Retraction

Retracted: Multiview Fusion 3D Target Information Perception Model in Nighttime Unmanned Intelligent Vehicles

Journal of Function Spaces

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] J. Xu and Z. Liang, "Multiview Fusion 3D Target Information Perception Model in Nighttime Unmanned Intelligent Vehicles," *Journal of Function Spaces*, vol. 2022, Article ID 9295395, 10 pages, 2022.

Research Article

Multiview Fusion 3D Target Information Perception Model in Nighttime Unmanned Intelligent Vehicles

Jijin Xu and **Zhen Liang**

Fujian Key Laboratory of Agricultural Information Sensoring Technology, College of Mechanical and Electrical Engineering, Fujian Agriculture and Forestry University, Fuzhou 350100, China

Correspondence should be addressed to Jijin Xu; michaelxu@fafu.edu.cn

Unmanned technology is an important development project of today's cutting-edge science and technology, which has a significant impact on social and economic development, national defense construction, and scientific and technological development. The rapid development of industrial information technology has driven the unmanned intelligent vehicle system to innovate and gradually enter the public's view, and at the same time, the driving safety of unmanned intelligent vehicles is also widely concerned. Target information perception system is the foundation of unmanned system and the fundamental guarantee of safety and intelligence of unmanned vehicles. There are three key problems of target recognition in unmanned driving, namely, target classification, localization, and attitude determination. In the implementation of a networked virtual environment system, a flexible and complete perception model is needed as the guiding model of the system. Since 3D point cloud data can provide more spatial information than 2D RGB image data, it is more beneficial to determine the target category, position, and pose in 3D. In this paper, based on the existing research of unmanned intelligent vehicle perception system, we combine the application of fusion of 3D target information perception model and develop a nighttime unmanned system based on multiview fusion of 3D target information perception model. This system can simultaneously perform the detection of multiple categories of objects and predict the center point, length, width, height, and orientation of the objects, so that the unmanned car can sense the location of the surrounding objects when driving in the actual scene.

1. Introduction

As a class of intelligent robots, the development of unmanned vehicles has a significant impact on social and economic development, national defense construction, and science and technology, and has become a strategic target for research in high-tech fields in various countries around the world [1]. Unmanned driving technology includes key technologies such as target information perception, path planning, vehicle control, and intelligent decision-making [2]. The multiview channels used in unmanned intelligent driving vehicles are visible light cameras, but ordinary visible light cameras cannot accurately obtain specific information in road scenes at night in dim and low-light environments

[3]. In the field of 3D spatial perception research, it is an important capability for robots to acquire 3D information through relevant sensors and to be able to understand the scene and interact with the surrounding environment through this data [4]. Among them, target information perception technology is the underlying module of unmanned technology and the basic guarantee for safe vehicle driving [5]. Through unmanned driving technology based on 3D target information perception model with multiview fusion, we can effectively control many unstable human factors, such as drunk driving and fatigue driving [6].

Unmanned vehicles traveling on the road at night need to sense the surrounding environment of vehicles, pedestrians, etc., and although the infrared camera-based

perception scheme can sense the obstacles in front of the vehicle, it cannot accurately obtain the location information of the obstacles and has limited accuracy [7]. Target information perception for unmanned vehicles is a unique machine-like human-like understanding of the vehicle's surroundings through some multiview channels, such as LIDAR and cameras [8]. 2D target detection is unable to obtain real-world target parameters such as physical dimensions and 3D coordinates due to the use of image information only and to identify target classes and locate targets in the image plane [9]. The basic requirement for 3D target recognition in a typical unmanned scenario is accurate classification, localization, and pose recognition of the target to be detected [10]. According to the differences in the acquired data, the existing 3D target detection algorithms can be divided into two categories: visual and laser point cloud.

Target detection and localization in three-dimensional space is an important basis for people to perceive the environment and has important applications in robotics, autonomous driving, security monitoring, industrial production, intelligent agriculture, and many other fields [11]. And because such sensors as cameras are vulnerable to external factors such as light and weather, and it is too difficult to infer three-dimensional information of the target from two-dimensional images, these algorithms are difficult to achieve satisfactory results in obtaining three-dimensional detection frames of the target [12]. In most of the night vision related research generally uses infrared camera instead of visible camera, the imaging principle of infrared camera is different from visible camera, which is able to transform the temperature distribution of the object in the surrounding scene into an image by detecting the object's own infrared radiation, and then through photoelectric transformation [13]. For safety reasons, autonomous driving is a scenario that requires extra precision as well as stability. In order to improve the performance of the algorithm, most of the current 3D target detection algorithms in autonomous driving scenarios use LiDAR as the main sensor. The chance of traffic accidents is much higher at night than during daytime; therefore, the 3D target information perception model using multiview fusion is important to study how unmanned intelligent vehicles can better perceive the surrounding environment information in the night scenario.

In this paper, a 3D target detection algorithm is proposed by fusing 3D target perception with LIDAR point cloud. The algorithm effectively improves the target detection accuracy of the unmanned intelligent vehicle in the night scene and makes the unmanned intelligent vehicle have better night target information perception ability. The innovative contribution lies in the research of driving behavior planning based on multiview fusion of three-dimensional target information perception model. It enables the driverless intelligent vehicle to make correct driving behavior planning in real time after acquiring the surrounding environment information at night. The basic data format of 3D target recognition algorithm is point cloud, which is relatively stable and not affected by illumination. It can truly reflect the position of objects in the scene and is very suitable for target recognition.

