



OPEN

Autonomous navigation and collision prediction of port channel based on computer vision and lidar

Zhan Zhang^{1✉}, NanWu Yang² & YiJian Yang²

This study aims to enhance the safety and efficiency of port navigation by reducing ship collision accidents, minimizing environmental risks, and optimizing waterways to increase port throughput. Initially, a three-dimensional map of the port's waterway, including data on water depth, rocks, and obstacles, is generated through laser radar scanning. Visual perception technology is adopted to process and identify the data for environmental awareness. Single Shot MultiBox Detector (SSD) is utilized to position ships and obstacles, while point cloud data create a comprehensive three-dimensional map. In order to improve the optimal navigation approach of the Rapidly-Exploring Random Tree (RRT), an artificial potential field method is employed. Additionally, the collision prediction model utilizes K-Means clustering to enhance the Faster R-CNN algorithm for predicting the paths of other ships and obstacles. The results indicate that the RRT enhanced by the artificial potential field method reduces the average path length (from 500 to 430 m), average time consumption (from 30 to 22 s), and maximum collision risk (from 15 to 8%). Moreover, the accuracy, recall rate, and F1 score of the K-Means + Faster R-CNN collision prediction model reach 92%, 88%, and 90%, respectively, outperforming other models. Overall, these findings underscore the substantial advantages of the proposed enhanced algorithm in autonomous navigation and collision prediction in port waterways.

Keywords Port channel, Autonomous navigation, Collision prediction, Computer vision, Laser radar

Port channel is the key hub of global trade, so it is very important to ensure the safe navigation and collision prediction of ships. Port navigation is highly complex, influenced by maritime traffic density and meteorological conditions, and there are various obstacles¹. The traditional navigation system relies on radar, Global Positioning System (GPS) and other technologies, but there are limitations in the complex port environment, such as difficulty in realizing high-precision perception and accurate collision prediction^{2,3}. In addition, at present, most of the methods for detecting obstacles in port waterways rely on machine vision or radar images, but the shortcomings of using machine vision or radar to detect obstacles alone are very obvious⁴. Compared with radar, machine vision has many advantages, such as a large detection range, rich original information, accurate angle measurement, and strong target classification ability. However, using machine vision sensors alone leads to poor detection results in extreme weather such as fog, which seriously interferes with the effective acquisition of visual images^{5,6}. When radar is used alone for target detection, although the speed of the target can be measured at the same time, it leads to low angular resolution, large signal loss and poor target classification ability. Therefore, no matter which target detection method is chosen, it is not the best choice, which greatly limits the ability of port channel to detect and identify obstacles. And it cannot meet the current needs of port channel obstacle identification^{7,8}.

Firstly, regarding the application of computer vision and laser radar in autonomous navigation and collision prediction in port waterways, compared to traditional technologies like radar and GPS, computer vision and laser radar can offer more detailed and high-resolution environmental perception. Laser radar provides precise distance measurements, while computer vision can perform real-time analysis of visual information around ships, allowing for a more comprehensive understanding of the relationship between ships and their environment.

¹Southwest Research Institute for Hydraulic and Water Transport Engineering, Chongqing Jiaotong University, Chongqing 400016, China. ²Guangxi Beigang Planning and Design Institute Co., Ltd, Nanning 530200, China. ✉email: 990020030767@cqjtu.edu.cn

Moreover, in handling complex scenes and multimodal information, computer vision demonstrates more robust capabilities than sensor technologies. Although laser radar is highly accurate in distance measurement, it may have limitations in certain situations, such as dense fog or intense light exposure. In contrast, computer vision systems, leveraging techniques like deep learning, can learn and adapt to various complex environments, enhancing perceptual performance under adverse weather conditions. In addition, Iran's "Morvarid" project uses positioning radar and detection equipment to research autonomous robots at sea. This also illustrates how radar technology assist in locating collisions within port navigation. Previous research has indicated that radar and laser monitoring technologies may face limitations in resolution and accuracy in certain scenarios. Particularly in complex environments, accurately identifying small-sized or densely packed targets may be challenging, impacting the performance of target detection and tracking. Radar systems may also experience significant performance degradation in adverse weather conditions like rain, snow, or fog due to electromagnetic wave interference. Laser monitoring systems may face limitations in atmospheric scattering and absorption. Moreover, radar and laser monitoring primarily rely on target reflection or scattering of electromagnetic waves, limiting their detection capabilities for non-cooperative targets such as potential threats or stealthy objects. In comparison to past research primarily based on a single sensor, this study achieves a more comprehensive perception of the port environment by integrating radar, laser, and image sensors, and enhancing the global awareness of the navigation system. In contrast to traditional RRT algorithms, this study introduces the artificial potential field method for optimization, generating safer and more effective navigation paths in dynamic waters. The incorporation of real-time decision-making algorithms enables the system to autonomously adjust navigation strategies based on real-time collision predictions. This is relatively uncommon in previous research and provides more flexible navigation strategies for the practical application of autonomous navigation systems.

The motivation of this study is to enhance the safety of port navigation by reducing collision accidents, lowering environmental risks, optimizing waterways, and ultimately increasing port throughput. The study focuses on three main objectives: firstly, an advanced path planning algorithm is employed to consider the dynamic characteristics of vessels and environmental conditions to determine the optimal navigation path. Secondly, a collision prediction model using historical data and real-time sensor information is established to forecast potential collision risks. Lastly, a real-time decision-making algorithm is designed to allow the system to autonomously adjust navigation strategies based on collision prediction results and avoid potential collisions. In order to achieve these goals, visual perception technology is applied to process lidar data and camera images for a better understanding of the port channel environment. Specific applications include point cloud data processing, target detection, and tracking. Considering vessel dynamics, channel complexity, and safety requirements, the artificial potential field method is introduced to enhance the capability of the Rapidly-Exploring Random Tree (RRT) in generating optimal navigation paths. In collision prediction, K-Means clustering improves the Fast Region-based Convolutional Network (Fast R-CNN) algorithm, constructing a robust collision prediction model. This model uses historical data and real-time sensor information for training to predict future trajectories of other vessels and obstacles.

The study utilizes a comprehensive fusion of radar, lidar, and image sensor data. Lidar scanning is employed to obtain a three-dimensional map of the port channel, followed by environmental perception through visual perception technology. This sensor data fusion enhances a comprehensive understanding of the port environment, facilitating more accurate navigation and collision prediction. The introduction of a real-time decision-making algorithm allows the system to autonomously adjust navigation strategies based on collision prediction results, promptly avoiding potential collision events. However, there are technical challenges, such as effectively integrating data from radar, lidar, and image sensors. It needs to overcome heterogeneity, noise issues, and real-time requirements of different sensors. High-precision collision prediction is crucial for avoiding potential collisions, particularly in complex water environments where accurately predicting the future trajectories of other vessels and obstacles is challenging. The study combines advanced image processing and autonomous navigation algorithms to improve the safety and efficiency of port navigation, reducing the risk of collision accidents and safeguarding lives and the environment. The innovation lies in the integration of multisensory data, enhanced path planning, and efficient collision prediction, providing a comprehensive and viable solution for autonomous navigation and collision prediction in port channels. The expected outcome is a substantial improvement in the safety and efficiency of port navigation.

