



## Autonomous port management based AGV path planning and optimization via an ensemble reinforcement learning framework

Xinqiang Chen<sup>a</sup>, Shuhao Liu<sup>b</sup>, Jiansen Zhao<sup>b</sup>, Huafeng Wu<sup>b,\*</sup>, Jiangfeng Xian<sup>a</sup>, Jakub Montewka<sup>c</sup>

<sup>a</sup> Institute of Logistics Science and Engineering, Shanghai Maritime University, Shanghai, 201306, China

<sup>b</sup> Merchant Marine College, Shanghai Maritime University, Shanghai, 201306, China

<sup>c</sup> Research Group on Maritime Transportation Risk and Safety, Gdynia Maritime University, Morska 81-87, 81-225, Gdynia, Poland



### ARTICLE INFO

#### Keywords:

Automated container terminal  
AGV transportation  
Path optimization  
APF algorithm  
TD3 algorithm

### ABSTRACT

The rapid development of shipping trade pushes automated container terminals toward the direction of intelligence, safety and efficiency. In particular, the formulation of AGV scheduling tasks and the safety and stability of transportation path is an important part of port operation and management, and it is one of the basic tasks to build an intelligent port. Existing research mainly focuses on collaborative operation between port equipment and path optimization under environmental perception, while there is relatively little research on optimization of path smoothness and safety. Therefore, we propose a path optimization model based on the artificial potential field and twin delayed deep deterministic policy gradient (APF-TD3) framework for the port environment. Firstly, we obtain the scheduling task plan of a single AGV by enumeration. Secondly, according to the artificial potential field (APF) algorithm to generate repulsion for obstacles in the harbor and attraction for container storage at the target point with the position information of the AGV as the input data of the reinforcement learning algorithm is inputted into the twin delayed deep deterministic policy gradient algorithm (TD3). Then TD3 selects the optimal action strategy for the AGV according to the input AGV state information and the designed reward mechanism as well as executes the action. Through repeated execution, the optimal action for the next step is selected at each point to generate a path with start and end points. We validate the model by simulating the scale of containerized cargo in the port i.e. small scale, medium scale and large scale scenes. The experimental results show that the method has the shortest path length of 27.519 m, 270.847 m, and 496.389 m compared to artificial potential field and deep deterministic policy gradient (APF-DDPG), APF, and rapidly-exploring random tree (RRT) algorithms, which also have significant advantages in terms of path security and path smoothness. This framework could respond to the scheduling and transportation tasks of single AGV in different environmental layouts and guarantee the smoothness and safety of the path based on the optimization of the path, which promotes the efficient operation and management of ports.

### 1. Introduction

As an important node of international trade, ports undertake the main tasks of cargo distribution, loading and unloading, and transportation. With the continuous development of global trade, the transportation demand of ports is also gradually increasing, which puts higher requirements on the efficiency and safety of port logistics (Chen et al., 2022a). As an efficient and flexible logistics equipment, automated guided vehicle (AGV) has become an indispensable part of port logistics system (Chen et al., 2022b; Li et al., 2021a). Port is a complex

industrial environment where AGV needs to reasonably plan its own driving path according to the handling demands and transportation tasks of containerized cargo. Path planning directly affects the transportation efficiency of the AGV and the overall operation of the port. the AGV can perceive the surrounding environment through laser radar, magnetic nails, video cameras and other sensors, and then use the built-in algorithms to plan the optimal path and automatic driving (Sun et al., 2023b). Li et al., used laser radar to construct environment maps in addition to robot kinematic and physical simulation models, and then combined the improved A\* algorithm and DWA algorithm to solve

\* Corresponding author.

E-mail address: [hfwu@shmtu.edu.cn](mailto:hfwu@shmtu.edu.cn) (H. Wu).

global and local path planning, which improved fusion algorithm outperforms the traditional fusion algorithm in terms of both accuracy and stability (Li et al., 2022). Hu et al., constructed a node network based on the characteristics of magnetic nails-guided AGVs and used the Gumbel-Softmax strategy to discretize the scenario created by the node network, and then solved for the shortest path using the multi-intelligent deep deterministic policy gradient method (Hu et al., 2023). Lian et al., proposed an improved heuristic path planning algorithm based on a visual sensor system, which could significantly improve the decoding speed and accuracy to achieve accurate position prediction and optimized path planning effectively reducing motion blur and corresponding attitude deviation (Lian et al., 2022). Zhu et al., utilized visual sensors and target detection algorithms to obtain the position information of obstacles and AGVs as well as introduced the position and attitude of physical AGVs into the virtual environment via visual sensors, which input the acquired images into a reinforcement-learning model to perform the next action and update the path in real time (Zhu et al., 2023).

At present, there are two ways of port path planning applications: one is the fixed path planning with known magnetic nails sensing, while the other is the autonomous path planning relying on GPS, laser radar and sensors to perceive the surrounding environment. The port layout under the fixed path can consider the overall resources of the port globally, and the AGV drives according to the instructions of the central control system, which can have a better overall optimization ability in tasks scheduling and path planning. Tao et al., developed a mathematical model with the objective of shortest transportation time which was solved by an improved particle swarm optimization algorithm to obtain the optimal path (Tao et al., 2021). Liu et al., considered the AGV path optimization and scheduling problem which has designed an improved genetic algorithm combined with the neighborhood search strategy for critical paths that could ensure the effectiveness of local search paths for AGV (Liu et al., 2023). Yue et al., fused an improved heuristic algorithm, Dijkstra, and Q-Learning to solve the AGV transportation cost model with time constraints to obtain the optimal scheduling and routing scheme (Yue and Fan, 2022). Zhong et al., established a mixed-integer planning model to minimize the delay time of AGVs based on shore bridges, yards and AGVs, and used a hybrid genetic algorithm-particle swarm optimization algorithm to solve the AGV conflict and deadlock problems for simulated path nodes, which has promoted the development of the application of automated container terminals (Zhong et al., 2020). Yang et al., created a two-layer programming model based on collaborative scheduling among port operation equipment which was solved based on a two-layer genetic algorithm with congestion prevention rules to improve the turnover efficiency of port operations (Yang et al., 2018). Although integrated scheduling and maximum utilization of equipment resources in the port can be achieved under fixed routes, the large construction cost and the irreversibility of the port layout might make it difficult to respond to the increasing demand for container cargo scheduling and the requirement for algorithm adaptability.

Autonomous path planning that relies on GPS, laser radar and sensors to perceive the surrounding environment. This means that each AGV has autonomous decision-making capabilities to make decisions on path planning based on its own perception and local information. Wang et al., proposed an artificial potential field method with interference perception to plan the path through the attraction of target point, repulsion of obstacles and repulsion of GPS jammer, which used simulation to optimize the path planning result and analyze the effect of interference power on the path planning result (Wang et al., 2023). Han et al., proposed a path planning algorithm that combined localization algorithms and mobile anchor nodes to trade-off between localization accuracy and energy consumption by reducing number of corners in the planning path and energy consumption of mobile anchor nodes (Han et al., 2020). Samal et al., equipped the vehicle with multiple sensors including laser, RGB, and radar so that it fused the multimodal information to improve the accuracy of the task and reduce the utilization of

resource-intensive sensors that facilitated the application of object tracking and path planning (Samal et al., 2022). Zhang et al., developed a bionic dynamic path planning algorithm based on line-of-sight tracking rules, which was capable of information sensing, obstacle avoidance, and target tracking in optimizing path safety (Zhang et al., 2023b). Dirik et al., used image vision information for data collection of target points, obstacle information and position information to determine the initial parameters and coordinates of the obstacle-free path, which was used to generate the optimal path with a novel kinematic control structure of hybrid algorithms (Dirik et al., 2020). Although the path planning research based on environment perception can reduce the complexity of the system with better real-time and flexibility and can better adapt to the dynamic changes within the port, but it will also have the problems of delayed data transmission, decreased efficiency of path generation, uncertainty of the path security and stability, etc.

Along with the development of automated container terminals towards intelligence, efficiency and multi-objective, deep reinforcement learning algorithms under non-fixed routes have gradually become the mainstream in path planning research. Ye et al. set the reward function with the main objective of minimizing the overall energy consumption of AGVs, and modified the action and state space to suit the settings of AGV activities with the multi-intelligent deep deterministic policy gradient algorithm. energy consumption results showed that the routes planned by this algorithm could effectively improve the efficiency of energy utilization (Ye et al., 2023). Guo et al., proposed an improved soft Actor-Critic reinforcement learning method with composite auxiliary reward structure to deal with real-time navigation, which could solve the inefficiency problem of calculating the optimal path to achieve the generation of optimal path in a short time (Guo et al., 2023). Xu et al., established a mixed-integer planning model for the terminal crane, AGV and double cantilever rail crane in loading and unloading modes, which was solved by a super-heuristic genetic algorithm with reinforcement learning to avoid conflicts among AGVs and improve the handling efficiency (Xu et al., 2021). Deep reinforcement learning algorithms can handle high-dimensional, complex input data as well as handle multi-objective, multi-constraint and uncertainty problems better than other path planning algorithms. However, it may also encounter instability problems that require tuning to solve.

Whatever the method to obtain environmental information needs to ensure the safety and stability of the path route. Currently, there are many studies on port path planning, in which traditional algorithms (such as Dijkstra's algorithm (Dudeja and Kumar, 2022)) have a high theoretical basis and reliability. But they are insensitive to environmental changes and have high computational complexity, making them unsuitable for large-scale path calculations in ports. Heuristic algorithms (A\* algorithm (Tang et al., 2021), ant colony algorithm (Hou et al., 2022) and genetic algorithm (Luan and Thinh, 2023), etc.) can consider the constraints of the actual problem and find a better path, but they need to design and select the appropriate heuristic information, which may not be suitable for all problems. So many researches are based on heuristic algorithms to consider realistic conditions and improve it. Multi-objective path planning can consider multiple objectives at the same time, which makes path planning more flexible and efficient, but the high computational complexity requires trade-offs and optimization of multiple objectives. In recent years machine learning methods have been widely used in the field of path planning. Although the optimal path strategy can be obtained through learning and exploration, which requires a large amount of data for training and the process of training is complex and time-consuming (Chen et al., 2023b). Currently, the path planning research of the AGV in the port has explored a wide range of issues such as energy consumption of transportation paths, scheduling efficiency between devices, port carbon emissions, and braking safety of the devices themselves, other than the algorithmic innovations (Li et al., 2021b). It mainly utilizes multi-objective constraints to build the objective function and solve it to get the optimal path scheme. Environmental perception under the port

through a combination of lidar, infrared, and cameras is a mainstream research trend at present (Chen et al., 2023a). Deep reinforcement learning algorithms under non-fixed routes have great potential in the application of automated container terminals, but there are challenges on planning the most secure and reasonable transportation routes according to the actual application of ports.

Path planning based on non-fixed routes usually involves continuous and complex actions in path selection, such as the moving trajectories of AGV in the real environment. However, APF-TD3 can effectively deal with this continuous action space, which makes it possible to better realize AGV accurate movement and path selection in the path planning problem. Moreover, APF-TD3 adopts the double-Q network and target policy noise, which helps to improve the stability and convergence of learning. This is significant for the practical application of the path planning problem. We have explored path planning for single AGV in port environments with different scale scenarios. The enumeration method is utilized to determine the scheduling transportation tasks of the single AGV thereby obtaining the start and end points of the path. We introduce the artificial potential field algorithm (APF) to construct the potential field of repulsion and attraction for the shore bridges, yards, and target point, while the state information of the AGV is used as the data input for the twin delayed deep deterministic policy gradient algorithm (TD3) and select the optimal action strategy through the reward and punishment mechanism. The modeling framework based on APF-TD3 can quickly identify the port environment and generate the shortest paths that is safe and smooth. Safe and reasonable planning of AGV driving paths can improve the cargo throughput efficiency of ports, reduce labor costs, and enhance operational efficiency and competitiveness. Adapting to different operational requirements and environmental conditions can also improve the adaptability of port logistics systems. The main contributions are summarized as follows: (1) We set different port sizes as experimental scenarios to verify the safety and smoothness of AGV driving paths under different path lengths. The results show that our algorithm is better than other algorithms in obtaining smoother and safer transportation paths. (2) The TD3-APF

integration framework can be adapted to the changes of the port environment. And the relationship among AGV speed, energy consumption, carbon emission, and safety can be set up in the future by using the reward function to enrich the management system of port operation. (3) We explored that the path of AGVs under non-fixed routes can provide theoretical basis and technical support for the research of environment perception and collaborative management in automated container ports. The content of this paper is organized as follows: section 2 introduces the research methodology, section 3 shows the model through simulation and experimental analysis, section 4 discusses the analysis, and section 5 gives the conclusion.

