length	Radial length of target
SNR	Signal-to-noise ratio
RCS	Measure of a target's ability to reflect electromagnetic wave

(1) Trajectory-based incident detection. This approach is mainly using the target trajectory to determine whether an incident has occurred, for example, whether the trajectory deviates from the lane.

(2) Inference-based incident detection. When the target is stationary, the target velocity is 0. In this case, the target cannot be distinguished from the ground by the difference in speed or distance. As a result, radar cannot reliably detect the stationary target. Therefore, incidents related to parking cannot be directly detected. A feasible way is to conduct indirect detection through the historical information of the target, such as the relationship between trajectory and velocity.

(3) Target classification-based incident detection. For example, the detection of pedestrian entering the highway is based on the results of radar target classification.

The traffic incidents to be detected in this work are shown in Table 2. There are a total of three categories including seven incident types. These incidents have important reference value in law enforcement at intersections and highway scenarios. For each incident, the application scenarios and sensor detection capabilities are given. For radar, due to the inability of stationary targets detection, inference-based incidents related to parking cannot achieve good performance. As for the camera, the velocity measurement of the target is less accurate due to the lack of depth information. It can be seen that the two sensors are complementary in event detection capability.

Specific incident detection methods based on radar tracking results are described in the rest of this section.

3.1. Lane Alignment.
After obtaining the target tracks by radar, it needs to map the target to the correct lane. This involves calibrating the position of the lanes relative to the radar. The typical installation of radar on roads is shown in Figure 3. The local coordinate system of the radar is XYZ, where the radar is located at the origin O. Radar antenna locates in XOZ plane, X is to the right, Y is the normal of the radar antenna, and Z-axis, X-axis, and Y-axis meet the right-hand rule.

Due to the existence of errors, after the radar is installed, the normal direction and the road direction are not parallel, which makes the target trajectory not in the correct lane. To get the correct lane information output by radar, it is necessary to offset the angle difference between the direction of radar antenna normal and the road. The calibration process is shown in Figure 4 and implemented as follows.

(1) Set a test vehicle to drive a straight distance along the lane line

(2) Record the trajectory of the test vehicle in the radar local coordinate system, as shown by the green point in Figure 4

(3) Use a straight line to fit the trajectory, as shown by the red solid line in Figure 4

(4) Calculate the angle α between the trajectory line (parallel to the red dotted line) and the Y-axis (parallel to the solid blue line with arrow), which is the correction value of the radar's normal direction

(5) Correct the radar normal direction to parallel the lane direction according α

3.2. Traffic Incident Detection by Radar Output.
After lane alignment, the lane label of each vehicle can be obtained by the x value of the radar output, that is, the lateral position of the target in the radar local coordinate system as shown in Figure 3. Based on lane information and vehicle trajectory, incident detection can be achieved.

3.3. Trajectory-Based Incident Detection.
Trajectory-based incidents mainly include speeding, retrograde vehicle, occupation of emergency lanes, and illegal lane changes. The detection approaches of these incidents are implemented as follows:

(1) Speeding and retrograde vehicle. For a vehicle target, when the current velocity is greater than the lane speed limit, the target is considered to be over speed. When the current velocity direction of the vehicle target is opposite to the lane direction, the target is considered to be retrograde.

(2) Occupy the emergency lane. As shown in Figure 3, the emergency lane is on the left side of the road. When the x coordinate of a certain vehicle is in the emergency lane, it is considered that the vehicle occupies the emergency lane.

(3) Illegal lane change. In this incident detection, the lane change prohibited section is set up firstly. In the process of radar tracking, if the lane label of a target changes within the area where lane change is prohibited, it is considered that an illegal lane change

TABLE 2: Introduction to traffic incidents, application scenarios, and sensor detection capabilities.

| Category	Incident	Radar	Camera	Applicable roadway
Trajectory-based	Speeding	Excellent	Fair	City crossroads, highway
	Retrograde vehicle	Excellent	Excellent	City crossroads, highway
	Occupation of emergency lanes	Excellent	Excellent	Highway
	Illegal lane changes	Excellent	Excellent	City crossroads, highway
Inference-based	Lane queuing	Fair	Excellent	City crossroads
	Abnormal stop	Fair	Excellent	City crossroads, highway
Target classification-based	Intrusion in forbidden area	Fair	Excellent	City crossroads, highway
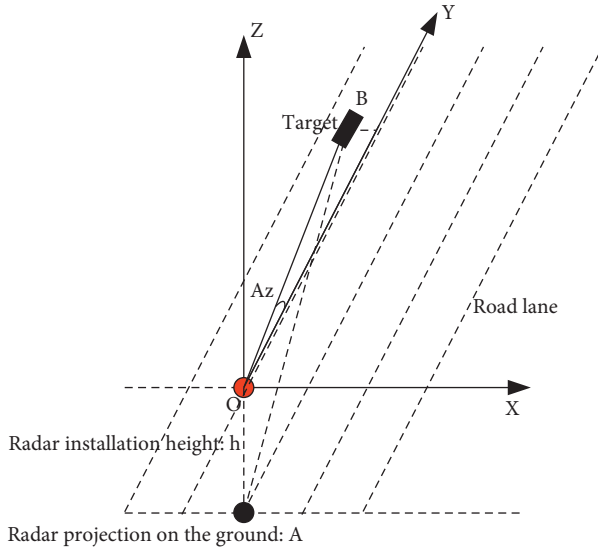
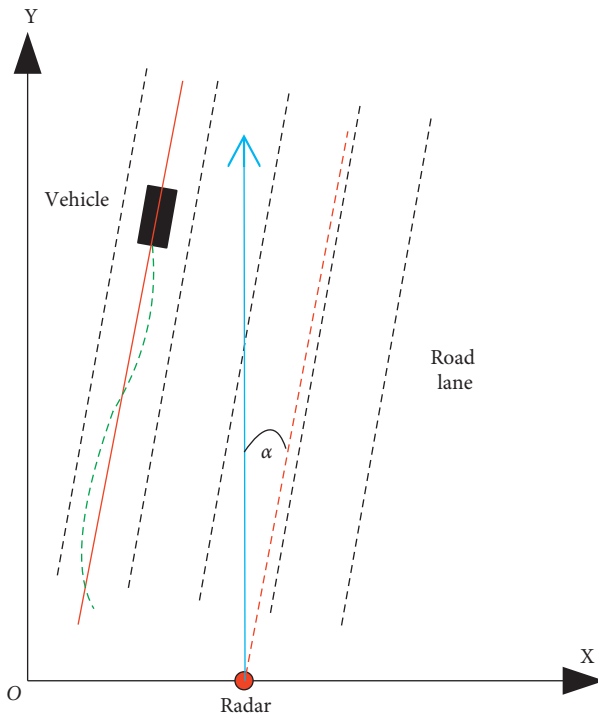FIGURE 3: Schematic diagram of radar installation position.

