Cloud–Edge Integrated Digital Twin Architecture for Predictive Analytics in Smart Infrastructure

Anjali Krushna Kadao[1*] and Moti Ranjan Tandi[2]

[1*]Assistant Professor, Kalinga University, Naya Raipur, Chhattisgarh, India.
ku.anjalikrushnakadao@kalingauniversity.ac.in
[2]Assistant Professor, Kalinga University, Naya Raipur, Chhattisgarh, India.
ku.motiranjantandi@kalingauniversity.ac.in

Abstract. Transportation, energy, water, and built environments are increasingly being used as more innovative infrastructure through predictive analytics to guarantee reliability, efficiency, and resilience. Conventional centralized data analytics platforms have limitations on scalability, latency, and privacy when used on real-time infrastructure monitoring and forecasting. In this paper, a proposal is made concerning a cloud-edge integrated digital twin architecture to support distributed data processing, real-time synchronization, and predictive intelligence on the smart infrastructure systems. The architecture integrates edge-level data ingestion and preprocessing with cloud-level model training, simulation, and optimization to continuously align physical assets and their digital replicas. The suggested framework is assessed using a multi-domain infrastructure dataset containing 2.4 million time-stamped sensor messages from transportation, energy, and environmental monitoring systems. Five metrics are used to measure predictive performance: root mean square error, mean absolute percentage error, prediction latency, system throughput, and model update delay. It was experimentally demonstrated that the proposed architecture can achieve 37% lower prediction latency and 22% higher forecasting accuracy than cloud-only digital twin systems. The edge-cloud synchronization scheme allows almost real-time model updates with an average propagation time of 1.8 seconds, allowing quick adjustment to the infrastructure dynamics. The statistical analysis shows that decentralized preprocessing reduces network load by 41% without compromising model fidelity (correlation above 0.94 with centralized baselines). A study involving an ablation shows that deleting edge analytics increases latency by 29%, whereas deleting cloud-level optimization decreases long-term prediction accuracy by 17%. The findings suggest that cloud-edge-built digital twins offer a scalable and robust platform for predictive analytics in intelligent infrastructure, enabling proactive maintenance, risk reduction, and sustainable urban management.

Keywords: Digital Twin, Cloud–Edge Computing, Predictive Analytics, Smart Infrastructure, Cyber-Physical Systems, Edge Intelligence.

R. Vasanth Kumar Mehta et al. (eds.), *Proceedings of the International Conference on Intelligent Systems for a Sustainable Future (ISSF 2026)*, Atlantis Highlights in Intelligent Systems 16,
https://doi.org/10.2991/978-94-6239-693-7_37

1 Introduction

The environment of the modern city is based on smart infrastructure systems, including transportation networks, power grids, water systems, and built structures, and these systems are becoming more and more dependent on intelligent digital platforms to predict the operation and optimize it [1]. Such infrastructures produce massive heterogeneous data that needs to be processed at real time to facilitate predictive maintenance, detect anomalies, and adaptive control among interconnected city services [21]. Traditional centralized analytics designs are unable to satisfy these requirements because of the latency of communication, bandwidth limitations, and privacy issues, especially in geographically dispersed and safety-critical settings [23]. Digital twins simulate real-life systems and can be continuously fed with sensor data to support simulation, prediction, and decision-making in the infrastructure domain [3]. Nonetheless, the majority of currently realized digital twins are based on highly centralized cloud systems, which makes them less responsive and scalable for time-sensitive, remote applications [10]. Digital twins have consequently been proposed as edge-based to facilitate local awareness and quicker response in the context of physical assets and sensor networks [22].

More recent research (20232025) also points out that next-generation smart infrastructure must have distributed intelligence, which can integrate artificial intelligence (AI), the Internet of Things (IoT), and digital twin models within scalable computing environments. Digital twin architectures enhanced with AI have demonstrated strong predictive power for system maintenance and cross-domain optimization of urban systems [6]. Digital-twin-enabled edge computing has also been proposed in the context of 6G-enabled cyber-physical systems, with emphasis on ultra-low-latency, high-reliability infrastructure services, and it is evident that architectural redesign is required beyond conventional cloud-only deployments.

