# Linkage Mechanism of Employee Performance Optimization and Functional Textile Quality Control Based on Improved DDQN

**Chunsun He[1], Lei Wang[2]**
[1]School of Urban Governance and Public Affairs, Suzhou City University, Suzhou 215104, Jiangsu, China
[2]Xianda College of Economics and Humanities, Shanghai International Studies University, Shanghai 200083, China
Corresponding author: Lei Wang (e-mail: reneewl@163.com)

**ABSTRACT** To address the challenge of jointly optimizing employee performance and functional textile quality in intelligent manufacturing environments, this study proposes a multi-objective collaborative optimization framework based on an improved Double Deep Q-Network (DDQN). A multidimensional state-space model is constructed by integrating employee behavioral indicators, production process parameters, and real-time quality inspection data, enabling coupled representation of human–machine interactions and manufacturing performance. The proposed method incorporates a dynamic reward function, prioritized experience replay, and multi-step temporal difference prediction to improve policy learning efficiency and long-term decision stability. Considering that modern textile manufacturing increasingly relies on heterogeneous sensing devices and distributed industrial information acquisition, the proposed state modeling and decision-making framework provides a scalable architecture for integrating multimodal production data and supporting intelligent monitoring systems. Experimental evaluations demonstrate that the improved DDQN reduces the average defect rate to 2.6% with a standard deviation of 0.4 while achieving a comprehensive performance score of 89.4, significantly outperforming conventional DDQN and static scheduling strategies. The framework effectively balances production efficiency, operational consistency, and quality stability, offering a practical solution for intelligent textile manufacturing and providing methodological insights for future interdisciplinary research on advanced sensing, industrial information fusion, and smart manufacturing infrastructures.

**INDEX TERMS** improved DDQN, multi-objective optimization, textile quality control, intelligent manufacturing, industrial sensing

## I. INTRODUCTION

In modern textile production, employee performance and product quality control are two core factors that ensure production efficiency and product pass rates [1,2]. However, in actual production environments, there is generally a lack of an effective linkage mechanism between the two. Operational efficiency and quality control often restrain or even conflict with each other, seriously affecting production stability and consistency. Especially in functional textiles, complex processes and dynamic requirements further magnify this contradiction, highlighting the urgent need for a collaborative mechanism integrating employee behavior–process parameters–quality feedback.

In recent years, deep reinforcement learning has shown great potential in solving such multi-objective optimization problems [3,4]. Double Deep Q-Network (DDQN) has advantages in dealing with dynamic and complex decision-making environments and can automatically learn optimal strategies to deal with uncertainty [5,6]. However, standard DDQN still cannot meet the needs of linkage optimization. Its static reward function and state modeling method cannot reflect the dynamic coupling relationship between employee behavior, process parameters, and quality indicators in real time [7,8]. At the same time, sampling efficiency and Q-value estimation stability are also insufficient [9,10].

In response to these problems, this paper proposes an improved DDQN method. This study constructs a multidimensional state-space model that integrates employee behavior, process parameters, and quality indicators to ensure comprehensive information coverage. Secondly, a dynamic weighted reward function is designed to adaptively adjust the strategy orientation based on performance and quality feedback. A prioritized experience replay mechanism is applied to increase the learning frequency of key experience samples, and a multi-step prediction method is combined to optimize long-term return estimation.

## II. RELATED WORK

In textile manufacturing, employee performance and product quality control are two core dimensions. The lack of a linkage mechanism between them often becomes an important bottleneck restricting production efficiency and product consistency [11]. Traditional performance management mainly relies on static indicator evaluation and manual intervention, and it is difficult to perceive the dynamic relationship between operational behavior and output quality in real time [12,13]. Similarly, in large-scale weaving processes, traditional quality control methods are also challenged to capture rapidly changing production fluctuations, resulting in feedback lags that further exacerbate the disconnect between operation and quality [14,15]. To address this problem, some studies have adopted multi-objective optimization methods and achieved initial results in manufacturing scheduling and quality control [16,17]. With the development of machine learning and computer vision technology, intelligent algorithms have been widely used in automated quality inspection through image recognition and data analysis, which can provide real-time feedback on textile quality status and improve quality control levels [18,19]. However, most intelligent quality control methods still focus on optimizing a single quality target, ignoring the multi-factor synergy between employee performance and quality targets [20].

As an important branch of reinforcement learning, DDQN improves strategy stability through a dual network structure and has been applied in industrial scenarios such as process optimization and equipment maintenance [21]. However, in textile tasks, DDQN struggles to reflect the dynamic trade-off between performance and quality, and it easily overlooks high-frequency variables such as employee behavior [22]. The sampling mechanism does not respond to sudden anomalies in a timely manner. Some existing studies focus on equipment parameter control and lack support for multi-objective coupling modeling in human– machine hybrid systems [23].

Therefore, this paper proposes a multi-objective linkage optimization method based on improved DDQN, aiming to address the performance–quality collaboration problem in textile manufacturing. It constructs a multidimensional state space that integrates employee behavior, process parameters, and quality indicators, and applies a dynamic weighted reward mechanism and a multi-step prediction strategy, improving the modeling ability in complex dynamic scenarios. Compared with existing work, this paper pays greater attention to modeling the linkage relationships between objectives and the mechanisms for real-time strategy adjustment, providing a feasible path for integrating intelligent performance management and quality control.

## III. CONSTRUCTION OF COLLABORATIVE OPTIMIZATION MECHANISM FOR EMPLOYEE PERFORMANCE AND TEXTILE QUALITY CONTROL

### *MULTIDIMENSIONAL STATE-SPACE MODELING*

The state space constructed in this study enables intelligent agents to fully perceive the textile production process in a reinforcement learning environment. It integrates employee behavior data, process parameters, and product quality control information. The state vector is represented as a fixed-length continuous numerical vector. The dimensional design is based on the key variables that have a significant impact on performance in the production process. Each dimension in the vector corresponds to a standardized continuous observation value.

The employee behavior indicator collection module integrates sensors and the Manufacturing Execution System (MES) data interface [24,25] to extract features including the number of tasks completed per unit time, consistency scores of operation steps, frequency of abnormal processes, and frequency of manual intervention requirements. All data are uploaded to the data center at fixed intervals, and the Z-score standardization method is used to normalize each indicator. The behavior data are weighted before being embedded in the state-space vector:

$$s_{\mathrm{emp},t} = \sum_{i=1}^{n} w_i \cdot \frac{b_{i,t} - \mu_i}{\sigma_i} \tag{1}$$

where $n$ is the total number of indicators; $b_{i,t}$ is the value of the $i$-th employee behavior indicator at time step $t$; $\mu_i$ and $\sigma_i$ are the mean and standard deviation of the indicator, respectively; $w_i$ is the weight of the $i$-th indicator.