## 2. APF-TD3 model for path planning

Path planning for AGV based on the APF-TD3 framework in a port environment has four steps in total: formulating AGV scheduling tasks, determining AGV state information, optimizing path decisions, and generating paths. Scheduling tasks are formulated for single AGV based on the input port layout information using enumeration method in the first step. In the second step, the APF algorithm is utilized to calculate the repulsive and attractive forces which are input to the TD3 algorithm as the state information of the AGV along with the port environment information. In the third step, the reward function of the TD3 algorithm is designed to select the optimal action decision for the AGV and return to the second step to execute the action. The second and third steps are executed iteratively until the AGV reaches the end point and stops, which eventually generates the complete path. The flowchart of the proposed framework is shown in Fig. 1.

### 2.1. Developing port AGV scheduling tasks

The AGV is required to transport containers from the shore bridge to the designated yard for temporary storage. Therefore, the scheduling task for the AGV needs to be formulated first in the port. The number of shore bridges is denoted by  $x_i^1$ , the number of yards is denoted by  $x_i^2$ , the

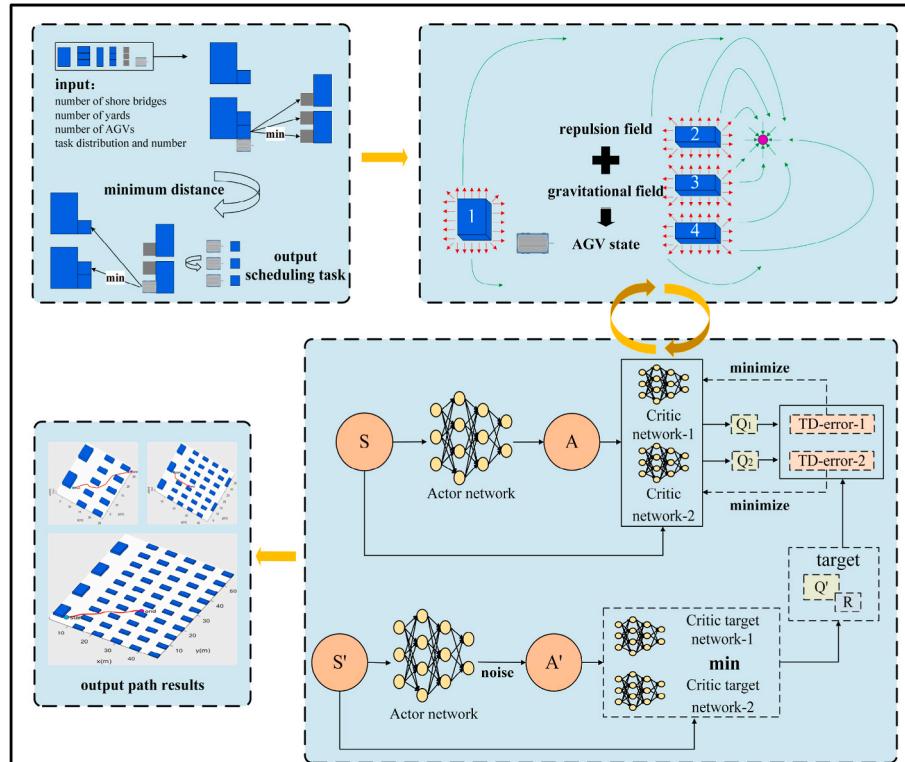


Fig. 1. APF-TD3 model framework.

number of container storage points is denoted by  $x_i^3$ , the number of AGVs is denoted by  $x_i^4$ , and the number of containers that need to be transported is denoted by  $x_i^5$  in which  $i$  denotes the number of these form the major part of the port operation. It is known that the number of scheduling task AGV is 1, which is  $x_1^4$ . We solve the path length  $d_{x_i^1-x_i^3}$  from the AGV starting position to the container storage point and  $d_{x_i^3-x_i^1}$  from the container storage point to the shore bridge by the APF-TD3 model. The path lengths from the start point of the shore bridge to all the container storage points are compared by enumeration, which selects the path with the minimum distance as the task priority to transport (Buermann and Zhang, 2022). The same is true for the container storage point to the shore bridge, as see Eq. (1).

Where  $i$  denotes the quantity and  $I$  denotes the selection of the optimal scheduling task. The tasks are completed when all the containers are transported to the storage points in the yard, and the driving records of the AGV is the scheduling tasks. Since the transportation task is performed for a single AGV so scheduling order of AGV can be quickly responded and calculated.

## 2.2. Determining the AGV state information

The state information of the AGV is based on the environment information obtained by the APF algorithm. APF defines the environment of an AGV in terms of potential field by imitating the electrostatic field model. The potential field is divided into attractive field and repulsive field where the attractive field is generated at the target point and the repulsive field is generated at the obstacles. The total potential field of the AGV environment is the superposition of the attractive and repulsive potential fields (Zhang et al., 2023a). The AGV is driven by the potential field force in a potential field environment. The attractive potential energy of the AGV is proportional to the distance of the target point. As the target point is farther away from the AGV the more potential energy it receives, and conversely the less potential energy it receives (Li et al., 2023). The attractive potential field function is shown in Eq. (2). The attractive force generated by the attractive potential field is the negative gradient of the attractive potential field, see Eqs. (3) and (4), where  $K_{att}$  is the attractive potential field gain factor, and  $\rho(q, q_g)$  is denoted as the Euclidean distance  $|q - q_g|$  between the two locations of the AGV current location  $q$  and the target point location  $q_g$ , with the direction pointing from the AGV location to the target point location.

$$U_{att}(q) = \frac{1}{2} K_{att} \rho^2(q, q_g) \quad (2)$$

$$F_{att}(q) = -\Delta U_{att}(q) \quad (3)$$

$$F_{att}(q) = K_{att}\rho(q, q_g) \quad (4)$$

The repulsive potential energy of the AGV is inversely proportional to the distance of the obstacle, when the distance between the obstacle and the moving AGV is farther away, in which case the potential energy suffered by it is less (Zhang et al., 2021). The potential energy suffered by an obstacle is more when its distance from the AGV is closer. The

repulsive potential field function is given by see Eq. (5). The repulsive force generated by the repulsive potential field is expressed by Eq. (6). Eq. (7) represents the total potential field function of the AGV, and Eq. (8) represents the combined force on the AGV.

$$U_{rep}(q) = \begin{cases} \frac{1}{2}K_{rep}\left(\frac{1}{\rho(q, q_0)} - \frac{1}{\rho_0}\right)^2, & 0 \leq \rho(q, q_0) \leq \rho_0 \\ 0, & \rho(q, q_0) > \rho_0 \end{cases} \quad (5)$$

$$F_{rep}(q) = \begin{cases} K_{rep} \left( \frac{1}{\rho(q, q_0)} - \frac{1}{\rho_0} \right) \frac{1}{\rho^2(q, q_0)} \nabla \rho(q, q_0), & 0 \leq \rho(q, q_0) \leq \rho_0 \\ 0, & \rho(q, q_0) > \rho_0 \end{cases} \quad (6)$$

$$N) N \in R \quad (1)$$

Where  $K_{rep}$  is the repulsive potential field gain factor.  $\rho_0$  is a constant

that represents the upper limit distance at which the obstacle effects the AGV.  $\rho(a, a_0)$  represents the distance between the current position  $a$  of

Where  $K_{rep}$  is the repulsive potential field gain factor.  $\rho_0$  is a constant that represents the upper limit distance at which the obstacle effects the AGV.  $\rho(q, q_0)$  represents the distance between the current position  $q$  of the AGV and the closest point  $q_0$  of the obstacle area around the AGV, which points to the direction from the obstacle to the AGV. The AGV position, the repulsive potential field of the obstacle, and the attractive potential field of the target point will all be input into the TD3 model as state information for optimal strategy selection, which will update the next position of the AGV.

### *2.3. Making decisions of the optimal path*

### 2.3.1. Markov decision process

The theoretical framework of reinforcement learning is usually based on Markov Decision Process (MDP). The AGV learns and explores in the port and gets the state information from the environment, meanwhile it evaluates the action according to the level of the reward provided by the environment, eventually the AGV chooses the action according to the size of the reward value (Savas et al., 2020). Markov decision process consists of  $(S, A, P, R, \gamma)$ , in which  $S$  is the set of states where AGV interacts with the environment,  $A$  is the set of behavioral actions of AGV,  $P$  is the state transition probability matrix, which denotes the probability  $P(s'|s, a)$  of AGV executing action  $a (a \in A)$  to reach  $s'$  in state  $s (s \in S)$ , and  $R$  is the reward function, where  $R(s, a, s')$  represents the reward value received by the AGV after taking action  $a$  to reach the next state  $s_{t+1}$  in state  $s$ .  $\gamma$  is the discount factor,  $\gamma \in (0, 1)$ . The executed action is defined by the policy  $\pi$ . The policy  $\pi(a|s)$  represents the probability distribution of the action  $a$  in the given state  $s$ . At any  $t$  moment AGV gets and executes action  $a_t$  from action set  $A$  according to policy  $\pi$  in state  $s_t$ , reaches the next state  $s_{t+1}$  and at the same time gets reward  $r(s_t, a_t)$ . After that, in state  $s_{t+1}$  execute action  $a_{t+1}$  according to policy  $\pi$  to reach state  $s_{t+2}$  and get reward  $r(s_{t+1}, a_{t+1})$ . This cycle is repeated and eventually AGV updates the policy to maximize its cumulative reward return  $G_t$ . Each state records the state information in the current environment therefore the current state does not have any relationship with the historical information. Both the transfer probability from the current state to the next state and the reward value are only related to the current state and

action (Tao et al., 2022).

