SIN-CHN SCIENTIFIC PRESS PTE.LTD.

Molecular & Cellular Biomechanics 2025, 22(3), 1325.
https://doi.org/10.62617/mcb1325

Article

Application of deep reinforcement learning under biomechanical load optimization in warehouse site selection and material transportation path

Xiaoming Zhang

School of Policing Equipment Technology, China People's Police University, Langfang 065000, China; ndzxm159357@163.com

CITATION

Zhang X. Application of deep reinforcement learning under biomechanical load optimization in warehouse site selection and material transportation path. Molecular & Cellular Biomechanics. 2025; 22(3): 1325.
https://doi.org/10.62617/mcb1325

ARTICLE INFO

Received: 8 January 2025
Accepted: 7 February 2025
Available online: 25 February 2025

Abstract: Path optimization of logistics and transportation systems has traditionally focused on the balance between efficiency and cost, but there is a lack of systematic research on the biomechanical load of transport personnel, which leads to fatigue accumulation and increased safety hazards. To fill this gap, this paper proposes a path optimization method based on deep reinforcement learning (DRL) based on biomechanical theory, aiming to combine biomechanical load management of transport personnel with logistics efficiency improvement. Firstly, a biomechanical load assessment system for transport personnel during long-distance driving is established using human kinematics and dynamics models, with quantitative indicators including muscle fatigue index, joint load and driving posture stability. Secondly, a national logistics transportation network is constructed based on a graph theory model, with transportation distance, time and biomechanical load as constraints for multi-objective optimization, and a Deep Q Network (DQN) is designed for path planning optimization. The calculation of fatigue index is combined with driving time, road section characteristics and individual biomechanical characteristics, and verified by biomechanical simulation tools. In order to improve the optimization efficiency, the simulated annealing algorithm is used to preliminarily screen the paths, and the DRL model is combined to achieve dynamic adjustment. The experimental results show that this method significantly reduces the biomechanical load of transport personnel in nationwide logistics scheduling (the fatigue index is controlled below 0.12), and at the same time reduces the accident rate caused by fatigue (reduced by 40%), and the transportation efficiency is superior to traditional research. The research results not only deepen the application of biomechanical theory in the field of long-distance transportation, but also provide theoretical support and technical reference for building a safe, efficient and intelligent logistics and transportation system, and promote the integrated development of biomechanics and artificial intelligence in complex engineering problems.

Keywords: deep reinforcement learning; biomechanical load; path optimization; fatigue index; logistics and transportation efficiency

1. Introduction

In modern logistics and transportation systems, route optimization [1–3] is considered to be the core of improving transportation efficiency and reducing transportation costs. However, traditional route optimization strategies usually focus on economic benefits and operational efficiency, and pay less attention to the health and safety of transportation personnel. Especially in long-distance transportation, the fatigue level of drivers has a significant impact on transportation safety and continuity. With the continuous growth of logistics demand and the frequent occurrence of related accidents, there is an urgent need for an optimization method that can take into account both biomechanical load and transportation efficiency, so as to improve overall benefits while ensuring safety.

In recent years, in the field of logistics scheduling, researchers have proposed a variety of models based on traditional optimization methods such as graph theory, genetic algorithms, and particle swarm optimization algorithms to improve transportation efficiency and reduce costs. Yucel et al. [4] used mixed integer linear models and adaptive large domain search algorithms to study a multi-period two-dimensional vehicle loading and scheduling problem with incompatibility constraints, and used real data examples to verify the effectiveness of the method. Lan et al. [5] studied the vehicle routing problem with time windows and proposed a decomposition-based multi-objective scheduling model and algorithm, which successfully solved the three-objective routing optimization problem and proved its effectiveness through experiments. Lei [6] used genetic algorithms as a solution algorithm in the intelligent transportation model decision-making in the logistics industry. Through a large number of experiments and data, it was shown that genetic algorithms can optimize the operation path and improve transportation efficiency. Cakmak et al. [7] employed the discrete binary particle swarm algorithm to address the issue in their research on mathematical model-based evaluation and multi-criteria methods for site selection. He observed that the algorithm reaches the optimal result more quickly as the test problem scale increases and confirmed the correctness of the optimal value through binary integer programming. Sihotang et al. [8] represented mining sites and processing plants as points in a graph and transportation routes as edges in a graph to formulate a path optimization problem. They used a graph-based optimization algorithm to reduce the transportation cost from 2100 units to 1700 units. However, these methods mainly focus on traffic flow and road conditions as core parameters, and pay less attention to the workload of transportation personnel, especially in fatigue management in long-term driving environments. Ignoring transport personnel load management will lead to accumulated fatigue, increased safety hazards, reduced transport efficiency, driver health problems, increased psychological stress and compliance risks, ultimately affecting the safety, efficiency and sustainability of the logistics system.

Due to its superior performance in dynamic decision-making, DRL has gradually become an important tool for logistics path optimization. Pan and Liu [9] proposed a novel DRL framework to solve the dynamic uncertain vehicle routing problem when studying real-time accurate tracking of real urban logistics, and proposed a cutting-edge reinforcement learning algorithm to control the value function in this problem. Zhang et al. [10] introduced a DRL-based hyper-heuristic framework to address the container terminal truck routing problem with uncertain service time. This framework improves the existing hyper-heuristic approach by leveraging powerful data-driven heuristic selection. Experiments show that this method has superior performance compared with existing solutions. In order to overcome the limitations of traditional research on vehicle routing problems, Zhao et al. [11] proposed a novel DRL model, combining the model with a local search algorithm to further improve the final quality of the solution. Liu et al. [12] proposed a reinforcement learning-based path planning solution for IoT drones, which plans hovering points for drones by learning the historical positions of cluster heads and uses simulated annealing to maximize the probability of encountering cluster heads, thereby planning the shortest path for drones to visit all hovering points. However, in existing studies, the introduction of biomechanical parameters is often one-sided and fails to be effectively integrated with the path optimization algorithm, which limits the effectiveness of the model in practical applications. Therefore, this paper adopts a multi-objective DRL method to solve the problem of insufficient consideration of the biomechanical load of transport personnel in existing studies.

Aiming at some problems existing in the transportation path in the current logistics transportation system on the basis of DRL [13–15], this paper studies the impact of biomechanical load [16,17] on transport personnel and establishes a biomechanical load assessment system. In this study, the graph theory [18,19] model was used to construct a logistics transportation network. The different components of the logistics transportation system were abstracted into nodes and edges in the network. The weight information of the edges included key information such as transportation distance and transportation time during logistics transportation. A multi-objective DQN [20–22] model was designed based on DRL. The model consists of two Q networks, which optimize the transportation timeliness and the biomechanical load of the transport personnel respectively, aiming to achieve the best balance between the two and design the optimal transportation route. At the same time, the simulated annealing algorithm [23–25] is used to preliminarily optimize the search space when designing the transportation route. The purpose is to optimize the transportation efficiency while reducing the fatigue of the transportation personnel. Finally, the designed multi-objective DQN model realizes the dynamic adjustment of the transportation route during transportation and verifies it using biosimulation software to cope with emergencies that may occur during the transportation process. Experiments have shown that the method used in this paper can be used to carry out logistics transportation across the country, with an average single transportation time of 32 h, and has achieved good control of the fatigue index of transportation personnel [26–28]. In the experiment, the fatigue index of transportation personnel was controlled below 0.12, which improved the safety of transportation personnel during transportation. In a certain area, the application of the method in this paper has alleviated the occurrence of accidents during transportation, especially accidents caused by fatigue, with the probability of occurrence reduced by 40%. This further demonstrates that the method presented in this paper enhances transportation safety while offering a practical solution for optimizing global logistics scheduling strategies. This contributes to the advancement of smart logistics systems [29,30] and supports broader industry applications.