2. Application of Three-Dimensional Target Information Perception Model in Unmanned Intelligent Vehicle at Night

2.1. Construction Method of Three-Dimensional Target Perception Model Based on Multiview Fusion. The 3D target detection model with multiview fusion consists of a feature extraction module, a candidate region generation module, and a channel fusion module. Unmanned vehicle is a comprehensive technology, and if we look at the functions of its parts, we can divide it into 3 parts: target information sensing, decision planning, and vehicle control [14]. However, the point cloud data is very sparse and uneven compared with image data and the data volume is large, which makes the data application and storage calculation face new problems. The 3D target recognition algorithm focuses on the design of algorithms for the above two problems, including various forms of reprocessing of the point cloud and clever design of the network structure to understand the geometric information in the point cloud [15]. The prediction-tracking relationship between the front and back frames of lane line detection is then established to further improve the accuracy and efficiency of lane line detection. The function of general tracking is to predict the location of road features in the next frame image, detect road features in a small range, and improve efficiency. If no road feature is detected within the prediction range, the position of the estimated or previous frame feature is used. If no road feature is detected for several consecutive frames, the full image road feature detection is started. The flow chart of the entire lane line detection algorithm is shown in Figure 1.

First, the feature extraction module acts on the encoded bird's eye view form and the front view form of the infrared image data and the LiDAR point cloud data, respectively. Each LiDAR scan is processed independently instead of creating a two-dimensional point cloud map. This reduces the complexity of the algorithm by analyzing only a single dimension of laser data, thus reducing the complexity of the algorithm. The target information sensing system acquires real-time and accurate information about the road environment through a multiview channel device. The threshold value T of the method is calculated by averaging E, squared difference between pixels P, and root mean square value Q between pixels for this window together, and the parametric equation for performing the threshold calculation is as follows:

$$T = a * E + b * p + c * Q. \tag{1}$$

The multiview channel data fusion discovers the existing and possible risk factors in the environment and reports them to the decision-making module in time. In the urban road environment, the main road types are tarmac, concrete, grass, and gravel roads, and the vehicles will have different characteristics when driving on different road types due to different friction coefficients and side
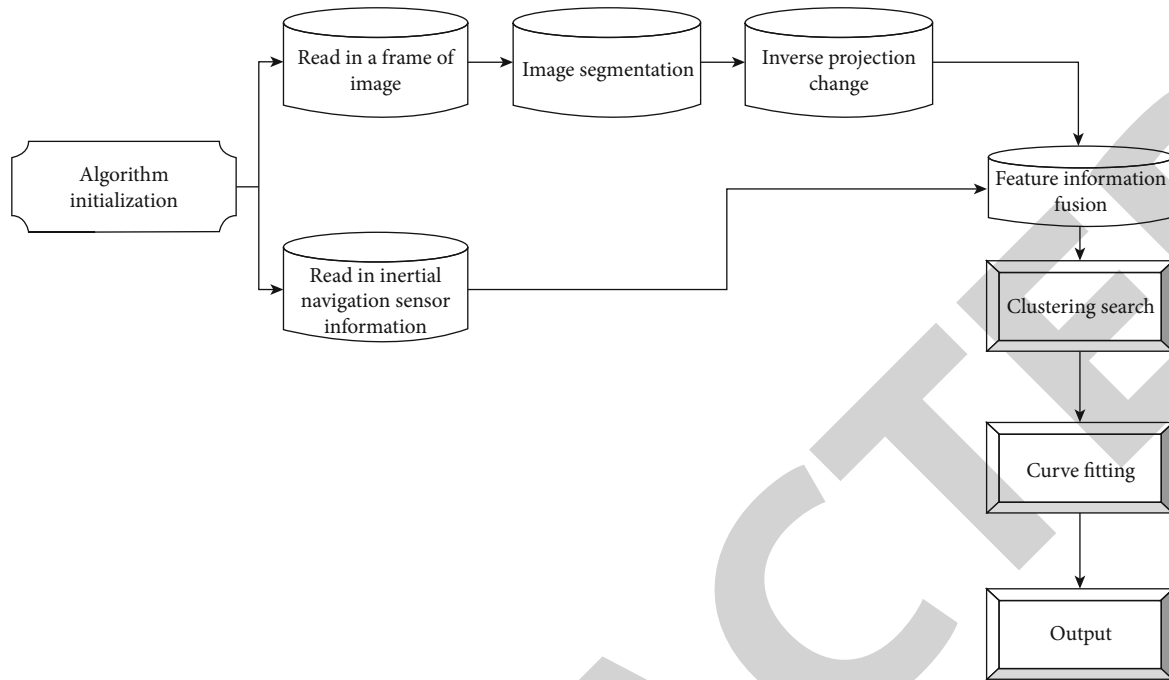
Figure 1: Flow chart of lane detection algorithm.

slip angles. The equations to calculate the position translation based on the unmanned vehicle speed and angular velocity are as follows:

$$E = B + V\Delta t \cos (w\Delta t), \qquad (2)$$

E–end position
B–starting position
V–speed of a motor vehicle
w–angular velocity
Δt–camera sampling period

Therefore, for the linear discrete system, it is required to satisfy the process model as a linear system with superimposed process excitation noise, the measurement model as a linear system with superimposed measurement noise, and the noise both obey normal distribution. Using the end-to-end convolutional neural network, the image obtained from the color camera is used as the input, and the steering wheel angle signal is output directly after the network decision, and the model schematic is shown in Figure 2.