A strategy  $\pi$  is a probability distribution that represents how an AGV chooses an action in the state of the environment at a particular moment in time. It is also defined as the mapping relation of the actions taken by the AGV in a particular state when a state transition occurs, which is denoted as  $\pi : S \rightarrow A$ . The policy determines the magnitude of the probability of the AGV taking different actions in various states, where the policy can also be the output of the algorithm. The randomness policy makes the output of the actions of the AGV in the same environmental state not unique, but satisfies a probability distribution of various actions. Deterministic policy means that the AGV will choose the same action every time in the same state. The policy  $\pi$  can be described as a function containing  $\theta$ , including the randomness policy  $\pi_\theta = P[a|s, \theta]$  and the deterministic policy  $a = \mu_\theta(s)$ . The reward value is the feedback that the AGV receives after executing an action during its interaction with the environment, which is called reward. The value size of the reward directly reflects how well the AGV selects an action in the current state. The reward is generally expressed as the expectation on the sum with the reward obtained from all possible actions in the current state. The cumulative reward in the markov process is expressed in terms of the reward  $G_t$  see Eq. (9). larger value of  $\gamma$  indicates that the AGV's action has a greater impact on the future, while  $\gamma = 0$  indicates that only the reward of the current state is considered (Jiang et al., 2022b).

$$G_t = R_{t+1} + \gamma R_{t+2} + \dots + \gamma^k R_{t+k+1} = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}, \gamma \in [0, 1] \quad (9)$$

### 2.3.2. Actor-critic framework

The Actor-Critic framework consists of value functions and policy functions in reinforcement learning. The policy function has the advantage of updating in a continuous action space while the value function is more efficient. The policy function acts as an Actor, which selects actions and interacts with the environment in a continuous action space, and the value function acts as a Critic, which evaluates the performance of the Actor (Chen and Wu, 2023). It is necessary to input the current state  $s_t$  and output the action  $a_t$  under the policy to reach the next new state  $s_{t+1}$  and get the reward  $r_{t+1}$  after interacting with the environment in Actor network. The Actor network calculates the value when executing action  $a_t$  in state  $s_t$  based on the neural network which is used to update the network parameters  $\theta$ . According to the reward and punishment mechanism, the network parameters are continuously optimized making the Actor network to get higher reward values after executing the actions. Critic network scores the values calculated in the Actor network and sends the scoring results back to the Actor network, which is updated according to the scoring results. Actor-Critic framework allows both updates to be done simultaneously, which significantly improves the efficiency of the algorithm. The policy controls the behavioral decisions of the AGV and the value function evaluates the merit of the policy behavior (Peng et al., 2023). The updating process is described in Eq. (10) and Eq. (11).

$$\Delta\theta = \alpha \nabla_\theta \log_{\pi_\theta}(S_t, A_t) Q_\omega(S_t, A_t) \quad (10)$$

$$\Delta\omega = \beta(R_{t+1} + \gamma Q_\omega(S_{t+1}, A_{t+1}) - Q_\omega(S_t, A_t)) \nabla_\omega Q_\omega(S_t, A_t) \quad (11)$$

Where  $\Delta\theta$  is the amount of updating of the policy parameters,  $\alpha$  and  $\beta$  is the learning rate parameter,  $\nabla_\theta \log_{\pi_\theta}(S_t, A_t)$  is the amount of change in the gradient of the policy function  $\pi_\theta$  in the state  $S_t$  and in the action  $A_t$ , and  $Q_\omega(S_t, A_t)$  is the estimate in the state  $S_t$  and in the action  $A_t$ . The policy function can be updated towards maximizing the value function  $Q_\omega(S_t, A_t)$ .  $\Delta\omega$  is the update of the weighting parameter,  $R_{t+1}$  is the reward at time step  $t+1$ ,  $Q_\omega(S_{t+1}, A_{t+1})$  is the estimator value function under the state  $S_{t+1}$  and the action  $A_{t+1}$ , and  $\nabla_\omega Q_\omega(S_t, A_t)$  is the gradient calculation of  $Q_\omega(S_t, A_t)$  with respect to the weight parameter  $\omega$ .

### 2.3.3. Twin delayed deep deterministic policy gradient algorithm

TD3 is a reinforcement learning algorithm using a deterministic

policy that has two Actor-Critic networks of the same structure. Each Critic network uses a double-network structure as to carry out the output as a result TD3 consists of a total of six neural networks including two Actor networks and four Critic networks (Hong et al., 2021).  $\theta_{i=1,2}$  and  $\theta'_{i=1,2}$  are the Critic network and Critic goal network parameters, and  $\varphi$  and  $\varphi'$  are the Actor network and Actor goal network parameters. Actor network outputs action  $a_t$  in state  $s_t$  and Actor goal network outputs action  $\tilde{a}_t$  in state  $s_{t+1}$ . Critic network samples a small batch of experiences  $(s_t, a_t, r_t, s_{t+1})$  from the experience pool and outputs  $Q_{\theta_i}(s_t, a_t)$  based on the state  $s_t$  and the action  $a_t$ . The Critic target network outputs  $Q_{\theta'_i}(s_{t+1}, \tilde{a}_t)$  based on the sampled states  $s_{t+1}$  and actions  $\tilde{a}_t$ . Since the Q values are maximized during the updating process, there exists some error in each update, which accumulated after many updates will lead to some states being overestimated (Zhang et al., 2022). So TD3 to solve the problem of error overestimation in each network uses two Critic networks to evaluate the Q-value and takes the smaller Critic network value as the update target during the update, as shown in Eq. (12).

$$y_{target} = r + \gamma \min_{i=1,2} Q_{\theta'_i}(s_{t+1}, \tilde{a}_t) \quad (12)$$

TD3 uses delayed updating in updating the network parameters by updating the Actor network parameters after every  $d$  times of updating the Critic network parameters. This avoids unnecessary repetitive updates and reduces the cumulative error of multiple updates which makes the network update more stable (Jiang et al., 2022a). The Critic network is updated by minimizing the function  $J(\theta)$  as shown in Eq. (13). The Actor network update uses a deterministic policy to update the policy gradient as shown in Eq. (14). The target network uses a soft update method, which updates the network parameters at every step in the main network and only updates a small part of the target network each time, so that the updated target does not change too much to increase the stability of training, as shown in Eqs. (15) and (16). Where  $N$  is the number of samples,  $t$  and  $i$  are the indexes of the samples,  $y_{target}$  is the target value,  $Q_{\theta_i}(s_t, a_t)$  is the value of Q calculated according to the parameter  $\theta_i$ ,  $Q_{\theta'_i}(s, a)$  is the value of Q calculated according to the parameter  $\theta'_i$ , and  $\pi_\varphi(s)$  is the policy calculated according to the parameter  $\varphi$ .  $\tau$  is a constant value, which usually takes the smaller value of  $\tau \in [0, 1]$  (Jeng and Chiang, 2023).

$$\nabla_{\theta_i} J(\theta_i) = N^{-1} \sum_t \nabla_{\theta_i} (y_{target} - Q_{\theta'_i}(s_t, a_t))^2 \quad (13)$$

$$\nabla_\varphi J(\varphi) = N^{-1} \sum_i \nabla_a Q_{\theta_1}(s, a) \Big|_{a=\pi_\varphi(s)} \nabla \pi_\varphi(s) \quad (14)$$

$$\theta'_i = \tau \theta_i + (1 - \tau) \theta'_i \quad (15)$$

$$\varphi' = \tau \varphi' + (1 - \tau) \varphi' \quad (16)$$

The TD3 algorithm stores the data from the exploration process in the experience pool and randomly samples all the data in the repository during updating that can well reduce the correlation between the samples and time. TD-error denotes the value of the difference between the current Q value and the target Q value, which also reflects the degree of learning needed by the AGV. The greater difference of TD-error indicates that the more the experience samples need to be updated. There is a problem of balancing exploration and utilization when AGV selects actions through policies.TD3 algorithm adds the degree of exploration to the policies of AGV action selection to change the locally optimal policy into the globally optimal policy. Noise is Gaussian noise that satisfies the distribution  $N(\mu, \sigma)$  with expectation  $\mu$  and standard deviation  $\sigma$  (Kim et al., 2020).

### 2.3.4. Transportation path generation

In path planning under the port, the APF framework is used to avoid collisions with obstacles and TD3 is used to optimize the path length and reward function. Specifically, the APF algorithm computes the repulsive

and attractive forces for each obstacle and the final total force which are used to calculate the AGV's actions to avoid collisions with obstacles. At each time step the AGV selects an action and collects state information and reward information which experiences are stored in a playback buffer used to update the AGV's network parameters. APF-TD3 framework can avoid collisions with obstacles while minimizing the path length and maximizing the reward function by combining both algorithms for cyclic updating.

**State Space:** The obstacles are mainly composed of rectangular objects in the environment and each obstacle has its own position with size parameter information. The state space includes the location and size of each obstacle, the location of the target point, and the location of the AGV.

**Action Space:** The action space design utilizes a continuous action space methodology. The dimensions of the action space correspond to the number of obstacles and each obstacle corresponds to a dimension indicating the action executed on that obstacle. The range of values for actions in each dimension is constrained to be between 0.1 and 3. This design is called boundary constraints in the action space, which allows the AGV to select actions in such a way that they cannot exceed a predetermined range and can avoid illogicality or too large action values (Ai et al., 2021).

**Reward Function:** Considering the shortest paths and not colliding with obstacles we set up a penalty and reward mechanism. Path length penalty: encouraging the search for shorter paths is represented by a negative value, which is 5 times the length of the path. Time penalty: the penalty is calculated based on the time consumed which is directly subtracted from the time value. Collision Penalty: If a collision occurs with obstacles and the penalty is a fixed value of -10, else we add rewards based on the reciprocal of the distance.

#### 2.4. Path evaluation indicators

In this paper we evaluate the generated paths using metrics: path length, Global Safety (GS) and Local Smoothness (LS). GS represents the average turn angle of a path, which is obtained by calculating the angle between each neighboring point on the path and averaging the values. GS can be used to describe the degree of curvature or frequency of turns in a path. Larger GS values indicate the path has more turns while smaller GS values indicate the path is more linear. By calculating the

average turning angle of a path it is possible to quantify and compare the curvilinear characteristics of the path. LS represents the change in steering angle of the AGV path that describes the degree of curvature of the path. It is obtained by calculating the difference between neighboring steering angles on the path and maximizing it by taking the absolute value. Larger LS values indicate that the steering angle of the path varies a lot and the AGV driving path needs to make steering operations which may be more zigzag and unstable. Therefore, smaller LS values indicate that the steering angle of the path varies less and the AGV driving path is smoother. GS and LS can be used to measure the safety and smoothness of the paths so that the path planning algorithms can be evaluated and optimized (Lv et al., 2019). GS and LS are calculated by Eqs. (17) and (18).

$$GS = \frac{1}{n-2} \times \sum_{i=2}^{n-1} \arccos\left(\frac{\text{dot}(v(i-1), v(i))}{\text{norm}(v(i-1)) \times \text{norm}(v(i))}\right) \quad (17)$$

$$LS = \max_{i=2}^{n-1} (\mu_1 \times |\beta_i - \beta_{i-1}|) \quad (18)$$

Where  $n$  is the number of points in the path,  $v(i)$  is the vector from the  $i$  point to the  $i+1$  point in the path,  $\text{dot}(v(i-1), v(i))$  is the dot product of the vector  $v(i-1)$  and the vector  $v(i)$ , and  $\text{norm}(v(i))$  is the length of the vector  $v(i)$ .  $\mu_1$  is the weighting factor for the change in steering angle, and  $\beta_i$  is the steering angle of the  $i$  point.