FIGURE 4: The schematic diagram of lane calibration.

incident has occurred, and the target ID, current coordinates, and current time are recorded.

3.4. Inference-Based Incident Detection. Due to the limitation of detection, i.e., radar cannot detect stationary targets effectively, radar can only detect incidents related to stationary targets indirectly through inference. Inference-based incidents include lane queuing and abnormal stop. The detection approaches of these incidents are implemented as follows.

(1) Vehicle stop detection. The detection of queuing and abnormal parking depends on the historical information of the target. Before incident detection, radar needs to judge whether the target changes from a moving state to a stopped state. To detect the stop state, radar tracking data is recorded for a certain period of time. When the target decelerates continuously over time and the target speed is below the stop threshold at a certain time, the target can be considered stopped.

(2) Lane queuing detection. During the tracking of vehicle targets in a lane, the radar updates the state of the vehicle in real time. When more than one vehicle stops behind the stop line in a lane and the spacing between the vehicles is below a distance threshold, it is considered that there was a queue phenomenon. Then the radar records the coordinates of the first/last car and calculates the current queue length from the position of the two cars. A schematic diagram of vehicles queuing is shown in Figure 5.

(3) Abnormal parking. For a certain vehicle, if the following conditions are met at the same time, it can be determined that abnormal parking occurs. The three conditions are as follows: (a) The target vehicle is stopped and the parking time exceeds a time threshold. (b) Space occupancy rate of the lane, where the parked vehicle is located, is less than a threshold. In this case, it indicates that the lane is not congested. (c) Other lanes in the same direction are not congested.

3.5. Target Classification-Based Incident Detection

(1) Intrusion into the forbidden area: It is generally considered the situation that pedestrians or non-motorized vehicles enter the highway lanes. The
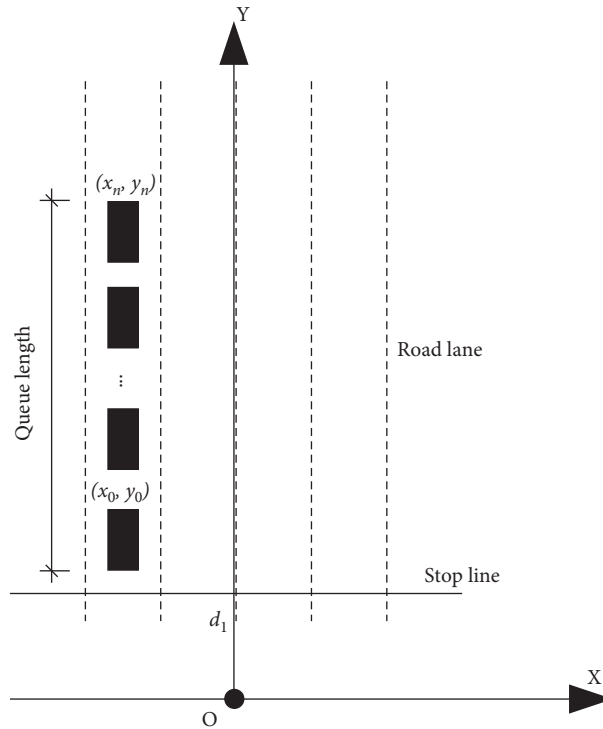
FIGURE 5: Schematic diagram of radar installation position.

TABLE 3: A comparison of camera and radar sensors [38].

Sensor characteristic	Camera	Radar
Range resolution	Fair	Excellent
Velocity detection	Limited	Excellent
Angle resolution	Excellent	Fair
Boundary	Excellent	Limited
Night operation	Fair	Excellent
Adverse weather	Limited	Excellent
Classification	Excellent	Fair
Stationary target detection	Excellent	Limited

performance of such incident detection depends on the target classification results. At present, the radar classification performance of traffic participants does not meet the demand [37].

4. Improved Radar Incident Detection Based on Vision Fusion

In the traffic perception, the characteristics of radar and camera are shown in Table 3. As mentioned in the previous section, the shortcomings of radar in incident detection include the following: (1) The radar cannot distinguish stationary targets from ground objects, so the detection of stationary targets is limited. When detecting incidents based on stationary targets, the effect will degrade. (2) The amount of information obtained by the radar is limited; thus the target classification is not accurate enough.

As a complementation, camera can make up for the shortcomings of radar [32, 38]. Due to rich details of targets, camera has advantages in target classification and stationary object detection. However, video-based traffic incident detection is sensitive to the environment. When the weather or light conditions are not good, the performance degrades seriously. In addition, the measurement of target distance and speed is also a weakness of video perception due to lack of depth information.

Overall, the camera and radar can form a good complement. Combining the two sensors, it is possible to obtain higher traffic incident detection performance.

The process of radar-camera fusion is shown in Figure 6, and the method details are described in the rest of this section.