Second, the candidate region generation module is highly effective in 3D target detection algorithms. The network draws on the ideas of the 3D target detection algorithm and uses a network instead of manual extraction to generate candidate regions of objects, using a bird's eye view form of LiDAR data as the input to the candidate region generation module. The input 3D image is given to the 3D detector for target detection. Using preprocessed dense/sparse labels, the nonroad labels in the sparse part are moved from the sparse points to the dense points closest to the vehicle perimeter. If a stochastic process is a Markov decision process, then for that stochastic process there must be a strategy that yields

better results than the other strategies in all cases, which is also called the optimal strategy, then.

$$\pi^*(s) = \arg \max Q(s, a), \qquad (3)$$

s–time
$\pi^*(s)$–optimal strategy
a–action

Then, the optimal value return function can be obtained:

$$Q^*(s, a) = \sum_{x_* \in s} T(s, a)(R(s, a) + \gamma V^*(s^*)), \qquad (4)$$

γ–influence factor

LIDAR emits tens or hundreds of thousands of laser pulses per second to the environment, and by measuring the return time of the pulses, information such as target distance and orientation can be calculated. Therefore, the perception model design of the network virtual environment must take into account the user requirements and the environment of the system, to be able to dynamically adjust the display and network transmission data according to the implementation conditions and to maintain the user's sense of reality. Instead of sticking to the optimization of the global model, the approach we have chosen is a transition from the local optimum to the global optimum. It can be used to describe, understand, analyze, and evaluate people's perception of information security, and can quantitatively analyze and compare the characteristics of different information security threats, thus establishing a theoretical basis for research in the field of information security and human-computer interaction. It requires loading some binary code into each client and then converting the JDBC calls from the loaded client API to Oracle, Sybase, Informix,
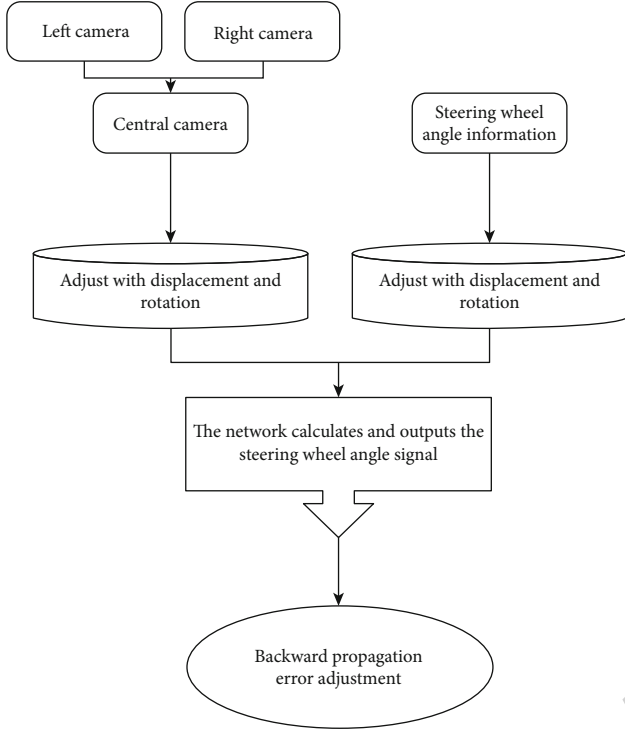
FIGURE 2: 3D target perception model.

DB2, or other DBMS transfers. Thus, given by two point sets:

$$A = \{(A_1, w(A_1), (A_2, w(A_2)), \cdots, (A_m, w(A_m)))\},$$
$$B = \{(B_1, w(B_1), (B_2, w(B_2)), \cdots, (B_m, w(B_m)))\}, \quad (5)$$

$w(A_i), w(B_i)$–weight

Finally, the channel fusion module uses a channel fusion network based on regional unification. This network effectively combines the features of each viewpoint, jointly classifies the targets on each viewpoint, and performs directional regression for the 3D candidate regions. Thus, road edge point detection is performed, and after processing such as filtering and fusion, a travelable area optimal selection algorithm is then designed to obtain the most passable area. Before filtering and fusion, the point cloud in the KITTI dataset is projected onto the RGB image plane to perform a matching verification operation with RGB. The purpose is to ensure the data match and to exclude the effect of bias caused by the data itself in the subsequent algorithm. The view cone of point cloud data is passed to the image-point cloud fusion segmentation network for front and background segmentation and to determine whether there is valid point cloud data, i.e., the scanned point cloud is obtained with the target object returned. Then, sparse points are removed from the classification list, and the final result list represents the location of possible road areas and surrounding geometric features.

2.2. Working Process of 3D Target Perception Model in Unmanned Driving. The application of 3D target perception

model in unmanned driving is divided into two parts: front-end data acquisition and data processing, with data transmission between the two parts via wireless communication. Using feature-centered Euclidean distance for target detection is the first method to use sparse convolutional layers and L1 regularization for processing large-scale 3D data. When the Euclidean distance is greater than a threshold, the point is considered a feature point; otherwise, it is a nonfeature point, and the screening formula is as follows:

$$p_i = \begin{cases} \text{True}, d_m > \sigma_h \\ \text{False}, d_m < \sigma_h \end{cases}, \quad (6)$$

σ_h–threshold value
P_i–a point in the set of points P

It is very necessary and important to study how to use 3D target perception model to process point cloud data efficiently and obtain accurate 3D target detection results for autonomous driving. The unmanned system generates many conditional branches for the information perceived by the multiview channels, and each conditional branch is divided into modules, making the whole system respond quickly to different information inputs. The unmanned behavior planning system is shown in Figure 3.

