System Engineering and Productivity/ 2026/ Vol.6/ No.3/ 287-301

System Engineering and Productivity

DOI: 10.22034/sep.2026.2077232.1426

Reinforcement Learning Framework for Workload Distribution and Scheduling in Multi-Model Production Lines

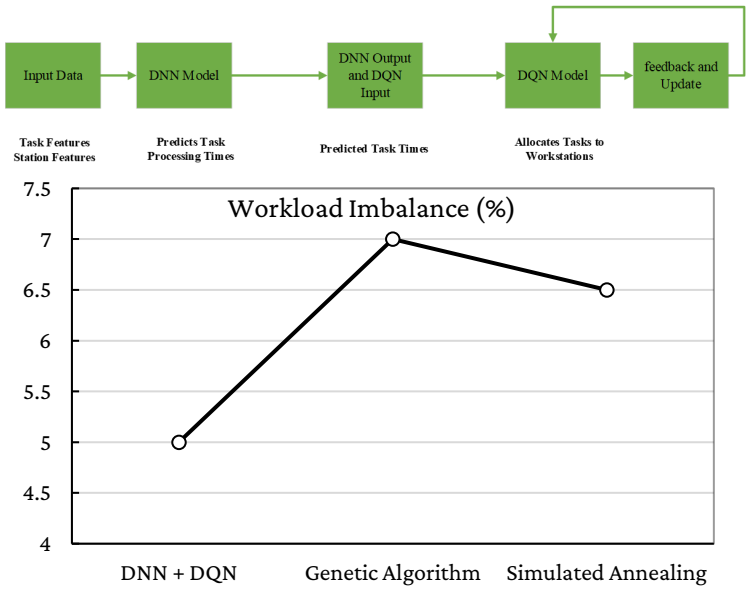
Fahimeh Tanhaie

Assistant Professor., Department of Industrial Engineering, Faculty of Basic Science and Engineering, Kosar University of Bojnord, Bojnord, Iran

HIGHLIGHTS

- Integrating deep learning and reinforcement learning to simultaneously solve balancing and sequencing problems in mixed-model assembly lines.
- Presenting a deep Q-network (DQN)-based model that dynamically and in real-time allocates tasks to workstations.
- Utilizing deep neural networks (DNNs) to accurately predict task processing times and facilitate intelligent task transfers between stations.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Article Type: Research paper
Received: 9 November 2025
Revised: 2 January 2026
Accepted: 31 January 2026
Available online: 7 February 2026
*Correspondence:
fahimeh.tanhaie@kub.ac.ir
How to cite this article:
Tanhaie, F. (2026). Reinforcement learning framework for workload distribution and scheduling in multi-model production lines. *System Engineering and Productivity*, 6 (3), 287-301.

Keywords:
Scheduling
Workload Distribution
Deep Neural Network (DNN)
Reinforcement Learning (RL)
Line Balancing

ABSTRACT

This research presents a novel approach for solving the balancing and sequencing problem in mixed-model assembly lines by leveraging the power of artificial intelligence. The proposed multi-model assembly line environment utilizes deep neural networks and reinforcement learning to properly consider both task balancing and sequencing. Production data—including task completion times, the total number of workstations, and precedence relationships between tasks—were collected and modeled. A reinforcement learning model was employed to learn the optimal task sequence and allocate them to workstations in real-time, aiming to achieve the shortest completion time and maximum assembly line efficiency. Furthermore, deep neural networks were used to predict the required processing times for tasks and to determine how tasks should be transferred between stations. Numerical evaluation based on real-world data demonstrates that the new method effectively reduces idle time, minimizes task waiting time, and improves production flow.

2538-2195/© 2026 The Authors. Published by University of Eyvanekey. This is an open-access article under the terms and conditions of the Creative Commons Attribution (CC BY 4.0) license.

1. Introduction

Mixed-model assembly lines (MMALs) play a crucial role in modern manufacturing systems by enabling the simultaneous production of multiple product variants on a shared assembly line, thereby enhancing flexibility and responsiveness to market demands. Two core decision problems in MMALs are assembly line balancing and task sequencing. Assembly line balancing focuses on assigning tasks to workstations such that workload distribution is as even as possible, while sequencing determines the optimal order of product models and tasks to maintain a continuous and efficient production flow. Although these problems are highly interdependent, most existing studies address them separately, which limits their applicability in complex and dynamic production environments (Boysen et al., 2009; Razali et al., 2019).

Previous research has predominantly applied heuristic and metaheuristic optimization techniques, such as genetic algorithms, simulated annealing, and tabu search, to solve the mixed-model assembly line balancing problem. Despite their effectiveness in generating near-optimal solutions, these methods often suffer from high computational cost, parameter sensitivity, and limited adaptability to dynamic and stochastic production conditions (Zhang et al., 2023).

2. Methodology

The DNN component serves as a predictive model to estimate task processing times at different workstations based on historical production data, task precedence relationships, and station characteristics. These predictions provide real-time decision support and enhance the system's responsiveness to dynamic production conditions. The architecture of the DNN model is shown in Figure 1.

The dataset used in this study was collected from a real mixed-model automotive assembly line over a 12-month period and includes task processing times, workstation capacities, task precedence constraints, and multi-model scheduling information.

A Deep Q-Network (DQN) is employed to learn an optimal sequencing and allocation policy through interaction with the environment. The reward function is designed to penalize makespan, idle times, and workload imbalance, while encouraging smooth task flow and balanced station utilization. The data exchange and interaction between the DNN and DQN components are depicted in Figure 2.

Model performance was evaluated separately for the prediction and decision-making components. The DNN was validated using five-fold cross-validation, with RMSE, MAE, and MAPE as accuracy metrics, while the DQN was tested over multiple independent runs with different random seeds to assess stability and robustness. Performance indicators included average reward, makespan, and idle time, and statistical comparisons with traditional optimization methods such as genetic algorithms (GA) and simulated annealing (SA) were conducted using paired t-tests.

3. Results and Discussion

The experimental results demonstrate that the proposed hybrid DNN–DQN approach consistently outperforms traditional optimization methods, including Genetic Algorithms (GA) and Simulated Annealing (SA), across all evaluated performance metrics. This superiority is primarily attributed to the continuous learning capability and real-time policy updates of the DQN agent, which dynamically adapts task allocation and sequencing decisions based on actual system states. As a result, the proposed method achieves lower makespan, reduced idle time, and improved workload balance among workstations (Figure 4). A comparative summary of the overall performance of all methods is provided in Table 1.

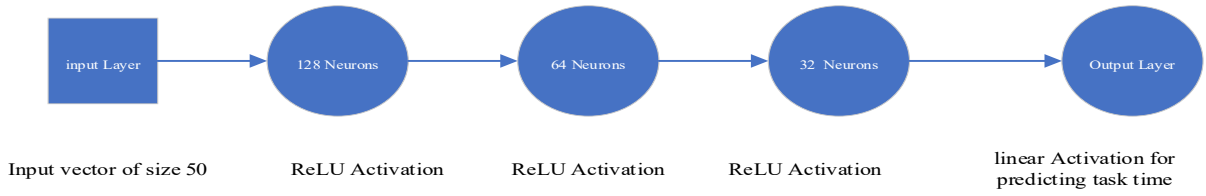


Figure 1. The architecture of the DNN.

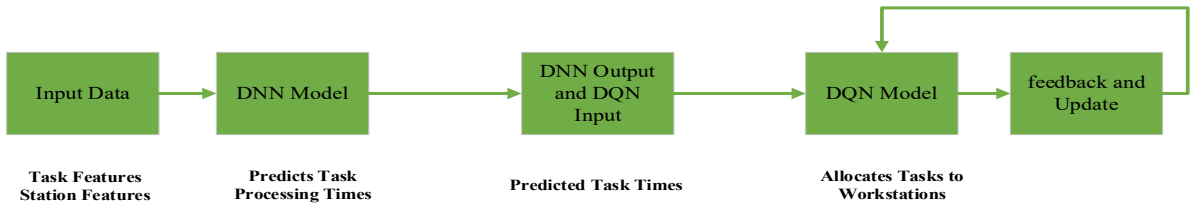


Figure 2. Flow of data between the DNN and DQN.

Table 1. Performance of the Proposed DNN and DQN Models Compared to Traditional Optimization Methods

Method	Makespan	Idle Time	Workload Imbalance
DNN + DQN	13,100	1,200	5%
(GA)	13,900	1,800	7%
(SA)	14,000	2,000	6.5%

From a methodological perspective, the hybrid framework benefits from both accurate prediction and adaptive decision-making. The DNN component provides reliable estimates of task processing times, forming a robust foundation for downstream scheduling decisions. Meanwhile, the DQN dynamically redistributes tasks to prevent workstation congestion and bottleneck formation, thereby improving line balance. Unlike GA and SA, which rely on static or offline optimization, the proposed approach operates in a fully dynamic and real-time manner, enabling higher flexibility and responsiveness to changes in the production environment. The comparison of makespan is illustrated in Figure 3.