4.1. Calibration of Coordinate between Radar and Camera.
In real scenario deployment, radar and camera are usually mounted in the same location. Therefore, the radar coordinate and the camera coordinate can be considered to coincide at the origin. The radar-camera coordinate system is shown in Figure 7. There are three coordinate systems; namely, radar coordinate system (RC) is XYZ, camera
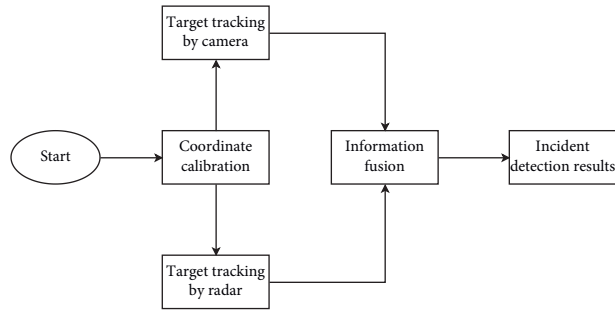
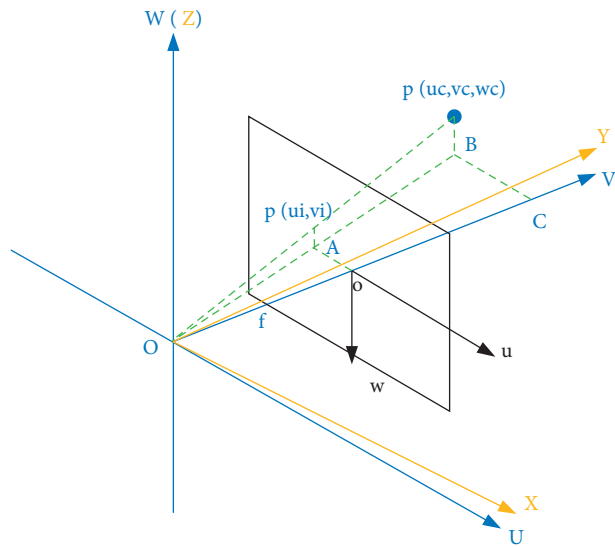FIGURE 6: Radar-camera incident detection system flow.
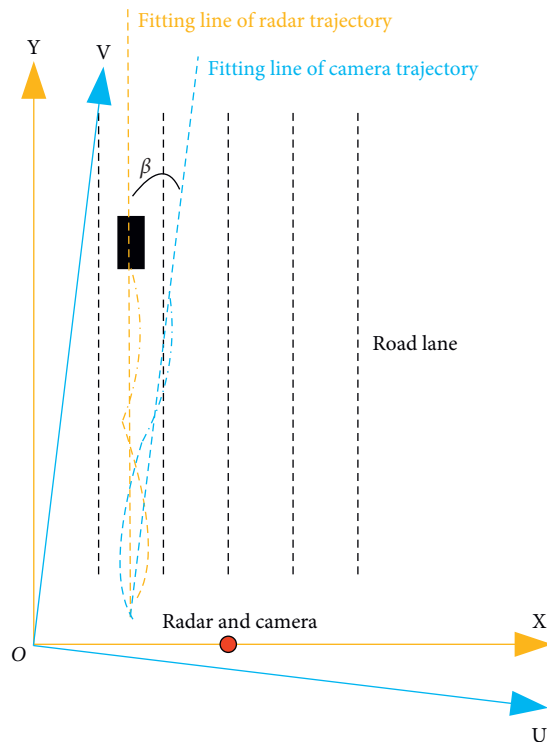

FIGURE 7: Radar-camera coordinate.

FIGURE 8: The schematic diagram of azimuth angle calibration between radar and camera.
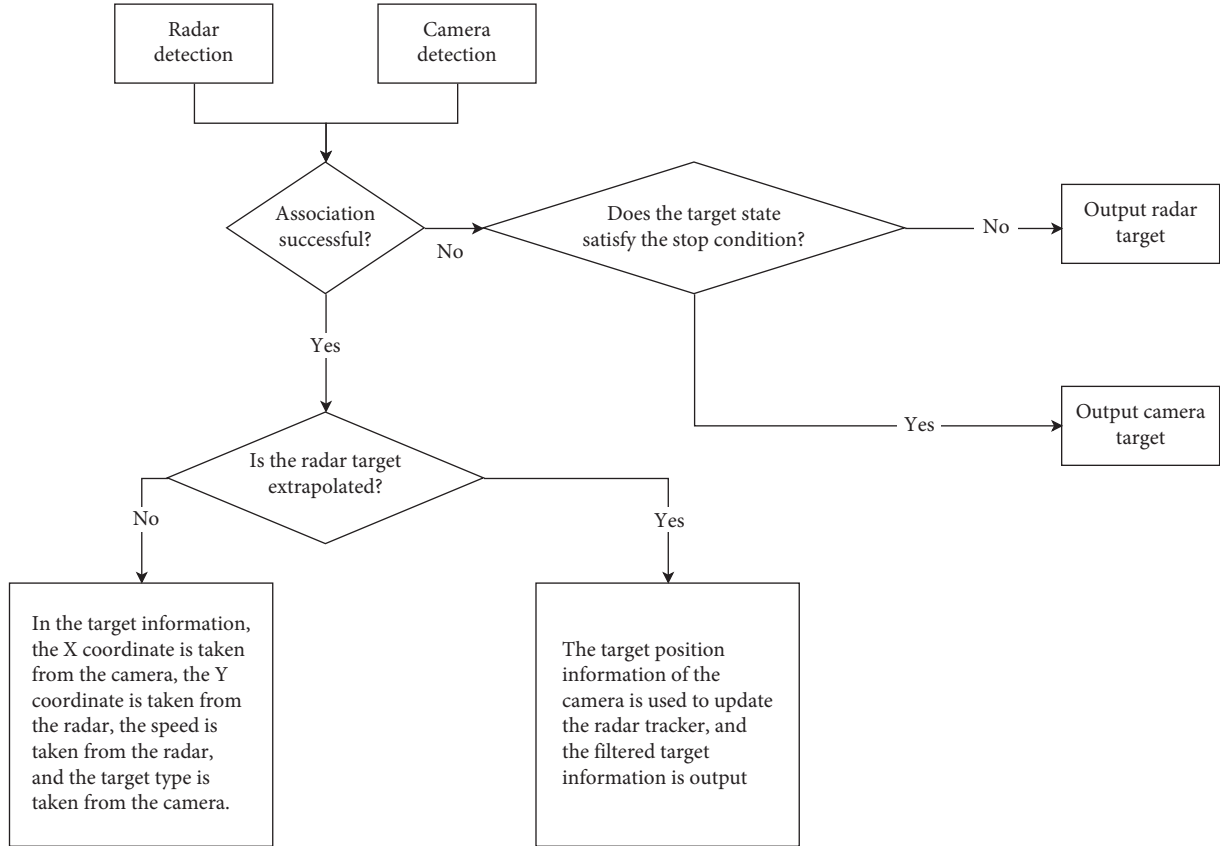
FIGURE 9: Radar-camera fusion flow.