First, the data from the network and the terminal are passed through a parameter conversion processor to complete the data separation. The acceleration data is preprocessed by means of a low-pass filter with a frequency of up to 100 Hz to process the raw acceleration data for non-true values and high-frequency noise. The same point clouds undergo certain rigid changes in space (rotation or translation) and their coordinate values change. However, we expect that the network's recognition of the point cloud should remain unchanged regardless of the rotation and translation of the point cloud. Then, the residuals squared at each point are summed as the total residuals for the whole fitted plane. The residuals from a point to a plane $aX + bY + cZ = 1$ are calculated as shown in the following equation.

$$\delta_i = \frac{|ax_i + by_i + cz_i + 1|}{^2\sqrt{a^2 + b^2 + c^2}}, i = (1, 2, \cdots, n). \quad (7)$$

It is then necessary to pre-set thresholds to determine the properties of the grid, and the size of the threshold selection reflects the ability of obstacle detection. It is assumed that the average distance calculated from all points conforms to the Gaussian distribution, and the mean and standard deviation can be obtained, and the points whose average distance is outside certain criteria are eliminated. A multilayer point deconvolution is built as the main structure of the segmentation network, and feature fusion with the corresponding point convolution layer of the feature network is added. By adding a classification loss function for each point, the final target segmentation network can predict a category for each point, thus achieving the effect of target segmentation. And after transforming into the form of bird's eye view data, the amount of data can be largely reduced, while the information is arranged in pixel form with a uniform and regular pattern. Suppose the
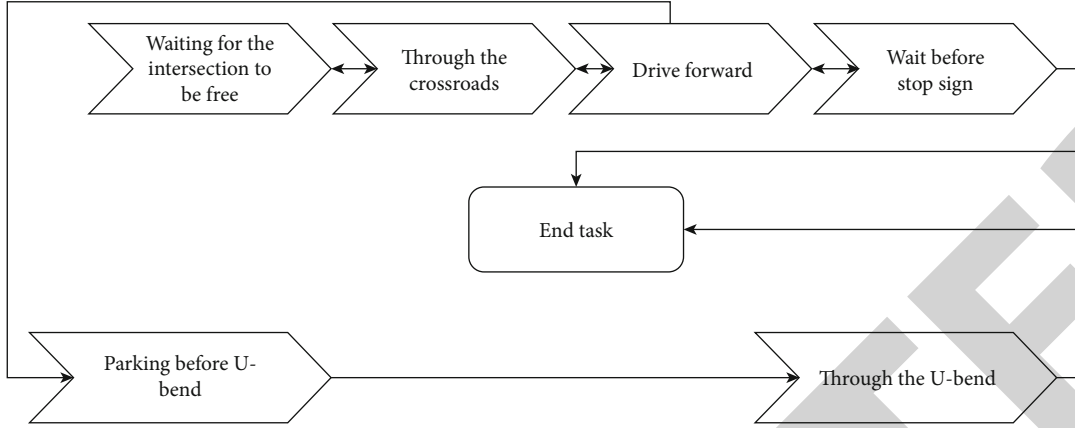
FIGURE 3: Unmanned driving behavior planning system.

system uses n multiview channels to measure a physical quantity, the ith multiview channel measurement is $X_i (i = 1, 2, \cdots, n)$, and the weighting factor of each multiview channel is w_i, then the final fusion result is:

$$\bar{X} = \sum_{i=1}^{n} w_i X_i. \tag{8}$$

Secondly, the separation is based on the data such as model position state obtained from the static model, separated into two parts, one part is used as parameters of the model data converter, and the other part modifies the dynamic data area in the model. The network structure of the auxiliary path is the same as the number of intermediate layers of the original network structure, and each intermediate layer in the auxiliary path shares the weights of the corresponding intermediate layers of the original network. In the backpropagation process, each auxiliary path also performs loss calculation, and the calculated loss on the auxiliary path is called the auxiliary loss, and the weights are also updated on the auxiliary path after calculating the auxiliary loss. Radius filtering is to define the points lacking enough nearest neighbors as isolated noise points. It is necessary to set the search radius and the threshold of the number of nearest neighbors, and if the set conditions are not satisfied, the corresponding points are removed. This part performs a convolution operation on the image block and the view cone point cloud data in parallel, where the image is convolved using a dimensionally invariant multiset convolution and the point cloud is convolved using a point convolution network and the features of the image are fused at each layer of the point convolution network. For this purpose, candidate point sets need to be constructed separately for the points in the point set P, and the candidate point sets in Q are constructed as follows.

$$C = \{c_i | c_i = \langle p_i, q_{i1}, q_{i2}, q_{i3}, \cdots, q_{ik}, 1 \le i \le 4 \rangle\}, \tag{9}$$

c_i–the ith point
p_i–candidate point set
k–number of candidate points

Finally, after adjustment, getting the adjustment and data acquisition strategy, the model conversion processor obtains the model data of the display from the static model data, then merges the dynamic model data, and finally sends the data to the display or network device. Point cloud noise refers to the existence of a certain probability that a point lies in a threshold radius around its original position. Compared with other similarity measures, the use of rank function is not so sensitive to image noise and luminance differences and is good in real time, and its calculation formula is:

$$\begin{cases} C_{\text{Rank}}(u, v, d) = g_1(u, v) - g_r(u + d, v) \\ g_1(u, v) = \sum_{(i,j)} g_1(u + i, v + i) < g_1(u, v) \end{cases}. \tag{10}$$

Also, there are outliers, i.e., random distributions of coordinates, in point clouds. Radar data are in the form of point clouds with varying amounts of data, but they all contain a large amount of noise, so the noise needs to be filtered out by a filtering algorithm. However, the sampling frequency of the horizontal axis of the road contour is no longer uniform due to the changing vehicle speed, and the sample interval of the road contour is small when the driving speed is slow and large when the driving speed is fast. So after data acquisition with noise removed, the system will fuse LiDAR features and image features and a priori knowledge of the road model to perform drivable area detection. For the occluded objects with different attributes, the size and arrangement of the objects described in the high-precision scene environment as well as are different. To make the acceleration data and velocity data sampled at the same frequency, the velocity data are interpolated.