The process parameter data come from the Programmable Logic Controller (PLC) control system and the Supervisory Control and Data Acquisition (SCADA) platform [26,27], which collect continuous variables such as spinning tension, deviation between the set value and the actual value of temperature and humidity control, equipment vibration frequency, and knitting speed in real time. To enhance the temporal expression ability of the state, a sliding window mechanism is applied to perform sequence interception on each variable, and the window length is set to five decision intervals (equivalent to a 50-second historical window) to form a temporal feature vector of the process parameters. After embedding this part of the feature, a 1D convolutional coding module is used for dimensional compression, and the output result is concatenated to the state vector.

Product quality control indicators are obtained through the quality inspection system and the visual recognition module, including weft skew error, yarn breakage rate, etc. Data preprocessing includes defect label normalization, confidence-weighted aggregation, and outlier processing. After removing abnormal samples identified by a Gaussian Mixture Model (GMM)-based outlier detection method [28], Principal Component Analysis (PCA) is applied to reduce the dimensionality to the range allowed by the state space.

The final state vector consists of three parts: the standardized employee behavior feature subvector, the convolutionally encoded process parameter time-series feature subvector, and the reduced-dimension vector of the quality control index. The detailed data source, feature type, processing method, and embedding method of the multidimensional state space are shown in Table 1, which presents the key indicators of each module and the corresponding preprocessing and feature encoding strategies.

**TABLE 1.** Multidimensional state-space feature composition and processing

| Module | Data Source | Feature Examples | Processing | Embedding Dimension |
|---|---|---|---|---|
| Employee Behavior | MES, sensors | Task count, consistency, anomaly frequency | Z-score normalization, weighted fusion | Direct concatenation, 32 |
| Process Parameters | PLC, SCADA | Tension, temp/humidity diff., vibration, knitting speed | Sliding window (5 steps), 1D conv. reduction | Conv. output concatenation, 48 |
| Quality Control | Vision system, QA database | Warp/weft error, break rate, surface defects, color match | GMM outlier removal, confidence aggregation, PCA | PCA output concatenation, 48 |
| Total State Vector | — | — | — | Concatenated for network input, 128 |

The entire state-space dimension is set to 128 to ensure dimensional compatibility and stable inputs for the DDQN framework. The reinforcement learning decision interval (also referred to as the agent's time step) is set to 10 seconds, and a buffer mechanism is combined to avoid policy fluctuations caused by state jitter. The state representation structure is mapped by a fully connected network at the front end of the neural network to enhance the nonlinear ability of feature expression.

The dimensional design of the state vector is based on the characteristic requirements of the textile production system. The 32 dimensions of employee behavior characteristics cover core indicators such as operational efficiency, standardization, and exception handling, which are key behavioral dimensions commonly monitored in MES systems. The 48 dimensions of process parameters are determined by the control variables of the spinning/weaving equipment, including the measurement requirements of continuous parameters such as tension and temperature at multiple process nodes. The 48 dimensions of quality control correspond to the main categories of textile defect detection and can fully characterize the morphology and distribution characteristics of fabric defects. The design of the 1D convolutional layer targets the temporal dependency of process parameters, and its hierarchical structure can effectively extract the local fluctuation patterns of the equipment signal.

### DYNAMIC REWARD FUNCTION DESIGN

The reward function adopts a dual-objective coupled structure [29], consisting of a performance sub-reward and a quality sub-reward, with dynamically adjusted weights to achieve real-time collaborative optimization of employee operational efficiency and textile quality. The overall reward is generated by combining the two submodules, each of which quantifies feedback from an independent data source and adjusts the strategy orientation according to actual production. The overall reward function is formalized as:

$$R_t = \omega_t^{(p)} \cdot R_t^{(p)} + \omega_t^{(q)} \cdot R_t^{(q)}, \quad \omega_t^{(p)} + \omega_t^{(q)} = 1 \tag{2}$$

$R_t$ represents the total reward at time step $t$; $R_t^{(p)}$ is the performance submodule reward; $R_t^{(q)}$ is the quality submodule reward; $\omega_t^{(p)}$ and $\omega_t^{(q)}$ represent the weights of the two, which are dynamically adjusted according to feedback in each round.

The reward value of the performance submodule is constructed based on the task completion rate $\hat{\eta}_t$ per unit time, the operation consistency score $\hat{\psi}_t$, and the number of manual interventions $\hat{\delta}_t$:

$$R_t^{(p)} = \alpha_1 \cdot \hat{\eta}_t + \alpha_2 \cdot \hat{\psi}_t + \alpha_3 \cdot (1 - \hat{\delta}_t), \quad \alpha_1 + \alpha_2 + \alpha_3 = 1 \tag{3}$$

$\alpha_1$, $\alpha_2$, $\alpha_3$ are the weight coefficients of the three standardized performance indicators, which are used to adjust the relative importance of task completion rate, operation consistency, and manual intervention frequency in the comprehensive reward.

The task completion rate is derived from the ratio of the output quantity to the target capacity within the time period recorded by the production execution system; the operation consistency is calculated by the consistency score obtained by comparing the operation trajectory with the standard operation process; the frequency of manual intervention is reflected by the statistical values of abnormal operation, equipment alarms, and records that require manual confirmation. The above three indicators are standardized and weighted to form instantaneous performance feedback at the employee behavior level.

The reward value of the quality submodule is composed of three dimensions: product defect changes $\hat{\phi}_t$, batch consistency fluctuations $\hat{\sigma}_t$, and final pass rate $\hat{\gamma}_t$:

$$R_t^{(q)} = \beta_1 \cdot (1 - \hat{\phi}_t) + \beta_2 \cdot (1 - \hat{\sigma}_t) + \beta_3 \cdot \hat{\gamma}_t, \quad \beta_1 + \beta_2 + \beta_3 = 1 \tag{4}$$

$\beta_1$, $\beta_2$, $\beta_3$ are the weight coefficients corresponding to quality-related indicators, which act on the defect rate change, batch consistency fluctuation, and batch qualification rate, respectively.

Defect changes are evaluated based on the defect type and frequency changes recorded by the online visual inspection system; consistency fluctuations use the standard deviation of thickness, color difference, and weaving density data collected in real time to evaluate product stability; the pass rate is calculated by the proportion of qualified products in the current batch and historical cycle output by the quality inspection system. All quality feedback data are compressed to a unified dimension through the distribution normalization method to form a quality control sub-reward signal.