coordinate system (CC) is UVW, and image coordinate system (IC) is uw. During installation, the horizontal error of the sensors can be compensated by the level meter; namely, the XOY and UOV planes are overlapped. As a result, the RC and the CC only deviate in the azimuth angle; i.e., there is a fixed angle between Y and V axis. When this angle error is calibrated, the radar and camera have the same coordinate to detect traffic targets.

Assuming there is a target at point P as shown in Figure 7, its coordinates is (u_c, v_c, w_c) in CC and is (u_i, w_i) in IC. According to the triangle relationship, u_c, w_c are

$$\begin{cases} u_c = v_c \dfrac{u_i}{f}, \\[2mm] w_c = v_c \dfrac{-w_i}{f}, \end{cases} \tag{5}$$

where f is the focal length of the camera. For traffic application, the height of the radar above the ground can be measured. Therefore, w_c is known in CC. The target position in IC can be determined according to the computer vision methods. For example, we use the YOLO algorithm to detect the vehicle target in the video, and then the target position (u_i, w_i) in IC is defined as the center position of the lower frame of the target bounding box. Thus the unknowns of the equation set in (5) are u_c and v_c. Through the above equation set in (5), the target position in CC can be solved. Rewrite (5) as follows; we can get u_c and v_c,

$$\begin{cases} u_c = u_i \dfrac{w_c}{w_i}, \\[2mm] v_c = -f \dfrac{w_c}{w_i}. \end{cases} \tag{6}$$

In order to calibrate the azimuth angle between the radar and the camera, the steps are implemented as follows:

(1) Set a testing vehicle traveling in a straight line, the tracking results of radar and camera are recorded at the same time as shown by the yellow and blue point for radar and camera, respectively, in Figure 8

(2) Map the camera tracking trajectory to the radar local coordinate system according to (6)

(3) straight line fitting to the trajectory of radar and camera, respectively, as shown by the yellow and blue dashed line in Figure 8

(4) Calculate the angle β between the two fitted lines

(5) Compensate the output target position of the camera by β to obtain the spatially aligned data to radar

4.2. Traffic Incident Detection by Radar-Camera Fusion. After calibration, the CC and the RC system coincide, and the detection results of the radar and the camera are spatial alignments. In a unified coordinate system, the fusion

Figure 10: Crossroad scenario in day time.

process of radar and camera is implemented as shown in Figure 9, and the details are described as follows.

(1) Record the track list of radar targets. The target track results of radar are obtained by the Kalman filter as described in Section 2, and the output data format is shown in Table 1.

(2) Record the track list of camera targets. The target track results of camera are realized by the SORT algorithm which uses a rudimentary combination of the Kalman filter and Hungarian algorithm [39, 40]. This tracking process is very similar to that of radar and the performance is good enough in general traffic scenario, in which vehicles and pedestrians are usually in regular motion. The output of the target track includes the horizontal and vertical pixel location of the center of the target, the scale (area) and the aspect ratio of the target's bounding box in IC, and the target category. Correspondingly the target position in unified coordinate can be obtained by (5) and (6).

(3) Associate and match the radar target position with the camera target position. Due to the difference between the track results of the two sensors, target association is required. Hungarian algorithm is used to match the camera target to the radar target [41].

(4) According to the assignment results, the target information of the radar and the camera is fused according to the process shown in Figure 9.

When the fusion information is obtained, the incident detection can be achieved by the method proposed in Section 3. Unlike the single radar detection, the target information at this time already includes the camera information, such as the location of the static target and the target category label. For the three types of incident detection mentioned in Section 3, the benefits of sensor fusion are as follows. (1) Trajectory-based incident detection: After the fusion of target position from camera and radar, the target track will be more robust, and the problem of target trajectory interruption is reduced. As a result, the reliability of incident detection is improved. (2) Inference-based incident detection: Due to camera's excellent detection performance for stationary targets, the information fusion can greatly improve the performance of inference-based detection; thus better accuracy for queued and parking incidents detection is obtained. (3) Target classification-based incident detection: Intrusion targets can easily be identified now due to
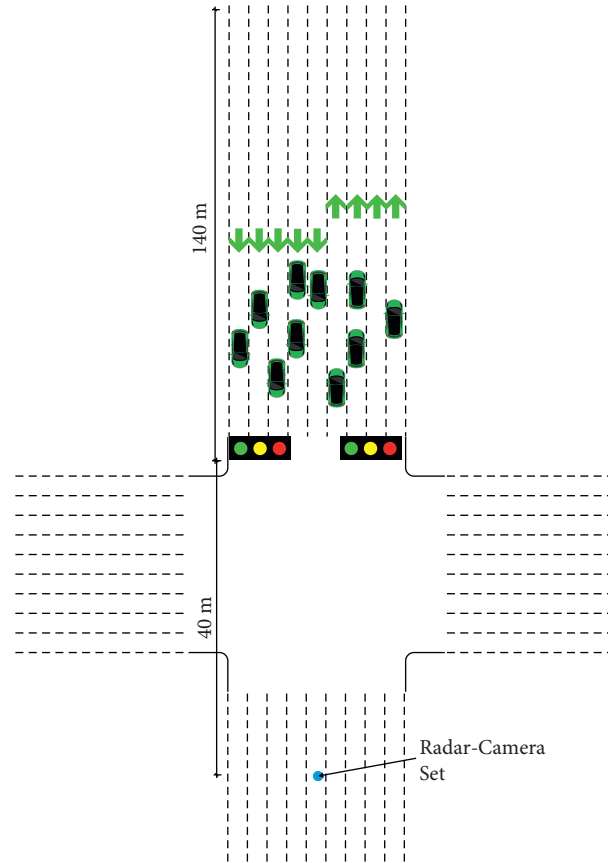
140 m