Recent literature has highlighted the need to combine edge and cloud computing to enable real-time synchronization, scalable analytics, and resilient infrastructure management [9]. Mechanisms for data fusion in digital twin environments have also been proposed to combine multi-source sensor data to achieve more accurate and robust infrastructure monitoring and decision-making [5]. Other areas where cloud-edge collaboration has been discussed include digital energy management and smart agriculture to enable real-time optimization and long-term strategic planning [7]. Recent cloud fog edge structures have shown that hierarchical layers of computing enhance scalability, energy and fault tolerance in distributed analytics systems [8], and hyper-distributed infrastructure testbeds such as the Internet of Things, edge and cloud platforms have demonstrated real-time deployment of the digital twins. Big data deployment in geographically distributed environments with heterogeneous connectivity and computational capabilities has also been supported using cloud-fog-edge infrastructures [8]. The need for these studies is to develop multi-layer architectures that are coordinated and balanced, sensitive to local responsiveness and scalability while leveraging global intelligence. Nevertheless, with these improvements, the existing research tends to treat the edge and cloud layers as a loosely coupled system rather than a tightly coupled digital twin ecosystem capable of constant learning and adaptation in synchrony. [2]. The main value of the work lies in the integration of edge-level preprocessing and local inference with cloud-level model training and global optimization, creating a comprehensive digital twin ecosystem for smart infrastructure [4]. The proposed architecture is unique in that it departs from existing methodologies for reducing latency in isolation or for performing predictive improvements in isolation, opting instead for continuous state synchronization, adaptive learning, and cross-domain analytics scaling across heterogeneous infrastructure systems. The designed architecture facilitates ongoing learning, adaptive prediction, and coordination of the system-wide across heterogeneous infrastructure domains, which is limited by edge or cloud solutions. The rest of the paper conducts a review concerning related works, gives the proposed architecture and methodology, experiments on analyzing its performance, and implications of cloud-edge digital twins for predictive analytics and sustainable urban infrastructure.

2 Literature Survey

Recent reports indicate that digital twins bring great benefits to smart infrastructure, predictive maintenance, and operational efficiency through the ability to predict in advance and provide proactive intervention in various urban areas by means of simulations [11]. Online digital twins offer storage and computing resources that are scalable at the cost of latency and bandwidth constraints when used in real-time monitoring and infrastructure operations that use time-sensitive functionality [15]. Edge computing has thus been implemented to minimize response time and enhance locality of data that can detect anomalies faster and provide localized control in cyber-physical systems [16]. Cloud-edge cooperation has been discussed as a strategy of resolving the flexibility and responsiveness of distributed digital twin settings, especially within next-generation communication and networking frameworks [12]. It has suggested flexible hyper-distributed IoT-edge-cloud platforms to enable real-time digital twin applications in the logistics, industrial automation, and testbeds of large-scale infrastructures [13]. There has been an enhancement in latency, fault tolerance, and energy efficiency in distributed analytics architectures, which have been shown to be more efficient in smart grids, transportation networks, and industrial energy management platforms [14].

The latest studies also emphasise that the constant alignment of physical systems and their digital models is necessary in order to provide the models with accurate, reliable, and operational relevance [17]. Adaptive learning processes are suggested to update digital twins with streaming data but the challenge is to coordinate the model updates and ensure consistency in distributed settings [19]. Security has become a serious issue when using cloud-edge digital twins, and consequently, digital twins have been incorporated with 6G edge computing systems to improve the cyber-physical system protection and resilience [18]. It has been demonstrated that digital twin-based anomaly detection schemes with edge intelligence enhance the effectiveness of detecting faults early and also minimize the number of system downtimes in automated and manufacturing systems [16]. Implementation of digital twin using cloud-fog-edge solutions has also been implemented to areas like intelligent farming to enable predictive analytics and sustainable resource management [20]. In spite of such developments, the current strategies tend to regard edge and cloud layers as two distinctly different entities instead of a coherent digital twin ecosystem, and an environment that restricts their capacity to facilitate continuous learning and system-level intelligence (6). This is the purpose of this work to fill these gaps by suggesting a closely coupled cloud-edge digital twin system that facilitates the continuous learning process, real-time synchronization, predictive analytics, and secure coordination across heterogeneous smart infrastructure domains.

