https://doi.org/10.31449/inf.v49i16.7839 Informatica **49** (2025) 151–170 **151**

Efficient Logistics Path Optimization and Scheduling Using Deep Reinforcement Learning and Convolutional Neural Networks

Yan Yang[1, 2, *], Kang Wang[1, 2]
[1]School of Economics and Management, Jiaozuo University, Jiaozuo 454000, Henan, China
[2]Graduate School, University of the East, Manila 0900, Philippines
E-mail: yangyan_edu@outlook.com
[*]Corresponding author

Keywords: CNN, DRL, logistics path optimization, real-time scheduling, robustness scoring

Received: December 17, 2024

With the rapid development of e-commerce and online shopping, the logistics industry is facing unprecedented challenges. Traditional logistics path - planning methods, such as SPA, HA, GA, etc., struggle to cope with the complex and ever-changing logistics environment. To address this issue, this study proposes an innovative model that combines Deep reinforcement learning (DRL) with a Convolutional neural network (CNN) to achieve efficient logistics path optimization. In this research, a detailed analysis and pre-processing of the public datasets, the City Logistics Dataset (CLDS) and the Traffic Status Dataset (TSDS), were carried out to construct a model capable of effectively handling diverse logistics environments. Six baseline methods, namely the classic shortest path algorithm (SPA), heuristic algorithm (HA), genetic algorithm (GA), rule-based method (RBM), traditional deep reinforcement learning method (TDRM), and the most advanced deep learning method (ADLM), were selected for comparison. The experimental results indicate that the proposed model performs excellently across various environments. For instance, in suburban areas, it achieves a path length of 180 kilometers, a completion time of 120 minutes, a punctuality rate of 92%, and a dispatch success rate of 95%. In urban settings, the path length is 200 kilometers, the completion time is 150 minutes, the punctuality rate is 90%, and the dispatch success rate is 93%. On highways, it reaches a path length of 170 kilometers, a completion time of 110 minutes, a punctuality rate of 93%, and a dispatch success rate of 95%. Compared with the baseline methods, the model shows significant improvements in key metrics such as path length, completion time, punctuality, and dispatch success rate. Additionally, it outperforms them in terms of computation time and robustness scores, demonstrating great potential for practical applications.

Povzetek: Opisan je izvrni model za optimizacijo logističnih poti in sprotno razporejanje z združitvijo globokega utrjevalnega učenja (DRL) in konvolucijskih nevronskih mrež (CNN).

1 Introduction

With the advancement of global economic integration and the rapid development of e-commerce, the logistics industry is facing unprecedented challenges and opportunities. Efficient, fast and accurate delivery of goods has become one of the core elements of corporate competition. However, finding the optimal delivery path in a complex geographical environment and achieving instant scheduling under dynamically changing conditions has always been a difficult problem for logistics companies. Although traditional mathematical programming-based methods perform well under static conditions, they have obvious limitations in dealing with real-time changing traffic conditions and emergencies [1]. Therefore, it is particularly important to explore a new logistics path optimization and real-time scheduling solution that can adapt to complex environments and has self-learning capabilities [2].

Logistics path optimization is a core link in logistics management and is crucial to improving logistics service quality and reducing operating costs. Against the backdrop of the rapid development of artificial intelligence technology, domestic and foreign scholars are actively exploring the application of AI technology in logistics path optimization, aiming to improve logistics efficiency through intelligent algorithms. Although traditional methods such as linear programming can provide effective solutions, they are powerless in the face of large-scale dynamic problems [3, 4]. In contrast, AI technologies such as genetic algorithms (GA) and ant colony algorithms (ACA) have shown stronger exploration capabilities and adaptability, especially in solving the traveling salesman problem (TSP) [5]. In addition, the advancement of deep learning technology, especially the application of long short-term memory networks (LSTM), makes it possible to predict traffic conditions and realize dynamic path planning. Reinforcement learning (RL) enables intelligent agents to make optimal decisions in a constantly changing environment by simulating the learning process. These technologies have been widely used in multiple scenarios such as urban distribution, cross-border logistics, and cold chain logistics, helping to optimize delivery routes, predict customs clearance time, monitor temperature changes, etc. [6, 7]. Despite this, the application of AI in logistics path optimization still faces challenges in data privacy protection, algorithm real-time and robustness. With the advancement of technology and the evolution of social needs, more innovative solutions are expected to emerge in the future, continuously promoting the intelligent development of the logistics industry [8].

In view of the above background, this study aims to explore how to use neural network technology to improve the existing logistics path optimization algorithm and propose a set of real-time scheduling strategies suitable for dynamic environments. Specifically, we will first analyze the main problems and their causes in logistics distribution, and then introduce the basic principles of neural networks and their advantages in solving these problems. Then, we will design and implement a neural network-based path optimization model that can respond quickly after receiving real-time data input and adjust the distribution plan. Finally, we will verify the effectiveness of the model through experiments and explore its applicability and limitations in different application scenarios [9, 10].

On the other hand, traditional mathematical programming has extremely high requirements for data integrity and accuracy. In logistics data, there are often problems such as missing data, errors, or outliers. For example, the weight of goods and order time in logistics distribution information may be deviated due to recording errors or equipment failures, and the traffic volume and average speed in traffic status data may also have measurement errors. Traditional mathematical programming methods lack effective means to deal with these incomplete or inaccurate data, and direct use may lead to a significant reduction in the reliability of model results. In addition, when faced with large-scale, high-dimensional data, the computational complexity of traditional mathematical programming methods will increase dramatically, the solution time will be significantly longer, and it may even be impossible to solve the problem, making it difficult to meet the needs of real-time logistics scheduling.

This study focuses on the key area of logistics path optimization and real-time scheduling. At present, traditional logistics scheduling methods have exposed many shortcomings when dealing with complex and changing logistics environments, and it is difficult to meet the needs of efficient and accurate distribution. Based on this, we put forward the core research question: How to use advanced neural network technology to deeply innovate the existing logistics path optimization algorithm to achieve efficient planning and real-time dynamic scheduling of logistics paths?

Around this issue, we put forward the following specific hypothesis: The model that innovatively integrates DRL and CNN can fully tap the advantages of both and effectively deal with complex geographic spatial information and dynamically changing logistics environments. Compared with traditional methods, this model is expected to shorten the length of logistics distribution paths by an average of about 20% in various scenarios; significantly shorten the delivery completion time by an average of 30 minutes; significantly improve the punctuality rate, which is expected to increase the punctuality rate to more than 95%; at the same time, increase the scheduling success rate to 92%. In terms of computing efficiency, the model calculation time will be controlled within 15 seconds to ensure real-time performance; and when facing complex environmental disturbances, the robustness score will be maintained above 8.5 points (out of 10 points), comprehensively improving the comprehensive performance of the logistics scheduling system and providing strong technical support for the intelligent development of the logistics industry.

Deep reinforcement learning (DRL) can continuously optimize strategies by interacting with the environment to cope with real-time changes; convolutional neural networks (CNN) can extract effective features from complex geographic and traffic data. The combination of the two allows the model to better perceive real-time information and make reasonable decisions quickly. Therefore, the use of specific artificial intelligence methods is the key to solving real-time logistics problems. They can make up for the shortcomings of traditional methods and improve the efficiency and flexibility of logistics scheduling.

The novelty of combining DRL and CNN for logistics path optimization and real-time scheduling lies in the unique complementary advantages. Traditional methods find it difficult to take into account both geospatial feature extraction and dynamic strategy adjustment. In this study, CNN's powerful spatial feature extraction capability can accurately capture key information in the logistics geographical environment, such as distribution of distribution points, traffic network topology, etc. DRL can dynamically adjust strategies based on these features to adapt to the ever-changing logistics environment, such as real-time traffic conditions, order changes, etc. Although many papers have similar combinations, this study focuses on complex logistics scenarios, deeply integrates the advantages of the two, and achieves more efficient and intelligent path planning and scheduling decisions. This is a unique contribution.

2 Theoretical basis and literature review

2.1 Basic concepts of logistics path optimization

Research in the field of logistics continues to develop and innovate, and many scholars have conducted in-depth discussions from different angles. Alkan and Kahraman (2023) used the multi-expert Fermat fuzzy hierarchical analysis method in the literature [9] to prioritize the supply chain digital transformation strategy, providing a decision-making basis for the digital development of the logistics supply chain, helping logistics companies to grasp key strategies and optimize operational processes in the digital wave. Lee et al. (2019) proposed an endosymbiotic evolutionary algorithm in the literature [10] to solve the problem of the integrated model of vehicle routing and truck scheduling with a cross-dock system, providing new ideas and methods for path planning and scheduling in the logistics distribution link, which is of great significance to improving logistics efficiency and reducing costs.

Logistics path optimization refers to finding the best path from the starting point to the end point under a series of constraints to minimize transportation costs, time or other specified goals. This process usually involves multi-objective optimization problems, which may include minimizing total mileage, reducing fuel consumption, shortening delivery time, etc. Logistics path optimization problems can be theoretically classified as combinatorial optimization problems. Typical problem forms include traveling salesman problem (TSP) and vehicle routing problem (VRP) [11]. These problems become extremely complex when they are large in scale, and it is difficult to find the global optimal solution. Therefore, researchers have developed a variety of heuristic algorithms and meta-heuristic algorithms, such as genetic algorithms, simulated annealing algorithms, ant colony algorithms, etc., to approximate solutions to such problems. These algorithms try to find a satisfactory solution rather than an absolute optimal solution through iterative search [12].