Literature review

The discipline of ship navigation is becoming more and more interested in autonomous navigation technologies. A new autopilot system that connected wave height prediction and ship driving was put forth by Lou et al. in⁹. The wave height may be precisely anticipated using Long Short-Term Memory (LSTM), and the ship can change its path in real-time to always travel in the region with the lowest wave height⁹. Gucma¹⁰ used the technique of computer-simulating ship traffic flow and split the approach channel into one-way and two-way segments in the best possible way. A unique two-stage simulation optimization method was used to find the ideal port entrance, steering pool, and port pool characteristics, and simulation tests were run on a three-dimensional visual ship manoeuvring simulator. The study has established the outer container terminal in Swinnuisi's best specifications, and the terminal's anticipated annual container handling capacity was 1.5 million TEUs. An ocean-going container ship¹⁰ with a length of 400 m and a width of 60 m was anticipated to be operated by the port. In their investigation of a reliable control system for unmanned surface vehicles in urban waterways, Cortes-Vega et al. suggested using a visual odometer to assess the position of the vehicles rather than conventional sensors¹¹. Mansuy et al. simulated the turning manoeuvres of two typical inland river shipping vessels under different hydrometeorological conditions in real time, and proposed a step-by-step method to select the optimized turning pool geometry according to the field conditions. This step-by-step design of the turning pool method can reduce

the real-time simulation required for upgrading the waterway network¹². Sinohara et al. analyzed the specific needs of autonomous ships for external and environmental information in restricted pilotage waters, and put forward targeted technical solutions in Paranagua and Antonina, Brazil¹³. Nzengu et al. analyzed the regulatory framework related to the operation and testing of unmanned Inland Waterway (IWW) ships in Flanders. A three-stage strategy is put forward as a road map for formulating the regulatory framework to adapt to IWW autonomous ships more widely¹⁴.

Computer vision technology and lidar have made remarkable progress in the field of navigation. Hu et al. used lidar to monitor the visibility of sea fog in Beilun area of Zhoushan Port in Ningbo, China, and compared the data of lidar with that of forward scattering visibility sensor. The results showed that the visibility lidar instrument had advantages in sea fog monitoring, and the correlation between lidar instrument data and forward scattering sensor data proved the practicability and potential of lidar in sea fog detection¹⁵. Lu et al. used Automatic Identification System (AIS), video surveillance, laser radar and other intelligent sensing technologies to realize automatic and accurate collection of channel traffic data, and built an integrated platform of existing ship channel traffic monitoring system in Yancheng, which integrated multi-dimensional sensing, fusion processing and statistical analysis¹⁶. Tak et al. monitored the beach width and beach profile along the east coast of Korea by unmanned aerial vehicles and ground-based lidar. The results showed that the plane layout of port facilities concentrated waves and increased the number of sediments northward¹⁷. Hake et al. used multi-sensor system to scan the above-water and underwater port structures, and used Visual Geometry Group19 (VGG19) deep neural network and local abnormal factors to identify the grid network of point clouds on the steel sheet pile wall. The results showed that the accuracy of VGG19 deep neural network was 8.95%¹⁸. Marchel et al. used extended Kalman filter and two-dimensional range bearing to evaluate the positioning accuracy of ships following a constant course and speed in the port approach channel, which showed that the adopted algorithm could be successfully used to plan their deployment to ensure the minimum accuracy requirements of navigation marking service in positioning navigation marks on the port approach channel and under restricted conditions¹⁹.

Channel collision is a serious safety problem, so it is very important to study collision prediction and prevention methods. Upadhyay et al. used the collaborative method of computer vision to track the target according to the specific position of interest in the image. Compared with the actual measurement, the test results on the framework of quadrotor UAV achieved 99% positioning accuracy²⁰. Padmaja et al. proposed a collision warning system for self-driving cars based on a new point-to-pixel multi-sensor data fusion algorithm, and used MobileNet SSD to classify targets. The results showed that the root mean square error and mean absolute error of the proposed fusion algorithm were 2.93 mm and 802.83 mm lower than those of the stereo camera and the two-dimensional lidar sensor respectively²¹. Miao et al. proposed a UAV obstacle identification based on airborne lidar and an improved density-based noisy application spatial clustering algorithm, and the experiments proved the effectiveness of the proposed algorithm in identifying the invading mobile state²². Guan et al. introduced a new multi-model full traffic trajectory data, and measured the fluctuation of pedestrian speed by computer lidar and computer vision respectively. Compared with the data based on computer vision, the current trajectory data based on lidar showed a wider detection range and was less affected by poor lighting conditions²³.

The above literature shows that autonomous navigation technology, computer vision and lidar technology are widely used in the field of navigation, which provides strong support for improving the safety and efficiency of ship navigation. These studies cover wave height prediction, channel optimization, unmanned surface vehicle control, collision prediction and prevention, environmental perception and obstacle detection. However, there are still some challenges to be solved, such as environmental changes and sensor errors, to further improve the performance of the navigation system. This study continues to explore innovative methods to promote the development of navigation field and improve the robustness and applicability of navigation system.

Research methodology

Visual perception technology and port channel environment perception

In port channel perception, point cloud data are obtained from lidar sensors. These data include a large number of discrete three-dimensional points, and each point represents a position in space. Point cloud data need to be processed and analyzed to create a three-dimensional map of the port channel. The steps of processing point cloud data include data acquisition and preprocessing, point cloud segmentation and grouping, and 3D map generation. The original point cloud data is obtained from the lidar sensor, and then the preprocessing steps such as denoising, filtering and coordinate transformation are carried out to prepare the data for subsequent processing. According to the characteristics of point cloud, point cloud data is divided into different objects or features, such as water surface, other ships, shore and obstacles. The processed point cloud data are combined to generate a three-dimensional map of the port channel, including the location information of waterways, port facilities and other targets^{24,25}. Figure 1 shows the three-dimensional map of the port channel.

Finding the whereabouts of other ships and obstructions in the port channel is crucial. Target recognition and tracking technology is used in this study to keep an eye on nearby ships and objects in real time. The Single Shot MultiBox Detector (SSD) technique is used in this procedure. SSD is a convolutional neural network-based target detection technique. To detect objects of various sizes and forms, it employs multi-scale feature maps, and to detect targets of various shapes, it employs multiple anchor frames. By performing a convolution operation on feature maps with various levels, SSD achieves multi-scale target detection²⁶. The SSD structure is shown in Fig. 2.