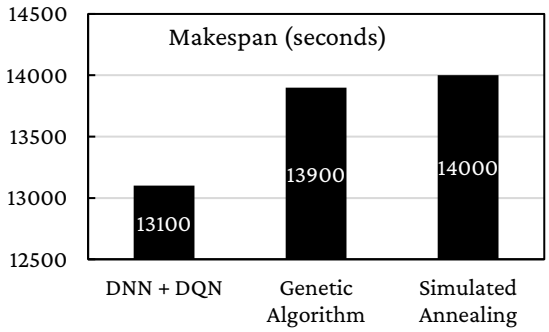


Figure 3. Comparison of makespan for different models.

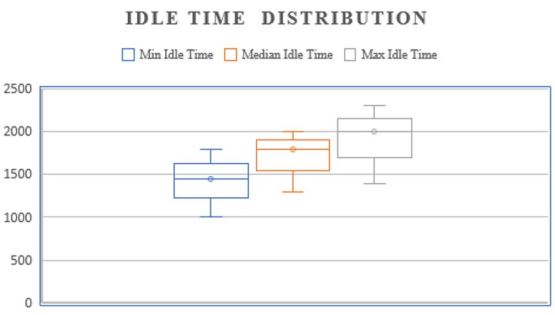


Figure 4. Idle time distribution.

4. Conclusions

This study proposes a novel integrated DNN–DQN framework for optimizing task allocation and scheduling in mixed-model assembly lines (MMALs). The DNN component accurately predicts task processing times, while the DQN agent dynamically learns optimal task sequencing and allocation policies in real time based on current production conditions. Experimental results demonstrate that the proposed approach significantly outperforms traditional optimization methods such as Genetic Algorithms (GA) and Simulated Annealing (SA), achieving a 33% reduction in idle time, a 10% reduction in makespan, and a 28% improvement in workload balance. Unlike static optimization approaches, the DNN–DQN model continuously adapts to system dynamics, enabling real-time control and enhanced flexibility.

Funding

This research received no external funding.

Author contributions

All parts of the research, including conceptualization, study design and implementation, data collection, analysis and interpretation, writing and final review of the article, were carried out by the author.

Conflicts of interest

There are no conflicts of interest associated with this research.

Acknowledgments

The author would like to thank the anonymous reviewers for their constructive comments and valuable suggestions, which helped improve the quality and clarity of this manuscript.

References

Boysen, N., Fliedner, M., & Scholl, A. (2009). Sequencing mixed-model assembly lines: Survey, classification and model critique. *European journal of operational research*, *192*(2), 349-373. https://doi.org/10.1016/j.ejor.2007.09.013

Razali, M. M., Kamarudin, N. H., Ab. Rashid, M. F. F., & Mohd Rose, A. N. (2019). Recent trend in mixed-model assembly line balancing optimization using soft computing approaches. *Engineering Computations*, *36*(2), 622-645. https://doi.org/10.1108/EC-05-2018-0205

Zhang, Z., Tang, Q., Chica, M., & Li, Z. (2023). Reinforcement learning-based multiobjective evolutionary algorithm for mixed-model multimanned assembly line balancing under uncertain demand. *IEEE Transactions on Cybernetics*, *54*(5), 2914-2927. https://doi.org/10.1109/TCYB.2022.3229666

مهندسی سیستم و بهره‌وری/ سال ۱۴۰۵/ دوره ۶/ شماره ۳/ صفحه ۲۸۷-۳۰۱

نشریه مهندسی سیستم و بهره‌وری

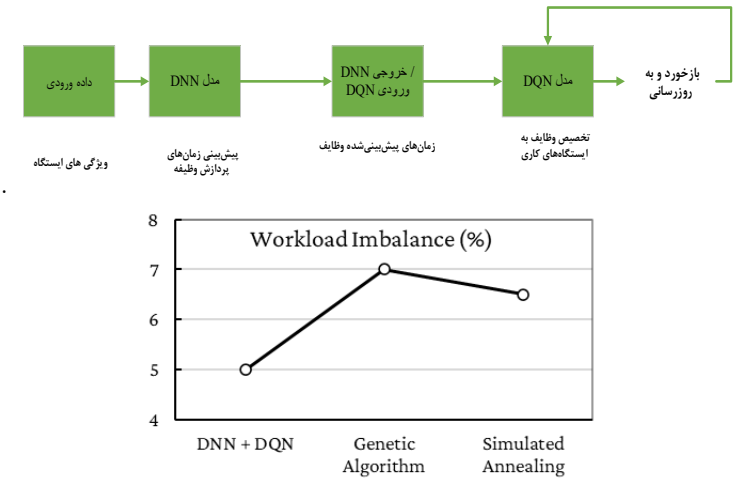
DOI: 10.22034/sep.2026.2077232.1426

چارچوبی ترکیبی مبتنی بر یادگیری تقویتی برای توزیع بار کاری و زمان‌بندی در خطوط مونتاژ مدل ترکیبی

فهیمه تنهایی

استادیار، گروه مهندسی صنایع، دانشکده فنی مهندسی، دانشگاه کوثر بجنورد، بجنورد، ایران

چکیده گرافیکی



برجسته‌ها

- ترکیب یادگیری عمیق و یادگیری تقویتی برای حل هم‌زمان مسائل تعادل‌یابی و توالی‌سازی در خطوط مونتاژ چندمدلی.
- ارائه یک مدل مبتنی بر شبکه Q عمیق (DQN) که به‌صورت پویا و در زمان واقعی وظایف را به ایستگاه‌های کاری تخصیص می‌دهد.
- استفاده از شبکه‌های عصبی عمیق (DNN) برای پیش‌بینی دقیق زمان پردازش وظایف و تسهیل جابه‌جایی هوشمند وظایف بین ایستگاه‌ها.

چکیده

این مطالعه، رویکردی نوآورانه برای مواجهه با چالش‌های دوگانه «بالانس خط» و «توالی‌سازی» در خطوط مونتاژ ترکیبی، از طریق ادغام هوش مصنوعی ارائه می‌دهد. با تلفیق شبکه‌های عصبی عمیق و یادگیری تقویتی، چارچوب پیشنهادی به‌طور هم‌زمان توزیع بار کاری ایستگاه‌ها و ترتیب اجرای وظایف را بهینه می‌سازد. داده‌های تاریخی تولید شامل زمان‌های پردازش وظایف، تعداد کل ایستگاه‌ها و وابستگی‌های بین وظایف، جمع‌آوری و در قالب مدل‌های پیش‌بینی‌کننده ساختاردهی شدند. یک عامل یادگیری تقویتی مبتنی بر شبکه **Q** عمیق توسعه داده شد تا وظایف را به‌طور پویا و در زمان واقعی به ایستگاه‌ها تخصیص دهد و ترتیب اجرای آن‌ها را مشخص کند، با هدف حداقل‌سازی زمان کل تولید و حداکثرسازی کارایی کلی خط. به‌موازات، از شبکه‌های عصبی عمیق برای پیش‌بینی مدت زمان پردازش وظایف و ارزیابی امکان جابه‌جایی وظایف بین ایستگاه‌ها استفاده شد. آزمایش‌های عددی با استفاده از داده‌های واقعی تولید نشان می‌دهد که روش پیشنهادی به‌طور چشمگیری زمان بیکاری را کاهش می‌دهد، دوره‌های انتظار وظایف را کوتاه می‌کند و تداوم جریان کار را بهبود می‌بخشد. افزون بر این، مقایسه با روش‌های بهینه‌سازی مرسوم مانند الگوریتم‌های ژنتیک و تَبرید شبیه‌سازی‌شده مزایای این راهبرد مبتنی بر یادگیری ماشین را برجسته می‌سازد؛ به‌ویژه در دستیابی به راه‌حل‌های نزدیک به بهینه با سرعت بیشتر و انعطاف‌پذیری بالاتر.

مشخصات مقاله

تاریخچه مقاله:

نوع مقاله: پژوهشی

دریافت: ۱۴۰۴/۰۸/۱۸

بازنگری: ۱۴۰۴/۱۰/۱۲

پذیرش: ۱۴۰۴/۱۱/۱۱

ارائه برخط: ۱۴۰۴/۱۱/۱۸

*نویسنده مسئول:

fahimeh.tanhaie@kub.ac.ir

کلیدواژه‌ها:

زمان‌بندی

توزیع بار کاری

شبکه‌های عصبی عمیق

یادگیری تقویتی

بالانس خط

۲۱۹۵-۲۵۳۸/© ۲۰۲۶ نویسندگان. منتشرشده توسط دانشگاه ایوان‌کی. این یک مقاله با دسترسی آزاد تحت شرایط و ضوابط مجوز Creative Commons Attribution (CC BY 4.0) است.

