

Application of digital twin technology in smart warehousing and distribution network optimization

Jianing Wang

University of Calgary, Calgary, Canada

wangjianing511@gmail.com

Abstract. With the rapid development of smart logistics, the integration of digital twin technology into warehousing and distribution network optimization has emerged as a promising solution to improve operational efficiency, reduce costs, and enhance decision-making capabilities. This study explores how a digital twin framework can be constructed for smart warehousing and how it can be leveraged to optimize distribution networks. We review existing literature on digital twin foundations, smart warehousing systems, and distribution optimization methods, highlighting research gaps and practical challenges. A conceptual three-layer architecture is proposed, comprising the perception layer, data layer, and application layer, along with data flow and functional modules including warehousing simulation, dynamic routing optimization, and predictive maintenance. The discussion elaborates on the advantages, limitations, and managerial implications of the proposed framework. Future research directions include empirical validation, integration with blockchain, and applications in green logistics.

Keywords: digital twin, smart logistics, smart warehousing, distribution network optimization

1. Introduction

The rapid advancement of digital technologies, coupled with the continued growth of e-commerce, has transformed the logistics industry into a highly dynamic, data-driven ecosystem [1]. Smart logistics, characterized by the integration of automation, real-time data analytics, and intelligent decision-making, has become a key enabler for meeting increasingly demanding customer delivery demands, including faster, more reliable, and more economical delivery. Within this model, smart warehousing plays a key role, leveraging automated systems such as Automated Guided Vehicles (AGVs), Automated Storage and Retrieval Systems (AS/RS), and robotic sorting technology to streamline inbound, outbound, and outbound processes [2]. Meanwhile, distribution network optimization strives to ensure the efficient flow of goods from warehouses to end consumers through optimized location planning, route design, and dynamic scheduling.

However, despite significant technological advances, the logistics industry continues to face persistent challenges. Fragmented data systems often result in suboptimal resource utilization in warehouses, while distribution networks must adapt to unpredictable demand fluctuations, traffic congestion, and the complexities of last-mile, point-to-point delivery [3]. These challenges highlight the urgent need for an integrated approach that enables real-time situational awareness, accurate operational forecasting, and proactive decision-making across warehousing and distribution functions.

Digital twin technology has emerged as a promising solution to these challenges. Initially proposed and widely adopted in the manufacturing and aerospace industries, digital twins create high-fidelity virtual replicas of physical systems, enabling real-time monitoring, simulation, and optimization. By coupling physical assets with their digital counterparts through continuous data exchange, organizations can predict equipment failures, simulate alternative operating strategies, and rapidly respond to disruptions. In logistics, digital twins can replicate entire warehouse environments and distribution networks, enabling accurate simulation of order picking strategies, routing alternatives, and maintenance schedules—capabilities that are particularly critical for enhancing the resilience and efficiency of complex logistics operations.

Although some logistics giants, such as Amazon, JD.com, and Cainiao, have experimented with digital twin-like systems in their logistics operations, most existing applications remain limited in scope, typically confined to warehouse automation or transportation optimization [4]. Challenges such as real-time data integration, cross-system interoperability, scalability, and cost-effectiveness have hindered the development of holistic digital twin frameworks that seamlessly integrate warehousing and distribution functions. Furthermore, academic research on logistics digital twins has primarily focused on specific use-case implementations, resulting in a gap in the literature regarding scalable architectures and network optimization methods.

To address this gap, this study proposes a conceptual framework for applying digital twin technology to the optimization of smart warehousing and distribution networks. Specifically, this study aims to design a three-layer digital twin architecture that integrates IoT sensing, AI-driven data analytics, and cloud-based visualization to enable real-time operational simulation and dynamic routing decisions. This research addresses the following questions:

1. How can a digital twin system architecture for smart warehousing be constructed to achieve accurate and real-time operational simulation?
2. How can digital twin technology be used to optimize distribution networks to improve cost efficiency, delivery timeliness, and resource utilization?

2. Literature review

By addressing these questions, this study aims to develop a unified data-driven decision-making support framework for logistics companies, with potential application to large-scale e-commerce operations, third-party logistics providers, and urban distribution systems. This section will introduce some advanced technologies, their current applications in enterprises and the challenges they need to address, as well as an analysis of the deep integration relationships between these technologies and industries.

