

Research Article

An Improved VM Obstacle Identification Method for Reflection Road

Guoxin Jiang , **Yi Xu** , **Xiaoqing Sang** , **Xiaotong Gong** , **Shanshang Gao** ,
Ruoyu Zhu , **Liming Wang** , and **Yuqiong Wang** 

School of Transportation and Vehicle Engineering, Shandong University of Technology, Zibo 255000, China

Correspondence should be addressed to Yi Xu; xuyisdut@163.com

Received 23 January 2022; Accepted 18 February 2022; Published 14 March 2022

Academic Editor: L. Fortuna

Copyright © 2022 Guoxin Jiang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

An obstacle detection method based on VM (VIDAR and machine learning joint detection model) is proposed to improve the monocular vision system's identification accuracy. When VIDAR (Vision-IMU-based detection and range method) detects unknown obstacles in a reflective environment, the reflections of the obstacles are identified as obstacles, reducing the accuracy of obstacle identification. We proposed an obstacle detection method called improved VM to avoid this situation. The experimental results demonstrated that the improved VM could identify and eliminate unknown obstacles. Compared with more advanced detection methods, the improved VM obstacle detection method is more accurate. It can detect unknown obstacles in reflection, reflective road environments.

1. Introduction

Obstacle detection has become a major concern in the field of driver assistance systems due to the complexity of the outdoor environment. Cameras (monocular, binocular, infrared, etc.), lidar, and millimeter-wave radar are all examples of obstacle identification equipment. While lidar and millimeter-wave radar are highly accurate at detecting obstacles, their high cost limits their use in low-end vehicles [1–3]. Due to the low cost, high detection accuracy, and speed of vision-based obstacle identification equipment, it has become more suitable for various vehicles [4, 5]. The vision-based sensor used in this study is a camera. Camera, GPS, and IMU constitute an innovative sensor combination. Compared with single sensor, the application of multisensor information fusion technology can improve the reliability of the whole system, enhance the reliability of data, improve the accuracy, and increase the information utilization rate of the system in solving the problems of detection, tracking, and target recognition.

Machine learning is the process of training and identifying images using deep convolutional neural networks. Compared with other image recognition technologies,

machine learning has an extremely high recognition rate for specific images. While machine learning is capable of accurate classification, it can only be used to identify known obstacles. While the vehicle is in motion, using machine learning to identify unknown obstacles may result in misidentification, posing a serious risk to the vehicle's safety (Figure 1). As a result, a method for detection and ranging using vision and an IMU (inertial measurement unit) has been proposed [6]. Given that VIDAR requires more time to run than machine learning, a method combining VIDAR and machine learning to detect obstacles has been proposed (called the VM method). Machine learning is used to identify known obstacles in the proposed method, while VIDAR is used to detect unknown obstacles.

To avoid the situation in which VIDAR detects the reflection as an obstacle when used in a reflective environment (Figure 2) and improve detection accuracy, a VIDAR-based pseudo-obstacle detection method (called improved VIDAR) has been proposed. This method's identification procedure is as follows. The rectangle of the obstacle was determined. The width of the obstacle rectangle was calculated using the transformation relationship between pixel coordinates and world coordinates, and then the height of

the obstacle rectangle was calculated using the transformation relationship between pixel coordinates and world coordinates. The true obstacle is determined by the fact that the actual height of the obstacle rectangle remains constant throughout the ego-vehicle movement. If the obstacle is a real one, tracking is continued.

To accelerate the detection speed of improved VIDAR, we combined it with machine learning (this article uses the faster RCNN algorithm) to identify known obstacles, which we refer to as improved VM. The improved VM obstacle detection method can quickly and accurately detect obstacles on reflection roads. The enhanced VM obstacle detection procedure is as follows: first, machine learning is used to identify known obstacles; second, the identified obstacles are removed from the background area; and finally, pseudo-obstacles are eliminated through the use of enhanced VIDAR.

2. Related Work

As the core section of automobile-assisted driving, obstacle detection has emerged as a critical area of research in recent years. Due to its simple ranging principle, the monocular vision sensor has become the primary obstacle identification equipment in obstacle identification. Many scholars have conducted related research on obstacle identification to accelerate the process. Traditional image object classification and detection algorithms and strategies are difficult to meet the requirements of image and video big data in terms of processing efficiency, performance, and intelligence. Deep learning establishes the mapping from low-level signals to high-level semantics by simulating the hierarchical structure similar to the human brain, so as to realize the hierarchical feature expression of data, and has powerful visual information processing capabilities. Therefore, in the field of machine vision, the representative of deep learning-convolutional neural network (CNN) is widely used [7, 8]. Convolutional neural networks are also called cellular nonlinear networks. Arena et al. have stressed the universal role that cellular nonlinear networks (CNNs) are assuming today. It is shown that the dynamical behavior of 3D CNN-based models allows us to approach new emerging problems, to open new research frontiers [9]. Shustanov and Yakimov proposed an implementation of the traffic signs recognition algorithm using a convolution neural network; training of the neural network is implemented using the TensorFlow library and massively parallel architecture for multithreaded programming CUDA; and the experiment proves the high efficiency of this method [10]. Zhu et al. have proposed a novel image classification framework that combines CNN and KELM (kernel extreme learning machines). They extracted feature categories using DenseNet as a feature extractor and radial basis function kernel ELM as a classifier to improve image classification performance [11]. Wang et al. proposed the occlusion-free road segmentation network, a fully convolutional neural network. Through foreground objects and visible road layouts, this method can predict roads in the semantic domain [12]. The accuracy of obstacle identification is also continuously improved as new

machine learning concepts such as SegNet, YOLO v5, faster RCNN, BigGAN, and mask RCNN are developed [13–20]. While machine learning is capable of accurate classification, it can only be used to identify known obstacles. Unknown obstacles may be missed while the vehicle is moving, which will cause a serious impact on the vehicle’s safety.

Generally, obstacles are detected using significant information such as color and prior shape. Zhu et al. proposed a method for detecting vehicles based on their edges and symmetry characteristics [21]. They hypothesized the vehicle’s location based on the image’s detected symmetric regions. The vehicle’s bounding box is determined using the projected image of the enhanced vertical and horizontal edges. Zhang et al. [21–23] used color information for background removal and shadow detection to improve object segmentation and background updating. This method is capable of rapidly and precisely detecting moving objects. Zhang et al. [24] introduced Deep Local Shapes (DeepLS), which are high-quality 3D shapes that can be encoded and reconstructed without requiring an excessive amount of storage. This local shape of the scene decomposition simplifies the prior distribution that the network must learn and accelerates and accurately detects obstacles. However, in an environment with reflections, the reflections contain significant information about the obstacles they use, reducing the accuracy of obstacle detection.

The vehicle’s position is generally determined by high-light information, such as the highlighted area and contour features. Park and Song [25] proposed a front vehicle identification algorithm based on contrast enhancement and vehicle lamp pairing. Lin et al. [26] discovered that the characteristics of headlights were more distinctive than the contours of vehicles and had a greater identification effect and thus proposed the use of lamps as a sign for vehicle identification at night. The Hough transform was proposed by Dai et al. [27] as a method for intelligent vehicle identification at night. This method divides the extracted lamps into connected domains, extracts the lamps’ edges, and then identifies the circle using the Hough transform. Finally, by pairing the lamps, the vehicle’s location is determined. Kavva et al. [28] proposed a method for detecting vehicles based on the color of the brake lamp during braking in the captured color image. The feature information required for the above identification method in a reflective environment will detect lamp reflections. The vehicle’s lamps will also be paired, which will reduce the vehicle’s accuracy of identification. We used a modified VM to detect obstacles, allowing for eliminating obstacles in the reflection, thereby increasing obstacle detection accuracy.

