

# Hybrid Deep Learning Approach for Ship Navigation in Curved River Sections Using PPO and CNN

Jianxia Wan<sup>1,\*</sup>, Sukui Zhang<sup>2</sup>

<sup>1</sup>Navigation College, Jiangsu Maritime Institute Nanjing 211170, China

<sup>2</sup>Digital Engineering Technology Research and Development Center for Maritime Safety and Security, Jiangsu Maritime Institute Nanjing 211170, China

Email: wanjianxia001@126.com, zhangshkfy@163.com

\*Corresponding author

**Keywords:** deep learning, PPO, CNN, path planning, hybrid controller, curved river

**Received:** August 16, 2024

*Curved river sections have complex water flow characteristics and difficulties in maneuvering ships through bends, which pose significant challenges to path planning and ship navigation control. The current path research algorithms still have limitations in dealing with curved and complex waterways. Given this, a convolutional neural network control model based on a hybrid controller and near-end strategy optimization is proposed. This model realizes the path and navigation planning of ships in curved river sections through the hybrid controller. This model utilizes convolutional neural networks to extract channel image features of curved river sections and plans the path through proximal strategy optimization algorithms. In the experiment, high-performance computer processors were used to accelerate the model's training, and the model was validated in a simulation environment. The results showed that when the research model reached 200 iterations in the simulated curved river section, the average reward value was 0.0323, 19.36% higher than the average reward value of other algorithms. The average instantaneous reward of the research model in path planning was 7.95, which was 3.69 and 1.58 higher than the proximal policy optimization model and the convolutional neural network model based on proximal policy optimization, respectively. The success rate of path planning in complex curved river sections was 82%, significantly higher than the other two algorithms, verifying its effectiveness and superiority in complex path planning tasks. Therefore, this study contributes to improving the safety, efficiency, and economic benefits of ship navigation, and promoting the intelligent and automated growth of the shipping industry*

*Povzetek: Razvili so hibridni model globokega učenja, ki združuje konvolucijske nevronske mreže (CNN) in optimizacijo bližnje strategije (PPO), za načrtovanje poti in nadzor navigacije ladij v ukrivljenih rečnih odsekih.*

## 1 Introduction

With the advancement of national strategies such as “Transportation Power” and “Yangtze River Economic Belt”, the importance of the Yangtze River Golden Waterway is increasingly prominent [1]. In recent years, with the opening of the 12.5-meter-deep water channel in the Yangtze River, the trend of large-scale ships has significantly accelerated, and the number of ships entering the Yangtze River has increased significantly [2]. However, as a typical section of this waterway, the curved section of the Yangtze River has a complex and variable water flow environment, and the channel is narrow and curved [3], which to some extent limits the navigation capacity of ships. When ships navigate in these areas, they need to deal with sharp turns, narrow passages, and variable water flows [4], which puts high demands on traditional navigation and control systems.

The traditional Ship Navigation Control (SNC) system mainly relies on experienced captains and crew members to operate. However, when facing complex and constantly changing curved river sections, it is often difficult to make optimal decisions, leading to accidents. Therefore, it is imperative to address the urgent need to develop effective and reliable autonomous navigation solutions for ships navigating dynamic and uncertain curved river sections.

Deep Learning (DL) is an algorithm that allows computers to simulate the learning and thinking processes of the human brain [5]. Alaba S Y et al. proposed a normalized difference vegetation index method based on long short-term memory DL. The method analyzed and predicted the time series data of normalized difference index by introducing the long short-term memory DL algorithm, and used grid search optimization to improve the prediction performance of the model. The

experimental results showed that the prediction performance of the method was higher than the current state-of-the-art methods, and the root mean square error was smaller, which is suitable for environmental monitoring climate change assessment [6]. Kumar et al. proposed a Convolutional Neural Network (CNN) hierarchical DL network for detecting leaf diseases. This method introduced intuitionistic fuzzy local binary patterns to extract leaf features and then used CNN for disease detection and classification. The evaluation results showed that this method helped prevent leaf diseases, thereby increasing potato yield [7]. Uyulan et al. proposed a diagnostic model for severe depression based on electroencephalography and CNN. This method combined computational neuroscience and deep CNN architecture to model and analyze the collected EEG data in major frequency bands to differentiate patients with severe depression. The model achieved classification accuracy of 89.33% and 92.66% on the data, significantly improving diagnostic speed and accuracy [8]. Minagawa et al. proposed a method for diagnosing skin tumors using Deep Neural Networks (DNNs). This method trained DNNs using images from the International Skin Image Dataset and the Shinshu Dataset in Japan and compared them with the diagnostic results of Japanese dermatologists. This indicated that DNN could improve the accuracy of dermatologists in detecting skin tumors in non-local populations in clinical practice [9]. Rezaee K et al. proposed a UAV crowd sensing and DL-based path selection method for emergency medical vehicles. The method analyzed the collected public transport video frames by DL, extracted the characteristic paths on different routes, and calculated the most suitable route for the paramedic vehicle to transfer the patient, avoiding congested places in the traffic. Experimental results showed that the method was more accurate and computationally efficient than current state-of-the-art methods [10].

There are also many studies on SNC. Tang et al. designed an automatic tracking control system for estimating the curvature of a narrow waterway with small curvature through an observer. They manipulated the unmanned fleet to flexibly pass-through narrow waterways in series through the control system. The effectiveness of the system in controlling unmanned

vessels through narrow waterways within the laboratory has been verified [11]. Z. Zeng et al. proposed an adaptive sampling tree algorithm for underwater vehicle path planning. The algorithm fused a point selection sampling strategy and an information heuristic search process as an algorithmic framework and introduced a fast exploration random tree algorithm to sample the path region. Simulation results showed that the algorithm had higher performance, path planning solution speed, and stability than current underwater vehicle planning algorithms [12]. J. Zhang et al. presented an adaptive surface path planning method based on heterogeneous autonomous underwater robots. The method incorporated meta-heuristics to balance global and local path exploration of underwater paths by integrating two optimizers, used conditional convergence factors to avoid the problem of the method falling into a local optimum, and considered the effects of sea currents at different locations. The validation results showed that the method converged well and had significant advantages in underwater path planning in complex marine environments [13]. Yu et al. proposed an economically efficient and safety driven path planning method for unmanned ships. This method utilized the Internet of Things, artificial intelligence simulation technology, and geographic information systems, combined with maritime transportation navigation assistance and decision support systems, to reduce human errors related to maritime accidents. This method improved navigation safety and operational efficiency in curved river sections and seaports [14]. J. Wang et al. proposed a path planning method for unmanned surface ships based on the Artificial Potential Field Method (APFM) and GPS. The method employed a signal strength detection process to ascertain the occurrence of signal interference in the surrounding area. This was followed by the utilization of the APFM to determine the position of the source of the interference. Finally, the path was re-planned based on the combined force around the target. The experimental results showed that this method could effectively solve the interference of the positioning system on the path planning and improve the efficiency of planning [15].

The literature review table summarized in this research is shown in Table 1.

Table 1: Summary table of literature reviews.

