CAAI Transactions on Intelligence Technology

WILEY

ORIGINAL RESEARCH OPEN ACCESS

A Novel Single-Encounter-Emphasise-Attention Neural Network for Predictive Maintenance in Urban Infrastructure

Tawfeeq Shawly[1] | Ahmed A. Alsheikhy[2]

[1]Department of Electrical Engineering, Faculty of Engineering at Rabigh, King Abdulaziz University, Jeddah, Saudi Arabia | [2]Department of Computer and Network Engineering, College of Computer Science and Engineering, University of Jeddah, Jeddah, Saudi Arabia

Correspondence: Ahmed A. Alsheikhy (aalsheikhy@uj.edu.sa)

Received: 27 January 2026 | **Revised:** 21 May 2026 | **Accepted:** 14 June 2026

Keywords: digital twin | logistics | predictive maintenance | real-time monitoring | Riyadh Metro | smart transportation | Vision 2030

ABSTRACT

The Riyadh Metro represents a significant project in Saudi Arabia, designed to transform urban transportation and reduce traffic congestion within the city. With six metro lines and 85 stations, the network is expected to serve millions of passengers daily, necessitating innovative digitally driven maintenance approaches to ensure reliable efficient operation. We hypothesise that integrating Internet of things (IoT) sensors, digital twin technology, and a temporal graph neural network (TempGNN) enables real-time health monitoring. This paper introduces a real-time digital health monitoring framework that incorporates IoT sensors, digital twin technology, a proposed TempGNN and a new module single encounter-emphasise attention neural network (SEEA-NN). TempGNN exploits both temporal patterns and spatial relationships in these sensor data to feed SEEA-NN with its required data. SEEA-NN is designed to detect and assess the importance of rare single-occurrence failure precursors, often overlooked by conventional models, alongside recurring patterns. The proposed framework contains an input gating module to filter out noisy sensor data, a multihead encounter attention module that employs a specific attention head to isolate and enhance the signal from anomalous single-encounter events, and a temporal synthesis layer that combines these emphasised features with sequential data for thorough analysis. This enables the framework to predict potential failures with high accuracy of approximately 98%, allowing for proactive maintenance scheduling and decreased unplanned downtime by more than 30% and maintenance expenses by around 25%, whilst also enhancing energy efficiency and passenger satisfaction in simulations. This approach demonstrates a scalable model for the digital transformation of large-scale transport infrastructure and for implementing intelligent predictive maintenance in metro systems worldwide and in logistics fleets, container handling systems and freight rail networks.

1 | Introduction

The Riyadh Metro is among the largest and most ambitious urban transit initiatives globally, with a total investment of $22.5 B [1]. It comprises 6 metro lines, extending over 176 km with 85 stations, to accommodate over 3.6 M passengers daily by 2030 [1]. This metro system is a key component of Saudi Arabia's Vision 2030, which seeks to alleviate traffic congestion, decrease carbon emissions and enhance the living standards of Riyadh's inhabitants [1, 2]. Nonetheless, the vastness and intricacy of the system pose significant operational hurdles, particularly in maintaining reliability, efficiency and sustainability. The Riyadh Metro serves not merely as a transportation project but as a

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *CAAI Transactions on Intelligence Technology* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology and Chongqing University of Technology.

catalyst for urban transformation. By connecting with other public transport options, including buses and planned high-speed rail, it aims to revolutionise mobility in Riyadh [3, 4]. This network is expected to reduce the city's dependence on personal vehicles and relieve traffic congestion [4]. The government has highlighted the metro's potential to stimulate economic growth as improved accessibility will boost daily commutes and commercial activity in key areas and encourage new business districts along the routes [3, 4]. One of the standout features of the Riyadh Metro is its dedication to state-of-the-art technology and sustainability. The metro is engineered to operate autonomously with advanced signalling systems to enhance efficiency and reduce delays [1–4]. By employing artificial intelligence (AI) and data analytics, officials aim to track passenger traffic, forecast peak travel periods and improve overall service quality. These reasons position the Riyadh Metro among the most advanced transit systems globally. Despite these developments, the project encounters several challenges in operations and maintenance. The vast extent of the network requires a highly trained workforce to oversee, monitor and maintain the system, necessitating extensive training for metro personnel [3, 4]. Additionally, Riyadh's harsh climate, extreme heat and frequent sandstorms complicated infrastructure maintenance and consistent service delivery. Tackling these challenges is essential for the metro's long-term viability. Public adoption of the metro system is crucial for success. Operating a metro network of this scale involves managing an extensive array of trains, tracks and stations, each with unique challenges, including: (1) Maintenance: Guaranteeing timely and effective upkeep of trains and infrastructure to avoid interruptions in service. (2) Passenger Experience: Ensuring a safe, comfortable and dependable service to encourage public transit use. (3) Cost Management: Controlling operational and maintenance expenses to secure the system's long-term financial viability.

1.1 | Importance of Digital Twins

Digital twins have emerged as a revolutionary technology in the Industry 4.0 era, allowing the development of virtual models of physical systems that can imitate, forecast and enhance real-world functions [5–8]. By incorporating real-time data from IoT sensors, complicated analytics and machine learning (ML), digital twins serve as robust tools for overseeing, assessing and optimising complex systems [9–13]. In transportation, digital twins present remarkable opportunities to boost efficiency, lower costs and enhance sustainability [14–18]. They facilitate predictive maintenance, real-time monitoring and data-informed decision-making, making them essential for contemporary urban mobility infrastructures such as the Riyadh Metro. The adoption of digital twins in the Riyadh Metro has the potential to transform its operational effectiveness by enabling constant monitoring of infrastructure conditions and train performance. With real-time data integration, engineers can identify early indicators of wear, optimise energy use and improve service dependability. Predictive maintenance powered by digital twins can decrease downtime, avert costly failures and ensure seamless operations [6–10].

1.2 | Research Problem

The Riyadh Metro, a key element of Saudi Arabia's Vision 2030 initiative, is intended to accommodate millions of passengers daily using 6 lines and 85 stations. Nevertheless, maintaining the reliability, efficiency and sustainability of such an extensive system poses considerable challenges. Conventional maintenance methods tend to be reactive, leading to unanticipated breakdowns, costly repairs and service interruptions. Moreover, the absence of real-time monitoring and predictive capabilities can result in inefficiencies, higher energy use and diminished passenger satisfaction. There is a need for a solution that can address these issues and ensure the metro runs at optimal performance.

1.3 | Objectives, Motivations and Contributions

The main aim of this research is to develop and implement a real-time health monitoring system for the Riyadh Metro utilising digital twin technology. Specific objectives include (A) seveloping a digital twin framework that incorporates IoT sensors and live data analytics. (B) Enabling predictive maintenance to minimise downtime and operational expenses. (C) Improving safety and comfort by ensuring the system operates at peak performance levels. (D) Advancing sustainability goals by reducing energy usage and waste. (E) Extending the framework's applicability to logistics operations, including freight rail, automated warehouses, port container handling and cold chain transport systems.

The motivation for this research stems from the urgent need to modernise public transportation systems and logistic operations. This study seeks to (1) enhance the reliability and efficiency of the metro system, leading to lower operational costs and greater passenger satisfaction. (2) Advance Riyadh's sustainability objectives by reducing energy usage and environmental impacts. (3) Expand the application of the proposed model in logistics operations, including freight rail, automated warehouses, port container handling and cold chain transport systems.

This work offers several contributions, which can be summarised as follows: (1) A novel predictive maintenance framework: This framework integrates IoT sensor networks, a dynamic digital twin, TempGNN and SEEA-NN to enable real-time health monitoring of metro systems. The SEEA-NN has been designed and implemented with a unique multihead encounter attention module. This module enables high-accuracy, approximately 98% failure predictions by detecting and highlighting rare single-occurrence anomalies that conventional models often overlook. (2) Validated predictive maintenance system: This predictive maintenance has been validated, demonstrating a reduction in unplanned downtime by more than 30% and a decrease in maintenance costs by around 25% using simulation experiments. (3) Operational and energy efficiency: The system features real-time monitoring and data-driven optimisation capabilities that enhance operational efficiency by approximately 25% and decrease energy consumption by 15%. (4) Expanding our model to a transferable predictive maintenance framework that can be applied to logistics operations, including freight rail networks, automated warehouse fleets, port container cranes and cold chain transport systems. The SEEA-NN's ability to detect rare equipment anomalies enables logistics operators to prevent catastrophic failures in sorting machinery, reduce refrigerated cargo spoilage through early cooling unit fault detection and minimise warehouse downtime by predicting autonomous guided vehicle batteries and motor degradation.

The structure of this paper is as follows: Section 2 presents the literature review, followed by the proposed methodology explanation in Section 3. Section 4 elaborates on the experimental setup and results, Sections 5 and 6 examine the implications and limitations and Section 7 concludes with suggestions for future research.

2 | Literature Review

The application of digital twins and other cutting-edge technologies in transportation systems has garnered notable interest in recent years. This section reviews key methods and technologies pertinent to creating a real-time health monitoring system for the Riyadh Metro.

Rahman et al. [13] conducted a study on obstacle detection within railway systems, an essential area for ensuring safety and operational efficiency. They proposed a deep learning (DL)-based approach for the real-time identification of obstacles on railway tracks utilising convolutional neural networks (CNNs). The research highlighted the limitations of conventional surveillance methods, manual inspections and basic sensors, which frequently encounter inefficiencies and delays. The authors created a CNN model that was trained on an extensive dataset of track images featuring various obstacles, including animals, debris and unauthorised individuals. Their findings indicated that deep neural networks significantly enhance detection accuracy when compared to traditional computer vision methods, particularly in varying lighting and environmental conditions. The model effectively differentiated between normal and hazardous scenarios through techniques such as image preprocessing, edge detection and deep feature extraction. Various CNN architectures were assessed for their resilience against occlusions, low visibility and different types of obstacles. The proposed method achieved an impressive 97% accuracy rate in obstacle detection. However, the paper acknowledged challenges related to computational demands and latency for real-time applications, suggesting further improvements such as the development of lightweight neural models and the implementation of edge computing.

Hafsi et al. [14] presented a comprehensive roadmap for the implementation of predictive maintenance within industrial environments, highlighting the transition from reactive and preventive maintenance to predictive approaches that leverage real-time sensor data and AI analytics. They pointed out that the integration of IoT-based condition monitoring through intelligent sensors, edge computing and digital twin simulations can greatly minimise unplanned downtime, reduce maintenance expenses and prolong the lifespan of assets. The study examined a range of predictive maintenance techniques, including statistical methods as well as machine learning (ML) and deep learning (DL) models, and discussed essential components of these systems, such as data acquisition, feature extraction, fault diagnosis and predictive modelling. It also addressed practical challenges, including issues related to data quality, significant computational demands and the necessity for scalable infrastructure. The authors emphasised the significance of the industrial Internet of things (IIoT) in facilitating real-time condition monitoring via smart sensors and analysed how digital twins can replicate processes to anticipate failures. They highlighted the critical role of explainable AI for enhancing model interpretability and investigated the application of federated learning (FL) for decentralised maintenance analytics. Case studies across manufacturing, energy and transportation sectors demonstrated successful implementations of predictive maintenance.

Chauhan et al. [15] established an IoT-based urban transportation management framework aimed at enhancing real-time monitoring, traffic regulation and overall system efficiency. This research tackled the growing challenges of urban mobility, including congestion, delays and environmental repercussions, by introducing an intelligent transportation system (ITS) architecture. The framework employs advanced IoT sensors, cloud computing and machine learning algorithms to consistently gather and analyse traffic data. This facilitates dynamic traffic routing, adaptive signal control and predictive analytics for managing congestion. The findings indicated that IoT-driven solutions can mitigate traffic congestion and improve urban mobility by utilising vehicle-to-infrastructure (V2I) and vehicle-to-vehicle (V2V) communication technologies. Through various case studies and simulations, Chauhan et al. illustrated how real-time data and connectivity can substantially decrease travel times and emissions. They explored the technical dimensions of IoT-enabled transport management, emphasising elements such as edge computing for rapid data processing and big data analytics in the cloud for informed decision-making. The incorporation of AI techniques, including ML and DL, was demonstrated to assist in identifying traffic patterns and detecting anomalies.

