

**Paper 96-2** has been designated as a Distinguished Paper at Display Week 2025. The full-length version of this paper appears in a Special Section of the *Journal of the Society for Information Display (JSID)* devoted to Display Week 2025 Distinguished Papers. This Special Section will be freely accessible until December 31, 2025 via:

<https://sid.onlinelibrary.wiley.com/doi/full/10.1002/jsid.2086>

Authors that wish to refer to this work are advised to cite the full-length version by referring to its DOI:

<https://doi.org/10.1002/jsid.2086>



# Optimize Manufacturing Operations with Digital Twin and Deep Q-Network

**Seki Park\***, **Sanghwa Lee\***, **Hyeonji Son\***, **Junghoon Kim\***, **Dongwon Han\***  
**Junwoo Lim\*\***, **EunYoung Choi\*\***, **Seoyoung Jo\*\***, **Hyun Seok Lee\*\***,  
**Whan Lee\*\***, **Sang Do Noh\*\***

\*Mobile Display Technology Innovation Team, Samsung Display, Chungcheongnam-do, Korea

\*\*Department of Industrial Engineering, Sungkyunkwan University, Gyeonggi-do, Korea

## Abstract

*The display manufacturing industry is exploring operational optimization strategies using digital twins to address the challenges of complex production lines and rapidly changing market demands. In this study, we explored an optimal solution to released equipment constraints in manufacturing lines using digital twin technology and deep reinforcement learning. As a result, the number of released equipment constraints was reduced by 27% compared to the conventional approach, and production volume increased by 4%. To facilitate the interaction between digital twins containing on-site reference information and deep reinforcement learning, the study adhered to the international ISO 23247 standard. This research marks the world's first attempt at a digital transformation study focused on manufacturing line operations using this approach.*

## Author Keywords

digital twin; manufacturing operation optimal solution; deep Q-network; digital twin based on international standards

## 1. Introduction

Digital twin is playing an increasingly important role in the manufacturing industry and holds the potential to revolutionize production efficiency and quality, especially when combined with AI. A digital twin is a virtual model of a physical system, used to monitor, analyze, and optimize the physical system in real time based on data. With recent advancements in AI technology, digital twins go beyond simple data visualization to perform advanced functions such as anomaly detection, predictive maintenance, and process optimization by leveraging machine learning and deep learning algorithms.

In particular, AI-based digital twins have opened new possibilities for processing real-time data and enhancing production process efficiency through predictive analytics in manufacturing. Lee et al. suggest that AI technologies, through digital twins, can contribute to overall manufacturing process optimization and reduced downtime [1, 2]. Abdelmoula et al. applied digital twins to a plastic injection molding factory, conducting simulations based on real-time data. They analyzed various process settings and outcomes using deep reinforcement learning to derive optimal process parameters [3]. Yongkui et al. utilized digital twins and the deep reinforcement learning algorithm DQN to optimize gripping tasks for industrial robots. A digital twin model was used to build a virtual robotic system, integrating real-world data to train the robot using the DQN algorithm. This enabled the robot to learn the optimal position and orientation for grasping objects, which was then implemented in the physical robot [4]. João V. A. Cabral et al. developed a digital twin architecture for metal additive manufacturing robotic cells based on the ISO 23247 standard. They proposed an approach to optimize processes through real-time data collection and 3D simulation [5].

This study focuses on a display manufacturing line. The equipment in this line operates under certain flow-controlling constraints, managed either by automated systems or on-site workers. However, these constraints are often overridden based on the subjective judgment of workers, leading to bottlenecks in the process flow. To address this issue, we explored a solution combining digital twin simulations with the Deep Q-Network (DQN) algorithm. This approach aims to optimize production output and enhance the intelligent operation of the manufacturing line while considering the constraints of each piece of equipment.

## 2. Research Methodology

### 2.1 The Concept of Digital Twin and AI-based Operation Optimization

A digital twin is a technology that virtually implements physical system, reflecting data in real time while enabling analysis and prediction, thereby facilitating efficient management of manufacturing processes. When combined with AI, digital twins can derive operational strategies that learn and execute optimal management plans, even under various constraints. This study proposes an operational optimization methodology that integrates simulation and AI technologies to effectively manage the constraints of actual display manufacturing processes and maximize productivity. This system collects real-time data from the manufacturing site to perform simulations within the digital twin. The results of these simulations are then processed by an AI model, which learns and derives optimal operational strategies. The derived strategies are applied to the manufacturing site to support or execute the decision-making required for actual operations. The following section details the specific methods used to implement this approach.