3. Application Analysis of 3D Target Detection Algorithm in Unmanned Driving

3.1. Analysis of 3D Target Detection Algorithm. The 3D target detection algorithm can obtain the classification result of the target object and the coordinate information in 3D space by processing the 3D data. Due to the binary output characteristics, the cross-entropy function is more suitable
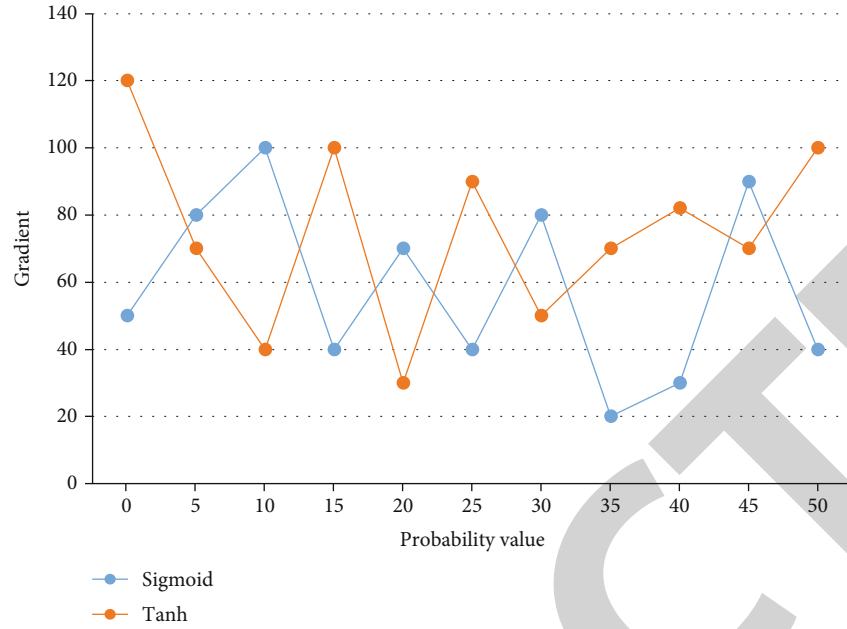
Figure 4: Comparison of gradients produced by different activation functions.

for pavement type classification than the squared error cost function, and the cross-entropy function is closer to the criteria of real classification. The 3D target detection algorithm separates the point clouds in space according to the foreground background based on the class prediction of each point by the target segmentation network. And the activation function is used to extract the points of the foreground object category out of them, so as to complete the extraction of the target object. The activation function must be nonlinear, and the activation function must be derivable everywhere to ensure computability in gradient calculation. The comparison of the gradients generated by different activation functions is shown in Figure 4.

First, after the depth information of each pixel on the image is determined by monocular or binocular vision, each coordinate is assigned to the corresponding depth information. The corresponding depth map is obtained and then transformed into a point cloud by geometric principles. The input point cloud data needs to be gridded, and the resolution of the gridding will determine the operation efficiency and detection accuracy of the algorithm later, too large may lead to slow operation and overfitting of the algorithm, too small will not be able to accurately detect the target location. Reasonable grid node distribution plays an important role in the calculation of nonlinear hyperbolic conservation law equations. The numerical solution of hyperbolic conservation law equations has always been one of the important research topics in computational fluid dynamics. In the numerical solution of nonlinear hyperbolic conservation law equations, if the mesh is uniformly divided, in the region where the properties of the solution change little, the sparse mesh distribution can obtain more accurate results. When there is a vehicle signal, the sampled signal will have greater fluctuation, and the median filtering method can effectively remove the impulse interference signal, so as to get a smooth waveform signal; Figure 5 shows the comparison of the signal after AMR-X and AMR-Y filtering.

Due to the limitation of the working principle of LiDAR, the point clouds collected have the characteristics of near-dense and far-sparse, and the high density of near points easily leads to the slow operation of the clustering algorithm. Therefore, voxel filtering of the point cloud data is needed first to improve the real-time performance of the algorithm. Compared with the basic detection network architecture such as VGG, Darknet53 has a residual module that allows the network to be optimized more easily at the same depth. Therefore, the residual module allows to increase the network depth to improve the network performance without degradation. After the channel fusion network fuses the features of each viewpoint channel, the fusion results are used to perform regression correction on the 3D candidate regions generated by the candidate region module.