2. Optimization path planning based on biomechanical load

2.1. Establishment of the biomechanical load assessment system

The construction of the biomechanical load assessment system is based on the kinematic and dynamic models of the human body and is used to quantify the biomechanical load that transportation personnel may encounter during long-distance driving. This system provides data support for driver load optimization in route planning by comprehensively considering factors such as muscle fatigue, joint load, and driving posture stability.

Fatigue index is a key indicator for evaluating driver fatigue, which is mainly calculated by factors such as muscle energy consumption, activity duration, and posture stability.

During long-term driving, the driver's muscles need to bear a certain load. The muscle model Hill's Muscle Model is used to calculate the muscle's power output and energy consumption during the duration. The calculation formula is:

$$F_{muscle} = \int_0^T (\tau(t) \cdot v(t)dt \tag{1}$$

Among them, $\tau(t)$ represents muscle force, $v(t)$ represents muscle contraction speed, and T is driving time. Through this formula, the total power consumption of muscles during driving can be obtained, and then the degree of fatigue can be calculated.

Joints are important load-bearing parts of the human body. Especially during long-term driving, the knee joints, waist and spine bear large mechanical and dynamic loads. The joint load model is used to evaluate the pressure and shear force of each joint. The evaluation formula is:

$$\sigma_{\text{joint}} = \frac{F_{\text{joint}}}{A_{contact}} \tag{2}$$

Among them, F_{joint} represents the force borne by the joint, and $A_{contact}$ represents the contact surface area of the joint.

The driver's sitting posture has an important impact on the load, especially the posture stability of the waist and spine. During driving, the driver needs to maintain a certain posture. If the posture is unstable, it will increase the burden on the muscles and lead to accelerated fatigue. The posture stability model is used to evaluate the driver's stability during driving. The calculation formula is:

$$Stability = \frac{1}{\int_0^T (\theta(t)^2)dt} \tag{3}$$

Among them, $\theta(t)$ represents the change in the driver's sitting angle.

The health status of each driver is different, which affects the fatigue accumulation of the driver under long-term work. By collecting the health records of the drivers, the personal health status coefficient of the driver is designed as:

$$H = 1 + \beta_1 \cdot WH + \beta_2 \cdot HH + \ldots + \beta_n \cdot OH \tag{4}$$

are represented by $\beta_1, \beta_2, \ldots, \beta_n$ *WH*, *HH*, and *OH*. If the driver is in poor health, his health factor *H* will be greater than 1, causing the fatigue index to increase.

The temperature and humidity of the in-car environment will affect the driver's comfort and physical exertion, and then affect the fatigue index. By collecting data on the temperature and humidity in the car through sensors, the influence coefficient of the in-car environment on the driver's fatigue index can be calculated:

$$E_{env} = \eta_1 \cdot T_{cat} + \eta_2 \cdot H_{car} \tag{5}$$

coefficient η_1, η_2 representing the influence of temperature and humidity inside the car on the driver's fatigue index.

Through the analysis of psychological assessment tools, there is a certain relationship between psychological load and physiological load. Higher psychological load will lead to increased physiological load, thereby accelerating the accumulation of fatigue. The calculation formula for the influence coefficient of psychological load is:

$$E_{psych} = \delta \cdot P \tag{6}$$

where δ represents the influence coefficient of psychological load on fatigue, and P is the driver's psychological load index evaluated by a psychological assessment tool.

of muscle fatigue, joint load, posture stability, personal health factor, in-vehicle environmental impact factor, and psychological load index is used to form a comprehensive fatigue index to describe the overall fatigue level of the driver:

$$f = w_1 \frac{F_{muscle}}{F_{\max}} + w_2 \frac{\sigma_{joint}}{\sigma_{\max}} + w_3 \frac{1}{Stability} + w_4 \cdot H + w_5 \cdot E_{env} + w_6 \cdot E_{psych} \tag{7}$$

Among them, $w_1, w_2, w_3, w_4, w_5, w_6$ the weight factors of F_{max} muscle fatigue, joint load, posture stability, personal health coefficient, in-vehicle environment impact coefficient and psychological load index represent the maximum bearing capacity of the muscles and σ_{max} the maximum bearing value of the joint load.

Based on the calculation of the above fatigue index, the evaluation results of the biomechanical load of the driver for each transportation route can be obtained as the basis for optimizing route selection.

2.2. Logistics and transportation network

When optimizing warehouse site selection and material transportation routes, the first thing to do is to build a logistics and transportation network. The network constructed in this paper relies on a graph theory model, abstracting the various components of the logistics system into nodes and edges in the graph. Each node represents a warehouse, logistics center or city, and each edge represents a transportation route. On the basis of this network, factors such as traffic flow, road conditions, and logistics demand are further considered to assign weights to each edge, providing a basis for subsequent path optimization and biomechanical load calculation.

In the logistics transportation network designed in this paper, each node represents a warehouse. Assuming there are N nodes, the node representation is:

$$V = \{v_1, v_2, \dots, v_N\} \tag{8}$$

Among them, v_i represents the i-th warehouse, and each warehouse v_i has specific attributes, such as location (latitude and longitude), storage capacity, transportation connection relationship with other nodes, etc.

The edge segment between two nodes in the network represents the transportation path between the two nodes. The edge between v_i and v_j is represented as (v_i, v_j), and the weight of the edge is represented as v_{ij}. This weight reflects the "cost" of the transportation path from v_i to v_j. When constructing the edges in the network, the following aspects can be considered:

(1) Transportation distance includes transportation distance, transportation mode, and the actual conditions of the transportation section.

(2) Time cost: the travel time on the path, which is limited by factors such as road conditions, traffic flow, and vehicle speed.

(3) Biomechanical load: The biomechanical load on the transport personnel during the transportation process, which mainly refers to the fatigue index of the transport personnel in this paper.

Therefore, the weight of each edge can be expressed as:

$$w_{ij} = \alpha_1 \cdot d_{ij} + \alpha_2 \cdot t_{ij} + \alpha_3 \cdot f_{ij} \tag{9}$$

Among them, d_{ij} is the transportation distance from node v_i to v_j, in km, t_{ij} is the transportation time from point v_i to v_j (unit: h), and f_{ij} is the fatigue index of the transport personnel from node v_i to v_j. α_1, α_2, α_3 are the weight coefficients corresponding to the three aspects, indicating the importance of the three in the total weight.