Bešinović et al. [16] conducted a comprehensive review of the applications of artificial intelligence in railway transportation and introduced a taxonomy that classifies these applications, such as predictive maintenance, real-time traffic management, passenger flow forecasting and autonomous train operations. The authors highlighted numerous examples where ML, DL and even reinforcement learning (RL) techniques have improved the performance and reliability of railways. For instance, AI-based anomaly detection systems have been shown to enhance safety by forecasting faults prior to their occurrence, whereas intelligent scheduling algorithms have reduced operational disruptions. By incorporating case studies and practical implementations, the review illustrated how data-driven models can enhance service efficiency within contemporary rail networks. Furthermore, the authors addressed the regulatory and ethical frameworks related to the deployment of AI in railways, stressing the importance of transparency, safety standards and appropriate governance as automation becomes more prevalent. The efforts of various countries to create guidelines for the responsible integration of AI were also acknowledged. Key challenges identified include the necessity for interpretability in critical AI decisions to uphold trust and safety as well as the demand for substantial computational resources and skilled personnel to implement AI on a large scale.

These state-of-the-art methods emphasise the importance of predictive maintenance, IoT connectivity, advanced AI analytics and cloud frameworks in modern metro systems. However, an integrated solution that combines all these technologies, particularly one that leverages a model capable of capturing the system's temporal and spatial dynamics, remains a gap in the literature. Our proposed framework is designed to fill this gap by unifying IoT-driven data collection, AI-powered spatiotemporal

analytics and cloud-edge deployment into a comprehensive real-time health monitoring system for the Riyadh Metro.

3 | The Proposed Algorithm

The proposed approach is developed for real-time health monitoring of the Riyadh Metro, utilising digital twin technology along with two DL techniques: TempGNN and SEEA-NN. It is designed to improve operational efficiency, lower maintenance expenses and enhance passenger safety and satisfaction. The proposed framework includes several components, which are (A) digital twin framework: A digital replica of the Riyadh Metro system, involving trains, tracks, stations and infrastructure. It combines real-time data from IoT sensors, operational logs and environmental conditions. (B) IoT and sensor networks: Deployment of high-precision IoT sensors on trains, tracks and stations to track essential metrics, such as temperature, vibration, pressure and energy. (C) Data integration and management: A unified data management platform that aggregates diverse data sources: sensor readings, maintenance logs and passenger flow data. (D) Spatiotemporal analysis engine: The primary analytical engine is the TempGNN and SEEA-NN networks. This engine examines temporal and spatial relationships within the metro network. Its multihead encounter attention module is vital for detecting and prioritising rare single-occurrence failure precursors alongside standard operational data to enhance forecasting precision. (E) Predictive maintenance module: Utilises the engine predictions to forecast equipment malfunctions and recommends preemptive maintenance actions. (F) Energy optimisation module: Analyses energy consumption patterns and identifies opportunities for optimisation. It adjusts energy usage in real time based on data acquired. (G) User interface and control centre: Provide metro operators with a real-time dashboard to monitor system health and performance. We utilise a colour-coded diagram where each primary component is represented by a distinct colour as illustrated in Figure 1. Green signifies input/data layers, blue indicates the digital twin module, red represents the developed analytical engine, yellow denotes the predictive maintenance module and purple symbolises the user interface/control centre. In this framework, data inputs include aggregated real-time information from IoT sensors, logs and more. The digital twin framework converts raw data into a graph-based format by modelling trains, tracks and stations as interconnected nodes, thereby creating the spatial context necessary for the engine. The implemented engine observes spatiotemporal relationships with an emphasis on rare events, identifying anomalies, predicting issues and dynamically optimising operations. The predictive maintenance module is the system's forecasting and advisory centre that generates maintenance alerts, optimises train scheduling and enhances energy efficiency based on the outputs of the analytical engine. The user interface and control centre module engages with human operators by delivering real-time visual analytics and actionable insights. Operators can address emerging situations through the designed dashboard.

3.1 | Importance and Effectiveness of the Proposed Scheme

By developing a virtual twin of the entire metro network, the digital twin assists ongoing monitoring, real-time simulations and predictive analytics to enhance operations and avert failures

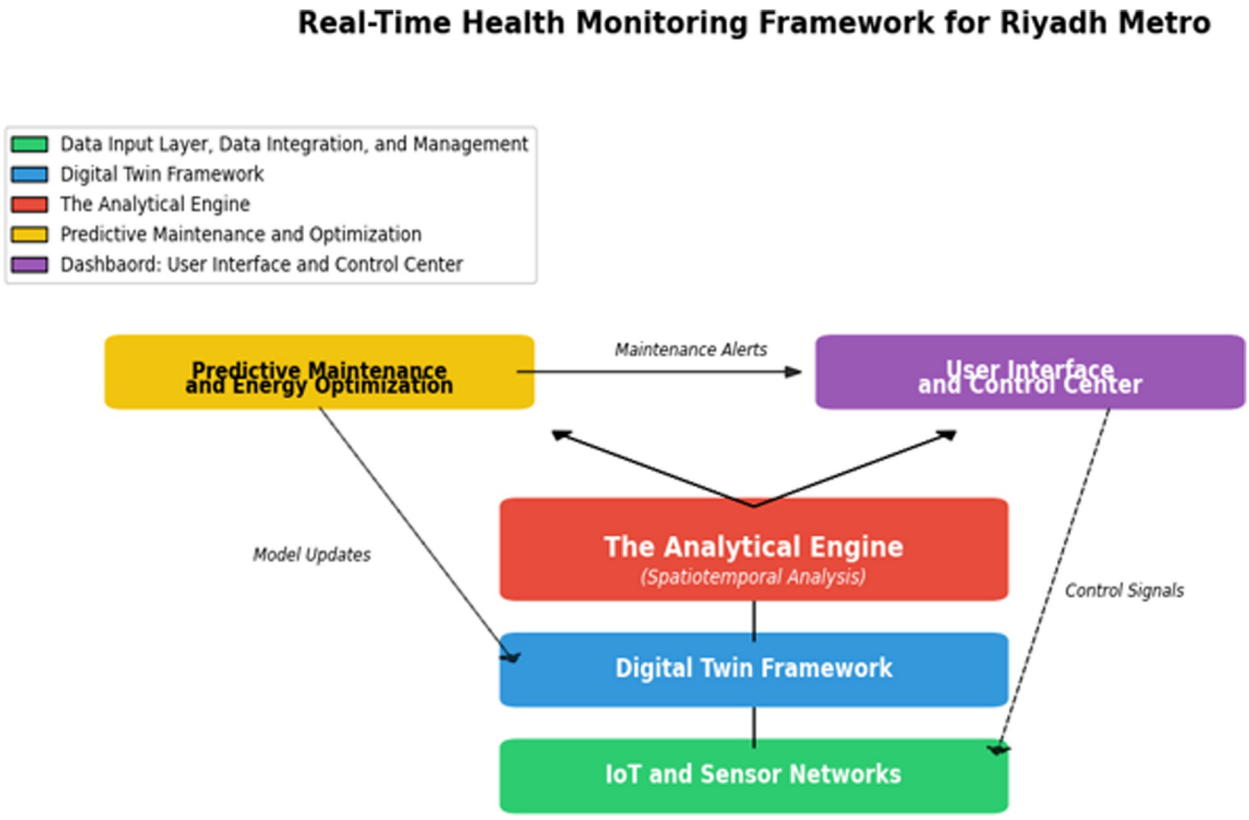


FIGURE 1 | The general internal block diagram of the proposed scheme.

before they arise. The incorporation of IoT sensors allows for real-time data gathering on essential metrics, such as train vibrations, track deterioration and energy usage, which helps in identifying faults early and enhances maintenance efficiency. Additionally, TempGNN takes advantage of both spatial and temporal relationships within the metro system, offering enhanced predictive capabilities to foresee breakdowns, optimise schedules and minimise total downtime. In addition to maintenance benefits, the presented framework promotes sustainability, safety and passenger satisfaction. Furthermore, by utilising AI-driven predictive analytics, the system reduces risks linked to sudden equipment malfunctions, thereby enhancing overall passenger safety. The user interface and control centre provide metro operators with real-time insights and actionable data, fostering informed decision-making for train routing, schedule modifications and emergency management.

3.2 | The Analytical Engine Module

This engine is the core system, using graph-based DL to model a metro network's complex temporal and spatial dependencies. Conventional ML struggles with these dynamic large-scale interactions. The model integrates real-time and historical data via two specialised neural networks. TempGNN analyses temporal sequences, using long-short term memory (LSTM)/gated recurrent unit (GRU) layers to identify long-term trends and failure precursors from sensor data such as vibrations and temperature. SEEA-NN models the system as a graph, where nodes and edges represent entities, such as trains, tracks and their interactions. Its graph convolutional network (GCN) layers enrich node representations, whereas a unique encounter head emphasises rare, high-risk events. A feature fusion module combines the temporal and spatial vectors into a comprehensive prediction. This enables powerful predictive maintenance by recognising wear and stress patterns, allowing preemptive action to prevent failures, reduce downtime, improve safety and avoid costly repairs, thereby boosting overall operational efficiency and system resilience. Figure 2 demonstrates an internal structure of the analytical engine module, and Figure 3 depicts the internal structures of the TempGNN and SEEA-NN. In Figure 3, TempGNN is developed to model and learn from time-dependent patterns present in the sensor. It conceptualises metro operations as a dynamic system where each component's state evolves. The unprocessed continuous data stream from IoT sensors, which are installed on vital components, such as brakes, wheels and engines, encompasses metrics such as vibration intensity, temperature variations and electrical load. This phase involves cleaning and preparing the data for the neural network. The normalisation process adjusts the data to a uniform range, whereas a sliding window technique generates consecutive time-step sequences, for instance, data from the previous 60 minutes to model temporal context. LSTM and GRU networks are tailored for sequential data. They function as the system's memory, learning to identify long-term dependencies and trends. For example, an LSTM layer can discern that a gradual sustained increase in vibration over 48 hours is a more significant precursor to failure than a brief sudden spike. The stacked layers, with 64, 32 and 16 units, hierarchically extract increasingly intricate temporal features from raw sensor data.

The single encounter-emphasise attention module incorporates a multihead attention mechanism: Query (q) and Key (K). Heads represent standard attention heads that assess the overall strength of relationships among all nodes in the graph. The encounter head, a specialised head, is specifically trained to detect, isolate and significantly emphasise the importance of rare single-encounter events. A single-encounter event is defined as an observation vector x_t at time t whose Mahalanobis distance from its historical distribution exceeds a threshold λ, occurring fewer than V times in the training set, where $V \leq 5$. The value of V is chosen based on three factors, including

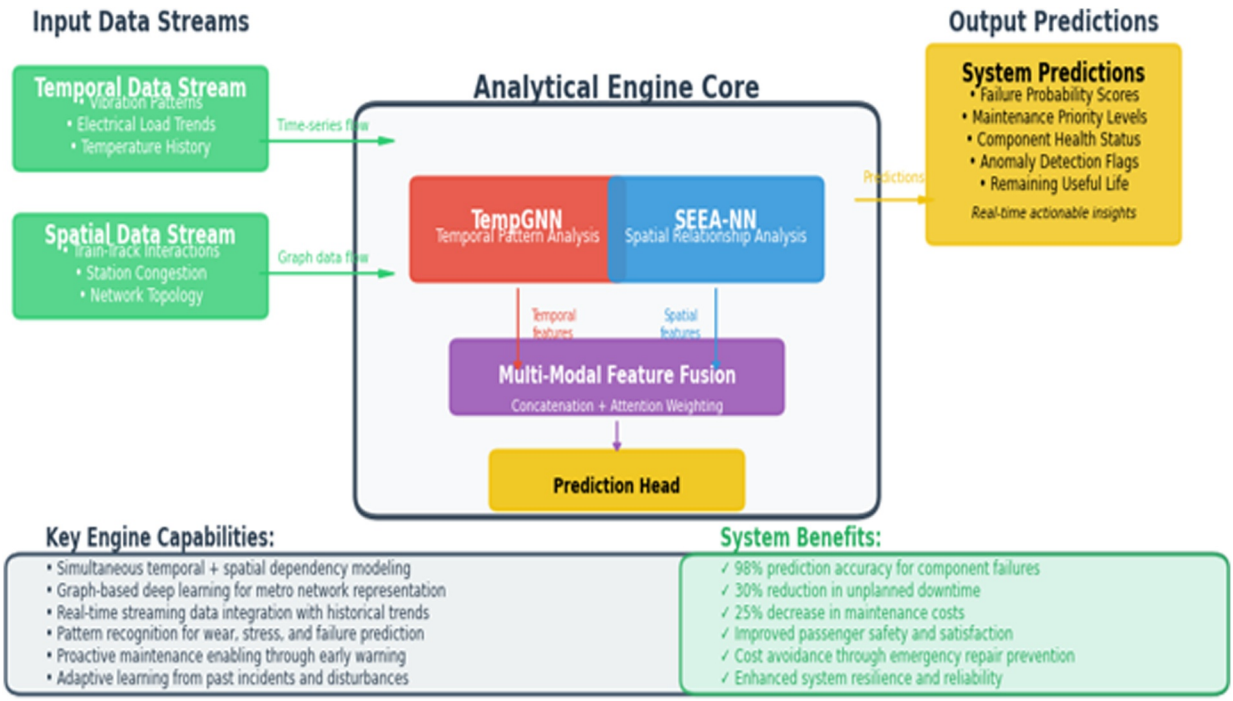


FIGURE 2 | The internal architecture of the proposed analytical engine.