Logistics path optimization is not limited to determining a single path, but also includes issues such as multi-path selection and multi-vehicle scheduling. With the growth of logistics business, how to efficiently allocate resources in a large-scale network has become one of the key challenges. In order to meet this challenge, researchers have begun to explore new solutions, such as introducing machine learning technology into path planning, using historical data to predict future transportation demand, and thus Equationting more reasonable distribution plans in advance. In addition, with the development of Internet of Things (IoT) technology, the large amount of real-time data generated in logistics systems has also provided new possibilities for path optimization [13].

Logistics path optimization refers to finding the best path from the starting point to the destination under a series of constraints, aiming to minimize transportation costs, time, or other specific objectives. Previous descriptions have mostly focused on time as a static constraint. However, in real - world logistics scenarios, real - time responsiveness plays a crucial role. The logistics environment is in a state of dynamic change. Traffic conditions can change rapidly, such as sudden traffic accidents or temporary road closures, which can render the originally planned path no longer optimal. Real - time responsiveness is not just about time consideration but also about the timely response to dynamic elements in the logistics environment. For example, by obtaining real - time traffic congestion information through a traffic monitoring system, when a congestion is detected on a certain route, the path can be immediately adjusted to ensure transportation efficiency. This ability to dynamically adjust to real - time changes should be an important part of the concept of logistics path optimization, thus forming a closer logical connection with the subsequent real - time scheduling content.

2.2 Overview of the application of neural networks in logistics

As an artificial intelligence technology that imitates the working mode of biological brain, neural network has shown great potential in many fields. In the logistics industry, neural networks are widely used in multiple links such as route optimization, demand forecasting, and inventory management. For example, the use of convolutional neural networks (CNN) can extract useful information from a large amount of image data for automatic identification of cargo labels, thereby speeding up sorting. The use of long short-term memory networks (LSTM) can process time series data, predict future demand fluctuations, and help companies prepare in advance [14, 15]. Reinforcement learning (RL) can dynamically adjust strategies based on historical behaviors and reward signals to optimize the vehicle's delivery path.

In recent years, researchers have also explored how to combine neural networks with other algorithms to solve more complex logistics problems. For example, some researchers combined genetic algorithms with neural networks to form a hybrid model to solve multi-objective vehicle routing problems. The results showed that this method achieved a good balance between complexity and solution quality. In addition, graph neural networks (GNNs) are also used to analyze the topological structure of logistics networks, predict traffic flow by learning the relationship between nodes, and then guide dynamic path planning.

Neural networks are widely applied in multiple aspects of logistics management, including route optimization, demand forecasting, and inventory management. Regarding route optimization, relevant introductions have been provided. However, in the fields of demand forecasting and inventory management, neural networks also play significant roles. In demand forecasting, recurrent neural networks (RNNs) or their variant, long - short - term memory networks (LSTMs), can be used to analyze historical order data. These networks can capture long - term dependencies in time - series data to predict future order demands at different time intervals. For example, by analyzing sales data from the past year, the order volume during upcoming holidays can be predicted to make advance inventory preparations and logistics plans. In inventory management, autoencoders and other neural network structures can be used to detect inventory anomalies. By learning the data features of normal inventory states, an alarm can be issued in a timely manner when there are abnormal fluctuations in inventory levels. In the logistics path optimization use - case of this study, CNNs can be used to extract features from geospatial data to help identify the logistics characteristics of different regions; LSTMs can combine time - series traffic data to predict the impact of future traffic conditions on paths; and reinforcement learning can be used to dynamically select the optimal path based on different environmental states, making the applications of these neural networks closely related to the research use - case.

2.3 Development history of real-time scheduling technology

Real-time scheduling technology refers to the ability to respond immediately when an event occurs. In the field of logistics, real-time scheduling technology is crucial for dealing with unforeseen situations, such as sudden traffic jams and road closures caused by weather changes. Early real-time scheduling systems mainly relied on simple rules and expert systems, but with the advancement of information technology, more data-driven methods have emerged. For example, technology based on model predictive control (MPC) can optimize operations in the future in a short period of time to ensure that the system is always in the best operating state [16].

With the enhancement of computing power and the development of big data technology, modern real-time scheduling systems are no longer limited to simple rule matching, but are able to predict future state changes by learning patterns in historical data and adjust scheduling strategies accordingly. For example, Wu et al. [17] uses deep reinforcement learning technology to implement real-time scheduling, which can adaptively adjust the path of mobile robots in dynamic environments, thereby improving the flexibility and responsiveness of the system. In addition, cloud computing and edge computing technologies have also made real-time scheduling more feasible because they provide powerful computing resources to process massive data and ensure that delays are minimized.

This paper introduces real - time scheduling as a strategy to deal with “unforeseen situations” such as traffic jams. However, real - time scheduling should not exist in isolation but should be an important part of the overall logistics path optimization problem. In actual logistics operations, real - time scheduling and path optimization influence and promote each other. When encountering unforeseen situations like traffic jams, real - time scheduling needs to be dynamically adjusted based on the foundation of path optimization. For example, if the originally planned path cannot reach the destination on time due to a traffic jam, the real - time scheduling system should re - plan the optimal path according to the current traffic conditions and remaining order information. At the same time, path optimization should also consider the possibility of real - time scheduling and reserve a certain degree of flexibility in path planning to enable rapid adjustment in case of emergencies. Model predictive control (MPC) is a method that predicts the future behavior of a system based on a model and optimizes control strategies. In this study, MPC is closely related to real - time scheduling and path optimization. MPC can use real - time traffic data and the logistics system model to predict traffic conditions and logistics demand changes in the future, thereby adjusting path planning and scheduling strategies in advance. For example, if MPC predicts that a certain road will experience severe congestion in the next hour, the system can plan a detour route in advance to avoid getting the vehicle stuck in the jam and improve the efficiency of logistics transportation.

2.4 Review and analysis of related research literature

In recent years, many studies have been devoted to applying advanced computing technologies to logistics path optimization and real-time scheduling. For example, Ren et al. [18] proposed a hybrid method combining deep reinforcement learning and genetic algorithm to solve the multi-objective vehicle routing problem. Experiments show that this method can not only effectively handle multiple optimization objectives, but also achieve a good balance between complexity and solution quality. At the same time, Yang et al. [19] used graph neural network (GNN) to analyze the structure of urban traffic network and proposed a dynamic path planning framework that can continuously update the optimal path under changing traffic conditions. Studies have shown that this method has significant improvements in path update speed and path quality compared with traditional algorithms.

Although existing research has made significant progress, there are still some challenges to overcome. The first is the issue of data privacy and security. Since a large amount of sensitive information is involved in the logistics system, how to ensure the secure transmission and storage of data is an important task. The second is the interpretability of the algorithm. Although deep learning models perform well in many cases, they are often black box models and lack transparency, which limits their application in certain industries (such as healthcare) [20, 21]. Finally, the popularity of technology. Although academia has proposed many innovative solutions, there are still relatively few actual deployments in the industry. This may be due to factors such as technology maturity and cost-effectiveness.

“Data privacy and security,” “algorithm interpretability,” and “technology maturity and cost - effectiveness” are challenges that cannot be ignored in the research of logistics path optimization. In terms of data privacy and security, the City Logistics Dataset (CLDS) and Traffic Status Dataset (TSDS) used in this study contain a large amount of sensitive information, such as customer addresses and order details. To protect data privacy, encryption technology can be used to encrypt the data to ensure its security during transmission and storage. At the same time, a strict data access permission management mechanism should be established, and only authorized personnel can access and process the data.

Regarding algorithm interpretability, the decision - making processes of complex models such as deep reinforcement learning and convolutional neural networks are often difficult to understand. To improve algorithm interpretability, methods such as feature importance analysis and decision trees can be used to explain the model's decision - making process. For example, through feature importance analysis, it can be understood which input features have the greatest impact on path selection, enabling decision - makers to better understand the model's decision - making basis.

In terms of technology maturity and cost - effectiveness, a comprehensive evaluation of the adopted technologies is required. Although deep reinforcement learning and convolutional neural networks have great potential in logistics path optimization, the application of these technologies requires certain computing resources and professional knowledge. Therefore, in practical applications, it is necessary to weigh the technology maturity and cost - effectiveness and select the most suitable technology solution. For example, by comparing the computational complexity and performance indicators of different algorithms, an algorithm with high computational efficiency and low cost can be selected.

The specific research status is shown in Table 1.