The transport distance d_{ij} is determined based on the latitude and longitude of geographical locations, with the Haversine formula [31,32] applied to compute the spherical distance between two network nodes:

$$d_{ij} = R \cdot \arccos(\sin\psi_i \sin\psi_j + \cos\psi_i \cos\psi_j \cos(\lambda_i - \lambda_j)) \tag{10}$$

ψ_i, λ_i and ψ_j, λ_j are the longitude and latitude of nodes v_i and v_j respectively, and R is the radius of the earth, which is 6371 km.

The derivation process of the Haversine formula is as follows:

(1) Convert longitude and latitude into spherical coordinate system.

(2) The central angle between two points on the sphere reflects the angle between the center of the sphere and the line connecting the two points. This angle is calculated by spherical trigonometry (click formula).

(3) Introduce the Haversine function to avoid the problem of numerical instability.

Finally, the great circle distance is obtained through the relationship between the spherical angle and the spherical radius, which is the required spherical distance.

The transportation time t_{ij} is affected by factors such as road conditions, traffic flow, and road speed limits. Assuming that the average speed of each edge is v_{avg}, the transportation time is calculated as follows:

$$t_{ij} = \frac{d_{ij}}{v_{avg}} \tag{11}$$

Once the transport distance, transport time and fatigue index are calculated, the edge weight w_{ij} can also be determined, and the entire logistics transport network is represented by a graph as follows:

$$G = (V, E) \tag{12}$$

V is the set of nodes, *E* is the set of edges between nodes, and each edge (v_i, v_j) has a corresponding weight w_{ij}. An example diagram of a logistics transportation network designed between seven nodes *A*, *B*, *C*, *D*, *E*, *F*, and *G* is shown in **Figure 1**.

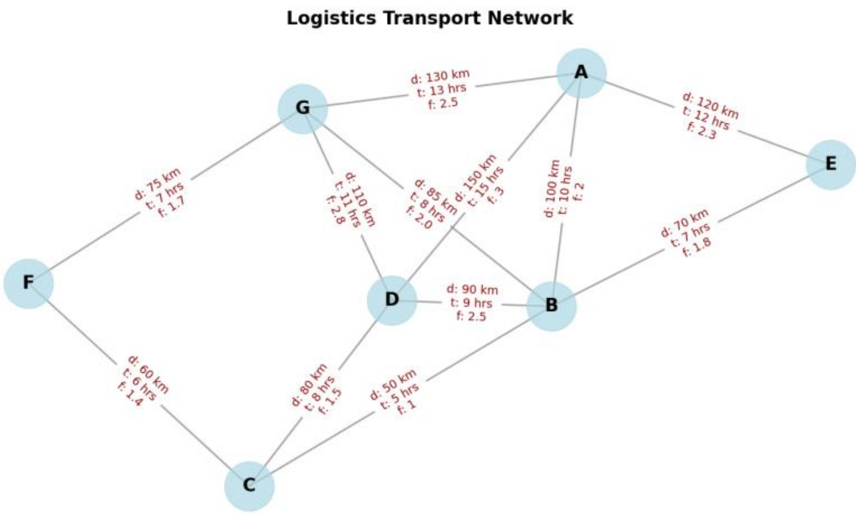

Figure 1. Example diagram of a transportation network.

2.3. Design of multi-objective DQN

DQN is based on a reinforcement learning framework, which learns how to choose the best actions to maximize the cumulative reward through interaction with the environment. In the multi-objective DQN designed in this paper, the timeliness of logistics transportation (minimizing transportation time) needs to be considered as a reward, and the biomechanical load of the transport personnel (minimizing the fatigue index) needs to be considered as an additional goal. The purpose of designing a multi-objective DQN is to optimize the timeliness of logistics and the biomechanical load of the transport personnel at the same time, achieving a balance between the two.

Compared with the traditional Q-network design that updates the expected cumulative reward for a state-action pair based on the Q value, this paper makes some adjustments and designs two independent Q-networks to estimate the timeliness of transportation and the biomechanical load of transportation personnel respectively. These two Q-networks are the Q-timeliness network and the Q-fatigue network. This method can more clearly separate the optimization process of different objectives, and deal with the two objectives of logistics timeliness and biomechanical load of transport personnel separately, avoiding their mutual interference in a single Q-network.

2.3.1. *Q* Timeliness Network

In the Q Timeliness Network, logistics timeliness is optimized specifically, focusing on how to choose the path to minimize the transportation time. Since the goal is to minimize the transportation time, the reward of the Q Timeliness Network is designed to be negatively correlated with the transportation time:

$$R_{efficiency} = -T_{path} \tag{13}$$

Among them, T_{path} is the transportation time under the current path, which is calculated using the transportation distance and average speed under the current path.

The Q value update formula of the Q time-sensitive network is:

$$Q_{efficiency}(s_t, a_t) = Q_{efficiency}(s_t, a_t) + \phi(R_{efficiency} + \gamma \max_{a_{t+1}} Q_{efficiency}(s_{t+1}, a_{t+1}) - Q_{efficiency}(s_t, a_t)) \tag{14}$$

s_t represents the state at the current time, a_t represents the action at the current time,

ϕ is the learning efficiency, γ is the discount factor, and $\max_{a_{t+1}} Q_{efficient}(s_{t+1}, a_{t+1})$ represents the maximum Q value corresponding to the action a_{t+1} selected in state s_{t+1}.

2.3.2. *Q* fatigue network

In the *Q*-fatigue network, the fatigue index of the transport personnel is optimized specifically, focusing on how to select the path to minimize the biomechanical load of the transport personnel. Since the goal is to minimize the fatigue index of the transport personnel, the reward of the *Q*-fatigue network is designed to be negatively correlated with the fatigue index:

$$R_{fatigue} = -F_{driver} \tag{15}$$

Among them, F_{driver} is the calculated fatigue index of the transport personnel, which is calculated based on the driving time and the complexity of the road section.

The *Q* value update formula of the *Q* fatigue network is:

$$Q_{fatigue}(s_t, a_t) = Q_{fatigue}(s_t, a_t) + \eta(R_{fatigue} + \kappa \max_{a_{t+1}} Q_{fatigue}(s_{t+1}, a_{t+1}) - Q_{fatigue}(s_t, a_t)) \tag{16}$$

Among them, η is the learning efficiency, κ is the discount factor, and $\max_{a_{t+1}} Q_{fatigue}(s_{t+1}, a_{t+1})$ represents the maximum *Q* value corresponding to the action a_{t+1} selected in state s_{t+1}.

2.3.3. Comprehensive reward and strategy selection

When calculating the comprehensive reward, an exponential function is used to weight the reward, making the multi-objective DQN more flexible in adjusting the weight of the target. The calculation formula is:

$$R_{composite} = f(\omega_1 \cdot R_{efficiency}) + f(\omega_2 \cdot R_{fatigue}) \tag{17}$$

$f(x)$ is an exponential function, and ω_1 and ω_2 are the weight ratios of the *Q*-timeliness network and the *Q*-fatigue network. The weight ratio is adaptively adjusted during the training process to enhance the influence of timeliness while suppressing the influence of the fatigue target.