3. Methodology of Improved VIDAR’s Pseudo-Obstacle Detection

The monocular visual identification method, based on machine learning, is limited to identifying previously identified obstacles. A vehicle collision accident may occur when unknown obstacles are present on the road. When VIDAR is used to detect obstacles, pseudo-obstacles in the reflection environment are mistaken for real obstacles. Thus,

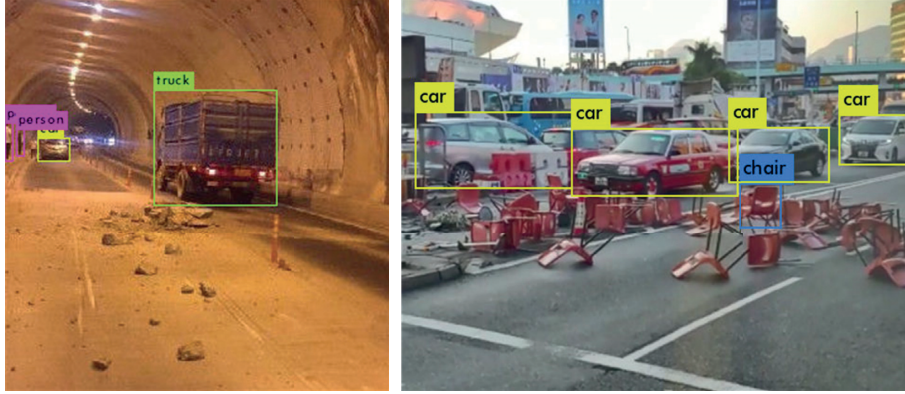


FIGURE 1: The results of obstacle identification based on faster RCNN.



FIGURE 2: The results of car detection based on VIDAR on reflection roads.

to increase the speed and accuracy of obstacle detection, we use an improved VM.

3.1. Transformation from World Coordinates to Pixel Coordinates. The camera can project objects in the three-dimensional world into a two-dimensional image by capturing an image. In reality, the imaging model establishes a projection mapping relationship between three-dimensional and two-dimensional space. The coordinate transformation is required to convert the world coordinate system's coordinates to the camera coordinate system's coordinates. A rigid body transformation is used to convert the world coordinate system to the camera coordinate system. It is determined by the camera's external parameters. The camera coordinate system to pixel coordinate system transformation converts three-dimensional coordinates to two-dimensional plane coordinates, as determined by the camera's internal parameters. Although both the pixel and image coordinate systems are located on the imaging plane, their origins and units of measurement are distinct. The origin of the image coordinate system is the point at which the camera's optical axis intersects the imaging plane, which is typically the imaging plane's midpoint.

Suppose the internal parameters matrix is M . Project $Q(X, Y, Z)$ in the physical world to the image plane $q(x, y)$. By adding a dimension w to $q(x, y)$, which is the expansion of $q(x, y, w)$, obtain $w = Z$. Point q is in the form of

homogeneous coordinates. Combine the rotation matrix R and the offset matrix T to obtain the external parameter matrix K , where c_x and c_y are the offsets. The coordinate transformation is shown in Figure 3.

The internal parameters of the camera are obtained by Zhang Zhengyou demarcate to determine the transformation relationship between world coordinates and pixel coordinates.

$$q = MKQ. \quad (1)$$

$$\text{Among them: } q = \begin{bmatrix} x \\ y \\ w \end{bmatrix}, \quad M = \begin{bmatrix} f_x 0 c_x \\ 0 f_y c_y \\ 0 0 1 \end{bmatrix},$$

$$K = \begin{bmatrix} R_{3 \times 3} & t_{3 \times 1} \\ 0 & 1 \end{bmatrix}, \text{ and } Q = \begin{bmatrix} X \\ Y \\ W \end{bmatrix}, \begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} f_x 0 c_x \\ 0 f_y c_y \\ 0 0 1 \end{bmatrix} \begin{bmatrix} R_{3 \times 3} t_{3 \times 1} \\ 0 1 \end{bmatrix}$$

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}.$$

3.2. Obstacle Ranging Model. The range of obstacles is as follows (Figure 4). Let f be the focal length of the camera; h be the installation height of the camera; μ be the pixel size; ∂ be the camera pitch angle; (x_0, y_0) be the intersection of the image plane and the optical axis of the camera; set to $(0, 0)$; (x, y) be the coordinates of intersection points of obstacles and pavement plane set to P ; and the horizontal distance d between the object point and the camera is

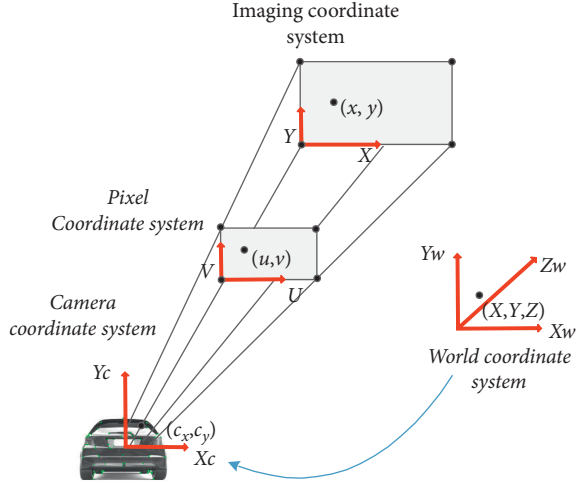


FIGURE 3: Schematic diagram of coordinate transformation.

$$d = \frac{h}{\tan(\theta + \arctan[(y_0 - y)\mu/f])} \quad (2)$$

Assume that Y_1 is the Y -axis in the previous image, Y_2 is the Y -axis in the previous image. When the camera moves from Y_1 to Y_2 on the axis of the imaging plane (Figure 5), let A be an imaging point for obstacle's top in the subsequent image, B be the same imaging point for obstacle's top in the latter image, A' is the object point of A , and B' is the object point of B . d_1 is the horizontal distance between A' and the camera; similarly, d_2 is the horizontal distance between B' and the camera. d_1 and d_2 can be obtained from Equation (3). The camera moved a certain distance Δd during the time between the previous and subsequent images; $d_1 = d_2 + \Delta d$, but $d_1 = d_2 + \Delta d + \Delta l$ actually. As a result, the A' and B'

have height if $d_1 \neq d_2 + \Delta d$, and static obstacles can be identified by Δl if Δd is known.

Additionally, if the obstacle is moving (as illustrated in Figure 6), the Δl can also be used as an obstacle judgment. The verification process has been shown in the paper [6].

3.3. Static Obstacle Identification Model. There are two types of static obstacles: real static obstacles and static pseudo-obstacles. Static real obstacles refer to actual road obstacles. The reflections identified as real obstacles during the obstacle identification process are called static pseudo-obstacles. It is a type of pseudo-obstacle that reflects some road obstacles but does not affect the vehicle's driving safety. To improve the accuracy of obstacle identification, we must identify and remove static pseudo-obstacles.