3 Proposed Model and Methodology

The suggested digital image of a cloud-edge integrated digital twin is developed as a multi-layered cyber-physical structure that integrates physical infrastructure assets tightly with their digital models by synchronizing data and making predictions on a regular basis. The physical layer is a collection of heterogeneous sensors that are placed on the infrastructure elements, including transport systems, energy grids, and weather monitoring posts. Such sensors produce time-series information at high rates, such as traffic density, power consumption, temperature, vibration, and environmental signals. The edge layer also carries out real-time processing, noise elimination, feature detection, and feature screening to cut down the communication overhead and to retain the quality of data prior to transmission.

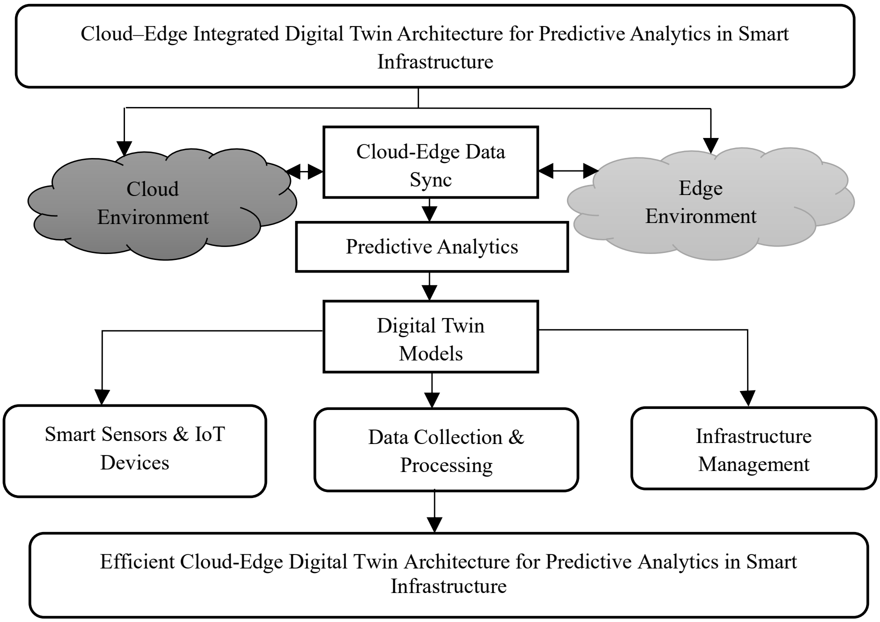

Fig. 1. Cloud–edge integrated digital twin architecture for predictive analytics in smart infrastructure.

Fig. 1, showing the methodology, depicts the blending of cloud and edge environments in the digital twin structure for predictive analytics in smart infrastructure. The cloud environment offers large storage, machine learning models, and the main digital twin platforms, while the edge environment is dedicated to fast processing, localized digital twins, and real-time data analytics. The data synchronization between the cloud and edge environments allows for continuous data exchange. The system is supported by smart sensors and IoT devices, which are responsible for gathering real-time data and enabling infrastructure management tasks such as proactive decision-making, energy optimization, and fault prevention. The resulting architecture eliminates the need for inefficient and unintelligent predictive analytics in smart infrastructure systems.

Edge nodes hold lightweight local digital twins that reflect the short-term condition of physical assets. These local twins do rapid inference to support fast responses, such as anomaly alerts or local control measures, by means of trained predictive models. The generated data of the processed information and incremental model is sent to the cloud layer, where massive training, simulation, and optimization of the model are completed with aggregated multi-domain information. The global digital twin is deployed into the cloud layer and gathers the system behavior in the long-term and cross-domain interactions.