### 3. Simulation experiment and result analysis

#### 3.1. Simulation layout

With the development of containerized shipping industry, port management planning is also on the rise. Existing port terminal types include vertical terminal layout, parallel terminal layout, and U-shaped terminal layout, which includes AGVs, shore bridges, containers, yards, etc. Parallel distribution of shore bridges and yards is a parallel terminal layout, while vertical distribution is a vertical terminal layout. The U-shaped terminal is a U-shaped yard that can be used for container loading and unloading at the same time. Fig. 2 simulates the port environment according to the vertical and parallel layout that contains the facilities in the port. The AGV will need to transport containers from the location where the containers are stored next to the shore bridge to

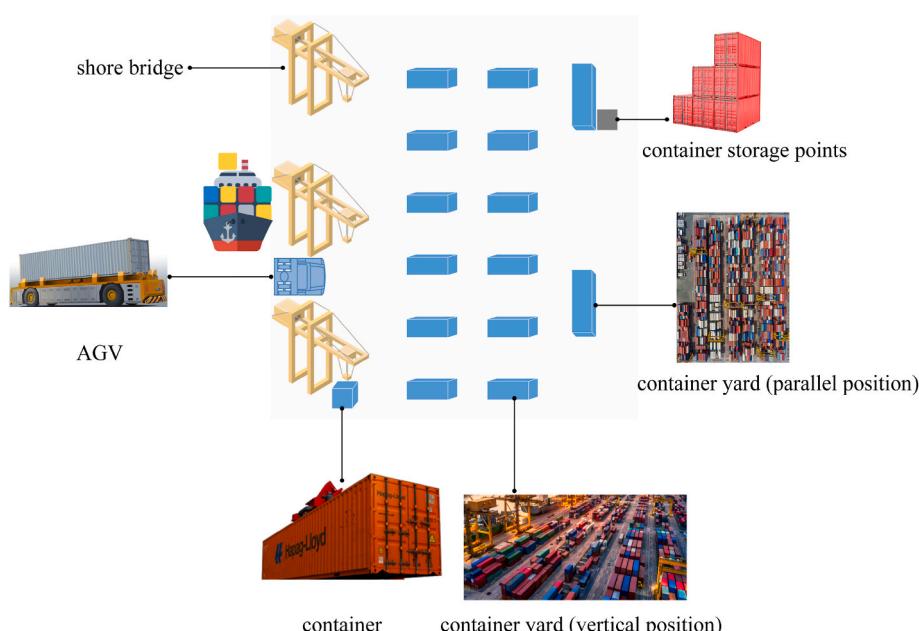


Fig. 2. Layout of the port environment.

the point where the containers are stored next to the yard. The work of the shore bridge is responsible for loading and unloading containers from the ship and temporary storage of containers in the yard. Fig. 2 contained the shore bridge, ship, AGV, container, container storage point and container yard. We use  $4 \times 6$  cell size blue rectangle as the shore bridge facility,  $4 \times 2$  cell size blue rectangle as the vertical layout of the yard facility,  $2 \times 6$  cell size blue rectangle as the parallel layout of the yard facility,  $2 \times 2$  cell size square as the container and the container storage point where gray color is the container storage point in simulation experiments. The length of the cell is 1 m in this case. In the port environment both the shore bridge and the yard will be used as obstacles in the path planning of the AGV. Therefore, the port environment can be complicated by increasing the number of shore bridges and yards as well as AGV tasks number will increase. Because in this paper we are considering single AGV path planning that means without other AGVs as obstacles. The layout of the port environment based on the change in the number of shore bridges and yards can verify the AGV driving path situation in different environments.

### 3.2. Experimental settings

The proposed framework was suitable for path planning of a single AGV in a harbor, which had been described in detail in the above section. Our framework was developed on Win11 operating system with 16G RAM and 1.7 GHz CPU. the GPU version was NVIDIA GeForce RTX 3050 Ti with 8 GB of RAM. Besides, the simulation platform was implemented on Python (version 3.9) and Matlab (version 2020b). The distribution of the number of yards in the port would be different according to the size and needs of the port. The number of yards might be distributed in several different areas such as south, north, east, west and other different directions, which should be determined according to the actual situation of each port. Therefore, we set up a small, medium and large port environment with different number of scales for simulation experiments, which was shown in Table 1. The small-scale scene was a vertically layout container port with 2 shore bridges, 10 yards, 1 transportation task and 1 container storage point. The medium-scale scene was a container port with parallel and vertical layouts of yards, with 3 shore bridges, 30 yards, 5 transportation tasks and 3 container storage points. The large-scale scene was a vertically layout container port with 4 shore bridges, 50 yards, 10 transportation tasks and 5 container storage points. Moreover, the environmental setting only considered the tasks of unloading containers without loading, which meant that the AGV was empty from the yard to the shore bridge. The algorithms used in this comparison experiment were the APF algorithm, the RRT algorithm, and the APF-DDPG algorithm which had the same reward function settings as the APF-TD3 algorithm proposed in this paper. The specific parameter configurations of the four algorithms were shown in Table 2.

The discount factor  $\gamma$  was used to discount future rewards to balance the importance of current and future rewards. soft update frequency  $\tau$  which denoted the frequency of updating parameters from the target network to the main network. actor network learning rate  $\alpha$  which

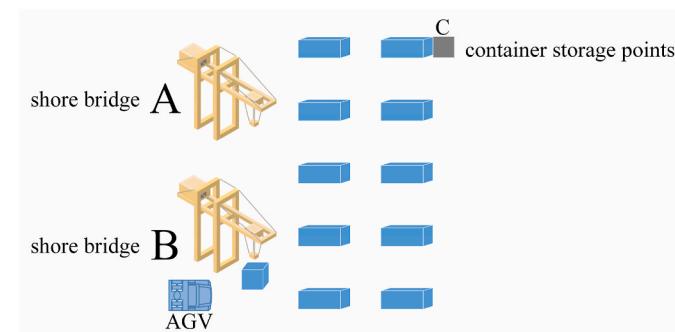
**Table 2**  
Four algorithmic parameters of configuration.

parameter	value
discount factor $\gamma$	0.99
soft update frequency $\tau$	0.003
actor network learning rate $\alpha$	0.001
critic network learning rate $\beta$	0.001
explore noise	0.15
policy noise	0.2
network update frequency	2
attraction factor $K_{att}$	0.8
repulsion factor $K_{rep}$	0.5
upper limit of distance $\rho_0$	5
steering angle change factor $\mu_1$	0.5
step size	0.2

denoted the actor network learning rate in the reinforcement learning actor-critic framework to control the speed of updating the network weights. critic network learning rate  $\beta$  which denoted the learning rate of the critic network in the reinforcement learning actor-critic framework, and controlled the speed of updating the value function. explore noise was used to introduce randomness into the exploration of the environment by AGV to better explore the state space. policy noise was used to introduce randomness into the policy update process. network update frequency denoted the frequency of neural network parameter update. The above were the settings for the parameters of the reinforcement learning algorithm. The attraction factor  $K_{att}$  represented the attraction of the target point to the AGV in the APF algorithm. The repulsion factor  $K_{rep}$  represented the repulsion force of the obstacle to the AGV in the APF algorithm. The upper limit of distance  $\rho_0$  was used to define the influence range of the APF algorithm. The steering angle change factor  $\mu_1$  indicated the flexibility of AGV steering in the APF algorithm. step size denoted the distance parameter for each time a new node was added to the tree during the growth of the seeded tree in the RRT algorithm.

### 3.3. Experimental results for small-scale scenes

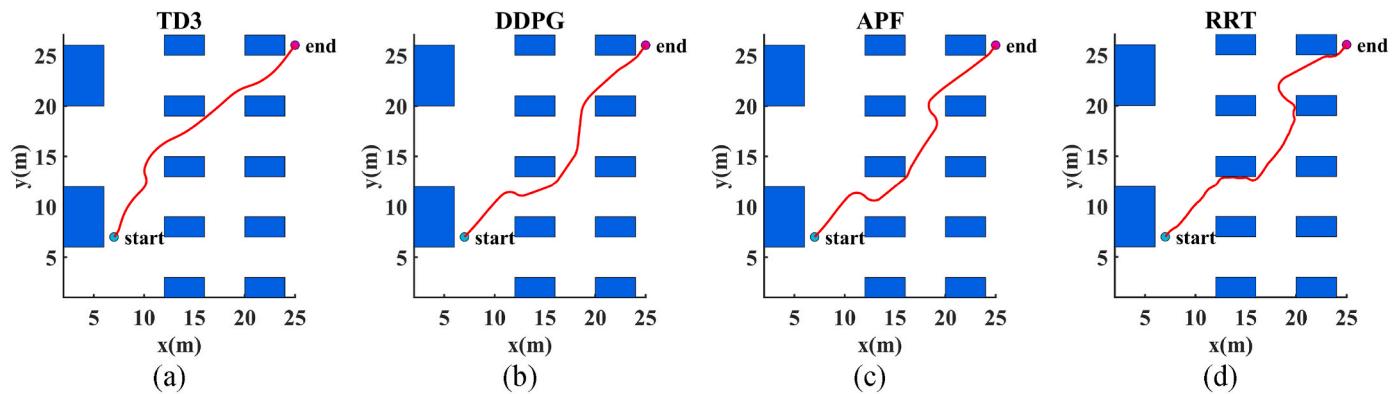
The small-scale scenes had only one transportation task, which required the container to be transported from the shore bridge marked with the letter B to the yard marked with the letter C. The layout of the scene had 2 shore bridges and 10 yards as shown in Fig. 3. Fig. 4 and Table 3 represented the path and data information of the four algorithms where the green circle indicated the position of the AGV starting from the shore bridge and the red circle indicated the position of the AGV arriving at the yard. Red path showed the trajectory path of AGV driving path in the harbor and green path showed the AGV trajectory path obtained after smoothing the route derived from the RRT algorithm. The results showed that the APF-TD3 algorithm proposed in this paper not only outperformed the APF-DDPG reinforcement learning algorithm but also outperformed the traditional path planning algorithms. It



**Fig. 3.** Distribution of small-scale scenes port.

**Table 1**  
Port layout information for different size scenes.

experimental scenes	number of shore bridges	number of yards	number of tasks	container storage points	yard layout
small-scale	2	10	1	1	vertical layout
medium-scale	3	30	5	3	parallel and vertical layout
large-scale	4	50	10	5	vertical layout



**Fig. 4.** Path results plotted for four algorithms in a small-scale scene: (a) path results generated by the APF-TD3 algorithm; (b) path results generated by the APF-DDPG algorithm; (c) path results generated by the APF algorithm; and (d) path results generated by the RRT algorithm.

**Table 3**  
Comparison of experimental data with different algorithms in small-scale scenes.

transportation path (B-C)	path length (meter)	GS	LS	calculation time (second)
APF-TD3	<b>27.519</b>	1.055°	5.000°	0.190
APF-DDPG	28.353	1.068°	10.489°	0.590
APF	29.393	1.529°	11.727°	0.070
RRT	33.628	11.680°	45.309°	31.709

significantly outperformed the other algorithms in both path length and degree of twists and turns but the difference in average turning angle with APF-DDPG was minimal. In the small-scale scenes, the RRT algorithm differed significantly from the other algorithms in all three evaluation indexes of path effect, which differed from the APF-TD3 algorithm by 6.109 m in path length, 10.625° in average turning angle, and 40.309° in path twisting degree. The APF-DDPG algorithm optimized the path length by 1.04 m based on the APF algorithm with an average turn angle optimized by 0.461° and the path twist optimized by 1.238°. The APF-TD3 algorithm optimized the path length by 0.834 m with an average turn angle optimized by 0.013° and a path twist optimized by 5.489° based on the APF-DDPG algorithm. Calculation time of APF-TD3 outperforms APF-DDPG and traditional RRT algorithms with a

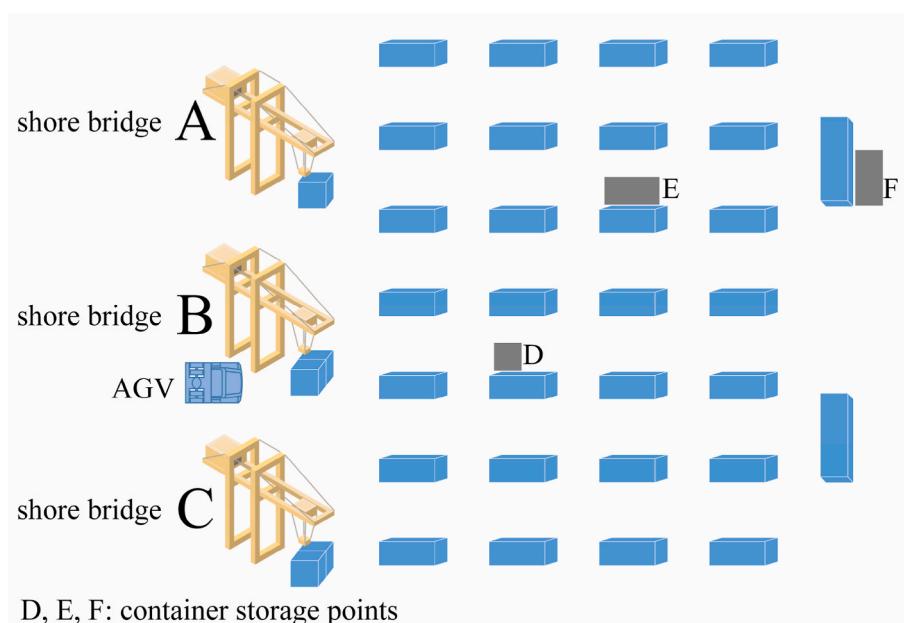
mere 0.19 s. The calculation time obtained by the reinforcement learning algorithm indicated the testing time rather than the training time. In small-scale scenes, our model was able to obtain the optimal path effect to ensure the safety of AGV driving path in the harbor.