In Fig. 2, SSD is based on Visual Geometry Group 16 (VGG16) network, and the fully connected layer is modified as convolution layer, and then four convolution layers are added, namely, conv6, conv7, conv8_2, conv9_2, conv10_2, conv11_2, conv4_3, conv7, conv8_2 and conv. Conv10_2 and conv11_2 are used as the detection heads of the network, and the detection results are obtained through the final Non-Maximum Suppression (NMS) non-maximum suppression. SSD starts from Conv4_3 and extracts feature maps. The number



Figure 1. Three-dimensional map of port channel.

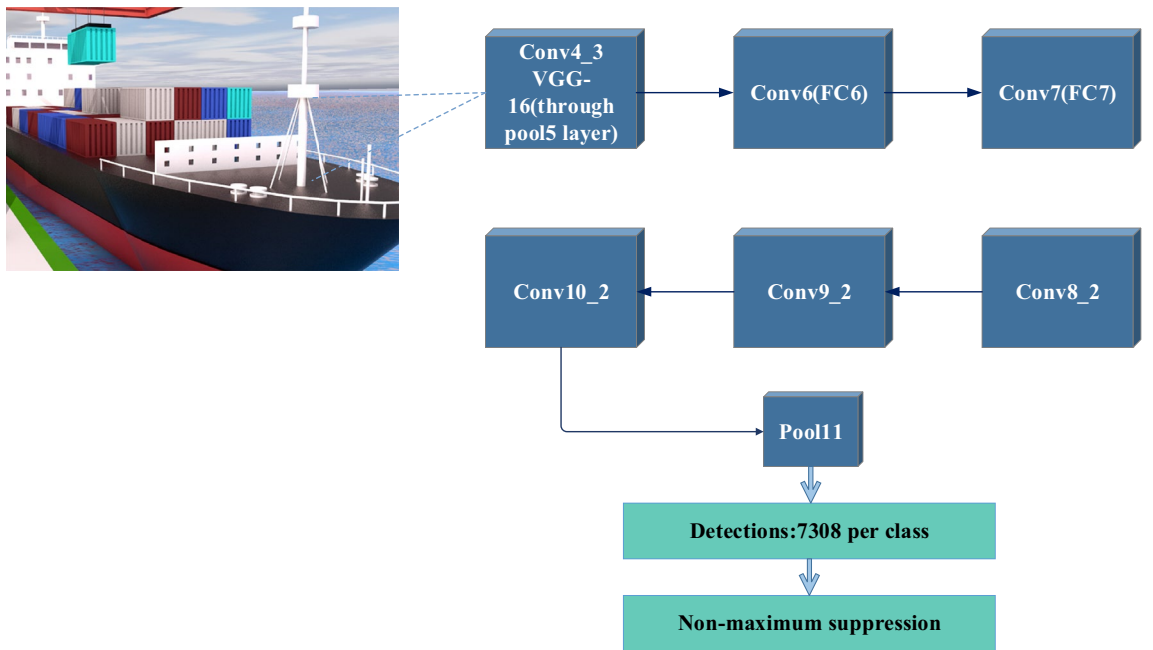


Figure 2. SSD structure.

of prior frames set on each feature map is different, and the setting of prior frames follows the rule of linear increase, that is, the size of feature map decreases and the scale of prior frames increases linearly. Equation (1) shows the setting of prior frames:

$$s_k = s_{min} + \frac{s_{max} - s_{min}}{m - 1} (k - 1), k \in [1, m] \tag{1}$$

In Eq. (1), m is the number of feature graphs, s_k is the ratio of the prior frame size to the picture, and s_{min} and s_{max} are the minimum and maximum values of the ratio respectively. The center point of the prior box of each pixel is distributed in the center of each pixel, and the calculation is shown in Eq. (2):

$$\left(\frac{i + 0.5}{|f_k|}, \frac{j + 0.5}{|f_k|} \right), i, j \in [0, |f_k|] \quad (2)$$

In Eq. (2), $|f_k|$ represents the size of the feature map. In the process of prediction, the model is predicted by using bounding_box. The purpose of bounding_box regression is: given (P_x, P_y, P_w, P_h) , find the mapping f , so that $f(P_x, P_y, P_w, P_h) = (G_x^\wedge, G_y^\wedge, G_w^\wedge, G_h^\wedge)$ can be obtained. G is ground truth and G^\wedge is bounding box. Eqs. (3–6) show the border regression process:

$$G_x^\wedge = P_w \frac{(G_x^\wedge - P_x)}{P_w} + P_x, \quad (3)$$

$$G_y^\wedge = P_h \frac{(G_h^\wedge - P_y)}{P_h} + P_y, \quad (4)$$

$$G_w^\wedge = P_w \exp\left(\log \frac{(G_w^\wedge)}{P_w}\right), \quad (5)$$

$$G_h^\wedge = P_h \exp\left(\log \frac{(G_h^\wedge)}{P_h}\right), \quad (6)$$

In the above equation, $(G_x^\wedge, G_y^\wedge, G_w^\wedge, G_h^\wedge)$ is the position offset and scale transformation of Bounding_box relative to prior_box. In an ideal state, the position offset and scale transformation of bounding_box relative to prior_box is shown in Eqs. (7–10):

$$t_x = \frac{(G_x - P_x)}{P_w}, \quad (7)$$

$$t_y = \frac{(G_y - P_y)}{P_h}, \quad (8)$$

$$t_w = \log \frac{(G_w)}{P_w}, \quad (9)$$

$$t_h = \log \frac{(G_h)}{P_h}, \quad (10)$$

Through the above equation, the Loss function can be obtained, through reducing the loss, the convergent position offset and scale transformation can be finally obtained, and the final predicted bounding_box can be obtained through decoding with prior_box. Equation (11) shows the loss function:

$$\text{Loss}(x, c, l, g) = \frac{1}{N} (L_{\text{conf}}(x, c) + \alpha L_{\text{loc}}(x, l, g)) \quad (11)$$

In Eq. (11), N is the number of positive samples in the prior frame, c is the predicted value of category confidence, l is the predicted value of the position corresponding to bounding_box in the prior frame, g is the position parameter of ground truth, and α is taken as 1 through cross-verification. The position function of $L_{\text{loc}}(x, l, g)$ is shown in Eqs. (12, 13):

$$L_{\text{loc}}(x, l, g) = \sum_{i \in \text{pos}} \sum_{m \in \{cx, cy, w, h\}} x_{ij}^p \text{Smooth}_{L1} \left(l_i^m - \widehat{g}_j^m \right) \quad (12)$$

$$\text{Smooth}_{L1}(x) = \begin{cases} 0.5x^2, & \text{if } |x| < 1 \\ |x| - 0.5, & \text{otherwise} \end{cases} \quad (13)$$