When designing the strategy selection, the multi-objective optimization [33] strategy with constraints is used to constrain the fatigue index. The specific approach is to set an upper limit for the fatigue index to ensure that the fatigue index of the transport personnel can not exceed a certain safety threshold when selecting the route. The constraint conditions are defined as:

$$F_{driver} \leq F_{max} \tag{18}$$

F_{max} is the maximum tolerance value of fatigue. Considering that in some cases, timeliness may be a hard condition and fatigue index is a soft constraint, in the optimization process, priority is given to ensuring the completion of timeliness goals and reducing fatigue index while minimizing fatigue constraints.

2.3.4. Target network and experience replay

The target network serves as a delayed updated copy of the current network for *Q* value calculation. Using this target network to compute the *Q* value for the next state helps reduce overestimation and enhances training stability. Assuming two *Q* networks, the current network and the target network, the process for updating the target network is as follows:

A. Copy the parameters of the current network to the target network every several time steps:

$$\hat{Q}(s,a) \leftarrow Q(s,a) \tag{19}$$

$\hat{Q}(s,a)$ represents the target network and $Q(s,a)$ represents the current network.

B. The Q value in the Q network is computed using the target network, which updates the target value in the equation:

$$\hat{Q}(s_{t+1},a_{t+1}) = \max_{a_{t+1}} Q(s_{t+1},a_{t+1}) \tag{20}$$

Experience replay is a strategy that uses historical experiences of the Q network's interaction with the environment and randomly extracts these experiences for training. Whenever the Q network interacts with the environment, the four-tuple (s_t,a_t,r_t,s_{t+1}) of state, action, reward, and next state is stored in the experience replay pool, and then a small batch of samples are randomly extracted from the experience replay pool for training each time. Through experience replay, the time dependence in reinforcement learning can be effectively broken and the efficiency of learning can be improved. At the same time, reusing the data in the experience replay pool can enable the Q network to learn more fully.

The steps to implement replay are as follows:

1) Storing experience: Each time the Q network interacts with the environment, a four-tuple experience (s_t,a_t,r_t,s_{t+1}) is stored.

2) Randomly extracting samples: Small batches of data can be randomly extracted from the experience replay pool, and the number of samples extracted is fixed.

3) Q value update: The Q value function can be updated with randomly extracted samples, with the goal of minimizing the loss function, which is:

$$\Gamma(\theta) = E_{(s_t,a_t,r_t,s_{t+1})\sim D}[(r_t + (\gamma/\kappa)\max_{a_{t+1}}\hat{Q}(s_t,a_{t+1}) - Q(s_t,a_t))^2] \tag{21}$$

(γ/κ) is the discount factor of the Q-time network and the Q-fatigue network, and D is the experience replay pool.

The architecture of the multi-objective DQN model is shown in **Figure 2**.

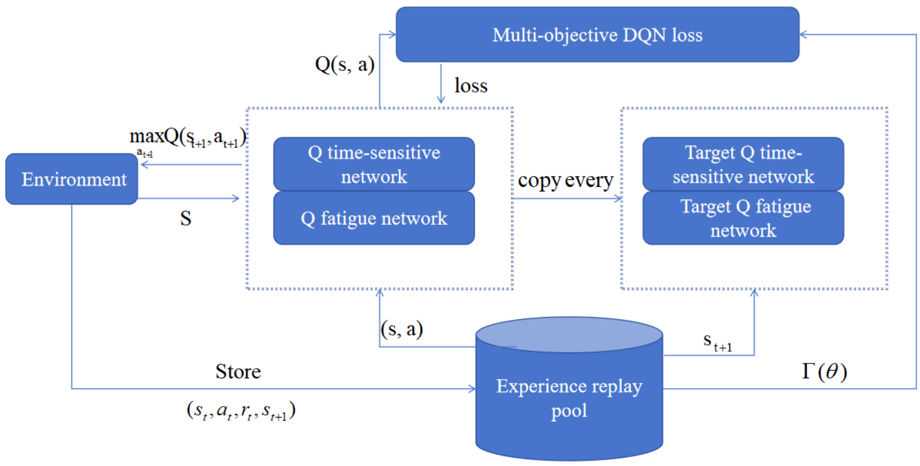

Figure 2. Multi-objective DQN architecture.

2.4. Preliminary optimization of simulated annealing algorithm

In this paper, the simulated annealing algorithm is used to perform preliminary optimization in the path search space, with the goal of reducing the fatigue burden of transport personnel while ensuring transportation efficiency. Specifically, simulated annealing regards the process of finding a path as a series of transformations and adjustments. Each transformation can lead to a slight change in the path. After a certain number of random searches and probability acceptance, a path that meets the requirements of multi-objective optimization is finally found.

The specific preliminary optimization steps are:

(1) Generate an initial solution: The first step of the simulated annealing algorithm involves creating an initial solution to serve as the starting point for the search. This paper employs a greedy algorithm to generate the initial warehouse and transportation path.

(2) Generate domain solutions: The core of the simulated annealing algorithm is to generate and evaluate domain solutions. Domain solutions are generated by making small perturbations to the current path. This paper designs three main perturbation operations. The first is the exchange operation, which randomly selects two transportation routes from the current path and swaps their order. For example, the current path is:

$$[A, B, C, D] \tag{22}$$

After swapping B and C, the new path can be got:

$$[A, C, B, D] \tag{23}$$

The second is the insertion operation, which randomly selects a warehouse from the current path and inserts it to another position in the path. For example, under the path of formula (16), selecting warehouse C and inserting it to the front of the path becomes:

$$[C, A, B, D] \tag{24}$$

The third is the reversal operation, which randomly selects a segment of the path and reverses the order of the segment. For example, under the path of formula (16), the reversed path of A, B, and C becomes:

$$[C, B, A, D] \tag{25}$$

(3) Energy calculation: In simulated annealing, "energy" measures solution quality. The energy of a path consists of two parts: logistics timeliness and transportation personnel fatigue index. Therefore, it is necessary to calculate the energy value of each field solution to evaluate the quality of the solution. The energy calculation formula is:

$$E = \varphi_1 T_{path} + \varphi_2 F_{driver} \tag{26}$$

Among them, E is the total energy of the current path, T_{path} is the total transportation time of the current path, F_{driver} is the fatigue index of the transport personnel, φ_1 and φ_2 are the weight coefficients of the transportation time and fatigue index, respectively.

(4) Acceptance criterion: The key to simulated annealing is the acceptance criterion, which is used to decide whether to accept a new domain solution. It is worth mentioning that the simulated annealing algorithm does not always accept only better solutions. It occasionally allows for the acceptance of worse solutions, which can jump out of the local optimal solution. The energy change between the domain solution and the current solution is calculated as:

$$\Delta E = E_{new} - E_{current} \tag{27}$$

E_{new} represents the objective function value of the domain solution, and $E_{current}$ represents the objective function value of the current solution. If $\Delta E \leq 0$, the domain solution is accepted. If $\Delta E > 0$, the probability of accepting the domain solution is determined by the temperature. The calculation formula for the acceptance probability is:

$$P = exp(\frac{-\Delta E}{T}) \tag{28}$$

Among them, P represents the probability of accepting a new solution, ΔE represents the energy difference, and T represents the current temperature. The temperature T gradually decreases during the execution of the algorithm.