#### 3.4. Experimental results for medium-scale scenes

The medium-scale scenes had five transportation tasks distributed in

**Table 4**  
AGV transportation paths in medium-scale scenes.

medium-scale	transportation path	AGV transportation container (1 denotes transportation)
1	B-D	1
2	D-B	
3	B-E	1
4	E-A	
5	A-E	1
6	E-C	
7	C-F	1
8	F-C	
9	C-F	1



**Fig. 5.** Distribution of medium-scale scenes port.

one container at shore bridge A, two containers at shore bridge B, two containers at shore bridge C, and one storage location at yard D, two storage locations at yard E, and two storage locations at yard F. The AGV started to transport containers to the yard using the B shore bridge as the starting point. The scene layout had 3 shore bridges and 30 yards as shown in Fig. 5. Because it was a single AGV transportation tasks and we calculated the path length utilizing APF-TD3 which was used as the main criterion for scheduling decision. Scheduling tasks order for the medium-scale experimental scenes was calculated by the enumeration method as shown in Table 4, which needed to calculate a total of 8 paths. Firstly, the AGV completed the task transportation of one container in yard D, then completed the task transportation of two containers in yard E, and finally completed the task transportation of two containers in yard F. The length of the transportation paths showed that the B-D and D-B paths were short distances with a length of less than 20 m, the B-E, E-A and A-E paths were medium distances with a length of less than 30 m, and the E-C, C-F and F-C paths were long distances with a length of more than 30 m.

Medium-scale scenes compared to the small-scale scenes whose paths were distributed in a wider range of eight transportation paths as shown in Fig. 6. From the experimental results data in Table 5, we could find APF-TD3 algorithm was able to adapt well on short, medium, and long distances that not only the path length was the shortest, but also the average turning angle and the degree of twists and turns were the smallest values. APF-DDPG algorithm minimized the difference with APF-TD3 algorithm in short distances but existed transportation paths A-E and C-F with larger difference to APF-TD3 algorithm in medium and long distances. The RRT algorithm in transportation paths A-E and C-F outperformed the APF-DDPG algorithm in path length by 0.206 m and 2.002 m, but was significantly different from the APF-DDPG algorithm in average turning angle and degree of twists and turns. It was shown that along with the longer driving path and the change of environment complexity would impact the solution of both APF-DDPG algorithm and RRT algorithm. Although RRT algorithm might have the shortest path solution but it could not satisfy the requirement of turning angle and twisting degree of the path in driving safety. The APF-DDPG algorithm could outperform the APF algorithm in the metrics of average turning angle for transportation paths in medium-scale scenes. However, the APF algorithm also outperformed the APF-DDPG algorithm in the metrics of path length for the C-F path in the long distances, as well as the metrics of degree of twists and turns in the medium distances B-E and E-A and in the long distances E-C and F-C. It was shown that the APF based DDPG reinforcement learning algorithm was not fully guaranteed to outperform the basic APF algorithm in calculating the optimal path length and degree of twists and turns when faced with the complex port environment of long distances. The APF-TD3 algorithm outperformed the reinforcement learning algorithm APF-DDPG and the traditional algorithm RRT in medium-scale scenes with different path lengths in terms of calculation efficiency.

Three sets of transportation paths with interchangeable starting and ending points, such as B-D and D-B, A-E and E-A, and C-F and F-C. Their driving path plots, path lengths, average turning angles of the paths, and the degree of path twists and turns were different, showing that results of running algorithms with different environments differed even on the same road segments. Under the optimal algorithm APF-TD3 short distances B-D and D-B showed a difference of 0.125 m in the path length, a difference of 0.206° in the average turning angle, and a difference of 1.355° in the degree of twists and turns. The medium distances A-E and E-A under the optimal algorithm APF-TD3 showed a difference of 1.101 m in the path length, a difference of 0.147° in the average turning angle, and a difference of 3.269° in the degree of twists and turns. Long distances C-F and F-C under the optimal algorithm APF-TD3 showed a difference of 0.33 m in the path length, a difference of 0.189° in the average turning angle, and a difference of 0.61° in the degree of twists and turns. The largest difference in path length and average turn angle in the B-D and D-B transportation paths was in the APF algorithm with

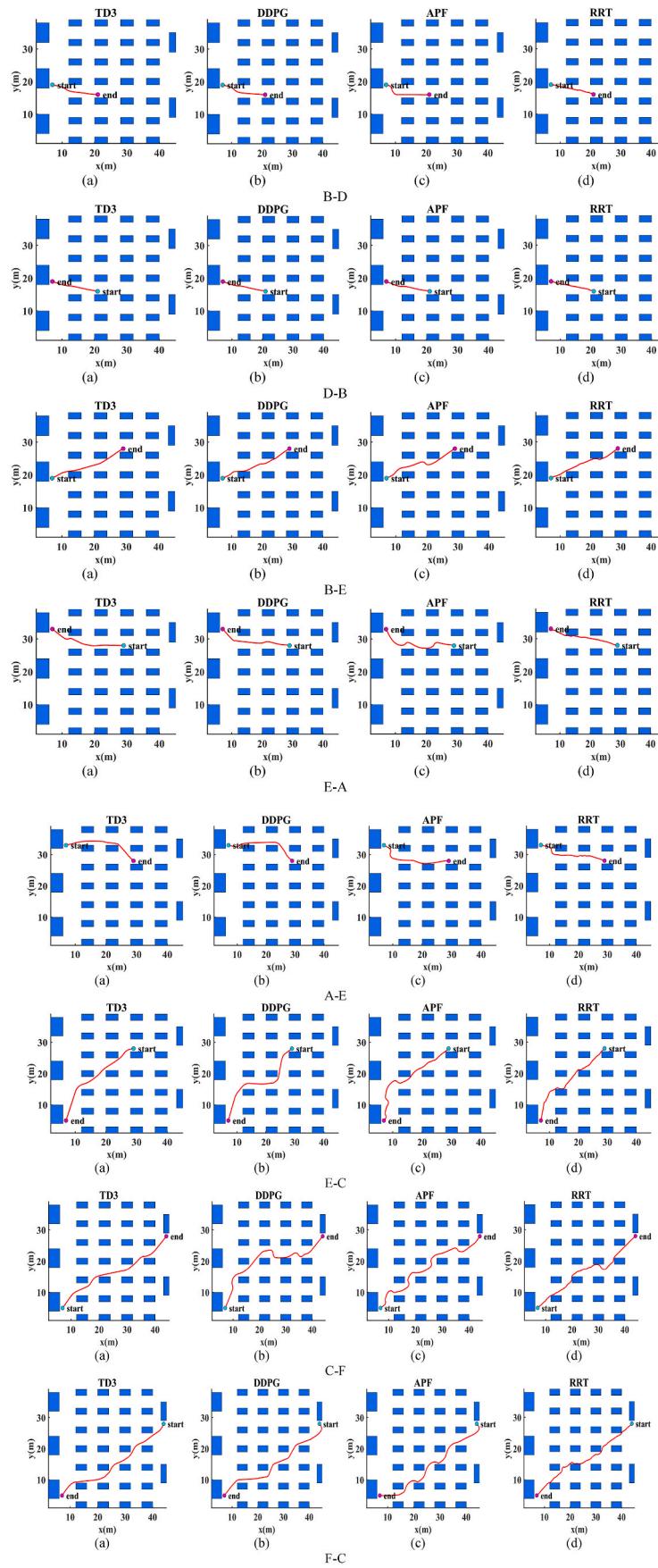
1.116 m and 0.979°, and RRT was the algorithm with the largest difference in the degree of twists and turns of the paths with a difference of 5.04°. In A-E and E-A transportation paths, the APF-DDPG algorithm calculated the largest difference in path length of 1.388 m, the RRT algorithm calculated the largest difference in the average turning angle of the path of 2.756°, and the APF algorithm calculated the largest difference in the degree of twists and turns in the path of 6.37°. In C-F and F-C transportation paths, the APF-DDPG algorithm calculated the maximum difference among the length of the path, the average turning angle, and the degree of twists and turns, which were 3.926 m, 0.621°, and 9.65°. The calculation efficiency varied for different transportation paths because the longer the transportation path the more calculation time, which could be quickly responded by our proposed APF-TD3 algorithm. Although APF had shortest calculation time, but its path length was not the shortest and it was not to ensure the safety and smoothness of the path. The experimental results described above showed that our proposed APF-TD3 model had better stability of paths during transportation tasks with interchangeable starting and ending points.

Based on our proposed APF-TD3 framework to solve the AGV scheduling transportation tasks in a medium-scale scenes, which total path length as shown in Table 6. The optimal path length for total driving of the framework was 270.847 m which was followed by RRT algorithm with a difference of 19.224 m. The maximum difference between the path length calculated by the APF algorithm and this framework was 22.54 m while the APF-DDPG algorithm optimized the path length of 2.311 m compared to the APF algorithm. Although the RRT algorithm had better results than the APF and APF-DDPG algorithms in the path length metrics, but it could not ensure the safety and stability of the AGV driving path. The APF-TD3 algorithm and the APF-DDPG algorithm had the same reward and penalty settings and parameter settings for algorithm operation. However, APF-DDPG algorithm did not have all the evaluation metrics of the transportation paths that were less different from the results of APF-TD3 algorithm, which would be reflected in the length and complexity of the transportation paths. We used the average occupancy of CPU and GPU during the model run as a manifestation of the algorithm's complexity. Since the complexity statistic was stochastic, therefore we used the mean of the average occupancy of CPU and GPU for all the transportation paths for the medium-scale scenes as a representation of the complexity. The complexity of the reinforcement learning algorithm was manifested mainly for the test set runs. Table 6 showed that the APF-TD3 algorithm had smaller average occupancy of CPU and GPU than the other compared algorithms with a value of 1.575%.

### 3.5. Experimental results for large-scale scenes

The large-scale scenes had 10 tasks distributed in one container at shore bridge A, three containers at shore bridge B, four containers at shore bridge C, two containers at shore bridge D, as well as two storage locations at yard E, three storage locations at yard F, two storage locations at yard G, one storage location at yard H, and two storage locations at yard I. The AGV started to transport containers to the yard using the C shore bridge as the starting point. The scene layout had 4 shore bridges and 50 yards, which was shown in Fig. 7. The calculated scheduling tasks order of the large-scale scenes in the port was shown in Table 7, which needed to calculate a total of 13 paths. The AGVs were used to transport two containers in yard E, three containers in yard F, one container in yard H and one container in yard G, two containers in yard I, and one container in yard G. In large-scale scenes the short distances had four transportation paths in total, which were C-E, E-D, D-E, and H-A, and the medium distances had four transportation paths in total, which were D-F, F-C, C-F, and C-H, and the long distances had five transportation paths in total, which were A-G, G-B, B-I, I-B, and B-G.

The large-scale scenes had strong container turnaround capacity and transportation capability. Fig. 8 showed the trajectory of 13 transportation paths. The results in Table 8 showed that APF-TD3 model we



**Fig. 6.** Path results of four algorithms under different transportation paths in a medium-scale scene: (a) path results generated by the APF-TD3 algorithm; (b) path results generated by the APF-DDPG algorithm; (c) path results generated by the APF algorithm; and (d) path results generated by the RRT algorithm.