۱- مقدمه

مسئله طراحی و بهره‌برداری از خطوط مونتاژ در شرکت‌های تولیدی از اهمیت بالایی برخوردار است. خطوط مونتاژ مدل ترکیبی که قابلیت تولید هم‌زمان انواع مختلف یک محصول (مانند مدل‌های گوناگون خودرو) را دارند، نقشی کلیدی در پاسخ‌گویی به نیاز بازارهای پویا و مشتریان خواهان محصولات متنوع و سفارشی ایفا می‌کنند. در این نوع خطوط، مدل‌های مختلف یک محصول به‌صورت هم‌زمان و در یک خط تولید مشترک مونتاژ می‌شوند، به‌گونه‌ای که جریان تولید بدون وقفه و با کارایی بالا ادامه یابد. ازاین‌رو، دو مسئله اساسی یعنی بالانس خط مونتاژ و توالی‌سازی وظایف اهمیت ویژه‌ای در عملکرد بهینه این سیستم‌ها دارند. هدف اصلی در تعادل‌یابی خط مونتاژ، تقسیم فعالیت‌ها میان ایستگاه‌های کاری به‌گونه‌ای است که زمان صرف‌شده در هر ایستگاه تا حد امکان متعادل باشد. در مقابل، توالی‌سازی وظایف به تعیین ترتیب بهینه اجرای عملیات در خط تولید مربوط می‌شود. مدیریت هم‌زمان و کارآمد این دو جنبه، برای دستیابی به بیشترین بهره‌وری همراه با کمترین هزینه، امری حیاتی به شمار می‌آید.

روش‌های پیشین برای حل مسئله تعادل‌یابی خط مونتاژ چندمدلی عمدتاً بر پایه رویکردهای بهینه‌سازی ابتکاری و فرا ابتکاری مانند الگوریتم ژنتیک (GA)، شبیه‌سازی تبرید (SA) و جستجوی ممنوعه (Tabu Search) توسعه یافته‌اند. هرچند این روش‌ها در یافتن پاسخ‌های نزدیک به بهینه عملکرد قابل‌توجهی داشته‌اند، اما معمولاً زمان‌بر بوده، وابستگی بالایی به تنظیم پارامترها دارند و در مواجهه با پیچیدگی و عدم‌قطعیت بالای سیستم‌های واقعی تولید با چالش روبه‌رو می‌شوند. افزون بر این، بیشتر این رویکردهای کلاسیک ماهیتی ایستا (static) دارند و قادر به در نظر گرفتن پویایی، تغییرپذیری و رفتار تصادفی سیستم‌های تولیدی مدرن نیستند. در سال‌های اخیر، رویکردهای نوین یادگیری ماشین (Machine Learning) نظیر یادگیری عمیق (Deep Learning) و یادگیری تقویتی (Reinforcement Learning) به‌عنوان افق‌هایی نو در حل مسائل پیچیده بهینه‌سازی در صنعت تولید مطرح شده‌اند. شبکه‌های عصبی عمیق (Deep Neural Networks; DNNs) به‌دلیل توانایی بالای خود در پیش‌بینی دقیق نتایج بر پایه داده‌های تاریخی موردتوجه قرار گرفته‌اند، درحالی‌که یادگیری تقویتی در قالب شبکه Q عمیق به مدل امکان می‌دهد در یک محیط پویا و تعاملی، استراتژی‌های تصمیم‌گیری بهینه را از طریق تجربه و بازخورد بیاموزد.

این رویکردهای یادگیری ماشین با هدف غلبه بر محدودیت‌های روش‌های سنتی بهینه‌سازی معرفی شده‌اند، زیرا می‌توانند تصمیمات تطبیقی و هوشمندانه را در طول فرآیند تولید اتخاذ کنند و بدین‌ترتیب با پیچیدگی، پویایی و تنوع خطوط مونتاژ چندمدلی سازگار شوند. در این پژوهش چهارچوبی ارائه می‌شود که از پتانسیل شبکه‌های عصبی عمیق (DNN) برای پیش‌بینی میزان تلاش موردنیاز در فرآیند بالانس و از توانایی یادگیری تقویتی (RL) برای مدیریت توالی‌سازی وظایف در خطوط مونتاژ ترکیبی بهره می‌گیرد. بر این اساس، جدول زمانی تولید بر مبنای زمان‌های پردازش استخراج‌شده و الگوریتم یادگیری تقویتی مبتنی بر DQN به‌عنوان سازوکاری برای تخصیص وظایف و زمان‌بندی پویا به کار گرفته می‌شود (Abolghasemina et al., 2024؛ Montazeri Najafabadi & Shams, 2026؛ Kashanian Monfared et al., 2025).

لازم به تأکید است که مدل‌های موجود برای خطوط مونتاژ مدل ترکیبی عمدتاً روی بالانس یا توالی‌سازی به‌صورت مجزا تمرکز دارند و کارهایی که این دو مسئله را در قالب یک چارچوب یکپارچه و تطبیقی حل کنند، هنوز محدود هستند. نتایج این پژوهش نشان می‌دهد که ادغام یک مدل پیش‌بینی‌کننده (DNN) با یک عامل یادگیری تقویتی (DQN) می‌تواند به‌عنوان یک مسیر نویدبخش برای توسعه مدل‌های ترکیبی بالانس–توالی‌سازی در محیط‌های تولید واقعی مطرح شود. اهداف اصلی این پژوهش دو مورد است: کمینه‌سازی زمان کل تکمیل کار (Makespan)، یعنی کوتاه‌ترین زمان ممکن برای تولید تمامی محصولات در سیستم و افزایش کارایی کلی خط تولید از طریق کاهش زمان‌های بیکاری و بهینه‌سازی خطوط بحرانی ناهمگون.

۲- مرور ادبیات

خطوط مونتاژ از الگوی تک‌مدلی که تنها یک محصول را به‌صورت انبوه تولید می‌کرد، به خطوط چندمدلی و

درنهایت به خطوط مونتاژ مدل ترکیبی (MMALs) تکامل یافته‌اند. در خطوط تک‌مدلی، تخصیص وظایف برای تولید یک محصول مشخص نسبتاً ساده بود و هدف اصلی بر کاهش زمان چرخه و زمان‌های بیکاری متمرکز می‌شد (Boysen et al., 2009). با ظهور خطوط چندمدلی، چالش جدیدی در زمینه ترکیب مدل‌های مختلف در یک خط تولید مشترک مطرح شد که نیازمند الگوریتم‌های پیشرفته توالی‌سازی بود. برای نمونه، کوچوکو (Kucukko et al., 2019) از الگوریتم کلونی زنبورعسل مصنوعی برای بهبود توزیع وظایف میان مدل‌های مختلف استفاده کرد.

در خطوط مدل ترکیبی (MMALs)، چندین مدل به‌طور هم‌زمان و بدون نیاز به توقف جهت تنظیم مجدد تجهیزات (retooling) تولید می‌شوند. در این نوع خطوط، دو مسئله تعادل‌یابی و توالی‌سازی به‌شدت به یکدیگر وابسته‌اند، زیرا نوسانات تقاضا و زمان‌های راه‌اندازی وابسته به مدل این دو فرایند را به هم پیوند می‌دهند. در سال‌های اخیر، رویکردهای پیشرفته‌تری همچون الگوریتم‌های تکاملی چندهدفه مبتنی بر یادگیری تقویتی (Zhang et al., 2024) و الگوریتم کلونی زنبورعسل چندهدفه (Saif et al., 2019) توسعه یافته‌اند که هدف آن‌ها کاهش هزینه‌ها، کمینه‌سازی زمان تکمیل کل (makespan) و افزایش انعطاف‌پذیری تولید در شرایط تقاضای نامطمئن است.

۲-۱- بالانس و توالی‌سازی در خطوط مدل ترکیبی

دو چالش اصلی در MMALs عبارت‌اند از: تعادل‌یابی بار کاری میان ایستگاه‌های کاری و توالی‌سازی مدل‌ها به‌گونه‌ای که از ایجاد گلوگاه‌ها و بیکاری در خط جلوگیری شود. جرنارد هانان و همکاران (Janardhanan et al., 2019) این مسئله را در بستر خطوط مونتاژ رباتیک بررسی کردند، جایی که تخصیص بهینه ربات‌ها نقشی کلیدی ایفا می‌کند.

دفرشا و محبی‌زاده (Defersha & Mohebalizadehgashti, 2018) یک الگوریتم ژنتیک دومرحله‌ای برای خطوط مونتاژ دستی ارائه کردند که از تکرار وظایف و طولانی‌شدن زمان ایستگاه‌ها جلوگیری می‌کرد. ژانگ و همکاران (Zhang et al., 2023) نیز در شرایط تقاضای نامطمئن از یادگیری تقویتی ترکیب‌شده با الگوریتم‌های تکاملی بهره گرفتند تا هزینه‌های تولید و جریمه‌های تأخیر را کاهش دهند.