40 m

Radar-Camera
Set

FIGURE 11: Crossroad typology.

TABLE 4: Incident detection results in crossroad scenario.

Category	Incident	Radar	Camera	Radar-camera fusion	Ground truth	RAR of radar	RAR of camera	RAR of fusion
Trajectory-based	Speeding	370	340	370	378	97.88	89.95	97.88
	Retrograde vehicle	399	399	399	399	100	100	100
	Occupation of emergency lanes	351	369	369	370	94.86	99.73	99.73
	Illegal lane changes	343	350	352	370	92.70	94.59	95.13
Inference-based	Lane queuing	237	380	380	385	61.56	98.70	98.70
	Abnormal stop	210	375	375	380	55.26	98.68	98.68
Target classification-based	Intrusion in forbidden area	531	374	374	380	60.26	98.42	98.42

accurate classification performance of camera, and the incident detection performance is improved.

5. Experimental Results

The real scenario experiment is used to verify the traffic incident detection performance proposed in this paper. Two scenarios are selected for incident detection validation. The first scenario is a daytime crossroad, in which there is a queue of vehicles due to traffic lights. Incidents related to urban traffic are verified in this scenario. The second scenario is chosen at night, on a long road. It mainly verifies the incident detection performance over long distances under low light conditions, which simulated the incident detection

scenario of highway. In each scenario, a radar and a camera installed at the same location are used for traffic data collection.

The relative accuracy rate (RAR) is used to evaluate the performance of incident detection. In this experiment, for each type of incident, more than 300 incidents are recorded, and the RAR is calculated by the following formula:

$$RAR = \left(1 - \frac{|n - N|}{N}\right) \times 100\%, \tag{7}$$

where n is the number of detected incidents, and N is the number of incidents that actually occurred, i.e., the ground truth. As can be seen from (7), the RAR is a description of
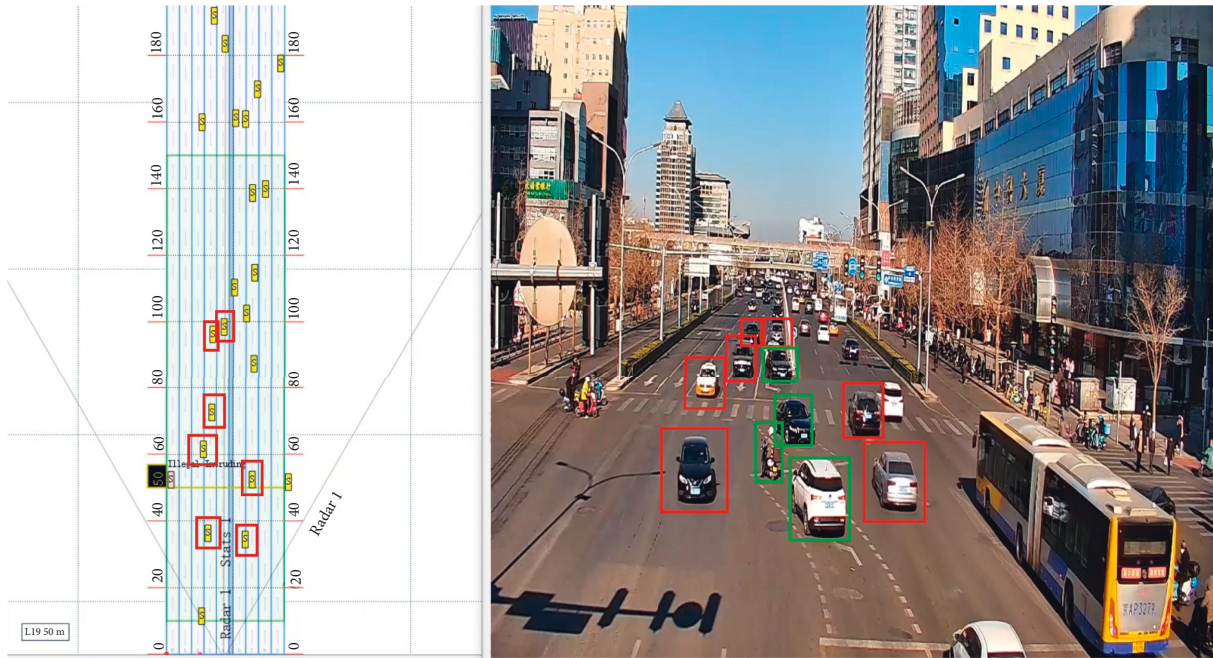
FIGURE 12: Tracking result of radar.

FIGURE 13: Tracking result of camera. Incident detection performance for long road scenario in night.

how similar the estimated value is to the true value, and it is equal to 1 minus the relative error. The relative error is the ratio of the absolute error to the true value, describing the difference between the estimated value and the true value. When the relative error is equal to 0, it means that the estimated value is equal to the true value, and the RAR is 100% in this case, which indicates that the system detects all incidents that occurred. Therefore, through RAR, we can quantitatively obtain the accuracy for traffic incident detection. The closer the RAR is to 100%, the higher the accuracy of event detection is.