For example, if the energy of the initial solution is $E(x_0) = 36$, x_{new} the energy of the new solution = 4 $E(x_{new}) = 25$ is, which is a decrease of 11, so this solution is accepted. However, if x_{new} the energy of the new solution = 5 is $E(x_{new}) = 38$, the energy increases instead, so this new solution is not accepted.

(5) Temperature decay: The search process of the simulated annealing algorithm shrinks the search space by gradually lowering the temperature. The initial temperature is high, and poor solutions can be accepted to explore more possibilities. As the temperature drops, the algorithm can progressively concentrate on the finer details of the optimal solution. After each iteration, the temperature decay is governed by the following formula:

$$T_{n+1} = \tau \cdot T_n \tag{29}$$

T_n is the temperature of the nth iteration, and τ is the temperature attenuation coefficient.

(6) Termination criteria: After 200 iterations, check if the output condition is satisfied. If it is, the algorithm stops. Otherwise, reduce the temperature and reset the iteration count to return to the solution generation phase.

The operation process of the simulated annealing algorithm is shown in **Figure 3**.

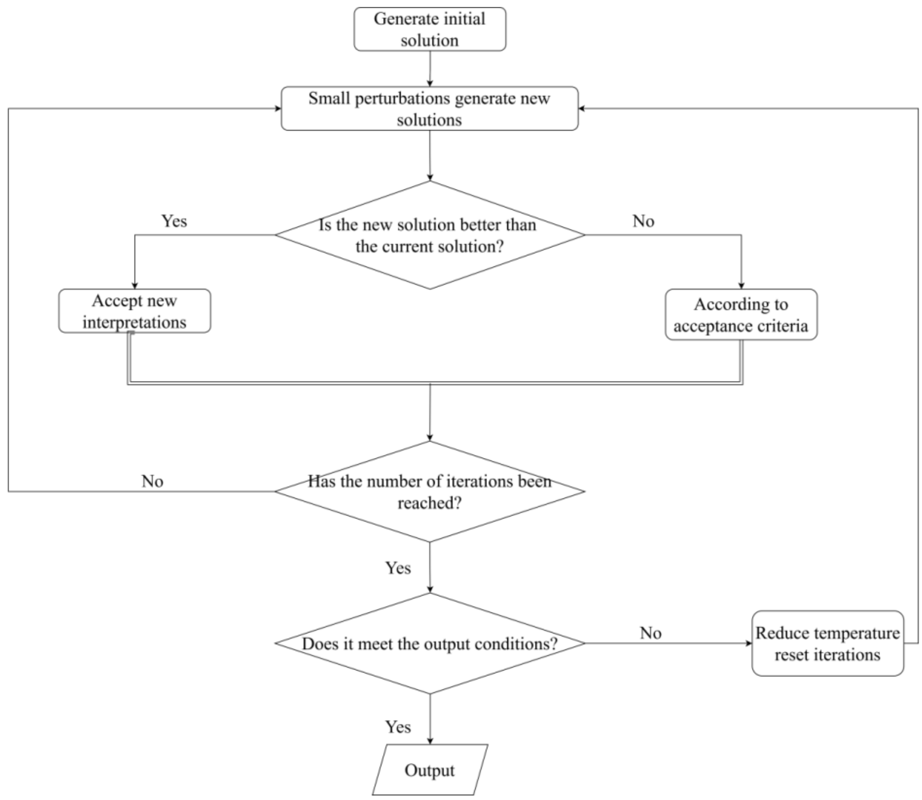

Figure 3. Simulated annealing algorithm flow.

2.5. Dynamic adjustment of transportation routes

During the logistics transportation process, emergencies (such as sudden traffic accidents, weather changes, equipment failures, etc.) may cause significant changes in the effectiveness and safety of the transportation route. In order to cope with these emergencies, this paper makes adjustments to emergencies based on the multi-objective DQN model.

First, in order to cope with emergencies, the state space update includes normal transportation information and adds emergencies. The state space of the model is adjusted to:

$$S_t = [f_t, t_t, c_t, r_t, T, \delta_{breakout}] \tag{30}$$

Among them, f_t represents the fatigue index of the transport personnel, t_t represents the current time, c_t represents the current traffic conditions, r_t represents the road section characteristics, T represents the remaining transport time, and $\delta_{breakout}$ represents the binary variable of the emergency, with a value of 1 if an emergency occurs and a value of 0 if there is no emergency.

Secondly, when an emergency occurs, the reward function needs to be dynamically adjusted according to the current environmental information. The reward ratio of the emergency is added to the comprehensive reward formula of formula (11). The adjusted formula is:

$$R_{total} = R_{composite} + \omega_3 \cdot v_{breakout} \quad (31)$$

Among them, ω_3 represents the weight coefficient of the emergency situation, and $v_{breakout}$ represents the emergency time. This additional term can avoid choosing unsafe or high-fatigue paths when an emergency occurs.

In emergencies, path selection must consider short-term rewards and also have long-term adaptability. Through a continuous learning process, multi-objective DQN continuously adjusts the path selection strategy based on feedback after emergencies. After each emergency, the model can be updated based on the real-time status and rewards to optimize future path selection. Through this feedback mechanism, multi-objective DQN can adaptively adjust the path strategy to ensure that each emergency can be reasonably responded to.

At the same time, in order to improve the applicability and practicality of the path planning model in different physical scenarios, the objective function and constraints in the optimization model are flexibly adjusted according to the needs of different physical tasks to meet the special requirements of different types of transportation tasks. Specifically, in the high-frequency and small-batch transportation scenarios of express logistics, the optimization objective function focuses on the shortest time and efficient distribution to meet the needs of fast delivery; in cold chain logistics, temperature control constraints and penalty mechanisms are added to ensure that the temperature during transportation always meets the prescribed standards to avoid quality risks. In addition, for special scenarios, such as the transportation of dangerous goods and bulk materials, the model can be customized according to safety, road carrying capacity and other characteristics. Through these flexible optimization adjustments, the model can achieve efficient, safe and practical path planning according to the characteristics of different logistics tasks.

The optimized path is verified using AnyBody Modeling System biomechanical simulation software. The verification steps are as follows:

(1) Modeling the driver's human body structure: Create a three-dimensional human body model of the driver, including bones, muscles, and joints.

(2) Setting the driver's scene and action: Set according to the driver's actual situation and the path optimization results designed in this paper to simulate the real environment.

(3) Applying external loads and forces: Define external load sources (such as road slope, vibration inside the car, etc.) in the software to simulate the driver's real driving environment.

(4) Calculate the biological load to verify the results and further verify the results.

3. Experimental design and result analysis

3.1. Experimental environment

The main hardware and software equipment used in this experiment are shown in **Table 1**.

Table 1. Main equipment.

category	Equipment	Specifications/Versions
Hardware	processor	AMD Ryzen 7 5800X
	Memory	16 GB DDR4
	storage	512 GB SSD
	Graphics Processing Unit	NVIDIA GeForce GTX 1660 Ti
	Network Interface	1000 Mbps Ethernet
Software Environment	Development Language	Python 3.10
	Main library	TensorFlow 2.6
	database	MySQL 8.0
	Programming Tools	Visual Studio Code
	Version Control Tools	Git 2.32
Experimental Tools	Simulated annealing algorithm implementation	SciPy 1.7
	Deep Learning Frameworks	PyTorch 1.9
	Optimization Tools	Scipy.optimize (for numerical optimization)

3.2. Transportation time analysis

Transportation time is one of the core indicators for measuring logistics efficiency. One of the main goals of optimizing routes and scheduling strategies is to reduce transportation time and improve the response speed of the overall logistics system. Shorter transportation time improves resource utilization efficiency and better meets customer needs and expectations.