3.3.1. Static Real Obstacle Identification. First, we used the VIDAR to detect stereo obstacles and determine which object point on the obstacle is the furthest away in the horizontal and vertical directions to construct a rectangle (the obstacle rectangle, Figure 7). Let $A(x_1, y_1)$ be the first imaging point for the identification of the width of the rectangular road surface of the obstacle, $B(x_2, y_1)$ be the other imaging point for the identification of the width of the rectangular road surface of the obstacle, and A' be the object point of A ; similarly, B' is the object point of B . The horizontal distances d_1 between A' and the camera can be calculated by Equation (2). Similarly, the horizontal distances d_2 between B' and the camera can also be calculated. The width of the obstacle rectangle can be calculated using the pinhole imaging principle and the geometrical relationship between cameras.

$$w = \sqrt{2h^2 + d_1^2 + d_2^2 - \frac{(2f^2 + 2y_1^2 + x_1^2 + x_2^2 - \mu^2(x_2 - x_1)^2)\sqrt{h^2 + d_1^2}\sqrt{h^2 + d_2^2}}{\sqrt{f^2 + x_1^2 + y_1^2}\sqrt{f^2 + x_2^2 + y_1^2}}} \quad (3)$$

When the camera moves from Y_1 to Y_2 on the axis of the imaging plane (Figure 8), let $C(x_3, y_2)$ be an imaging point for identifying the width of the opposite side of the obstacle, $D(x_4, y_2)$ be another imaging point for the identification of

the width of the opposite side of the obstacle, C' be the object point of C , and D' be the object point of D . (4) and (5) have the same width. The height of the pseudo-obstacle rectangle is calculated.

$$\frac{2f^2 + 2y_2^2 + x_3^2 + x_4^2 - \mu^2(x_4 - x_3)^2}{\sqrt{f^2 + x_3^2 + y_2^2}\sqrt{f^2 + x_4^2 + y_2^2}} = \frac{2(h - h_v)^2 + d_3^2 + d_4^2 - w^2}{\sqrt{(h - h_v)^2 + d_3^2}\sqrt{(h - h_v)^2 + d_4^2}} \quad (4)$$

3.3.2. Static Pseudo-Obstacle Identification. The procedure for identifying pseudo-obstacles is similar to the procedure for identifying real obstacles. However, when obstacles are detected using VIDAR, the object points of the obstacles are different from their actual positions (Figure 9). We detect

pseudo-obstacles using VIDAR and construct a rectangle (pseudo-obstacle rectangle) from the object points on the pseudo-obstacle (the object points are the farthest in the horizontal and vertical directions). Let A be the first imaging point for pseudo-obstacle width identification with (x_1, y_1) ,

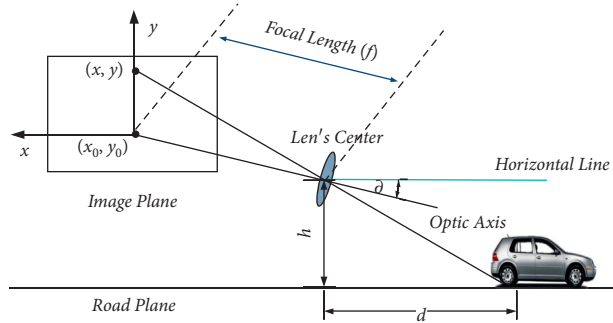


FIGURE 4: Schematic diagram of the horizontal distance between the object point and the camera.

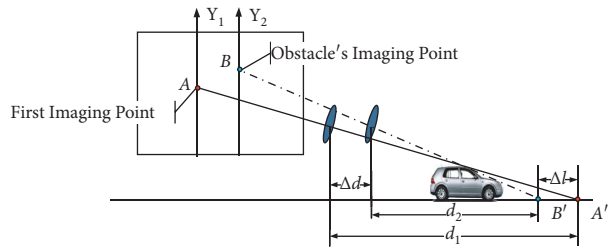


FIGURE 5: Static obstacle imaging.

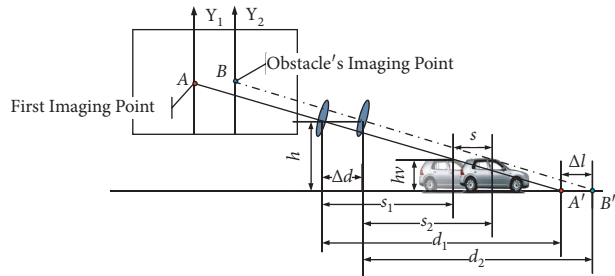


FIGURE 6: Moving obstacle imaging.

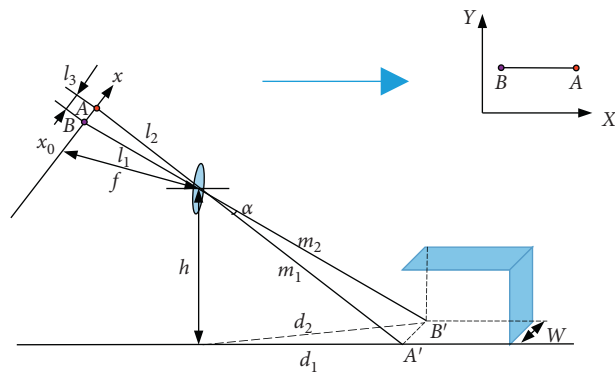


FIGURE 7: Schematic diagram of static obstacle.

B be the other imaging point for pseudo-obstacle width identification with (x_2, y_1) , A' be the object point of point A, and B' be the object point of point B. The horizontal distances d_1 between A' and the camera can be calculated using (2). Similarly, the horizontal distances d_2 between B' and the

camera can also be calculated using (2), and the width of the pseudo-obstacle rectangle can be calculated by the pinhole imaging principle and the geometrical relationship between cameras.

$$w = \sqrt{2h^2 + d_1^2 + d_2^2 - \frac{(2f^2 + 2y_1^2 + x_1^2 + x_2^2 - \mu^2(x_2 - x_1)^2)\sqrt{h^2 + d_1^2}\sqrt{h^2 + d_2^2}}{\sqrt{f^2 + x_1^2 + y_1^2}\sqrt{f^2 + x_2^2 + y_1^2}}}. \quad (5)$$

When the camera moves from Y_1 to Y_2 on the axis of the imaging plane (Figure 10), after the pseudo-obstacle moved, let C (x_3, y_2) be an imaging point of the rectangular width of the identified pseudo-obstacle and D (x_4, y_2) be another imaging point of the rectangular width of the pseudo-obstacle, and C' be the object point of C and D' be the object

point of D. At this point, the width of the object point of the pseudo-obstacle changes from W_1 to W_2 . Similarly, W_2 can be solved using the pinhole imaging principle and the geometrical relationship between cameras. The rectangular height of the pseudo-obstacle can be obtained by (5) and (6) and the triangle similarity principle.

$$\frac{2f^2 + 2y_2^2 + x_3^2 + x_4^2 - \mu^2(x_4 - x_3)^2}{\sqrt{f^2 + x_3^2 + y_2^2}\sqrt{f^2 + x_4^2 + y_2^2}} = \frac{2(h - h_v)^2 + d_3^2 + d_4^2 - (h/h + h_v)^2 w^2}{\sqrt{(h - h_v)^2 + d_3^2}\sqrt{(h - h_v)^2 + d_4^2}}. \quad (6)$$

3.4. Static Obstacle Identification Model. Moving obstacles are classified as either moving real obstacles or moving pseudo-obstacles. Moving real obstacles refers to obstacles on the road. The reflections identified as real obstacles during the obstacle identification process are referred to as moving pseudo-obstacles. It is a type of pseudo-obstacle that replicates some road obstacles but does not destroy the vehicle's driving safety. We must identify and remove moving pseudo-obstacles to improve accuracy when identifying obstacles.