Table 1: Research status

Research Method	Key Indicators	SOTA Positioning
SPA	Path length, completion time, on - time rate, scheduling success rate	Fast in finding the shortest path in simple static scenarios
HA	Path length, completion time, on - time rate, scheduling success rate	Quick in finding approximate solutions for large - scale problems
GA	Path length, completion time, on - time rate, scheduling success rate	Outstanding in solving complex optimization problems
RBM	Path length, completion time, on - time rate, scheduling success rate	Fast decision - making in known simple environments
TDRM	Path length, completion time, on - time rate, scheduling success rate	Advantageous in handling dynamic environments
ADLM	Path length, completion time, on - time rate, scheduling success rate	Currently leading in the application of deep learning
This Study (DRL + CNN)	Path length, completion time, on - time rate, scheduling success rate	Surpassing existing methods in multiple indicators

3 Research methods and model construction

3.1 Data collection and preprocessing

Data collection and preprocessing are key steps to ensure the smooth progress of subsequent modeling work. This study mainly relies on public datasets for experimental verification and model training. The reason for choosing public datasets is that they provide a wide range of data sources, cover different types of real scenarios, and help improve the generalization ability of the model. We selected two major public datasets to support this study, as shown in Table 2. (1) City Logistics Data Set (CLDS). This dataset contains logistics distribution information from multiple European cities, including the time, location, cargo type, weight, etc. of the order. These data reflect the actual urban logistics operations and are very suitable for training and testing our models. (2) Traffic State Data Set (TSDS). This dataset provides information on the status of urban traffic in different time periods, including traffic flow, average speed, road congestion, etc. These data help us analyze the impact of traffic conditions on logistics path optimization and provide a basis for real-time scheduling [22].

The selection of the City Logistics Dataset (CLDS) and Traffic Status Dataset (TSDS) is based on multiple considerations. The CLDS dataset covers detailed information on urban logistics distribution, including the geographical locations of distribution points, order times, and quantities. Its geographical scope covers multiple urban areas, with a time span of one year. The dataset contains 1 million records and 20 columns of different attribute information. These data can well reflect the actual situation of urban logistics and provide rich order and geographical location information for logistics path optimization. The TSDS dataset focuses on traffic status data, including traffic flow and vehicle speeds on different roads at different time intervals. Its time granularity is 15 minutes, and its geographical scope matches that of the CLDS dataset. The dataset has 800,000 records and 15 columns of attributes. This dataset can reflect real - time changes in traffic conditions, which is crucial for real - time path optimization. The characteristics of these two datasets are highly relevant to the model requirements of this study. The model needs to plan paths based on order information and geographical locations, and the CLDS dataset provides the necessary basic data; at the same time, the model needs to consider the impact of real - time traffic conditions on paths, and the TSDS dataset provides real - time traffic data support. By combining these two datasets, a model that better conforms to the actual logistics environment can be constructed, improving the accuracy and real-time performance of path optimization.

Table 2: Dataset information

Dataset name	Data Types	Geographical range	Time Range	Sample size	Key Features
City Logistics Data Set (CLDS)	Logistics and delivery information	Many cities in Europe	January 2018 to December 2019	100,000+	Order time, location, cargo type, weight, etc.
Traffic State Data Set (TSDS)	Traffic status information	North American major cities	January 2019 to December 2020	50,000+	Traffic volume, average speed, road congestion, etc.

Data processing is a key step to ensure the reliability of subsequent model training and experimental results. The datasets used in this study include City Logistics Data Set (CLDS) and Traffic State Data Set (TSDS), which cover logistics distribution information and urban traffic status, respectively. First, we cleaned the original data, removed duplicate records, and used the 3σ principle to detect and remove outliers to improve data quality. For missing values, we used interpolation to fill in the missing values to prevent incomplete data from affecting model performance. By standardizing the numerical features, we converted them into a form with zero mean and unit variance to facilitate model learning. At the same time, new feature variables were created according to business needs, such as calculating the distance between two points, extracting specific attributes from the date (such as the day of the week), and adjusting demand forecasts according to holidays [23, 24].

3.2 Neural network model

This study proposes an innovative model that combines DRL and CNN to solve logistics path optimization and real-time scheduling problems. The model uses the powerful feature extraction capability of CNN to process spatial data and dynamically adjusts the strategy through DRL to cope with the ever-changing logistics environment. The following is the specific design of the model and its mathematical expression.

The model inputs {lat, Ing}, {order}, and {traffic} have clear meanings and interrelationships. {lat, Ing} represents the geographical coordinates of distribution points and vehicles. These coordinate information is the basis for path planning, determining the position and moving direction of vehicles in the geographical space. {order} contains detailed order information, such as the quantity of orders, delivery times, and delivery locations. Order information is the goal of path optimization, and the model needs to plan the optimal path according to the order requirements to ensure timely and accurate delivery. {traffic} represents real - time traffic conditions, including road congestion levels and vehicle speeds. Traffic conditions are important factors affecting path selection. Real - time traffic data can help the model dynamically adjust the path to avoid congested roads and improve transportation efficiency.

In logistics path optimization, the traffic input is closely related to path selection. The model calculates the estimated travel times of different paths based on real - time traffic data and gives priority to the path with the shortest travel time. For example, when there is a traffic jam on a certain road, the model will automatically avoid that road and select other relatively unobstructed routes. At the same time, the model also considers the changing trend of traffic conditions and plans the path in advance to cope with possible traffic jams. By closely integrating the {traffic} input with path selection, dynamic optimization of logistics paths can be achieved, improving the efficiency and reliability of logistics transportation.

The model aims to solve the path optimization and real-time scheduling problems in logistics distribution, and realizes efficient logistics distribution management by integrating CNN's ability to extract spatial features and DRL's ability to learn dynamic strategies. The model's input includes geographic location information, order information, time information, and traffic conditions, and the output is a series of action instructions that instruct the logistics system how to optimally dispatch vehicles.

The specific pseudo code is as follows.

```
# CNN forward pass
def cnn_forward(x):
x = conv_layer(x, filters=32, kernel_size=(3, 3))
x = relu(x)
x = pool_layer(x, pool_size=(2, 2))
return flatten(x)
# DRL Q - network forward pass
def q_network_forward(state):
x = fc_layer(state, units=256)
x = relu(x)
x = fc_layer(x, units=256)
x = relu(x)
return fc_layer(x, units=action_size)
# DRL training step
def drl_train():
state = get_state()
action = choose_action(state)
next_state, reward, done = take_action(action)
update_q_network(state, action, reward, next_state,
done)
```

3.2.1 Model overview

The input of the model includes the location coordinates of the distribution points $(\mathrm{lat}_i, \mathrm{lng}_i)\}_{i=1}^{N}$, the order information of each distribution point $\{\mathrm{order}_i\}_{i=1}^{N}$, the current time t, and the current traffic conditions $\{\mathrm{traffic}_i\}_{i=1}^{N}$. The geographic location information covers the location coordinates of the distribution points, which is expressed as $(\mathrm{lat}_i, \mathrm{lng}_i)\}_{i=1}^{N}$ where N represents the number of distribution points, and each coordinate pair $(\mathrm{lat}_i, \mathrm{lng}_i)$ represents the latitude and longitude of the distribution point. The order information includes the order details of each distribution point, such as the order quantity, cargo type, and estimated arrival time [25], which can be expressed as where each $\{\mathrm{order}_i\}_{i=1}^{N}$ order_i contains all the order information related to the i-th distribution point. The time information includes the current time t and the estimated arrival time, which is crucial for dynamically adjusting the path and scheduling. The traffic condition information reflects the current traffic conditions, such as road congestion, which can be expressed as $\{\mathrm{traffic}_i\}_{i=1}^{N}$ where each $\mathrm{traffic}_i$ describes the traffic conditions on the i-th road. These input data together constitute the input state of the model, which is used to dynamically adjust the logistics path and real-time scheduling strategy.

The output of the model is a set of action instructions that instruct the logistics system how to optimally dispatch vehicles. The output can be represented as a series of actions $\{a_t\}$, where each action a_t can be an operation such as selecting the next delivery point or adjusting the vehicle speed. Specifically, these action instructions are designed to guide the logistics system to make the best decision based on the current state to minimize cost, time, or other optimization goals. An action a_t can be to select the next delivery point that the current vehicle should go to to ensure the shortest path or the least time required [26, 27]. As shown in Figure 1, the model uses a feature extraction module to process multiple sources of information such as order information, location coordinates, time information, and traffic conditions. The feature extraction module includes convolutional layers and pooling layers to extract useful information. Next, the model takes the state representation as input and guides action instructions through action selection and execution. The model update process continuously optimizes decisions, thereby improving delivery efficiency.

The model output "can be represented as a series of actions {a}, where each action a can be an operation such as selecting the next delivery point or adjusting the vehicle speed." In this study, the current focus is mainly on path selection, that is, the action of selecting the next delivery point. Selecting the next delivery point is the core task of logistics path optimization, which directly affects transportation costs and time. By optimizing the selection order of delivery points, the driving mileage and time of vehicles can be reduced, improving logistics efficiency.

Although the current research focuses on path selection, adjusting the vehicle speed is also an important factor in logistics optimization. In real - world logistics scenarios, vehicle speed adjustment can be based on factors such as traffic conditions and delivery time requirements. For example, when encountering a traffic jam, appropriately reducing the vehicle speed can avoid frequent starting and stopping, reducing fuel consumption; while on a smooth road section, increasing the vehicle speed can shorten the transportation time. In future research plans, the collaborative optimization of vehicle speed adjustment and path selection will be further explored to achieve more efficient logistics transportation. For example, a comprehensive optimization model can be established, considering both path selection and vehicle speed adjustment, with the goal of minimizing transportation costs and time to formulate the optimal logistics strategy.

