Substation Equipment-Terrain Map Structure Modeling and Chassis Dynamic Control Based on Graphormer

Jian Duan[1], Yinghui Lu[2], Jianxun Zhao[2], Yongchao Zhu[2]
[1]State Grid Hebei Electric Power Co., Ltd, Shijiazhuang 050000, Hebei, China
[2]State Grid Hebei Electric Power Co., LTD, Hengshui Power Supply Branch, Hengshui 053000, Hebei, China

Corresponding author: Yinghui Lu (e-mail: luyinghui2024@163.com)

ABSTRACT To address the limitations of insufficient dynamic correlation modeling between complex terrain and substation equipment layouts and the low efficiency of real-time multimodal sensor fusion, this paper proposes a collaborative framework for graph structure modeling and chassis dynamic control based on Graphormer. The proposed approach enables dynamic representation learning of equipment–terrain topological relationships and multimodal data-driven stability optimization, providing enhanced environmental perception and autonomous navigation capabilities for intelligent inspection systems operating in complex electromagnetic power infrastructures. An equipment–terrain graph is constructed by incorporating centrality encoding, spatial encoding, and edge feature encoding to capture global topological dependencies. A cross-modal attention mechanism is integrated with a reinforcement learning strategy to achieve multimodal sensor fusion and adaptive chassis stability control. Experimental results demonstrate that the proposed method achieves a positioning error standard deviation of only 1.48 cm and obtains the optimal node-to-ground-truth matching performance. The average path success rate across ten simulation scenarios reaches 91.6%, outperforming mainstream approaches, while the fluctuation energy of pitch and roll angles within the 4–8 Hz frequency band is reduced to 14.0% and 15.5%, respectively. The proposed framework provides an effective graph neural network solution for autonomous substation inspection robots and offers technical support for reliable perception, navigation, and intelligent operation in modern electromagnetic energy systems.

INDEX TERMS graphormer, substation inspection robot, equipment-terrain graph modeling, chassis dynamic control, intelligent electromagnetic systems

I. INTRODUCTION

With the acceleration of smart grid construction, substation inspection robots can replace manual high-risk operations, improve operation and maintenance efficiency and power system stability in complex scenarios, and become a key technical support for ensuring power grid security, which is crucial for industrial zones containing high-load sectors like manufacturing, where power stability directly impacts continuous operation and prevents damage to sensitive machinery [1,2]. The demand for autonomous navigation in complex terrain and dense equipment environments is becoming increasingly urgent. However, existing studies generally have problems such as insufficient dynamic correlation modeling between equipment layout and terrain features, making it difficult to quantify the topological relationship between equipment and the impact of terrain mutations on path planning, as well as low efficiency in real-time fusion of multimodal sensor data (lidar, IMU, vision), which causes the chassis control strategy to lag behind environmental changes [3-5]. These problems limit the adaptability of inspection robots in industrial scenarios, and a collaborative framework that takes into account both global topological representation and real-time response capabilities is urgently needed.

This paper focuses on the collaborative optimization of Equipment-Terrain Graph modeling and chassis dynamic control. The equipment-terrain map needs to integrate the physical connection relationship of the equipment (such as power topology) and terrain attributes (slope, obstacle distribution) to form a dynamically updated topological representation [6]. Chassis control needs to balance path planning efficiency and posture stability in complex terrain [7,8]. Equipment-terrain map and chassis control jointly determine the robot's ability to operate in the narrow space and unstructured terrain of the substation, and are the core technical bottlenecks for realizing autonomous inspection [9]. In view of the defects of existing inspection robots such as low battery capacity, weak obstacle crossing ability, slow movement speed, and large inspection blind spots, Ren Y et al. [10] proposed a pan-tilt control and visual sensing framework to establish an automatic camera system. It can partially perform photography and video tasks autonomously in actual application scenarios, improving the efficiency and quality of existing intelligent inspection systems, but does not consider the impact of dynamic terrain changes.

Existing research on the device-terrain relationship modeling problem mostly uses traditional local neighborhood aggregation graph neural networks (GNNs), such as GCN (Graph Convolutional Network) and GAT (Graph Attention Network). However, they cannot capture the functional dependencies of long-distance devices and have difficulty adapting to dynamic terrain changes [11-13]. Zhao Z et al. [14] proposed a GCNbased equipment topology modeling method to accurately distinguish risk areas, but it is only applicable to static scenes. Liu H et al. [15] used deep neural networks to build a specific infrared image processing framework that can achieve automatic fault diagnosis, but it is still limited to local neighborhood information aggregation. In terms of multimodal fusion, existing research mostly relies on the general multi-coordinated data fusion module (MCDFM) framework of real-time heterogeneous data sources [16], but its computational overhead is large. Yang Q et al. [17] solved the problem of time series asynchrony by fusing two-dimensional lidar data with IMU (Inertial Measurement Unit) data to achieve multi-sensor information fusion.