۲-۲- روش‌های سنتی بهینه‌سازی

در دو دهه گذشته، الگوریتم‌های ابتکاری و فرا ابتکاری ابزارهای اصلی در حل مسائل بالانس و توالی‌سازی خطوط مونتاژ بوده‌اند. به‌عنوان نمونه، (Akpınar et al., 2011) از الگوریتم ژنتیک ترکیبی (HGA) برای کاهش زمان‌های بیکاری استفاده کردند. تانگ و همکاران (Tang et al, 2016) نیز از شبیه‌سازی تبرید (SA) برای مدل‌سازی زمان‌های راه‌اندازی وابسته به توالی بهره بردند.

۲-۳- کاربرد یادگیری ماشین در بالانس و توالی‌سازی

در سال‌های اخیر، یادگیری ماشین (ML)، به‌ویژه یادگیری تقویتی (RL) و یادگیری عمیق (DL)، به‌عنوان راه‌حلی پویا، تطبیقی و داده‌محور برای این مسائل مطرح شده است. آیوق و خورشیدوند (Ayough & Khorshidvand, 2023) ترکیبی از شبکه عصبی عمیق (DNN) و الگوریتم کرم شب‌تاب (Dragonfly Algorithm) را برای کاهش مصرف انرژی و زمان اجرای وظایف ارائه کردند و نشان دادند که DNN می‌تواند در پیش‌بینی زمان پردازش وظایف عملکرد دقیقی داشته باشد.

تنهایی (Tanhaie, 2024) از الگوریتم چندهدفه ازدحام ذرات بهبودیافته (AMOPSO) برای خطوط انسان–ربات استفاده کرد و با بهره‌گیری از یادگیری تطبیقی، عملکرد هم‌زمان تعادل و توالی را بهبود بخشید. صیف و همکاران (Saif et al., 2019) با استفاده از الگوریتم کلونی زنبورعسل چندهدفه موفق به کاهش makespan و نوسان مصرف مواد شدند. رویکردهای ترکیبی: یادگیری ماشین و روش‌های ابتکاری ادغام قابلیت‌های یادگیری ماشین با الگوریتم‌های کلاسیک بهینه‌سازی، موجب افزایش چشمگیر کارایی مدل‌ها شده است. تانگ و زمزم (Taha & Zamzam, 2022) یک الگوریتم ابتکاری دومرحله‌ای متشکل از الگوریتم ژنتیک (GA) و روش‌های

ساختاریافته را برای خطوط چندنفره ارائه داد. سیل و کی‌زیلا (Çil & Kizila, 2020) نیز با ترکیب الگوریتم ازدحام ذرات (PSO) و یک الگوریتم ابتکاری ساختاریافته، توانستند هزینه‌های ایستگاه‌ها و مصرف منابع را کاهش دهند؛ این روش نسبت به جستجوی ممنوعه (Tabu Search) و الگوریتم فاخته (Cuckoo Algorithm) عملکرد بهتری نشان داد.

درمجموع، با وجود موفقیت روش‌های کلاسیک نظیر الگوریتم‌های ژنتیک (GA) و بهینه‌سازی کلونی مورچگان (ACO)، ادغام یادگیری ماشین و هوش مصنوعی تطبیقی در حال تحول بنیادین در بهینه‌سازی خطوط مونتاژ چندمدلی ترکیبی است. این رویکردها نه‌تنها موجب افزایش کارایی و انعطاف‌پذیری سیستم‌های تولید می‌شوند، بلکه پاسخ‌گویی مؤثر به نوسانات و پویایی تقاضای بازار را نیز امکان‌پذیر می‌سازند.

۳- اهداف و چارچوب روش‌شناختی

در این بخش، رویکرد روش‌شناختی مورد استفاده برای پاسخگویی به چالش‌های بالانس و توالی‌سازی در خطوط مونتاژ مدل ترکیبی (MMALs) تشریح می‌شود. هدف اصلی این پژوهش، شناسایی الگوی تخصیص وظایف (بین شغل‌ها و عملیات‌ها) و بهینه‌سازی جریان کار در یک خط مونتاژ ترکیبی است تا زمان کل تکمیل (makespan) کمینه و کارایی خط تولید بیشینه شود.

برای این منظور، چارچوبی نوین مبتنی بر ترکیب شبکه‌های عصبی عمیق (DNN) و یادگیری تقویتی (RL) ارائه شده است که امکان تخصیص و زمان‌بندی پویای وظایف را در بستر شرایط واقعی تولید فراهم می‌کند.

DNN به‌عنوان یک مدل پیش‌بینی‌کننده برای برآورد زمان پردازش وظایف در ایستگاه‌های مختلف استفاده می‌شود. این مدل با آموزش بر روی داده‌های تاریخی تولید، وابستگی‌های بین وظایف، توانمندی‌های ایستگاه‌ها و محدودیت‌های تولید را یاد می‌گیرد و زمان وظایف را با دقت بالا پیش‌بینی می‌کند. این قابلیت، تصمیم‌گیری را در زمان واقعی بهبود می‌بخشد.

از سوی دیگر، یادگیری تقویتی مبتنی بر شبکه Q عمیق (DQN) برای مدیریت پویای توالی و تخصیص وظایف به کار گرفته شده است. این مدل، خط مونتاژ را به‌عنوان یک سیستم پویا در نظر می‌گیرد و با تعامل مستمر با محیط تولید، از طریق دریافت پاداش یا جریمه، توالی بهینه وظایف را یاد می‌گیرد. تابع پاداش طوری طراحی‌شده که makespan را کاهش دهد، زمان بیکاری را به حداقل برساند و بار کاری را به‌صورت عادلانه بین ایستگاه‌ها توزیع کند. ترکیب این دو روش یادگیری ماشین فراتر از روش‌های سنتی مانند الگوریتم ژنتیک (GA) و شبیه‌سازی تبرید (SA) عمل می‌کند؛ چراکه از انعطاف‌پذیری بیشتری در مواجهه با سیستم‌های تولید تصادفی (SPS) برخوردار است، همگرایی سریع‌تری دارد و قادر به واکنش بلادرنگ به تغییرات محیطی است. در بخش‌های بعدی، جزئیات پیش‌پردازش داده‌ها، ساختار مدل، فرآیند آموزش و معیارهای ارزیابی ارائه می‌شود. همچنین، نتایج شبیه‌سازی بر روی داده‌های واقعی تولید و مقایسه با روش‌های بهینه‌سازی سنتی گزارش خواهد شد.

۳-۱- جمع‌آوری و پیش‌پردازش داده‌ها

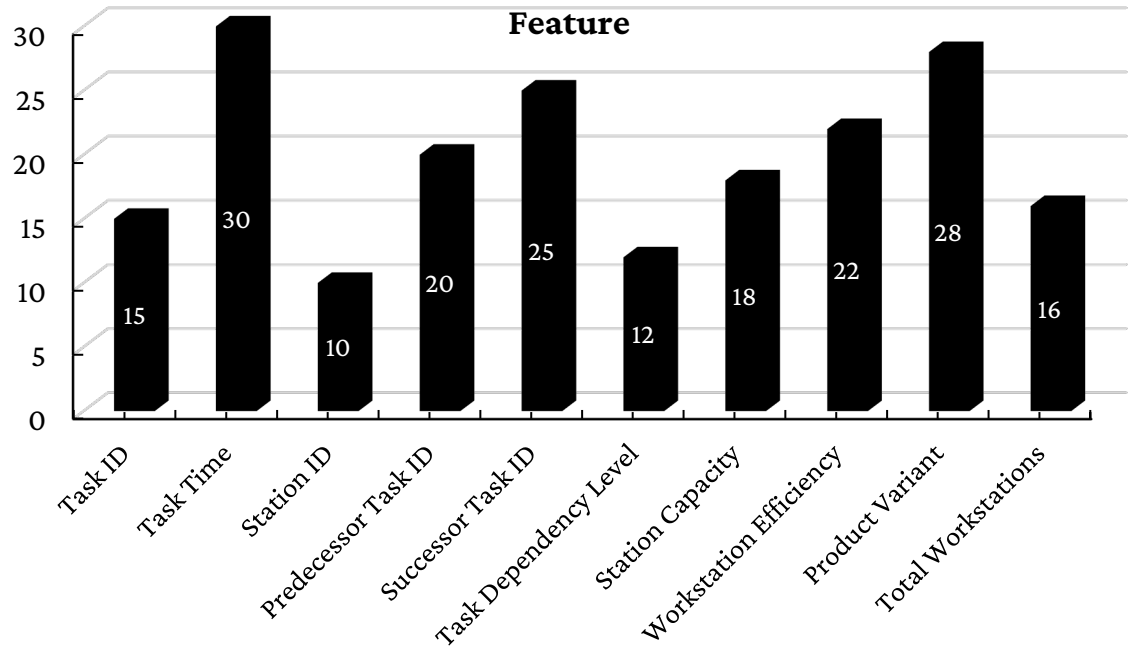
داده‌های این پژوهش از یک خط مونتاژ مدل ترکیبی در صنعت خودروسازی استخراج شده‌اند و شامل اطلاعات تاریخچه تولید طی یک دوره ۱۲ ماهه است. این داده‌ها، ویژگی‌هایی مانند زمان پردازش وظایف، تعداد و ظرفیت ایستگاه‌ها، روابط پیش‌نیازی بین وظایف و انعطاف‌پذیری زمان‌بندی در محیطی که چندین محصول روی یک خط تولید مونتاژ می‌شوند را منعکس می‌کنند. دو مدل اصلی این خط تولید با روابط پیش‌نیازی در جدول **۱** دیده می‌شود.