Research	Literature	Research worker	Key findings	Latent deficiency
DL Algorithms	6	Alaba S Y et al.	A normalized vegetation prediction method based on long and short-term memory was designed to improve the prediction performance with less error	Limited applicability for Normalized Vegetation Index (NVI) predictions only
	7	Kumar A et	Proposed CNN-based leaf disease	May not be able to handle

		al.	detection combined with fuzzy local patterns to enhance disease detection capability	complex and diverse disease types
	8	Uyulan C et al.	Using EEG with CNN for depression diagnosis, CNN is effective in differentiating depressed patients with high accuracy rates	Reliance on EEG data, possible data bias
	9	Minagawa A et al.	Using DNNs for skin tumour diagnosis, DNNs improve diagnostic accuracy for non-native skin tumours	Only applicable to specific populations, need to expand sample
	10	Rezaee K et al.	Designing DL and UAV-based crowd-aware medical path selection, DL to improve the accuracy and efficiency of medical vehicle path selection	Effectiveness in severe weather conditions to be verified
<b>Vessel navigation control</b>	11	Tang C et al.	Designed an automatic tracking control system for narrow waterways with small curvature, which can flexibly control unmanned vessels through narrow waterways.	Experiments are limited to the laboratory environment
	12	Z. Zeng et al.	Designed adaptive sampling tree algorithm for underwater vehicle path planning with higher performance, path planning solving speed and stability	Validation of applicability to other environments is required
	13	J. Zhang et al.	Designed an adaptive surface path planning method for heterogeneous autonomous underwater robots to improve the efficiency of path planning for complex marine environments by fusing meta-heuristic methods	Unable to handle boundary identification in extreme weather
	14	Yu H et al.	Economic efficiency and safety-driven system-driven vessel path planning improves navigational safety and operational efficiency in curved river sections	Adaptability to complex weather and traffic conditions needs to be verified
	15	J. Wang et al.	Propose a path planning method for unmanned surface ships based on APFM and global positioning system to reduce the interference of positioning system on path planning and improve the efficiency of planning	Need to validate its planning effect in curves

In summary, the DL algorithm is extensively utilized in life. There have been many studies on the navigation of ships in curved river sections by researchers, but there is still relatively little research on the autonomous navigation of ships in curved sections. To improve the navigation efficiency and safety of ships in curved river sections, a more intelligent and dynamic control strategy is needed. Based on this, this study proposes a ship

navigation path planning model for curved river sections, namely the HC-PPO-CNN. It is expected to address the complex navigation challenges in curved river sections and improve the safety and efficiency of water transportation.

The innovation of this study lies in (1) The first point is to combine Proximal Policy Optimization (PPO) and CNN model for path planning, extract channel image

features through CNN, and optimize them by combining the PPO algorithm; (2) The second point is to introduce hybrid control to achieve dynamic weight adjustment to improve the adaptability in complex weather and curved waterways. The research content has three parts. Part 1 elaborates on the construction of the new model, Part 2 is the performance testing of the model, and Part 3 is the discussion and analysis of experimental results.

## 2 Methods and materials

### 2.1 Construction of ship curved river path planning model based on PPO-CNN

When ships navigate curved river sections, they often face complex water flow environments and narrow waterways, which pose challenges to navigation control. The constantly changing water flow velocity and direction in curved river sections make the navigation environment more complex [17]. To address these challenges, this study presents a path planning model based on PPO. The PPO is known for its excellent stability and efficiency in deep reinforcement learning and can achieve policy optimization in complex environments [18]. By combining PPO and CNN, this study constructs a hybrid model, PPO-CNN, which enables intelligent path planning and navigation in dynamically changing curved river sections, effectively

adapting to changing river conditions. This study is based on the application concept of DL and designs a navigation framework and path planning, as shown in Figure 1.

The framework of Figure 1 consists of two parts, namely, using reinforcement learning methods to make ship navigation decisions and using DL algorithms to perceive the environment for ship navigation. In the perception stage, neural networks process environmental information obtained from ship sensors, extract key features, and fit policy and value functions in reinforcement learning. In the decision-making stage, the main task is to collect data through the interaction between the ship and the environment, learn strategies based on the accumulated experience during exploration, and finally find the converged optimal strategy. The navigation and planning of ship navigation paths belong to long-term continuous decision-making problems [19]. To avoid the impact caused by delayed rewards, a dense reward function as shown in equation (1) is set.

$$r_t = \begin{cases} 100, d_g^t < d_a \\ 50, o_{\min} < 0.46 \\ (d_g^t - d_g^{t-1}) \times 100 + t_p, \text{otherwise} \end{cases} \quad (1)$$

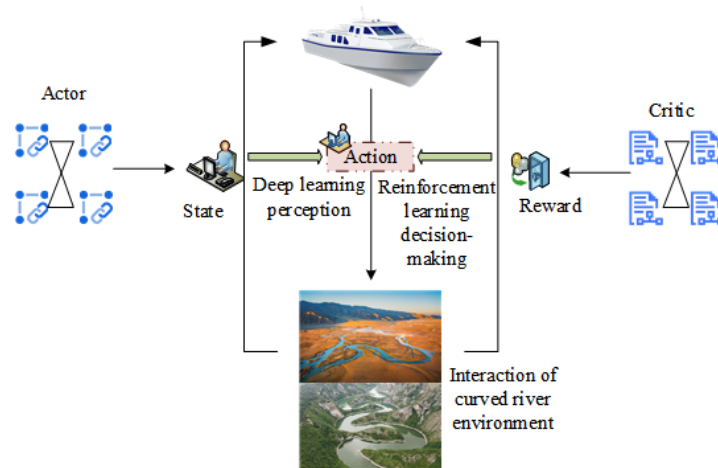


Figure 1: DL path planning framework.

In equation (1),  $o_{\min}$  is the threshold distance for obstacles.  $r_t$  is the value of the intensive reward function.  $d_a$  is the length of the path to the vessel to the target threshold.  $t_p$  is a time penalty item to encourage ships to complete tasks in a short period of time.  $d_g^{t-1}$  is the distance between the ship and the target point in the previous time.  $d_g^t$  is the actual distance between the ship and the target point. Through the continuous iterative interaction between ships and the environment, plenty of sample data will be generated. This study proposes the

concept of importance sampling to improve the utilization of interactive data in PPO algorithm. The expression is shown in equation (2).

$$E_{x \sim p}[f(x)] = E_{x \sim q}\left[f(x) \frac{p(x)}{q(x)}\right] \quad (2)$$

In equation (2),  $\frac{p(x)}{q(x)}$  represents the importance

weight.  $E_{x \sim p}[f(x)]$  is the expected value of random variables  $x$  and  $p$  under probability distribution  $f(x)$ . If a reasonable  $f(x)$  is chosen so that  $\frac{p(x)}{q(x)}$  is not too large or too small, it can significantly improve the efficiency and accuracy of the data. The formula for applying importance sampling in PPO is shown in equation (3).

$$\nabla \bar{R}_\theta = E_{\tau \sim \pi_\theta} \left[ \frac{p_\theta(a_t / S_t)}{p_\theta(a_t / S_t)} A^\theta(S_t, a_t) \nabla \log p_\theta(a_t / S_t) \right] \quad (3)$$

In equation (3),  $S_t$  is the environment state at the time of  $t$ .  $S_t^n$  is the environment state at the time of  $n$ .  $\nabla \bar{R}_\theta$  is the importance sampling value.  $a_t$  is the action output.  $a_t^n$  is the output value of the action output at the time of  $n$ .  $E_{\tau \sim \pi_\theta}$  is the expected value of the trajectory  $\tau$ .  $p_\theta$  is the value of the instantaneous reward function.  $p_{\theta'}$  is the value of the rewards generated by the next state.  $A$  is the vessel's sailing space.  $\theta$  is the strategy parameter.  $A^\theta(S_t, a_t)$  is the dominance function.  $\pi$  is the strategy value.  $\pi_\theta$  is the value of the training strategy. Due to the uncertainty of the strategy, the reward value has randomness in the interaction process, that is, the state value function is taken to evaluate the value of the current state, as shown in equation (4).

$$\pi(a/s) = P(A_t = a / S_t = s) \quad (4)$$

In equation (4),  $\pi(a/s)$  is the value of state value function.  $P$  is the probability of ship trajectory.  $a$  is the ship action.  $s$  is the current state value. To prevent the large difference between  $\pi_\theta$  and  $\pi$  from causing excessive variance between  $f(x)$  and  $f(x) \frac{p(x)}{q(x)}$ , this

study introduces a constrained trust region policy optimization algorithm based on equations (3) and (4) to constrain them, as shown in equation (5).