5.1. Incident Detection Performance in Real Crossroad Scenario.

The test scenario is selected at a real crossroad as shown in Figure 10 and the corresponding road topology is shown in Figure 11. There are a total of nine lanes in the scenario, of which five are coming direction and four are going directions. We chose one direction of the intersection as our test scenario. Along the radar line of sight, it is 140 m from the stop line to the farthest end of the road that can be detected, which is a typical urban intersection. Traffic signal lights are installed at the intersection. When the red light is on, cars will queue up from the stop line. Besides vehicle

Figure 14: Long road scenario in night.

targets, there are other traffic targets such as motorcycles, bicycles, and pedestrians. In this experiment, radar and camera are installed on the overbridge facing the test road and deployed at the same location as shown in Figure 10.

Total seven traffic incidents are evaluated, which include speeding, retrograde vehicle, emergency lane occupation, illegal lane changes, queuing, abnormal stop, and illegal intrusion.

For speeding, retrograde vehicle, emergency lane occupation, and illegal lane changes, which are judged by target trajectory, we manually configure certain thresholds to make the targets meet the incident detection conditions to count the detection accuracy. For example, we set the speed threshold to 20 km/h; then the most of the vehicles on road can be considered above the threshold. For retrograde, we set the direction of one lane to be opposite to the real direction, so that all the targets in this lane meet the retrograde condition. For the emergency lane occupation, we set a certain lane as the emergency lane, so that the targets in this lane meet occupied condition. Similarly, for illegal lane changes, setting a certain lane as the attribute that does not

allow lane change, then the lane-changing vehicle in that lane can be regarded as the illegal lane change target.

For inferred-based incidents such as queuing and abnormal stops, we use real incidents to test. Every time the traffic light turns to red, there may be car queuing on the corresponding road and this queuing incident is used to evaluate the detection performance. Unlike queuing incidents, the abnormal stops are not common on actual roads, so we use the bus station on the right side of the scenario shown in Figure 10 to test abnormal stops. We set a short time threshold of abnormal stops, so that the detection condition can be met when the bus stops.

For illegal intrusion incident based on target classification, we use pedestrians on the road for testing. The sidewalk area at the intersection is set as a forbidden area, and the pedestrians crossing the street meet the incident detection condition.

We counted the incident detection results of radar only and camera only, respectively, and the incident detection results of radar-camera fusion. The results are shown in Table 4.
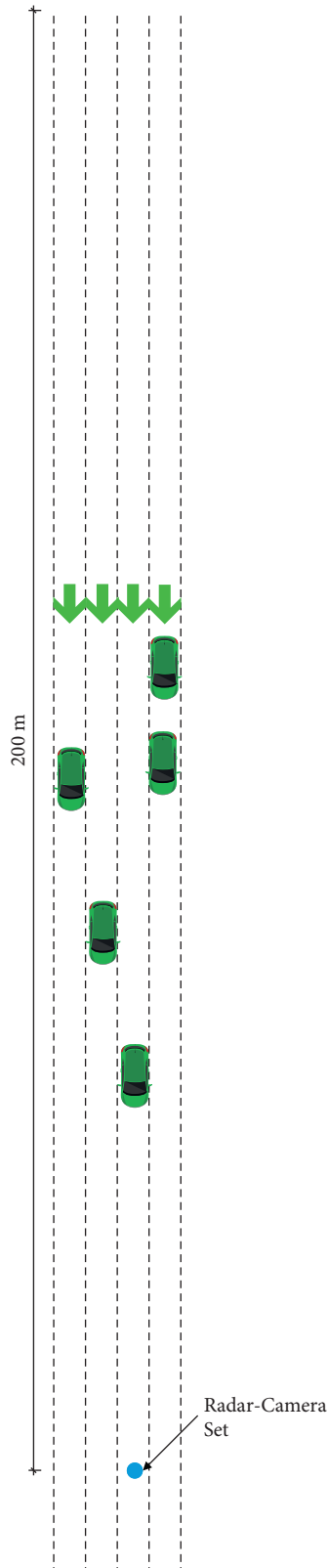
FIGURE 15: Long straight road typology.

Since radar has better tracking accuracy for targets, incident detection based on trajectory has achieved better accuracy when using radar only. After fusion with video, the accuracy of incident detection is slightly improved compared to single radar. This is due to the limited information increment that video brings in terms of target trajectory. It should be noticed that the accuracy of speeding detected by a single camera is not high, since the camera has low accuracy of target velocity measurement due to the lack of depth information.

For the inference-based methods, the incident detection performance of the fusion system is more greatly improved than that of radar only in queuing and parking, since the camera has a good detection effect on stationary targets. The tracking results of radar and camera on the targets are shown in Figures 12 and 13, respectively. We manually marked the tracking results of the radar with rectangular boxes of different colours as shown in Figure 12. It can be seen that the radar has a better tracking effect for moving targets (red boxes) on the lane. But for stationary targets (green boxes), the radar did not output the corresponding trajectory. In addition, it can be seen that the camera has a good tracking effect on the targets as shown in Figure 13, including the parking targets to be turned at the intersection.

As to target classification-based detection, such as pedestrian intrusion, due to the limited amount of information obtained by the radar, the performance of target classification is poor, and the false alarm rate for pedestrian intrusion incident is high, i.e., some left-turn vehicles with lower radial speed were identified as pedestrians by radar. After fusion with the camera, the accuracy of pedestrian intrusion is greatly improved. The same result can be seen from Figure 13. Due to the advantages of video in the information of target details, the traffic target category can be well recognized. Pedestrian and nonmotor vehicle targets are marked as red boxes by the YOLO algorithm, while vehicles are marked as green.

In order to simulate the highway scenario, a long road without traffic lights is selected. In addition, considering verifying the system's working ability throughout the day, the experiment is tested at night. In this scenario, the farthest point of the road is 200 m, and the location about 180 m is used for performance evaluation, which is shown by the red line in Figure 14. The topological structure of the long straight road is shown in Figure 15. There are a total of four lanes with a coming direction. The system is installed on an overpass, facing the middle of the road. Since there are no traffic lights and no pedestrians on this road, we do not evaluate the inference-based and classification-based incident detection performance in this experiment.