مجموعه داده شامل ۱۰ ویژگی کلیدی است که برای فرآیند بالانس و توالی‌سازی ضروری هستند. در این پژوهش، به‌جای تمرکز بر رتبه‌بندی تک‌تک ویژگی‌ها، از روش پیچیده (wrapper method) و الگوریتم حریصانه پیش‌رو (greedy forward) برای انتخاب زیرمجموعه‌هایی از ویژگی‌ها استفاده شد که بیشترین تأثیر را بر دقت پیش‌بینی (با معیار RMSE) داشته باشند. شبکه عصبی عمیق (DNN) برای پیش‌بینی زمان انجام وظایف به کار رفت. شبکه Q عمیق (DQN) برای یادگیری توالی بهینه وظایف در زمان واقعی استفاده شد. شکل **۱** نشان می‌دهد که کدام ویژگی‌ها در پیش‌بینی تخصیص وظایف یا زمان توالی‌سازی، بیشترین نقش را داشته‌اند.

جدول ۱. داده‌های مدل خط تولید و روابط پیش‌نیازی

Table 1. Production Line Model Data and Task Precedence Relationships

وظایف	زمان مدل ۱	روابط	زمان مدل ۲	روابط
۱	۱۲	_	۱۰۲	_
۲	۶۰	_	۳۴	۱
۳	۵۴	۱	۳۶	۱
۴	۲۴	۱	۳۲	۲
۵	۵۴	۱	۳۳	۳
۶	۴۸	۲	۲۰	۴/۵
۷	۶	۴	۱۳۰	۴/۵
۸	۶	۵	۲۱	۶
۹	۴۲	۵	۳۶	۸
۱۰	۷۲	۶/۷	۷۸	۷
۱۱	۶۰	۸/۹	۲۰	۷/۱۰
۱۲	۴۸	۱۰	۷۸	۷/۹
۱۳	۲۴	۱۰	۲۱	۱۱
۱۴	۱۲	۱۱	۳۶	۱۵/۱۲
۱۵	۳۶	۱۲/۱۳	۳۶	۱۳
۱۶	۳۶	۱۴	۲۴	۱۴
۱۷	۹۰	۱۵/۱۶	_	۱۵/۱۶
۱۸	۱۸	۱۷	_	_
۱۹	۴۸	۱۷	_	_
۲۰	۵۴	۱۸/۱۹	_	_



شکل ۱. فراوانی هر ویژگی.

Figure 1. frequency of each feature.

برای بهبود فرآیند تخصیص و توالی وظایف در یک خط مونتاژ چندمدلی، این مقاله ماتریس شباهت وظایف را پیشنهاد می‌دهد. این ماتریس، شباهت بین وظایف را بر اساس سه ویژگی کلیدی زمان انجام وظیفه، سطح وابستگی (روابط پیش‌نیازی) و نیاز به ظرفیت ایستگاه محاسبه می‌کند.

به‌کمک این ماتریس، مدل شبکه Q عمیق (DQN) می‌تواند وظایف دارای ویژگی‌های مشابه را در زمان واقعی به ایستگاه‌های مناسب تخصیص دهد تا از بیکاری کارگران جلوگیری شود و کارایی کل سیستم افزایش یابد.

در این ماتریس شباهت بین کار i و کار j همیشه برابر با شباهت بین کار j و کار i است (فرمول ۱ و ۲) و میزان شباهت کلی یک کار هدف t با سایر کارها با جمع کردن مقادیر شباهت آن با بقیه کارها در فرمول ۳ محاسبه می‌شود. شبکه عصبی عمیق (DNN) از این ماتریس برای پیش‌بینی دقیق‌تر زمان انجام وظایف مشابه استفاده می‌کند.

$$d_\{ij\} = d_\{ji\} \quad (۱)$$

$$d_{ij} \le d_{ik} + d_{kj} \quad \text{for all } k \quad (۲)$$

$$S(t) = \sum s\,(t,i) = \frac{1}{N_i}\sum_{j=t} s\,(j,i) \quad (۳)$$

شبکه عصبی عمیق (DNN) برای پیش‌بینی زمان انجام وظایف، شبکه Q عمیق (DQN) برای یادگیری تقویتی در تخصیص و توالی وظایف. این دو مدل به‌صورت هم‌زمان و مکمل یکدیگر، چالش‌های بالانس و توالی‌سازی در خط مونتاژ چندمدلی را هدف قرار می‌دهند.

لایه ورودی: برداری با ۵۰ ویژگی (ترکیبی از داده‌های one-hot و نرمال‌شده)

لایه‌های پنهان (سه لایه متراکم)؛ لایه ۱: ۱۲۸ نورون با تابع فعال‌سازی ReLU

لایه ۲: ۶۴ نورون با ReLU

لایه ۳: ۳۲ نورون با ReLU

لایه خروجی: ۱ نورون با فعال‌سازی خطی (برای رگرسیون)

تنظیمات آموزش: بهینه‌ساز: Adam

نرخ یادگیری: ۰/۰۰۱

اندازه دسته (Batch size): ۳۲

تعداد دوره‌ها (Epochs): ۱۰۰

تابع خطا میانگین مربع خطا (MSE) است که در فرمول ۴ محاسبه شده است.

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (۴)$$

ساختار کامل DNN در شکل **۲** نمایش داده شده است.

خط مونتاژ مدل ترکیبی به‌صورت یک فرایند تصمیم‌گیری مارکوفی مدل‌سازی شده است. هر حالت در زمان t باs_t نشان داده می‌شود و شامل اطلاعات زیر است: وضعیت هر ایستگاه کاری (مشغول/آزاد بودن، زمان باقی‌مانده تا اتمام وظیفه جاری)، مجموعه وظایف مجاز با توجه به روابط پیش‌نیازی (یعنی وظایفی که تمام پیش‌نیازهای آن‌ها انجام شده است)، زمان‌های پردازش پیش‌بینی‌شده برای هر زوج وظیفه؛ ایستگاه توسط مدل DNN، شاخص‌های موقت بار کاری و میزان بیکاری در سطح خط مونتاژ، اطلاعات مربوط به مدل محصول (نوع مدل در خطوط چندمدلی) و وضعیت فعلی توالی.

فضای اقدام A به‌صورت مجموعه‌ای از تصمیمات تخصیص وظیفه مجاز به یک ایستگاه مشخص تعریف می‌شود. به‌طور مشخص، هر اقدام متناظر با انتخاب یک وظیفه از میان وظایف مجاز و تخصیص آن به یکی از ایستگاه‌های موجود است. اقداماتی که با قیود سیستم (مانند پیش‌نیازی، ظرفیت ایستگاه یا محدودیت‌های زمان‌بندی) ناسازگار باشند، با استفاده از مکانیزم Action Masking از فضای اقدام حذف می‌شوند تا عامل تنها از بین اقدامات معتبر انتخاب کند. تابع پاداش به‌گونه‌ای طراحی شده است که کاهش زمان کل تکمیل، کاهش زمان‌های بیکاری و بهبود تعادل بار کاری را تشویق کرده و برای ایجاد گلوگاه‌ها، بیکاری زیاد یا توالی‌های نامطلوب جریمه در نظر بگیرد.

به‌روزرسانی مقادیر Q و تابع پاداش: مقادیر Q با استفاده از معادله بلمن مانند فرمول ۵ به‌روزرسانی می‌شوند.

$$Q(s,a) = r + \gamma\, max_{\ a'}\, Q(s',a') \quad (۵)$$

تابع پاداش (Reward Function): تابع پاداش در فرمول ۶ به‌گونه‌ای طراحی‌شده که توزیع مناسب وظایف و توالی بهینه اجرای آن‌ها را تشویق کند و از وجود شکاف‌ها و عدم تعادل در خط تولید جلوگیری نماید.

$$r_t = -\text{makespan} - \lambda(\text{idle time}) + \mu(\text{task balance}) \quad (۶)$$

پیش‌بینی توسط DNN: برای هر وظیفه، زمان پردازش بر اساس شرایط فعلی وظیفه و ایستگاه برآورد می‌شود.

تخصیص توسط DQN: DQN از این زمان‌های پیش‌بینی‌شده و وضعیت فعلی خط مونتاژ برای تخصیص هوشمند وظایف به ایستگاه‌ها استفاده می‌کند.