**Table 5**

Comparison of experimental data with different algorithms in medium-scale scenes.

transportation path	comparison algorithm	path length (meter)	GS	LS	calculation time (second)
B-D	APF-TD3	<b>14.479</b>	<b>0.448°</b>	<b>2.497°</b>	<b>0.190</b>
	APF-DDPG	14.609	0.513°	3.433°	0.570
	APF	15.517	1.341°	5.000°	0.070
	RRT	14.897	5.879°	43.224°	10.620
D-B	APF-TD3	<b>14.354</b>	<b>0.242°</b>	<b>1.142°</b>	<b>0.210</b>
	APF-DDPG	14.355	0.250°	1.192°	0.540
	APF	14.401	0.362°	1.445°	0.070
	RRT	14.828	6.454°	38.183°	8.784
B-E	APF-TD3	<b>24.065</b>	<b>0.278°</b>	<b>2.398°</b>	<b>0.310</b>
	APF-DDPG	24.202	0.770°	10.924°	0.670
	APF	24.924	0.971°	5.000°	0.120
	RRT	25.155	6.278°	43.397°	20.566
E-A	APF-TD3	<b>23.447</b>	<b>0.662°</b>	<b>5.000°</b>	<b>0.320</b>
	APF-DDPG	23.629	0.685°	9.533°	0.730
	APF	25.681	1.449°	5.000°	0.110
	RRT	23.990	8.153°	40.364°	35.178
A-E	APF-TD3	<b>24.548</b>	<b>0.515°</b>	<b>8.269°</b>	<b>0.320</b>
	APF-DDPG	25.017	0.787°	10.640°	0.650
	APF	25.132	1.149°	11.378°	0.120
	RRT	24.811	5.397°	41.937°	15.999
E-C	APF-TD3	<b>33.666</b>	<b>0.557°</b>	<b>3.299°</b>	<b>0.460</b>
	APF-DDPG	36.666	0.744°	10.942°	0.970
	APF	37.035	1.897°	8.234°	0.170
	RRT	37.147	9.755°	45.823°	36.997
C-F	APF-TD3	<b>45.319</b>	<b>0.583°</b>	<b>7.452°</b>	<b>0.700</b>
	APF-DDPG	52.175	1.429°	11.115°	1.470
	APF	50.581	2.021°	17.170°	0.230
	RRT	50.173	8.851°	44.960°	35.700
F-C	APF-TD3	<b>45.649</b>	<b>0.772°</b>	<b>8.062°</b>	<b>0.600</b>
	APF-DDPG	48.249	0.808°	20.763°	1.450
	APF	49.535	1.492°	11.305°	0.240
	RRT	48.897	8.826°	45.274°	50.168

**Table 6**

Total length of transportation paths with different algorithms in medium-scale scenes.

medium-scale	path length (meter)	average occupancy of CPU and GPU
APF-TD3	<b>270.847</b>	<b>1.575%</b>
APF-DDPG	291.075	2.063%
APF	293.386	3.100%
RRT	290.070	3.600%

proposed was able to optimize the path length, average turning angle, and degree of twists and turns for different transportation paths by increasing the complexity of the port environment. The APF-DDPG algorithm was able to optimize the path length metrics or the difference was very small compared to the APF algorithm. The RRT algorithm outperformed the APF algorithm but was inferior to the APF-DDPG algorithm for all transportation paths in long distances. The short distances E-D and D-E transportation paths were calculated in the APF-TD3 framework with a difference of 0.011 m in path length, a difference of 0.474° in average turning angle and a difference of 2.684° in the degree of twists and turns. The medium distances F-C and C-F transportation paths had a difference of 0.003 m in path length, 0.001° in average turning angle, and 0.757° in degree of twists and turns. The long distances G-B and B-G transportation paths had a difference of 0.327 m in path length, 0.429° in average turning angle, and 3.258° in degree of twists and turns. The path length of long distances B-I and I-B transportation paths had a difference of 0.061 m, the average turning angle had a difference of 0.014°, and the degree of twist had a difference of 1.018°. We could observe that transportation paths with interchangeable starting and ending positions under the APF-TD3 algorithm such as F-C and C-F with B-I and I-B had relatively low environmental complexity of the transportation paths, and the difference among the

path metrics was smaller, but the B-I path was longer than the F-C path.

We were able to classify 13 AGV driving paths according to the complexity of the paths based on the large-scale port environment. The classification was based on the number of obstacles avoided by the AGV as a criterion for the complexity of the environment. The environment complexity of F-C, C-F, C-H, H-A, B-I and I-B was lower in which fewer obstacles needed to be avoided, and D-F, G-B and B-G were higher in which more obstacles needed to be avoided. According to Table 8 we could observe that the APF-TD3 algorithm could solve optimal solutions of any complexity. APF-DDPG algorithm could satisfy the path length better than APF algorithm and RRT algorithm but average turning angle and degree of twists could not be better than APF algorithm in all environments. RRT algorithm was a sample extended search algorithm which was better than APF algorithm in environments where it could search for more obstacles such as transportation paths in A-G, B-I and B-G. The APF algorithm had relative advantages in calculation efficiency, but the APF-TD3 algorithm not only had absolute advantages in path length and route safety in the calculation efficiency of the response time was relatively fast.

Our APF-TD3 framework still showed its advantage in large-scale scenes, where the total path length results were shown in Table 9. The optimal AGV driving path length for total AGV driving was 496.389 m, which was followed by APF-DDPG algorithm with a difference of 4.672 m from it. The APF-DDPG algorithm saved 20.563 m in total path length than the APF algorithm, and the difference in total path length between the APF algorithm and the RRT algorithm was 2.543 m. As the complexity of the port environment increases, the RRT algorithm had an advantage over the APF algorithm in total path length, but it had a large difference with the APF in both the average turn angle and the degree of twists and turns. APF-DDPG algorithm had some advantages over APF algorithm, but APF-TD3 algorithm showed obvious advantages over APF-DDPG algorithm. The complexity of the APF-TD3 algorithm showed a significant advantage in the average occupancy of its CPU and GPU was only 2.177%. Therefore, our framework could be applied not only to short, medium and long distances, but also to deal with the complexity of the port scale, which could guarantee the safety of the AGV and reduce the power consumption during the driving process.

#### 4. Discussion

AGV undertake the main tasks of container transportation in port operations and transportation path optimization is also the key to the efficient operation of automated container terminals. As the main transportation tool of port operational the AGV driving distance, driving speed, energy consumption, scheduling efficiency, fault response, and emergency situations all of which affect the turnover rate of the port (Sun et al., 2023a). Where speed, energy consumption, and efficiency are directly or indirectly affecting the driving distance of AGV, and fault response and emergencies are affecting the driving safety of AGV. Path planning based on port environment is usually used to study the driving distance and emergency situations, which is committed to minimize the driving distance and reach the end point safely. Currently, many scholars have constructed the connection between speed, energy consumption, and efficiency of AGVs through multiple constraints and objectives, which is finally reflected by the driving distance. From the overall operation and management of the port that is a process of continuous enrichment, how to manifest the safety and efficiency of the operation from multiple perspectives (Shu et al., 2023b). Moreover, AGV itself also exists malfunctioning situation, how to carry out fault detection and rapid repair in the unmanned port is also one of the important issues in the automated container port advancement. Multi-AGV coordination is also the norm in port operations, our algorithmic framework is based on obstacle avoidance based on repulsive force between obstacles and AGVs, where other moving AGVs need to be considered as obstacles for repulsive force calculation. Combining our algorithmic framework with lidar, infrared, and camera for mobile

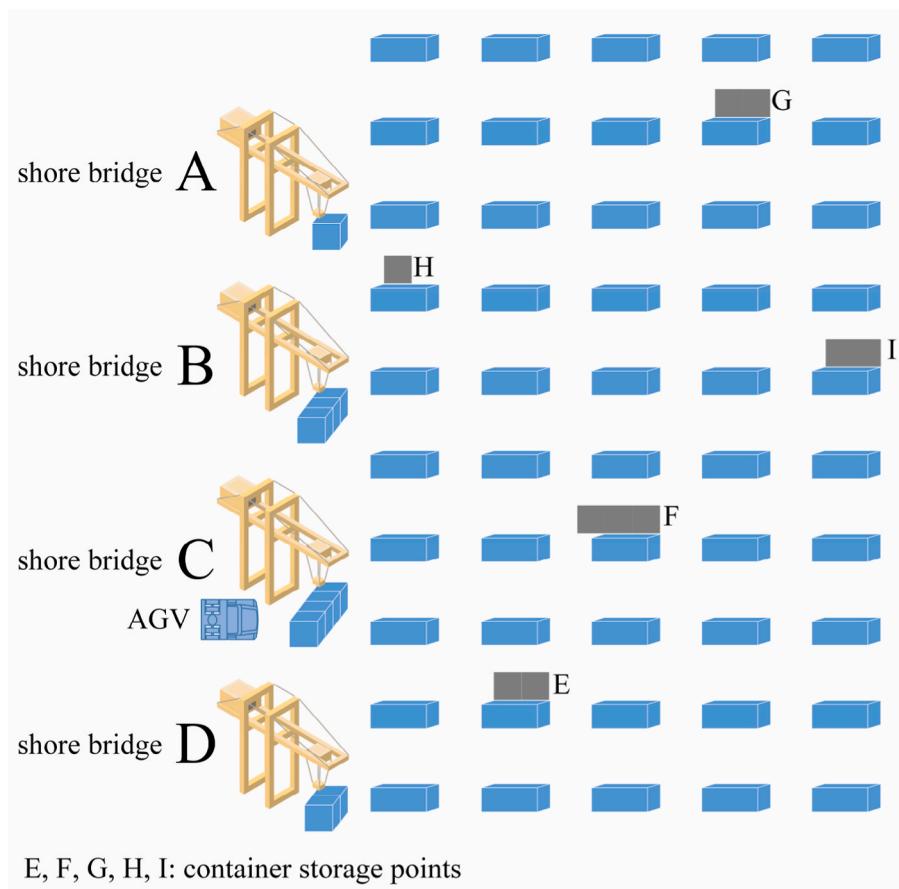


Fig. 7. Distribution of large-scale scenes port.

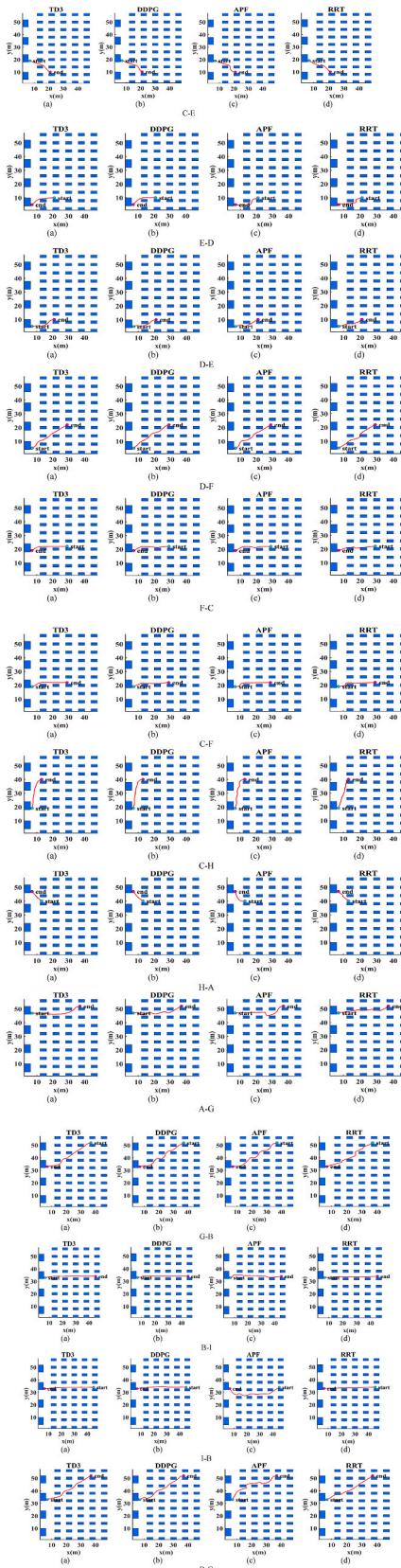
**Table 7**  
AGV transportation paths in large-scale scenes.