The experiment results are shown in Table 5. Because of the poor lighting conditions, the backlight effect caused by car light, and the reduction of pixels of long-distance targets, the target detection performance of camera degrades

TABLE 5: Incident detection results in night long road scenario.

Category	Incident	Radar	Camera	Radar-camera fusion	Ground truth	RAR of radar	RAR of camera	RAR of fusion
Trajectory-based	Speeding	310	172	310	330	93.94	52.12	93.94
	Retrograde vehicle	317	270	317	317	100	100	100
	Occupation of emergency lanes	324	281	326	357	90.76	78.71	91.32
	Illegal lane changes	321	263	322	355	90.42	74.08	90.70

significantly, which further affects the incident detection performance. In this case, the performance of the fusion system is approximately equal to that of a single radar, that is, the camera does not bring more performance improvement.

6. Conclusions

In this paper, a traffic incident detection method based on radar-camera fusion is proposed to improve the reliability of traffic incident detection in urban intersection and highway scenarios. To the shortcomings of radar perception, the detection accuracy of lane queuing, abnormal stop, and intrusion in forbidden area are improved by the fusion system, since the lack of radar's inability to detect stationary targets and weak target classification capabilities are made up. In the real scenario experiment, the RAR of lane queuing detection was increased from 61% of single radar perception to 98%, the RAR of abnormal stop detection was increased from 55% of single radar perception to 98%, and the RAR of intrusion in forbidden area detection was increased from 60% of single radar perception to 98%. To the shortcomings of camera perception, the fusion system improves the speed measurement accuracy of the target, which makes up for the disadvantage that the camera cannot obtain depth information. Meanwhile, the ability to work all day is obtained by the fusion process compared to that of single camera, which makes up for the problem of camera performance degradation at night. In the real scenario experiment, the RAR of speeding detection is increased from 52% of single camera to 94%. The fusion incident detection system combines the advantages of radar and camera, which can provide more stable and reliable perception data and improve the safety of the intelligent transportation system.

The experimental results based on the real scenario show the performance improvement of the fusion system in traffic incident detection compared with the single sensor approach. However, in the scenario with large traffic flow and dense target, there will be target matching error during the fusion process. In response to this, we will study the accuracy of target association in the future. In addition, the process proposed in this paper is the fusion of two sensors at the target trajectory level. In the future, we will explore the process of fusing information at more primitive levels, such as at sensor detections level or at digital signal level. This allows more useful information about the target to be preserved and fused before tracking processing, so as to further improve the performance of incident detection.

Data Availability

The data used to support the findings in this study are available from the corresponding authors upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] Z. Cakici and Y. S. Murat, "A differential evolution algorithm-based traffic control model for signalized intersections," *Advances in Civil Engineering*, vol. 2019, Article ID 7360939, 16 pages, 2019.

[2] Y. He, Z. Liu, X. Zhou, and B. Zhong, "Analysis of urban traffic accidents features and correlation with traffic congestion in large-scale construction district," in *Proceedings of the 2017 International Conference on Smart Grid and Electrical Automation (ICSGEA)*, pp. 641–644, Changsha, China, May 2017.

[3] R. Arnott and E. Inci, "An integrated model of downtown parking and traffic congestion," *Journal of Urban Economics*, vol. 60, no. 3, pp. 418–442, 2006.

[4] J. Golias, G. Yannis, and C. Antoniou, "Classification of driver-assistance systems according to their impact on road safety and traffic efficiency," *Transport Reviews*, vol. 22, no. 2, pp. 179–196, 2002.

[5] M. Shengdong, X. Zhengxian, and T. Yixiang, "Intelligent traffic control system based on cloud computing and big data mining," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 12, pp. 6583–6592, Dec. 2019.

[6] D. Zhao, Y. Dai, and Z. Zhang, "Computational intelligence in urban traffic signal control: a survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, pp. 485–494, 2011.

[7] Y. Yang, K. He, Y.-p. Wang, Z.-z. Yuan, Y.-h. Yin, and M.-z. Guo, "Identification of dynamic traffic crash risk for cross-area freeways based on statistical and machine learning methods," *Physica A: Statistical Mechanics and Its Applications*, vol. 595, p. 127083, 2022.

[8] Y. Yang, K. Wang, Z. Yuan, and D. Liu, "Predicting freeway traffic crash severity using XGBoost-Bayesian network model with consideration of features interaction," *Journal of Advanced Transportation*, vol. 19, p. 4257865, 2022.

[9] Y. Yongchang Ma, M. Chowdhury, A. Sadek, and M. Jeihani, "Real-time highway traffic condition assessment framework using vehicle-infrastructure integration (VII) with artificial intelligence (AI)," *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, no. 4, pp. 615–627, 2009.

[10] X. Dai, D. Liu, L. Yang, and Y. Liu, "Research on headlight technology of night vehicle intelligent detection based on

hough transform," in *Proceedings of the 2019 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, pp. 49–52, Changsha, China, January 2019.

[11] V. Mandal, A. R. Mussah, P. Jin, and Y. Adu-Gyamfi, "Artificial intelligence-enabled traffic monitoring system," *Sustainability*, vol. 12, no. 21, p. 9177, 2020.

[12] M. Akhtar and S. Moridpour, "A review of traffic congestion prediction using artificial intelligence," *Journal of Advanced Transportation*, vol. 2021, Article ID 8878011, 18 pages, 2021.

[13] S. R. E. Datondji, Y. Dupuis, P. Subirats, and P. Vasseur, "A survey of vision-based traffic monitoring of road intersections," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 10, pp. 2681–2698, 2016.

[14] R. Weil, J. Wootton, and A. Garcia-Ortiz, "Traffic incident detection: sensors and algorithms," *Mathematical and Computer Modelling*, vol. 27, pp. 257–291, 1998.