In the above equation, l_i^m and \widehat{g}_j^m are the position parameters after encode, x_{ij}^p is the i th prior_box matching with the j th gt_box, and the category of gt_box is p , with a value of 1, otherwise with a value of 0. The position loss function is only for positive samples. For every prior_box matching gt_box, the difference between the offset and scaling scale of bounding_box and that of gt_box is calculated by using $\text{Smooth}_{L1}(x)$ loss, and the optimization is achieved by reducing its value. L2 regularization in Conv4_3 is used, as shown in Eq. (14):

$$y_i = L_2Norm(x) = \frac{x_i}{\sqrt{\sum_{k=1}^n x_k^2}} \quad (14)$$

The Conv4_3 layer has a different feature scale compared with other layers. L2 regularization technology is used to normalize the feature of each pixel in conv4_3 feature map to 20 to ensure that there is little difference with the following layers. In the process of identifying the positions of other ships and obstacles, SSD algorithm inputs the preprocessed point cloud data and camera images into SSD algorithm. SSD algorithm extracts features from input data through CNN for target detection. The feature map is used with the anchor frame to locate the target and determine the locations of other ships and obstacles. The target tracking algorithm updates the target's position information in real-time for path planning and collision prediction²⁷.

Path planning algorithm and optimal navigation path generation

In the aspect of path planning, considering the dynamic characteristics of the ship, the complexity of the channel environment and the safety requirements, the improved Rapidly-Exploiting Random Trees (RRT) is adopted to generate the best navigation path^{28,29}. Table 1 shows the principle of RRT algorithm.

In Table 1, *RANDOM_STATE()* function generates random points within the set environment, *NEAREST_NEIGHBOR()* function traverses the random tree to find the node closest to the random point. *SELECT_INPUT()* function expands the random tree according to the set value, *NEW_STATE()* function generates x_{new} , *judge* (x_{new}) function determines whether the newly generated node satisfies non holonomic constraints, *T.add_Vertex ()* insert x_{new} , *T.add_Edge ()* adds an edge between x_{near} and x_{new} , do not add new nodes in this loop. Regenerate x in the next loop *New*, and then make a judgment if it belongs to X_{free} , then keep the new node. In summary, after adding x_{new} , when adding a new node, it needs to be judged twice, namely obstacle detection and non-holonomic constraint detection. Only when both meet the requirements can a new node be added³⁰.

However, the initial path generated by RRT algorithm is completely random, and it may not be possible to search for navigation tracks in maps with many obstacles. In this paper, the concept of artificial potential field method is combined, and the generated track is more in line with the requirements of safety and smoothness through the action of repulsion and gravity. By deleting redundant nodes, the efficiency and feasibility of the track are further improved. Figure 3 shows the autonomous navigation process of port channel based on improved RRT algorithm.

Figure 3 shows an updated RRT method that starts from a beginning point, creates additional nodes in a random manner, and tries to connect the new nodes to the preexisting tree structure. The artificial potential field approach is used to expand nodes, and each node is subjected to the attraction of the target point and the repulsion of barriers, ensuring that the node proceeds to the target point along the safest path. Through the process of node expansion, the algorithm gradually generates a series of connected nodes, forming the initial track. In order to make the track smoother and safer, some redundant nodes are deleted by line-of-sight algorithm. The line-of-sight algorithm checks the nodes on the track. If there is no obstacle between two nodes, the intermediate node between them can be deleted. Through the above steps, the improved RRT algorithm of potential field is obtained. This algorithm generates a smooth, safe flight path with fewer nodes, which can be used for navigation and path planning. Finally, the improved potential field RRT algorithm is applied to the three-dimensional map to plan the three-dimensional track of the autonomous navigation process of the port channel.

Collision prediction model and training process

In the autonomous navigation system of port channel, the construction and training process of collision prediction model is very important to ensure that ships can safely avoid collision and plan the best path. In this study,

GENERATE_RRT($x_{init}, K, \Delta t$)
1. T.init(x_{init});
2. For k=1 to K do
3. If ($\ x_{new} - x_{goal}\ < d$)
4. Break;
5. $x_{rand} \leftarrow RANDOM_STATE()$;
6. $x_{near} \leftarrow NEAREST_NEIGHBOR(x_{rand}, T)$;
7. $u \leftarrow SELECT_INPUT(x_{rand}, x_{near})$;
8. $x_{new} \leftarrow NEW_STATE(x_{near}, u, \Delta t)$
9. Judge (x_{new});
10. If (judge (x_{new}) == false and collision free() == false)
11. Continue;
12. T.add_vertex (x_{new});
13. T.add_edge (x_{near}, x_{new}, u);
14. Return T

Table 1. Principle of RRT algorithm.

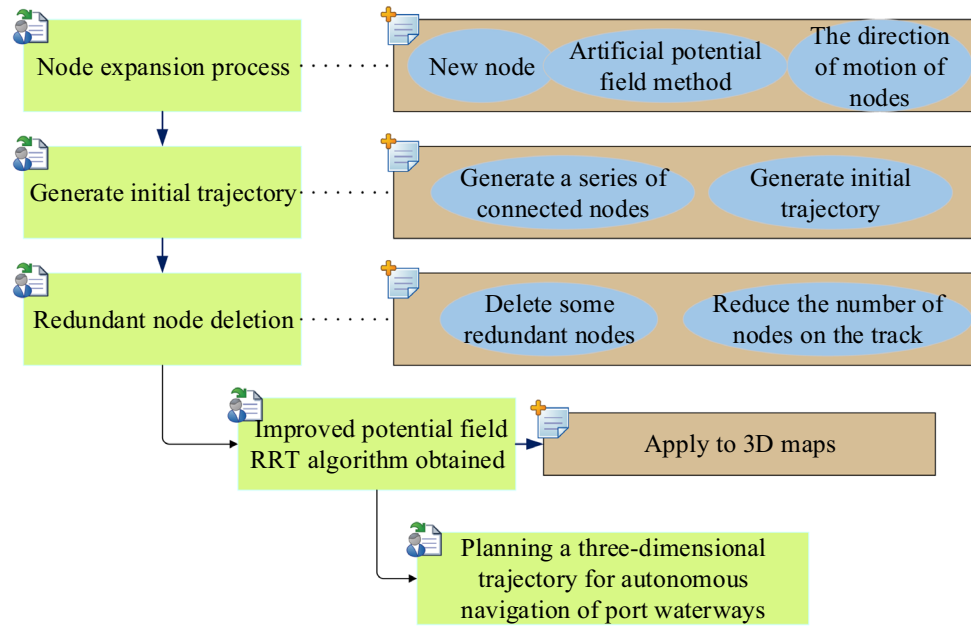


Figure 3. Autonomous navigation process of port channel based on improved RRT algorithm.