The dynamic weight mechanism of the reward function is designed to achieve policy-guided adaptive adjustment. When the performance improvement rate in the current task cycle is greater than the quality improvement rate, the

$$\Delta_t^{(p)} = \frac{R_t^{(p)} - R_{t-1}^{(p)}}{R_{t-1}^{(p)}} > \Delta_t^{(q)} = \frac{R_t^{(q)} - R_{t-1}^{(q)}}{R_{t-1}^{(q)}} \tag{5}$$

system reduces $\omega_t^{(p)}$ and increases $\omega_t^{(q)}$ to suppress over-reliance on short-term output efficiency; conversely, if the quality is stable but the efficiency is improving slowly, the performance weight is increased to guide the strategy to migrate in a more efficient direction. To prevent the system from overreacting to short-term improvement fluctuations in a steady-state environment, a mechanism based on absolute values or sliding averages is applied. If the performance and quality improvement rates are low and the system is close to steady state, the dynamic weights are adjusted based on their historical averages or sliding averages, thereby reducing overcorrections caused by fluctuations in a single improvement rate.

To avoid the fluctuation of immediate rewards interfering with the stability of the strategy, the reward used for value function regression is processed by a sliding window and defined as follows:

$$\widetilde{R}_t = \frac{1}{3}\left(R_t + R_{t-1} + R_{t-2}\right) \tag{6}$$

$\widetilde{R}_t$ is the smoothed reward value, covering the current and previous two time steps, and is used as the expected reward input in the reinforcement learning objective function.

Figure 1 shows the trend of the combined impact of the two dimensions of performance and quality on the reward value under different combinations. As product quality improves, the weight of quality feedback in the reward function gradually increases; when the quality is low, the system tends to amplify the performance factor.

This reward structure is calculated once at each reinforcement learning step, and the result is directly used for policy update in the target value feedback stage. To avoid policy instability caused by reward fluctuations, all instantaneous reward data are smoothed by a time sliding window, and the sliding window covers the current and previous two decision cycles. The sliding value is used as the final input to participate in the value function regression of the improved DDQN algorithm.

## IMPROVED DDQN STRUCTURE

To improve the learning efficiency of key behavior patterns in reinforcement learning under a multi-target linkage environment, the improved DDQN structure adopts a dual-structure deep network and a temporal-difference multi-step prediction mechanism combined with a prioritized experience replay strategy based on Temporal Difference (TD) error, to improve the response ability of the policy network to high-value state–action pairs [30,31].

The behavior network and the target network share an identical architecture. The input features (fused employee behavior, process parameters, quality signals, and historical data) are first processed by two 1D convolutional layers for local pattern extraction in the temporal dimension. The output of the convolutional layers is then flattened and passed through a fully connected network consisting of three hidden layers with 256, 128, and 64 nodes, respectively, all utilizing ReLU activation functions. Finally, a linear output layer estimates the Q-values for all possible actions. The behavior network selects the action with the maximum Q-value, and the target network is used for target calculation and periodic synchronization.

A five-step multi-step prediction strategy is adopted, and multi-step cumulative rewards are used instead of single-step rewards to improve the strategy's ability to handle delay effects and long sequence dependencies. The step size for multi-step prediction is selected based on a combination of reinforcement learning theory and the characteristics of textile production. In action-value estimation, a moderate step size balances immediate rewards with long-term benefits. The continuous nature of textile production dictates that its state transitions exhibit specific Markov properties. Analysis shows that a span of five decision intervals (covering 50 seconds of historical data) ensures that the state transition matrix approaches a steady-state distribution. The target value is estimated by the output action of the target network, and the behavior network selects the action index to reduce overestimation and update oscillation.

Experience replay applies TD error priority sampling and adopts a proportional priority mechanism and cumulative probability tree for sample indexing to improve the utilization rate of key state samples. After the error is updated, the priority is dynamically adjusted, and the importance sampling weight is applied to correct the loss and offset the deviation.

The loss function combines multi-step TD errors and weighting coefficients for behavior network updates. The target network is synchronized at a fixed step, and the experience pool capacity covers the state changes within the process cycle, ensuring the breadth and representativeness of the sample distribution and maintaining training stability. Table 2 lists the core components of the improved DDQN architecture and their key parameter configurations.

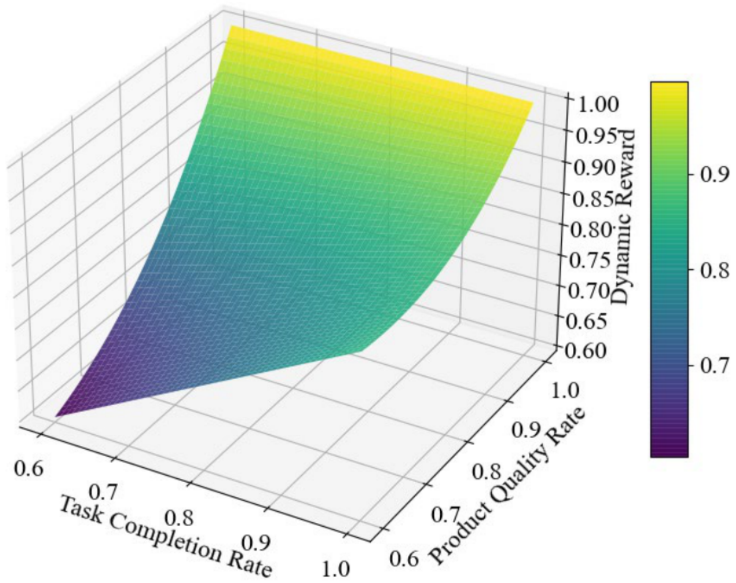


**FIGURE 1.** Dynamic reward function response surface based on performance and quality feedback

**TABLE 2.** Key modules and parameter configuration of the improved DDQN structure

| Module | Core Content | Key Parameters / Configurations |
|---|---|---|
| Network Structure | Two convolutional layers + two fully connected layers | Input feature dimensions, number of convolutional kernels, activation functions, etc. |
| Multi-step Prediction | 5-step temporal difference cumulative reward | Step length = 5, cumulative reward as target |
| Experience Replay | Prioritized experience replay with proportional sampling mechanism | Replay buffer size, priority update |
| Weight Adjustment | Importance sampling weight correction | Weight calculation formula, sampling probability |
| Target Network Update | Parameter synchronization at fixed intervals | Synchronization interval steps |
| Loss Function | Multi-step TD error + weighted coefficients | Loss function structure, weighting ratios |

To enhance the interpretability of model decisions, this study embedded an attention mechanism within the agent network structure. This mechanism automatically learns and quantifies the contribution of different features in the state space to the final decision. By analyzing the distribution of attention weights, production managers can trace the underlying rationale for specific decisions and understand the model's perception– diagnosis–decision logic. This transforms the model's optimization decisions from a black box into reliable, traceable, and supplementary recommendations based on management experience.