Experimental design: The experiment uses a nationwide logistics transportation network dataset to simulate 10 transportation tasks, including the distribution of materials from warehouses to destinations. The control group uses traditional algorithms (Dijkstra [34–36], *A**) for shortest path planning, while the experimental group uses the research method in this paper to dynamically adjust the path, taking into account factors such as real-time traffic flow and road conditions. In the experiment, the same route can be tested multiple times and the transportation time of each transportation task can be recorded. Finally, the average transportation time (unit/h) of each transportation task can be calculated. The results are shown in **Table 2**.

Table 2 shows the transportation time results of 10 transportation task experiments conducted across China, comparing the optimization paths of the proposed method with those of the traditional Dijkstra and *A* algorithms. For the 10 different transportation tasks, the Dijkstra and *A* algorithms performed similarly, and because they failed to consider dynamic factors such as real-time traffic flow and road conditions, the transportation time fluctuated greatly. In contrast, the approach presented in this paper leverages dynamic data effectively and adjusts the path according to real-time conditions, thereby maintaining a small fluctuation in transportation time in different transportation tasks, and the average time for a single transportation is 32 h. This result shows that the method proposed in this paper can provide a more accurate and efficient path optimization strategy compared to the traditional Dijkstra and *A** algorithms in the face of complex and changing transportation environments, thereby improving the efficiency and reliability of logistics scheduling.

Table 2. Analysis of transportation time results.

Transport mission ID	This paper studies	Dijkstra	A^*
1	31.6	41.5	42.2
2	33.3	55.2	57.8
3	30.9	48.1	50.5
4	32.4	37.9	39.2
5	33.8	61.7	64.3
6	31.4	49.3	51.1
7	32.7	52.8	54.2
8	33.1	40.6	41.7
9	30.5	49.9	52.4
10	34.0	46.4	48.5

3.3. Fatigue of transport personnel

Fatigue driving is a major cause of traffic accidents. By continuously monitoring the fatigue index, potential accident risks can be effectively predicted and avoided, thereby reducing the probability of accidents. Proper management of fatigue conditions can improve transportation efficiency and reduce delays and accidents caused by driver fatigue. In addition, the analysis of fatigue index helps ensure compliance with safety regulations, reduce health risks, and improve employees' occupational health and job satisfaction.

Experimental design: In order to verify the effectiveness of this research method in controlling the fatigue index of transportation personnel, the experimental design uses this research method and traditional research methods (Dijkstra, A^*) for path planning for the same transportation task. In the experiment of this research method, both timeliness and fatigue index are optimized at the same time. In the traditional research method experiment, only timeliness is optimized without considering fatigue factors. Finally, the fatigue index of transport personnel in different methods is calculated, and a line graph of the change of fatigue index of transport personnel over time in the transport task is drawn, as shown in **Figure 4**.

Figure 4 is an analysis of the fatigue index of transport personnel using different methods for path planning in the transport task in the experiment. In the experiment, the fatigue index of the transport personnel at the beginning of the transport task was basically the same, and different methods for path planning in the same transport task had different transport times, which is shown in the legend. Analyzing the data in **Figure 4**, during the transportation process, as the transportation task continues, the fatigue index of the three methods increases. However, for the transportation personnel who use the research method in this paper for path planning, the fatigue index remains in a low range throughout the transportation process, always below 0.12. The transporter who used the research method in this paper to plan the route was the fastest among the three transporters to complete the transport task and was also the only one whose fatigue index decreased in the later period. This shows that the optimization of biomechanical load in route planning using the research method in this paper can ensure the optimization of the transport task while effectively reducing the fatigue of the transporter.

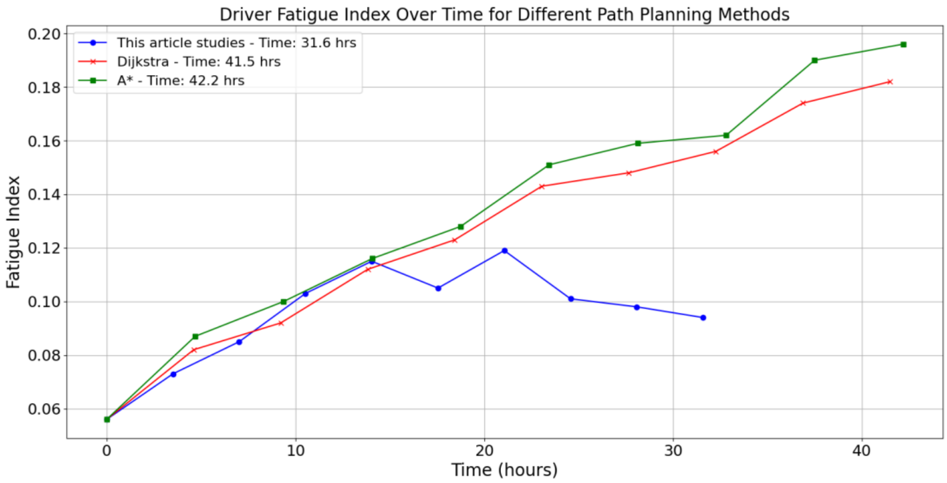

Figure 4. Analysis of the fatigue index of transport personnel.

3.4. Accident analysis

Analyzing accidents during transportation is the key to verifying the effectiveness of the optimization method in this paper. Long driving hours can cause fatigue in transport personnel and increase the risk of accidents. By optimizing routes and controlling fatigue index, traffic accidents caused by fatigue can be effectively reduced. The analysis of accident rates not only helps to evaluate the effectiveness of fatigue control and ensure the safety of transport personnel, but also verifies the comprehensive effect of multi-objective optimization models in balancing logistics efficiency and safety. In addition, accident analysis provides feedback for further improving optimization strategies, enhancing transportation safety, and promoting the logistics industry to pay attention to the health and safety management of transportation personnel.

Experimental design: In order to analyze the occurrence of accidents before and after the application of the research method in this paper, an experiment was conducted in a certain area. First, the accident data of the area one month ago was collected, and then the accident data of the month was collected one month after the method in this paper was applied. The two sets of data were compared and analyzed, and a bar chart was drawn as shown in **Figure 5**.

Figure 5 is an image showing the accident situation before and after the application of the research method in this paper. It mainly conducts statistical analysis on five types of accidents, namely, accidents caused by fatigue driving, traffic violations, accidents caused by bad weather, accidents caused by overloading, and accidents caused by problems with road facilities. The number of accidents involving fatigue driving dropped from 120 to 72, the number of accidents involving traffic violations dropped from 76 to 62, the number of accidents caused by bad weather dropped from 63 to 54, the number of accidents caused by overloading dropped from 47 to 41, and the number of accidents caused by road facility problems dropped from 54 to 39. Overall, the number of accidents has decreased after applying the research method in this paper, indicating that the combination of path optimization based on deep reinforcement learning and biomechanical load management in this paper can effectively reduce traffic accidents caused by factors such as driver fatigue, speeding, and bad weather. At the same time, through the refined management of the driver's biomechanical load, it is shown that the optimization method studied in this paper is not aimed at a single factor, but improves the safety of the transportation process from multiple aspects. In particular, the accident rate caused by fatigue driving has decreased by 40%, reflecting the significant advantages of fatigue control in the transportation process studied in this paper.