Digital Twin (DT) technology, first conceptualized in the early 2000s within the field of manufacturing and product lifecycle management, refers to a dynamic, high-fidelity virtual representation of a physical asset, process, or system that is continuously updated through real-time data exchange [5]. A digital twin typically integrates three essential components: the physical entity, such as a warehouse or distribution fleet; the virtual model, which simulates the structure and behavior of its physical counterpart; and the data interaction mechanism, which enables continuous, bidirectional synchronization between the physical and virtual systems. The implementation of DT relies on a range of enabling technologies. The Internet of Things (IoT) facilitates data acquisition from physical assets through sensors, Radio Frequency Identification (RFID) tags, and connected devices. Artificial Intelligence (AI) supports predictive analytics, demand forecasting, and operational optimization [6]. For example, artificial intelligence has been widely used in many fields and has achieved very good results [7, 8].

Meanwhile, fifth-generation (5G) mobile networks provide ultra-low latency and high-bandwidth communication, ensuring timely synchronization [9]. Cloud computing offers scalable data storage and processing capabilities for complex simulations [10]. In the logistics context, these technologies collectively transform DT into a real-time decision-support system capable of integrating operational data with predictive modeling to enhance performance, efficiency, and resilience.

The evolution of smart warehousing has paralleled advances in automation, analytics, and intelligent control systems. Modern warehouses increasingly adopt AGVs for driverless goods transportation, AS/RS for high-density storage and rapid retrieval, and robotic sorting systems for high-speed, accurate order fulfillment [11]. Complementing these mechanical innovations are information systems such as Warehouse Management Systems (WMS), which integrate order processing, inventory tracking, and operational analytics into a centralized platform, and RFID technology, which enables real-time tracking of goods and assets. While these innovations have significantly improved throughput and reduced labor dependency, existing research highlights persistent challenges in layout optimization, human-machine collaboration, and operational adaptability to rapidly changing demand patterns, indicating the need for more adaptive and integrated solutions [12].

Parallel to warehouse automation, distribution network optimization remains a critical research domain in logistics management. Strategic network design involves determining the optimal locations and capacities of warehouses, cross-docks, and delivery stations to balance cost, efficiency, and service quality. Mathematical models such as mixed integer linear programming are widely applied for these large-scale planning problems [13]. At the operational level, vehicle routing problems and their numerous variants, including capacity-constrained and time-window-based formulations, aim to identify the most efficient delivery paths under real-world constraints. In recent years, dynamic optimization and real-time scheduling have gained prominence, driven by the need to adapt to live traffic data, last-minute orders, and vehicle breakdowns. Emerging research increasingly integrates predictive analytics into routing algorithms, shifting from reactive responses to proactive, data-driven decision-making [14].

Although digital twin technology has been successfully deployed in manufacturing and aerospace industries, its application in logistics is still at an early stage. Pioneering companies have initiated digital twin-like solutions, with JD.com employing IoT-enabled warehouse twins for real-time simulation of storage layouts and picking strategies, Amazon utilizing AI-powered virtual models of distribution networks for predictive demand modeling and capacity planning, and Cainiao integrating DT systems with 5G communication for multi-warehouse coordination. These implementations demonstrate tangible benefits such as reduced delivery times, improved resource utilization, and lower operating costs. However, significant challenges remain. Achieving real-time data synchronization requires advanced infrastructure and low-latency networks [15, 16]. Data fusion across heterogeneous IoT devices, WMS, and transportation platforms presents considerable complexity. Interoperability between legacy systems and modern DT platforms remains problematic, while the high capital and computational requirements limit adoption among small and medium-sized enterprises [17]. These problems underscore the necessity for academic research to move beyond isolated

warehouse or transport applications, toward the development of holistic, integrated DT frameworks that unify warehousing and distribution network optimization while ensuring real-time responsiveness, interoperability, and scalability.