### *DECISION-MAKING AND EXECUTION PROCESS OF LINKAGE OPTIMIZATION STRATEGY*

The optimal strategy output by the improved DDQN is converted into a specific execution operation path in the actual textile production system, involving multidimensional collaborative linkage of personnel scheduling, process parameter adjustment, and quality abnormality warning. The strategy module receives real-time multidimensional state input and outputs action instruction sets through the strategy network. The action space covers employee job allocation adjustment, process parameter fine-tuning, and dynamic setting of quality monitoring thresholds. The action instructions are converted into executable control signals by the action parsing unit and executed based on the manufacturing execution system interface. Figure 2 illustrates the decision-making, execution, and feedback process for the linked optimization strategy. It clarifies how employee behavior, process parameters, and quality data serve as state inputs, generating action instructions through the policy network. The execution system then completes the linked operations of personnel scheduling, process adjustments, and quality alerts. The feedback mechanism updates the strategy through a dynamic reward function, achieving closed-loop optimization.

Personnel scheduling is based on employee performance and real-time status, dynamically adjusting task allocation and team structure in combination with production plans. Scheduling instructions are issued by the production system, and behavioral feedback is used to update status and correct strategies.

Process parameter adjustment covers variables such as yarn tension, weaving speed, and dyeing temperature. The strategy instructions are transmitted to the equipment control module through the industrial control system, and parameter adjustment feedback is used to update status and affect strategy evaluation, realizing strategy adaptation based on quality feedback.

The quality abnormality warning is triggered based on

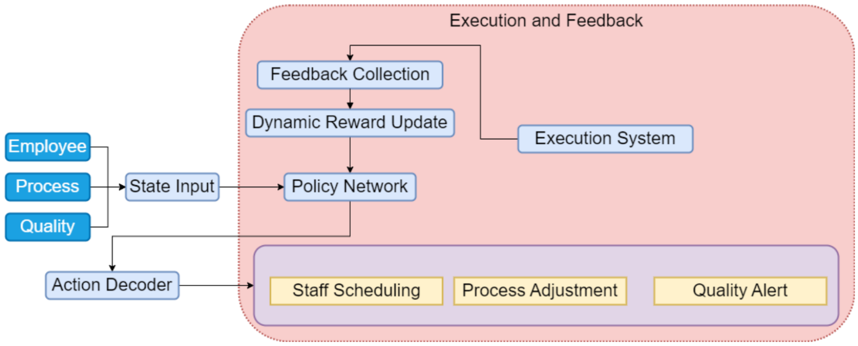


FIGURE 2. Decision-making, execution, and feedback process for the linked optimization strategy.

the risk signal identified by the strategy, including line stop, detection, and personnel intervention, and responds to defects in combination with visual and sensor data. Abnormal data feedback adjusts the reward function and Q-value update direction, improves the priority of quality-oriented actions, and builds a performance–quality linkage mechanism.

During execution, the action scheduling module monitors the operation status in real time, and the status is transmitted back through the industrial network for strategy update. The order of action execution is dynamically sorted by Q-value to improve adaptability under multi-objective conflict.

To ensure the feasibility and safety of all decisions, the action space is predefined with strict physical and regulatory boundaries during the design phase. Any action instructions generated by the policy network must pass a rule-based safety verification module before execution. This module verifies whether the action exceeds the device's safe operating range, complies with personnel skill certification rules, and violates quality standards. The system is equipped with action verification and boundary detection mechanisms and execution logs. The execution path can be adjusted in real time, and the strategy can be updated in combination with performance and quality feedback to achieve closed-loop optimization and dynamic linkage.

## IV. EXPERIMENTAL DESIGN

### DATA SET AND EXPERIMENTAL ENVIRONMENT

The experimental data were collected from the production site of a large functional textile manufacturer. Three parallel production lines were selected, covering 15 production shifts and the production records of 45 operators in total. After data cleaning, the valid data included 320 complete production batches. The employee behavior indicator data were collected jointly by the Radio Frequency Identification (RFID) identifier and the MES system, including the number of tasks completed per unit time, operation consistency score, manual intervention frequency, and abnormal process records. The process parameters were recorded by the PLC control unit and the local SCADA terminal, including continuous variables such as tension control error, equipment vibration frequency, temperature and humidity setting and measured deviation, knitting line speed, etc. The sampling period was set to two seconds. The quality control indicators were provided by a system consisting of a charge-coupled device (CCD) industrial camera, an online thickness gauge, and a color difference analyzer. The indicators covered weft error, yarn break frequency, fabric defect detection score, thickness mean and variance, color difference matching, etc.

The experimental platform was built on a workstation equipped with an Intel Core i7-11700 CPU, 32 GB memory, and an NVIDIA RTX 3060 graphics card. The operating system was Windows 10; the software environment was Python 3.10; the deep learning framework used TensorFlow 2.13; and the CUDA and cuDNN versions were 11.8 and 8.6, respectively. The entire training process was completed in a local stand-alone environment without multi-node parallel acceleration.

### EXPERIMENTAL PROCESS AND IMPLEMENTATION METHODS

The state vector was updated every 10 seconds (i.e., per decision interval) and contained three types of subvectors. After normalization, the employee behavior characteristics were weighted by the historical performance regression coefficient to generate a 32-dimensional behavior vector. The process parameters were extracted through a sliding window to extract sequence features, input into a three-layer 1D convolutional network (kernel sizes were 3, 3, and 2; channels were 32, 16, and 8; activation function was ReLU), and output 48-dimensional features. The quality data were filtered by removing samples with a confidence level lower than 0.7 through a Gaussian mixture model, and PCA compressed the data into a 48- dimensional quality vector.

The three types of features were concatenated into a 128-dimensional state vector and input into the improved DDQN policy network. The action space was a combination of discrete employee scheduling and process fine-tuning, with action selection based on the maximum Q-value. The policy network was a fourlayer structure (256, 128, 64 hidden nodes, ReLU); the output layer was a linear function, with the dimension consistent with the action space. The target network structure was the same, synchronized every 1,000 steps, and the policy network was updated every 250 steps.