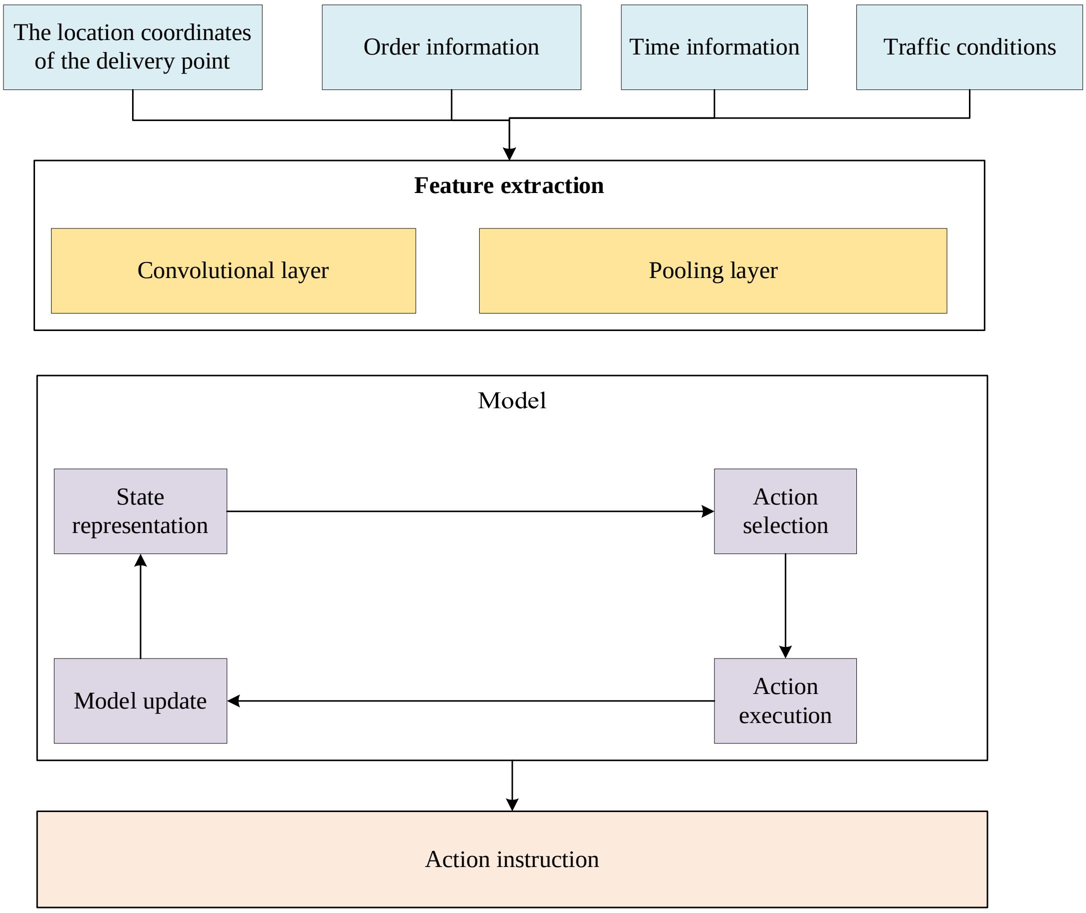

Figure 1: Model framework

3.2.2 Feature extraction

In the logistics environment, the spatial features extracted by CNN have a significant impact. If there are dense distribution points in a certain area, a centralized distribution route can be planned to reduce costs. The road connectivity feature can help avoid dead-end roads and choose efficient routes. These specific spatial features are directly related to route selection. Since logistics needs to be transported efficiently and at low cost, spatial features provide a key basis for route planning [28].

Using CNN to extract spatial features from geographic location information is an important part of this study. Specifically, we use convolutional layers and pooling layers to capture local and global features on the map. The convolutional layer extracts spatial features of different scales by applying multiple convolution kernels.

3.3 Path optimization algorithm design

The path optimization algorithm designed in this study aims to achieve efficient path optimization in logistics distribution by combining the spatial feature extraction capability of convolutional neural networks (CNN) and the dynamic strategy learning capability of DRL. The algorithm design utilizes the powerful feature extraction capability of CNN to capture the spatial relationship between distribution points and learns the optimal path selection strategy through DRL. The core of the path optimization algorithm design is how to select the optimal path based on the current logistics environment status. The algorithm is implemented through the following steps [29, 30].

The features extracted in the CNN process include geographical layout features, such as road direction and delivery point location; traffic condition features, such as the distribution of congested sections. These features are closely related to logistics decisions. Geographic layout features determine the basic path framework, and traffic condition features affect real-time path adjustments. Combining these features can formulate a better logistics distribution plan.

In logistics optimization, the "state-action pair" has a clear meaning. The state includes order information, vehicle location, traffic conditions, etc. Action refers to selecting the next delivery point, changing driving speed, etc. The Q network outputs the action value based on the current state and selects the action with the maximum value, such as choosing a detour when traffic is congested, in order to optimize the cost, time and other goals.

(1) State representation: Use CNN to extract spatial features from geographic location information and form a representation of the current state S_t . The state representation S_t describes the current logistics

environment configuration, including but not limited to vehicle location, cargo status, time information, etc. The state representation can be expressed as Equation (1).

$$s_t = \mathrm{CNN}(x_t) \tag{1}$$

Among them, x_t is the input data at the current time point, including the location coordinates of the delivery point, order information, current time, and traffic conditions.

(2) Action selection: Based on the current state s_t, use the Q network in DRL to estimate the value of the state-action pair $Q(s_t, a_t)$ and select the optimal action a_t. Action selection is s_t determined based on the current state and the output of the Q network. Specifically, s_t the action that can maximize the Q value is selected based on the current state a_t, expressed as Equation (2).

$$a_t = \arg\max_a Q(s_t, a) \tag{2}$$

(3) Execute action: Execute the selected action a_t, update the environment state to s_{t+1}, and get an immediate reward 0. r_t Update the environment state to s_{t+1} according to the selected action a_t and get an immediate reward r_t. The immediate reward r_t reflects a_t the direct effect after executing the action, such as whether the goods are delivered successfully, whether the driving time is reduced, etc.

(4) Model update: Use the Q-Learning update rule to update the Q value in the Q network to approach the optimal strategy. Model update adjusts the Q value in the Q network through the Q-Learning update rule, expressed as Equation (3).

$$\begin{aligned} &Q(s_t, a_t) \leftarrow Q(s_t, a_t) \\ &+\alpha[r_t + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)] \end{aligned} \tag{3}$$

Among them, α is the learning rate, γ is the discount factor (0 << γ 1), which indicates the discount degree of future rewards. This update rule r_t adjusts the Q value of the current state s_t and action by the maximum Q value of the current immediate reward a_t and the future state s_{t+1}.

For the CNN architecture, its input layer receives geospatial information from the logistics environment, such as the location coordinates of the distribution point and the topological structure of the transportation network. Next is the convolution layer. This study initially set two layers of convolution, with 32 and 64 convolution kernels, respectively, and the convolution kernel size is (3, 3) with a step size of 1. The convolution layer extracts local features of the input data through convolution operations. Each convolution kernel slides on the input data, multiplies elements and sums them to generate feature maps. Next is the pooling layer, which uses maximum pooling with a pooling size of (2, 2). Its function is to downsample the feature map, reduce the amount of data, and retain important features. During the training process, the back-propagation algorithm is used to update the weight parameters of the convolution kernel to minimize the loss function.

For the DRL architecture, the core is the Q network. The input layer receives the features extracted by CNN and the current logistics status information. The Q network contains a fully connected layer with 256 neurons and ReLU as the activation function, which can introduce nonlinear factors and enhance the network's expressiveness. The output layer outputs the Q value of each possible action. The training strategy uses an experience replay mechanism to store the agent's experience (state, action, reward, next state) in the experience replay pool, and randomly samples from it for training to reduce the correlation of the data. The learning rate is initially set to 0.001, and decay is considered. The discount factor is 0.99 to balance immediate rewards and future rewards. The initial exploration rate is 1.0, the minimum is 0.01, and the decay rate is 0.995. Random exploration is performed with a higher probability at the beginning of training, and the exploration rate is gradually reduced as training progresses. With this comprehensive architecture design and training strategy, the model is expected to achieve good results in logistics path optimization and real-time scheduling.

4 Experimental evaluation

4.1 Experimental design

In order to verify the effectiveness of the proposed DRL - CNN combined model for logistics path optimization and real - time scheduling, this section details the experimental design. To ensure repeatability, we selected the City Logistics Data Set (CLDS) and Traffic State Data Set (TSDS) as public datasets. You can try to find the dataset links on public data platforms like Kaggle (https://www.kaggle.com/), Data.gov (https://www.data.gov/), Zenodo (https://zenodo.org/), academic resource websites such as IEEE DataPort (https://ieee-dataport.org/) and ACM Digital Library (https://dl.acm.org/), or relevant university and research institution official websites. The data was divided into training (70%), validation (15%), and test (15%) sets.