To more comprehensively illustrate the technological chain relationships within smart warehousing, Figure 1 presents a comprehensive diagram of the Digital Twin in Logistics. It depicts the overall architecture and data flow of digital twin technology in the logistics sector, demonstrating how it connects the physical and digital worlds in real time and drives optimization. The left side of the diagram represents the physical entity layer, encompassing operational objects such as warehouses, transport vehicles, distribution networks, and cargo. These entities continuously collect real-time status data (such as inventory levels, location, temperature and humidity, vehicle speed, and equipment health) through sensors, GPS, RFID, and IoT devices. This data enters the digital space through the data collection and transmission layer, where it is cleaned, integrated, and modeled at the data processing and integration layer, forming a highly accurate and interactive digital twin model. This model simulates the operational status of the logistics system in real time and, combined with predictive analytics, optimization algorithms, and machine learning, provides intelligent decision support for warehouse operation optimization (such as picking routes and inventory layout), transportation scheduling optimization (such as route planning and load distribution), and equipment maintenance prediction. Digital twins not only passively reflect reality but also push optimization suggestions or automated scheduling instructions back to the physical system through feedback loops, completing a closed loop of real-time monitoring, prediction, optimization, and execution. This reduces costs, shortens delivery times, improves resource utilization, and enhances supply chain resilience. Overall, this diagram emphasizes the core value of digital twins in logistics: through the bidirectional mapping of digital and physical data, a logistics management model with global visualization, real-time intelligence, and adaptive optimization is achieved.

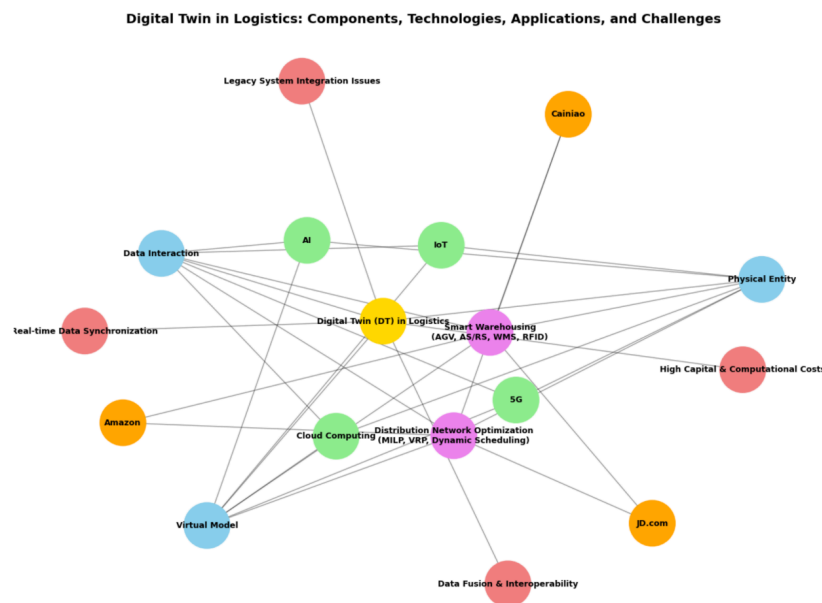


Figure 1. The comprehensive diagram of the digital twin in logistics

3. Methodology

The overall architecture of this system adopts a three-layer structure consisting of a perception layer, a data layer, and an application layer. The perception layer is responsible for real-time collection of multi-source information, including the operating status of warehouse equipment, operational process parameters, and environmental monitoring data. It leverages IoT sensors, RFID recognition technology, and automated detection devices to achieve precise perception of the work site. The data layer undertakes the core functions of data storage, processing, and analysis. It centrally manages and performs high-speed computations on the information collected by the perception layer through a cloud computing platform. Combining big data analysis with machine learning algorithms, it provides reliable data support for upper-layer applications. Based on the processing results from the data layer, the application layer constructs a visual management interface and decision support tools, enabling intelligent scheduling, optimization, and monitoring of the entire warehousing and distribution process.

In terms of data flow and information exchange mechanisms, the system's operational process can be summarized into four steps: "equipment data collection - cloud-based data processing - virtual simulation - optimization feedback." First, the devices and sensors in the perception layer collect on-site operational data in real time and upload it to the cloud via secure data transmission channels. Subsequently, the data layer processes this data using modeling and algorithmic analysis, generating

virtual simulation scenarios to simulate the implementation effects of different decision-making scenarios. Finally, based on the simulation results, the system feeds optimization solutions back to the site, enabling dynamic adjustments to warehousing operations and delivery routes, thus forming a closed-loop information exchange system.

In terms of functional module design, the system primarily includes three core functions: warehousing operation simulation, dynamic delivery route optimization, and equipment predictive maintenance. The warehousing operation simulation module covers key processes such as warehousing, picking, and outbound delivery, recreating actual operational processes in a virtual environment to facilitate process optimization and personnel training. The dynamic delivery route optimization module utilizes real-time traffic and order data, combined with a path planning algorithm, to generate optimal transportation routes, improving delivery efficiency and reducing operating costs. The equipment predictive maintenance module uses machine learning models based on equipment operating data and failure history to proactively identify potential failure risks, enabling preventive maintenance, reducing equipment downtime and extending service life.