This paper aims to construct a collaborative framework based on Graphormer to solve the problem of equipment-terrain dynamic association modeling and multimodal real-time fusion. This paper implements global topological learning of device-terrain maps by designing spatial encoding and edge feature encoding modules. Based on the cross-modal attention mechanism, lidar, IMU and visual data are mapped to a unified graph space. This paper combines the PPO (Proximal Policy Optimization) reinforcement learning framework with the MPC (Model Predictive Control) path planning algorithm to form a closed-loop control strategy, and integrates device-terrain edge features through Graphormer's Edge Gate mechanism to improve topological representation capabilities. This paper designs a multimodal feature alignment module to reduce the computational overhead caused by the heterogeneity of sensor data, and finally combines the knowledge distillation technology to compress the model to ensure real-time performance, which is vital for guaranteeing the immediate response capabilities needed by high-stakes industrial applications, such as the monitoring and control of critical power equipment in a factory setting. The core contribution of this paper is to unify the spatial and functional relationship between substation equipment and terrain through graph neural network, and to drive multi-sensor fusion and chassis real-time control, forming a closed-loop system of perception-modeling-control.

II. MULTIMODAL COLLABORATIVE MODELING AND CHASSIS CONTROL OPTIMIZATION FRAMEWORK BASED ON GRAPHORMER

SUBSTATION EQUIPMENT-TOPOGRAPHIC MAP STRUCTURE MODELING

Based on the equipment layout and terrain features in the substation environment (the image comes from the Electric Substation: https://www.kaggle.com/datasets/aakashmodi/electric-substation), this paper constructs a graph structure with multi-attribute information to achieve topological representation of complex scenes. The substation equipment-topography map structure diagram is constructed through graph structure construction, multi-scale feature encoding and dynamic topology learning.

The three-dimensional coordinates of the device are obtained from the LiDAR point cloud and SLAM (Simultaneous Localization and Mapping) positioning system, and the terrain elevation data is extracted in combination with the digital elevation model [18,19] (DEM); the physical connection relationship between the devices is obtained based on the CAD (Computer Aided Design) drawings. The geometric center of the equipment and the terrain feature points are used as nodes of the graph, where the equipment node contains attributes such as type, height, and volume, and the terrain node contains information such as slope and roughness. The establishment of edges is based on the principles of spatial adjacency and functional association. For nodes whose Euclidean distance is less than the set threshold, it is considered that there is a spatial adjacency relationship. For equipment nodes with power connections or logical function dependencies, they are defined as functional association edges, and the edge weight is set to the inverse of the distance or the connection strength. Finally, the initialization of the graph structure is completed, and Open3D is used for point cloud processing and DEM generation. Figure 1 shows the constructed equipment-terrain structure diagram.

This paper introduces centrality encoding, spatial encoding and edge feature encoding methods in the Graphormer framework [20]. Standard Graph Neural Networks (GNNs) are constrained by local neighborhood aggregation, making it difficult to capture global functional dependencies across regional devices. Conventional time series Transformers, however, lack explicit modeling capabilities for graph topological structures such as edge types and node centrality. Graphormer, by encoding centrality, spatial features, and edge characteristics, enables end-to-end learning of non-local topological semantics, making it particularly suitable for representing device-terrain coupling relationships. All centrality metrics are computed based on the initial static graph structure (comprising device geometric centers, terrain feature points, and their spatial/functional connections) to initialize node embedding, independent of the subsequent Graphormer dynamic update process. Define graph $G = (V, E)$, where V is the node set, $|V| = N$, and $E \subseteq V \times V$ represents the edge set. For each node $i \in V$, calculate its