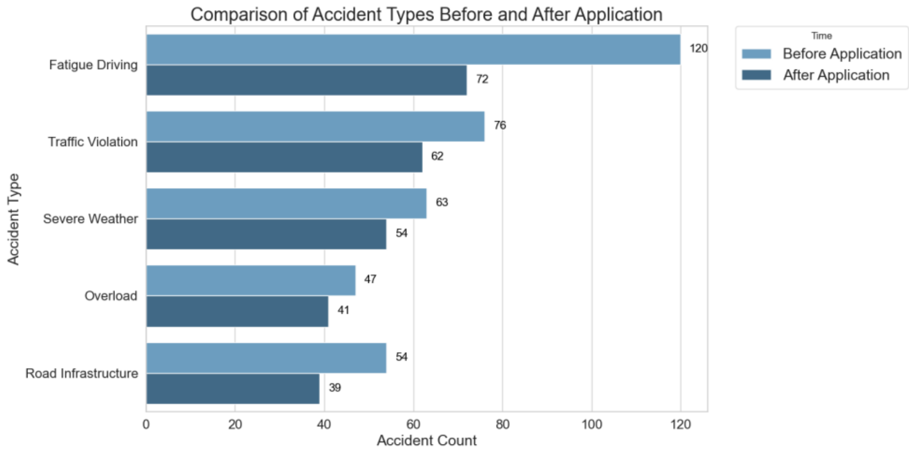

Figure 5. Accident analysis before and after application.

4. Conclusions

This paper studies the problem of biomechanical load of transport personnel in the current logistics system, which is neglected. A DRL-based method is used to fully consider the biomechanical load of transport personnel in the transportation process, and a biomechanical load evaluation system is established to focus on analyzing the fatigue indicators of transport personnel. This study introduces a graph theory model to model nationwide logistics transportation, and designs a multi-objective DQN network model to optimize the timeliness of logistics transportation and the biomechanical load of transportation personnel. When optimizing the transport path, the simulated annealing algorithm can be used to preliminarily optimize the path search space, realize dynamic adjustment of the path during the transport process, and use biosimulation software to verify the results. Experiments have proved that:

(1) The research method in this paper can provide accurate and efficient transportation path solutions and realize efficient and reliable logistics scheduling.

(2) When optimizing the transportation path, the research method in this paper can not only ensure the timeliness of logistics transportation, but also effectively reduce the fatigue of transportation personnel.

(3) The application of the research method in this paper can effectively reduce the occurrence of accidents in the logistics transportation process, especially accidents caused by fatigue.

Although this study provides an effective solution for path optimization combined with biomechanical load in logistics dispatch, there are still some shortcomings:

1) In the study, the fatigue of transport personnel was quantified through factors such as transportation time and road section characteristics. This quantification method ignores the dynamic changes of the actual fatigue state of transport personnel to a certain extent. In fact, there may be more factors affecting the fatigue state of transport personnel. These factors may include individual differences of drivers, such as physical condition and driving skills, and psychological factors, such as the driver's work pressure and emotional state, etc. These factors may not be quantifiable.

2) Although this paper has conducted some experiments on the application of the research method, different regions and different logistics and transportation systems may face different challenges and constraints, such as differences in traffic regulations in different regions and uneven distribution of logistics demand, and this study has not yet analyzed more situations.

In view of the above-mentioned shortcomings, future research and optimization can be carried out from the following aspects:

1) Future research directions should focus on individual differences of drivers and more refined psychological factors, and analyze multi-dimensional data such as biomechanical characteristics, driver's health status, driving level, work pressure and driver's emotions. Real-time monitoring and quantification of driver fatigue can be carried out through personalized dynamic models, thereby improving the accuracy of driver fatigue assessment.

2) Future research needs to focus on analyzing the differences in different traffic regulations, road conditions and logistics needs based on multi-regional experiments, and building a more adaptable path optimization model to ensure that the optimization strategy in different environments can effectively reduce driver fatigue and improve transportation efficiency.

Funding: Funded by Science Research Project of Hebei Education Department: Research on design and coordination optimization of the "Beijing-Tianjin-Hebei" integrated police epidemic prevention equipment support system (ZC2023011).

Ethical approval: Not applicable.

Data availability: The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflict of interest: The author declares no conflict of interest.

References

1. Leng K, Li S. 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
2. Yu H, Murray AT, Fang Z, et al. Ship Path Optimization That Accounts for Geographical Traffic Characteristics to Increase Maritime Port Safety. IEEE Transactions on Intelligent Transportation Systems. 2022; 23(6): 5765-5776. doi: 10.1109/tits.2021.3057907
3. Zhang D, Luo R, Yin Y bo, et al. Multi-objective path planning for mobile robot in nuclear accident environment based on improved ant colony optimization with modified A∗. Nuclear Engineering and Technology. 2023; 55(5): 1838-1854. doi: 10.1016/j.net.2023.02.005