Experience replay adopted a priority mechanism; the buffer capacity was 20,000; the sample priority was determined by the TD error; the importance sampling correction factor $\beta$ increased linearly from 0.4 to 1.0; the batch size was 64.

A five-step Q-value prediction was applied, consistent with the state window, and the discount factor $\gamma$ was set to 0.99. The reward function was the weighted sum of performance and quality sub-rewards. The weight was dynamically adjusted according to the volatility within the 50-step window. The exponential weighted moving average was used, and the reward was normalized to [-1, 1].

The training used the Adam optimizer; the initial learning rate was 0.0003; the gradient clipping upper limit was 5; and the loss function was Huber loss. The training consisted of 5,000 steps per round, for a total of 30 rounds, without learning rate decay, and the training log was recorded through TensorBoard.

## V. LINKAGE OPTIMIZATION EFFECT AND EXPERIMENTAL RESULTS

### *EMPLOYEE PERFORMANCE OPTIMIZATION*

To verify the actual application effect of the improved DDQN algorithm in the textile production environment, this study designed a comprehensive experiment for task completion efficiency, system response speed, and personnel operation consistency. Through the continuous execution of multiple batches of production tasks, the experiment collected the number of task completions under different strategies, monitored the average response time of the system in each training stage, and evaluated the behavioral consistency of employees during operation. The experiment was compared with three types of methods: a static strategy based on fixed rules, a basic linkage optimization method based on Deep Q-Network (DQN), and a linkage strategy based on standard DDQN. The results are shown in Figure 3.

The results in Figure 3 show that the improved DDQN consistently outperforms other methods in terms of the number of task completions, reaching an average of 139, and the average response time is 23.32 s. In terms of operation consistency, the median score of the improved DDQN is about 0.81, with small fluctuation, demonstrating a more stable behavior pattern. This performance advantage is due to the improved DDQN's enhanced state recognition and response capabilities through multi-source feature fusion and time-series convolution. Quality data denoising and the weighting mechanism optimize input and enhance decision stability.

### *QUALITY CONTROL OPTIMIZATION*

To quantify the effectiveness of each strategy in production quality control, the experiment focused on the defect rate. By statistically modeling the defect rate data collected in a large sample batch, the ability and stability of each strategy to control production fluctuations were studied to reflect the consistency control ability of the intelligent strategy on the quality of finished products. This evaluation was conducted across 100 simulation runs to robustly assess the performance of various strategies under varying initial conditions. Each simulation run was initialized by randomly sampling a starting state from historical data. Subsequently, the process parameters and potential noise in each simulation step were generated based on statistical properties derived from real data. Figure 4 shows the probability density estimation and histogram distribution of the cloth defect rate of the four methods in 100 simulation runs. Figure 5 shows the distribution of the pass rate of each method in different batches, as well as the numerical relationship between the pass rate and the defect rate.

Figure 4 shows the performance of different scheduling strategies in controlling cloth defect rates. The static strategy has the widest defect rate distribution, mainly concentrated at 7%, with large fluctuations; the DQN strategy has slightly better convergence, mainly concentrated between 4% and 6%; DDQN further narrows the distribution range, and the average defect rate drops to about 3.7%; the improved DDQN defect rate is concentrated between 2% and 3%, with an average of 2.4%. This trend reflects the difference in the ability of each strategy to respond to dynamic states. The improved DDQN integrates multi-source information and timing features to achieve more precise suppression of defects.

As shown in Figure 5, the improved DDQN performs stably in all batches, with a pass rate of $\geq$92.2%, while the lowest batch pass rate of the traditional static method is close to 74.2%. The improved DDQN points are concentrated in areas with low defect rates and high pass rates, showing superior performance. This difference lies in the different responsiveness of each strategy to state feedback. The improved DDQN alleviates the problem of overestimation of actions and enhances the stability of strategy updates.

### *IMPACT OF LINKAGE MECHANISM ON PRODUCTION EFFICIENCY*

To evaluate the performance of the linkage mechanism in terms of production efficiency, this experiment compares and analyzes the execution effect of each method in the manufacturing process from two aspects: unit output time and unit time output. The results are shown in Figure 6.

It can be seen that the static method has the highest unit output time of 9.6 minutes, and the output efficiency is only 6.1 pieces/hour; while the improved DDQN unit output time is 7 minutes, and the unit time output is increased to 8.3 pieces/hour. This efficiency improvement results from the enhanced intelligence of the scheduling strategy. The improved DDQN optimizes state expression and the reward mechanism, improves the ability of scheduling actions to grasp long-term benefits, and achieves a better balance between time control and output efficiency.

To evaluate the comprehensive performance of different strategies in production line operation, this section con-

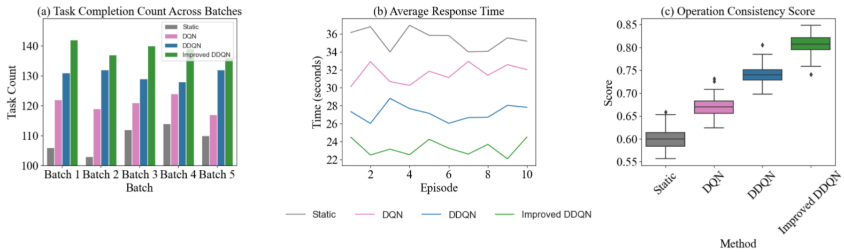


FIGURE 3. Number of task completions, response time, and operation consistency performance of strategy execution in the textile production system