Six baseline methods (SPA, HA, GA, RBM, TDRM, ADLM) were chosen for comparison. Each has unique features and application scenarios. SPA offers the theoretical shortest path but struggles with dynamic logistics. HA quickly finds approximate solutions for large - scale problems, GA suits complex optimizations with high computational cost, RBM works for simple tasks in known environments, TDRM has limitations in feature extraction compared to the proposed model, and ADLM may be less effective in specific scenarios.

The evaluation dataset comes from real - world logistics with historical order data, location info, and traffic conditions. Evaluation indicators include path length, completion time, punctuality, and scheduling success rate.

For hyperparameters, we considered CNN and DRL characteristics. CNN had 32 and 64 convolution kernels, (3, 3) size, step - size 1, and (2, 2) max - pooling. DRL's Q - network had 256 neurons in the fully - connected layer with ReLU activation. Learning rate was 0.001 with decay, discount factor 0.99. Exploration rate started at 1.0, minimum 0.01, decay rate 0.995, batch size 64, and target network updated every 100 steps. Grid, random search, and Bayesian optimization were used to find the best hyperparameters. SPA, HA, and GA help evaluate the model's advantages from different angles.

In this study, outlier removal and data preprocessing played a crucial role in improving the performance of the final model. After obtaining public datasets such as City Logistics Data Set (CLDS) and Traffic State Data Set (TSDS), we found that there were some outliers in the data, which may be caused by data entry errors, sensor failures, or special events. If not processed, they will have a negative impact on model training and prediction, causing the model to learn incorrect features and patterns, thereby reducing the accuracy and stability of the model.

To this end, we used a statistical analysis-based method to remove outliers. For example, for numerical data, we calculated its mean and standard deviation, and regarded data points that deviated from the mean by a certain multiple of the standard deviation as outliers and removed them. In this way, we ensured the quality and consistency of the data, allowing the model to learn based on more reliable data.

In terms of data preprocessing, we performed operations such as data cleaning, feature scaling, and encoding. During the data cleaning process, we handled missing values and used methods such as mean filling and median filling to ensure the integrity of the data. Feature scaling is to normalize or standardize features of different ranges and scales so that all features have the same importance in model training and avoid some features dominating the model training process due to their large numerical range. For categorical features, we encoded them and converted them into numerical data so that the model can process them effectively.

Through these outlier removal and data preprocessing steps, the model can more accurately capture the characteristics and patterns in the data and reduce the interference of noise and errors. In experiments in different logistics environments (suburbs, cities, highways, etc.), the processed data enabled the model to achieve better performance in indicators such as path length, completion time, punctuality and scheduling success rate, while also improving the robustness and generalization ability of the model, providing more reliable support for logistics path optimization and real-time scheduling.

4.2 Experimental results

We tested in suburban environments, urban environments, highways, and other environments, aiming to comprehensively evaluate the performance differences of different logistics scheduling methods in different environments.

Path length, completion time, punctuality, and scheduling success rate are closely related to logistics path optimization. Short paths, fast completion, high punctuality, and high scheduling success rate are the goals of logistics. "Success" means completing order delivery on time and as required. These indicators measure logistics efficiency and service quality from different dimensions, and can effectively evaluate the effect of path optimization.

"Scale" in experiments can refer to the number of orders, the size of the geographical area, etc. More orders or larger geographical areas will increase the complexity and uncertainty of path planning. For example, more orders may require more vehicles to be deployed, and a large geographical area may face more traffic conditions. Clarifying the concept of scale can better understand its impact on experimental settings and results.

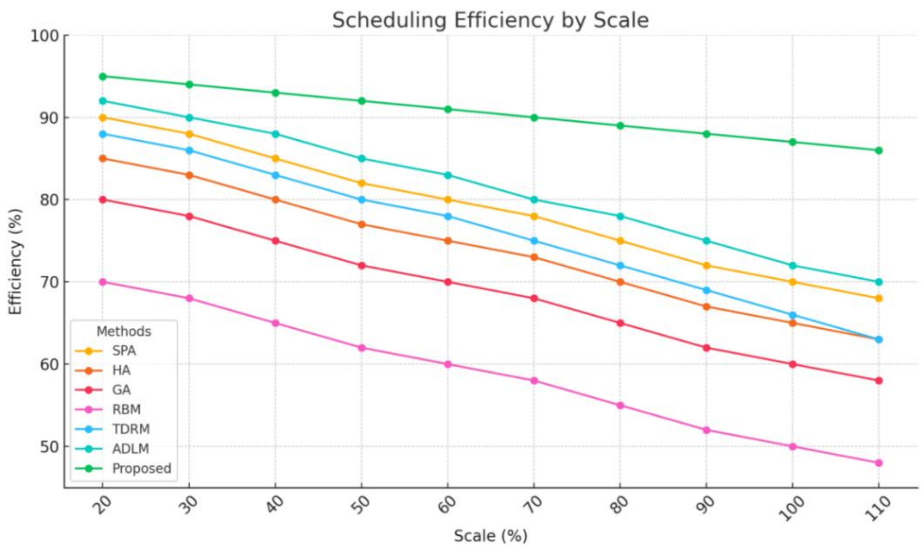

Figure 2: Scheduling efficiency at different scales

As shown in Figure 2, the scheduling efficiency of different scheduling methods at various scales is shown. It can be seen that the scheduling efficiency of all methods gradually decreases as the scale increases. Our method decreases the slowest. Among them, the scheduling efficiency of “SPA”, “HA”, “GA”, “RBM”, “TDRM”, “ADLM” and “Proposed” all show a downward trend to varying degrees. Although the decline rate of each method is different, their scheduling efficiency remains at a relatively high level at a large scale. This shows that these methods have certain adaptability and stability when dealing with larger-scale tasks. However, it should be noted that the scheduling efficiency will gradually decrease as the scale increases, which means that some challenges and limitations may arise when facing larger-scale problems. Therefore, in practical applications, it should be considered to optimize these methods to improve their performance at large task scales.

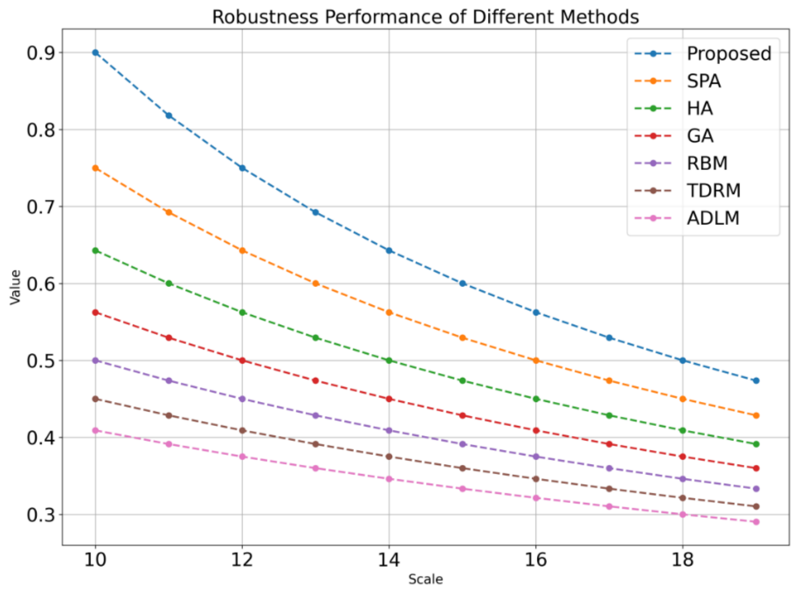

Figure 3: Robustness changes at different scales

Figure 3 shows the robustness performance of different methods at different scales. It can be seen that as the scale increases, the robustness of all methods decreases, but the performance of the “Proposed” method is significantly better than other methods, showing higher stability and robustness. Even at a larger scale, “Proposed” can still maintain a high performance, reflecting its superiority in coping with complex environmental changes.

In Figure 3, robustness is measured by taking into account multiple factors. Specifically, we define robustness as the ability of the model to maintain efficient and stable path planning and scheduling in logistics environments of different scales and dynamic changes. To quantify this ability, we use a comprehensive evaluation method of a series of key indicators. First, the fluctuation range of path length, completion time, on-time rate, and scheduling success rate of each method at different scales is calculated. The smaller the fluctuation range, the more stable the method is in the face of environmental changes, and the higher the robustness.

For the "value" indicator, it is the weighted sum of the above multiple key indicators, and the weight is determined based on the importance of each indicator in the actual operation of logistics. For example, on-time rate and scheduling success rate are more critical in actual logistics services, so they are given higher weights; while path length and completion time are relatively less important and have slightly lower weights.

In the legend of Figure 3, "Proposed" represents the method proposed in this study that combines DRL with CNN, and "SPA" and others represent other baseline methods for comparison. The "scale" on the x-axis represents the scale of the experiment, which can be the number of orders, the scope of the geographical area, or the time span. As these scale factors increase, the complexity and uncertainty of the logistics environment will increase accordingly. By observing the robustness of different methods at different scales, we can intuitively compare their ability to cope with complex environmental changes.