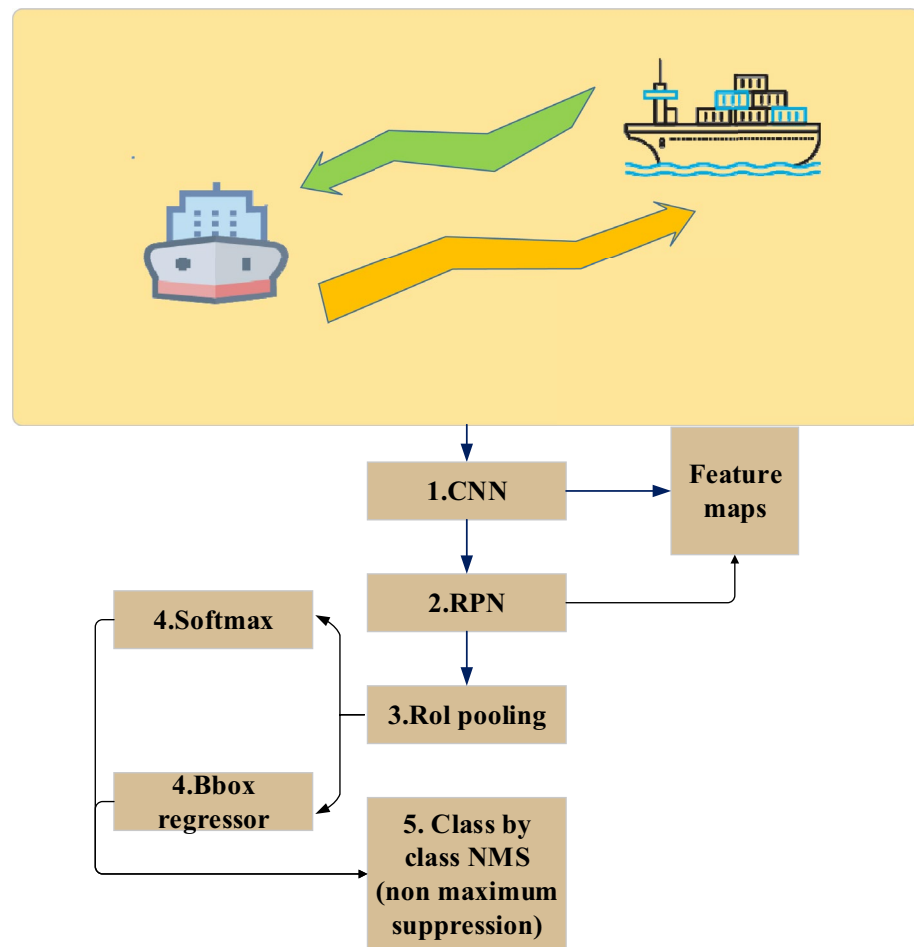


Figure 4. Structure of Faster R-CNN.

a collision prediction model is created using an upgraded FAST Region-based Convolutional Neural Network (R-CNN). A popular deep learning model for target identification and object recognition is Faster R-CNN, which can identify objects in photos and pinpoint their locations. The Faster R-CNN's construction is depicted in Fig. 4.

In Fig. 4, Fast R-CNN consists of two main components: CNN and Region Proposal Network (RPN). CNN is used to extract features from input images. These feature maps contain different levels of information in the image and are used for subsequent target detection tasks. RPN is used to generate candidate regions. It slides the window on the convolution feature map of the backbone network and outputs the suggested target box through classification and regression header. Each suggestion box is accompanied by a candidate box score for subsequent screening. The RoI pooling layer is used to cut and standardize candidate frames of different sizes into feature maps of the same size for input into the subsequent classification and regression head network. These networks receive the characteristic map of RoI pool as input, and carry out target classification and position regression. The classification header is used to determine whether the candidate frame contains the target object, and the regression header is used to adjust the position of the candidate frame. In the training process of Faster R-CNN, the loss calculation of the network is shown in Eqs. (15, 16):

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (15)$$

$$L_{reg}(t_i, t_i^*) = \sum_{i \in \{x, y, w, h\}} smooth_{L1}(t_i - t_i^*) \quad (16)$$

In the above equation, i is the anchors index. p_i is the positive softmax probability. p_i^* is the corresponding GT predict probability. t is the predict bounding box, and t_i^* is the corresponding GT box of the positive anchor. L_{cls} is the softmax loss calculated by RPN_cls_loss layer, which is used to classify anchors as positive and negative network training, and L_{reg} is the smooth L1 loss calculated by RPN_loss_bbox layer, which is used to train the bounding box regression network. In the actual port navigation collision prediction, Faster R-CNN usually needs large-scale tag data for training, but in a specific port channel environment, it may need more model adaptability to adapt to different meteorological conditions, port structures and ship types. In this study, K-Means clustering algorithm is used to improve Faster R-CNN algorithm for port channel collision prediction, and K-Means algorithm is used to cluster lidar data and image data, and the data points are divided into different clusters. These clusters can represent different types of ships and obstacles. In Faster R-CNN, the results of K-Means clustering are used to define the category and location information of target detection. This can improve the accuracy of target detection. Finally, the improved Faster R-CNN model is used for training, so that the model can detect