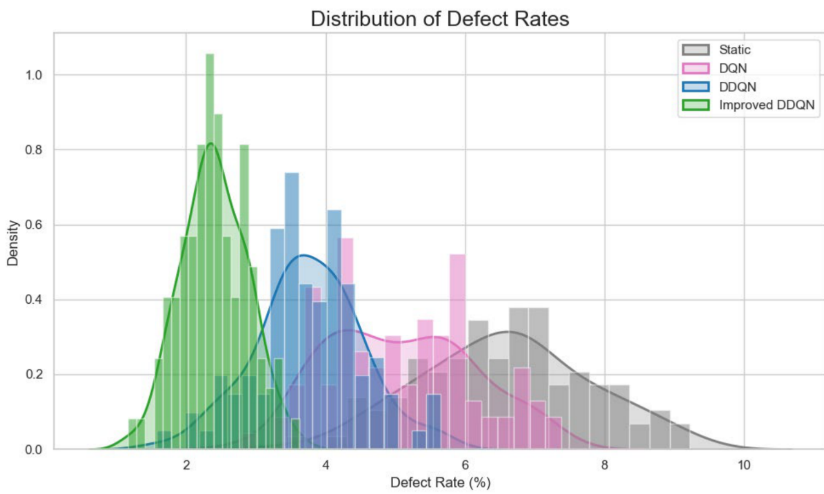


FIGURE 4. Probability density distribution of defect rates of each strategy

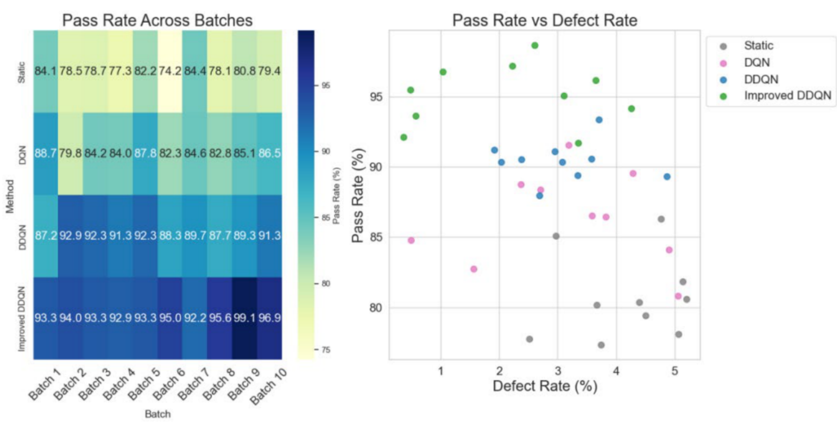


FIGURE 5. Distribution of pass rates of each strategy and its correlation with defect rates

ducts quantitative analysis from multiple key dimensions such as total output, quality stability, and overall performance, and constructs an evaluation system including output, defect rate, and scoring indicators (the scoring is a weighted comprehensive performance score of three standardized indicators: output efficiency, quality stability, and operation consistency) to measure the execution efficiency and improvement potential of each strategy in a complex production environment. The calculation formula for the comprehensive performance score C is as follows:

$$C = \omega_e \hat{E} + \omega_s \hat{S} + \omega_c \hat{C} \tag{7}$$

$\hat{E}$, $\hat{S}$ and $\hat{C}$ represent the minimum–maximum standardized values of output efficiency, quality stability, and operational consistency scores, respectively. The weight coefficients are set to $\omega_e = 0.4$, $\omega_s = 0.4$, and $\omega_c = 0.2$. The results are shown in Table 3.

The static method has a total output of 10,200 pieces, a high defect rate (6.5%), and a performance score of 72.3. DQN increases the output to 11,800 pieces, reduces the

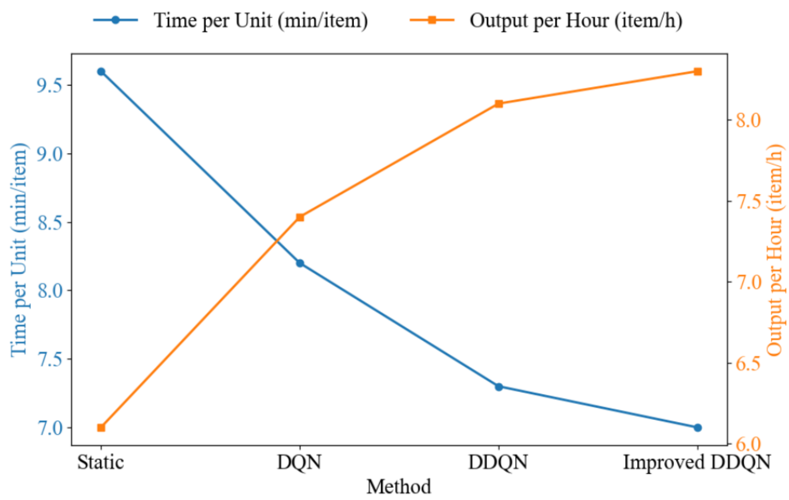


**FIGURE 6.** Comparison of performance of each method in unit output time and unit time output

**TABLE 3.** Comprehensive comparison of production performance of different strategies

| Method | Total output (pieces) | Improvement rate (%) | Average defect rate (%) | Standard deviation of defect rate | Comprehensive performance score |
|---|---|---|---|---|---|
| Static | 10,200 | — | 6.5 | 1.1 | 72.3 |
| DQN | 11,800 | +15.7 | 5.1 | 0.95 | 78.9 |
| DDQN | 12,700 | +24.5 | 3.9 | 0.62 | 83.6 |
| Improved DDQN | 13,800 | +35.3 | 2.6 | 0.40 | 89.4 |

defect rate to 5.1%, and improves performance to 78.9. DDQN further improves output and quality, with an output of 12,700 pieces, a defect rate of 3.9%, and a score of 83.6. The improved DDQN performs best, with an output of 13,800 pieces, a defect rate of only 2.6%, and a score of 89.4. This result is due to the fact that the improved DDQN combines multiple optimization strategies to improve learning efficiency and decision accuracy, achieving a double breakthrough in output and quality.

To quantitatively evaluate the performance improvement of the improved DDQN strategy compared to the baseline method, this study used an independent sample t-test for inter-group statistical comparison and used the comprehensive performance score as the evaluation indicator to analyze the degree of difference between each strategy and the improved DDQN. The specific statistical results are shown in Table 4.

The improved DDQN outperformed all compared methods with an overall performance score of 89.4, with a standard deviation of only 1.1, demonstrating outstanding stability. Statistical results showed that the differences between the groups were significant ($p < 0.001$), with the largest mean difference compared to the static strategy. This advantage stems from the synergy of multiple underlying mechanisms in the model: a 128-dimensional state space incorporates multidimensional production characteristics; a dynamic reward function uses a weighted adaptive mechanism to balance efficiency and quality objectives in real time; and prioritized experience replay and multi-step forecasting enhance the algorithm's ability to capture long-term decision relationships, ultimately achieving improved overall performance.

### *CONVERGENCE OF THE IMPROVED DDQN ALGORITHM*

This experiment compares and analyzes the changes in the loss of various models in consecutive training rounds, aiming to reveal the optimization efficiency of the model and the fluctuation characteristics during the training process. By monitoring the relationship between the training round and the loss value, the effect of parameter updates and the stabilization speed of each algorithm during the learning process are evaluated. Figure 7 shows the trend of the training round and the loss value.

Overall, the training loss curve of the improved DDQN demonstrates superior convergence properties compared to the baseline algorithms. It not only starts from a lower initial loss but also achieves a stable plateau significantly faster, with its final loss value being an order of magnitude lower than that of DQN and DDQN. This accelerated and more stable convergence pattern indicates higher learning efficiency and optimization effectiveness of the proposed improvements.