بازخورد و به‌روزرسانی: پس از هر تخصیص، پاداش محاسبه می‌شود، حالت سیستم به‌روز می‌شود و مدل DQN بر اساس پاداش دریافتی تنظیم می‌گردد (بدون تغییر در سیاست کلی).

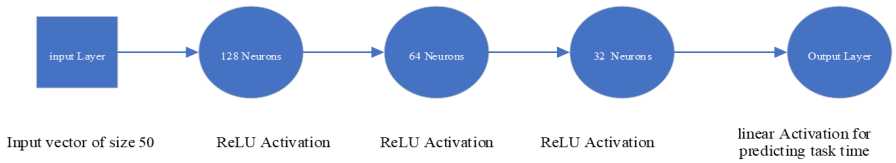
تبادل داده بین DNN و DQN در شکل **۳** نمایش داده شده است .برای ارزیابی دقت مدل‌های پیشنهادی، فرآیند اعتبارسنجی به‌صورت جداگانه برای مؤلفه‌های DNN و DQN انجام شد. در بخش پیش‌بینی زمان انجام وظایف، مدل DNN با استفاده از روش اعتبارسنجی متقاطع پنج‌تایی آموزش داده شد و شاخص‌های خطا شامل RMSE، MAE و MAPE محاسبه گردید. در بخش یادگیری تقویتی، مدل DQN در ۲۰ اجرای مستقل با بذرهای تصادفی متفاوت ارزیابی شد و میانگین پاداش، میانگین زمان تکمیل و زمان بیکاری محاسبه شد. نتایج حاصل، کارایی و پایداری قابل‌قبول مدل را نشان می‌دهد.

جدول **۲** خلاصه نتایج اعتبارسنجی را نشان می‌دهد. نتایج نشان می‌دهد که مدل DNN در پیش‌بینی زمان انجام وظایف از دقت مناسبی برخوردار است و مقدار RMSE کمتر از ۱۳ ثانیه عملکرد قابل‌قبول شبکه را تأیید می‌کند. همچنین مدل DQN در ۲۰ اجرای مستقل پایداری بالایی نشان داده و تغییرات کم در معیارهای زمان تکمیل و زمان بیکاری بیانگر ثبات سیاست یادگیری است. مقادیر p کمتر از ۰/۰۵ نیز نشان می‌دهد که تفاوت عملکرد مدل پیشنهادی در مقایسه با روش‌های GA و SA ازنظر آماری معنادار است.

جدول ۲. اعتبارسنجی مدل‌های DNN و DQN

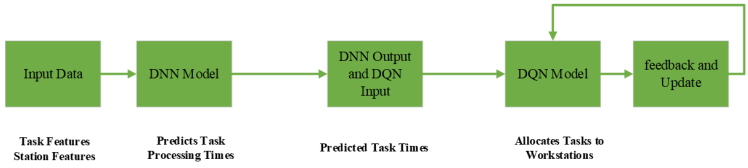
Table 2. Validation of DNN and DQN Models

مدل	(میانگین ± انحراف معیار)	شاخص اعتبارسنجی
DNN	۱۲/۴	**RMSE (ثانیه)**
DNN	۸/۷	**MAE (ثانیه)**
DNN	۶/۳	**MAPE (٪)**
DQN	۱۸۷/۵	**میانگین پاداش**
DQN	۱۳/۲۱	**زمان تکمیل (Makespan)**
DQN	۲۶۰/۱	**زمان بیکاری (Idle Time)**
مقایسه آماری	۰/۰۱۸	**مقدار p نسبت به GA**
مقایسه آماری	۰/۰۱۱	**مقدار p نسبت به SA**



شکل ۲. معماری شبکه‌ی عصبی عمیق.

Figure 2. The architecture of the DNN.



شکل ۳. جریان داده بین شبکه‌های DNN و DQN.

Figure 3. Flow of data between the DNN and DQN.

۴- نتایج پژوهش

یافته‌های این پژوهش نشان می‌دهد که روش پیشنهادی ترکیبی DNN + DQN در مقایسه با روش‌های سنتی بهینه‌سازی (مانند الگوریتم ژنتیک و شبیه‌سازی تبرید) عملکرد بهتری دارد. این برتری عمدتاً به‌دلیل یادگیری پیوسته و به‌روزرسانی بلادرنگ مدل DQN در پاسخ به تخصیص واقعی وظایف است که منجر به کاهش زمان کل تکمیل (makespan)، زمان بیکاری و بهبود تعادل بار کاری بین ایستگاه‌ها می‌شود.

انعطاف‌پذیری: مدل DQN به‌صورت پویا و در زمان واقعی تصمیم می‌گیرد، درحالی‌که روش‌های سنتی مبتنی بر مدل‌های ایستا هستند و برای محیط‌های ثابت طراحی شده‌اند.

دقت پیش‌بینی: مدل DNN با پیش‌بینی دقیق زمان انجام وظایف، پایه‌ای محکم برای تصمیم‌گیری‌های بهینه در تخصیص وظایف فراهم می‌کند.

توزیع مناسب بار کاری: DQN از تمرکز بیش‌ازحد وظایف در برخی ایستگاه‌ها جلوگیری کرده و بدین‌ترتیب گلوگاه‌ها و هدررفت زمان را کاهش می‌دهد.

درمجموع، مدل DNN+DQN ازنظر انعطاف‌پذیری، کارایی و معیارهای عملکردی، برتری روشنی نسبت به روش‌های ثابت دارد. لازم به ذکر است که ترکیب DNN و DQN اگرچه ازنظر مفهومی چندلایه است، اما ازنظر محاسباتی منجر به افزایش غیرقابل‌کنترل پیچیدگی نمی‌شود. پیچیدگی زمانی مؤلفه DNN عمدتاً وابسته به هزینه یک عبور روبه‌جلو در شبکه است که با $O(C\phi)$ نشان داده می‌شود. این مقدار با توجه به ساختار شبکه‌ها (سه لایه مخفی با تعداد نورون محدود) بسیار کوچک و ثابت است. مؤلفه DQN، بیشترین هزینه محاسباتی مربوط به انتخاب اقدام (action selection) و به‌روزرسانی Q-value است که به ترتیب برابر است با $O(|A| \times C\phi)$ و $O(B \times C\phi)$

با توجه به اینکه تعداد ایستگاه‌ها محدود است و فضای اقدام در خطوط مونتاژ نسبتاً کوچک باقی می‌ماند، ترکیب DNN و DQN ازنظر محاسباتی مقیاس‌پذیر و سبک است. در عمل نیز زمان اجرای کل سیستم برای هر اپیزود کمتر از چند ده میلی‌ثانیه بوده و تأثیر قابل‌توجهی بر کارایی عملیاتی ندارد؛ بنابراین، ترکیب DNN و DQN منجر به افزایش پیچیدگی نمی‌شود و ازنظر محاسباتی برای کاربردهای صنعتی با اندازه متوسط کاملاً مناسب است.

معیارهای ارزیابی: برای سنجش کارایی مدل‌های پیشنهادی در مقایسه با روش‌های سنتی (GA و SA)، از سه معیار اصلی بهینه‌سازی تولید استفاده شد:

۱) زمان کل تکمیل (Makespan): مجموع زمان لازم برای تکمیل تمام وظایف در یک چرخه تولید. کاهش makespan، مستقیماً نرخ تولید (throughput) را افزایش می‌دهد.

۲) زمان بیکاری (Idle Time): مجموع زمانی که ایستگاه‌ها بدون فعالیت می‌مانند توسط فرمول ۷ محاسبه گردیده است. کاهش این زمان، کارایی خط تولید را بهبود می‌بخشد و از ضرر اقتصادی ناشی از بیکاری جلوگیری می‌کند.

$$I = \sum_{k=1}^{K} \left(T - \sum_{i=1}^{n_k} t_{ik} \right) \quad (۷)$$

۳) عدم تعادل بار کاری (Task Imbalance): این شاخص مانند فرمول ۸ نشان می‌دهد که آیا بار کاری بین ایستگاه‌ها به‌صورت یکنواخت توزیع شده است یا خیر. مقدار پایین‌تر این شاخص نشان‌دهنده تعادل بهتر و کاهش گلوگاه است.

$$B = \frac{1}{K} \sum_{k=1}^{K} |W_k - \bar{W}| \quad (۸)$$

$$\bar{W} = \frac{1}{K} \sum_{k=1}^{K} W_k$$

نتایج کامل در جدول **۳** ارائه شده است.

زمان کل تکمیل (Makespan): مدل DNN + DQN با ۱۰۰/۱۳ ثانیه، نسبت به الگوریتم ژنتیک ٪10 (GA) و نسبت به شبیه‌سازی تبرید (SA) به‌طور چشمگیری عملکرد بهتری داشت (شکل **۴**).