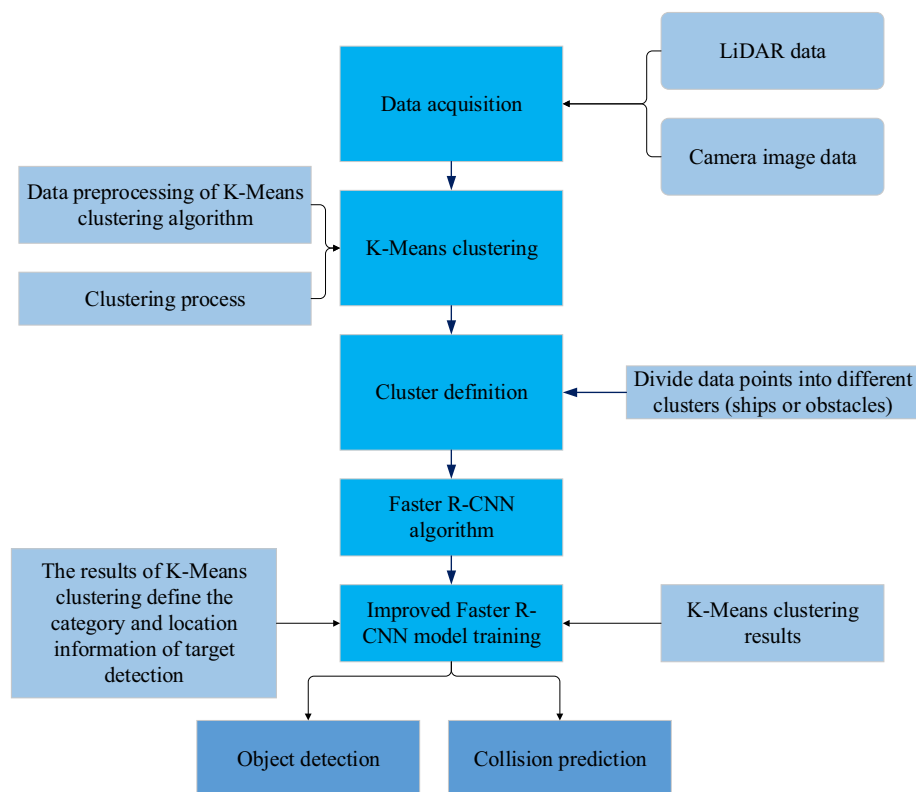


Figure 5. Collision prediction model based on improved Faster R-CNN algorithm.

targets and predict collisions according to the K-Means clustering results. Figure 5 shows the collision prediction model of the improved Faster R-CNN algorithm.

Results and discussion

The result of SSD algorithm to identify the position of ships and obstacles

Figure 6 shows the comparison results between SSD algorithm and actual results in identifying the positions of ships and obstacles. In Fig. 6, in the case of testing data point 1 to data point 3, the performance of using visual perception + lidar + SSD algorithm is better than using lidar only. The relative errors are 2.33%, 2.5% and 3.13% respectively. This shows that SSD algorithm has achieved good results in identifying the positions of ships and obstacles, and its accuracy is relatively higher. With the increase of data points, the relative error increases gradually. In the case of data point 5, the relative error reaches 5.56%. This may be because the performance of visual perception + lidar + SSD algorithm is challenged in a longer distance or in a more complex environment, and the error increases slightly. Generally speaking, visual perception + lidar + SSD algorithm performs well in ship and obstacle location recognition, and its performance is better than that of using lidar only.

The experimental results indicate that the combination of visual perception, lidar, and SSD algorithm performs better in short distances compared to using only lidar. This suggests that the integrated use of multisensor data (visual and lidar) enhances the accuracy of ship and obstacle position identification. The effective fusion of information from multiple sources allows the model to more accurately capture target positions in short distances. Although there is a slight increase in relative errors in some cases, the overall performance remains superior to the scenario using only lidar. Future improvements could involve introducing more training data to enhance generalization performance or adjusting algorithm parameters to balance performance under different distances and environmental conditions.

Improvement of navigation path generated by RRT by artificial potential field method

The outcome of enhancing RRT to generate a navigation path using an artificial potential field approach is shown in Fig. 7. Figure 7 illustrates a considerable improvement in average path length and average time consumption between the RRT improved by the artificial potential field method and the conventional RRT algorithm. Average time consumption decreases from 30 to 22 s, and average journey length decreases from 500 to 430 m. Thus, the enhanced algorithm suggested in this study provides shorter and quicker travel paths and boosts navigational effectiveness. The RRT improved by artificial potential field method also shows obvious advantages in the maximum collision risk. The maximum collision risk is reduced from 15% of the traditional RRT algorithm to 8%. This shows that the algorithm in this study has made remarkable progress in reducing collision risk and improving navigation safety. Compared with Dijkstra algorithm and A* algorithm, RRT improved by artificial potential field method is competitive in average path length and average time consumption, which provides more efficient and safer navigation path planning for autonomous navigation of port channels.

Compared to Dijkstra's algorithm and A* algorithm, the RRT algorithm improved through the artificial potential field method demonstrates competitiveness in both average path length and average time consumption. This implies that the enhanced RRT algorithm is more advantageous in terms of efficiency for navigation path planning compared to traditional algorithms. This provides a more efficient and secure planning path for autonomous navigation in port channels. The RRT algorithm improved through the artificial potential field method exhibits enhancements in various aspects, including path length, time consumption, and collision risk. This comprehensive improvement positions the enhanced algorithm as a viable choice in autonomous navigation systems, particularly in scenarios where rapid, secure, and low collision risk navigation paths are required.

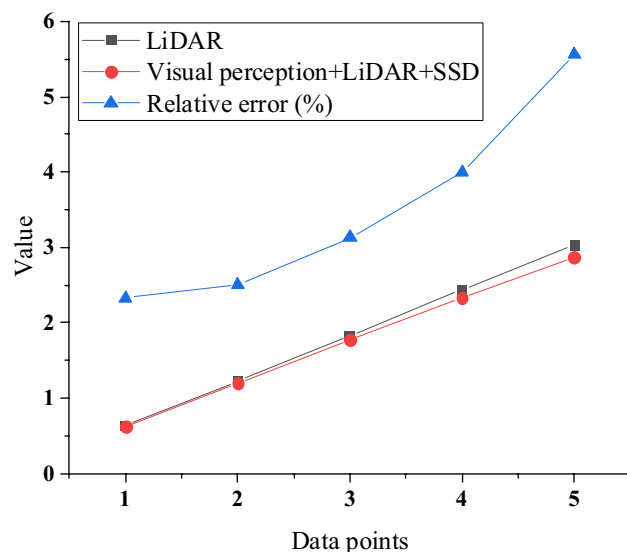


Figure 6. Comparison of SSD algorithm to identify the position of ships and obstacles with the actual results.