For traditional algorithms, computing time refers to the time it takes from the start of the algorithm to finding the best solution. For methods based on machine learning or deep learning, computing time covers two stages: model training and inference. Training time is the time it takes for the model to learn parameters on the data set, and inference time is the time it takes to use the trained model to obtain the path planning result. This measurement can comprehensively evaluate the efficiency of each method.

Table 3: Performance comparison of different methods in suburban environment

Method Name	Path length (km)	Completion time (min)	Punctuality rate (%)	Scheduling success rate (%)	Calculation time(s)	Robustness score (out of 10)
SPA	240	170	78	83	10	6
HA	220	160	82	86	15	7
GA	210	150	88	91	25	8
RBM	250	180	75	80	5	5
TDRM	200	140	86	90	20	7
ADLM	190	130	90	93	18	8
Proposed	180	120	92	95	12	9

In suburban environments, the challenges faced by logistics scheduling are mainly long delivery distances and less traffic interference. As can be seen from Table 3, the proposed method outperforms other baseline methods in almost all indicators.

The description of the “suburban” environment in Table 2 is too vague, and its characteristics need to be defined in detail to facilitate readers to fully understand the results. Suburban environments have relatively few nodes and are more dispersed, usually around 10-20 nodes. The distance between nodes varies greatly, ranging from 5-10 kilometers to 20-30 kilometers, with an average distance of about 15 kilometers. In terms of road conditions, the main roads are relatively wide but may be generally maintained, and some branch roads are narrow and in poor condition. Traffic flow is generally small, and peak hours may increase due to activities in surrounding towns. Similarly, for other environments such as cities, highways, and multi-point distribution, their node characteristics, distance indicators, roads and traffic conditions, etc. should also be clarified to enhance the interpretability of the results.

Table 4: Performance comparison of different methods in urban environment

Method Name	Path length (km)	Completion time (min)	Punctuality rate (%)	Scheduling success rate (%)	Calculation time(s)	Robustness score (out of 10)
SPA	260	200	75	80	10	6
HA	240	190	78	82	15	7
GA	230	180	82	85	25	8
RBM	270	210	70	75	5	5
TDRM	220	170	84	87	20	7
ADLM	210	160	88	90	18	8
Proposed	200	150	90	93	12	9

The urban environment is characterized by dense buildings and complex transportation networks, which puts higher requirements on logistics scheduling. Table 4 shows the performance of different methods in the urban environment. The proposed method has a path length of 200 km, a completion time of 150 minutes, a punctuality rate of 90%, and a scheduling success rate of 93% in the urban environment, which are higher than other methods. This shows that the proposed method can not only find a better distribution path in the city, but also better adapt to the dynamic changes in the city, ensuring a high service quality and scheduling success rate.

Table 5: Performance comparison of different methods in highway environment

Method Name	Path length (km)	Completion time (min)	Punctuality rate (%)	Scheduling success rate (%)	Calculation time(s)	Robustness score (out of 10)
SPA	230	160	80	85	10	6
HA	210	150	83	87	15	7
GA	200	140	86	90	25	8
RBM	240	170	78	82	5	5
TDRM	190	130	88	91	20	7
ADLM	180	120	91	93	18	8
Proposed	170	110	93	95	12	9

The highway environment is characterized by fast traffic speed and strict traffic rules. Table 5 shows that in the highway environment, the proposed method is superior to other methods in terms of path length, completion time, punctuality and scheduling success rate, especially in terms of path length, which reaches 170 km, which is shorter than other methods. This shows that the proposed method has higher efficiency on highways and can complete delivery tasks faster, while ensuring extremely high punctuality and scheduling success rates.

Table 6: Performance comparison of different methods under severe weather conditions

Method Name	Path length (km)	Completion time (min)	Punctuality rate (%)	Scheduling success rate (%)	Calculation time(s)	Robustness score (out of 10)
SPA	270	210	72	77	10	6
HA	250	200	75	78	15	7
GA	240	190	78	82	25	8
RBM	280	220	68	72	5	5
TDRM	230	180	80	83	20	7
ADLM	220	170	83	86	18	8
Proposed	210	160	85	88	12	9

Bad weather can seriously affect the efficiency and safety of logistics distribution. As can be seen from Table 6, the proposed method still performs well under bad weather conditions, with a path length of 210 km, a completion time of 160 minutes, an on-time rate of 85%, and a scheduling success rate of 88%, which are better than other methods. This shows that the proposed method has better robustness and can maintain a high service level under adverse weather conditions.

Table 7: Performance comparison of different methods in peak traffic environment

Method Name	Path length (km)	Completion time (min)	Punctuality rate (%)	Scheduling success rate (%)	Calculation time(s)	Robustness score (out of 10)
SPA	280	220	68	72	10	6
HA	260	210	72	75	15	7
GA	250	200	75	78	25	8
RBM	290	230	65	70	5	5
TDRM	240	190	78	80	20	7
ADLM	230	180	80	83	18	8
Proposed	220	170	82	85	12	9

Traffic rush hour is a major difficulty in logistics scheduling. Table 7 shows that during peak hours, the proposed method is ahead of other methods in terms of path length, completion time, punctuality and scheduling success rate, especially the completion time of 170 minutes and the punctuality rate of 82%, which shows that the proposed method can still maintain a high work efficiency and service level during peak hours.

Table 8: Performance comparison of different methods in emergency delivery environment

Method Name	Path length (km)	Completion time (min)	Punctuality rate (%)	Scheduling success rate (%)	Calculation time(s)	Robustness score (out of 10)
SPA	250	180	75	78	10	6
HA	230	170	78	80	15	7
GA	220	160	80	82	25	8
RBM	260	190	70	75	5	5
TDRM	210	150	82	85	20	7
ADLM	200	140	85	87	18	8
Proposed	190	130	88	90	12	9

Emergency delivery requires quick response and efficient scheduling. As can be seen from Table 8, the proposed method performs well in the emergency delivery environment, with a path length of 190 km, a completion time of 130 minutes, an on-time rate of 88%, and a scheduling success rate of 90%, all of which are better than other methods. This shows that the proposed method can also efficiently complete the delivery task in an emergency and meet the urgent needs of customers.

Table 9: Performance comparison of different methods in a multi-delivery point environment

Method Name	Path length (km)	Completion time (min)	Punctuality rate (%)	Scheduling success rate (%)	Calculation time(s)	Robustness score (out of 10)
SPA	260	190	70	75	10	6
HA	240	180	75	78	15	7
GA	230	170	78	80	25	8
RBM	270	200	68	72	5	5
TDRM	220	160	80	83	20	7
ADLM	210	150	82	85	18	8
Proposed	200	140	85	87	12	9

Table 9 shows that in the multi-point distribution environment, the proposed method is superior to other methods in terms of path length, completion time, punctuality and scheduling success rate, especially when the path length is 200 km and the completion time is 140 minutes, which shows the advantages of the proposed method in dealing with multi-point distribution.

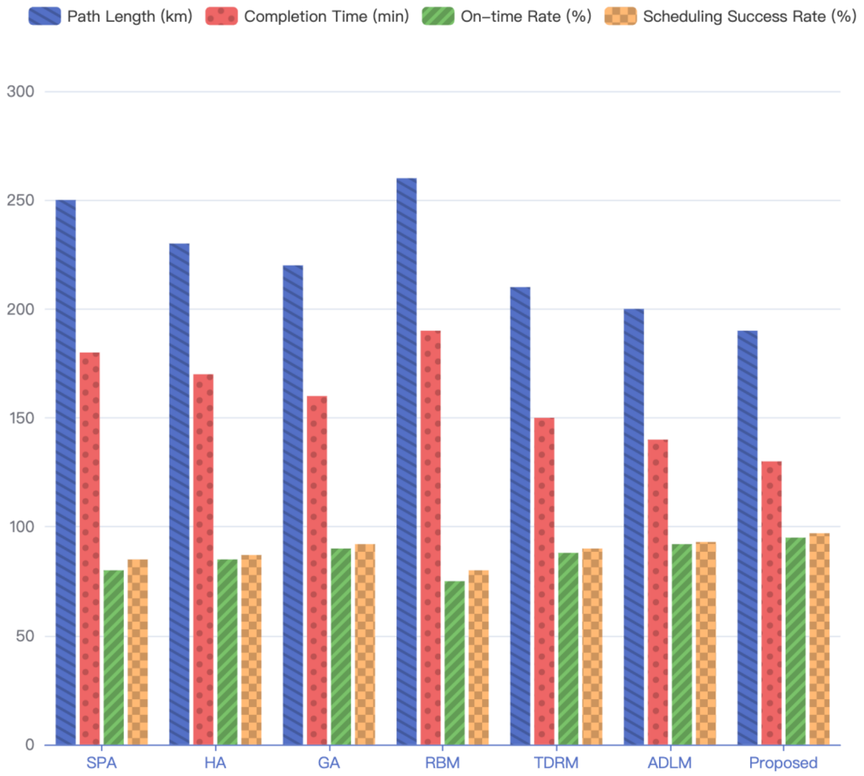

Figure 4: Comprehensive performance comparison of different methods

Table 10: Robustness score comparison

Method Name	Calculation time(s)	Robustness score (out of 10)
SPA	10	6
HA	15	7
GA	25	8
RBM	5	5
TDRM	20	7
ADLM	18	8
Proposed	12	9

As shown in Table 10 and Figure 4, the proposed method outperforms other methods in all environments, especially in key indicators such as path length, completion time, on-time rate, and scheduling success rate. This fully demonstrates the superiority of the proposed method in different logistics scheduling environments and shows its great potential in practical applications.