$$J_{PPO}^{\theta^k}(\theta) = J^{\theta^k} - \beta KL(\theta, \theta^k) \quad (5)$$

In equation (5),  $J_{PPO}^{\theta^k}(\theta)$  is the Proximal policy optimization penalty function (PPO-penalty).  $\beta$  is the penalty coefficient of the Kullback-Leibler (KL) scatter.  $KL(\theta, \theta^k)$  is the value of the KL scatter, which is used to measure the difference between the old and the new policies.  $J^{\theta^k}$  is the value of the old policy the value of the expectation function generated. Proximal policy optimization clip function (PPO-clip) is calculated as equation (6).

$$J^{\theta^k}(\theta) = E_{(S_t, a_t) \sim \pi_{\theta^k}} \left[ \frac{P_\theta(a_t / S_t)}{P_{\theta^k}(a_t / S_t)} A^{\theta^k}(S_t, a_t) \right] \quad (6)$$

In equation (6),  $J^{\theta^k}(\theta)$  represents the PPO loss function value obtained through pruning method.  $P_\theta(a_t / S_t)$  is the probability of the current strategy  $\pi$  selecting action  $a_t$  in state  $S_t$  under parameter  $\theta$ .  $P_{\theta^k}(a_t / S_t)$  is the policy from the last policy update. Equations (5) and (6) limit the differences between new and old strategies by introducing KL divergence, avoiding drastic changes in the strategy during a single update and enhancing the stability of the training process. The expression for  $\beta$  is shown in equation (7).

$$\beta = \begin{cases} \frac{\beta}{2}, & KL(\theta, \theta^k) > KL_{\min} \\ 2\beta, & KL(\theta, \theta^k) > KL_{\max} \end{cases} \quad (7)$$

In equation (7),  $\beta(\theta)$  is the new penalty coefficient.  $KL(\theta, \theta^k)$  denotes the KL scatter between the old and new strategies.  $KL_{\min}$  is the minimum target value of the KL scatter.  $KL_{\max}$  is the maximum target value of the KL scatter. A larger value of  $\beta(\theta)$  increases the penalty for KL scatter and limits the update magnitude of the policy [20]. While a smaller value of  $\beta(\theta)$  reduces this limitation and allows the policy to be updated more substantially. The structure of the studied PPO-CNN neural network is shown in Figure 2.

In Figure 2, when the river channel image

information is input into the CNN, convolution calculation is performed in the image area through the convolution kernel. Convolutional kernels multiply corresponding elements at each position in the image and accumulate them to generate new feature map elements, ultimately generating a feature map of curved rivers. The feature map is input into the PPO for training, and optimal path planning is achieved through exploration

and optimization. The ship path planning model based on PPO-CNN can effectively combine image feature extraction and reinforcement learning algorithms to generate efficient and safe ship path planning schemes through in-depth analysis and learning of curved river channel images.

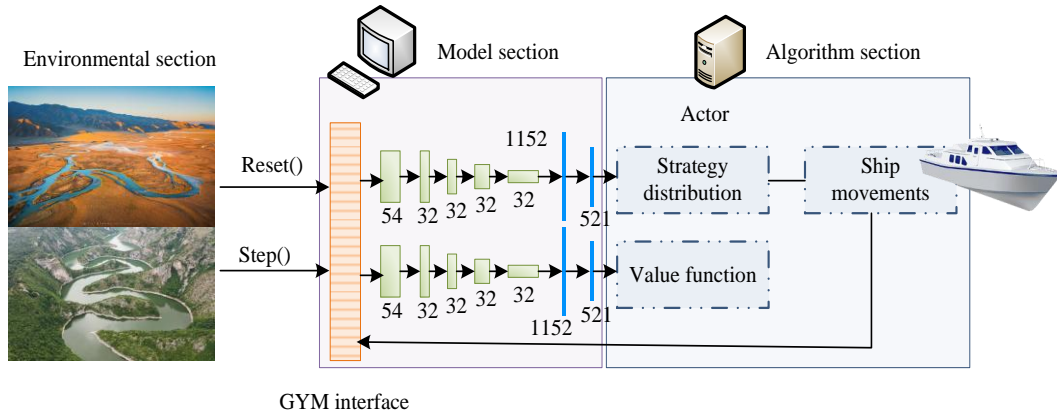


Figure 2: PPO-CNN neural network structure.

### 2.2 Construction of SNC model for curved river section combining hybrid controller and PPO-CNN

This study uses the reinforcement learning algorithm PPO to construct path planning and navigation tasks for ships navigating in curved river sections. Using the ship's camera to obtain image data from the environment as state input, a CNN-based model is built to process the input and achieve the mapping of visual images to ship navigation speed [21]. When a ship is in a river with poor weather conditions and complex curved sections, due to the reduced navigation field of view, the ship often cannot accurately find the yaw angle to avoid collisions in curved sections. To enable ships to adapt to unknown long-sequence complex environments and complete planned navigation tasks [22], it is necessary to improve the algorithm's ability to explore and generalize. This study proposes an algorithm that combines a Hybrid Controller (HC) and PPO-CNN, HC-PPO-CNN. It integrates target position, self-state, and LiDAR information into the state input through multiple HC, improving the perception ability of ships in curved river sections. Figure 3 shows the HC reinforcement learning framework of the algorithm.

This framework consists of two controllers. The one is the reinforcement learning algorithm controller  $\pi_\theta$ , and the other is the prior controller  $\pi_{prior}$ . To integrate the two controllers, it is necessary to label the output of

$\pi_\theta$  as a Gaussian distribution of the same dimension as the policy output. The calculation of the fusion method is given by equation (8).

$$\pi'(a/s) = \frac{1}{Z} (\pi_\theta(a/S)^{1-\alpha} \cdot \pi_{prior}(a/S)^\alpha) \quad (8)$$

In equation (8),  $\alpha$  is the gating function.  $\pi'(a/s)$  is the value of the fused output distribution.  $Z$  is the weighting coefficient.  $\pi_{prior}(a/S)^\alpha$  is the output distribution of  $\pi_{prior}$ .  $\pi_\theta(a/S)$  is the output distribution of  $\pi_\theta$ . The expression for the fusion distribution is  $\pi'(a/s) \sim N(\mu', \sigma'^2)$ . The expression of  $\mu'$  is given by equation (9).

$$\mu' = \frac{\mu_\theta \sigma_{prior}^2 (1-\alpha) + \mu_{prior} \sigma_\theta^2 \alpha}{\sigma_{prior}^2 (1-\alpha) + \sigma_\theta^2 \alpha} \quad (9)$$

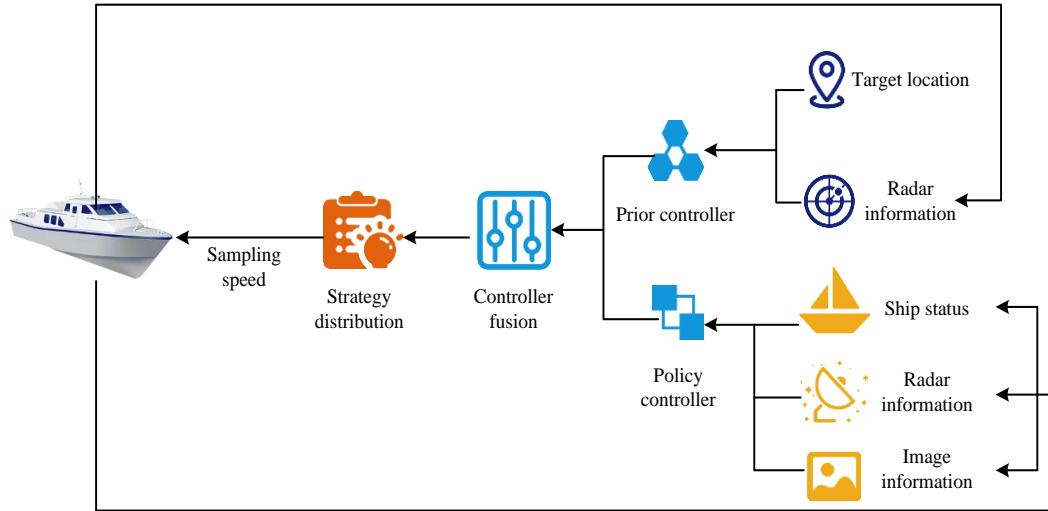


Figure 3: HC reinforcement learning framework.