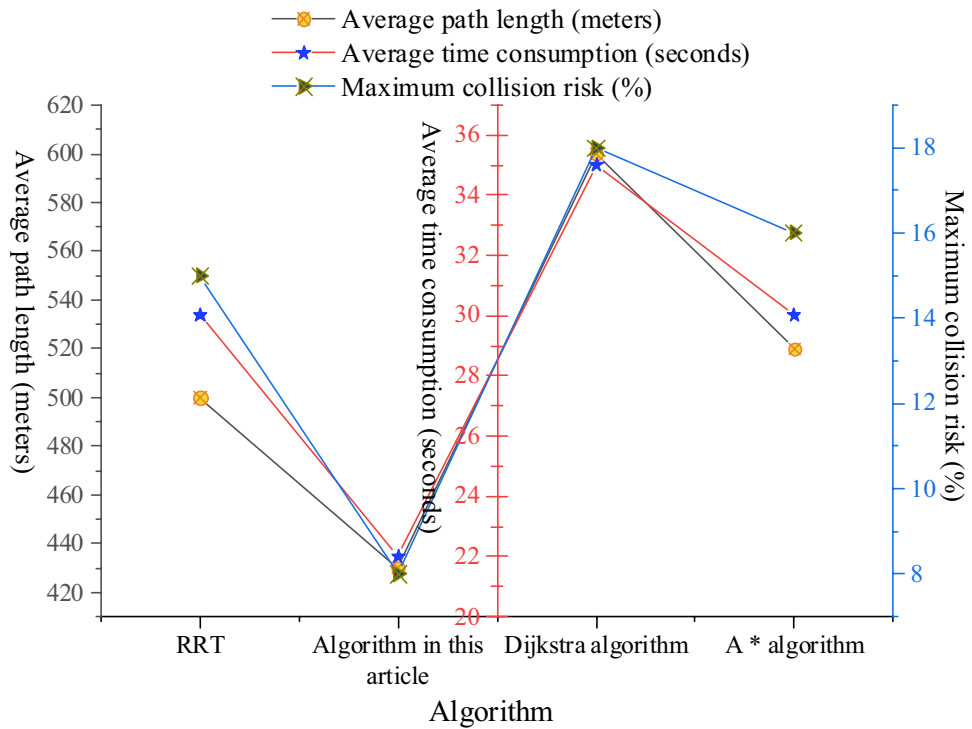


Figure 7. Improvement of navigation path generated by RRT by artificial potential field method.

Performance results of improving the collision prediction model of faster R-CNN based on K-means clustering

The study evaluates the model performance using accuracy, recall, and F1 score. Higher values for accuracy and recall, closer to 1, indicate better precision or recall rates. The F1 score ranges from 0 to 1, with 1 being the maximum value and 0 the minimum. A higher precision and recall are desirable. In the range of 0 to 1, a higher F1 score is preferable. Figure 8 shows the performance results of the collision prediction model improved by Faster R-CNN based on K-Means clustering. It shows that for the “buoy” type obstacle, the accuracy of the model is 95%, the recall rate is 92%, the F1 score is 93%, and the average calculation time is 33 ms. This shows that the model has high performance in detecting and locating buoy obstacles, and has high precision and recall rate. For the “container” type of obstacles, the performance of the model is relatively good, while for the “fishing boat” type of obstacles, the performance of the model is relatively low, with an accuracy of 88%, a recall rate of 85%, a F1 score of 87% and an average calculation time of 50 ms. This shows that the performance of the model

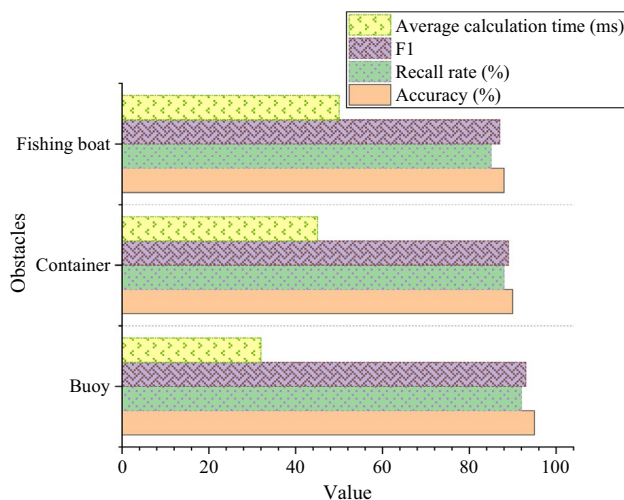


Figure 8. Correlation analysis results between digital platform and corporate social responsibility.

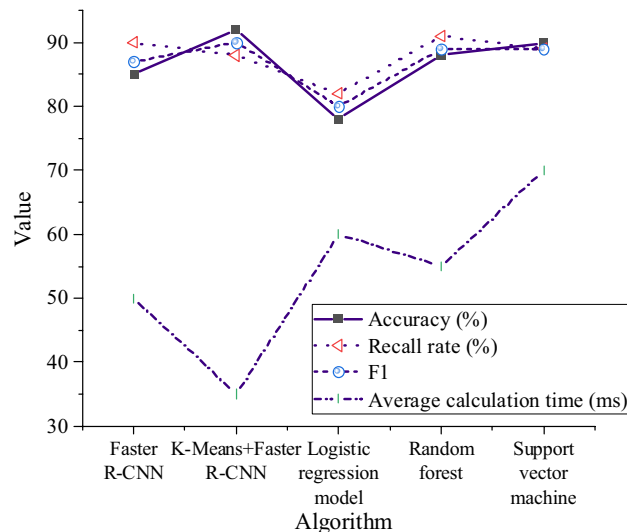


Figure 9. Performance comparison of different collision prediction models.

is relatively low in detecting and locating obstacles of fishing boats. Generally speaking, the improved Faster R-CNN collision prediction model based on K-Means clustering shows different performance levels under different types of obstacles. The model has excellent performance in detecting buoy obstacles.

Figure 9 shows the performance comparison of different collision prediction models. In Fig. 9, the K-Means + Faster R-CNN algorithm proposed in this paper shows significant advantages in accuracy, recall and F1 score, and has higher performance compared with other collision prediction models. Its accuracy is 92%, the recall rate is 88%, and the F1 score is 90%, so it is one of the best models. Random forest model and support vector machine model also show high performance in accuracy, recall and F1 score, but they are slightly lower than K-Means + Faster R-CNN model. The performance of these models is better than that of Faster R-CNN and Logistic regression models. Logistic regression model is relatively low in accuracy, recall and F1 score, and its performance is the worst. Meanwhile, its average calculation time is also long, 60 ms. On the whole, K-Means + Faster R-CNN algorithm has obvious advantages over other collision prediction models in terms of performance index and calculation efficiency, and it is a better collision prediction model.

Taking a comprehensive view, the K-Means + Faster R-CNN algorithm demonstrates superiority in both performance and computational efficiency, making it a more outstanding collision prediction model. While other models perform well, they may not surpass K-Means + Faster R-CNN in certain performance indicators. The experimental results in Figs. 8, 9 indicate that the Faster R-CNN model improved by K-Means clustering exhibits excellent performance under different types of obstacles, particularly showing significant improvement in collision prediction.

Conclusion

The successful integration of computer vision and Lidar technology enables effective perception and analysis of the port channel environment. The improved RRT algorithm, employing the artificial potential field method for path planning, significantly reduces average path length and time consumption, leading to a substantial improvement in navigation efficiency and safety. Additionally, the collision prediction model based on K-Means clustering and Faster R-CNN outperforms other models, demonstrating excellent adaptability to various obstacle scenarios. While the proposed algorithmic enhancements represent a significant breakthrough in the field of autonomous navigation and collision prediction in port channels, acknowledging the limitations of this study is crucial. Future studies should focus on refining path-planning algorithms to adapt to more complex environmental conditions and consider the dynamic characteristics of vessels. Furthermore, there is room for optimization in the collision prediction model to enhance its adaptability to a broader range of obstacles. Although this study marks significant progress, the identified constraints highlight areas for future exploration and improvement. Ongoing research can further refine the proposed algorithms to ensure their applicability to various dynamic maritime scenarios, ultimately advancing the field of autonomous navigation and collision prediction in port channels.