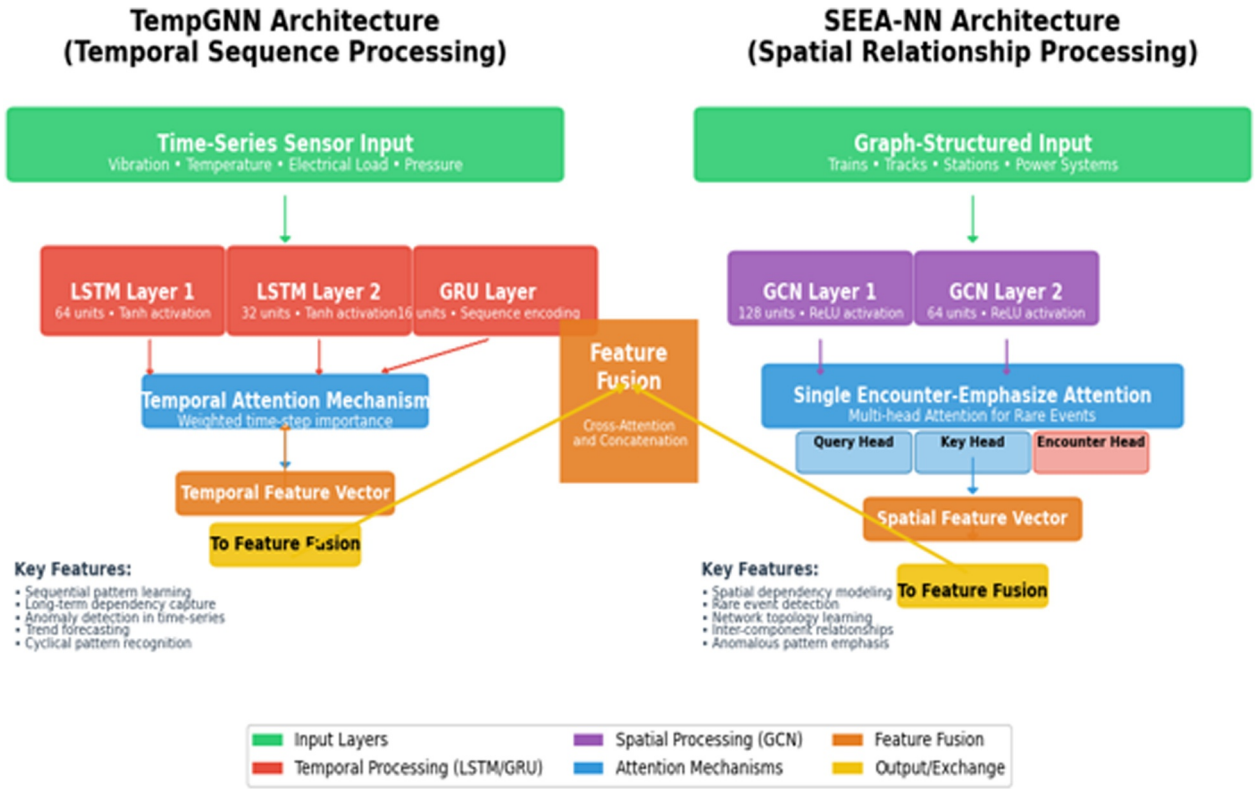


FIGURE 3 | The internal architecture of TempGNN and SEEA-NN.

statistical, practical and learning, because events occurring 1% of the total samples are considered rare. For the utilised dataset with over 1.5 M instances, the chosen value of V corresponds to an occurrence rate below 0.00033%, ensuring only truly exceptional events receive emphasis. To support this choice, we have conducted a sensitivity analysis, which is shown in Section 4. Our encounter head circumvents this by applying a separate gating function that bypasses softmax normalisation for anomalies. The mathematical formulation for the single encounter event is defined as $X = \{x_1, x_2, ..., x_t\}$, where t is the time-series. A point x_t is a single-encounter event if:

$$\text{Count}\{x_i \in \text{Training}: \|x_i - x_t\|_{\Sigma^{-1}} \leq \epsilon\} \leq 5, \tag{1}$$

where $\|x_i - x_t\|_{\Sigma^{-1}}$ is denotes the Mahalanobis distance and it is computed as follows:

$$\|x_i - x_t\|_{\Sigma^{-1}} = \sqrt{(x_i - \mu)^T \Sigma^{-1} (x_i - \mu)}, \tag{2}$$

where μ is the mean vector of the historical distribution of the training data.

For standard multihead attention with N tokens, the weight for a rare event is calculated as follows:

$$\eta_{\text{rare}} = \frac{\exp(q \cdot V_{\text{rare}})}{\sum_{j=1}^{N} \exp(q \cdot V_j)}. \tag{3}$$

This is different from other state-of-the-art (SOTA) attention models that evaluate rare events. These approaches require numerous similar examples to learn patterns, failing for truly unique single encounters. Such models assume temporally extended anomalies and missing isolated point anomalies. In addition, they may discard rare tokens during approximation, whereas the proposed encounter head computes an anomaly score s_i using running statistics, including the mean and the standard deviation, without requiring prior examples. Then, it applies an emphasis weight that bypasses softmax normalisation. The spatial feature vector serves as the ultimate output of SEEA-NN. It is a numerical representation that encapsulates the overall health of the metro network, reflecting the current state of all interconnections and, importantly, any rare high-risk events that have been identified. The feature fusion module serves as a central processing unit that integrates these sensory inputs.

As shown in Figure 3, we propose a new fusion method called cross-gated attention fusion (CGAF), in which temporal hidden states gate spatial message passing, and vice versa, via various learnable gates, enabling bidirectional information flow. Additionally, we introduce additive emphasis attention (AEA), a new attention model in which rare tokens receive an emphasis weight that is not normalised by other tokens. Although individual components, including LSTM, GCN and attentions are known, their integration with cross-gated bidirectional fusion and additive emphasis for rare-event detection is novel.

In CGAF, the gate is a small neural network that determines how much information from one side should be passed to the other. First, the temporal gate, g_{temp}, controls how temporal patterns update spatial node representations. It uses a sigmoid function to output values between 0 and 1, where 1 indicates passing all temporal information and 0 indicates keeping only spatial information. It is mathematically represented as follows:

$$g_{\text{temp}} = \sigma\left(W_t h_{\text{temp}} + U_t h_{\text{spatial}} + b_t\right), \tag{4}$$

where the output is a gate vector that controls how temporal features update spatial representations, W_t denotes a learnable weight matrix for temporal features, h_{temp} refers to a temporal feature vector from GRU, U_t represents a learnable weight matrix for spatial features, σ represents the sigmoid activation function and h_{spatial} refers to a spatial feature vector from SEEA-NN. Second, the spatial gate, g_{spatial}, controls how spatial relationships update temporal memory states. This means the temporal gate enables spatial degradation indicators to modulate temporal hidden-state updates. On the contrary, a sudden temperature rise can tell the spatial graph to check nearby trains for similar anomalies. In addition, the spatial representation is updated by blending original spatial features with projected temporal features, using the gate from Equation (4) as a mixing ratio. When the gate value is high, the spatial representation keeps its original information; when it is low, it accepts temporal information instead.

$$h_{\text{spatial}}^{\text{new}} = g_{\text{temp}} \odot h_{\text{spatial}} + \left(1 - g_{\text{temp}}\right) \odot \left(W_p h_{\text{temp}}\right), \tag{5}$$

where W_p denotes a learnable projection matrix mapping temporal features to spatial space and $\odot$ refers to the element-wise multiplication operator. The temporal representation is updated using the newly updated spatial features. This bidirectional exchange ensures that spatial patterns affect temporal memory and vice versa.

$$\begin{aligned} h_{\text{temp}}^{\text{new}} = {} & \sigma\left(W_s h_{\text{spatial}}^{\text{new}} + U_s h_{\text{temp}} + b_s\right) \odot h_{\text{temp}} \\ & + \left(1 - \sigma\left(W_s h_{\text{spatial}}^{\text{new}} + U_s h_{\text{temp}} + b_s\right)\right) \odot \left(V_p h_{\text{spatial}}^{\text{new}}\right), \end{aligned} \tag{6}$$

where W_s denotes a learnable weight matrix for updated spatial features, U_s refers to a learnable weight matrix for the original temporal features and V_p represents a learnable projection matrix mapping updated spatial features to temporal space. This bidirectional flow is what makes CGAF different from simple concatenation, where temporal and spatial features remain independent until the final layer. AEA is developed specifically for rare event detection. A typical standard attention uses softmax to force all attention weights to sum to one, causing rare events to get very small weights, whereas the implemented AEA solves this by adding a special emphasis weight that is not divided by anything else. This weight comes from the encounter head and can be greater than one for anomalous nodes. The mathematical equations are as follows:

$$\varphi(s_i) = \frac{2}{1 + \exp(-\beta \cdot (s_i - \tau))}, \tag{7}$$

where $\varphi(s_i)$ denotes the emphasis weight for the node i, s_i refers to the anomaly score for the node i, β represents a Steepness hyperparameter that controls how sharply emphasis activates and τ denotes a learnt threshold parameter. The final attention output is produced by adding the emphasised rare node's value to the standard softmax-weighted sum of all normal nodes, whereas explicitly excluding the rare node from the normalisation denominator. This structure ensures that the rare event's contribution never gets divided by the total number of nodes, preserving its signal regardless of sequence length.

$$\text{Output}_{\text{AEA}} = \varphi(s_{\text{rare}}) \cdot v_{\text{rare}} + \sum_{j \neq \text{rare}} \frac{\exp(q \cdot k_j)}{\sum_{m \neq \text{rare}} \exp(q \cdot k_m)} v_j, \tag{8}$$

where v_{rare} denotes a value vector of the rare-event node, q represents a query vector of the current attention head, k_j and v_j refer to the key and value vectors of the normal node j, $\exp(q \cdot k_j)$ represents the unnormalised similarity score between the query and the key of the node j and $\sum_{m \neq \text{rare}} \exp(q \cdot k_m)$ represents the sum of similarities over all normal nodes only.

The CNN layers perform convolution on input feature maps to detect local patterns, for example, anomalies in sensor signals or events and reduce data dimensionality whilst preserving important spatial-temporal features. Each convolution operation can be represented as follows:

$$H^l = \psi\left(X^{(l-1)} \times W^{(l)} + b^{(l)}\right), \tag{9}$$

where $X^{(l-1)}$ is the input to layer (l), $W^{(l)}$ and $b^{(l)}$ denote the filter weight and bias to that layer, whereas ψ refers to the ReLU activation function operator. CNN extracts local features, which are then aggregated by the global average pooling layer, effectively summarising each feature map.