3.4.1. Static Pseudo-Obstacle Identification. The steps for identifying real moving obstacles are identical to those for real static obstacles (Figure 11). VIDAR is used to detect stereo obstacles, construct obstacle rectangles, and calculate their width. After the ego-vehicle and obstacle have been moved, the width of the obstacle rectangle is recalculated and then the obstacle height is solved using the triangle similarity principle.

$$\frac{2f^2 + 2y_2^2 + x_3^2 + x_4^2 - \mu^2(x_4 - x_3)^2}{\sqrt{f^2 + x_3^2 + y_2^2}\sqrt{f^2 + x_4^2 + y_2^2}} = \frac{2(h - h_v)^2 + d_3^2 + d_4^2 - w^2}{\sqrt{(h - h_v)^2 + d_3^2}\sqrt{(h - h_v)^2 + d_4^2}}. \quad (7)$$

3.4.2. Moving Pseudo-Obstacle Identification. The steps for identifying moving pseudo-obstacles are identical to those for static pseudo-obstacles (Figure 12): detecting stereo obstacles with VIDAR, determining the pseudo-obstacle

rectangle, and calculating the pseudo-obstacle rectangle's width. Following the ego-vehicle and pseudo-obstacle movement, the height of the pseudo-obstacle is calculated using the width of the pseudo-obstacle's imaging point.

$$\frac{2f^2 + 2y_2^2 + x_3^2 + x_4^2 - \mu^2(x_4 - x_3)^2}{\sqrt{f^2 + x_3^2 + y_2^2}\sqrt{f^2 + x_4^2 + y_2^2}} = \frac{2(h - h_v)^2 + d_3^2 + d_4^2 - (h/h + h_v)^2 w^2}{\sqrt{(h - h_v)^2 + d_3^2}\sqrt{(h - h_v)^2 + d_4^2}}. \quad (8)$$

3.5. Removal Model of Pseudo-Obstacles. The ego-vehicle movement assesses the obstacle's authenticity. Using (3) and (5), the widths of obstacles and pseudo-obstacles when the vehicle moves for the first time are calculated. When the

vehicle resumes motion, the width and height of the obstacle and pseudo-obstacle are also calculated using (4) and (6), and the heights of the obstacle and pseudo-obstacle are determined by their widths. Compared with the calculated

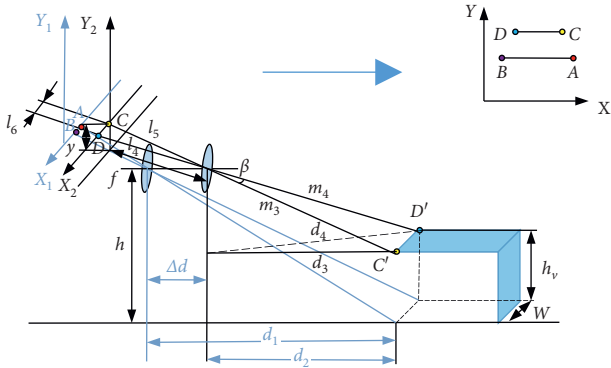


FIGURE 8: Schematic diagram of static pseudo-obstacles after the ego-vehicle moved.

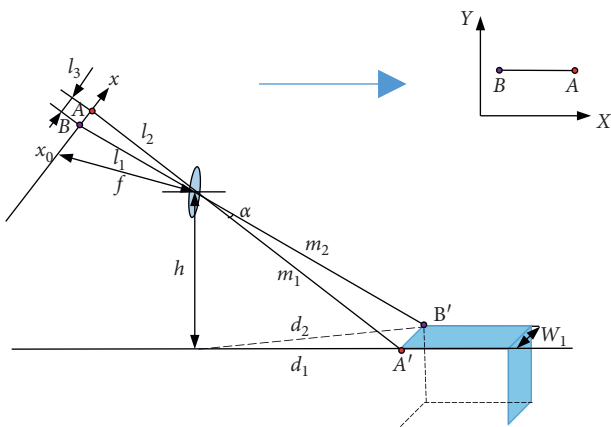


FIGURE 9: Schematic diagram of static pseudo-obstacle.

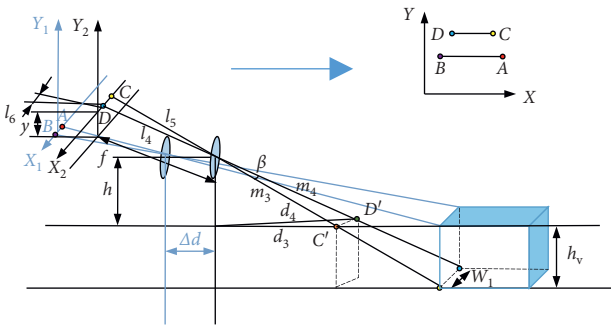


FIGURE 10: Schematic diagram of static pseudo-obstacle after the ego-vehicle moved.

results for obstacles, detected obstacles are those with a similar height.

The process of obstacle identification is as follows:

- (1) Confirm stereo unknown obstacles. Machine learning is used to identify known obstacles, obtain images after removing the known obstacles, and then screen out stereo unknown obstacles using VIDAR's obstacle detection principle.
- (2) Construct an obstacle rectangle. To construct a rectangle, locate the object points on the obstacle that

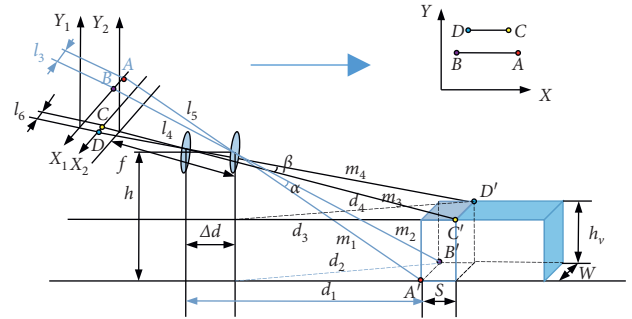


FIGURE 11: Schematic diagram of moving obstacle.

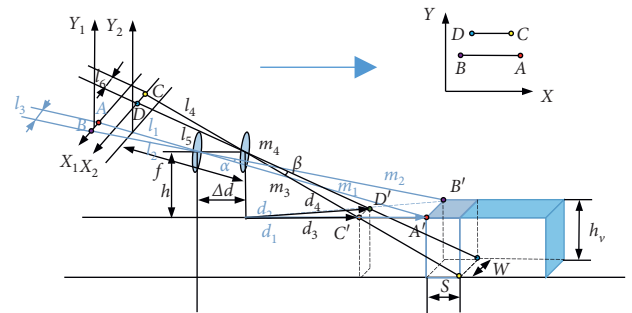


FIGURE 12: Schematic diagram of moving pseudo-obstacle after the ego-vehicle moved.

are the furthest apart in the horizontal and vertical directions (Figure 13 and Figure 14).

- (3) Calculate the horizontal distance. Determine the horizontal distance from the object point to the camera according to VIDAR.
- (4) Identify obstacles. First, calculate the width W of the obstacle rectangle. Second, determine the relationship between the height h_v and width W through the triangle similarity principle.
- (5) Calculate the rectangular height of real and pseudo-obstacles by using the same width when vehicles and obstacles move. Calculate the height value twice, compare the two height values, and determine the identified obstacle.

The overall flow of obstacle identification is shown in Figure 15.