large-scale	transportation path	AGV transportation container (1 denotes transportation)
1	C-E	1
2	E-D	
3	D-E	1
4	E-D	
5	D-F	1
6	F-C	
7	C-F	1
8	F-C	
9	C-F	1
10	F-C	
11	C-H	1
12	H-A	
13	A-G	1
14	G-B	
15	B-I	1
16	I-B	
17	B-I	1
18	I-B	
19	B-G	1

obstacle decision can be adapted to multi-AGV systems, which provides theoretical and technological support for the research of cooperative scheduling of multi-AGVs (Wei et al., 2023; Yang et al., 2023). Existing research mainly focuses on the path optimization under magnetic nail fixed route and environmental perception (Chen et al., 2024). The path under the magnetic nail fixed path is straight and the turning angle is 90°, which is usually optimized from the path length. Environmentally perception paths are curved and not only the path length but also the smoothness and safety of the paths need to be evaluated (Shu et al.,

2023a). There are few studies on optimizing the safety of path routes from the perspective of efficient port operations. We propose an APF-TD3 framework to optimize the whole process of scheduling planning for a single AGV in a port. It is evident from the simulation results that our proposed framework has significant advantages but the gap between its comparison algorithms is not the same in different environments. AGV driving path is more complex, and the gap between the comparison algorithms is larger, and the degree of path twist is also higher. The difference between the comparison algorithms and evaluation metrics is smaller for shorter path lengths. Optimizing the path length can improve the efficiency of port operations and make the AGV work better. Optimizing the average turning angle of the path can evaluate the safety of the whole path from the global perspective. Optimizing the degree of twists and turns of the path can improve the smoothness of the path in different road sections from the local perspective (Gan et al., 2022; Xiao et al., 2023). Optimizing the path length, the average turning angle of the path, and the degree of twists and turns of the path in real port transportation can reduce the energy consumption of AGV, which in turn attenuates the carbon emissions during transportation and makes a small contribution to the sustainability of the environment (Yin et al., 2021).

Along with the change of port size and environment complexity, in which our framework can adapt to different sizes of port environments and avoid obstacles successfully to accomplish scheduling tasks. Moreover, APF-TD3 achieves excellent results: path length, average turning angle of the path, and the degree of twists and turns of the path in different path lengths and different driving directions. It improves the efficiency of AGV arrival and the safety of AGV driving path under the environment perception. Complexity and calculation efficiency are outperformed over reinforcement learning algorithms with the same structure and RRT algorithms. It would help us to better understand the



**Fig. 8.** Path results of four algorithms under different transportation paths in a large-scale scene: (a) Path results generated by the APF-TD3 algorithm; (b) Path results generated by the APF-DDPG algorithm; (c) Path results generated by the APF algorithm; and (d) Path results generated by the RRT algorithm.

**Table 8**

Comparison of experimental data with different algorithms in large-scale scenes.

transportation path	comparison algorithm	path length (meter)	GS	LS	calculation time (second)
C-E	APF-TD3	<b>17.107</b>	<b>0.570°</b>	<b>8.636°</b>	<b>0.280</b>
	APF-DDPG	17.413	0.667°	10.383°	0.670
	APF	18.048	1.044°	9.686°	0.120
	RRT	18.932	9.153°	42.280°	14.292
E-D	APF-TD3	<b>15.661</b>	<b>0.335°</b>	<b>5.982°</b>	<b>0.310</b>
	APF-DDPG	15.960	0.384°	6.785°	0.810
	APF	16.956	0.848°	13.977°	0.110
	RRT	17.546	7.225°	45.815°	14.392
D-E	APF-TD3	<b>15.650</b>	<b>0.809°</b>	<b>3.298°</b>	<b>0.300</b>
	APF-DDPG	15.681	1.372°	5.000°	0.680
	APF	16.051	1.092°	5.000°	0.110
	RRT	16.098	8.360°	46.919°	28.898
D-F	APF-TD3	<b>28.268</b>	<b>0.571°</b>	<b>3.605°</b>	<b>0.520</b>
	APF-DDPG	28.394	0.727°	4.724°	0.920
	APF	28.829	0.794°	4.381°	0.200
	RRT	30.950	9.337°	42.847°	24.550
F-C	APF-TD3	<b>22.826</b>	<b>0.297°</b>	<b>2.921°</b>	<b>0.400</b>
	APF-DDPG	22.836	0.400°	9.403°	1.940
	APF	22.891	0.464°	7.269°	0.170
	RRT	23.782	7.516°	40.354°	40.220
C-F	APF-TD3	<b>22.823</b>	<b>0.298°</b>	<b>3.678°</b>	<b>0.470</b>
	APF-DDPG	22.825	0.640°	5.000°	0.860
	APF	23.079	0.599°	5.000°	0.160
	RRT	24.550	5.401°	47.421°	18.426
C-H	APF-TD3	<b>22.825</b>	<b>0.456°</b>	<b>3.112°</b>	<b>0.380</b>
	APF-DDPG	23.335	0.808°	11.725°	0.790
	APF	23.622	1.122°	5.000°	0.150
	RRT	24.364	9.714°	49.502°	20.136
H-A	APF-TD3	<b>9.243</b>	<b>0.136°</b>	<b>1.127°</b>	<b>0.190</b>
	APF-DDPG	9.337	0.553°	1.894°	0.580
	APF	10.375	0.882°	4.345°	0.070
	RRT	10.276	11.855°	45.727°	30.915
A-G	APF-TD3	<b>31.819</b>	<b>0.523°</b>	<b>4.464°</b>	<b>0.650</b>
	APF-DDPG	32.465	1.066°	5.000°	1.040
	APF	34.343	1.287°	9.023°	0.230
	RRT	33.252	6.992°	52.118°	32.839
G-B	APF-TD3	<b>37.042</b>	<b>1.077°</b>	<b>8.587°</b>	<b>0.670</b>
	APF-DDPG	38.284	1.227°	11.002°	1.900
	APF	38.139	1.342°	15.433°	0.260
	RRT	40.458	7.449°	38.931°	44.196
B-I	APF-TD3	<b>38.099</b>	<b>0.122°</b>	<b>3.032°</b>	<b>0.730</b>
	APF-DDPG	38.188	0.155°	3.252°	0.990
	APF	39.943	1.947°	5.000°	0.270
	RRT	38.354	1.573°	40.748°	25.557
I-B	APF-TD3	<b>38.038</b>	<b>0.108°</b>	<b>2.014°</b>	<b>1.030</b>
	APF-DDPG	38.218	0.338°	5.494°	1.630
	APF	42.855	1.101°	9.649°	0.260
	RRT	38.462	1.655°	33.461°	22.218
B-G	APF-TD3	<b>36.715</b>	<b>0.648°</b>	<b>5.329°</b>	<b>0.770</b>
	APF-DDPG	37.261	1.226°	14.836°	1.940
	APF	40.422	1.593°	8.485°	0.260
	RRT	38.125	7.831°	44.934°	49.055

**Table 9**

Total length of transportation paths with different algorithms in large-scale scenes.

large -scale	path length (meter)	average occupancy of CPU and GPU
APF-TD3	<b>496.389</b>	<b>2.177%</b>
APF-DDPG	501.062	2.577%
APF	524.168	2.900%
RRT	521.625	3.600%

path selection and movement decision-making process of AGV in the port environment to obtain safer and more efficient path planning results, which is significant for the optimized management of automated terminals. It can effectively combine and optimize with lidar, camera, infrared and other equipment to realize port intelligent detection and integrated scheduling optimization. It has a very significant role in

promoting the construction of automated container terminal optimization management. Path planning in more complex environments, collaborative path planning with multiple intelligences, and real-time path planning in dynamic environments have potential implications for the future.

## 5. Conclusion

Optimization of AGV driving path is a key step in promoting the development of automated container terminals, which ensures the safety and smoothness of driving path and is one of the necessary conditions for efficient operation of ports. We propose an APF-TD3 modeling framework for generating AGV driving paths, which can quickly respond to safe and smooth transportation paths under different environmental scales. The framework consists of four steps: the formulation of the AGV scheduling task, the potential field force constructed by the APF on the target point and obstacles, the selection of the optimal action strategy for TD3, and the generation of the optimal transportation path. We obtain the scheduling task list of a single AGV in different environment scales by selecting the shortest transportation distance as the next task execution in an enumeration manner sequentially. Second step based on the APF from the attraction to the target point and the repulsion of obstacles as well as the position information of the AGV itself will be used as the input data for TD3. Third step selects the optimal action strategy from all the actions and executes that based on the input state information and the designed reward mechanism. It is necessary to execute the second and third steps repeatedly until the AGV arrives at the container storage point. Finally, the transportation path from the starting point to the end point can be obtained, and we set up three scenes of small-scale, medium-scale and large-scale simulation experiments according to the environment of port transportation. We take APF-DDPG, APF, and RRT as the comparison algorithms, which show that the shortest transportation paths are obtained by our proposed framework in the scheduling and transportation tasks of small-scale, medium-scale, and large-scale scenes, which is 27.519 m, 270.847 m, and 496.389 m. It has significant advantages over the comparative algorithms in both safety and smoothness of the path, which helps to guarantee the efficiency of port operations and reduce the occurrence of dangerous accidents.

While our method shows satisfactory results in the safety and smoothness of AGV path optimization in port, but there are still some research works that can be further improved in the complex environment of port. Firstly, we consider the transportation routes of a single AGV, but the operation of a port is a multi-number AGVs environment. Along with the increased complexity of the environment and the requirement of equipment synergy, that the formulation of scheduling tasks will become a hot issue in future research. Secondly, combining sensors such as laser radar, infrared, and camera with APF can recognize static and dynamic obstacles which will guarantee the safety of AGV driving in extremely complex environments. TD3 is optimized for experience playback and noise introduction so that the AGV can select the optimal action more quickly, while balancing exploration and exploitation to achieve a better learning effect. China automated ports such as Shanghai Port and Qingdao Port are divided into different port areas according to the different cargoes loaded for transportation, in which the layout of the port area is mainly based on vertical layout terminals and parallel layout terminals. Our modeling framework matches the layout in real scenes from shoreline layout, but the considered objectives are not enough. Environmental perception and testing in realistic scenes can be performed based on the data of multiple devices (lidar, infrared, camera) under the port in the future.

## CRediT authorship contribution statement

**Xinqiang Chen:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Formal analysis,

Conceptualization. **Shuhao Liu:** Writing – original draft, Visualization, Validation, Methodology, Investigation. **Jiansen Zhao:** Formal analysis, Conceptualization. **Huafeng Wu:** Writing – review & editing, Supervision, Project administration, Methodology, Conceptualization. **Jiangfeng Xian:** Writing – review & editing, Methodology, Conceptualization. **Jakub Montewka:** Writing – review & editing, Data curation, Conceptualization.

## Declaration of competing interest

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

## Data availability

Data will be made available on request.

## Acknowledgments

This work was jointly supported by National Natural Science Foundation of China (52331012, 52102397, 52071200).