Next, the pseudo point cloud data corresponding to these target objects are found based on the pixel mapping of the target objects segmented on the image to the pseudo point cloud. After meshing the point cloud data, the features of each mesh need to be specified, which can be set artificially or extracted using deep learning. The front-end data acquisition part collects data from each lane of the traffic intersection and packages the data for transmission to the upper computer data processing host for calculation and processing via wireless sensor network according to the developed transmission protocol. The upper computer data processing host displays traffic information such as current traffic flow, speed, and time of obtaining vehicle information, and saves the traffic data. The decomposition of image basis functions with ICA is limited to the spatial domain only, and if temporal variations are considered, such as continuous horizontal motion, vertical motion, and rotational motion, image sequences in the time domain can be
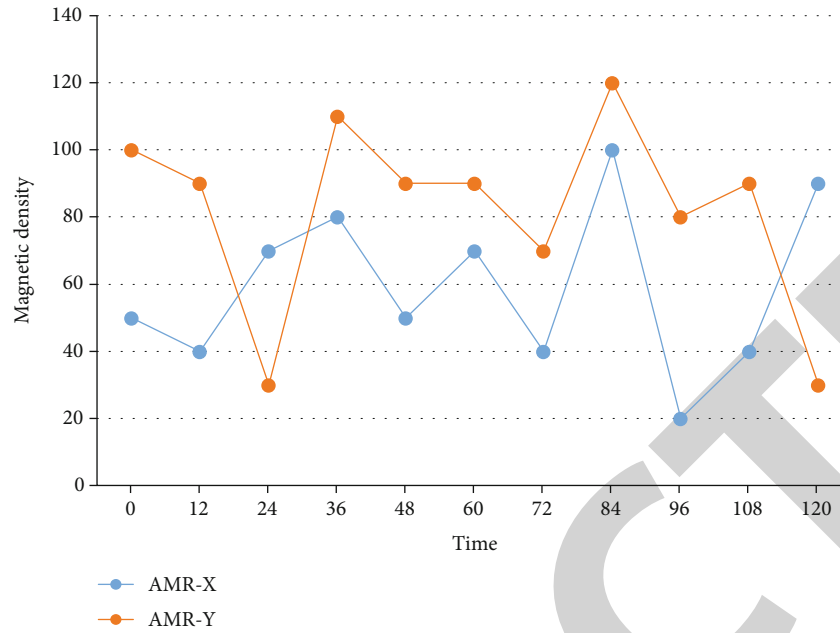
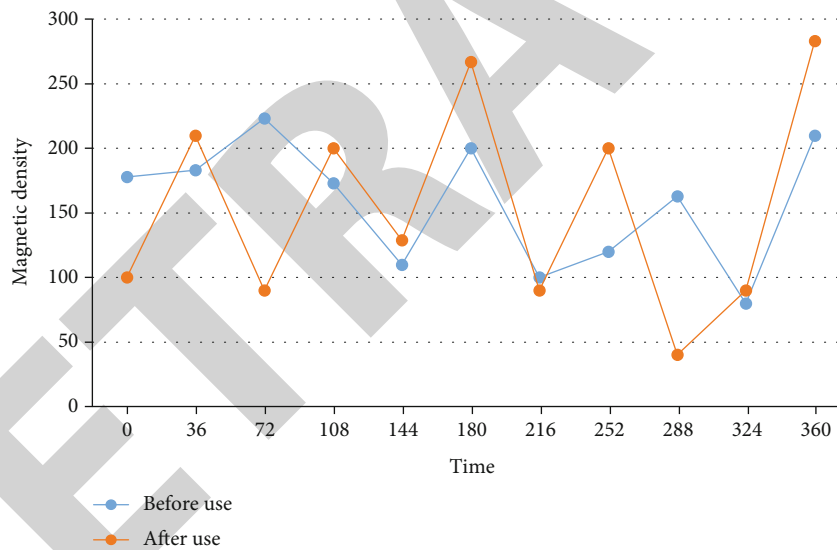Figure 5: Comparison of AMR-X and AMR-Y filtered signals.

Figure 6: Waveform changes of a driving vehicle before and after using 3D target detection algorithm.

composed. The classification and regression of regions are performed by two branches of the region recommendation network, and then feature extraction is performed on the feature map according to the size scaling of the filtered recommended regions and sent to the classification and regression branches to obtain the class and location of the target. In order to detect low-speed vehicles in congested lanes, the waveform change of a moving vehicle before and after using the 3D target detection algorithm based on the multistate machine vehicle detection algorithm is shown in Figure 6.

Finally, the pseudo point cloud in the candidate frame and the corresponding labeled frame of the candidate frame are fed into the second stage of the Point RCNN network for

candidate frame correction training. The 3D convolution of the gridded data is performed to extract the necessary features. Based on the point cloud data, the laser beam to which the point belongs and the distance from the radar center can be calculated. If the distance is less than the set distance of the laser beam to which it belongs, then the point is a nonground point; if the distance is not much different from the set distance of the laser beam to which it belongs, then the point is a ground point. The distance between the data points near the vehicle is still much smaller than the distance far away from the vehicle. If we simply use the same radius threshold for point cloud clustering, we will not be able to achieve good clustering effect in both near and far away. If the threshold is too large, the objects in the near area and
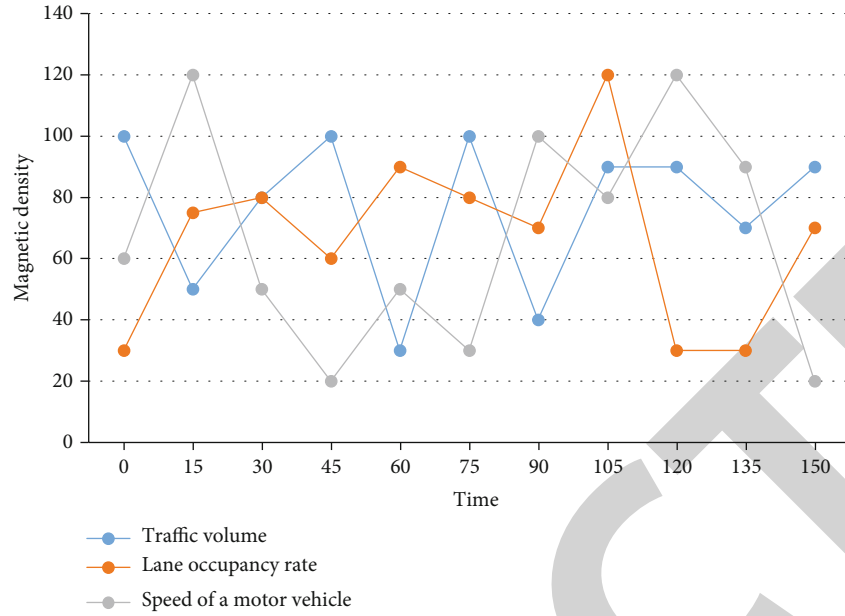
Figure 7: Vehicle identification results.

the environment will be grouped into the same cluster, while if the threshold is too small, the targets in the far area of the vehicle will not be clustered correctly. Therefore, one dimension can be reduced by setting the height span of the 3D convolution, fusing the dimension of height into the information of the channel, and then regressing the position and size of the target using the anchor frame and the final feature map as in the image detection algorithm.