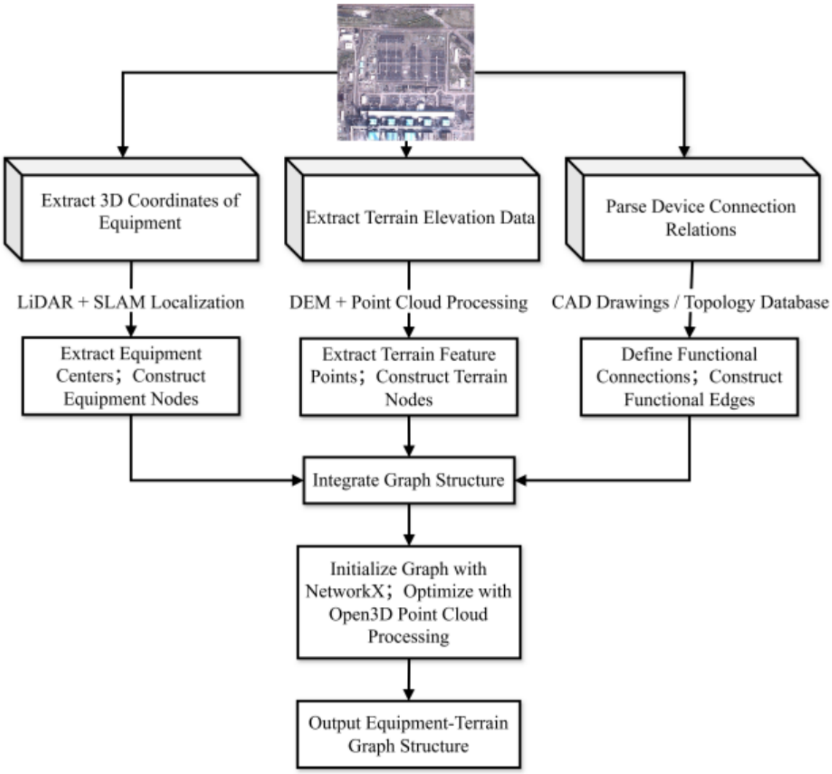

FIGURE 1. Equipment-Topography Map Structure

degree centrality $C_d(i)$ and PageRank centrality $R(i)$:

$$C_d(i) = \frac{|\{j \in V : (i,j) \in E\}|}{N-1} \tag{1}$$

$$R(i) = \frac{1-d}{N} + d \sum_{j \in N_{\text{in}}(i)} \frac{PR(j)}{|N_{\text{out}}(j)|} \tag{2}$$

where $d \in [0,1]$ is the damping factor (usually set to 0.85), $N_{\text{in}}(i)$ is the set of neighbor nodes pointing to node i, and $N_{\text{out}}(j)$ is the set of neighbors starting from node j.

The final centrality encoding vector is expressed as:

$$e_c^{(i)} = W_c \cdot [\text{ReLU}(W_d C_d(i)), \text{ReLU}(W_p PR(i))] \tag{3}$$

where $W_c \in \mathbb{R}^{d_c \times 2}$ and $W_d, W_p \in \mathbb{R}$ are learnable parameters, ReLU is the activation function, and d_c is the centrality embedding dimension.

Spatial encoding is based on the sine/cosine position encoding method in Transformer, which maps the spatial coordinates of the node to the embedding space. For the coordinate axis $k \in \{x,y,z\}$, the position encoding is defined as:

$$PE(x_{i,k}, 2m) = \sin\left(\frac{x_{i,k}}{10000^{2m/d_s}}\right) \tag{4}$$

$$PE(x_{i,k}, 2m+1) = \cos\left(\frac{x_{i,k}}{10000^{2m/d_s}}\right) \tag{5}$$

where $m = 0, 1, \cdots, d_s/2 - 1$, and d_s is the spatial encoding dimension.

The final spatial encoding vector is expressed as:

$$e_s(i) = \text{LayerNorm}\left(W_s \cdot \text{Concat}_{k=x,y,z} PE(x_{i,k})\right) \tag{6}$$

where $W_s \in \mathbb{R}^{d_s \times 3d'_s}$, Concat represents the concatenation operation, and LayerNorm is the layer normalization function.

Edge feature encoding uses a multi-layer perceptron (MLP) to perform nonlinear transformations on edge weights, types, etc., to generate a learnable edge embedding vector. Each edge $e_{ij} \in E$ defines its attributes as $f_{ij} = (w_{ij}, t_{ij})$, where w_{ij} is the weight (such as the inverse of the distance) and t_{ij} is the edge type (such as power connection, adjacency, etc.): where $W_1 \in \mathbb{R}^{h_e \times d_f}$, $W_2 \in \mathbb{R}^{d_e \times h_e}$, d_f is the input edge feature dimension, d_e is the output edge embedding dimension, and h_e is the hidden layer dimension.