## References

- Ai, B., Jia, M.X., Xu, H.W., Xu, J.L., Wen, Z., Li, B.S., Zhang, D., 2021. Coverage path planning for maritime search and rescue using reinforcement learning. *Ocean Eng.* 241.
- Buermann, J., Zhang, J., 2022. Multi-robot adversarial patrolling strategies via lattice paths. *Artif. Intell.* 311.
- Chen, J.H., Ye, J., Zhuang, C.L., Qin, Q.D., Shu, Y.Q., 2022a. Liner shipping alliance management: Overview and future research directions. *Ocean Coast Manag.* 219.
- Chen, J.H., Zhuang, C.L., Xu, H., Xu, L., Ye, S.M., Rangel-Buitrago, N., 2022b. Collaborative management evaluation of container shipping alliance in maritime logistics industry: CKYHE case analysis. *Ocean Coast Manag.* 225.
- Chen, W.J., Chen, J.H., Geng, J.J., Ye, J., Yan, T., Shi, J., Xu, J.H., 2023a. Monitoring and evaluation of ship operation congestion status at container ports based on AIS data. *Ocean Coast Manag.* 245.
- Chen, X., Wang, M., Ling, J., Wu, H., Wu, B., Li, C., 2024. Ship imaging trajectory extraction via an aggregated you only look once (YOLO) model. *Eng. Appl. Artif. Intell.* 130, 107742.
- Chen, X.Q., Wu, H., Han, B., Liu, W., Montewka, J., Liu, R.W., 2023b. Orientation-aware ship detection via a rotation feature decoupling supported deep learning approach. *Eng. Appl. Artif. Intell.* 125.
- Chen, Y.H., Wu, S.F., 2023. Framework of active obstacle avoidance for autonomous vehicle based on hybrid soft actor-critic algorithm. *J. Transport. Eng. Part a-Systems* 149.
- Dirik, M., Kocamaz, A.F., Castillo, O., 2020. Global path planning and path-following for wheeled mobile robot using a novel control structure based on a vision sensor. *Int. J. Fuzzy Syst.* 22, 1880–1891.
- Dudeja, C., Kumar, P., 2022. An improved weighted sum-fuzzy Dijkstra's algorithm for shortest path problem (IWSFDA). *Soft Comput.* 26, 3217–3226.
- Gan, L.X., Yan, Z.X., Zhang, L., Liu, K.Z., Zheng, Y.Z., Zhou, C.H., Shu, Y.Q., 2022. Ship path planning based on safety potential field in inland rivers. *Ocean Eng.* 260.
- Guo, H.S., Ren, Z.G., Lai, J.L., Wu, Z.Z., Xie, S.L., 2023. Optimal navigation for AGVs: a soft actor-critic-based reinforcement learning approach with composite auxiliary rewards. *Eng. Appl. Artif. Intell.* 124.
- Han, G.J., Yang, X., Liu, L., Zhang, W.B., Guizani, M., 2020. A disaster management-oriented path planning for mobile anchor node-based localization in wireless sensor networks. *IEEE Transact. Emerg. Topics In Comput.* 8, 115–125.
- Hong, D., Lee, S., Cho, Y.H., Baek, D., Kim, J., Chang, N., 2021. Energy-efficient online path planning of multiple drones using reinforcement learning. *IEEE Trans. Veh. Technol.* 70, 9725–9740.
- Hou, W.B., Xiong, Z.H., Wang, C.S., Chen, H., 2022. Enhanced ant colony algorithm with communication mechanism for mobile robot path planning. *Robot. Autonom. Syst.* 148.
- Hu, H.T., Yang, X.R., Xiao, S.C., Wang, F.Y., 2023. Anti-conflict AGV path planning in automated container terminals based on multi-agent reinforcement learning. *Int. J. Prod. Res.* 61, 65–80.
- Jeng, S.L., Chiang, C.H., 2023. End-to-End autonomous navigation based on deep reinforcement learning with a survival penalty function. *Sensors* 23.
- Jiang, H.G., Esfahani, M.A., Wu, K.Y., Wan, K.W., Heng, K.K., Wang, H., Jiang, X.D., 2022a. iTD3-CLN: learn to navigate in dynamic scene through Deep Reinforcement Learning. *Neurocomputing* 503, 118–128.
- Jiang, W., Lyu, Y.X., Li, Y.F., Guo, Y.C., Zhang, W.G., 2022b. UAV path planning and collision avoidance in 3D environments based on POMDP and improved grey wolf optimizer. *Aero. Sci. Technol.* 121.

- Kim, M., Han, D.K., Park, J.H., Kim, J.S., 2020. Motion planning of robot manipulators for a smoother path using a twin delayed deep deterministic policy gradient with hindsight experience replay. *Appl. Sci.* Basel 10.
- Li, H., Peng, J.B.A., Wang, X., Wan, J.L., 2021a. Integrated resource assignment and scheduling optimization with limited critical equipment constraints at an automated container terminal. *IEEE Trans. Intell. Transport. Syst.* 22, 7607–7618.
- Li, M.Z., Li, B., Qi, Z.G., Li, J.S., Wu, J.W., 2023. Optimized APF-ACO algorithm for ship collision avoidance and path planning. *J. Mar. Sci. Eng.* 11.
- Li, X.D., Peng, Y., Huang, J., Wang, W.Y., Song, X.Q., 2021b. Simulation study on terminal layout in automated container terminals from efficiency, economic and environment perspectives. *Ocean Coast Manag.* 213.
- Li, Y.F., Li, J.B., Zhou, W.H., Yao, Q.W., Nie, J., Qi, X.C., 2022. Robot path planning navigation for dense planting red jujube orchards based on the joint improved A\* and DWA algorithms under laser SLAM. *Agriculture-Basel* 12.
- Lian, Y.D., Xie, W., Yang, Q.F., Liu, Y.R., Yang, Y.B., Wu, A.G., Eisaka, T., 2022. Improved coding landmark-based visual sensor position measurement and planning strategy for multiwarehouse automated guided vehicle. *IEEE Trans. Instrum. Meas.* 71.
- Liu, Q.H., Wang, C.Y., Li, X.Y., Gao, L., 2023. An improved genetic algorithm with modified critical path-based searching for integrated process planning and scheduling problem considering automated guided vehicle transportation task. *J. Manuf. Syst.* 70, 127–136.
- Luan, P.G., Thinh, N.T., 2023. Hybrid genetic algorithm based smooth global-path planning for a mobile robot. *Mech. Base. Des. Struct. Mach.* 51, 1758–1774.
- Lv, L.H., Zhang, S.J., Ding, D.R., Wang, Y.X., 2019. Path planning via an improved DQN-based learning policy. *IEEE Access* 7, 67319–67330.
- Peng, C., Liu, X.M., Ma, J.J., 2023. Design of safe optimal guidance with obstacle avoidance using control barrier function-based actor-critic reinforcement learning. *IEEE Transact. Syst. Man Cybernetics-Syst.* 53, 6861–6873.
- Samal, K., Kumawat, H., Saha, P., Wolf, M., Mukhopadhyay, S., 2022. Task-driven RGB-lidar fusion for object tracking in resource-efficient autonomous system. *IEEE Transact. Intellig. Vehicles* 7, 102–112.
- Savas, Y., Ornik, M., Cubuktepe, M., Karabag, M.O., Topcu, U., 2020. Entropy maximization for markov decision processes under temporal logic constraints. *IEEE Trans. Automat. Control* 65, 1552–1567.
- Shu, Y.Q., Hu, A.Y., Zheng, Y.Z., Gan, L.X., Xiao, G.N., Zhou, C.H., Song, L., 2023a. Evaluation of ship emission intensity and the inaccuracy of exhaust emission estimation model. *Ocean Eng.* 287.
- Shu, Y.Q., Zhu, Y.J., Xu, F., Gan, L.X., Lee, P.T.W., Yin, J.C., Chen, J.H., 2023b. Path planning for ships assisted by the icebreaker in ice-covered waters in the Northern Sea Route based on optimal control. *Ocean Eng.* 267.
- Sun, B.F., Zhai, G.S., Li, S., Pei, B., 2023a. Multi-resource collaborative scheduling problem of automated terminal considering the AGV charging effect under COVID-19. *Ocean Coast Manag.* 232.
- Sun, P.Z.H., You, J.P., Qiu, S.Q., Wu, E.Q., Xiong, P.W., Song, A.G., Zhang, H.Z., Lu, T., 2023b. AGV-based vehicle transportation in automated container terminals: a survey. *IEEE Trans. Intell. Transport. Syst.* 24, 341–356.
- Tang, G., Tang, C.Q., Claramunt, C., Hu, X., Zhou, P.P., 2021. Geometric A-star algorithm: an improved A-star algorithm for AGV path planning in a port environment. *IEEE Access* 9, 59196–59210.
- Tao, Q.Y., Sang, H.Y., Guo, H.W., Wang, P., 2021. Improved particle swarm optimization algorithm for AGV path planning. *IEEE Access* 9, 33522–33531.
- Tao, X.Y., Lang, N., Li, H.P., Xu, D.M., 2022. Path planning in uncertain environment with moving obstacles using warm start cross entropy. *IEEE-Asme Transact. Mechatron.* 27, 800–810.
- Wang, J., Xiao, Y., Chen, C.L.P., Li, T.S., 2023. A jamming aware artificial potential field method to counter GPS jamming for unmanned surface ship path planning. *IEEE Syst. J.* 17, 4555–4566.
- Wei, M.H., He, J.L., Tan, C.M., Yue, J.T., Yu, H., 2023. Quay crane scheduling with time windows constraints for automated container port. *Ocean Coast Manag.* 231.
- Xiao, G.N., Chen, L., Chen, X.Q., Jiang, C.M., Ni, A.N., Zhang, C.Q., Zong, F., 2023. A hybrid visualization model for knowledge mapping: scientometrics, SAOM, and SAQ. *IEEE Trans. Intell. Transport. Syst.*
- Xu, B.W., Jie, D.P., Li, J.J., Yang, Y.S., Wen, F.R., Song, H.T., 2021. Integrated scheduling optimization of U-shaped automated container terminal under loading and unloading mode. *Comput. Ind. Eng.* 162.
- Yang, X.R., Hu, H.T., Jin, J.A., 2023. Battery-powered automated guided vehicles scheduling problem in automated container terminals for minimizing energy consumption. *Ocean Coast Manag.* 246.
- Yang, Y., Zhong, M., Dessouky, Y., Postolache, O., 2018. An integrated scheduling method for AGV routing in automated container terminals. *Comput. Ind. Eng.* 126, 482–493.
- Ye, X.F., Deng, Z.Y., Shi, Y.J., Shen, W.M., 2023. Toward energy-efficient routing of multiple AGVs with multi-agent reinforcement learning. *Sensors* 23.
- Yin, C.Z., Ke, Y.D., Chen, J.H., Liu, M., 2021. Interrelations between sea hub ports and inland hinterlands: perspectives of multimodal freight transport organization and low carbon emissions. *Ocean Coast Manag.* 214.
- Yue, L.J., Fan, H.M., 2022. Dynamic scheduling and path planning of automated guided vehicles in automatic container terminal. *IEEE-CAA J. Automatica Sinica* 9, 2005–2019.
- Zhang, L., Mou, J.M., Chen, P.F., Li, M.X., 2021. Path planning for autonomous ships: a hybrid approach based on improved APF and modified VO methods. *J. Mar. Sci. Eng.* 9.
- Zhang, S.T., Li, Y.B., Dong, Q.H., 2022. Autonomous navigation of UAV in multi-obstacle environments based on a Deep Reinforcement Learning approach. *Appl. Soft Comput.* 115.
- Zhang, W., Wang, N.X., Wu, W.H., 2023a. A hybrid path planning algorithm considering AUV dynamic constraints based on improved A\* algorithm and APF algorithm. *Ocean Eng.* 285.
- Zhang, Z.C., Wang, S.B., Chen, J., Han, Y., 2023b. A bionic dynamic path planning algorithm of the micro UAV based on the fusion of deep neural network optimization/filtering and Hawk-Eye vision. *IEEE Transact. Syst. Man Cybernetics-Syst.* 53, 3728–3740.
- Zhong, M.S., Yang, Y.S., Dessouky, Y., Postolache, O., 2020. Multi-AGV scheduling for conflict-free path planning in automated container terminals. *Comput. Ind. Eng.* 142.
- Zhu, Q.X., Zheng, Z.L., Wang, C., Lu, Y.J., 2023. Research on AGV path tracking method based on global vision and reinforcement learning. *Sci. Prog.* 106.