## VI. CONCLUSION

This paper proposes a multi-objective linkage mechanism based on the improved DDQN for the collaborative optimization problem of employee performance and functional textile quality control in textile production. By constructing a multidimensional state-space model, integrating employee behavior, process parameters, and quality control data, and

**TABLE 4.** Analysis of the significance of the overall performance of different scheduling strategies

| Group | Improved DDQN | Static | DQN | DDQN |
|---|---|---|---|---|
| $n$ | 20 | 20 | 20 | 20 |
| Mean Score | 89.4 | 72.3 | 78.9 | 83.6 |
| SD | 1.1 | 3 | 2.5 | 2 |
| Mean Difference (95% CI) | — | -17.1 (-18.7, -15.5) | -10.5 (-11.8, -9.2) | -5.8 (-6.8, -4.8) |
| $t$ value | — | -21.39 | -15.87 | -11.43 |
| df | — | 28.4 | 33.2 | 33.8 |
| $p$ value | — | $< 0.001$ | $< 0.001$ | $< 0.001$ |

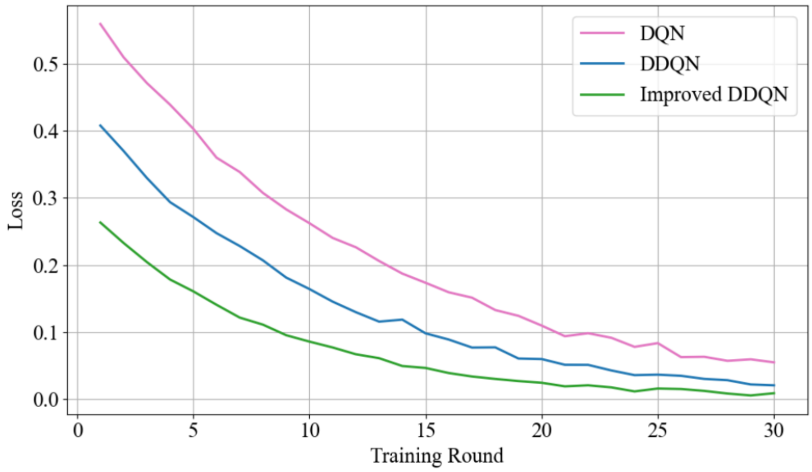


**FIGURE 7.** Changes in model training loss

combining dynamic reward functions and a prioritized experience replay mechanism, the collaborative optimization of production efficiency and product quality is achieved. Experimental results show that the improved DDQN is superior to traditional methods in terms of task completion number, defect rate, and comprehensive performance. In addition, the algorithm converges faster, and the training loss is stable at 0.02, verifying its efficiency and robustness in a dynamic production environment. This study provides an innovative multi-objective optimization solution for textile intelligent manufacturing, which has important industrial application value.

## AUTHOR CONTRIBUTIONS

Conceptualization – Chunsun He, Lei Wang; methodology – Chunsun He; investigation – Chunsun He; resources – Lei Wang; writing-original draft preparation – Chunsun He; writing-review and editing – Lei Wang. All authors have read and agreed to the published version of the manuscript.

## CONFLICTS OF INTEREST

The author declares no conflict of interest.

## FUNDING

This research received no external funding.

## ACKNOWLEDGEMENTS

Suzhou City University National-level Pre-research Project.

## REFERENCES

[1] H. Hatif, K. L. Lee, and G. Nawanir, “Improving employee and workplace performance via lean manufacturing practices: a case study in textile company,” International Journal of Industrial Management, vol. 18, no. 1, pp. 43-59, 2024, doi: 10.15282/ijim.18.1.2024.10425.
[2] M. Ewnetu and Y. Gzate, “Assembly operation productivity improvement for garment production industry through the integration of lean and work-study, a case study on Bahir Dar textile share company in garment, Bahir Dar, Ethiopia,” Heliyon, vol. 9, no. 7, pp. e17917, 2023, doi: 10.1016/j.heliyon.2023.e17917.
[3] M. R. Islam, M. Z. H. Zamil, M. E. Rayed, M. M. Kabir, M. F. Mridha, S. Nishimura, et al., “Deep Learning and Computer Vision Techniques for Enhanced Quality Control in Manufacturing Processes,” IEEE Access, vol. 12, pp. 121449-121479, 2024, doi: 10.1109/ACCESS.2024.3453664.
[4] M. P. Sikka, A. Sarkar, and S. Garg, “Artificial intelligence (AI) in textile industry operational modernization,” Research Journal of Textile and Apparel, vol. 28, no. 1, pp. 67-83, 2024, doi: 10.1108/RJTA-04-2021-0046.
[5] L. Chen, Q. Wang, C. Deng, B. Xie, X. Tuo, and G. Jiang, “Improved Double Deep Q-Network Algorithm Applied to Multi-Dimensional Environment Path Planning of Hexapod Robots,” Sensors, vol. 24, no. 7, pp. 2061, 2024, doi: 10.3390/s24072061.
[6] S. Zhang, L. Zhang, F. Xu, S. Cheng, W. Su, and S. Wang, “Dynamic deployment method based on double deep Q-network in UAV-assisted

MEC systems," Journal of Cloud Computing, vol. 12, no. 1, pp. 130, 2023, doi: 10.1186/s13677-023-00507-6.

[7] Y. Li, C. Yu, M. Shahidehpour, T. Yang, Z. Zeng, and T. Chai, "Deep Reinforcement Learning for Smart Grid Operations: Algorithms, Applications, and Prospects," Proceedings of the IEEE, vol. 111, no. 9, pp. 1055-1096, 2023, doi: 10.1109/JPROC.2023.3303358.

[8] N. Yang, S. Chen, H. Zhang, and R. Berry, "Beyond the Edge: An Advanced Exploration of Reinforcement Learning for Mobile Edge Computing, Its Applications, and Future Research Trajectories," IEEE Communications Surveys & Tutorials, vol. 27, no. 1, pp. 546-594, 2025, doi: 10.1109/COMST.2024.3405075.

[9] J. Zhou, Y. Tian, and M. Ren, "Data-efficient reinforcement learning by generalized value estimation," Machine Learning, vol. 114, no. 6, pp. 133, 2025, doi: 10.1007/s10994-025-06763-8.

[10] L. Zhao, T. Chang, L. Guo, J. Zhang, L. Zhang, and J. Ma, "An Overestimation Reduction Method Based on the Multistep Weighted Double Estimation Using Value-Decomposition Multi-agent Reinforcement Learning," Neural Processing Letters, vol. 56, no. 3, pp. 152, 2024, doi: 10.1007/s11063-024-11611-2.