3.2. Target Detection and Location Analysis. Different sensors output different forms of data, and the corresponding data processing processes are different. Before radar data can be processed, target detection and localization must be analyzed according to the radar protocol. After multiview channel calibration and data acquisition, the algorithm incorporates laser features, image features, and the necessary a priori knowledge to achieve driving area detection. Since the acquisition frequency of IR camera and LIDAR is not the same, this paper matches the IR image according to the LIDAR frame number, sets the threshold value to 20 frames, and looks for the image with the radar frame number less than the threshold value and closest to it as the corresponding image. After acquiring the parameters, the various data provided by the system implementation layer should be integrated, and the scheduling and adjustment should be carried out. The variance is multiplied as a weight factor by the original signal as the feature signal, so that the characteristics of the magnetic field strength signal can be maintained and the drift of the signal can be suppressed by the characteristics of the variance. The vehicle recognition results are shown in Figure 7.

First, a pseudo point cloud is generated and the data with height more than 2 m above the observation origin is removed, and the point cloud is centered on the ground. Since most of the information of the point cloud target is gathered at its edges, the features at the edges can be

extracted effectively by using deformable convolution to achieve the same effective accuracy, so the setting of the anchor frame is not necessary, and the target frame is reclustered for each new scene. A multilayer point convolutional network is built using point convolutional layers as the main structure of the localization network, and several branching networks are added after the main body of the network according to the prediction task as the head structure of the prediction network. When the vehicle passes above, the AMR sensor detects the information of the vehicle's perturbation to the earth's magnetic field in real time, and the subsequent processing circuit amplifies and filters this perturbation signal and then converts it into a digital signal by the controller's A/D sampling channel processing to analyze the sampled time series signal. The time synchronization algorithm is used to time synchronize the two sensor nodes, and the vehicle detection algorithm mentioned in this paper can be used to obtain the moment when the vehicle passes the sensor node, and the waveform perturbation schematic of the dual sensor node speed measurement process is shown in Figure 8.

Next, the left-eye images are put through Mask RCNN model and inference is performed. Then, the minimum enclosing box algorithm is used to calculate the minimum enclosing box of the point cloud to get the position, size, and direction of each target in the LIDAR column coordinate system. All the 3D radar data falling into the grid are calculated to get their maximum-minimum quotient difference, and if the difference is greater than the set idle value, the grid is an obstacle point, and vice versa; it is a ground point. The main concern is whether the three-dimensional detector has missed detection, because if it misses detection, the three-dimensional target location network is unable to retrieve the lost object again. By analyzing and processing the geomagnetic vehicle characteristic information collected by the receiving station, the real-time data waveform of each
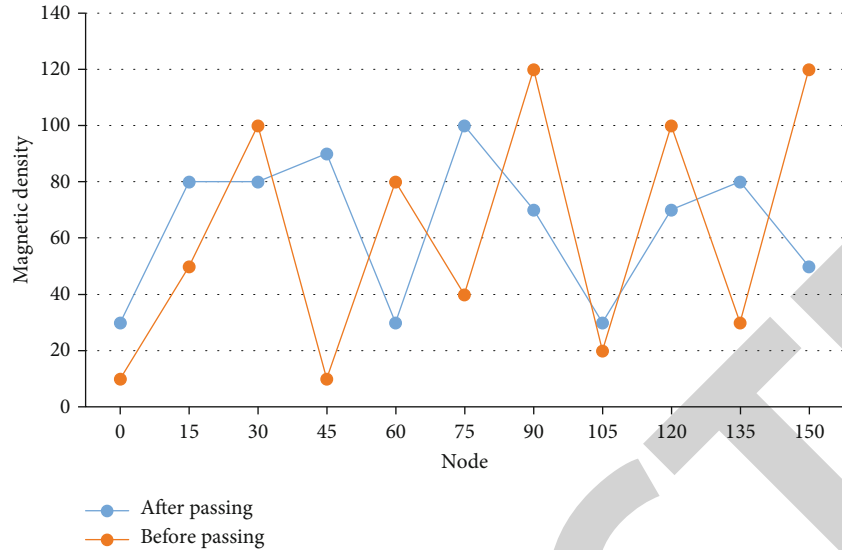
FIGURE 8: Schematic diagram of magnetic field waveform disturbance when a vehicle passes through a dual sensor node.

TABLE 1: Comparison of target test results.

	Fast RCNN	SSD	Mask RCNN
Voc index (mAP)	67.26	73.45	86.71
COCO (AP50) index	34.99	45.62	57.38

TABLE 2: Comparison of three-dimensional target detection results.

	VoxelNet	AVOD	MV3D
Model	LiDAR	RGB + LiDAR	RGB + LiDAR
Difficulty index	52.99	76.17	56.32

lane is displayed. By calculating and obtaining traffic information such as road traffic flow, vehicle speed, single vehicle length, vehicle running direction and vehicle type, reliable data are provided for the traffic information center. And for different scene obstacles, the final multipath effect is different because of their material, shape, and other property differences, and the reflected wave characteristics are also different.

Finally, the mask generated by Mask RCNN is combined with the pseudo point cloud to delineate the point cloud foreground and background and filter out the background information. In the process of updating the network parameters using stochastic gradient descent and chained derivatives, the gradient descent algorithm must accurately calculate the error of each layer before updating the parameters. The comparison of the target detection results of Fast RCNN, Mask RCNN, and SSD is shown in Table 1.

The average precision mean result of Mask RCNN is 13.26 higher than SSD on the PASCAL VOC test set and 19.45 better than Fast RCNN.