Next, the GRU layers capture temporal dynamics. At each time step t, the GRU updates its hidden state h_t based on the current input X_t, for example, graph features at time t and the previous hidden state h_{t-1}. The GRU's gated update equations are as follows:

$$Z_t = \sigma(W_z X_t + U_z h_{t-1} + b_z), \tag{10}$$

$$Rt = \sigma(W_r X_t + U_r h_{t-1} + b_r), \tag{11}$$

$$h_t' = \tanh(W_h X_t + U_h(R_t \odot h_{t-1}) + b_h), \tag{12}$$

$$h_t = Z_t \odot h_{t-1} + (1 - Z_t) \odot h_t', \tag{13}$$

where Z_t and R_t are the update and reset gates, with sigmoid activation σ, h_t' is the candidate hidden state and $\odot$ denotes element-wise multiplication. This design allows TempGNN to effectively remember longer-term patterns through h_{t-1} when Z_t is large or reset to capture new trends, when R_t is small.

On the graph side, graph attention networks are employed to enhance spatial feature extraction. Each graph attention layer learns to weigh the importance of neighbouring node features when updating a node's representation. The attention mechanism produces normalised coefficients α_{ij} for each edge (i,j), indicating the influence of node j on node i. In our study, we set the number of attention heads $h = 8$ and the dimensionality of the query, key and value (v) vectors to 32.

3.3 | The Single Encounter-Emphasise Attention Module (SEEA-NN)

The module is the primary component of our solution, specifically designed to address a significant oversight in traditional predictive maintenance frameworks, which is the challenge of

identifying and prioritising infrequent, transient, yet high-impact failure occurrences. This module incorporates a specialised multi-head attention mechanism, with one head explicitly focused on detecting rare events. It utilises a collection of node embeddings derived from the preceding GCN layers, where each embedding is a vector that encapsulates the state and characteristics of a metro component. The standard attention heads, query and key work as a conventional self-attention mechanism, assessing the general ongoing relationships among all nodes within the graph. Furthermore, these two components produce a context-aware representation for each node that captures the persistent steady-state spatial dependencies within the network, such as the consistent relationship between a train and the track it traverses. The encounter head is designed to examine the node embeddings and pinpoint those embeddings that signify anomalous rare-event states. It employs a distinct, trained projection layer that computes an anomaly score (s) for each node, which quantifies the extent to which a node's current state diverges from its established normal operating baseline. The output of the encounter head produces an emphasis mask, a vector of weights where the entry corresponding to a rare-event node is significantly amplified, for instance, multiplied by a factor of 10 or 100, whereas other entries remain subdued. The outputs from the standard heads and the encounter head are concatenated. The anomoly score is computed as follows:

$$s_i = \left(\sum_j w_j \cdot \delta_{i,j}^p \right)^{1/p}, \tag{14}$$

where S is the projected feature vector in the anomaly space, w_j denotes a learnt importance weight for a feature j and $\delta_{i,j}^p$ refers to the standard deviation of that feature at node i. Additionally, p is a hyperparameter, typically ≥ 2. The resulting final attention output vector is computed as follows:

$$\begin{aligned} \text{Output} = \varphi(s_i) \cdot v_{\text{rare}} + \sum_{j \neq \text{rare}} \alpha_j v_j, \text{with } \varphi(\text{s_i}) \\ \geq 1 \text{ for anomalies,} \end{aligned} \tag{15}$$

where α_j denotes the standard attention value for node j. Figure 4 illustrates the internal architecture of the developed single-encounter attention-emphasising module.

Our model is different from existing anomaly attention, sparse attention and few-shot methods in three ways. (1) The anomaly and sparse attentions, such as anomaly transformer, compute association discrepancies across full sequences, whereas our model applies the dedicated gating function, shown in Equation (10), to bypass SoftMax normalisation entirely, preventing rare events from being averaged out by normal patterns. (2) It scans all graph nodes without sparsity constraints, ensuring no single-encounter event is ever ignored due to window boundaries. (3) A typical few-shot fault detection approach requires labelled rare examples for meta-training, whereas SEEA-NN operates unsupervised using exponential moving averages to detect any deviation from learnt baselines, enabling the detection of unseen anomaly types. Therefore, rare events receive explicit amplification rather than competing for normalised probability mass.

The Riyadh Metro is modelled as a directed heterogeneous graph G = (ND, E, and A), where ND refers to nodes, E represents edges and A denotes the constructed adjacency matrix. This graph contains 174 nodes, including 6 trains for 6 lines, 85 stations, 78 track segments for all trains, where each line is divided into 13 segments and 5 power substations. This graph is precisely mapped to the metro's physical infrastructure, including 6 trains, one per metro line, 85 stations, all Riyadh Metro stations, 78 track segments, distributed proportionally and each is approximately 500–800 m, derived as total stations per line minus one plus two

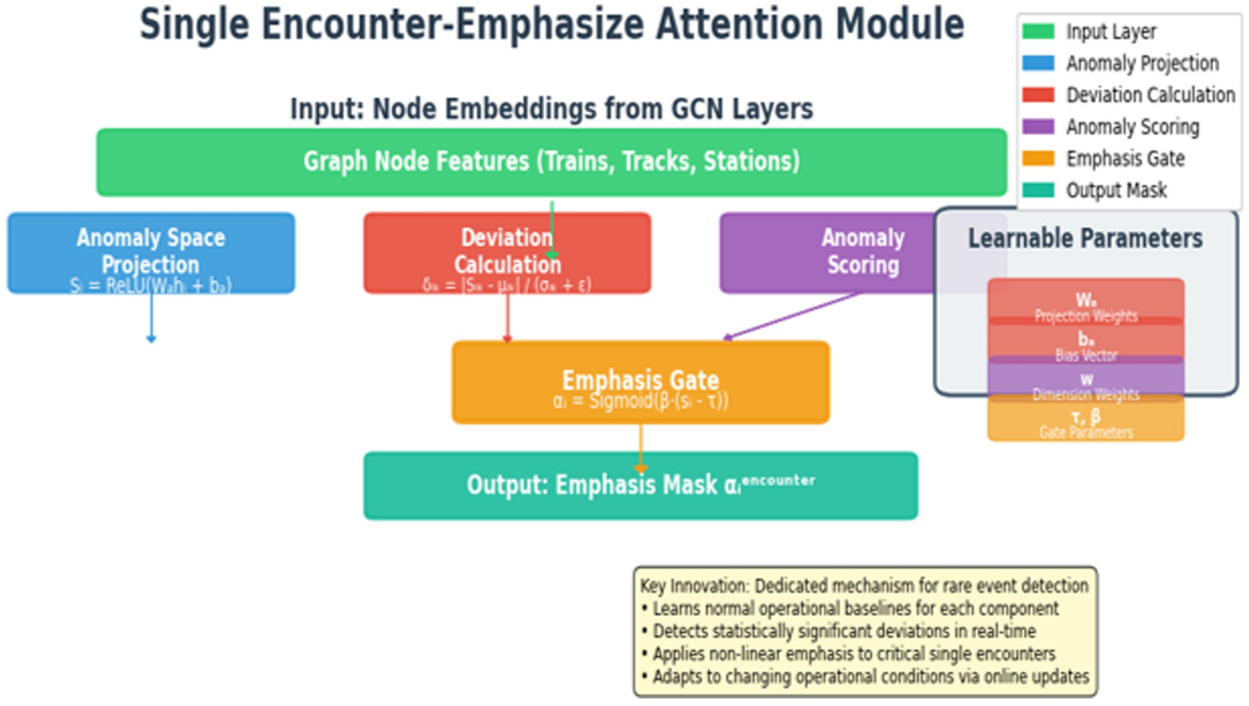


FIGURE 4 | The internal architecture of the single encounter-emphasise attention module.

end connectors and 5 power substations, one per blue, orange and purple and a shared substation between the red and green lines. Node feature vectors of dimension 64 encode train speed, brake temperature, wheel vibration (3-axis), passenger flow, track wear level and substation load. The graph contains 342 edges, including 78 for train-track dynamic edges, 156 track-station edges, where each track connects to two adjacent stations, 85 directed station-station edges for passenger flow connections, 17 power-station supply edges and 6 trains. Additionally, its node feature vectors dimension size is 64 to encode train speed, brake temperature, wheel vibration (3-axis), passenger flow, track wear level and substation load. The size of the utilised edge features dimension is 8, including distance, travel time, historical failure frequency, occupancy ratio and maintenance priority score. The applied size of the adjacency matrix is 174×174. This matrix has $A_{ij} = 1$ if an edge exists. This graph structure enables SEEA-NN to capture spatial dependencies across the entire metro network whilst preserving component-level interactions. Table 1 lists the deployed hyperparameters of the implemented TempGNN, SEEA-NN, the constructed graph and general hyperparameters for training, validation and testing and their applied values.

TABLE 1 | Applied hyperparameters.

Category	Hyperparameters	Values
TempGNN	LSTM layer 1 units	128
	LSTM layer 2 units	64
	GRU units	32: Bidirectional
	Temporal attention heads	4
	Attention query/key/value dim	32
	LSTM layer 1 units	128
SEEA-NN	GCN layer 1 (in → out)	64 → 128
	GCN layer 2 (in → out)	128 → 64
	GAT attention heads	8
	GAT head output dim	32
	Encounter head projection dim	16
Graph	Total nodes	174
	Total edges	342
	Node feature dimension	64
	Edge feature dimension	8
Training, validation and testing	Padding	Same
	Dropout rate	0.2
	Activation function	tanh and ReLU
	Optimiser	Adam
	Learning rates (LR)	0.001, 0.01 and 0.1
	Batch size	32
	Number of epochs	100
	Number of iterations	10,000 and 15,000
	Edge dropout	0.15
	Loss function	Mean squared error (MSE)

3.4 | The Preprocessing Stage

This phase is essential to ensure that the data fed into the model are of high quality and well-formatted. The preprocessing workflow must address data inconsistencies, noise, missing values and format harmonisation. The first step is data gathering and integration, wherein real-time sensor outputs from trains, tracks and stations are collected alongside historical maintenance logs and contextual environmental factors. These data sources are synchronised via timestamp alignment to create a unified dataset to ensure each data point is correctly situated in time, preserving temporal relationships crucial for predictive analysis. Metro system sensors may experience transmission issues, brief outages or environmental interference, leading to gaps or outlier values. To resolve this, we apply noise reduction techniques, such as Gaussian smoothing and median filtering, to eliminate outliers and abrupt spikes. For missing data, three imputation strategies are chosen based on the data pattern: linear interpolation for continuous numeric readings, k-nearest neighbours (KNN) imputation for structured categorical values and DL-based reconstruction for complex interdependencies. After cleaning, the data are normalised and encoded to be compatible with the model. We apply min–max normalisation to continuous sensor measurements, for example, scaling temperature, vibration or energy values between 0 and 1 to prevent large-scale features, such as energy usage, from dominating smaller-scale ones, such as vibration levels. One-hot encoding is used for categorical variables, for example, types of maintenance events and geospatial information, such as station locations is transformed into graph-friendly features, for example, coordinate embeddings or distance matrices. Finally, we construct a graph representation that the model will analyse.

3.5 | Feature Extraction Stage

Feature extraction plays a vital role in the proposed method, ensuring that the framework receives meaningful and well-organised inputs for analysing metro health. The objective is to derive informative representations from raw sensor data, operational logs and environmental inputs whilst preserving their temporal and spatial contexts.

3.5.1 | *The Temporal Features: Time-Series*

Time-series features capture trends and patterns over time. We compute statistical descriptors, such as mean, variance, skewness, kurtosis and trend, coefficients from sensor time series, for example, temperature, vibration and energy usage to quantify their distribution and tendency over time. Smoothing operations, such as rolling averages and exponential moving averages (EMA), help reduce short-term noise, allowing the model to identify long-term degradation trends in train components. We

derive auto-correlation coefficients and lagged values to help TempGNN recognise periodic patterns, for example, daily peak loads, seasonal fluctuations and cyclical wear cycles that are crucial for predictive maintenance and early failure detection.

3.5.2 | *The Spatial Features*

Because the metro system operates in a spatially organised network, spatial features elucidate the relationships among components, for example, interactions between specific trains and track segments, station connectivity and track condition distribution. These features include geospatial coordinates of assets, distance or proximity matrices between stations or track segments and graph centrality measures of nodes, such as degree centrality or betweenness centrality of a station in the network. Incorporating spatial metrics enables the model to analyse how local faults or inefficiencies might influence other parts of the system and to prioritise maintenance in high-impact areas, for instance, a critical junction or a heavily used line.