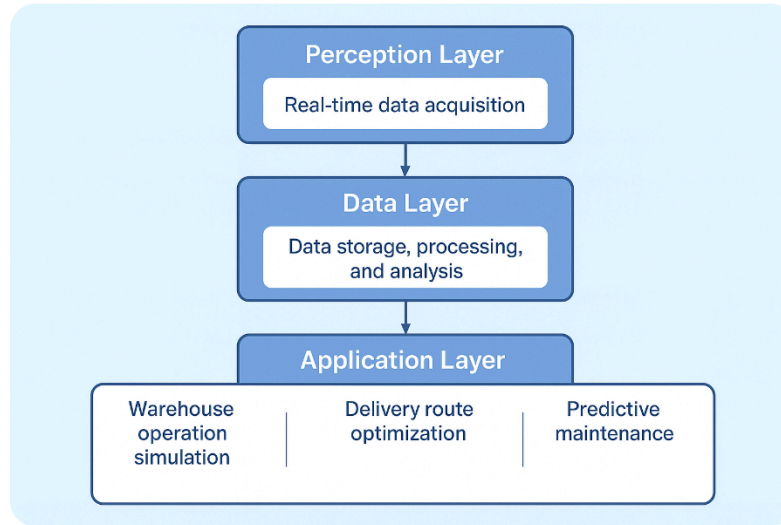


Figure 2. Digital Twin smart warehousing system architecture

Table 1 illustrates the mapping between each functional module of the system and its corresponding optimization objectives. The aim of this correspondence is to clarify how different system components contribute to overall performance enhancement. For example, the user interface module focuses on improving usability and providing a seamless user experience, while the content management module ensures the delivery of relevant, comprehensive, and accessible information. Similarly, data processing is optimized for accuracy and efficiency, and the interaction and feedback module is designed to foster active participation and responsiveness. The security and privacy protection module safeguards user data, and the system maintenance and upgrade module ensures long-term stability and scalability. This mapping provides a clear framework for targeted optimization during system design and implementation.

Table 1. Functional modules and their corresponding optimization objectives

Functional Module	Optimization Objective
User Interface Module	Improve usability and enhance user experience
Content Management Module	Ensure content relevance, richness, and accessibility
Data Processing Module	Enhance data accuracy, timeliness, and processing efficiency
Interaction and Feedback Module	Strengthen user engagement and timely feedback mechanisms
Security and Privacy Protection Module	Ensure data security and user privacy
System Maintenance and Upgrade Module	Improve system stability, scalability, and performance

4. Conceptual case study

4.1. Case background

This study uses a simplified, virtual e-commerce company as a representative example of a regional logistics layout. The company operates three primary regional warehouses, each responsible for inventory management and goods distribution within its region. These warehouses are supported by 15 delivery stations to provide multi-point coverage. This network configuration is designed to meet diverse order demand and cope with peak logistics pressures. Typical order distribution exhibits dynamic spatial and temporal variations, and the company operates a moderately sized delivery fleet. This setup creates a complex and challenging logistics network, providing a realistic foundation for the application and validation of the proposed system.

4.2. Application scenarios

The system focuses on two core optimization scenarios: warehouse operations and distribution network management. In warehouse operations, the system constructs a digital simulation environment to optimize picking routes, minimizing travel distance and time for workers. The system redesigns the inventory layout, placing high-frequency items closer to the picking area and low-frequency items farther away. This improves picking efficiency, optimizes space utilization, and supports virtual training for new employees, thereby reducing onboarding costs. Regarding distribution network optimization, the system integrates real-time traffic congestion data, order distribution, and fleet status, applying a dynamic route planning algorithm to generate efficient delivery routes. By adjusting delivery sequence and routes, total mileage and delivery time are minimized. Real-time scheduling capabilities enable the system to respond to unexpected changes in orders and traffic, ensuring delivery flexibility and accuracy.