In equation (9),  $\mu_{prior}$  is the expected value of the output distribution of  $\pi_{prior}$ .  $\alpha$  is the gating function.  $\sigma_{prior}^2$  is the output variance value of  $\pi_{prior}$ .  $\mu'$  is the variance of the fused distribution.  $\sigma_{prior}^2$  is the expected value of the fused control output.  $\sigma_{\theta}^2$  is the output variance of  $\pi_{\theta}$ .  $\mu_{\theta}$  is the expected output value  $\mu'$  of  $\pi_{\theta}$ , which comprehensively considers the output information of the prior controller and PPO-CNN model. The calculation of  $\sigma^{2'}$  is given by equation (10).

$$\sigma^{2'} = \frac{\sigma_{\theta}^2 \sigma_{prior}^2}{\sigma_{prior}^2 (1 - \alpha) + \sigma_{\theta}^2 \alpha} \quad (10)$$

In equation (10),  $\sigma^{2'}$  is the output variance of the PPO-CNN model, which indicates the degree of uncertainty of the integrated control output.  $\sigma_{\theta}^2$  is the variance of the output distribution of the strategy controller. The value of the gating function,  $\alpha$ , is employed to attain a more flexible control strategy. This enables the dynamic adjustment of the weight of the a priori controller and the PPO-CNN model in the fusion

process. The value of  $\alpha$  is calculated using the following equation (11).

$$\alpha = \begin{cases} 1, & t = 0 \\ 1 - 0.001t, & t \neq 0 \end{cases} \quad (11)$$

In equation (11),  $t$  is the number of times the ship avoids obstacles in curved river sections during simulation experiments. By calculation, the fused control output can comprehensively consider the advantages of the fused prior controller and PPO-CNN model, thereby improving the performance and stability of SNC. To enhance the prior guidance for  $\pi_{\theta}$ , the control output is modified to achieve a state of complementary advantages. The study adopts the APFM as the prior controller for  $\pi_{\theta}$ . APFM is a commonly used algorithm in path planning, which relies on the combined force of attractive and repulsive potential fields to guide ships from the starting point to the target position, while avoiding obstacles. Figure 4 is a schematic diagram of APFM for ship navigation.

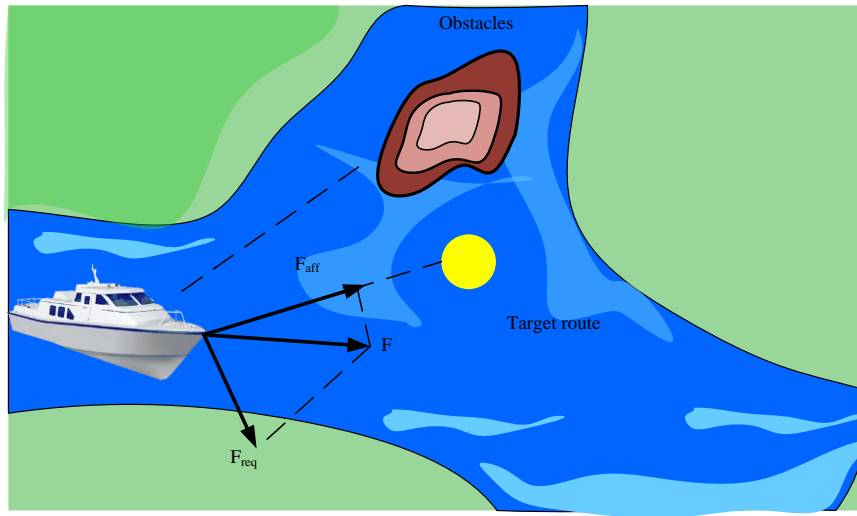


Figure 4: Schematic diagram of AFPM for ships.

In APFM, the attractive potential field function is used to generate attractive force [25], causing the ship to move towards the target position. The commonly used equation for the attractive potential field function is given by equation (12).

$$U_{att}(q) = \frac{1}{2} \eta \rho^2(q, q_g) \quad (12)$$

In equation (12),  $q$  is the current position of the vessel.  $\rho^2(q, q_g)$  is the distance between the ship and the obstacle avoidance target point in the curved river section.  $q_g$  is the location of the obstacle avoidance target point in the curved river section. The direction of  $\rho^2(q, q_g)$  is directed towards the target point by the line connecting  $q$  and  $q_g$ .  $\eta$  is the value of the stress gain coefficient. The magnitude of the gravitational force generated by the attractive potential field is given by equation (13).

$$F_{att} = \eta \rho(q, q_g) \quad (13)$$

In equation (13),  $\eta$  is the gravitational gain coefficient.  $F_{att}$  is the gravitational force that attracts the potential field.  $\rho(q, q_0)$  is the path length between the vessel position and the bend obstacle position. The magnitude of gravitational and repulsive forces is inversely proportional to the distance between the ship and the obstacle avoidance point in the curved river section. The potential energy of a ship is directly proportional to the cube of the distance between the ship and the center of gravity. A decrease in distance results in a corresponding decrease in potential energy, and 0 indicates that the ship has reached the target point. The formula for repulsive potential field is given by equation (14).

$$U_{rep}(q) = \begin{cases} \frac{1}{2} k \left( \frac{1}{\rho(q, q_0)} - \frac{1}{\rho_0} \right), & 0 \leq \rho(q, q_0) \leq \rho_0 \\ 0, & \rho(q, q_0) > \rho_0 \end{cases} \quad (14)$$

In equation (14),  $\rho_0$  is the threshold distance.  $k$  is the repulsive gain coefficient.  $F_{att}$  is the gravitational force that attracts the potential field. The magnitude of gravity is mainly related to the distance between ships and obstacle avoidance targets in curved river sections. The larger the distance, the higher the potential energy value that the ship experiences. From this, the global potential field function in the curved river environment is obtained, as shown in equation (15).

$$U(q) = U_{rep}(q) + U_{att}(q) \quad (15)$$

In equation (15),  $U(q)$  is the resultant force of the attractive potential field and the repulsive potential field. The introduction of a combined attractive and repulsive potential field allows ships to navigate around obstacles and reach their intended destination. This method is intuitive and straightforward to calculate, making it well-suited for real-time path planning. By introducing the fusion of prior controller outputs and  $\alpha$ -gate functions, dynamic weight adjustment has been achieved, further improving the response speed and adaptability of the ship navigation system. In addition, combined with APFM, it can effectively reduce the collision problem of obstacles in curved river sections during navigation, ensuring navigation safety.



### 3 Results

#### 3.1 Performance verification of PPO-CNN model in curved river sections

This study selects a certain type of ocean freight ship as the experimental object to verify the performance of the PPO-CNN model in curved river sections. The experimental environment is set as follows: The server model is XYZ-1234, the CPU is Intel Xeon E5-2670 v3, the memory is 16GB RAM, the operating system is Windows 10, the GPU is NVIDIA Tesla V100, the power supply is 800W-ATX. A simulation map is constructed using a section of the Yangtze River and a curved section of the Cangzhou Canal, as shown in Figure 5.