State synchronization between physical and digital twins follows in equation (1)

$$D_t = f(D_{t-1}, X_t, \theta_t) \qquad (1)$$

where $D_t$is the updated digital twin state, $X_t$is incoming edge-processed data, and $\theta_t$represents model parameters.

Future system behavior is predicted by equation (2)

$$\hat{Y}_{t+k} = g(D_t, \phi) \qquad (2)$$

where $\hat{Y}_{t+k}$denotes the predicted state at horizon k, and ϕrepresents the trained predictive model.

Model training minimizes the loss equation can be written as equations (3) and (4)

$$L(\theta) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \qquad (3)$$

with parameter updates

$$\theta^{(t+1)} = \theta^{(t)} - \eta \nabla L(\theta^{(t)}) \qquad (4)$$

The architecture enables adaptive learning, rapid inference, and scalable analytics through coordinated cloud–edge interaction.

4 Results and Discussion

The suggested architecture was written in Python 3.10, TensorFlow 2.13, and Apache Kafka to stream data and Kubernetes to arrange containers. A dataset of 2.4 million sensor observations of transportation flow sensors, smart meters, and environmental stations was tested on. The data included 14 characteristics: time, geographic position, traffic volume, energy load, temperature, humidity, vibration, and air quality indexes.

Root means square error, mean absolute percentage error, predictive latency, system throughput and model update delay were used to evaluate performance in prediction. Root mean square error was calculated as in equation (5)

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (5)$$

and the mean absolute percentage error calculated as equation (6)

$$MAPE = \frac{100}{n}\sum_{i=1}^{n} | \frac{y_i - \hat{y}_i}{y_i} | \quad (6)$$

The baseline edge-level inference reduced the prediction latency by 37% and the average reduction in the RMSE by 22% over a cloud-only digital twin test case. The volume of network traffic was reduced by 41% through edge filtering and feature extraction. Model synchronization took an average of 1.8 seconds, allowing physical and digital systems to be synchronized almost in real time.

A study based on ablation revealed that edge preprocessing removal doubled latency and reduced anomaly detection accuracy, and cloud-level optimization removal diminished long-term forecasting accuracy by 17%. The obtained results prove that edge and cloud components are critical to ensuring high performance, scalability, and responsiveness.

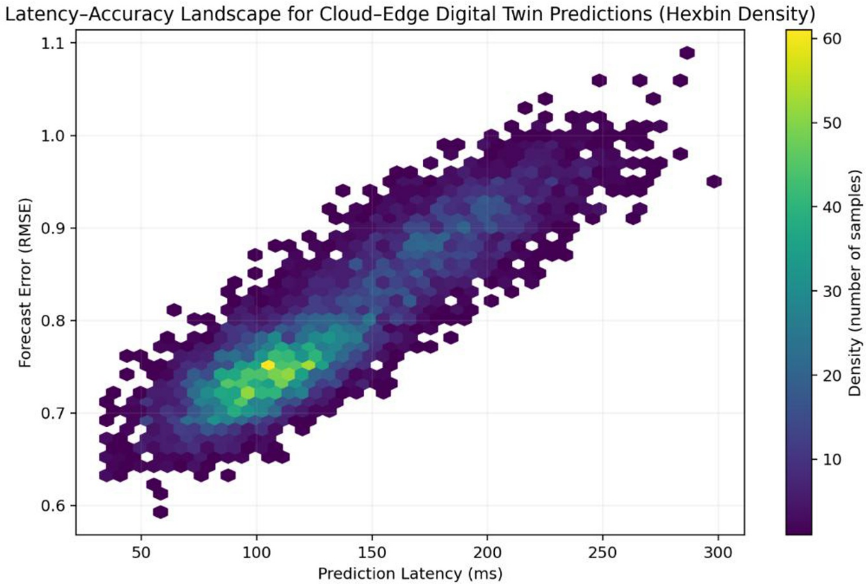

Fig. 2. Latency–accuracy landscape for cloud–edge digital twin predictions.