4. Obstacle Identification Experiment and Effect Analysis

We analyze the identification effect of the VM and improved VM in two environments. On the movable platform, the experimental equipment, including the camera unit and IMU, is installed (Figure 16(a)). A scale model of the vehicle is used to simulate a specific obstacle. To simulate the unknown obstacle, a beverage bottle cap is used (Figure 16(b)). The polished paper is used to create a reflection of the road (Figure 16(c)). The camera's video captured at a frame rate of 20 fps is utilized to generate an image sequence, and then

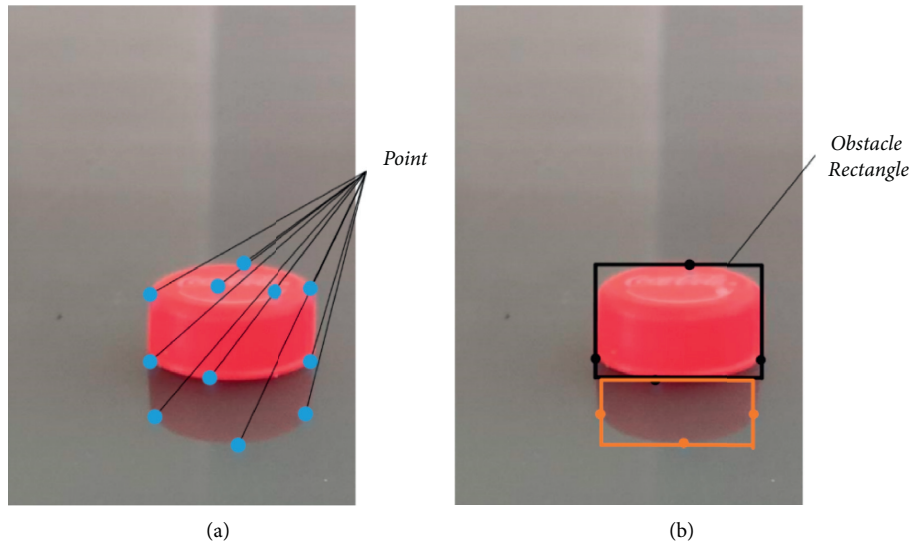


FIGURE 13: The first image. (a) The feature points. (b) Construction of the obstacle rectangle by the feature points.

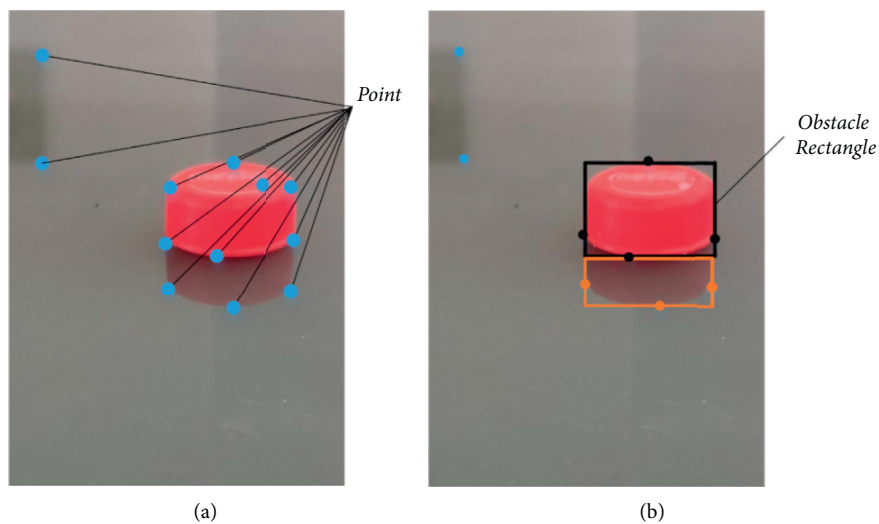


FIGURE 14: The second image. (a) The feature points. (b) Construction of the obstacle rectangle by the feature points.

obstacle detection on the generated image sequence is performed.

4.1. Improved VIDAR and Improved VM Simulation Experiments. A beverage bottle cap is used as an unknown obstacle, and the angular acceleration and acceleration of the ego-vehicle are obtained from the IMU installed in the ego-vehicle. The quaternion method is used to solve the camera attitude and update the camera pitch angle. The image is processed by a fast image region matching method based on MSER. Acceleration is used to calculate the horizontal distance between the vehicle and the obstacles. The height of the obstacle rectangle by keeping the actual width constant during the vehicles and obstacles movement is calculated, the authenticity of the identified obstacles by keeping the

height unchanged is confirmed, and the real obstacles are marked.

The previous image and latter image are used to judge whether the height of the obstacle rectangle has changed (Figure 17).

In the VM and improved VM comparison experiments, the faster RCNN is used to identify known obstacles and to identify known obstacles as background, while VIDAR and improved VIDAR are used to perform secondary detection on the background-removed image to identify unknown obstacles (Figure 18, Figure 19, and Figure 20).

The detection of unknown obstacles in this paper is shown in Figure 19 and Figure 20. While VIDAR in the VM is capable of identifying the bottle cap, the cap's reflection is also detected as an obstacle, resulting in low obstacle identification accuracy. When the improved VIDAR is used

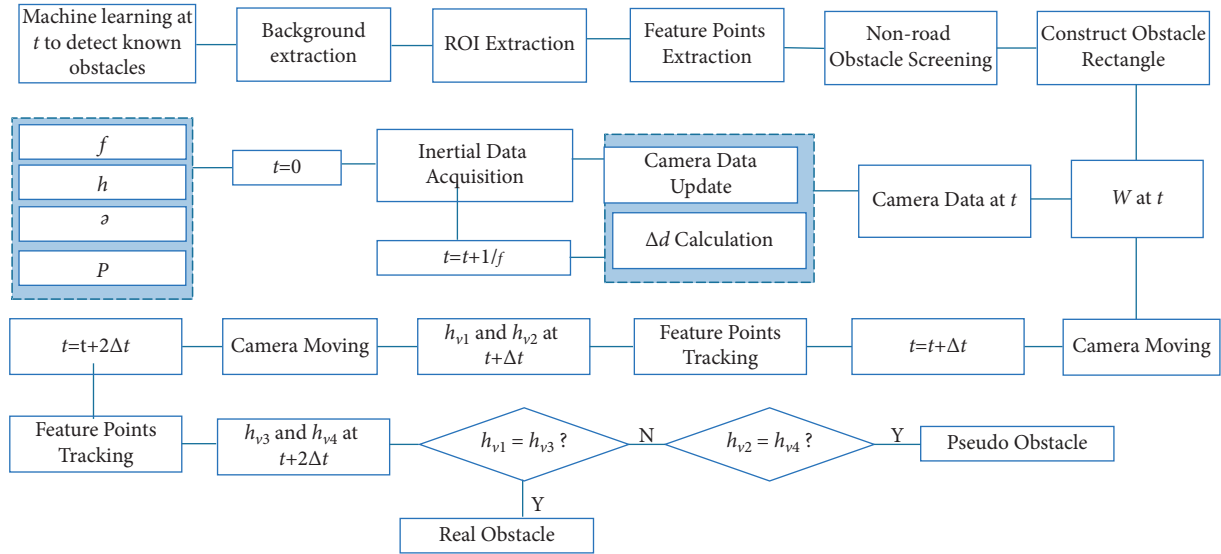


FIGURE 15: Improved VM pseudo-obstacle identification flowchart.