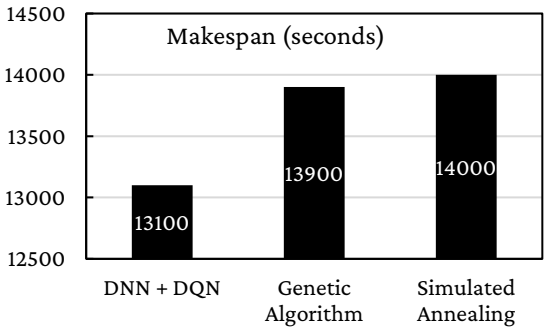
زمان بیکاری (Idle Time): زمان بیکاری در روش پیشنهادی ۲۰۰/۱ ثانیه بود که ٪۳۳ بهتر از GA و ٪۴۰ بهتر از SA است (شکل **۵**).

عدم تعادل بار کاری (Task Imbalance): میزان عدم تعادل در DNN + DQN برابر با ٪۵ گزارش شد که در مقایسه با GA (٪۷) و SA (٪۶/۵) بهبود قابل‌توجهی نشان می‌دهد (شکل **۶**).

جدول ۳. عملکرد مدل‌های پیشنهادی DNN و DQN در مقایسه با روش‌های بهینه‌سازی سنتی

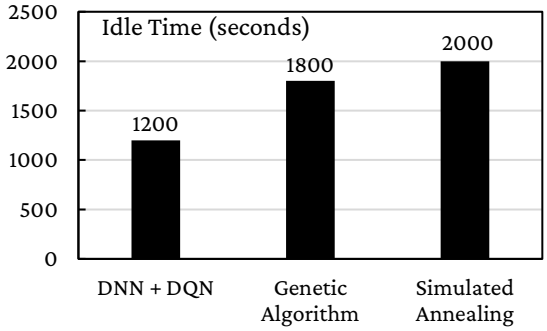
Table 3. Performance of the Proposed DNN and DQN Models Compared to Traditional Optimization Methods

	زمان کل تکمیل	زمان بیکاری	عدم تعادل بار کاری
DNN + DQN	۱۳/۱۰۰	۱/۲۰۰	٪۵
الگوریتم ژنتیک (GA)	۱۳/۹۰۰	۱/۸۰۰	٪۷
شبیه‌سازی تبرید (SA)	۱۴/۰۰۰	۲/۰۰۰	٪۶/۵



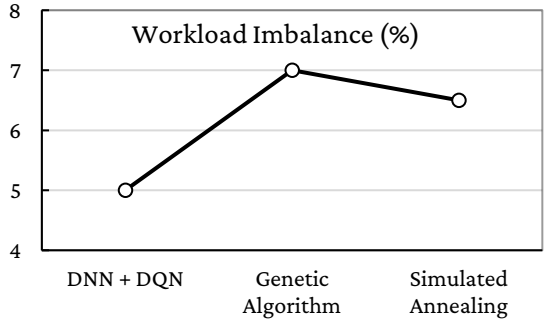
شکل ۴. مقایسه زمان تکمیل کل در مدل‌های مختلف.

Figure 4. Comparison of makespan for Different models.



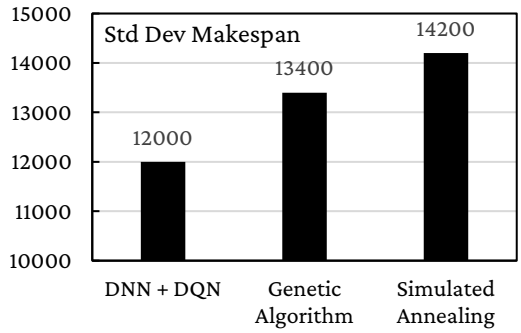
شکل ۵. مقایسه زمان بیکاری در مدل‌های مختلف.

Figure 5. Comparison of Idle Time for Different Models.



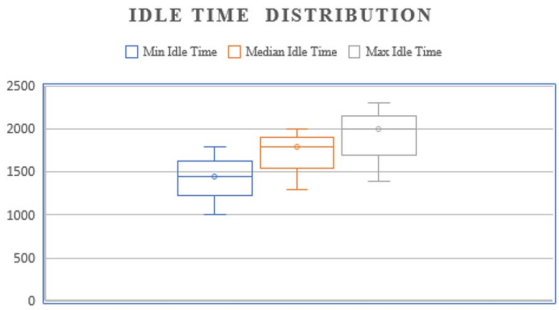
شکل ۶. مقایسه عدم‌تعادل بار کاری در مدل‌های مختلف.

Figure 6. Comparison of Workload Imbalance for Different Models.



شکل ۷. انحراف معیار زمان تکمیل کل.

Figure 7. standard deviation makespan.



شکل ۸. توزیع زمان بیکاری.

Figure 8. Idle Time Distribution.

این نتایج تأیید می‌کنند که مدل پیشنهادی به‌دلیل توانایی در واکنش پویا و تنظیم بلادرنگ توالی وظایف، عملکردی برتر نسبت به روش‌های سنتی دارد. با استفاده از تحلیل‌های آماری پیشرفته، می‌توان عمق بیشتری از عملکرد مدل پیشنهادی DNN + DQN در مقایسه با روش‌های سنتی (GA و SA) استنباط کرد:

۱) تحلیل زمان کل تکمیل (Makespan): مدل DNN + DQN نه‌تنها میانگین makespan کمتری داشت، بلکه پراکندگی (انحراف معیار) بسیار پایین‌تری نیز نشان داد (شکل **۷**). آزمون t نشان داد که تفاوت بین این مدل و روش‌های سنتی ازنظر آماری معنادار است ($p < 0.05$).

۲) تحلیل زمان بیکاری (Idle Time): نمودار جعبه‌ای در شکل **۸** نشان می‌دهد که توزیع

زمان بیکاری در مدل DNN + DQN بسیار متراکم‌تر و با داده‌های پرت (outliers) کمتری همراه است. در مقابل، روش‌های GA و SA دارای پراکندگی بالاتر و نوسان بیشتر هستند.

۳) تحلیل عدم تعادل بار کاری (Task Imbalance): مدل پیشنهادی با میانگین عدم تعادل ۵٪، نسبت به GA (۷٪) و SA (۶/۵٪) عملکرد بهتری داشت.

در این پژوهش، عامل یادگیری تقویتی بر پایه معماری DQN پیاده‌سازی شده است. برای کنترل تعادل بین اکتشاف و بهره‌برداری، از سیاست ε-greedy استفاده شده است؛ به‌طوری‌که مقدار ε در ابتدای آموزش برابر با ۰/۱ در نظر گرفته‌شده و به‌صورت خطی تا مقدار ۰/۰۵ در طول ۲۰٪ نخست اپیزودها کاهش می‌یابد و سپس ثابت می‌ماند. حافظه replay با ظرفیت ۵۰۰۰۰ انتقال تعریف‌شده و در هر گام، یک mini-batch با اندازه ۶۴ نمونه برای به‌روزرسانی شبکه انتخاب می‌شود. ضریب تنزیل پاداش‌ها γ برابر ۰/۹۹ و نرخ یادگیری بهینه‌ساز Adam برابر ۰/۰۰۱ در نظر گرفته شده است. برای بهبود پایداری آموزش، از شبکه هدف (Target Network) استفاده‌شده و پارامترهای آن هر ۱۰۰۰ گام یک‌بار با پارامترهای شبکه اصلی به‌روزرسانی می‌شوند. گرادیان‌ها نیز در بازه قابل‌قبول برش داده شده‌اند تا از ناپایداری عددی جلوگیری شود.

۵- نتیجه‌گیری

یافته‌های این پژوهش راه‌حلی نوین برای مسئله بهینه‌سازی تخصیص و زمان‌بندی وظایف در خطوط مونتاژ چندمدلی (MMALs) ارائه می‌دهد که بر پایه یک چارچوب یکپارچه DNN-DQN استوار است. رویکرد ما به‌طور دقیق زمان کل تکمیل (makespan) را کاهش می‌دهد، زمان بیکاری را کم می‌کند و یکنواختی بار کاری را افزایش می‌دهد. مؤلفه DNN، زمان لازم برای انجام هر وظیفه را با دقت بیشتری برآورد می‌کند، درحالی‌که مؤلفه DQN به‌صورت بلادرنگ، توالی و توزیع وظایف را بر اساس شرایط فعلی خط تولید یاد می‌گیرد. نتایج آزمایش‌ها نشان می‌دهد که در شرایط واقعی، چارچوب پیشنهادی DNN + DQN در مقایسه با روش‌های سنتی مانند الگوریتم ژنتیک (GA) و شبیه‌سازی تبرید (SA)، عملکرد بهتری دارد:

- کاهش ۳۳٪ زمان بیکاری
- کاهش ۱۰٪ زمان کل تکمیل (makespan)
- کاهش ۲۸٪ عدم تعادل بار کاری

این یافته‌ها، کاربردی بودن ترکیب الگوریتم‌های پیش‌بینی‌کننده و تطبیقی را در بهینه‌سازی خطوط مدل ترکیبی تأیید می‌کند و مدلی عملی برای پیاده‌سازی هوش مصنوعی در صنعت ارائه می‌دهد. کارهای آتی می‌توانند با اعمال این چارچوب به خطوط تولیدی گسترده‌تر و پیچیده‌تر، مقیاس‌پذیری و استحکام آن را ارزیابی کنند و محدودیت‌های واقعی مانند خرابی تجهیزات یا نوسان در دسترس‌پذیری نیروی انسانی را نیز در نظر بگیرند.