Figures 6 (a) and (b) show the average reward

and path length obtained by PPO-CNN during the training process of simulating curved river sections. In Figure 6 (a), as the amount of training iterations grows, the average reward number becomes positive at 75 rounds of algorithm iteration. This indicates that the model is beginning to obtain useful information and gradually optimizing its navigation strategy in the channel. When the algorithm is iterated to 100 times, the average reward is in the range of 0.02~0.05, indicating that the model can effectively handle the task of simulating curved river sections with fewer errors. In Figure 6 (b), when the iterations reach 80, the path length rapidly decreases. At iteration 200, the path length tends to flatten and fluctuates around 83 steps, indicating that the PPO-CNN algorithm is in a convergence state.

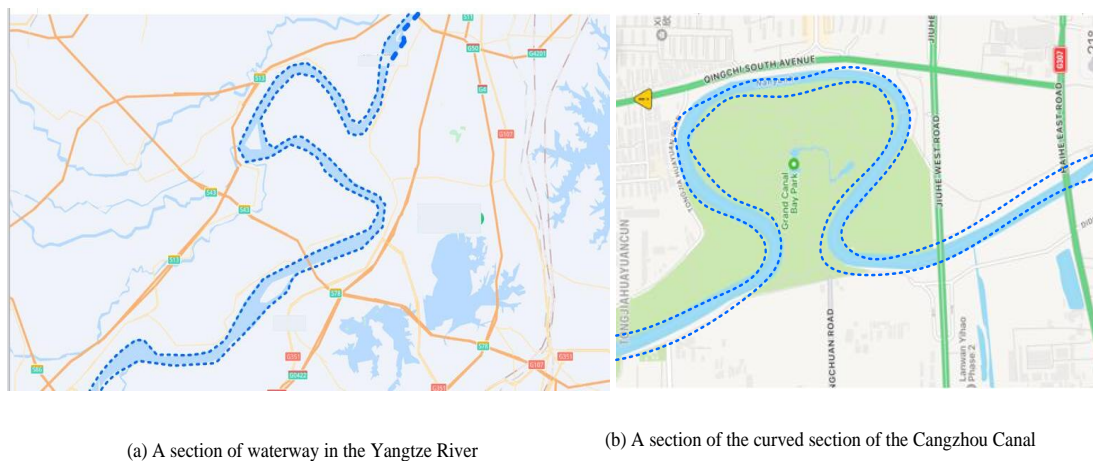


Figure 5: Simulated navigation map based on the Yangtze River and Cangzhou canal.

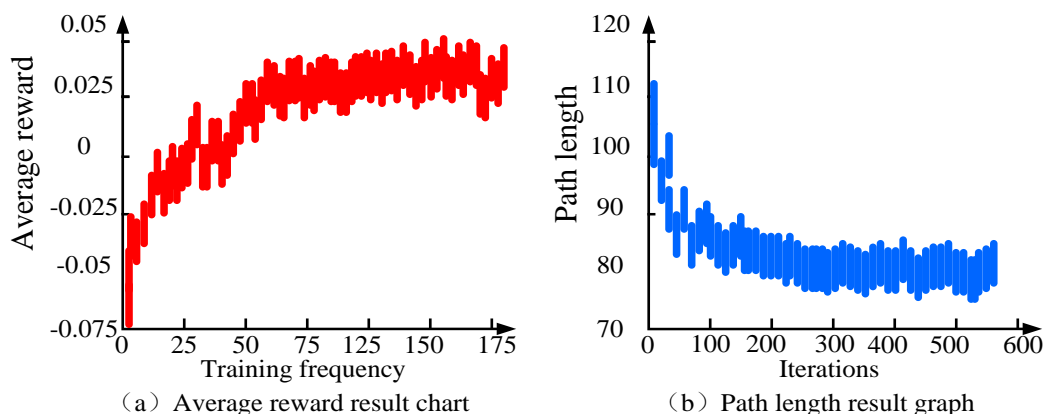


Figure 6: The average reward and path length results of PPO-CNN in simulating curved river sections.

Figure 7 shows the results of path planning for two algorithms in different winding river sections. In Figure 7 (a), the PPO-CNN algorithm plans a path for the winding

river section of the Yangtze River. The path in the figure shows that the route planned by the model in the winding river section is coherent and smooth, and the planned

navigation path is 660.2km. Figure 7(b) illustrates the path planning results obtained using the traditional Dijkstra algorithm for the winding river section of the Yangtze River. The path in the figure shows that the model's planned path in the winding river section is poor. The planned navigation path is 678.3km, and by calculating the length of the path, it is 2.73% longer than the path planned by the PPO-CNN algorithm. Figure 7 (c) shows the path planning result of the PPO-CNN algorithm in the winding section of the Cangzhou Canal. The PPO-CNN algorithm has a better route planning for the winding section of the river, and the navigable path after path planning is 397.42 km. Figure 7 (d) shows the path planning result of the Dijkstra algorithm for the winding section of the Cangzhou Canal. The length of the planned navigation path is 436.5 km. The length of the path planned by the PPO-CNN algorithm is 8.95% shorter than that planned by the Dijkstra algorithm. The results show that the PPO-CNN model is significantly

better than the Dijkstra algorithm for path planning in winding river sections.

Figure 8 (a) shows a comparison of the average rewards obtained from training models using traditional PPO, Dijkstra, and PPO-CNN in simulating fast-flowing curved river sections. PPO-CNN has relatively small fluctuations in the early stages, with positive values when the number of training rounds reaches 75-200, and an average reward value of 0.0323. Compared to other algorithms, the average reward value of PPO-CNN is 19.36% higher, indicating that the improved PPO-CNN can quickly find the target point of yaw in curved river sections. Figure 8 (b) shows a comparison of the path lengths obtained by three algorithms in training models for simulating turbulent water flow in curved river sections. When PPO-CNN trains for 150-240 rounds, the path length tends to be more stable

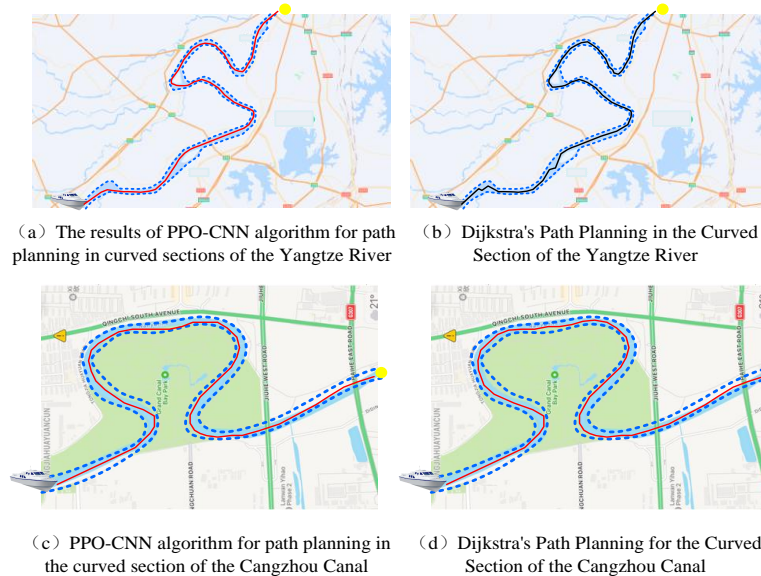


Figure 7: Results of the path length of the PPO-CNN simulation for the curved river section.

as a straight line, with an average value of 139.30. Other algorithms still exhibit significant data fluctuations, indicating that the PPO-CNN-trained model has better robustness and stronger generalization ability.