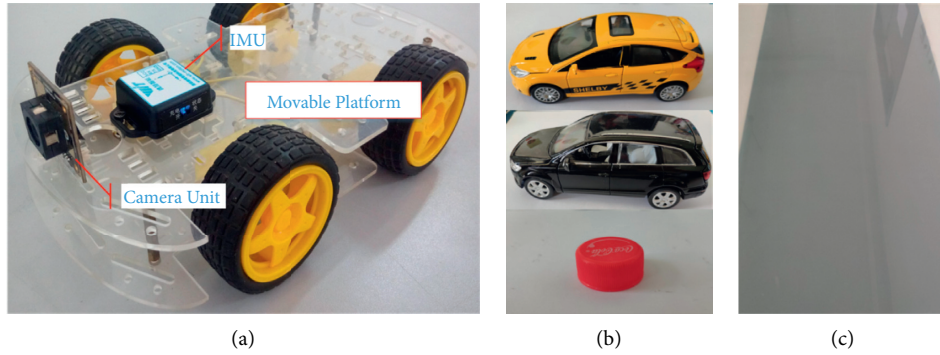


FIGURE 16: Experimental equipment. (a) Mobile platform, camera unit, IMU. (b) Vehicle scaled model (known obstacle) and unknown obstacle model. (c) Polished paper simulates reflection road.

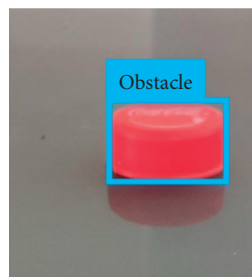


FIGURE 17: Detection result of bottle cap.

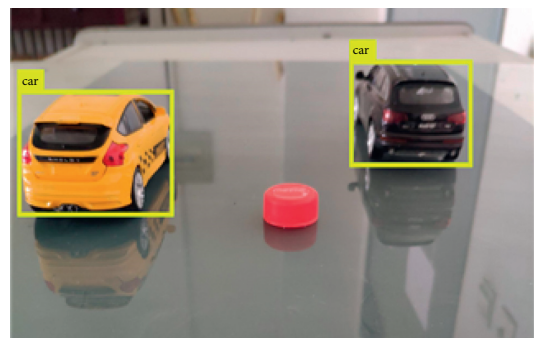


FIGURE 18: Faster RCNN identification effect diagram.

to detect unknown obstacles, the obstacles in the reflection without height can be eliminated, compensating for the unknown obstacles being misdetected in the reflective environment. As a result, the improved VM detects obstacles more precisely than the baseline VM.

4.2. Analysis of the Identification Result of Improved VM and Improved VIDAR. In the experimental test, a pure electric vehicle is used as the test vehicle (Figure 21). A

MV-VDF300SC industrial digital camera is used as a monocular vision sensor. This model camera adopts the USB 2.0 standard interface and has a high resolution, precision, and clarity. The camera’s performance parameters are listed in Table 1. The camera is installed at the height of 1.60 m and collects real-time environmental data (we only used the left camera). The HEC295 IMU is mounted on the bottom of the test vehicle and is used to locate and read the vehicle’s

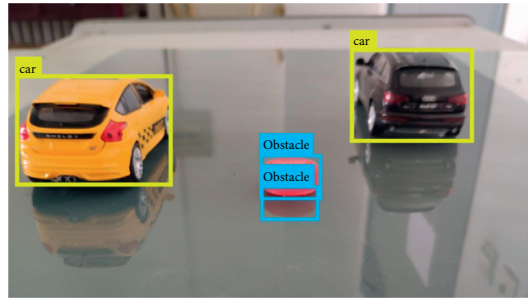


FIGURE 19: VM identification effect diagram.

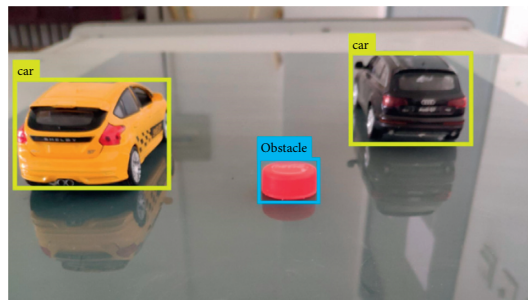


FIGURE 20: Improved VM identification effect diagram.

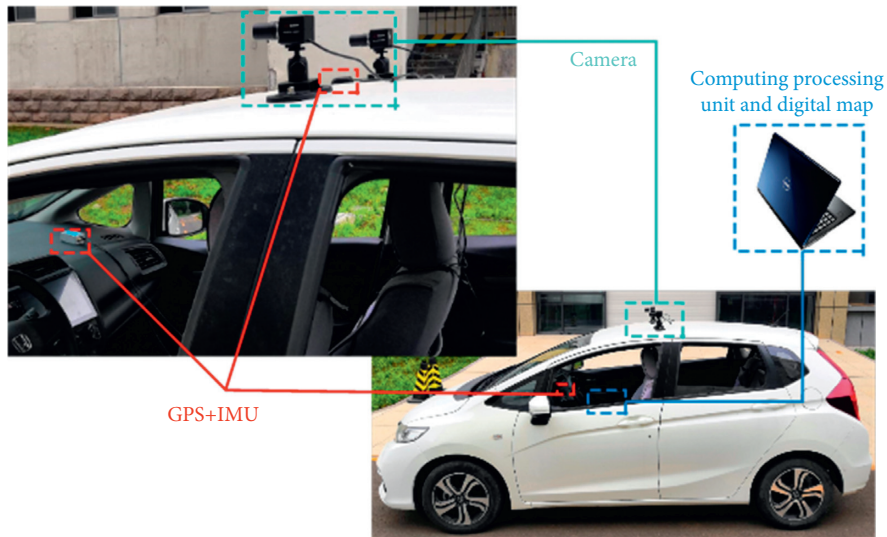


FIGURE 21: Schematic diagram of the test vehicle.

TABLE 1: Some performance parameters of the MV-VDF300SC camera.

MV-VDF300SC			
Highest resolution	2048 * 1536	Power requirements (V)	5
Output color	Color	Power consumption (W)	Rated <5
Frame rate (fps)	12	Operating temperature (°C)	0-60
Output method	USB 2.0	Dimensions (mm)	43.3 * 29 * 29

motion status in real time. GPS is used to determine a precise location. Digital maps are utilized to obtain precise road data, such as distance and slope. The computing unit is used

to perform real-time data processing. In the process of calculation, multisensor data processing is the combination and processing of multisource information, which is rather

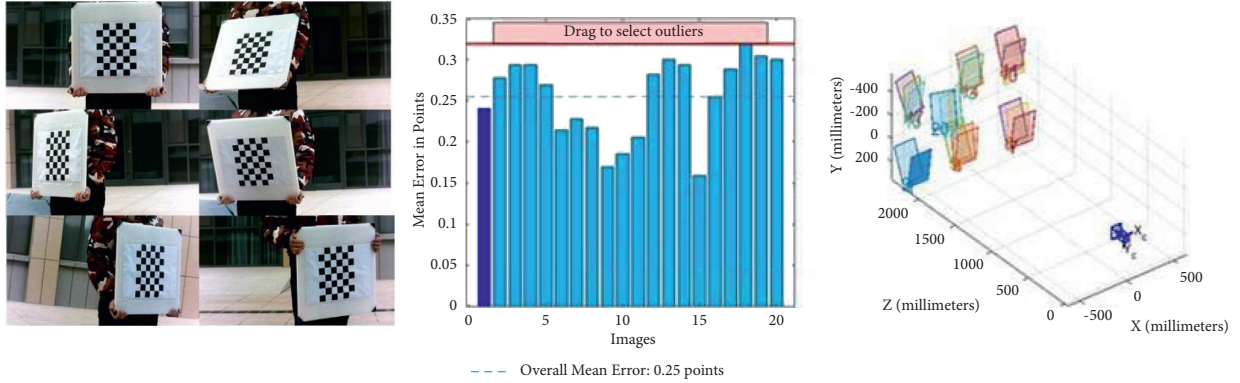


FIGURE 22: The camera calibration result.