The three encoding modules are integrated into the Graphormer backbone network to jointly complete the global feature extraction of the graph structure. After completing the initial graph construction and feature encoding, the multi-head attention mechanism of Graphormer is further used to dynamically update the graph structure and learn the topology. Through the self-attention mechanism, each node can aggregate information from other nodes in the entire graph. The Edge Gate mechanism is introduced to fuse edge features during the attention calculation process, so that the graph structure can dynamically respond to terrain changes and device status adjustments. In each round of attention calculation, the node representation can be weighted and updated according to the features of its neighboring nodes and their connecting edges, and finally the global embedding vector H of the graph can be output as the input of the subsequent chassis control module. Figure 2 shows the structure diagram of the dynamic topology learning mechanism.

MULTIMODAL SENSOR DATA FUSION

This paper aims to efficiently fuse heterogeneous data collected by lidar, IMU and visual sensors to improve terrain perception and support chassis dynamic control.

The lidar point cloud data is filtered and denoised through the PCL (Point Cloud Library), and the terrain elevation map is extracted to calculate the local curvature as the terrain roughness feature. The IMU raw data is processed by Kalman filtering to output the robot's current attitude angle, including the pitch angle θ and the roll angle ϕ, which are used for subsequent attitude compensation and stability evaluation. The visual module uses the ResNet-18 network to extract features from the image, obtains a 128-dimensional texture feature vector, and further reduces the dimension to 64 dimensions through principal component analysis (PCA) to reduce computational complexity and retain the main visual information. All sensor data are sent to the next stage for processing after timestamp alignment. An example of multimodal data input is shown in Figure 3.

To solve the problem of modal differences between different sensor data, a feature alignment module based on the cross-attention mechanism [21] is designed to map multi-source sensor data into a unified graph space. Specifically, the feature vectors from each sensor modality (LiDAR, IMU, and visual) serve as queries, while the device-terrain map node embeddings generated by Graphormer act as shared keys and values. Through the cross-attention mechanism, these multimodal observations are mapped and aligned into a unified semantic space of the graph structure. The gated recurrent unit (GRU) is introduced to update the state of graph nodes and fuse historical information in the time dimension to improve the system's adaptability to dynamic environments. The fused graph structure embedding vector H is input into the MLP classifier to predict the terrain stability risk probability, that is, the possibility of the robot becoming unstable when driving in a certain area. Unlike static approaches like feature stitching or post-processing fusion, our cross-modal attention mechanism employs graph embedding as semantic anchors to dynamically weight each modality's contribution across different graph nodes. This approach significantly improves semantic alignment accuracy and robustness for heterogeneous data in complex terrains. Suppose the fused graph structure embedding vector is $H \in \mathbb{R}^{N\times d_h}$, where N is the number of nodes and d_h is the embedding dimension of each node. For each node $i \in \{1,\ldots,N\}$, extract its embedding representation $h_i \in \mathbb{R}^{d_h}$ and input it into the multi-layer perceptron (MLP) for classification:

$$\begin{aligned} z_i^{(1)} &= W_1 h_i + b_1, \quad a_i^{(1)} = \mathrm{ReLU}(z_i^{(1)}), \\ z_i^{(2)} &= W_2 a_i^{(1)} + b_2, \quad a_i^{(2)} = \mathrm{ReLU}(z_i^{(2)}), \\ z_i^{(3)} &= W_3 a_i^{(2)}, \quad p_{\mathrm{risk},i} = \sigma(z_i^{(3)}). \end{aligned} \tag{7}$$

where $W_1 \in \mathbb{R}^{d_1\times d_h}$, $W_2 \in \mathbb{R}^{d_2\times d_1}$, and $W_3 \in \mathbb{R}^{1\times d_2}$ are learnable weight matrices; $b_1 \in \mathbb{R}^{d_1}$ and $b_2 \in \mathbb{R}^{d_2}$ are bias terms; $\mathrm{ReLU}(x) = \max(0,x)$ is the activation function; $\sigma(x) = 1/(1+e^{-x})$ is the Sigmoid activation function; $p_{\mathrm{risk},i} \in [0,1]$ represents the terrain stability risk probability of the area corresponding to node i.

During the training process, the binary cross entropy loss function [22] is used to optimize the model parameters, and the actual path stability test results of the labels are generated:

$$\begin{aligned} L_{BCE} = -\frac{1}{N}\sum_{i=1}^{N} \big[& y_i \log(p_{\mathrm{risk},i}) \\ & + (1-y_i)\log(1-p_{\mathrm{risk},i}) \big] \end{aligned} \tag{8}$$

where $y_i = 1$ indicates that there is a risk of instability in the area, and $y_i = 0$ indicates a safe area.