Data availability

All data generated or analysed during this study are included in this published article [and its [supplementary information files](#)].

Received: 15 November 2023; Accepted: 22 April 2024

Published online: 17 May 2024

References

- Zorzi, S. *et al.* Full-waveform airborne LiDAR data classification using convolutional neural networks. *IEEE Trans. Geosci. Remote Sens.* **57**(10), 8255–8261 (2019).
- Kakaletsis, E. *et al.* Computer vision for autonomous UAV flight safety: An overview and a vision-based safe landing pipeline example. *Acm Comput. Surv. Csur* **54**(9), 1–37 (2021).
- Nishio, T. *et al.* When wireless communications meet computer vision in beyond 5G. *IEEE Commun. Stand. Mag.* **5**(2), 76–83 (2021).
- Serghienco, O. Y. & Tyrsa, V. V. 3D optical machine vision sensors with intelligent data management for robotic swarm navigation improvement. *IEEE Sens. J.* **21**(10), 11262–11274 (2020).
- Shen, J. *et al.* Deep learning-based object identification with instance segmentation and pseudo-LiDAR point cloud for work zone safety. *Comput. Aided Civil Infrastruct. Eng.* **36**(12), 1549–1567 (2021).
- Wu, P. *et al.* Next-generation machine vision systems incorporating two-dimensional materials: Progress and perspectives. *InfoMat* **4**(1), e12275 (2022).
- Lin, C. *et al.* Background point filtering of low-channel infrastructure-based LiDAR data using a slice-based projection filtering algorithm. *Sensors* **20**(11), 3054 (2020).
- Wilson, A. N. *et al.* Embedded sensors, communication technologies, computing platforms and machine learning for UAVs: A review. *IEEE Sens. J.* **22**(3), 1807–1826 (2021).
- Lou, R. *et al.* Prediction of ocean wave height suitable for ship autopilot. *IEEE Trans. Intell. Transp. Syst.* **23**(12), 25557–25566 (2021).
- Gucma, S. Conditions of safe ship operation in seaports—optimization of port waterway parameters. *Pol. Marit. Res.* **26**(3), 22–29 (2019).
- Cortes-Vega, D., Alazki, H. & Rullan-Lara, J. L. Visual odometry-based robust control for an unmanned surface vehicle under waves and currents in a urban waterway. *J. Mar. Sci. Eng.* **11**(3), 515 (2023).
- Mansuy, M. *et al.* Optimized turning basin design for inland waterways. *J. Waterw. Port Coast. Ocean Eng.* **149**(4), 04023008 (2023).
- Sinohara, H. T. & Tannuri, E. A. Port's technologies for autonomous ships in restricted waters. *Marit. Bus. Rev.* **7**(1), 5–23 (2022).
- Nzengu, W. *et al.* Regulatory framework analysis for the unmanned inland waterway vessel. *WMU J. Marit. Aff.* **20**(3), 357–376 (2021).
- Hu, L. & Yang, H. Monitoring and analysis of sea fog in an offshore waterway using lidar. *Opt. Eng.* **60**(6), 064103–064103 (2021).
- Lu, Y. *et al.* Research on intelligent sensing technology of traffic volume in Yancheng inland trunk waterway. *Front. Comput. Intell. Syst.* **2**(3), 1–3 (2022).
- Tak, W. J. *et al.* Using drone and LiDAR to assess coastal erosion and shoreline change due to the construction of coastal structures. *J. Coast. Res.* **95**(SI), 674–678 (2020).
- Hake, F. *et al.* Using machine-learning for the damage detection of harbour structures. *Remote Sens.* **14**(11), 2518 (2022).
- Marchel, Ł., Naus, K. & Specht, M. Optimisation of the position of navigational aids for the purposes of SLAM technology for accuracy of vessel positioning. *J. Navig.* **73**(2), 282–295 (2020).
- Upadhyay, J., Rawat, A. & Deb, D. Multiple drone navigation and formation using selective target tracking-based computer vision. *Electronics* **10**(17), 2125 (2021).
- Padmaja, B. *et al.* Exploration of issues, challenges and latest developments in autonomous cars. *J. Big Data* **10**(1), 61 (2023).
- Miao, Y. *et al.* Airborne LiDAR assisted obstacle recognition and intrusion detection towards unmanned aerial vehicle: Architecture, modeling and evaluation. *IEEE Trans. Intell. Transp. Syst.* **22**(7), 4531–4540 (2020).
- Guan, F., Xu, H. & Tian, Y. Evaluation of roadside LiDAR-based and vision-based multi-model all-traffic trajectory data. *Sensors* **23**(12), 5377 (2023).
- Rodriguez-Conde, I., Campos, C. & Fdez-Riverola, F. On-device object detection for more efficient and privacy-compliant visual perception in context-aware systems. *Appl. Sci.* **11**(19), 9173 (2021).
- Chu, W. *et al.* Cloud control system architectures, technologies and applications on intelligent and connected vehicles: A review. *Chin. J. Mech. Eng.* **34**(1), 1–23 (2021).
- Hardstone, R. *et al.* Long-term priors influence visual perception through recruitment of long-range feedback. *Nat. Commun.* **12**(1), 6288 (2021).
- Xu, Y. *et al.* Deep coupled visual perceptual networks for motor fault diagnosis under nonstationary conditions. *IEEE/ASME Trans. Mechatron.* **27**(6), 4840–4850 (2022).
- Qi, J., Yang, H. & Sun, H. MOD-RRT*: A sampling-based algorithm for robot path planning in dynamic environment. *IEEE Trans. Ind. Electron.* **68**(8), 7244–7251 (2020).
- Le, A. V., Nhan, N. H. K. & Mohan, R. E. Evolutionary algorithm-based complete coverage path planning for tetraiamond tiling robots. *Sensors* **20**(2), 445 (2020).
- Luna, M. A. *et al.* Fast multi-UAV path planning for optimal area coverage in aerial sensing applications. *Sensors* **22**(6), 2297 (2022).

Author contributions

Zhan Zhang contributed to conception and design of the study. Nanwu Yang organized the database. Yijian Yang performed the statistical analysis. Zhan Zhang wrote the first draft of the manuscript. Nanwu Yang and Yijian Yang wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-60327-9>.

Correspondence and requests for materials should be addressed to Z.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024