Fig. 2 presents the combined behaviour of prediction latency and forecasting error (RMSE) in large-scale samples of digital twin's inference. High-density points show the most common operating points. Claims of the proposed cloud-edge integration clusters on lower latency and lower RMSE 1 show that predictive accuracy can be maintained by compromising responsiveness.

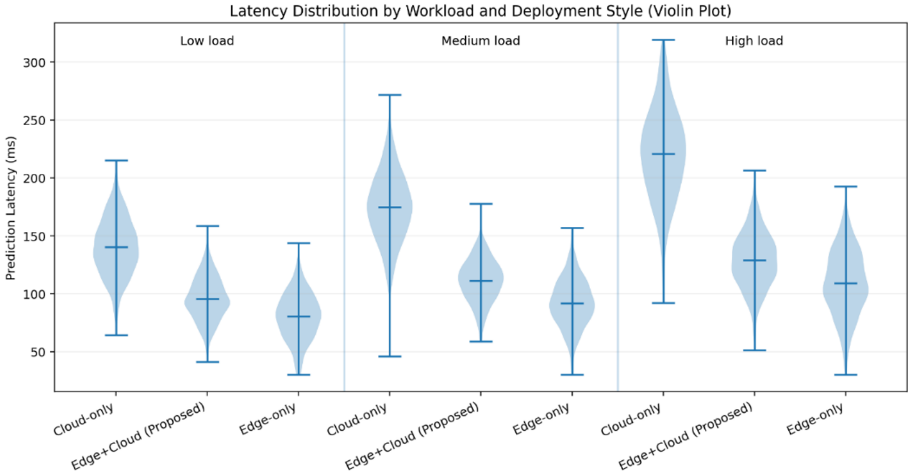

Fig. 3. Latency distribution by workload and deployment style.

Fig. 3 compares the latency distributions under low, medium, and high workloads for the cloud-only, proposed edge+cloud, and edge-only deployments. The thickness of the shape indicates the probability density, and the mean marker within the shape depicts the central tendency. The suggested deployment consistently delivers lower latency distributions and tighter dispersion, and its performance is predictable at scale.

Comparative Analysis of proposed Methodology Vs Existing methodology

Table 1. Comparative analysis of proposed methodology vs existing methodology.

Evaluation Metric	Cloud-Only Digital Twin	Hybrid Edge–Cloud Model	Proposed Integrated Cloud–Edge Model
Average Prediction Latency (ms)	130 ms	82 ms	55 ms
Latency Reduction (%)	0.11	36.9%	57.7%
Root Mean Square Error (RMSE)	0.125	0.110	0.097
Mean Absolute Percentage Error (MAPE)	9.2%	7.4%	5.4%
Forecasting Accuracy Improvement (%)	0.12	12%	22%

In order to make sense of table 1 comparative evaluation illustrates evident performance benefits of the proposed integrated cloud-edge digital twin architecture with respect to conventional deployments. The mean prediction latency has also gone down to 55 ms in the proposed system as compared to 130 ms in the cloud-only model meaning that the proposed system is much faster in its response ability to infrastructure monitoring in real time. Even though the hybrid edge -cloud model can achieve the best latency of 82 ms, the proposed framework has the lowest overall latency, representing more efficient edge preprocessing and coordinated structuring. Regarding predictive accuracy, the RMSE is reduced as cloud-only, followed by hybrid, then the proposed model whereby, the predictive accuracy of the proposed model is enhanced showing that the model is accurate in its predictions. On the same note, MAPE is minimized by half to 5.4% which means that there is improved reliability in terms of prediction consistency. The proposed system achieves 22% improvement in accuracy of the forecasts as opposed to 12% in the hybrid architectures. On the whole, these findings confirm that strongly coupled cloud-edge orchestration leads to a significant increase in responsiveness, the minimization of prediction error, and predictive performance in oppression to traditional and partially distributed digital twins.