TABLE 2: Calibration results of camera external parameters.

External parameter type	Parameter size
Pitch angle	1.25
Yaw angle	3.65
Rotation angle	2.45

complicated. Fuzzy logic can deal with complex systems [29]. It can coordinate and combine the acquired information to improve the efficiency of the system and effectively deal with the knowledge acquired in the scene.

Accurate calibration of camera parameters was a prerequisite for the whole experiment and is a very important task for obstacle detection methods. In this paper, Zhang Zhengyou's camera calibration method was adopted to calibrate the DaYing camera. First, the camera was fixed to capture images of a checkerboard at different positions and angles. Then, key points of the checkerboard were selected and used to establish a relationship equation. Finally, the internal parameter calibration was realized. The camera calibration process and result are shown in Figure 22.

Camera distortion includes radial distortion, thin lens distortion, and centrifugal distortion. The superposition of the three kinds of distortion results in a nonlinear distortion, the model of which can be expressed in the image coordinate system as follows:

$$\begin{cases} \delta_x(x, y) = s_1x(x^2 + y^2) + 2p_1xy + p_2y^3 + k_1x(x^2 + y^2) \\ \delta_y(x, y) = s_2y(x^2 + y^2) + 2p_2xy + p_1y^3 + k_1x(x^2 + y^2) \end{cases} \quad (9)$$

where s_1 and s_2 are the centrifugal distortion coefficients; k_1 and k_2 are the radial distortion coefficients, and p_1 and p_2 are the distortion coefficients of thin lenses.

Because the centrifugal distortion of the camera is not considered in this paper, the internal reference matrix of the camera can be expressed as shown in

$$M = \begin{bmatrix} 5.9774e + 03 & 0 & 949.8843 \\ 0 & 5.9880e + 03 & 357.0539 \\ 0 & 0 & 1 \end{bmatrix}. \quad (10)$$

The calibration of the camera's external parameters can be calculated by taking the edge object points of lane lines. The calibration results are shown in Table 2.

Due to a lack of reflection road images in the public data set and the fact that different camera parameters would affect range accuracy, we created a VIDAR-Reflection Road database (Figure 23) with a total of 2000 images. The MV-VDF300SC camera unit was used to record the experiment in its natural environment. The test roads were Xuezhai Road and Jiefang East Road in Jinan, Shandong Province, and traffic environment images were collected during rainy days from 10:00 to 11:00 and 19:00 to 20:00.

Figure 24 depicts the identification result for the two images. The VM and improved VM accuracy are compared by counting the number of TP, FP, TN, and FN obstacles in each image frame. Let a be an obstacle that is correctly identified as a positive example; b be an obstacle that is incorrectly identified as a positive example; c be an obstacle that is correctly identified as a negative example; and d be an obstacle that is incorrectly identified as a negative example. Then, $TP = \sum_{i=1}^n a_i$, $FP = \sum_{i=1}^n b_i$, $TN = \sum_{i=1}^n c_i$, and $FN = \sum_{i=1}^n d_i$. The comparison of the detection effects of VM and improved VM in the reflection environment is shown in Table 3.

In the results' analysis, accuracy (A), recall (R), and precision (P) were used as evaluation indices for the two obstacle detection methods, calculated through

$$A = \frac{TP + TN}{TP + TN + FP + FN}, \quad (11)$$

$$R = \frac{TP}{TP + FN}, \quad (12)$$

$$P = \frac{TP}{TP + FP}. \quad (13)$$

The accuracy, recall, and precision of the method proposed in this paper are shown in Table 4.

As demonstrated by the experimental results in Table 4, the accuracy of obstacle identification is increased when using the improved VM for obstacle identification in a reflective environment. Due to the weather and other factors, there are times when misidentification and missed



FIGURE 23: A partial sample of the VIDAR-Reflection Road data set.



FIGURE 24: Comparison of detection results of VM and improved VM in an environment with reflection.

identification occur during the experiment. However, the improved method proposed in this paper improves obstacle identification accuracy.

Additionally, we compared our method's detection accuracy to other commonly used target detection methods. Table 5 summarizes the detection results. It is obvious that the proposed obstacle detection method outperforms state-of-the-art methods in terms of accuracy.

The term "real time" refers to the processing of each image frame collected over time. In terms of detection speed, 2000 images were processed using improved VIDAR, VM, VIDAR, VM, and YOLO v5. Table 6 summarizes the average detection times for the five identification methods.

As shown in Table 6, an improved VM takes longer to determine the authenticity of obstacles than a VM. Similarly, improved VIDAR requires more time to determine the

TABLE 3: Comparison of the detection effect of VM and improved VM in an environment with reflection.

	Input	TP	FP	TN	FN
VM	7356	6591	365	279	392
Improved VM	7356	6832	224	252	243

TABLE 4: Evaluation indices of the proposed method.

Detection method	A (%)	R (%)	P (%)
VM	90.05	94.38	94.63
Improved VM	93.81	96.56	96.82

TABLE 5: Evaluation indices of advanced target detection methods and the proposed method.

Detection method	A (%)
Fast RCNN	76.53
SSD	71.28
Fast YOLO	78.95
SSD 300	73.96
YOLO v5s	81.73
Proposed method	90.36

TABLE 6: Evaluation indices of detections of YOLO v5s and proposed method.

	Improved VIDAR	VIDAR	Improved VM	VM	YOLO v5
Identification time (s)	0.627	0.538	0.372	0.296	0.273

authenticity of obstacles when compared with VIDAR. Due to the fewer feature points, improved VM detects faster than improved VIDAR, and less time is required. As a result, using an improved VM for obstacle detection takes not only advantage of machine learning's speed but also improves identification accuracy.

5. Conclusion

This paper first proposes an improved VIDAR method based on VIDAR and then combines machine learning to propose an improved method for VM obstacle identification. On the basis of machine learning to detect known obstacles, VIDAR is used to determine whether there is an obstacle with height by calculating the position of road imaging points, the obstacle rectangle is determined for nonroad obstacles, and then the obstacle height (including real obstacles and pseudo-obstacles) is calculated by using the obstacle imaging points of two frames before and after the vehicle moves. By calculating the height after moving again (including real obstacles and pseudo-obstacles), the two heights are compared to determine the authenticity of the obstacle, so as to realize the obstacle detection. This paper aims to show the effect of obstacle detection using improved VM in the environment with reflection. The experimental results indicate that when compared with VM, the improved VM method for obstacle detection is more accurate in a reflective environment. Because the method proposed in this paper

needs a lot of calculations, improving the efficiency of the proposed method will be the next research direction. In addition, obstacle detection is a prerequisite for obstacle avoidance, and an improved obstacle avoidance method is also a future research direction.

Data Availability

Data are available on request to the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 51905320, the China Postdoctoral Science Foundation under Grants 2018M632696 and 2018M642684, the Shandong Key R and D Plan Project under Grant 2019GGX104066, and SDUT and Zibo City Integration Development Project under Grant 2017ZBXC133.