In a simulated analysis environment, the system demonstrated significant performance improvements based on implemented optimization strategies. By optimizing picking and routing, labor and other resource consumption were reduced, resulting in a 15% reduction in costs. By dynamically adjusting routes and increasing delivery speeds, average delivery time was shortened by 12%. Furthermore, resource utilization improved by 10% compared to the baseline, including better warehouse space utilization and higher vehicle load factors. Figure 3 visually compares the routes of the distribution network before and after optimization, demonstrating a more centralized and coordinated routing pattern after optimization.

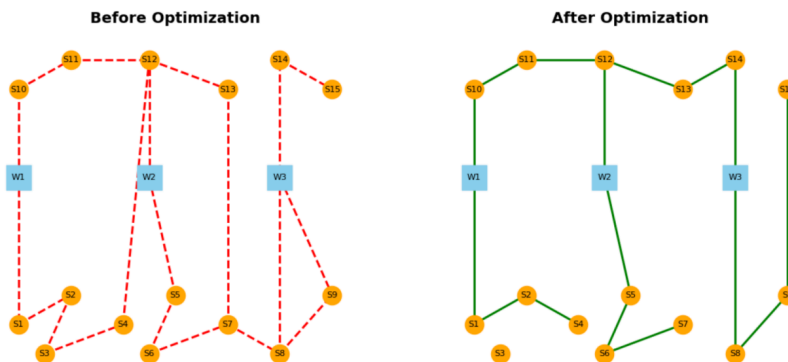


Figure 3. The routes of the distribution network before and after optimization

5. Discussion

The logistics optimization framework proposed in this study offers significant advantages in visualization, real-time performance, and simulation capabilities. Through advanced data visualization technology, managers can intuitively understand the operational status of their warehousing and distribution networks and promptly identify potential bottlenecks. The real-time data processing module enables the system to rapidly generate new optimization solutions as order demand or the external environment changes, improving decision-making flexibility and responsiveness. Furthermore, the simulation function not only supports the comparison of multiple solutions but also predicts optimization results before implementation, reducing trial-and-error costs.

However, the framework faces several challenges and limitations in practical application. First, the integrity and accuracy of real-time data directly impact optimization results, and many companies' data collection systems are still underdeveloped, which can lead to biased decisions. Second, system construction and maintenance require a high initial investment, including hardware deployment, software development, and personnel training—a significant consideration for small and medium-sized enterprises [18]. Finally, the complexity of technical integration cannot be ignored, especially when integrating the framework with existing

ERP, WMS , or Transportation Management System), which may require customized development and lengthy debugging cycles. In addition, it has been noted that artificial intelligence algorithms may play a good role in planning, as cases of artificial intelligence algorithms from different fields [19-21] have proven that they can play a certain role in product optimization.

From a management and practical perspective, the introduction of this framework can help improve operational transparency and scientific management in enterprises. Through real-time monitoring and data-driven decision-making, companies can more accurately allocate resources, reduce inefficient operations, and maintain supply chain stability amidst demand fluctuations. Simulation and forecasting capabilities can also inform strategic investments and market expansion, mitigating operational risks.

6. Conclusion and future work

This research constructs and validates a warehousing and distribution network optimization framework for virtual e-commerce companies. This framework integrates three core capabilities: data visualization, real-time scheduling, and simulation analysis. In case testing, the framework effectively reduces logistics costs and shortens average delivery times, demonstrating its high feasibility and application value in improving operational efficiency. For logistics companies, this framework can serve as a key tool for digital transformation. It is recommended that companies prioritize establishing high-quality data collection and management mechanisms during implementation to ensure real-time performance and accuracy. Furthermore, during system deployment, compatibility with existing information systems should be considered, and a phased implementation approach should be adopted to mitigate investment and technical risks. Furthermore, customized optimization based on the company's strategic goals will help maximize the framework's benefits.

Future research can be carried out in the following three directions: (1) Conduct larger-scale empirical research to verify the applicability and stability of the model in different business environments by comparing it with actual operational data; (2) Explore the combination of this framework with blockchain technology to enhance data transparency and traceability, especially in multi-party collaborative supply chain scenarios; (3) Expand its application in the field of green logistics, and incorporate sustainability indicators such as carbon emissions and energy consumption into optimization goals to support enterprises in achieving the dual goals of economic benefits and environmental responsibility.