تخصیص ناکارا وظایف ازنظر زمان و توالی، همچنان یکی از بزرگ‌ترین چالش‌های بهبود خطوط MMAL محسوب می‌شود. با ادغام DNN برای پیش‌بینی زمان وظایف و DQN (یادگیری تقویتی) برای توالی‌سازی و تخصیص پویا، پیشرفت چشمگیری در این حوزه حاصل شده است. برخلاف روش‌های سنتی که به دنبال یک بهینه سراسری ایستا هستند، مدل DNN + DQN قادر است به‌صورت پویا یاد بگیرد و در زمان واقعی تطبیق یابد. راه‌حل ارائه‌شده کلی و قابل تعمیم است و می‌تواند در صنایع مختلفی که نیاز به تولید مدل ترکیبی دارند به‌کار گرفته شود. ازآنجاکه این مدل امکان کنترل بلادرنگ را بر اساس تغییرات محیط تولید فراهم می‌کند، مزایای عمده‌ای در ابعادی مانند کاهش هزینه تولید، افزایش نرخ تولید (throughput) و انعطاف‌پذیری سیستم به همراه دارد.

۶- فهرست علائم

dij	فاصله (عدم شباهت) بین کار i و کار j
$S(t,i)$	شباهت بین کارi و هدف t
Nt	تعداد کارهای مرتبط با کار هدف
r	پاداش فوری
γ	ضریب تنزیل
I	زمان بیکاری کل خط تولید
T	زمان کل تکمیل
tik	زمان انجام کار i در ایستگاه k

nk تعداد کارهای تخصیص‌یافته به ایستگاه k

مشارکت‌های نویسندگان

کلیه مراحل انجام پژوهش ازجمله مفهوم‌سازی، طراحی و اجرای مطالعه، گردآوری، تجزیه‌وتحلیل و تفسیر داده‌ها، نگارش و بازبینی نهایی مقاله توسط نویسنده انجام شده است.

تضاد منافع

نویسنده اظهار می‌دارد که در ارتباط با این مقاله هیچ‌گونه منافع مالی، شخصی یا سازمانی که منجر به تضاد منافع شود، وجود ندارد و نتایج پژوهش به‌صورت مستقل و بی‌طرفانه گزارش شده است.

قدردانی

نویسنده از تمامی افرادی که به‌صورت غیرمستقیم با ارائه بازخوردهای علمی و اخلاقی در بهبود این پژوهش مؤثر بوده‌اند، قدردانی می‌نماید.

مراجع

Abolghasemian, M., Kheiri, A. O., & Saberifard, N. (2024). Prioritizing factors affecting the flexibility and performance of the digital supply chain system in the Iranian food industry. *System Engineering and Productivity*, *4*(1), 41-57. https://doi.org/10.22034/msb.2024.2025240.1194

Akpınar, S., & Bayhan, G. M. (2011). A hybrid genetic algorithm for mixed model assembly line balancing problem with parallel workstations and zoning constraints. *Engineering Applications of Artificial Intelligence*, 24(3), 449–457. https://doi.org/10.1016/j.engappai.2010.08.006

Ayough, A., & Khorshidvand, B. (2023). Robust optimization for the integrated worker-cell assignment and sequencing problem in a lean U-shaped assembly line. *Computers & Industrial Engineering,* 178, 109139. https://doi.org/10.1016/j.cie.2023.109139

Boysen, N., Fliedner, M., & Scholl, A. (2009). Sequencing mixed-model assembly lines: Survey, classification and model critique. *European journal of operational research*, *192*(2), 349-373. https://doi.org/10.1016/j.ejor.2007.09.013

Çil, Z. A., & Kizilay, D. (2020). Constraint programming model for multi-manned assembly line balancing problem. *Computers & Operations Research*, 124, 105069. https://doi.org/10.1016/j.cor.2020.105069

Defersha, F. M., & Mohebalizadehgashti, F. (2018). Simultaneous balancing, sequencing, and workstation planning for a mixed model manual assembly line using hybrid genetic algorithm. *Computers & industrial engineering*, *119*, 370-387. https://doi.org/10.1016/j.cie.2018.04.014

Gholamian, S. A. (2025). Evaluation and selection of sustainable suppliers by providing a decision support system based on a new data envelopment analysis model and cumulative star utility. *System Engineering and Productivity*, *4*(1), 1-13. https://doi.org/10.22034/msb.2024.2025845.1198

Guo, G., & Ryan, S. M. (2022). Sequencing mixed-model assembly lines with risk-averse stochastic mixed-integer programming. *International Journal of Production Research*, 60 (12), 3774–3791. https://doi.org/10.1080/00207543.2021.1931978

Janardhanan, M. N., Li, Z., Bocewicz, G., Banaszak, Z., & Nielsen, P. (2019). Metaheuristic algorithms for balancing robotic assembly lines with sequence-dependent robot setup times. *Applied Mathematical Modelling*, *65*, 256-270. https://doi.org/10.1016/j.apm.2018.08.016

Kashanian Monfared, N., Safaie, N., & Hosseininezhad, S. J. (2025). A decision-making model for the problem of designing the layout of medical centers considering uncertainty. *System Engineering and Productivity*, *5*(2), 97-118. https://doi.org/10.22034/sep.2025.2049327.1252

Kucukkoc, I., Buyukozkan, K., Satoglu, S. I., & Zhang, D. Z. (2019). A mathematical model and artificial bee colony algorithm for the lexicographic bottleneck mixed-model assembly line balancing problem. *Journal of Intelligent Manufacturing*, 30, 2913–2925. https://doi.org/10.1007/s10845-015-1150-5

Li, Y., Liu, D., & Kucukkoc, I. (2023). Mixed-model assembly line balancing problem considering learning effect and uncertain demand. *Journal of Computational and Applied Mathematics*, 422, 114823. https://doi.org/10.1016/j.cam.2022.114823

Montazeri Najafabadi, R., & Shams, M. (2026). Developing a deep learning-based power outage predictive model to improve resilience of power systems. *System Engineering and Productivity*.

https://doi.org/10.22034/sep.2026.2076246.1420

Razali, M. M., Kamarudin, N. H., Ab. Rashid, M. F. F., & Mohd Rose, A. N. (2019). Recent trend in mixed-model assembly line balancing optimization using soft computing approaches. *Engineering Computations*, 36 (2), 622–645. https://doi.org/10.1108/EC-05-2018-0205

Saif, U., Guan, Z., Zhang, L., Zhang, F., Wang, B., & Mirza, J. (2019). Multi-objective artificial bee colony algorithm for order oriented simultaneous sequencing and balancing of multi-mixed model assembly line. *Journal of Intelligent Manufacturing*, 30, 1195–1220. https://doi.org/10.1007/s10845-017-1316-4

Taha, R. B., & Zamzam, N. Z. (2022). A genetic algorithm for multi-manned multi-position assembly line under technological constraint. *Port-Said Engineering Research Journal*, 26 (3), 134–145. https://doi.org/10.21608/pserj.2022.74000.1107

Tang, Q., Liang, Y., Zhang, L., Floudas, C. A., & Cao, X. (2016). Balancing mixed-model assembly lines with sequence-dependent tasks via hybrid genetic algorithm. *Journal of Global Optimization*, 65 (1), 83–107. https://doi.org/10.1007/s10898-015-0316-1

Tanhaie, F. (2024). Applying a multi-objective particle swarm optimization algorithm for sequencing and balancing a mixed-model assembly line problem with setup times between tasks. *Journal of Applied Research on Industrial Engineering*, 11 (3), 350–368. https://doi.org/10.22105/jarie.2024.394265.1545

Zhang, X., Fathollahi-Fard, A. M., Tian, G., Yaseen, Z. M., Pham, D. T., Zhao, Q., & Wu, J. (2024). Human–robot collaboration in mixed-flow assembly line balancing under uncertainty: An efficient discrete bees' algorithm. *Journal of Industrial Information Integration*, 41, 100676. https://doi.org/10.1016/j.jii.2024.100676

Zhang, Z., Tang, Q., Chica, M., & Li, Z. (2023). Reinforcement learning-based multiobjective evolutionary algorithm for mixed-model multimanned assembly line balancing under uncertain demand. *IEEE Transactions on Cybernetics*, *54*(5), 2914-2927. https://doi.org/10.1109/TCYB.2022.3229666