[15] J. Xiao and Y. Liu, "Traffic incident detection using multiple-kernel support vector machine," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2324, no. 1, pp. 44–52, 2012.

[16] Z. Marszalek, W. Gawedzki, and K. Duda, "A reliable moving vehicle axle-to-axle distance measurement system based on multi-frequency impedance measurement of a slim inductive-loop sensor," *Measurement*, vol. 169, p. 108525, 2021.

[17] Q. Li, H. Cheng, Y. Zhou, and G. Huo, "Road vehicle monitoring system based on intelligent visual internet of things," *Journal of Sensors*, vol. 2015, Article ID 720308, 16 pages, 2015.

[18] U. P. Naik, V. Rajesh, and R. Kumar, "Implementation of YOLOv4 algorithm for multiple object detection in image and video dataset using deep learning and artificial intelligence for urban traffic video surveillance application," in *Proceedings of the 2021 Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, September 2021.

[19] R. Ke, Y. Zhuang, Z. Pu, and Y. H. Wang, "A smart, efficient, and reliable parking surveillance system with edge artificial intelligence on IoT devices," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, 2020.

[20] S. Sivaraman and M. M. Trivedi, "A general active-learning framework for on-road vehicle recognition and tracking," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 2, pp. 267–276, Jun. 2010.

[21] S. S. Teoh and T. Bräunl, "Symmetry-based monocular vehicle detection system," *Machine Vision and Applications*, vol. 23, no. 5, pp. 831–842, Sep. 2012.

[22] H. Zhu, K.-V. Yuen, L. Mihaylova, and H. Leung, "Overview of environment perception for intelligent vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 10, pp. 2584–2601, Oct. 2017.

[23] A. Mukhtar, L. Xia, and T. B. Tang, "Vehicle detection techniques for collision avoidance systems: a review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 5, pp. 2318–2338, Oct. 2015.

[24] M. Rezaei, M. Terauchi, and R. Klette, "Robust vehicle detection and distance estimation under challenging lighting conditions," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 5, pp. 2723–2743, Oct. 2015.

[25] L.-C. Liu, C.-Y. Fang, and S.-W. Chen, "A novel distance estimation method leading a forward collision avoidance assist system for vehicles on highways," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 4, pp. 937–949, Apr. 2017.

[26] A. Joglekar, D. Joshi, R. Khemani, S. Nair, and S. Sahare, "Depth estimation using monocular camera," *International Journal of Computer Science and Information Technology*, vol. 2, no. 4, pp. 1758–1763, 2011.

[27] S. Lessmann, M. Meuter, D. Muller, and J. Pauli, "Probabilistic distance estimation for vehicle tracking application in monocular vision," in *Proceedings of the IEEE Intell. Vehicles Symp*, no. IV, pp. 1199–1204, Gothenburg, Sweden, June 2016.

[28] S. Han, X. Wang, L. Xu, H. Sun, and N. Zheng, "Frontal object perception for intelligent vehicles based on radar and camera fusion," in *Proceedings of the 35th Chin. Control Conf. (CCC)*, pp. 4003–4008, Chengdu, China, July 2016.

[29] D. Gao, J. Duan, X. Yang, and B. Zheng, "A method of spatial calibration for camera and radar," in *Proceedings of the 8th World Congr. Intell. Control Autom*, pp. 6211–6215, Jinan, July 2010.

[30] Y. Feng, S. Pickering, E. Chappell, P. iravani PhD, and C. Brace, "Distance estimation by fusing radar and monocular camera with k filter," *SAE Technical Paper Series*, 2017.

[31] M. Nishigaki, S. Rebhan, and N. Einecke, "Vision-based lateral position improvement of RADAR detections," in *Proceedings of the 15th Int. IEEE Conf. Intell. Transp. Syst*, pp. 90–97, Anchorage, AK, USA, September 2012.

[32] Yu. Du, Ka L. Man, and E. G. Lim, "Image radar-based traffic surveillance system: an all-weather sensor as intelligent transportation infrastructure component," in *Proceedings of the 2020 International SoC Design Conference (ISOCC)*, October 2020.

[33] S. M. Patole, M. Torlak, D. Wang, and M. Ali, "Automotive radars: a review of signal processing techniques," *IEEE Signal Processing Magazine*, vol. 34, no. 2, pp. 22–35, 2017.

[34] H. Liu, N. Li, D. Guan, and L. Rai, "Data feature analysis of non-scanning multi target millimeter-wave radar in traffic flow detection applications," *Sensors*, vol. 18, no. 9, p. 2756, 2018.

[35] H. Rohling, "Radar CFAR thresholding in clutter and multiple target situations," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-19, no. 4, pp. 608–621, 1983.

[36] K. V. Ramachandra, *Kalman Filtering Techniques for Radar Tracking*, CRC Press, 2018.

[37] S. Heuel and H. Rohling, "Pedestrian classification in automotive radar systems," in *Proceedings of the 2012 13th International Radar Symposium*, pp. 39–44, Warsaw, Poland, May 2012.

[38] S. Alland, W. Stark, M. Ali, and M. Hegde, "Interference in automotive radar systems: characteristics, mitigation techniques, and current and future research," *IEEE Signal Processing Magazine*, vol. 36, no. 5, pp. 45–59, Sept. 2019.

[39] A. Bewley, Z. Ge, L. Ott, F. Ramos, and B. Upcroft, "Simple Online and Realtime Tracking," *Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP)*, vol. 64, pp. 3464–3468.

[40] N. Wojke, A. Bewley, and D. Paulus, "Simple online and realtime tracking with a deep association metric," in *Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP)*, pp. 3645–3649, Beijing, China, September 2017.

[41] H. W. Kuhn, "The Hungarian method for the assignment problem," *Naval Research Logistics Quarterly*, vol. 2, no. 1-2, pp. 83–97, 1955.