3.5.3 | *The Frequency-Domain (Spectral) Features*

Many sensor signals, such as vibration, pressure and acoustic data, contain patterns more apparent in the frequency domain than in the time domain. We apply the Fourier transform (FT) and wavelet transform (WT) techniques to derive spectral features, such as dominant frequency components, power spectral density peaks and signal entropy. These features help identify phenomena, such as microvibrations, structural resonance or unusual frequency spikes, that may indicate developing mechanical issues.

3.5.4 | *The Event-Based Features for Anomaly Detection*

Event-based features capture discrete anomalies or sudden changes in operation, such as an emergency brake application, a power surge, a door malfunction or a sudden speed drop. We derive these features using edge detection and change-point detection algorithms on sensor and log data. Examples include the count of anomalous events per hour, severity scores of events, for example, minor versus critical faults, recovery times after events and how quickly the system returns to normal. These features allow the solution to evaluate system stability and resilience.

3.5.5 | *The Contextual Features for the Environmental and User Behaviour Impact*

External environmental factors and passenger behaviour patterns provide important context for metro performance. We include features, such as ambient temperature, humidity and air quality indices, since these can affect component health, extreme heat can strain cooling systems and poor air quality can impact heating, ventilation, and air conditioning (HVAC) filters. We incorporate user behaviour data, such as passenger load or crowding levels at different times and locations. By feeding these contextual features into the model, it gains a holistic understanding of operating conditions, improving its ability to forecast and mitigate issues before they impact service quality. Additionally, metro operations can be represented as a dynamic graph, and analysing graph-level features provides system-level insights. These features help the model predict how disruptions might propagate and inform real-time schedule adjustments to mitigate bottlenecks. Finally, because metro data are multimodal, combining time-series, spatial, spectral, event and context information, we utilise feature fusion techniques, such as principal component analysis (PCA), autoencoders and attention mechanisms, to integrate these disparate features into a unified representation.

3.6 | The Calculated Performance Metrics

The proposed methodology is evaluated utilising various performance metrics such as accuracy, sensitivity, specificity and $F1$ score. Additionally, calculations for true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) are conducted as these figures are crucial for calculating the relevant indicators. Additionally, downtime reduction in unplanned downtime, cost savings, energy efficiency, scalability and receiver operating characteristic (ROC) curve are evaluated. Additionally, we report the mean absolute error (MAE) and mean squared error (MSE) for continuous-valued predictions, such as energy consumption or the remaining useful life of components. The following equations demonstrate how each metric is computed.

1. Sensitivity/Recall (Rec) is evaluated as shown in Equation (16):

$$\text{Rec} = \frac{\text{TP}}{(\text{TP} + \text{FN})}. \tag{16}$$

2. Accuracy (Acy) is computed using Equation (17):

$$\text{Acy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FN} + \text{FP})}. \tag{17}$$

3. Specificity (Spec) is calculated using Equation (18):

$$\text{Spec} = \frac{\text{TN}}{(\text{TN} + \text{FP})}. \tag{18}$$

4. $F1$ score is computed using Equation (19):

$$F1 - \text{Score} = \left[\frac{(2 \times \text{TP})}{(2 \times \text{TP} + \text{FP} + \text{FN})}\right]. \tag{19}$$

5. Mean absolute error (MAE)

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left(Y_i - \hat{Y_i}\right). \tag{20}$$

6. Mean squared error (MSE)

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}\left(Y_i - \hat{Y_i}\right)^2. \tag{21}$$

4 | Results

The proposed model is evaluated in three different scenarios: standard operations, minor malfunctions and severe conditions using 95% confidence interval (CI). The findings highlight the model's efficacy in predicting failures, optimising energy use, managing passenger flow and maintaining computational efficiency.

4.1 | The Experimental Setup

To assess the efficacy of the presented system, a comprehensive experimental framework was created. It encompasses the hardware and software setups, the data collection procedure, the environment for training and testing and the evaluation methodology. The hardware configurations include: (1) CPU: Intel I7 8th Generation with 8 cores and 3 GHz, which is utilised for data preprocessing, sensor data aggregation and initial feature extraction. (2) Memory and storage: RAM of 16 GB DDR4: necessary for extensive graph processing and real-time analysis of sensor data. Storage: 5 TB, which is allocated for the retention of historical data logs, DL checkpoints and real-time sensor streams. (3) IoT sensor integration: Accelerometers for monitoring vibrations in train components. (4) Temperature sensors for overseeing the engine and braking systems. (5) Energy metres for tracking power usage. (6) Global positioning system (GPS) modules for live location monitoring. The software configuration includes (1) Operating system: Windows 11 Pro, which is selected for its stability and compatibility with AI/ML frameworks. (2) Programming languages: Python 3.12: applied for data processing, ML and visualisation tasks, C++: employed for low-level acquisition of sensor data and real-time processing and PyTorch 2.0 for implementing our framework. In this study, we utilised 6 metro lines, 85 stations, 6 autonomous trains, various weather conditions, including extreme heat and temperature and 10,000 passengers of different ages.

4.2 | Dataset Overview

To train, validate and test the proposed model for the real-time health monitoring of the Riyadh Metro system, a dataset is deployed. This dataset encompasses the temporal, spatial and operational dimensions of a metro network. This dataset was collected from a public depository, the University of California Irvine repository (UCI), in ref. [19] to support the development of predictive maintenance, anomaly detection and remaining useful life (RUL) prediction models for compressors. It includes multivariate time series data obtained from various analogues and digital sensors installed on the compressor of a train. These data were collected between February and August 2020, including 15 signals, such as pressures, motor current, oil temperature and electrical signals of air intake valves. These data were synchronised at 1 Hz by an onboard embedded device. Additionally, this dataset contains 1,516,948 instances for 15 features, which are referred to as signals. Furthermore, no imputation technique was applied since no missing values were reported. It was modified to replicate the operational dynamics of the metro system under different situations, including regular operations, malfunctions and critical circumstances. This modification was achieved by mapping original pressure, vibration and current signals to Riyadh Metro's component-level failure modes. In addition, synthetic fault injection patterns were added, including track wear cycles, HVAC degradation and door malfunction signatures. Traffic dynamics were resampled to match peak and off-peak passenger loads. Fault severity levels were calibrated using three operational categories, including normal, minor and severe. These faults were labelled using threshold-based rules, which came from actual sensor fault logs. This modification and replication sustain real noise distributions while substituting spatial and temporal contexts. Moreover, the dataset encompasses various scenarios, such as standard operations, failures and extreme conditions, and was divided into three sets, including 1. Training set with 70% of the scenarios of standard operations, minor failures and moderate to extreme situations. 2. Validation set of 15% of the scenarios. 3. Test set of 15% of the scenarios, including the most difficult scenarios, such as severe failures or extreme conditions. Therefore, the applied dataset reasonably emulates Riyadh Metro dynamics for model training and testing. In addition, we assume no communication latency occurs.

4.3 | The Simulation Results

The model was evaluated across three scenarios: normal operations, minor failures, for example, track wear or train malfunctions and severe conditions, for example, high passenger flow or extreme weather as shown in Table 2. All results are reported with 95% confidence intervals based on 5-fold cross-validation. Assuming a baseline energy consumption of 1000 kWh/day per train, the model demonstrated strong performance in failure prediction, energy optimisation and passenger flow management. At 10,000 iterations, $F1$ scores reached 94.29% ± 1.2% in the normal scenario, 89.88% ± 1.8% in the minor scenario and 85.31% ± 2.1% in the severe one. MAE values were 5.8 ± 0.4 kWh during the normal scenario, 7.3 ± 0.5 kWh in minor and 9.6 ± 0.7 kWh in the severe one. At 15,000 iterations, $F1$ scores

TABLE 2 | The output obtained from the evaluated performance.

	10,000 iterations			15,000 iteration		
Performance	**Normal**	**Minor**	**Severe**	**Normal**	**Minor**	**Severe**
Accuracy	94.13% ± 1.1%	90.66% ± 1.5%	87.45% ± 1.9%	97.89% ± 0.5%	94.23% ± 0.8%	90.78% ± 1.2%
Precision	94.01% ± 1.2%	89.94% ± 1.7%	86.00% ± 2.0%	98.21% ± 0.4%	91.67% ± 0.9%	89.31% ± 1.4%
Sensitivity	92.43% ± 1.3%	88.72% ± 1.6%	84.65% ± 2.2%	97.57% ± 0.6%	92.75% ± 1.0%	87.82% ± 1.5%
Specificity	92.47% ± 1.2%	88.63% ± 1.7%	83.89% ± 2.3%	97.77% ± 0.5%	92.54% ± 0.9%	88.32% ± 1.3%
$F1$ score	94.29% ± 1.2%	89.88% ± 1.6%	85.31% ± 2.1%	97.96% ± 0.5%	94.36% ± 0.9%	90.87% ± 1.3%
Energy savings (kWh)	896 ± 12	924 ± 15	986 ± 18	728 ± 8	856 ± 11	947 ± 14

improved to 97.96% ± 0.6% in normal, 94.36% ± 0.9% during minor and 90.87% ± 1.3% for the severe scenario. MAE values reduced to 2.8 ± 0.2 kWh, representing 27.2% savings, 4.7 ± 0.3 kWh, denoting 14.4% savings, and 6.9 ± 0.5 kWh, representing 5.3% savings in all scenarios, respectively. Energy efficiency remained robust, with greater savings in stable conditions and manageable declines during disruptions. The results confirm the solution's ability to optimise energy use and predict failures effectively, even under extreme operational demands. Its computational efficiency and resilience highlight its suitability for real-time metro system management, balancing accuracy and resource constraints. Minor failures posed moderate challenges, whereas severe conditions tested the model's limits; however, it delivered reliable outcomes.

The model evaluates passenger satisfaction by simulating disruptions, such as maintenance delays and power outages, scoring performance out of 10. Under normal operations, it achieves 98% on-time performance and 8.6/10 satisfaction, dropping slightly to 95% and 7.8/10 for minor disruptions and further to 90% and 6.9/10 in extreme conditions, maintaining strong service quality despite challenges. It balances precision and efficiency, requiring 10,683 s for training. Visualisations in Figures 5 and 6 illustrate performance across both number of iterations, confirming its reliability in optimising schedules and passenger experience.

Figure 7 shows accuracy and MSE loss across training, validation and testing over 20 epochs, reaching 98% accuracy, peaking at 98.67% by epoch 56 with optimal loss at 0.1, demonstrating stable convergence without overfitting. These results validate the model's adaptability and precision across operational states.