### 3.2 Performance verification of SNC model for curved river sections using HC-PPO-CNN

To investigate whether the HC-PPO-CNN algorithm can complete the planned navigation task in unknown long sequence complex environments, obstacles are set in the training scenario in the previous section to create a relatively narrow navigation area. By limiting the

viewing angle range of ships and reducing the viewing angle range by 45 degrees, the environment of poor river weather conditions is simulated. Figure 9 shows the performance comparison of three path planning algorithms based on long sequence complex environments.

Figure 9 (a) shows the average path planning of PPO, PPO-CNN, and HC-PPO-CNN algorithms on simulated maps. The average path of HC-PPO-CNN is 1.996, which is lower than the 2.426 of PPO and the 2.865 of PPO-CNN. HC-PPO-CNN performs the best on simulated maps with the shortest average path length. This indicates that the combination of HC and PPO-CNN significantly improves the effectiveness of path planning in complex curved river sections. Figure 9 (b) shows the total number of

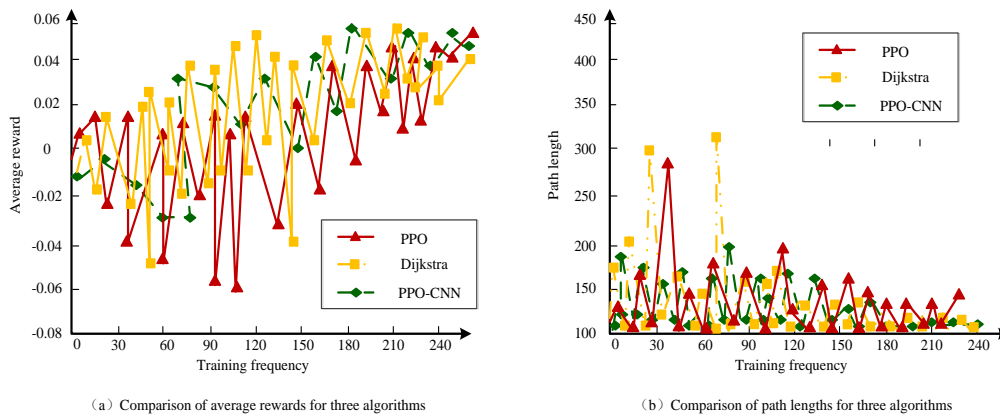


Figure 8: Comparison of average reward and path length of three algorithms in curved river sections under high flow velocity.

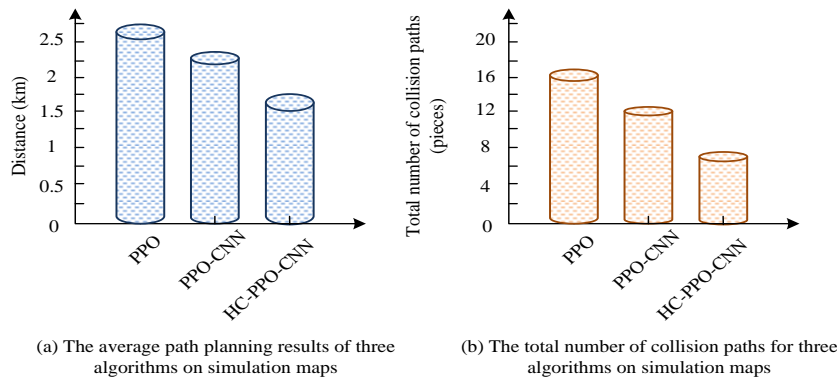


Figure 9: Performance comparison of three path planning algorithms based on long sequence complex environments.

collision paths for three algorithms. The total number of collision paths for HC-PPO-CNN is 7, which is lower than the 11 paths for PPO and the 16 paths for PPO-CNN. This indicates that HC-PPO-CNN exhibits strong adaptability in complex curved river environments, and can effectively cope with diverse curved river sections and adverse weather conditions, reducing collisions.

Figure 10 (a) shows a comparison of the average reward values of PPO, PPO-CNN, and HC-PPO-CNN algorithms under limited field of view. HC-PPO-CNN gradually converges at around 5,000 steps, with an

average reward value between 300-500 between 10,000 and 20,000 steps, which is higher than the other two algorithms. Figure 10 (b) shows a comparison of three algorithms in a long sequence environment. When HC-PPO-CNN training reaches 2,000 steps of falling, it tends to flatten out, with an average value of 448.3. PPO and PPO-CNN both tend to flatten out at 3,500 steps. This indicates that the HC-PPO-CNN algorithm exhibits good convergence and stability in long-sequence environments.

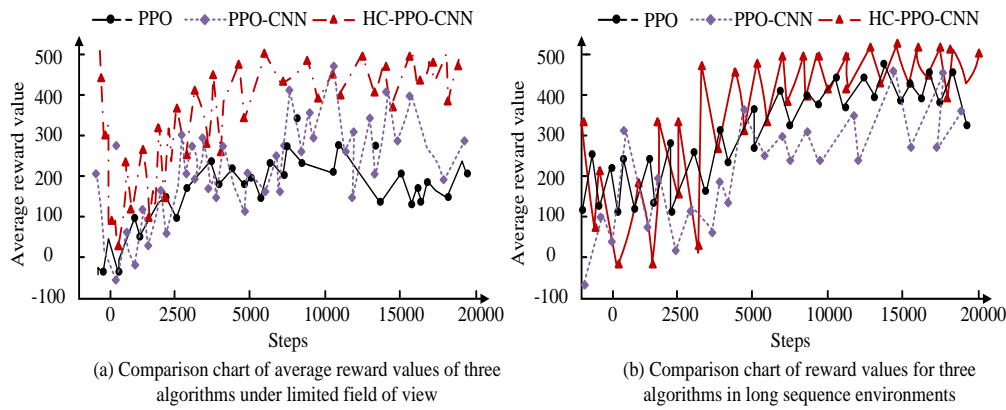


Figure 10: Comparison chart of average reward values of algorithms under limited field if view and reward values in long sequence environments.

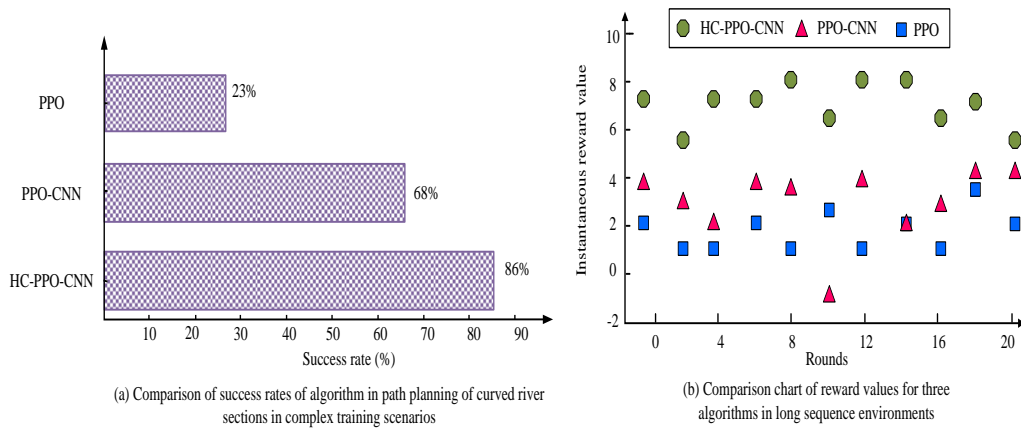


Figure 11: The success rate and IR map of three algorithms for path planning in curved river sections in complex training scenarios.

Table 2: Experimental results of HC-PPO-CNN model ablation.