[11] D. Hussain and M. C. Figueiredo, "Improving the time-based performance of the preparatory stage in textile manufacturing process with value stream mapping," Business Process Management Journal, vol. 29, no. 3, pp. 801-837, 2023, doi: 10.1108/BPMJ-08-2022-0366.

[12] P. Kalia, M. Singla, and R. Kaushal, "Human resource management practices and employee retention in the Indian textile industry," International Journal of Productivity and Performance Management, vol. 73, no. 11, pp. 96-121, 2024, doi: 10.1108/IJPPM-01-2022-0057.

[13] J. M. X. Casavilca, B. A. Y. Pedraza, C. B. A. Meneses, and R. J. A. Zapana, "Human talent management and the work performance of employees in a textile company in Comas," Southern Perspective / Perspectiva Austral, vol. 1, pp. 5, 2023, doi: 10.56294/pa20235.

[14] Y. Liu, C. Zhang, and X. Dong, "A survey of real-time surface defect inspection methods based on deep learning," Artificial Intelligence Review, vol. 56, no. 10, pp. 12131-12170, 2023, doi: 10.1007/s10462-023-10475-7.

[15] S. Zhou, J. Zhao, Y. Shi, Y. Wang, and S. Mei, "Research on improving YOLOv5s algorithm for fabric defect detection," International Journal of Clothing Science and Technology, vol. 35, no. 1, pp. 88-106, 2023, doi: 10.1108/IJCST-11-2021-0165.

[16] A. Metin and T. T. Bilgin, "Automated machine learning for fabric quality prediction: a comparative analysis," PeerJ Computer Science, vol. 10, pp. e2188, 2024, doi: 10.7717/peerj-cs.2188.

[17] T. Liang, T. Liu, J. Wang, J. Zhang, and P. Zheng, "Causal deep learning for explainable vision-based quality inspection under visual interference," Journal of Intelligent Manufacturing, vol. 36, no. 2, pp. 1363-1384, 2025, doi: 10.1007/s10845-023-02297-9.

[18] T. Meeradevi, S. Sasikala, S. Gomathi, and K. Prabakaran, "An analytical survey of textile fabric defect and shade variation detection system using image processing," Multimedia Tools and Applications, vol. 82, no. 4, pp. 6167-6196, 2023, doi: 10.1007/s11042-022-13575-8.

[19] R. S. Sabeenian, E. Paul, and C. Prakash, "Fabric defect detection and classification using modified VGG network," The Journal of The Textile Institute, vol. 114, no. 7, pp. 1032-1040, 2023, doi: 10.1080/00405000.2022.2105112.

[20] X. Luo, "Reshaping Coordination Efficiency in the Textile Supply Chain Through Intelligent Scheduling Technologies," Economics and Management Innovation, vol. 2, no. 4, pp. 1-9, 2025, doi: 10.71222/ww35bp29.

[21] M. Chen, Y. Kang, K. Li, P. Li, and Y. Zhao, "Deep reinforcement learning for maintenance optimization of multicomponent production systems considering quality and production plan," Quality Engineering, vol. 37, no. 2, pp. 219-230, 2025, doi: 10.1080/08982112.2024.2373362.

[22] O. Nocentini, J. Kim, Z. M. Bashir, and F. Cavallo, "Learning-based control approaches for service robots on cloth manipulation and dressing assistance: a comprehensive review," Journal of NeuroEngineering and Rehabilitation, vol. 19, no. 1, pp. 117, 2022, doi: 10.1186/s12984-022-01078-4.

[23] H. Xu, J. Zheng, L. Huang, J. Tao, and C. Zhang, "Solving Dynamic Multi-Objective Flexible Job Shop Scheduling Problems Using a Dual-Level Integrated Deep Q-Network Approach," Processes, vol. 13, no. 2, pp. 386, 2025, doi: 10.3390/pr13020386.

[24] A. B. Mizhir, "Analysis of the Impact of Implementation of Integrated Manufacturing Information Systems (MES) in Increasing Production Efficiency in the Electronics Industry," Jurnal Mekintek: Jurnal Mekanikal, Energi, Industri, Dan Teknologi, vol. 16, no. 1, pp. 9-18, 2025, [Online]. Available: https://www.ejournal.isha.or.id/index.php/Mekintek/article/view/377.

[25] M. M. Mohamed and A. S. Alraddadi, "Agile Factor Flexible Industry 4.0 with MES Manufacturing Execution System Along with ERP Back-End Integration: Ready-Made Garments as a Case Study," American Journal of Operations Research, vol. 14, no. 4, pp. 137-150, 2024, doi: 10.4236/ajor.2024.144007.

[26] Y. Chen, W. Chiu, Q. Wang, J. Chen, and H. Zhao, "Energy-Efficient Prediction in Textile Manufacturing: Enhancing Accuracy and Data Efficiency WITH Ensemble Deep Transfer Learning," IEEE Access, vol. 13, pp. 57177-57190, 2025, doi: 10.1109/ACCESS.2025.3551798.

[27] A. A. Amin, N. Alvi, M. Irfan, M. Abdullah, and H. U. Manzoor, "Design of a Low-Cost Efficient IoT Based SCADA System for Automating Different Textile Machines and Controllers for Central Monitoring," The Journal of Engineering, vol. 2025, no. 1, pp. e70089, 2025, doi: 10.1049/tje2.70089.

[28] F. He, B. Xin, Z. Zhan, W. Yang, and J. Liu, "A review of fabric defect detection in textile manufacturing," The Journal of The Textile Institute, pp. 1-19, 2025, doi: 10.1080/00405000.2025.2502188.

[29] D. Zhao, Z. Ding, W. Li, S. Zhao, and Y. Du, "Cascaded Fuzzy Reward Mechanisms in Deep Reinforcement Learning for Comprehensive Path Planning in Textile Robotic Systems," Applied Sciences, vol. 14, no. 2, pp. 851, 2024, doi: 10.3390/app14020851.

[30] J. Chang, D. Yu, Z. Zhou, W. He, and L. Zhang, "Hierarchical Reinforcement Learning for Multi-Objective Real-Time Flexible Scheduling in a Smart Shop Floor," Machines, vol. 10, no. 12, pp. 1195, 2022, doi: 10.3390/machines10121195.

[31] L. Song, Y. Li, and J. Xu, "Dynamic Job-Shop Scheduling Based on Transformer and Deep Reinforcement Learning," Processes, vol. 11, no. 12, pp. 3434, 2023, doi: 10.3390/pr11123434.