### 2.2 Digital Twin with AI and ISO 23247

This study aims to develop an operational optimization system that integrates AI models with digital twins for a single process in display manufacturing, considering equipment constraints. By combining AI with a digital twin that incorporates constraint-based reference information and the operational logic of the display manufacturing site, the system provides a consistent decision-making framework to release or adjust equipment constraints effectively.

Figure 1 illustrates the framework of the system proposed in this study. Using operational and reference information collected during production, the system conducts simulations that reflect the equipment constraints and production processes of the actual manufacturing site. The data generated from these simulations is passed to an AI model, which uses it to derive operational optimization strategies while considering equipment constraints. The AI learns to achieve the highest possible production output while considering the constraints of each piece of equipment. It receives greater rewards when production output is high and the

number of constraint releases is minimized. Throughout this process, the transfer and timing of data between the simulation and the AI are managed by a ‘Behavior Model’ defined in compliance with ISO 23247. This ensures the reusability and scalability of the digital twin.

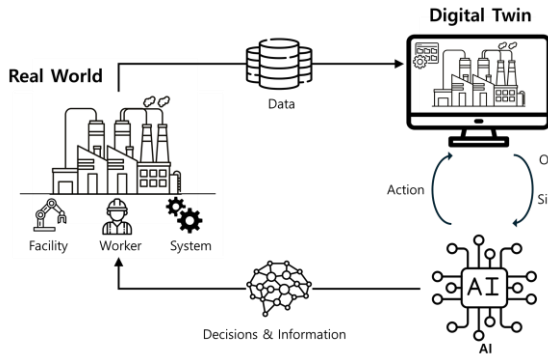


Figure 1. Concept map of digital twin and AI-based operation optimization

The data exchange system is designed with a modular (model) structure to enable the implementation of digital twins in a display manufacturing environment. It facilitates interaction and data exchange between AI model and digital twin model (Figure 2). In this system, data exchange is conducted between the ISO 23247 module and the digital twin model, as well as between the ISO 23247 model and the AI model, using interfaces. The interfaces support data integration between the models and are designed to allow seamless data transmission through a unified interface. This ensures that data can be exchanged in a consistent format, even between different systems. At the core of this interaction is the state transition functionality of the behavior model, which enables efficient integration of data between models based on input and output mappings of key data.

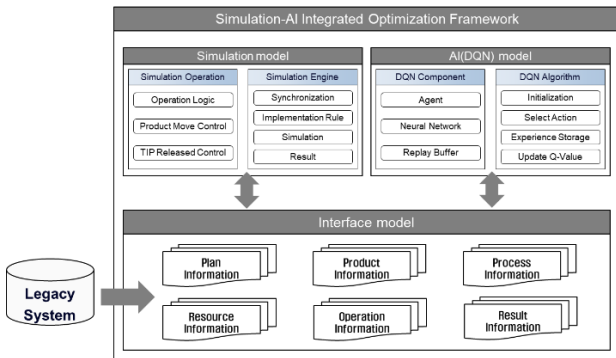


Figure 2. Digital twin and AI convergence system framework

In the case of the simulation model, it performs simulations based on various input data from the manufacturing environment, such as equipment, planning, logistics, processes, and operations, and outputs results like production volume and constraints. On the other hand, the AI model utilizes data such as state information controlling AI learning and behavior, reward information, and learning parameters to ultimately output behavioral information,

which represents AI's decision-making outcomes. This data flow is integrated by the behavior model, where information is exchanged through input/output mappings between the AI model and the simulation model. Once certain conditions are met, state transitions occur. Through this process, the system maintains consistency across different systems using a standardized data structure, enabling diverse systems to establish a unified data exchange framework with ease (Figure 3).

Simulation model	Behavior model	AI model
<b>Input</b> Equipment info. Planning info. Logistics info. Process info. Operation info.	<b>Data Input/Output Mapping</b> Data to be sent Data to be received  <b>State Transition</b> Status-change condition  <b>External Interaction Transition</b> an interaction condition btw. system	<b>Input</b> AI status info. AI compensation info. Logistics info. Learning parameter  <b>Output</b> AI behavior information
<b>Output</b> # of BP Glass Production Work in Process Turn Around Time # of TIP-Released facilities		

Figure 3. Models required for common interface defined in ISO 23247 of the international standard for digital twin

### 2.3 Reinforcement Learning Method

The applied algorithm chosen for this system is the Deep Q-Network (DQN), capable of learning in complex environments. DQN is a reinforcement learning algorithm that combines Q-learning with neural networks, enabling the agent to learn optimal behaviors using a deep neural network (DNN). Q-learning estimates the Q-value, which represents the expected reward (return) obtained by the agent when selecting action  $a$  in state  $s$ , and it uses this to learn the optimal policy. The optimal Q-value function  $Q^*(s, a)$  is calculated as the value satisfying the Bellman equation. The Bellman equation provides a formula for updating the state-action value function  $Q^*(s, a)$ .

$$Q^*(s, a) = E [r + \gamma * \max_{a'} Q^*(s', a')] \quad (1)$$

**E** (Expectation): Calculates the average expected value based on the probability distribution of the environment.

**r**: The immediate reward obtained by performing action  $a$  in the current state.

**s'**: The next state reached after performing action  $a$

**$\gamma$**  (Discount Factor): A value between 0 and 1 that determines the importance of future rewards. A value of 0 means the agent focuses solely on immediate rewards, while a value of 1 treats future rewards as equally important as the current ones.