Model version	Accuracy (%)	LOSS	F1 score	Recall(%)	Mean absolute error (MAE)
Complete model	92.5	0.23	0.91	93.2	0.15
L-HC	89.3	0.31	0.88	90.1	0.21
L-CNN	81.7	0.45	0.79	82.4	0.37
L-PPO	85.6	0.37	0.84	86.5	0.29

Figure 11 (a) shows a comparison of the success rates of three algorithms for curved river path planning in complex training scenarios. The success rate of HC-PPO-CNN is 82%, while PPO (63%) and PPO-CNN (28%) differ from HC-PPO-CNN by 19%

and 54%, respectively. This indicates that HC-PPO-CNN has higher processing capability and path planning effectiveness in complex environments than the other two algorithms. Figure 11 (b) shows the Instantaneous Reward (IR) maps obtained by each algorithm for each action taken during path planning for curved river sections. Compared to PPO and PPO-CNN,

HC-PPO-CNN achieves higher IR scores in each path planning process. The mean IR score of HC-PPO-CNN is 7.95, which is 3.69 and 1.58 higher than others. This proves that the research algorithm can make better choices at every step of the planning process, demonstrating its immediate decision-making advantage in handling complex path planning tasks. To verify the impact of different modules of the HC-PPO-CNN model, ablation experiments are conducted on the model to remove the HC module (L-HC), the CNN module (L-CNN), and the PPO module (L-PPO), respectively. The experimental results of the accuracy are shown in Table 2.

Table 2 shows the results of removing different modules on the HC-PPO-CNN model. Among the results of accuracy, loss value, F1-score loss value, recall rate, and MAE, the model data with the CNN module removed is poorer, with values of only 81.7%, 0.45, 0.79, 82.4%, and 0.37. This indicates that after the CNN module is removed, the model loses its ability to extract features from images of curved river sections. Based on this, the input image cannot be effectively processed, and the classification ability of positive and negative samples is reduced. The results verify that the CNN is an indispensable part of the model. The L-PPO model's accuracy, F1 score, and recall rate are 85.6%, 0.84, and 86.5%, respectively, with performance only lower than the complete model and the L-HC model. This indicates that PPO plays an important role in proximal strategy optimization, but the impact on model performance after removal is smaller than that of L-CNN. The accuracy, loss value, F1 score, recall rate, and MAE values of L-HC model are 89.3%, 0.31, 0.88, 90.1%, and 0.21, respectively. Its experimental values are only lower than those of the complete model. This indicates that after the HC module is removed, there is a small impact on the model's feature extraction ability and sample recognition, but its performance does not differ greatly after the CNN module is removed. The experimental results show that the CNN module plays a more important role in the model and is the core module for processing river images and assisting navigation planning.

## 4 Discussion

The proposed HC-PPO-CNN model has unique advantages in the planning of ship navigation paths in winding river sections. In the experiment, the PPO-CNN algorithm reduced the planned path in the winding section of the Yangtze River by 2.73% compared to the path planned by the Dijkstra algorithm. The PPO-CNN algorithm reduced the planned path in the winding section of the Cangzhou Canal by 8.95% compared to the Dijkstra algorithm. HC-PPO-CNN had a success rate of 82% in path planning for complex winding sections and could effectively reduce collisions in winding sections.

The experimental results of the path planning length show that the research method has a significant advantage in reducing the length of the ship's navigation in the curved section. The advantage of the model lies in its combination of HC and PPO-CNN. HC can express superior adaptability in dynamic environments, particularly in curved sections. In contrast, static algorithms such as Dijkstra are often unable to respond to rapidly changing environments in curved sections promptly. In addition, PPO does not perform well when dealing with the complex dynamic environment of the winding river. The research introduces a CNN to extract and process the characteristics of the winding river section, which is then input into the PPO algorithm for a strategy update. This effectively improves the navigation accuracy of the ship in the winding river section and reduces collisions. Due to the nonlinear dynamic characteristics and sudden environmental changes of the winding river section, algorithms such as Dijkstra and PPO are more dependent on route planning in static or partially dynamic environments, making it difficult to respond and handle the winding river section scenario promptly. The HC-PPO-CNN model can respond to changes in water flow and curvature in winding river sections promptly through dynamic adjustments of different modules. HC controls real-time adjustments, while PPO-CNN ensures the robustness of path planning decisions during navigation. These operations can effectively reduce errors in path planning, so that HC-PPO-CNN can achieve a path planning success rate of 82% in complex winding river sections. The research model has reduced the length of the planned path by 15% compared to the traditional method in the planning of curved navigation sections, and the collision rate has been reduced by more than 20%. It has effectively improved navigation performance and reduced collisions.

Compared with the model in reference [13], the research design uses PPO-CNN for ship navigation, which improves the adaptability of path planning and global exploration capabilities. This is because PPO can avoid the problem of the model getting stuck in a local optimum and can exhibit better performance and stability in complex dynamic environments. Compared with the model of reference [15], the model designed in this study achieves higher path planning accuracy and stability in curved river navigation through PPO-CNN. This is because PPO can perform effective strategy optimization updates and CNN has a high perception of environmental features. Compared with path planning based only on APFM, the model designed in this study has higher adaptability and robustness in complex environments.

In summary, the research optimizes the PPO-CNN-based model by introducing an HC module, which effectively improves the success and efficiency of the model in navigation path planning in curved river

sections. The research contributes to the development of intelligent shipping.

## 5 Conclusion

The experimental analysis verified the effectiveness of the HC-PPO-CNN model in navigating curved river sections. The average reward value and path length data evaluation of the model for path planning in simulated curved river sections showed that the trained model had better robustness and stronger generalization ability. This model reduced the number of paths for ships to travel in curved river sections. This study also combined HC to plan curved river sections for ships in long sequence complex environments. This enabled ships to optimize their navigation paths based on actual conditions even in difficult weather conditions and complex curved river sections. As a result, the reliability and efficiency of autonomous navigation for ships in complex environments were improved.

To make accurate path planning for SNC in curved river sections and ensure the safe operation of ships in curved river sections, this study proposed the HC-PPO-CNN. To verify its effectiveness, this study conducted relevant experiments. In a long sequence environment, the average reward value of HC-PPO-CNN stabilized at 448.36 after 2000 steps, while PPO and PPO-CNN tended to flatten out after 3500 steps, respectively. In complex training scenarios, the success rate of HC-PPO-CNN for curved river path planning was 82%, significantly higher than PPO's 63% and PPO-CNN's 28%. The average IR value of HC-PPO-CNN in each path planning was 7.95, which was 3.69 higher than PPO and 1.58 higher than PPO-CNN. The average path length of HC-PPO-CNN on the simulation map was 1.996, and the number of collision paths was 7, both lower than PPO and PPO-CNN algorithms. In summary, HC-PPO-CNN has significant effectiveness and robustness in SNC of complex curved river sections. The limitation of this study is that the actual application effect of the model depends on a lot of input data and high-performance computing resources. This may limit its application in projects with limited resources. In the future, lightweight versions of the model will be developed to reduce the dependence on computer resources.

## 6 Funding

The research is supported by Jiangsu Province University “Qinglan Project” Excellent Young Backbone Teacher Training Funding Project; Jiangsu Province Vocational Education “Dual-Qualified Teacher” Famous Teacher Studio Training Project (2022-09).