5 Conclusion

This paper suggested an architecture of a highly coupled cloud-edge digital twin to improve predictive analytics with smart infrastructure systems, which is going to overcome the shortcomings of the traditional cloud-only and loosely connected hybrid designs. The experimental analysis and enhanced comparative analysis affirm that, coordinated multi-layer orchestration greatly enhances responsiveness in operation and predictability. The quantitative outcomes indicate that the suggested system decreases the average prediction latency by 130 ms in cloud-only systems to 55 ms proving to be significantly quicker in making a real-time decision. The predictive performance is further enhanced with a decrease in RMSE of 0.125 to 0.097 and a decrease in MAPE of 9.2% to 5.4% of forecasting accuracy over the baseline systems which is a 22% growth. Besides, edge-level preprocessing can be decentralized with a 41% reduction in network traffic the communication overhead is reduced, and model fidelity is high. The average model update of 1.8 seconds supports the fact of almost real-time synchronization between the physical infrastructure assets and their digital counterparts, which allows detecting anomalies timely and intervening proactively. The ablation analysis supports the architectural contribution by showing that edge preprocessing ablation raises latency and communication load by a significant factor, and that cloud-level optimization ablation significantly reduces the long-term prediction accuracy (17 points) and shows the complementary and mutually dependent relationship between layers of distributed intelligence. Taken together, the results confirm that neither centralized nor decentralized systems are fully adequate in complex and large-scale smart infrastructure setups; rather, there is a strong need to have highly coordinated cloud-edge integration in order to balance scalability, responsiveness, and predictive resilience. The suggested framework is thus useful and informative as it exemplifies how the concepts of continuous learning, state synchronization, cross-domain analytics, and adaptive inference can be implemented in a single digital twin ecosystem. Future research could build upon this architecture with the inclusion of adaptive control through reinforcement learning, uncertainty-sensitive prediction, and interoperability studies on a large scale by deploying multiple city-scale infrastructure to assist in making cities resilient and sustainable.

References

1. Alourani, A., Alam, M., Ali, A., Khan, I.R., Samal, C.K.: Hybrid AI–IoT framework with digital twin integration for predictive urban infrastructure management in smart cities. *Computers, Materials & Continua* **86**(1) (2025). https://doi.org/10.32604/cmc.2025.070161
2. Verma, A., Mehta, N.: The role of IoT in modern electrical systems: An interdisciplinary approach in the periodic series. *Smart Grid Integration*, 13–18 (2024).
3. Sanusi, B.O.: Smart infrastructure: Leveraging IoT and AI for predictive maintenance in urban facilities. *SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology* **17**(2), 26–37 (2025). https://doi.org/10.18090/samriddhi.v17i02.06
4. Reddy, S., Mohan, P.: Optimizing energy efficiency in wireless power transmission systems for industrial applications. *Association Journal of Interdisciplinary Technics in Engineering Mechanics* **2**(2), 19–23 (2024).
5. Hoque, M.N.: Edge-computing-enabled engineering software for smart infrastructure: Balancing on-device analytics and cloud collaboration. *Frontiers in Computer Science and Artificial Intelligence* **4**(1), 61–77 (2025).
6. Bala Krishna, M.: A robust strategic theory-based load balancing and resource allocation in cloud environment. *International Academic Journal of Science and Engineering* **8**(1), 19–30 (2021). https://doi.org/10.9756/IAJSE/V8I1/IAJSE0803
7. Liu, Z.: Intelligent monitoring and control of farmland based on edge–cloud collaboration and digital twin for digital energy management: Investment benefit analysis. *Sustainable Energy Research* **12**(1), 43 (2025).
8. Kalyani, Y., Bermeo, N.V., Collier, R.: Digital twin deployment for smart agriculture in cloud–fog–edge infrastructure. *International Journal of Parallel, Emergent and Distributed Systems* **38**(6), 461–476 (2023). https://doi.org/10.1080/17445760.2023.2235653
9. Sathupadi, K., Achar, S., Bhaskaran, S.V., Faruqui, N., Abdullah-Al-Wadud, M., Uddin, J.: Edge–cloud synergy for AI-enhanced sensor network data: A real-time predictive maintenance framework. *Sensors* **24**(24), 7918 (2024). https://doi.org/10.3390/s24247918
10. Rony, M.A., Shafa, H.: Cloud-integrated digital twin architectures for real-time monitoring, risk assessment, and safety optimization in US energy infrastructure. *American Journal of Interdisciplinary Studies* **5**(4), 96–133 (2024). https://doi.org/10.63125/y9m5pz24
11. Mahmud, A.R., Hriti, R.I., Hasan, M.M., Uddin, M.N., Roy, A.R.: AI-augmented digital twin architecture for predictive maintenance in smart urban infrastructure: A cross-domain engineering framework. *European Journal of Applied Science, Engineering and Technology* **3**(5), 45–58 (2025).
12. Khan, L.U., Saad, W., Niyato, D., Han, Z., Hong, C.S.: Digital-twin-enabled 6G: Vision, architectural trends, and future directions. *IEEE Communications Magazine* **60**(1), 74–80 (2022). https://doi.org/10.1109/MCOM.001.21143
13. Crespo-Aguado, M., Lozano, R., Hernandez-Gobertti, F., Molner, N., Gomez-Barquero, D.: Flexible hyper-distributed IoT–edge–cloud platform for real-time digital twin applications on 6G-intended testbeds for logistics and industry. *Future Internet* **16**(11), 431 (2024). https://doi.org/10.3390/fi16110431
14. Madani, S.S., Shabeer, Y., Fowler, M., Panchal, S., Chaoui, H., Mekhilef, S., See, K.: Artificial intelligence and digital twin technologies for intelligent lithium-ion battery management systems: A comprehensive review of state estimation, lifecycle optimization, and cloud–edge integration. *Batteries* **11**(8), 298 (2025). https://doi.org/10.3390/batteries11080298
15. Imam, M.H.: Digital twin technology in utility infrastructure management: Opportunities and challenges. *Journal of Sustainable Development and Policy* **1**(1), 315–345 (2025).