If the selected training data does not fully contain all the feature information that the test data has, it can seriously affect the classification results. Therefore, a sampling window of size 5×5 pixels is randomly placed at a pixel location in a large image. It is important to note that the entire sampling window boundary should be within the boundary of the large image, cropping a small piece of the same image size. The real regions of the three categories of vehicles and pedestrians and cyclists are also marked on each copy of data by manual annotation. A bottom-up candidate region search network is used to extract the candidate regions from the point cloud. The detection results of each 3D target detection algorithm MV3D, AVOD, and VoxelNet are shown in Table 2.

The foreground points are marked in green in the whole point cloud scene to observe the performance of the 3D segmentation network. And the results of the 3D prediction module are represented by 3D boxes to observe whether the network predicts the 3D pose correctly. The purpose is to increase the nonlinear representation capability of the network to obtain more refined image features. Also, the design of filtering the reflected waves by means of antenna polarization, etc., in order to eliminate the negative impact of the localization process due to multipath effect is applied in many receivers.

4. Conclusions

Unmanned self-driving vehicle is also known as intelligent vehicle. It uses on-board multiview channels to sense the vehicle's surroundings and controls the vehicle's steering and speed based on the road, vehicle attitude, and obstacle information obtained from the sensing, thus enabling the vehicle to travel safely and reliably on the road. If RGB data is used, although the 2D recognition network is more mature, it will face the problems of dark light failure and the presence of strong occlusion. In addition, the method of classifying different road types using acceleration multiview channels alone is relatively ineffective for classifying and recognizing tarmac and concrete roads with similar vibration characteristics. In contrast, using a 3D target information perception model with multiview fusion, after

obtaining 2D detection results, a 3D target localization network based on image-point cloud fusion can provide accurate and fast estimation of the spatial location, size, and pose of surrounding target objects in an unmanned environment. In this paper, we take the autonomous and safe driving of unmanned intelligent vehicles in the night environment as the application background, and propose the idea of applying 3D target information perception model in unmanned intelligent vehicles at night. It also analyzes the application of 3D target detection algorithm in unmanned driving and investigates how to improve the information perception ability and driving behavior planning ability of unmanned intelligent vehicles in the surrounding environment at night. The research of this paper further uses the good detection and segmentation results of 3D target detection for foreground background segmentation of pseudo point clouds, which significantly reduces the amount of operations in the subsequent steps. However, this paper does not discuss and analyze the problem of low detection accuracy when the existing technology realizes three-dimensional target detection. Therefore, further development analysis is needed in future research.

Data Availability

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

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] W. Yulan, S. Shaoyuan, L. Zhiyi, and B. U. Defei, "Nighttime three-dimensional target detection of driverless vehicles based on multi-view channel fusion network," *Applied Optics*, vol. 41, no. 2, pp. 296–301, 2020.

[2] Q. Xue and X. Wu, "Vehicle detection for unmanned systems based on multimodal feature fusion," *Journal of Guangxi Normal University: Natural Science Edition*, vol. 12, no. 12, p. 6198, 2022.

[3] W. Zhou, L. Lu, and J. Wang, "A review of research on multisensor information fusion in autonomous driving," *Automotive Digest*, vol. 1, p. 7, 2022.

[4] J. Liu, "Autonomous driving perception based on Mask R-CNN network model," *Automotive Practical Technology*, vol. 7, p. 2, 2019.

[5] H. Zhang, G. Wang, and P. Sun, "Research on target detection system based on information fusion," *Electronic Measurement Technology*, vol. 44, no. 19, p. 8, 2021.

[6] J. Shi, L. Zhou, T. Liu et al., "Multiple-view D^{2}NNs array: realizing robust 3D object recognition," *Optics Letters*, vol. 46, no. 14, pp. 3388–3391, 2021.

[7] L. Binping and Z. Yue, "A novel anchor-free 3D object detector," *China Stereology and Image Analysis*, vol. 1, p. 7, 2020.

[8] J. Chen, "A single image 3D target tracking and positioning method based on color difference model," *Journal of the Yellow River University of Science and Technology*, vol. 22, no. 11, p. 5, 2020.

[9] G. Wang and P. Wang, "Research on 3D object detection method based on deep learning," *Computer Application and Software*, vol. 37, no. 12, p. 5, 2020.

[10] J. Zhang and J. Liu, "Design and application of joint environment perception model based on blockchain technology," *Computer and Modernization*, vol. 279, no. 11, pp. 56–59, 2018.

[11] B. Wang, Y. Han, D. Tian, and T. Guan, "Sensor-based environmental perception Technology for Intelligent Vehicles," *Journal of Sensors*, vol. 2021, Article ID 8199361, 14 pages, 2021.

[12] Z. Ouyang, J. Cui, X. Dong, Y. Li, and J. Niu, "SaccadeFork: a lightweight multi-sensor fusion-based target detector," *Information Fusion*, vol. 77, pp. 172–183, 2022.

[13] Q. Chen, Y. Xie, S. Guo, J. Bai, and Q. Shu, "Sensing system of environmental perception technologies for driverless vehicle: a review of state of the art and challenges," *Sensors and Actuators A: Physical*, vol. 319, article 112566, 2021.

[14] A. Chetouane, S. Mabrouk, I. Jemili, and M. Mosbah, "Vision-based vehicle detection for road traffic congestion classification," *Concurrency and Computation: Practice and Experience*, vol. 34, no. 7, article e5983, 2022.

[15] Q. Liu, Z. Li, S. Yuan, Y. Zhu, and X. Li, "Review on vehicle detection technology for unmanned ground vehicles," *Sensors*, vol. 21, no. 4, p. 1354, 2021.