References

- [1] Madanchian, M. (2024). The role of complex systems in predictive analytics for e-commerce innovations in business management. *Systems*, 12(10), 415.
- [2] Sodiya, E. O., Umoga, U. J., Amoo, O. O., & Atadoga, A. (2024). AI-driven warehouse automation: A comprehensive review of systems. *GSC Advanced Research and Reviews*, 18(2), 272–282.
- [3] Brdulak, H., & Brdulak, A. (2021). Challenges and threats faced in 2020 by international logistics companies operating on the Polish market. *Sustainability*, 13(1), 359.
- [4] Lindstrom, E. (2023). Feasibility of a Human Capital Digital Twin (Master's thesis). Massachusetts Institute of Technology.
- [5] Lim, K. Y. H., Zheng, P., & Chen, C.-H. (2020). A state-of-the-art survey of Digital Twin: Techniques, engineering product lifecycle management and business innovation perspectives. *Journal of Intelligent Manufacturing*, 31(6), 1313–1337.
- [6] Fuller, A., Fan, Z., Day, C., & Barlow, C. (2020). Digital twin: Enabling technologies, challenges and open research. *IEEE Access*, 8, 108952–108971.
- [7] Luo, Y. T., Liu, T., Pang, P. C.-I., Wang, Z., & Chan, K. I. (2025). Exploring information interaction preferences in an LLM-assisted learning environment with a topic modeling framework. *Applied Sciences*, 15(13), 7515.
- [8] Luo, Y., & Wang, Z. (2024). Feature mining algorithm for student academic prediction based on interpretable deep neural network. In *2024 12th International Conference on Information and Education Technology (ICIET)* (pp. 1–5). IEEE.
- [9] Kelechi, A. H., Alsharif, M. H., Ramly, A. M., Abdullah, N. F., & Nordin, R. (2019). The four-C framework for high capacity ultra-low latency in 5G networks: A review. *Energies*, 12(18), 3449.
- [10] Byrne, J., et al. (2017). A review of cloud computing simulation platforms and related environments. In *International Conference on Cloud Computing and Services Science* (Vol. 2, pp. 679–691). SciTePress.
- [11] Torchio, F. (2023). Survey on automated systems for smart warehouses (Master's thesis). Politecnico di Torino.
- [12] Islam, M. T., Sepanloo, K., Woo, S., Woo, S. H., & Son, Y.-J. (2025). A review of the industry 4.0 to 5.0 transition: Exploring the intersection, challenges, and opportunities of technology and human-machine collaboration. *Machines*, 13(4), 267.
- [13] Ho, W., Lee, C. K., & Ho, G. T. S. (2010). Multiple criteria optimization of contemporary logistics distribution network problems. *OR Insight*, 23(1), 27–43.
- [14] Cowling, P., & Johansson, M. (2002). Using real time information for effective dynamic scheduling. *European Journal of Operational Research*, 139(2), 230–244.
- [15] Liß, C., Ulbricht, M., Zia, U. F., & Müller, H. (2017). Architecture of a synchronized low-latency network node targeted to research and education. In *2017 IEEE 18th International Conference on High Performance Switching and Routing (HPSR)* (pp. 1–7). IEEE.
- [16] Shukla, S., Hassan, M. F., Tran, D. C., Akbar, R., Paputungan, I. V., & Khan, M. K. (2023). Improving latency in Internet-of-Things and cloud computing for real-time data transmission: A systematic literature review (SLR). *Cluster Computing*, 26(5), 2657–2680.

- [17] Ponnusamy, S., Assaf, M., Antari, J., Singh, S., & Kalyanaraman, S. (2024). Digital twin technology and AI implementations in future-focused businesses. IGI Global.
- [18] Love, P. E., Irani, Z., & Edwards, D. J. (2004). Industry-centric benchmarking of information technology benefits, costs and risks for small-to-medium sized enterprises in construction. *Automation in Construction*, 13(4), 507–524.
- [19] Xu, J., & Wang, Y. (2025). FMT: A multimodal pneumonia detection model based on stacking MOE framework. In *2025 8th International Conference on Information and Computer Technologies (ICICT)* (pp. 517–521). IEEE.
- [20] Niu, T., Liu, T., Luo, Y. T., Pang, P. C.-I., Huang, S., & Xiang, A. (2025). Decoding student cognitive abilities: A comparative study of explainable AI algorithms in educational data mining. *Scientific Reports*, 15(1), 26862.
- [21] Du, H., Zhuang, Y., & Xu, J. (2024). Symbiotic fusion of evolving anatomical landscapes for intelligent diagnostic narratives. In *2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (pp. 4513–4520). IEEE.