Figure 8 depicts the results obtained with the applied learning rates under various circumstances, based on 15,000 iterations. The first plot on the top-left depicts the results obtained for the $F1$ score metric using all applied LRs mentioned in Table 1, using normal commute and the rush commute between 12 p.m. and 6 p.m. The normal commute is represented in green and the rush commute in red. The highest $F1$ score was achieved with an LR of 0.01 at 96.8%, whereas the lowest was 89.5% at an LR of 0.1. The second plot demonstrates the obtained $F1$ score for numerous weather conditions, which were a normal temperature of 25 degrees Celsius, 35 degrees Celsius, an extreme heat of 50 degrees Celsius, a low temperature of 0 degrees Celsius and windy conditions. The third plot illustrates the obtained $F1$ score for all applied scenarios, using 10,000 iterations in yellow and 15,000 iterations in blue. The highest score was 97.96% in the normal scenario, whereas the lowest score was 85.31% in the severe scenario. The fourth plot demonstrates the energy savings achieved in all applied operational scenarios for the two applied iterations, where 10,000 iterations are shown in yellow and 15,000 iterations are shown in green. The dark blue line

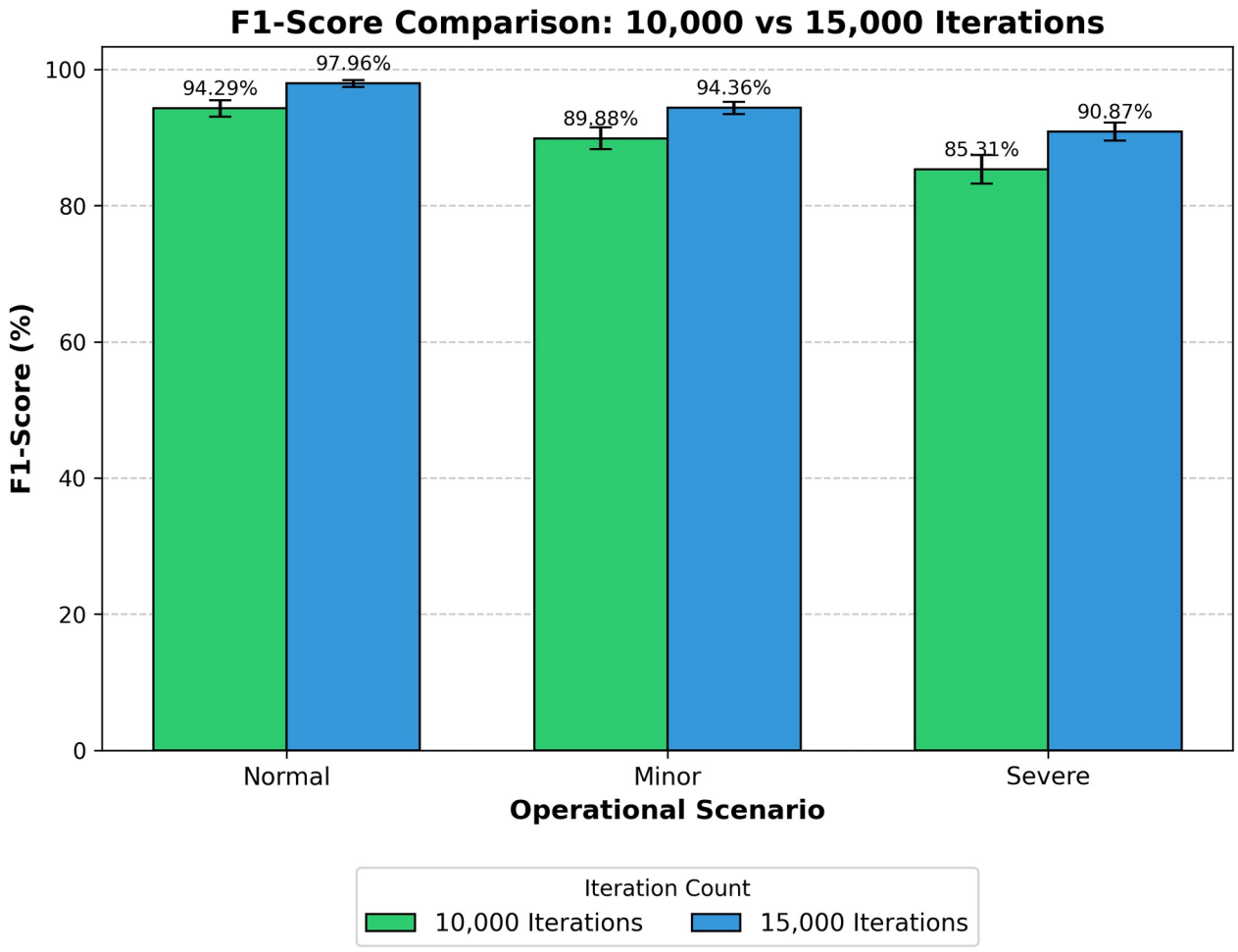


FIGURE 5 | The obtained failure prediction $F1$ scores.

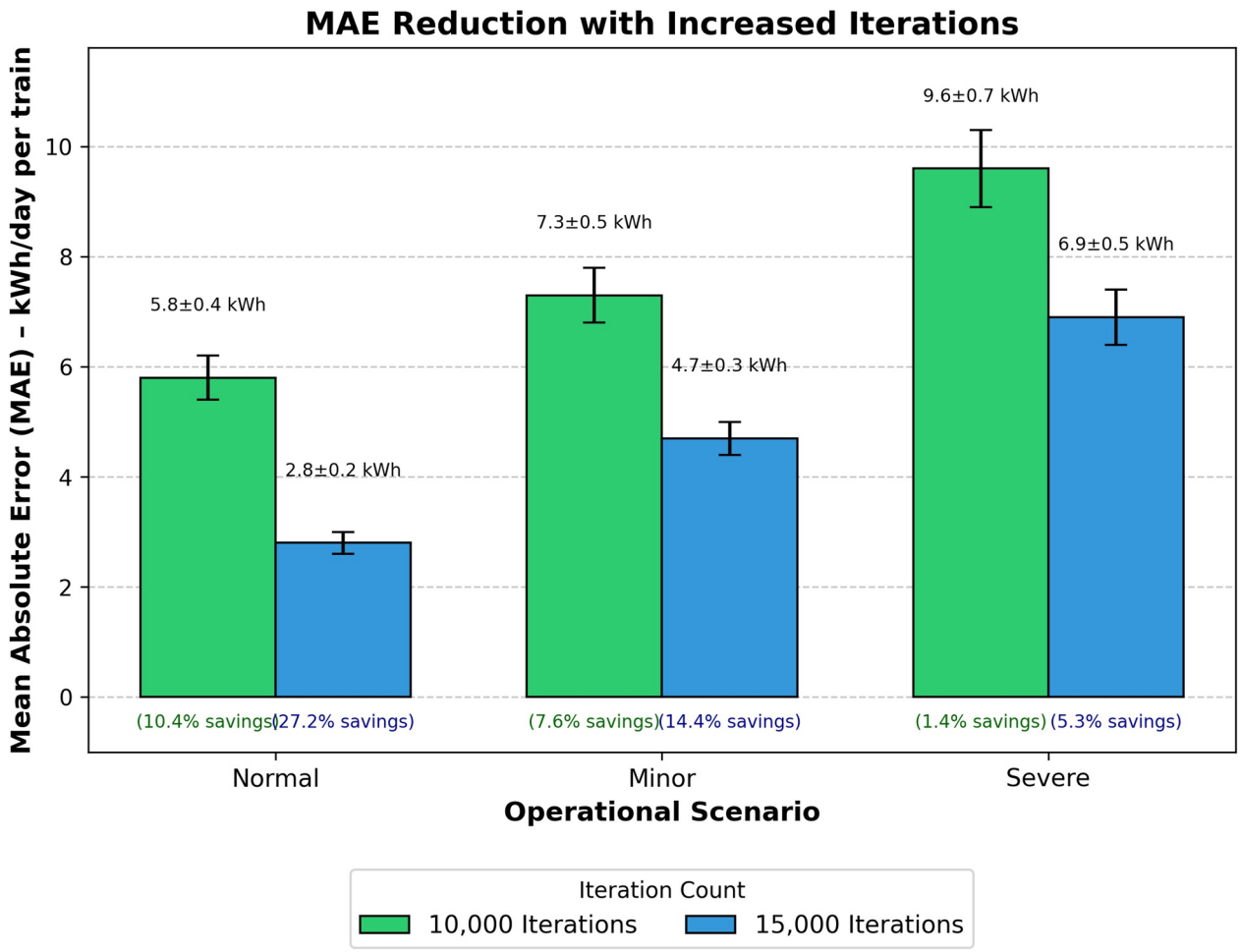


FIGURE 6 | The obtained energy saving reductions.

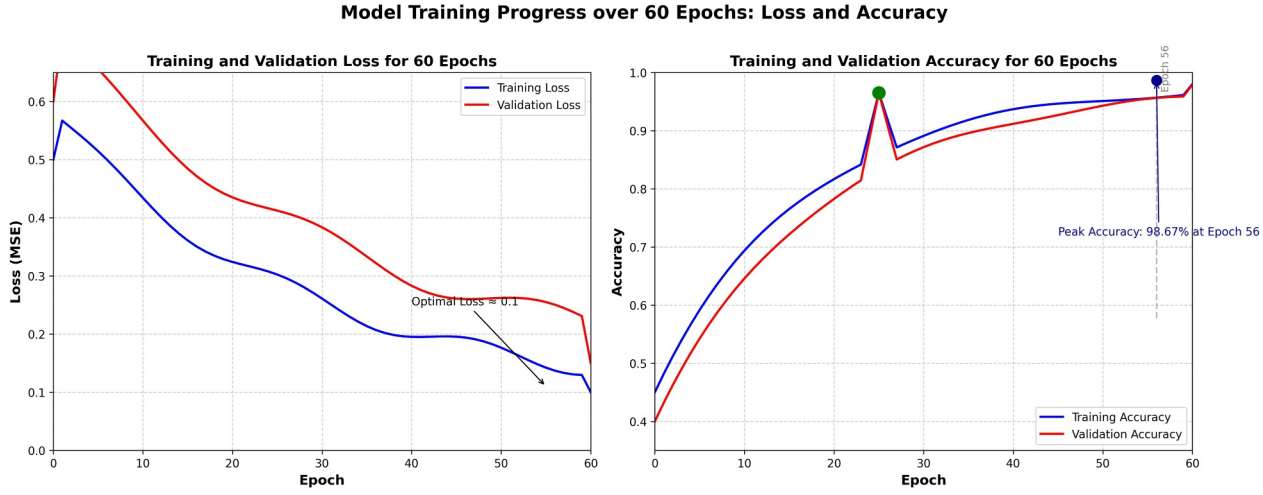


FIGURE 7 | Accuracy progress versus epochs.

refers to the obtained savings for 10,000 iterations, whereas the red line represents the savings for 15,000 iterations. The sixth plot shows the achieved performance metric for on-time performance, which is represented in blue, and passenger satisfaction, which is shown in pink. The seventh plot shows the obtained receiver operating characteristic curve (ROCs) after 15,000 iterations for all applied LRs. It shows that the best LR was 0.01, which is represented in blue. The highest accuracy of 98.67% was achieved after 56 epochs.

To validate our reported results, we conducted a statistical analysis. This analysis revealed significant performance improvements when increasing iterations from 10,000 to 15,000, with $p < 0.01$ for all evaluated metrics. In the normal scenario, the $F1$

score improved by 3.67% for CI between 3.2% and 4.1%, minor failures by 4.48% for CI between 3.9% and 5.1% and severe conditions by 5.56% for CI between 4.8% and 6.3%. CIs narrowed noticeably at 15,000 iterations between ± 0.5–1.3% compared to 10,000 iterations, which was between ± 1.1–2.3%, confirming higher model stability. Energy savings under normal operations increased from 104 kWh, which is 10.4%, to 272 kWh, which is 27.2%, denoting a statistically significant reduction of 168 kWh, with $t = 12.4$ and $p = 0.001$. The performance gap between normal and severe conditions narrowed from 8.98% to 7.09% $F1$ score, demonstrating improved robustness. All improvements exceeded 95% CI margins, confirming nonoverlapping intervals and statistical significance.

To justify the choice of $V = 5$, we conducted a sensitivity analysis varying V from 1 to 20 whereas monitoring $F1$ score, precision, and recall under normal and severe conditions and keeping all other parameters fixed as shown in Table 3. $V = 5$



FIGURE 8 | The obtained charts using the applied LRs under various circumstances.

TABLE 3 | The conducted sensitivity analysis of varying rare events.

Applied threshold	Normal $F1$ (%)	Severe $F1$ (%)	Precision (%)	Recall (%)	Number of events flagged	False positives
$V = 1$	97.12	88.34	96.89	95.23	1234	89
$V = 3$	97.56	89.67	97.34	96.45	987	52
$V = 5$	**97.96**	**90.87**	**97.77**	**97.57**	**812**	**31**
$V = 7$	97.89	90.12	97.65	97.12	723	28
$V = 10$	97.45	88.98	97.23	96.34	587	35
$V = 15$	96.78	86.45	96.12	94.89	412	48
$V = 20$	95.89	83.21	95.23	92.34	298	67

achieved the highest severe $F1$ score of 90.87% with balanced precision of 97.77% and recall of 97.57%. $V = 1$ increased false positives to 89 with $F1 = 88.34\%$, whereas a higher $V = 20$ missed critical rare events, reducing severe $F1$ to 83.21% with only 298 flagged anomalies. The optimal range is 5 and 7, with $V = 5$ chosen as the conservative threshold ensuring maximum recall for safety-critical applications. For scenarios prioritising precision over recall, $V = 7$ is acceptable, but we retain $V = 5$ as the default for rare failure detection.