References

- [1] D. Gusland, B. Torvik, E. Finden, F. Gulbrandsen, and R. Smestad, "Imaging radar for navigation and surveillance on an autonomous unmanned ground vehicle capable of

- identifying obstacles obscured by vegetation,” in *Proceedings of the 2019 IEEE Radar Conference (RadarConf)*, pp. 1–6, Boston, MA, USA, 2019.
- [2] M. P. Muresan, S. Nedeveschi, and I. Giosan, “Real-time object identification using a sparse 4-layer LIDAR,” in *Proceedings of the 2017 13th IEEE International Conference on Intelligent Computer Communication and Processing (ICCP)*, pp. 317–322, Cluj-Napoca, 2017.
 - [3] M. Cho, “A study on the obstacle identification for autonomous driving RC car using LiDAR and thermal infrared camera,” in *Proceedings of the 2019 Eleventh International Conference on Ubiquitous and Future Networks (ICUFN)*, pp. 544–546, Zagreb, Croatia, 2019.
 - [4] S. Wang and X. Li, “A real-time monocular vision-based obstacle identification,” in *Proceedings of the 2020 6th International Conference on Control, Automation and Robotics (ICCAR)*, pp. 695–699, Singapore, 2020.
 - [5] W. Song, Y. Yang, M. Fu, Y. Li, and M. Wang, “Lane identification and classification for forward collision warning system based on stereo vision,” *IEEE Sensors Journal*, vol. 12, pp. 5151–5163, 2018.
 - [6] X. Yi, G. Song, and T. Derong, “Fast road obstacle identification method based on maximally stable extremal regions,” *International Journal of Advanced Robotic Systems*, vol. 15, no. 1, Article ID 1729881418759118, 2018.
 - [7] A. Sharif Razavian, H. Azizpour, and J. Sullivan, “CNN features off-the-shelf: an astounding baseline for recognition,” *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 806–813, 2014.
 - [8] S. Y. Wang, O. Wang, and R. Zhang, “Cnn-generated images are surprisingly easy to spot for now,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 8695–8704, Seattle, WA, USA, June 2020.
 - [9] P. Arena, M. Bucolo, S. Fazzino, L. Fortuna, and M. Frasca, “The CNN paradigm: shapes and complexity,” *International Journal of Bifurcation and Chaos*, vol. 15, no. 07, pp. 2063–2090, 2005.
 - [10] A. Shustanov and P. Yakimov, “CNN design for real-time traffic sign recognition,” *Procedia Engineering*, vol. 201, pp. 718–725, 2017.
 - [11] X. Zhu, Z. Li, X.-Y. Zhang, P. Li, Z. Xue, and L. Wang, “Deep convolutional representations and kernel extreme learning machines for image classification,” *Multimedia Tools and Applications*, vol. 78, no. 20, pp. 29271–29290, 2019.
 - [12] K. Wang, F. Yan, B. Zou, L. Tang, Q. Yuan, and C. Lv, “Occlusion-free road segmentation leveraging semantics for autonomous vehicles,” *Sensors*, vol. 19, no. 21, p. 4711, 2019.
 - [13] G. D. Guo and N. Zhang, “A survey on deep learning based face identification,” *Computer Vision and Image Understanding*, vol. 189, Article ID 102805, 2019.
 - [14] W. K. Jia, Y. Y. Tian, R. Luo, Z. H. Zhang, J. Lian, and Y. J. Zheng, “Identification and segmentation of overlapped fruits based on optimized mask R-CNN application in apple harvesting robot,” *Computers and Electronics in Agriculture*, vol. 172, Article ID 105380, 2020.
 - [15] S. H. Wang, J. D. Sun, I. Mehmood, C. C. Pan, Y. Chen, and Y. D. Zhang, “Cerebral micro-bleeding identification based on a nine-layer convolutional neural network with stochastic pooling,” *Concurrency and Computation: Practice and Experience*, vol. 1, no. 2, Article ID e5130, 2020.
 - [16] B. Saritürk, B. Bayram, Z. Duran, and D. Z. Seker, “Feature extraction from satellite images using segnet and fully convolutional networks (FCN),” *International Journal of Electronic Governance*, vol. 3, no. 5, pp. 138–143, 2020.
 - [17] K. H. Lin, H. M. Zhao, J. J. Lv et al., “Face identification and segmentation based on improved mask R-CNN,” *Discrete Movings in Nature and Society*, vol. 2020, Article ID 9242917, 2020.
 - [18] Y. Tian, G. Yang, Z. Wang, E. Li, and Z. Liang, “Instance segmentation of apple flowers using the improved mask R-CNN model,” *Biosystems Engineering*, vol. 193, pp. 264–278, 2020.
 - [19] X. X. Zhang and X. Zhu, “Moving vehicle identification in aerial infrared image sequences via fast image registration and improved YOLOv3 network,” *International Journal of Remote Sensing*, vol. 11, no. 41, pp. 4312–4335, 2020.
 - [20] T. Haas, C. Schubert, M. Eickhoff, and H. Pfeifer, “BubCNN: bubble identification using Faster RCNN and shape regression network,” *Chemical Engineering Science*, vol. 216, Article ID 115467, 2020.
 - [21] M. Zhang, W. Zhao, and X. Li, “Shadow detection of moving objects in traffic monitoring video,” vol. 9, pp. 1983–1987, in *Proceedings of the 2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, vol. 9, IEEE, Chongqing, China, December 2020.
 - [22] D. J. R. Del Carmen and R. D. Cajote, “Assessment of vision-based vehicle tracking for traffic monitoring applications,” in *Proceedings of the 2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pp. 2014–2021, IEEE, Honolulu, Hawaii, November 2018.
 - [23] S. H. Shaikh, K. Saeed, and N. Chaki, “Moving object detection using background subtraction,” *Moving Object Detection Using Background Subtraction*, Springer, Cham, pp. 15–23, 2014.
 - [24] R. Chabra, J. E. Lenssen, E. Ilg et al., “Deep local shapes: learning local sdf priors for detailed 3d reconstruction,” in *Computer Vision - ECCV 2020*, pp. 608–625, Springer, Cham, 2020.
 - [25] J. Park and B. C. Song, “Night-time vehicle identification using low exposure video enhancement and lamp identification,” in *Proceedings of the 2016 International Conference on Electronics, Information, and Communications (ICEIC)*, pp. 1–2, Da Nang, Danang, Vietnam, January 2016.
 - [26] H. P. Lin, P. H. Liao, and Y. L. Chang, “Long-distance vehicle identification algorithm at night for driving assistance,” in *Proceedings of the 3rd IEEE International Conference on Intelligent Transportation Engineering*, pp. 296–300, Singapore, September 2018.
 - [27] X. Dai, D. Liu, L. Yang, and Y. Liu, “Research on headlight technology of night vehicle intelligent identification based on Hough transform,” in *Proceedings of the 2019 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, pp. 49–52, Changsha, China, 2019.
 - [28] T. S. Kavya, E. Tsogtbaatar, Y. Jang, and S. Cho, “Night-time vehicle identification based on brake/tail light color,” in *Proceedings of the 2018 International SoC Design Conference (ISOCC)*, pp. 206–207, Daegu, Korea (South), 2018.
 - [29] M. Bucolo, L. Fortuna, and M. LaRosa, “Complex dynamics through fuzzy chains,” *IEEE Transactions on Fuzzy Systems*, vol. 12, no. 3, pp. 289–295, 2004.