**$\max Q^*(s', a')$** : The maximum Q-value among the actions available in the next state  $s'$ , representing the highest expected value of future rewards.

To enhance the stability and efficiency of learning, DQN introduces the experience replay technique. All experiences during the learning process—comprising state, action, reward, and next state—are stored in a ‘Replay Buffer’. Mini-batches are then randomly sampled from this buffer for training. This approach prevents correlated experiences from excessively

influencing the neural network, ensuring more stable and efficient learning. Additionally, DQN employs a ‘Target Network’ to mitigate the instability of Q-learning updates. The target network in DQN periodically copies the weights from the main network (Q-Network) and keeps them fixed for a set duration. It is then used to predict the Q-values for the next state. This mechanism helps apply the Bellman equation stably by reducing fluctuations in the learning process.

The loss function of DQN is designed to minimize the Temporal Difference (TD) Error, which represents the difference between the Q-values predicted by the main network and the target Q-values provided by the target network. The loss function is defined as follows in equation 2. :

$$L(\theta) = E_{(s,a,r,s')} [(r + \gamma * \max Q_{target}(s', a'; \theta^-) - Q(s, a; \theta))]^2 \quad (2)$$

$\theta$  represents the weights of the current Q-Network, while  $\theta^-$  represents the weights of the target network.

### 3. Experiment and Result Analysis

#### 3.1 Case Study: Target Process and Problem Statement

As a case study, this research focuses on the DRY process within a display manufacturing line, addressing issues related to equipment selection and the release of TIP (Track In Prevent) constraints. TIP refers to an operational constraint in production processes that limits product input to specific equipment. This constraint is imposed due to maintenance or quality issues and can only be lifted when certain predefined conditions are satisfied. During the process operation, equipment is selected based on one of four equipment operation rules. If the selected equipment is under TIP, the system evaluates whether the conditions for TIP release are met. When the conditions for TIP release are satisfied, TIP is lifted, allowing equipment to operate. This improves equipment utilization and maximizes the efficiency of the entire process flow. In this study, digital twins and AI were utilized to perform operational optimization, considering TIP release, with the goal of maximizing production output. To develop the system proposed in the study, simulation model, data interface, and AI models were required. The simulation model was built using Siemens' Plant Simulation 22.1 tool, while the data interface model and AI model were developed using the Spyder 5.5.0 tool with Python 3.11.6.

#### 3.2 Simulation Model

Figure 4 shows the digital twin of our factory, created using Siemens' Technomatix Plant Simulation. In order to apply AI to the model, we worked to link the behavior model and the simulation model using standards. Before conducting parameter optimization using the entire twin, the research was initially carried out using a single process.

The DQN agent observes the state of the simulation and selects an operating rule based on the defined rules. The data received as the simulation state includes daily production indicators such as the number of work-in-progress items, the number of equipment with TIP released, and production volume, and so on. At fixed intervals, the agent selects an action and updates the state based on the simulation results. The selection of the next equipment was based on operating rules. Although the next equipment was

selected, if any constraints were imposed on the equipment, the logic was designed to resolve these constraints. The reward was based on daily production volume. This series of tasks was performed through a 10-day predictive simulation. To find the optimal solution, a total of 500 episodes were executed, with the rules for TIP release being adjusted for each episode to optimize the process.



Figure 4. The factory snapshot of target process for optimization

#### 3.3 Digital Twin with Reinforcement Learning

The DQN process flow used in this study is shown in Figure 5. The network consists of three fully connected layers, and the agent was trained through 500 simulation experiments.