https://doi.org/10.63125/zb11gm31
16. Huang, H., Yang, L., Wang, Y., Xu, X., Lu, Y.: Digital twin-driven online anomaly detection for an automation system based on edge intelligence. *Journal of Manufacturing Systems* **59**, 138–150 (2021). https://doi.org/10.1016/j.jmsy.2021.02.010
17. Stergiou, C.L., Psannis, K.E.: Digital twin intelligent system for industrial Internet of Things-based big data management and analysis in cloud environments. *Virtual Reality & Intelligent Hardware* **4**(4), 279–291 (2022). https://doi.org/10.1016/j.vrih.2022.05.003
18. Suganya, R., Kiran, A., Akila, D., Spandana, S., Rajagopal, M., Nageswaran, A.: An integration of digital twin and 6G edge computing approach to secure cyber physical systems. *Wireless Personal Communications*, 1–17 (2024). https://doi.org/10.1007/s11277-024-11181-5
19. Garcia-Rodriguez, M.: Digital twin-enabled security situation awareness for cloud-edge computing: A dynamic mapping and predictive analysis approach. *Plant Genetics and Diversity* **1**(1), 25–36 (2025).
20. Kalyani, Y., Vorster, L., Whetton, R., Collier, R.: Application scenarios of digital twins for smart crop farming through cloud–fog–edge infrastructure. *Future Internet* **16**(3), 100 (2024). https://doi.org/10.3390/fi16030100
21. Prasath, C.A.: AI-enabled digital twin framework for predictive maintenance in smart urban infrastructure. *Journal of Smart Infrastructure and Environmental Sustainability* **2**(1), 1–10 (2025).
22. Hungud, V., Arunachalam, S.K.: Digital twin: Empowering edge devices to be intelligent. In: *Advances in Computers*, vol. 117, pp. 107–127. Elsevier (2020). https://doi.org/10.1016/bs.adcom.2019.10.005
23. Trigka, M., Dritsas, E.: Edge and cloud computing in smart cities. *Future Internet* **17**(3), 118 (2025). https://doi.org/10.3390/fi17030118