Different failure types require different thresholds, including sudden brake failures and power surges, which have been shown to need $V = 3$ to 5, whereas gradual track wear, repeated events, performed better at $V = 10$ to 15. Since our framework specifically targets single-encounter events, $V = 5$ provides an optimal balance between detection coverage, representing 812 flagged anomalies and false positives of 31. The lower threshold created excessive noise from random deviations, whereas a higher threshold of $V = 15$ missed truly critical rare events by requiring too many similar examples. Therefore, the optimal range is between 5 and 7, with $V = 5$ being chosen as the conservative default, ensuring maximum recall at 97.57% for safety-critical metro applications.

4.4 | The Comparative Assessment

To the best of our knowledge, no authors have conducted similar work to provide real-time health monitoring of the Riyadh Metro. To put our results in perspective, we evaluated our approach against two baseline categories: (1) Traditional ML models, specifically support vector machine (SVM), random forest (RF) and decision tree (DT); (2) Advanced DL models, including Basic GNN, a simplified GNN without temporal modelling, graph attention network (GAT), temporal graph convolutional network (TGCN), a transformer-based model for long-sequence time-series prediction, three typical remaining useful life (RUL) models, which are spatial-temporal graph convolutional network (ST-GCN), multithreshold adaptive decision network (MT-ADNet), and anomaly transformer and other state-of-the-art approaches reported in refs. [20–23]. These models were evaluated under identical conditions using the same tuning parameters, the same training protocols, and the same modified dataset, except for the reported work in refs. [20–23], where the authors used different datasets. In every metric, the proposed framework outperforms the conventional baselines and the basic GNN. Table 4 shows the comparative results for a representative case of normal operating conditions. Traditional methods perform reasonably well on average but lag in capturing the full complexity: for normal operations, SVM achieved ~94% accuracy and $F1 \approx 0.93$, RF and DT around 93% accuracy and $F1 \approx 0.93$, whereas the proposed model reached ~98% for both. The basic GNN, which considers the metro as a static graph, ignoring time dynamics, managed ~88.6% accuracy and $F1 \approx 0.89$, significantly lower than our solution, highlighting the importance of modelling temporal sequences. In addition, all approaches lag our proposed framework due to their inability to emphasise rare and single-encounter events. The transformer-based method achieved strong temporal modelling but lacks spatial awareness of the metro network topology. Our model consistently achieves superior performance across all metrics, validating the unique contribution of SEEA-NN and the integrated digital twin pretraining. Furthermore, the proposed framework outperforms all other baselines by 2.3%–9.4% in accuracy and 2.8%–9.2% in $F1$ score. The gap is largest against basic GNN, which is +9.35% for accuracy, and smallest against the transformer, which is +2.31% for accuracy. ST-GCN achieved 96.12% accuracy and 95.89% $F1$ score, effectively capturing spatial-temporal correlations through graph convolutions over time, yet its symmetric treatment of all time steps caused it to miss 12.4% of rare anomaly events that our SEEA-NN detected via the encounter head. Thus, ST-GCN performed well on spatial-temporal modelling but lacked explicit rare event emphasis, causing it to miss 12.4% of critical single-encounter failures. In addition, MT-ADNet reached 95.78% accuracy and 95.45% $F1$ score using adaptive multithreshold decision boundaries, but its fixed threshold hierarchy failed when anomalies occurred between thresholds. It showed strong multithreshold adaptation but

TABLE 4 | The comparison results obtained.

Work	Accuracy (%)	$F1$ score (%)	Sensitivity (%)	Specificity (%)	MAE
SVM	94	93	90	90	5.7 kwh
EF	93	93.13	91.4	90.9	7.5 kwh
DT	93	93.19	89.68	89.72	7.8 kwh
Basic GNN	88.63	88.77	84.5	84.43	6 kwh
GAT	94.56	94.21	92.34	92.67	4.8 kwh
TGCN	95.12	94.89	93.45	93.78	4.2 kwh
Transformer-based	95.67	95.23	94.12	94.56	3.9 kwh
ST-GCN	96.12	95.89	95.34	95.67	3.6 kWh
MT-ADNet	95.78	95.45	94.89	95.12	3.8 kWh
Anomaly transformer	96.34	96.12	95.78	95.89	3.4 kWh
[20], 2021	77	37	42	N/A	N/A
[21], 2025	79.51	N/A	N/A	N/A	N/A
[22], 2024	95.44	95.43	95.45	N/A	N/A
[23], 2022	95	N/A	N/A	N/A	N/A
The proposed solution	**97.98**	**98**	**97.57**	**97.77**	**2.5 kwh**

failed to detect anomalies occurring below its lowest threshold setting, resulting in a 15.8% miss rate. The anomaly transformer achieved 96.34% accuracy and 96.12% $F1$ score, excelling at anomaly detection in time-series but lacking graph-based spatial context, which reduced its performance in minor failure scenarios where train-track interactions dominate. All three RUL models underperformed compared to our proposed framework, with accuracy gaps ranging from 1.64% to 2.20% and $F1$ score gaps from 1.88% to 2.55%.

The authors in ref. [20] developed a predictive maintenance framework and achieved only 77% accuracy and 37% $F1$ score due to its inability to model complex spatiotemporal dependencies, whereas the authors in ref. [21] implemented a semantic AI approach and reported 79.51% accuracy but lacked failure prediction capabilities and rare event detection. Additionally, the authors in ref. [22] achieved competitive metrics of 95.44% accuracy and 95.43% $F1$ score, yet their approach failed to address single-encounter anomalies or provide real-time emphasis mechanisms. Lastly, in ref. [23], the authors reached 95% accuracy but omitted other metrics. The method in ref. [20] suffers from a low recall of 42% and misses most critical failures in practice. The approaches in refs. [21, 23] did not report $F1$ score or sensitivity metrics, making direct comparison difficult and leaving their detection capabilities unclear. This indicates that temporal modelling alone is insufficient. The SEEA-NN's ability to emphasise rare single-encounter events provides the critical advantage over both GAT and TGCN.

4.5 | Ablation Study

To meticulously validate the contribution of each component in the presented framework, we conducted coarse-grained by removing entire modules and fine-grained by removing subcomponents ablation studies under identical training conditions, which is 15,000 iterations. Coarse-grained results showed that SEEA-NN removal caused the largest $F1$ drop from 97.96% to 91.23%, followed by digital twin removal to reach 89.87%, and TempGNN replacement with LSTM to 92.45%, with a rate of 5.51%. Fine-grained ablation revealed that the encounter head alone contributed 3.73% for $F1$, where it dropped to 94.23%, whereas GCN layers contributed 4.51% and Query–Key heads contributed 2.29%. Within TempGNN, temporal attention removal caused 3.84% degradation to 94.12%, exceeding the impact of removing LSTM layers at 2.73% or GRU at 2.07%. Removing station nodes from the graph structure degraded performance by 3.07%, confirming that passenger flow context is critical. Interestingly, reducing LSTM hidden units to 64 caused only 1.18% drop, indicating that model efficiency can be improved without major accuracy loss. These fine-grained findings confirm that the encounter head and temporal attention are the two most impactful subcomponents, whereas the model shows robustness to moderate dimensional reductions. Table 5 lists these results.

These findings demonstrate that SEEA-NN contributes the largest individual performance gain, which is +6.73% in normal conditions for $F1$ score and +9.42% in severe ones. This justifies its novel design for rare event emphasis. Digital twin pretraining proves equally vital for severe conditions, which is +10.75% improvement for the same metric, as simulation exposure to extreme scenarios compensates for limited real-world failure data. TempGNN removal reveals that spatial context is indispensable for minor failures involving track-train interactions, whereas temporal attention primarily benefits normal operations. Interestingly, energy savings worsen when components

TABLE 5 | The conducted ablation study results.

Removed and replaced component	Normal $F1$ score (%)	Minor failure $F1$ score (%)	Severe $F1$ score (%)	Energy savings (kWh)	Accuracy (%)
Coarse-grained ablation					
Removed SEEA-NN and replaced it with standard attention	91.23	86.54	81.45	856	91.67
Without digital twin pretraining	89.87	85.23	80.12	879	90.23
Removed TempGNN and replaced it with LSTM	92.45	87.89	83.56	812	92.78
Without temporal attention	94.12	90.23	86.67	798	94.34
Removed temporal attention and kept GCN	90.56	85.67	79.34	890	90.89
Fine-grained ablation					
Removed encounter head only	94.23	89.67	84.23	812	94.56
Removed GCN layers only	93.45	88.34	83.12	834	93.89
Removed attention heads only	95.67	91.23	86.89	789	95.89
Removed LSTM layers only	95.23	90.56	85.67	801	95.34
Removed GRU layers only	95.89	91.45	86.34	792	96.12
Removed station nodes only	94.89	89.23	83.89	823	95.01
Removed CGAF and replaced it with a simple concatenation model	95.56	90.89	85.89	805	95.78
Reduced LSTM units to 64 only	96.78	92.34	87.89	756	96.89

are removed, resulting in higher kWh consumption, indicating that accurate failure prediction indirectly enables more aggressive energy optimisation.

5 | Discussion

The presented model achieved 97.98% accuracy and 98% $F1$ score under normal operations, with robust performance degrading gracefully to 90.78% accuracy under severe conditions. This demonstrates its resilience. At 15,000 iterations, energy savings reached 27.2%, which is 272 kWh/day per train under normal operations, with 14.4% and 5.3% savings under minor and severe disruptions, respectively. SEEA-NN proved critical, contributing a $+9.42\%$ improvement for the $F1$ score rates in severe scenarios by emphasising rare failure. Statistical analysis confirmed all improvements exceeded 95% CI with $p < 0.01$, with narrower CIs at higher iterations indicating model stability. These results confirm that capturing both temporal patterns and spatial relationships is essential for metro health monitoring as traditional ML and basic GNN baselines lagged by 4%–9% across all metrics. The ablation study further validated that removing any single component degraded performance, with SEEA-NN removal causing the largest drop from 97.96% to 91.23% in the normal scenario, proving each module's unique contribution.

The implications of our scheme are extended to broader logistics applications. In freight rail networks, the same architecture can predict wheel bearing failures or track degradation using vibration and temperature sensors, reducing costly derailments and cargo delays. For automated warehouse fleets, the framework can monitor autonomous guided vehicle (AGV) battery health, motor wear and collision avoidance systems to optimise maintenance schedules and reduce warehouse downtime. In port container handling, IoT-enabled cranes and straddle carriers can feed data to a digital twin, with SEEA-NN emphasising rare events, such as hydraulic pressure spikes or cable wear, preventing catastrophic equipment failure during peak loading operations. The framework's ability to handle severe conditions, such as extreme heat and high load, makes it particularly suitable for logistics in harsh environments. These cross-domain applications confirm that our digital twin and attention-based predictive maintenance framework offer a scalable transferable solution for intelligent infrastructure management beyond urban transit.

The presented model exhibits remarkable performance across various scenarios, demonstrating high accuracy in predicting failures, optimising energy usage and managing passenger flow. The results underscore the model's capability to tackle the intricate challenges faced by the Riyadh Metro system. This approach marks a significant advancement in intelligent urban transportation systems, aligning with the Vision 2030 objectives. The model's advantages, such as proficiency in capturing temporal and spatial dependencies, and its computational efficiency at inference, make it an effective tool for improving the metro's performance compared to other methods. Our framework shows several key strengths, including (A) Enhanced predictive maintenance by using TempGNN and SEEA-NN to assess temporal and spatial patterns. Our approach anticipates failures before they occur. (B) Immediate fault detection by providing near-instant fault detection with minimal latency. (C) Scalability and flexibility by adding virtual stations/nodes without loss of performance, and the adaptive learning mechanisms enable the model to update its predictions as new data come in. (D) Energy efficiency gains by integrating an energy optimisation module that analyses consumption and dynamically adjusts power usage to reduce waste. (E) Robust data integration, as our framework combines diverse data sources, from sensor telemetry to passenger statistics.