## References

- [1] Yunping Yang, Ming Li, Wanli Liu, Yuanfang Chai, Jie Zhang, and Wenjun Yu. Relationship between potential waterway depth improvement and evolution of the Xigjiang Reach of the Yangtze River in China. *Journal of Geographical Sciences*, 33(3):547-575, 2023. <https://doi.org/10.1007/s11442-023-2096-8>
- [2] Yunping Yang, Hong Yin, Ming Li, Wanli Liu, Kanyu Li, and Wenjun Yu. Effect of water depth and waterway obstructions on the divergence and confluence areas of Dongting Lake and the Yangtze River after the operation of the Three Gorges Project. *River*, 2(1):88-108, 2023. <http://dx.doi.org/10.1002/rvr2.31>
- [3] Yong Hu, Jinyun Deng, Yitian Li, Congcong Liu, and Zican He. Flow resistance adjustments of channel and bars in the middle reaches of the Yangtze River in response to the operation of the Three Gorges Dam. *Journal of Geographical Sciences*, 32(10):2013-2035, 2022. <https://doi.org/10.1007/s11442-022-2034-1>
- [4] Inderveer Solanki. Manoeuvrability of vessels in inland waterways and safety of navigation. *Maritime Affairs: Journal of the National Maritime Foundation of India*, 17(2):107-121, 2021. <https://doi.org/10.1080/09733159.2022.2026496>
- [5] Rushit Dave, and Joy Purohit. Leveraging deep learning techniques to obtain efficacious segmentation results. *Archives of Advanced Engineering Science*, 1(1):11-26, 2023. <https://doi.org/10.47852/bonviewAAES32021220>
- [6] Simegnew Yihunie Alaba, and John E. Ball. Deep learning-based image 3-D object detection for autonomous driving: review. in *IEEE Sensors Journal*, 23(4):3378-3394, 2023. <https://doi.org/10.1109/JSEN.2023.3235830>
- [7] Alok Kumar, and Vijesh Kumar Patel. Classification and identification of disease in potato leaf using hierarchical based deep learning convolutional neural network. *Multimedia Tools and Applications*, 82(20):31101-31127, 2023. <https://doi.org/10.1007/s11042-023-14663-z>
- [8] Caglar Uyulan, Turker Tekin Ergüzel, Huseyin Unubol, Merve Cebi, Gorben Hizli Sayar, Mahdi Nezhad Asad, and K. Nevzat Tarhan. Major depressive disorder classification based on different convolutional neural network models: Deep learning approach. *Clinical EEG and Neuroscience*, 52(1):38-51, 2021. <https://doi.org/10.1177/1550059420916634>
- [9] Akane Minagawa, Hiroshi Koga, Tasuku Sano, Kazuhisa Matsunaga, Yoshihiro Teshima, Akira Hamada, Yoshiharu Houjou, Ryuhei Okuyama.

- Dermoscopic diagnostic performance of Japanese dermatologists for skin tumors differs by patient origin: A deep learning convolutional neural network closes the gap. *The Journal of Dermatology*, 48(2):232-236, 2021. <https://doi.org/10.1111/1346-8138.15640>
- [10] Khosro Rezaee, Mohammad R. Khosravi, Hani Attar, Varun G. Menon, Mohammad Ayoub, Haitham Issa, and Lianyong Qi. IoMT-assisted medical vehicle routing based on UAV-Borne human crowd sensing and deep learning in smart cities. *IEEE Internet of Things Journal*, 10(21):18529-18536, 2023. <https://doi.org/10.1109/JIOT.2023.3284056>
- [11] Chuancong Tang, Haitao Zhang, and Jun Wang. Flexible formation tracking control of multiple unmanned surface vessels for navigating through narrow channels with unknown curvatures. *IEEE Transactions on Industrial Electronics*, 70(3):2927-2938, 2022. <https://doi.org/10.1109/TIE.2022.3145641>
- [12] Zheng Zeng, Chengke Xiong, Xinyi Yuan, Hexiong Zhou, Yuing Bai, Yufei Jin, Di Lu, and Lian Lian. Information-driven path planning for hybrid aerial underwater vehicles. in *IEEE Journal of Oceanic Engineering*, 48(3):689-715, 2023. <https://doi.org/10.1109/JOE.2022.3212356>
- [13] Jie Zhang, Zhengxin Wang, Guangjie Han, Yujie Qian, and Zhenglin Li. A collaborative path planning method for heterogeneous autonomous marine vehicles. in *IEEE Internet of Things Journal*, 11(1):1465-1480, 2024. <https://doi.org/10.1109/JIOT.2023.3224967>
- [14] Hongchu Yu, Alan T. Murray, Zhixiang Fang, Jingxian Liu, Guojun Peng, Mohammad Solgi, and Weilong Zhang. "Ship path optimization that accounts for geographical traffic characteristics to increase maritime port safety," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 6, pp. 5765-5776, 2022. <https://doi.org/10.1109/TITS.2021.3057907>
- [15] Jia Wang, Yang Xiao, Tieshi Li, and C. L. Philip Chen. A jamming aware artificial potential field method to counter GPS jamming for unmanned surface ship path planning. *IEEE Systems Journal*, 17(3):4555-4566, 2023. <https://doi.org/10.1109/JSYST.2023.3237613>
- [16] Inderveer Solanki. Manoeuvrability of vessels in inland waterways and safety of navigation. *Maritime Affairs: Journal of the National Maritime Foundation of India*, 17(2):107-121, 2021. <https://doi.org/10.1080/09733159.2022.2026496>
- [17] Kai Zhang, Min Hu, Fuji Ren, Yanwei Bao, Piao Shi, and Daoyang Yu. River boundary detection and autonomous cruise for unmanned surface vehicles. *IET Image Processing*, 17(11):3196-3215, 2023. <https://doi.org/10.1049/ipr2.12848>
- [18] Yang Gu, YuhuCheng, C. L. Philip Chen, and Xuesong Wang. Proximal policy optimization with policy feedback. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 52(7):4600-4610, 2021. <https://doi.org/10.1109/TSMC.2021.3098451>
- [19] Azzeddine Bakdi, and Erik Vanem. Fullest COLREGs evaluation using fuzzy logic for collaborative decision-making analysis of autonomous ships in complex situations. *IEEE Transactions on Intelligent Transportation Systems*, 23(10): 18433-18445, 2022. <https://doi.org/10.1109/TITS.2022.3151826>
- [20] Zhijie Xie, and Shenghui Song. FedKL: Tackling data heterogeneity in federated reinforcement learning by penalizing KL divergence. *IEEE Journal on Selected Areas in Communications*, 41(4):1227-1242, 2023. <https://doi.org/10.1109/JSAC.2023.3242734>
- [21] Xiaolong Chen, Xiaoqian Mu, Guan Jian, Ningbo Liu, and Wei Zhou. Marine target detection based on Marine-Faster R-CNN for navigation radar plane position indicator images. *Frontiers of Information Technology & Electronic Engineering*, 23(4):630-643, 2022. <https://doi.org/10.1631/FITEE.2000611>
- [22] Sriranga Narasimha Gandhi Aryavalli, and G. Hemantha Kumar. Futuristic vigilance: empowering chipko movement with cyber-savvy IoT to safeguard forests. *Archives of Advanced Engineering Science*, 1(8):1-16, 2023. <https://doi.org/10.47852/bonviewAAES32021480>
- [23] Wei Li, and Rong Xiong. A hybrid visual servo control method for simultaneously controlling a nonholonomic mobile and a manipulator. *Frontiers of Information Technology & Electronic Engineering*, 22(2):141-154, 2021. <https://doi.org/10.1631/FITEE.1900460>
- [24] Lisha Dong. Improved a algorithm for intelligent navigation path planning. *informatica*, 2024(10):181-194, 2024. <https://doi.org/10.31449/inf.v48i10.5693>
- [25] Tran Van Phong, Binh Thai Pham, and Phan Trong Trinh, Hai-Bang Ly, Quoc Hung Vu, Lanh Si Ho, Hiep Van Le, Lai Hop Phong, Mohammadtaghi Avand, and Indra Prakash. Groundwater potential map using GIS-based hybrid artificial intelligence methods. *Groundwater*, 59(5):745-760, 2021. <https://doi.org/10.1111/gwat.13094>