4. Yucel E, Salman FS, Erdoğan G. Optimizing two-dimensional vehicle loading and dispatching decisions in freight logistics. European Journal of Operational Research. 2022; 302(3): 954-969. doi: 10.1016/j.ejor.2022.01.021
5. Lan YL, Liu F, Ng WWY, et al. Decomposition Based Multi-Objective Variable Neighborhood Descent Algorithm for Logistics Dispatching. IEEE Transactions on Emerging Topics in Computational Intelligence. 2021; 5(5): 826-839. doi: 10.1109/tetci.2020.3002228
6. Lei N. Intelligent logistics scheduling model and algorithm based on Internet of Things technology. Alexandria Engineering Journal. 2022; 61(1): 893-903. doi: 10.1016/j.aej.2021.04.075
7. Cakmak E, Önden İ, Acar AZ, et al. Analyzing the location of city logistics centers in Istanbul by integrating Geographic Information Systems with Binary Particle Swarm Optimization algorithm. Case Studies on Transport Policy. 2021; 9(1): 59-67. doi: 10.1016/j.cstp.2020.07.004
8. Sihotang HT, Riandari F, Sihotang J. Graph-based Exploration for Mining and Optimization of Yields (GEMOY Method). Jurnal Teknik Informatika CIT Medicom. 2024; 16(2): 70-81. doi: 10.35335/cit.vol16.2024.777.pp70-81
9. Pan W, Liu SQ. Deep reinforcement learning for the dynamic and uncertain vehicle routing problem. Applied Intelligence. 2022; 53(1): 405-422. doi: 10.1007/s10489-022-03456-w
10. Zhang Y, Bai R, Qu R, et al. A deep reinforcement learning based hyper-heuristic for combinatorial optimisation with uncertainties. European Journal of Operational Research. 2022; 300(2): 418-427. doi: 10.1016/j.ejor.2021.10.032
11. Zhao J, Mao M, Zhao X, et al. A Hybrid of Deep Reinforcement Learning and Local Search for the Vehicle Routing Problems. IEEE Transactions on Intelligent Transportation Systems. 2021; 22(11): 7208-7218. doi: 10.1109/tits.2020.3003163
12. Liu R, Qu Z, Huang G, et al. DRL-UTPS: DRL-Based Trajectory Planning for Unmanned Aerial Vehicles for Data Collection in Dynamic IoT Network. IEEE Transactions on Intelligent Vehicles. 2023; 8(2): 1204-1218. doi: 10.1109/tiv.2022.3213703
13. Wang X, Wang S, Liang X, et al. Deep Reinforcement Learning: A Survey. IEEE Transactions on Neural Networks and Learning Systems. 2024; 35(4): 5064-5078. doi: 10.1109/tnnls.2022.3207346
14. Wang H, Liu N, Zhang Y, et al. Deep reinforcement learning: a survey. Frontiers of Information Technology & Electronic Engineering. 2020; 21(12): 1726-1744. doi: 10.1631/fitee.1900533
15. Gronauer S, Diepold K. Multi-agent deep reinforcement learning: a survey. Artificial Intelligence Review. 2021; 55(2): 895-943. doi: 10.1007/s10462-021-09996-w
16. Vinstrup J, Jakobsen MD, Madeleine P, et al. Biomechanical load during patient transfer with assistive devices: Cross-sectional study. Ergonomics. 2020; 63(9): 1164-1174. doi: 10.1080/00140139.2020.1764113
17. Verheul J, Nedergaard NJ, Pogson M, et al. Biomechanical loading during running: can a two mass-spring-damper model be used to evaluate ground reaction forces for high-intensity tasks?. Sports Biomechanics. 2019; 20(5): 571-582. doi: 10.1080/14763141.2019.1584238
18. Ali SS, Kaur R, Khan S. Identification of innovative technology enablers and drone technology determinants adoption: a graph theory matrix analysis framework. Operations Management Research. 2023; 16(2): 830-852. doi: 10.1007/s12063-023-00346-3
19. Ekhlasi A, Nasrabadi AM, Mohammadi M. Analysis of EEG brain connectivity of children with ADHD using graph theory and directional information transfer. Biomedical Engineering/Biomedizinische Technik. 2022; 68(2): 133-146. doi: 10.1515/bmt-2022-0100
20. Oroojlooyjadid A, Nazari M, Snyder LV, et al. A Deep Q-Network for the Beer Game: Deep Reinforcement Learning for Inventory Optimization. Manufacturing & Service Operations Management. 2022; 24(1): 285-304. doi: 10.1287/msom.2020.0939
21. Tao X, Hafid AS. DeepSensing: A Novel Mobile Crowdsensing Framework With Double Deep Q-Network and Prioritized Experience Replay. IEEE Internet of Things Journal. 2020; 7(12): 11547-11558. doi: 10.1109/jiot.2020.3022611
22. Talaat FM. Effective deep Q-networks (EDQN) strategy for resource allocation based on optimized reinforcement learning algorithm. Multimedia Tools and Applications. 2022; 81(28): 39945-39961. doi: 10.1007/s11042-022-13000-0
23. Abdel-Basset M, Ding W, El-Shahat D. A hybrid Harris Hawks optimization algorithm with simulated annealing for feature selection. Artificial Intelligence Review. 2020; 54(1): 593-637. doi: 10.1007/s10462-020-09860-3
24. Fontes DBMM, Homayouni SM, Gonçalves JF. A hybrid particle swarm optimization and simulated annealing algorithm for the job shop scheduling problem with transport resources. European Journal of Operational Research. 2023; 306(3): 1140-

1157. doi: 10.1016/j.ejor.2022.09.006
25. Shi K, Wu Z, Jiang B, et al. Dynamic path planning of mobile robot based on improved simulated annealing algorithm. Journal of the Franklin Institute. 2023; 360(6): 4378-4398. doi: 10.1016/j.jfranklin.2023.01.033
26. Wang YD, Underwood BS, Kim YR. Development of a fatigue index parameter, Sapp, for asphalt mixes using viscoelastic continuum damage theory. International Journal of Pavement Engineering. 2020; 23(2): 438-452. doi: 10.1080/10298436.2020.1751844
27. Bonetti A, Bonetti L, Čipčić O. Self-Assessment of Vocal Fatigue in Muscle Tension Dysphonia and Vocal Nodules: A Preliminary Analysis of the Discriminatory Potential of the Croatian Adaptation of the Vocal Fatigue Index (VFI-C). Journal of Voice. 2021; 35(2): 325.e1-325.e15. doi: 10.1016/j.jvoice.2019.08.028
28. Küçükakgün H, Tulek Z, Kılıçaslan K, et al. Validation of the Turkish version of the Neurological Fatigue Index for Stroke. Cognitive Neuropsychiatry. 2024; 29(2): 141-153. doi: 10.1080/13546805.2024.2337155
29. Ding Y, Jin M, Li S, et al. Smart logistics based on the internet of things technology: an overview. International Journal of Logistics Research and Applications. 2020; 24(4): 323-345. doi: 10.1080/13675567.2020.1757053
30. Cimini C, Lagorio A, Romero D, et al. Smart Logistics and The Logistics Operator 4.0. IFAC-PapersOnLine. 2020; 53(2): 10615-10620. doi: 10.1016/j.ifacol.2020.12.2818
31. Mughni MD, Putri ARD. Implementation Of Teacher Presence System Using Mobile-Based Geofencing & Haversine Formula Methods. Applied Technology and Computing Science Journal. 2023; 6(1): 31-40. doi: 10.33086/atcsj.v6i1.4119
32. Mahatmi MF, Hasanuddin T, Umar F. Implementasi Metode Haversine Formula Untuk Menentukan Jarak Terdekat Pada Pengantaran Air Galon Depot Anantama Berbasis Android. Buletin Sistem Informasi dan Teknologi Islam. 2022; 3(1): 69-78. doi: 10.33096/busiti.v3i1.1098
33. Sharma S, Kumar V. A Comprehensive Review on Multi-objective Optimization Techniques: Past, Present and Future. Archives of Computational Methods in Engineering. 2022; 29(7): 5605-5633. doi: 10.1007/s11831-022-09778-9
34. Sari IP, Fahroza M, Fahri MM. et al. Implementation of Dijkstra's Algorithm to Determine the Shortest Route in a City. Journal of Computer Science, Information Technology and Telecommunication Engineering. 2021.
35. Wayahdi MR, Ginting SHN, Syahputra D. Greedy, A-Star, and Dijkstra's Algorithms in Finding Shortest Path. International Journal of Advances in Data and Information Systems. 2021; 2(1): 45-52. doi: 10.25008/ijadis.v2i1.1206
36. Akram M, Habib A, Alcantud JCR. An optimization study based on Dijkstra algorithm for a network with trapezoidal picture fuzzy numbers. Neural Computing and Applications. 2020; 33(4): 1329-1342. doi: 10.1007/s00521-020-05034-y