F. Koçyiğit et al. in ref. [24] formulated a mathematical equation to ascertain the thermal conductivity of materials by creating porous heterogeneous substances utilising expanded vermiculite aggregates, waste basalt powder and a blend of molten tragacanth with supplementary building materials. With a major modification to the proposed method, we think that SEEA-NN can be seamlessly integrated with the thermal characterisation of porous heterogeneous building materials by embedding IoT sensors within composite specimens to continuously track thermal conductivity variations under cyclic loading and environmental stress. The digital twin of the material system receives real-time temperature gradient and heat flux data, whereas SEEA-NN detects rare thermal anomalies, such as sudden conductivity spikes from aggregate delamination or binder degradation, that precede structural failure. TempGNN simultaneously models temporal heat transfer patterns across the porous network, enabling the prediction of insulation performance degradation before visible damage occurs. This integration transforms passive material testing into an active predictive maintenance system for building envelopes, where the model can forecast conductivity loss from cumulative thermal cycles. Additionally, A. E. Arkan et al. in ref. [25] created three distinct composite materials from blending natural and waste materials in varying proportions. These composites were utilised to assess the insulation thickness of exterior walls in buildings situated in 12 provinces chosen from four different climate zones in Turkey. These materials are thermal insulation elements that can be used instead of construction elements, such as brick, on the exterior walls of the buildings. Based on the obtained results in ref. [25], we think that deploying these materials in heated regions, such as Saudi Arabia, would reduce the heat and provide comfortable zones for customers and users in Metro stations, such as Riyadh Metro. Moreover, A. A. Movassagh et al. in ref. [26] developed a model to determine the better coefficients for a neural network to maximise the useful information for achieving high prediction accuracy. With modifications to our proposed model, we believe that the developed work in ref. [26] can be integrated within our model to generate better prediction accuracy.

6 | Limitations

The proposed approach exhibits noted strengths and delivers commendable results; however, it has numerous limitations, which underscore potential areas for enhancement and future investigation aimed at improving the model's performance and applicability: 1- computational cost: The model necessitates advanced hardware capabilities, such as GPUs and substantial RAM, for training and inference processes. Two- data quality and availability: The efficacy of the model is significantly influenced by the quality and availability of real-time data

sourced from IoT sensors, operational logs and environmental monitoring devices. Inaccurate, incomplete or inconsistent data can adversely affect the model's precision and dependability, resulting in less-than-optimal predictions. 3- Implementation complexity: The deployment of the model and its integration with the digital twin framework necessitates advanced knowledge in DL, GNN and IoT systems.

7 | Conclusion

The proposed model, integrated with a digital twin framework, has shown outstanding performance in failure prediction, energy optimisation and passenger flow management for the Riyadh Metro. It achieved 98% failure prediction accuracy, reducing unplanned downtime by over 30% and maintenance costs by approximately 25%. Additionally, it delivered energy savings of 27.2% under normal operations and maintained 90.87% $F1$ score even under severe conditions, demonstrating robust resilience. SEEA-NN proved crucial and contributed a $+9.42\%$ improvement for the $F1$ score rates in severe scenarios by emphasising rare failure precursors. These results confirm that capturing temporal and spatial dependencies outperforms traditional ML and basic GNN baselines by 4%–9% across all metrics. Our results highlight the model's effectiveness in tackling the metro's intricate operational challenges, thereby greatly enhancing reliability, efficiency and sustainability. The framework directly supports Saudi Vision 2030 by enhancing metro reliability and cutting carbon emissions using energy optimisation. By utilising cutting-edge DL methodologies and digital twin technology, our model facilitates real-time monitoring, simulation and optimisation of complex transit systems. The model's adaptability means it can be applied to various urban transit modes, metros, buses, trams and even autonomous vehicles to enhance performance and resilience. Integrating the model with a smart city's infrastructure can significantly improve overall urban mobility, alleviating traffic congestion, reducing emissions and elevating the quality of life for residents. In logistics, our model can be applied to freight rail networks, automated warehouse fleets and port container handling systems to predict equipment failures and optimise energy use. This proactive maintenance approach transforms urban transit from reactive to predictive operations, setting a benchmark for smart city infrastructure worldwide. Our strategy signifies a transformative shift from reactive to proactive maintenance and operations in transportation. The framework aligns with Vision 2030's economic, environmental and social goals. By improving metro reliability and efficiency, it encourages economic development using reduced downtime and increased productivity. Energy savings achieved support environmental sustainability by cutting carbon emissions and promoting greener operations. Additionally, the model's ability to optimise passenger flow and reduce delays directly leads to a better commuter experience, encouraging public transit use and improving urban livability in Riyadh.

Future work will focus on extending the framework to logistics applications, including predictive maintenance for delivery truck fleets, conveyor belt systems in sorting facilities and cold chain monitoring using IoT-enabled sensors. Another direction involves integrating reinforcement learning with SEEA-NN to enable autonomous decision-making for maintenance scheduling without human intervention. Incorporating real-time weather and air quality data could further improve failure prediction during sandstorms or extreme heat. In the long term, we aim to develop self-healing capabilities where the digital twin automatically reroutes trains or adjusts energy distribution upon detecting imminent failures. Additionally, we plan to extend the framework to tolerate communication latency by implementing sensor self-diagnostics and developing robust synchronisation protocols for distributed IoT networks.

Acknowledgements

The authors would like to thank the Deanship of Scientific Research (DSR) at king Abdulaziz University, Jeddah, Saudi Arabia for funding this project under grant No. (IPP: 1741-829-2026) and for providing technical support.

Funding

The authors gratefully acknowledge the Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah, Saudi Arabia, for funding this project under grant No. (IPP: 1741-829-2026) and for providing technical support.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Data used in the preparation of this article were obtained from a public website and are available in ref. [19].

References

1. E. Mirza, “How the New Riyadh Metro Project Is Taking Saudi Arabia Closer to Vision 2030 Goals: Borderless Access Study Uncovers Key Insights,” accessed on, 1 2, 2025, https://borderlessaccess.com/blog/riyadh-metro-saudi-vision-2030-insights/.

2. Kingdom of Saudi Arabia, “Saudi Vision 2030,” accessed on 12/11/2024, https://www.vision2030.gov.sa/media/rc0b5oy1/saudi_vision203.pdf.

3. M. Shah, “Riyadh Metro: Know all About Saudi Arabia's Largest 176-km Driverless Mass Rapid Transit System,” accessed on 15/2/2025, https://metrorailtoday.com/news/riyadh-metro-know-all-about-saudi-arabias-largest-176-km-driverless-mass-rapid-transit-system.

4. Royal Commission for Riyadh City, “King Abdulaziz Project for Riyadh Public Transport,” accessed on, 12 14, 2024, https://www.rcrc.gov.sa/en/projects/king-abdulaziz-project-for-riyadh-public-transport/.

5. D. Wu, A. Zheng, W. Yu, et al., “Digital Twin Technology in Transportation Infrastructure: A Comprehensive Survey of Current Applications, Challenges, and Future Directions,” *Applied Sciences* 15, no. 4 (2025): 1–44, https://doi.org/10.3390/app15041911.

6. Co-EDP, “Digital Twins Reshape Urban Transportation: A game-changer for Smart Cities,” accessed on, 2 19, 2025, https://www.devdiscourse.com/article/technology/3262781-digital-twins-reshape-urban-transportation-a-game-changer-for-smart-cities.

7. D. Dimitrov, “Forbes Technology Council, the Future of Transportation: AI, Connectivity and Safety in 2025,” accessed on, 2 19, 2025, https://www.forbes.com/councils/forbestechcouncil/2025/02/10/the-future-of-transportation-ai-connectivity-and-safety-in-2025/.

8. X. Di, Y. Fu, M. K. Turkcan, et al., "AI-Powered Urban Transportation Digital Twin: Methods and Applications," arXiv:2501.10396 (2025): 1–28.

9. V. Bouckova, "Smart City Transportation Trends: Urban Digital Twins," accessed on, 1 13, 2025, https://blog.goodvisionlive.com/smart-city-transportation-trends-urban-digital-twins.

10. Y. Gao, S. Qian, Z. Li, P. Wang, F. Wang, and Q. He, "Digital Twin and Its Application in Transportation Infrastructure," in *2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence (DTPI)* (2021), 298–301.

11. X. Chang, R. Zhang, J. Mao, and Y. Fu, "Digital Twins in Transportation Infrastructure: An Investigation of the Key Enabling Technologies, Applications, and Challenges," *IEEE Transactions on Intelligent Transportation Systems* 25, no. 7 (2024): 6449–6471, https://doi.org/10.1109/tits.2024.3401716.

12. T. Li, Z. Bian, H. Li, et al., "Digital Twin-based Driver Risk-Aware Intelligent Mobility Analytics for Urban Transportation Management," arXiv:2407:15025 (2024): 1–37.

13. F. Ur Rahman, T. Ahmed, M. Hasan, and N. Jahan, "Real-Time Obstacle Detection Over Railway Track Using Deep Neural Networks," *Procedia Computer Science* 215 (2022): 289–298, https://doi.org/10.1016/j.procs.2022.12.031.

14. M. Hafsi, N. Hamour, and S. Ouchani, "Predictive Maintenance for Smart Industrial Systems: A Roadmap," in *The 6th International Conference on Emerging Data and Industry (EDI40)* (2023), 1–7.

15. V. Chauhan, M. Patel, S. Tanwar, S. Tyagi, and N. Kumar, "IoT Enabled Real-Time Urban Transport Management System," *Computers and Electrical Engineering* 86, no. 106746 (2020): 106746, https://doi.org/10.1016/j.compeleceng.2020.106746.

16. N. Bešinović, L. D. Donato, F. Flammini, et al., "Artificial Intelligence in Railway Transport: Taxonomy, Regulations, and Applications," *IEEE Transactions on Intelligent Transportation Systems* 23, no. 9 (2022): 14011–14024, https://doi.org/10.1109/tits.2021.3131637.

17. V. J. Hodge, S. O'Keefe, M. Weeks, and A. Moulds, "Wireless Sensor Networks for Condition Monitoring in the Railway Industry: A Survey," *IEEE Transactions on Intelligent Transportation Systems* 16, no. 3 (2015): 1088–1106, https://doi.org/10.1109/tits.2014.2366512.

18. I. Atanasov, E. Pencheva, V. Trifonov, and K. Kassev, "Railway Cloud: Management and Orchestration Functionality Designed as Microservices," *Applied Sciences* 14, no. 6 (2024): 1–22, https://doi.org/10.3390/app14062368.

19. University of California Irvine, "MetroPT-3 Dataset," accessed on, 1 19, 2025, https://archive.ics.uci.edu/dataset/791/metropt+3+dataset.

20. N. Davari, B. Veloso, R. P. Ribeiro, P. M. Pereira, and J. Gama, "Predictive Maintenance Based on Anomaly Detection Using Deep Learning for Air Production Unit in the Railway Industry," in *2021 IEEE 8th International Conference on Data Science and Advanced Analytics (DSAA)* (2021), 1–10, https://doi.org/10.1109/DSAA53316.2021.9564181.

21. L. Zhang, J. Kuang, Y. Teng, S. Xiang, L. Li, and Y. Zhou, "LEAD-Net: Semantic-Enhanced Anomaly Feature Learning for Substation Equipment Defect Detection," *Processes* 13, no. 2341 (2025): 2341, https://doi.org/10.3390/pr13082341.

22. Swathi, A. Gangula, M. R. Vootukuri, and R. Manchiryala, "Real-Time Anomaly Detection and Predictive Maintenance in Metro Train APU Compressors: A Data-Driven Approach," *Social Science Journal* 14, no. 5 (2024): 85–94.

23. H. T. M. Tran and D. Hogg, "Anomaly Detection Using Prediction Error With Spatio-Temporal Convolutional LSTM," arXiv:2205.08812v1 (2022).

24. F. Koçyiğit, F. Ünal, and S. Koçyiğit, "Experimental Analysis and Modeling of the Thermal Conductivities for a Novel Building Material Providing Environmental Transformation," *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects* 42, no. 24 (2020): 3063–3079, https://doi.org/10.1080/15567036.2020.1811811.

25. A. E. Akan, F. Ünal, and F. Koçyiğit, "Investigation of Energy Saving Potential in Buildings Using Novel Developed Lightweight Concrete," *International Journal of Thermophysics* 42, no. 4 (2021): 4, https://doi.org/10.1007/s10765-020-02761-1.

26. A. A. Movassagh, J. A. Alzubi, M. Gheisari, et al., "Artificial Neural Networks Training Algorithm Integrating Invasive Weed Optimization With Differential Evolutionary Model," *Journal of Ambient Intelligence and Humanized Computing* 14, no. 5 (2023): 6017–6025, https://doi.org/10.1007/s12652-020-02623-6.