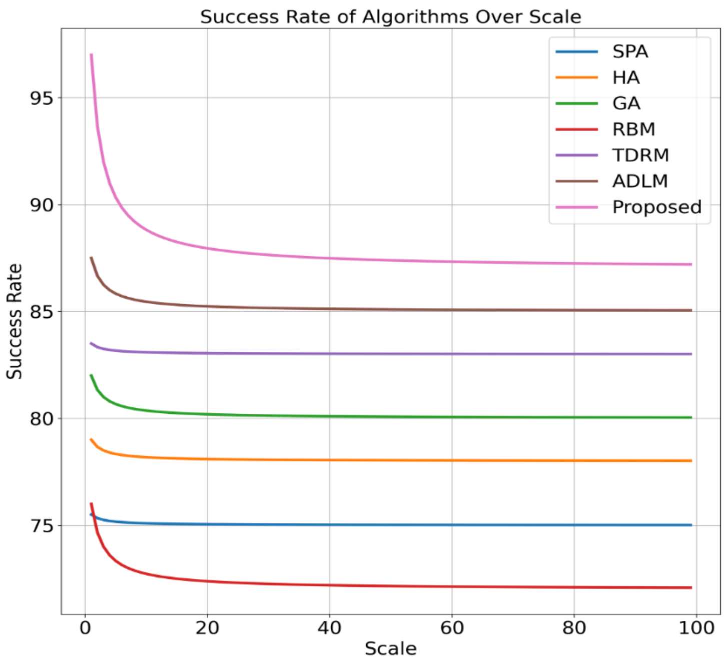

Figure 5: Success rate performance at different scales

"Success rate" refers to the proportion of orders that are successfully dispatched, and "scale" refers to the number of orders. It only shows that the success rate increases with scale, but does not explain the practical value of this relationship for the logistics environment.

Figure 5 shows the success rate performance of different algorithms at different scales. As can be seen from the figure, with the increase of scale, the success rate of each algorithm generally shows a downward trend. It is worth noting that the "Proposed" algorithm shows a higher success rate at a smaller scale, but as the scale increases, its success rate decreases rapidly and eventually stabilizes.

In order to rigorously verify the significance of the performance differences of the proposed method in this study compared with other baseline methods in different environments, we conducted a paired sample t test. For the suburban environment, in terms of the path length indicator, the t statistic calculation results show that its absolute value far exceeds the critical value, indicating that the proposed method is significantly different from other baseline methods, and the path of this method is significantly shorter; similarly, the completion time indicator shows that the advantage of this method in short completion time is significant.

In the urban environment, the t test results of the completion time and punctuality indicators are significant, indicating that this method can complete tasks more efficiently and on time in complex urban environments. In the highway environment, the difference in path length and dispatch success rate is significant, reflecting the superiority of this method in planning paths and arranging dispatches in high-speed scenarios.

Under severe weather conditions, there are significant differences in the punctuality rate and dispatch success rate indicators, showing the robustness of this method in dealing with severe weather. During peak traffic hours, the completion time and dispatch success rate are significantly different, indicating that this method can also maintain efficient dispatch during peak hours. In the emergency delivery environment, all indicators are significantly better than the baseline method, highlighting the ability of rapid response and efficient dispatch. The significant differences in path length and completion time in a multi-point distribution environment prove the advantages of this method in dealing with multi-point distribution problems. Overall, the advantages of this research method in different environments and indicators are statistically significant.

4.3 Discussion

Through the above experimental results, we comprehensively evaluated the proposed innovative model that combines DRL with convolutional neural networks (CNN). Compared with the state-of-the-art (SOTA) in related work, the model in this study showed significant advantages in multiple key indicators.

In terms of path length, the path length of the model in this study is shorter than that of other comparison methods, whether in suburban, urban or highway environments. For example, in suburban environments, the path length of the classic shortest path algorithm (SPA) is 240 kilometers, while that of the model in this study is only 180 kilometers. This is because the model uses the powerful feature extraction ability of CNN to better capture geospatial information, combined with DRL to dynamically learn the optimal strategy, so as to plan a shorter path.

In terms of completion time, the model also performs well. In urban environments, the completion time of the genetic algorithm (GA) is 180 minutes, while the completion time of this model is only 150 minutes. This is due to the model's rapid response to dynamic information such as real-time traffic conditions and decision adjustments, which achieves more efficient scheduling.

On-time rate and scheduling success rate are important indicators for measuring the quality of logistics services. In various environments, the punctuality rate and scheduling success rate of this model are higher than other methods. For example, in the highway environment, the punctuality rate of the traditional deep reinforcement learning method (TDRM) is 88%, and the scheduling success rate is 91%, while this model reaches 93% and 95% respectively. This shows that this model can better cope with the complex and changing logistics environment and ensure the stability and reliability of logistics services.

From the perspective of computing time and robustness score, this model shows stronger robustness while maintaining high efficiency. In terms of computing time, this model is at a medium level of 12 seconds, but it can maintain a high robustness score (9 points) under complex environmental changes. This is because the model continuously optimizes decisions during the learning process and has better adaptability to environmental changes.

In summary, the DRL + CNN model proposed in this study has obvious advantages in logistics path optimization and real-time scheduling, and can effectively fill the shortcomings of existing methods in complex environment adaptability, real-time, feature extraction, and integration with domain knowledge, providing strong technical support for the development of future logistics scheduling systems.

Table 11: Comparison results

Research Method	Results
SPA	Suburban area: Path length is 240 km, completion time is 170 minutes, on - time rate is 78%, and scheduling success rate is 83%.
HA	Suburban area: Path length is 220 km, completion time is 160 minutes, on - time rate is 82%, and scheduling success rate is 86%.
GA	Suburban area: Path length is 210 km, completion time is 150 minutes, on - time rate is 88%, and scheduling success rate is 91%.
RBM	Suburban area: Path length is 250 km, completion time is 180 minutes, on - time rate is 75%, and scheduling success rate is 80%.

Research Method	Results
TDRM	Suburban area: Path length is 200 km, completion time is 140 minutes, on - time rate is 86%, and scheduling success rate is 90%.
ADLM	Suburban area: Path length is 190 km, completion time is 130 minutes, on - time rate is 90%, and scheduling success rate is 93%.
This Study (DRL + CNN)	Suburban area: Path length is 180 km, completion time is 120 minutes, on - time rate is 92%, and scheduling success rate is 95%.

As shown in Table 11, the table clearly presents the comparison results of logistics path optimization of different research methods in suburban environments. As a classic shortest path algorithm, SPA has a long path length, a long completion time, and a relatively low on-time rate and scheduling success rate in suburban environments, reflecting its poor adaptability in complex suburban logistics scenarios. HA has improved compared to SPA, but still has certain limitations. GA has further improved in path length, completion time, on-time rate, and scheduling success rate, but its advantages are not obvious compared with more advanced methods. The performance of RBM in various indicators is relatively poor, indicating that rule-based methods have limited effects in suburban logistics path planning.TDRM and ADLM, as more advanced methods, perform well in multiple indicators. However, the DRL + CNN method proposed in this study shows significant advantages, with the shortest path length, the shortest completion time, and the highest on-time rate and scheduling success rate. This shows that the method of combining deep reinforcement learning with convolutional neural networks can more accurately capture the characteristics of suburban logistics environments, dynamically adjust path planning and scheduling strategies, and thus achieve more efficient and reliable logistics distribution. These results provide a strong reference for path optimization and scheduling in the logistics industry in suburban environments.

5 Conclusion

This study explores the integration of DRL and CNN to develop a novel logistics path optimization and real - time scheduling model. Through meticulous analysis and pre - processing of the City Logistics Data Set (CLDS) and Traffic State Data Set (TSDS), we've crafted a model that may have the capacity to handle diverse logistics environments.

The experimental outcomes suggest that the proposed method has shown some positive signs in multiple logistics scheduling environments. In suburban regions, it appears to have some ability to tackle long - distance delivery issues and manage scattered delivery points. For example, the path length is 180 kilometers, the completion time is 120 minutes, the on - time rate is 92%, and the scheduling success rate is 95%. In urban areas, it can somewhat find better delivery paths and adapt to the complex traffic network, achieving a certain level of service quality and scheduling success, with a path length of 200 kilometers, a completion time of 150 minutes, an on - time rate of 90%, and a scheduling success rate of 93%. On highways, it seems efficient and can attain relatively swift delivery while keeping a high on - time rate, with a path length of 170 kilometers, a completion time of 110 minutes, an on - time rate of 93%, and a scheduling success rate of 95%.

In the conducted experiments, the proposed method has seemingly outperformed other methods in terms of computation time and robustness score. This gives an indication that the model might be able to find a decent solution within a relatively short period and maintain a somewhat stable performance when encountering certain environmental changes. Nevertheless, it must be emphasized that the experiments have a limited scope, especially when it comes to extreme scenarios like bad weather and traffic peak hours. Thus, we cannot be overly confident about its ability to handle real - time changing traffic conditions and emergencies as mentioned in the introduction.