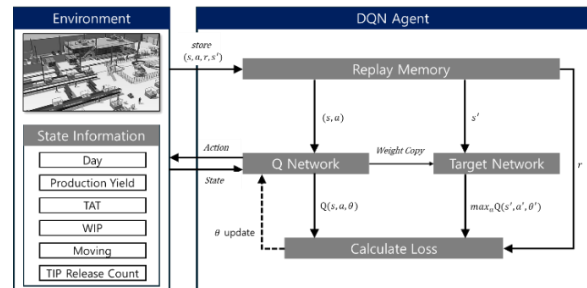


Figure 5. TIP Release Operation Optimization Methodology Using DQN

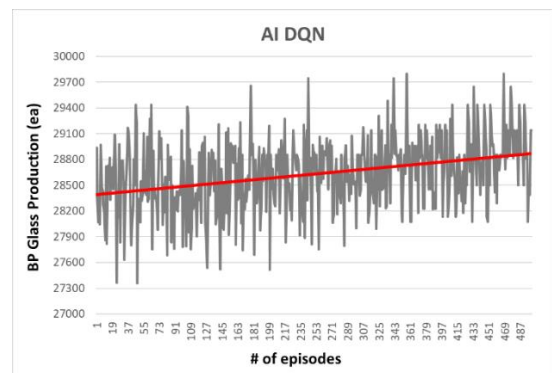


Figure 6. The DQN learning result graph. It can be seen that reward (production) increases as episodes increase. The red line is the trend.

Figure 6 shows the DQN learning result, where the production volume corresponding to the reward gradually increases over the course of 500 episodes. Among these episodes, the 357th episode proposed the most optimal TIP release operating rule, which is presented in Table 1. As shown in Table 1, episode 357 followed operation rule #4 on the first day of the simulation, applied operation rule #1 from days 2 to 5, and followed operation rule #3 from days 6 to 10, except for day 7. This combination was found to be the optimal condition, resulting in the highest production output.

**Table 1.** Daily TIP switching operation rule optimized through DQN learning

Day	1	2	3	4	5	6	7	8	9	10
Epi48	3	4	1	1	1	2	2	1	3	3
Epi302	4	1	1	1	1	1	1	4	1	1
Epi357	4	1	1	1	1	3	2	3	3	3

Rule #1: A higher score is assigned when the number of lots in the queue is smaller.

Rule #2: A score of 1 is given if on the same floor, and 0 if it is on a different floor.

Rule #3: A higher score is assigned when the distance between the current stocker and the next stocker is shorter.

Rule #4: The sum of all the above scores.

The objective of this experiment was to optimize both production volume and the number of released equipment constraints. The results comparing the case with AI application and the case without AI are presented in Table 2. When AI was applied, production increased by 1,133 units compared to when it was not applied, representing a 4.01% improvement. Furthermore, the number of released constraints was reduced by approximately 27%, allowing for an increase in production with fewer constraints being lifted. These findings confirm that AI can effectively optimize production volume while managing equipment constraints more efficiently.

**Table 2.** Daily TIP switching operation rule optimized through DQN learning

No	Rule	# of TIP releases facilities(ea)	# of BP production(ea)
0	No Release	0	26,706 (5.53%↓)
1	LOT Queue Score (Current Rule)	64	28,270 (Baseline)
2	Same Floor Score	32	28,039 (0.82%↓)
3	Transfer Time Score	26	27,700 (2.02%↓)
4	<b>DQN Result</b>	<b>47</b>	<b>29,403</b> <b>(4.01%↑)</b>

#### 4. Conclusions

This paper proposes an operational optimization method that takes equipment constraints into account by integrating digital twins and reinforcement learning. The DQN based reinforcement

learning agent used in this study was designed to find the best operating strategies while considering various equipment constraints. Additionally, a data exchange system following ISO 23247 standards was implemented to ensure smooth communication between different systems, keeping the structure of distributed data consistent and enabling technological integration across different manufacturing elements. In the future, the optimization will extend beyond a single unit process to the entire manufacturing line, with the goal of increasing production by more than 5%.

#### 5. Acknowledgements

This work was supported by SAMSUNG DISPLAY Co. Ltd.

#### 6. References

- Lee J, Davari H, Singh J, Pandhare V. Industrial AI: Applications with digital twin. Springer International Publishing; 2020. Available from: <https://doi.org/10.1007/978-3-030-26586-8>
- Rathore MM, Shah SA, Shukla D, Bentafat E, Bakiras S. The role of AI, machine learning, and big data in digital twinning: A systematic literature review, challenges, and opportunities. IEEE Access. 2021;9:32030–52.
- Khdoudi A, Masrour T, El Hassani I, El Mazgualdi C. A deep-reinforcement-learning-based digital twin for manufacturing process optimization. Systems. 2024;12(2):38.
- Liu Y, Xu H, Liu D, Wang L. A digital twin-based sim-to-real transfer for deep reinforcement learning-enabled industrial robot grasping. Robotics and Computer-Integrated Manufacturing. 2022;78:102365.
- Cabral JVA, Álvares AJ, de Carvalho GC. Digital twin implementation for an additive manufacturing robotic cell based on the ISO 23247 standard. IEEE Latin America Transactions. 2024;22(8):651–8.