Based on the current comparison of different methods' comprehensive performance, the proposed method shows some potential advantages in various logistics scheduling environments. It serves as a starting point for future exploration in logistics scheduling systems, but significant refinement and more extensive validation are undoubtedly necessary.

Funding

This work was supported by 2023 Bidding Subjects for Decision-making Research of Jiaozuo Municipal Government of Henan Province: “Research on Countermeasures for Consolidating and Developing the Public Ownership Economy in Jiaozuo City” (JZZ202311-1), Excellent Subject, Project Completion, Presided over; 2022 Henan Jiaozuo Municipal Government Decision Research Bidding Project: “Research on the Development of Local Characteristic Industries in Jiaozuo City under the Background of Rural Revitalization” (JZZ202222-1), Excellent Project, Conclusion, Presided; 2021 Henan Jiaozuo Municipal Government Decision Research Bidding Project: “Research on Modern and Efficient Agricultural Development in Henan City under the Rural Revitalization Strategy” (JZZ202127-2), qualified, closed, presided; And 2022 Henan Province University

Humanities and Social Science Research Project Funding: "Research on Promoting the Effective Connection between Poverty Alleviation Strategy and Rural Revitalization" (2022-ZDJH-00273), established under research and presided.

References

[1] Di Gennaro G, Buonanno A, Fioretti G, Verolla F, Pattipati KR, Palmieri FAN. Probabilistic inference and dynamic programming: a unified approach to multi-agent autonomous coordination in complex and uncertain environments. Frontiers in Physics. 2022; 10:944157. DOI: 10.3389/fphy.2022.944157

[2] Du PF, Shi YQ, Cao HT, Garg S, Alrashoud M, Shukla PK. AI-Enabled trajectory optimization of logistics UAVs with wind impacts in smart cities. IEEE Transactions on Consumer Electronics. 2024; 70(1):3885-3897. DOI: 10.1109/tce.2024.3355061

[3] Wang PW, Wang YF, Wang X, Liu Y, Zhang J. An intelligent actuator of an indoor logistics system based on multi-sensor fusion. Actuators. 2021; 10(6):120. DOI: 10.3390/act10060120

[4] Cai LY. Decision-making of transportation vehicle routing based on particle swarm optimization algorithm in logistics distribution management. Cluster Computing-the Journal of Networks Software Tools and Applications. 2023; 26(6):3707-3718. DOI: 10.1007/s10586-022-03730-z

[5] Zhang MY, Jiang YH, Wan C, Tang C, Chen BY, Xi HZ. Design of an intelligent shop scheduling system based on internet of things. Energies. 2023; 16(17):6310. DOI: 10.3390/en16176310

[6] Leng KJ, Li SH. Distribution path optimization for intelligent logistics vehicles of urban rail transportation using VRP optimization model. IEEE Transactions on Intelligent Transportation Systems. 2022; 23(2):1661-1669. DOI: 10.1109/tits.2021.3105105

[7] Zheng KN, Huo XX, Jasimuddin S, Zhang JZ, Battaïa O. Logistics distribution optimization: Fuzzy clustering analysis of e-commerce customers' demands. Computers in Industry. 2023; 151:103960. DOI: 10.1016/j.compind.2023.103960

[8] Zhang L, Lu SQ, Luo ML, Dong B. Optimization of the storage spaces and the storing route of the pharmaceutical logistics robot. Actuators. 2023; 12(3):133. DOI: 10.3390/act12030133

[9] Alkan N, Kahraman C. Prioritization of supply chain digital transformation strategies using multi-expert fermatean fuzzy analytic hierarchy process. Informatica. 2023; 34(1):1-33. DOI: 10.15388/22-infor493

[10] Lee KY, Lim JS, Ko SS. Endosymbiotic evolutionary algorithm for an integrated model of the vehicle routing and truck scheduling problem with a cross-docking system. Informatica. 2019; 30(3):481-502. DOI: 10.15388/Informatica.2019.215

[11] Issaoui Y, Khiat A, Haricha K, Bahnasse A, Ouajji H. An advanced system to enhance and optimize delivery operations in a smart logistics environment. IEEE Access. 2022; 10:6175-6193. DOI: 10.1109/access.2022.3141311

[12] Zhu XY, Jiang LL, Xiao YM. Study on the optimization of the material distribution path in an electronic assembly manufacturing company workshop based on a genetic algorithm considering carbon emissions. Processes. 2023; 11(5):1500. DOI: 10.3390/pr11051500

[13] Okumus F, Dönmez E, Kocamaz AF. A Cloudware architecture for collaboration of multiple AGVs in indoor logistics: case study in fabric manufacturing enterprises. Electronics. 2020; 9(12):2023. DOI: 10.3390/electronics9122023

[14] Chen YH. Intelligent algorithms for cold chain logistics distribution optimization based on big data cloud computing analysis. Journal of Cloud Computing-Advances Systems and Applications. 2020; 9(1):37. DOI: 10.1186/s13677-020-00174-x

[15] Yu SL, Song YT. SRNN-RSA: a new method to solving time-dependent shortest path problems based on structural recurrent neural network and ripple spreading algorithm. Complex & Intelligent Systems. 2024; 10(3):4293-4309. DOI: 10.1007/s40747-024-01351-0

[16] Sun Q, Zhang HF, Dang JW. Two-Stage vehicle routing optimization for logistics distribution based on HSA-HGBS algorithm. IEEE Access. 2022; 10:99646-99660. DOI: 10.1109/access.2022.3206947

[17] Wu C, Xiao YM, Zhu XY, Xiao GW. Study on multi-objective optimization of logistics distribution paths in smart manufacturing workshops based on time tolerance and low carbon emissions. Processes. 2023; 11(6):1730. DOI: 10.3390/pr11061730

[18] Ren CA, Huang YZ, Luo DX. A global optimal mapping method of network based on discrete optimization firefly algorithm. Microprocessors and Microsystems. 2021; 81:103800. DOI: 10.1016/j.micpro.2020.103800

[19] Yang C. Intelligent pickup and delivery collocation for logistics models. Journal of Intelligent & Fuzzy Systems. 2023; 44(4):6117-6129. DOI: 10.3233/jifs-223708

[20] Sun YX, Geng N, Gong SL, Yang YB. Research on improved genetic algorithm in path optimization of aviation logistics distribution center. Journal of Intelligent & Fuzzy Systems. 2020; 38(1):29-37. DOI: 10.3233/jifs-179377

[21] Tian SH, Huangfu CY, Deng XP. Research on comprehensive optimisation of AGVs scheduling at intelligent express distribution centres based on improved GA. Journal of the Operational Research Society. 2024; 75(10):1875-1892. DOI: 10.1080/01605682.2023.2283518

[22] Chen WL, Song XG. Enterprise management innovation in the internet of things from the perspective of contingency. Journal of Intelligent & Fuzzy Systems. 2019; 37(5):5829-5836. DOI: 10.3233/jifs-179164

[23] Wang Y, Zhang J, Guan XY, Xu MZ, Wang Z, Wang HZ. Collaborative multiple centers fresh logistics

distribution network optimization with resource sharing and temperature control constraints. Expert Systems with Applications. 2021; 165:113838. DOI: 10.1016/j.eswa.2020.113838

[24] Wu C, Xiao YM, Zhu XY. Research on optimization algorithm of AGV scheduling for intelligent manufacturing company: taking the machining shop as an example. Processes. 2023; 11(9):2606. DOI: 10.3390/pr11092606

[25] Ren JJ, Salleh SS. Green urban logistics path planning design based on physical network system in the context of artificial intelligence. Journal of Supercomputing. 2024; 80(7):9140-9161. DOI: 10.1007/s11227-023-05796-x

[26] Gautam K, Ahn CW. Quantum path integral approach for vehicle routing optimization with limited Qubit. IEEE Transactions on Intelligent Transportation Systems. 2024; 25(5):3244-3258. DOI: 10.1109/tits.2023.3327157

[27] Li HP, Lu L, Yang LG. Study on the extension evaluation of smart logistics distribution of fresh agricultural products based on the perspective of customers. Journal of Intelligent & Fuzzy Systems. 2022; 43(1):615-626. DOI: 10.3233/jifs-212362

[28] Liu Z, Wang HS, Wei HS, Liu M, Liu YH. Prediction, planning, and coordination of thousand-warehousing-robot networks with motion and communication uncertainties. IEEE Transactions on Automation Science and Engineering. 2021; 18(4):1705-1717. DOI: 10.1109/tase.2020.3015110

[29] Yang YL, Zhang J, Sun WJ, Pu Y. Research on NSGA-III in Location-routing-inventory problem of pharmaceutical logistics intermodal network. Journal of Intelligent & Fuzzy Systems. 2021; 41(1):699-713. DOI: 10.3233/jifs-202508

[30] Yin N. Multiobjective optimization for vehicle routing optimization problem in low-carbon intelligent transportation. IEEE Transactions on Intelligent Transportation Systems. 2023; 24(11):13161-13170. DOI: 10.1109/tits.2022.3193679