CROSS-SCALE FEATURE FUSION MODULE

This section designs a chassis dynamic control strategy based on the equipment-terrain map embedding vector and multimodal sensor data to achieve stable motion control in a complex substation environment. The chassis control strategy $\pi(a|s)$ is trained using the proximal policy optimization (PPO) reinforcement learning framework, where the state s is composed of the graph embedding vector H output by Graphormer and the IMU attitude angle (pitch θ, roll ϕ). Action a is defined as the motor torque increment $\Delta\tau$, which is used to adjust the power distribution of each wheel of the chassis. The control strategy parameter optimization table is shown in Table 1. The reward function design takes into account both stability and energy efficiency:

$$R = w_1 \cdot \cos(\theta) + w_2 \cdot \cos(\phi) - w_3 \cdot \|\Delta\tau\| \tag{9}$$

where w_1, w_2, and w_3 are weight coefficients, which emphasize posture stability and energy minimization respectively.

During the training process, various terrain scenarios are simulated. The final strategy model is deployed in the control module as the basic posture control mechanism. Based on the global graph structure, the improved A* algorithm [23] is used to adapt to the dynamic terrain risk perception requirements. The cost function is defined as:

$$C = \alpha \cdot d_{\mathrm{path}} + \beta \cdot \int p_{\mathrm{risk}}(x)\, dx \tag{10}$$

where d_{path} represents the path length, the integral term $\int p_{\mathrm{risk}}(x)\, dx$ reflects the cumulative terrain risk probability on the path, and α and β are trade-off coefficients.

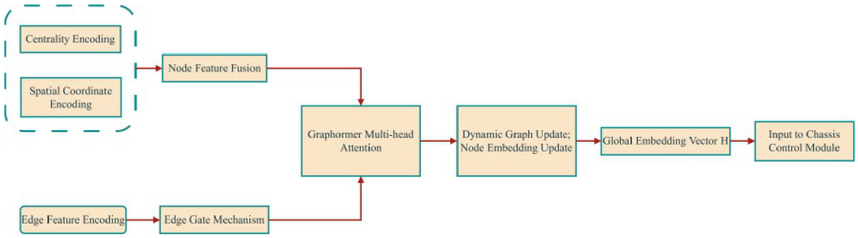

FIGURE 2. Dynamic topology learning mechanism structure

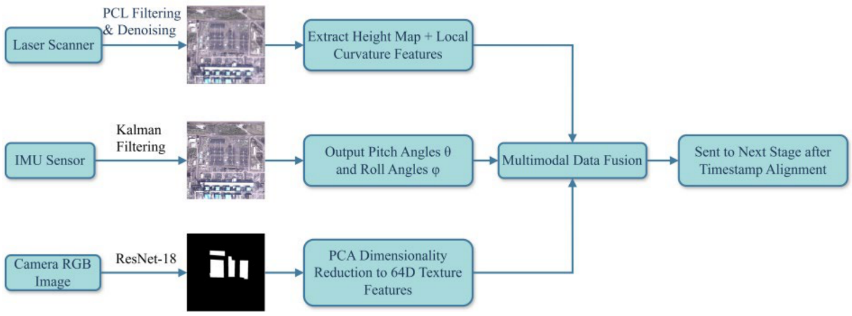

FIGURE 3. Example of multimodal data input

TABLE 1. Control strategy parameter optimization table

Parameter Category	Parameter Name	Initial Value	Optimized Value	Optimization Method
PPO Reinforcement Learning	Learning Rate	1e-4	3e-4	Grid Search + Early Stopping
PPO Reinforcement Learning	Batch Size	32	64	Adaptive Adjustment
PPO Reinforcement Learning	Discount Factor (γ)	0.9	0.99	Fixed Empirical Value
PPO Reinforcement Learning	GAE (Generalized Advantage Estimation) Coefficient (λ)	0.9	0.95	Bayesian Optimization
PPO Reinforcement Learning	Clip Range (ϵ)	0.1	0.2	Dynamic Decay Strategy
Reward Function Weights	Stability Weights (w_1, w_2)	0.3, 0.3	0.5, 0.5	Multi-Objective Pareto Optimization
Reward Function Weights	Energy Penalty Weight (w_3)	0.05	0.1	Gradient Ascent
MPC Planning Parameters	Rolling Window Length	10 steps	15 steps	Simulation Validation
MPC Planning Parameters	Cost Function Weights (α; β)	0.8, 0.5	1, 0.8	Genetic Algorithm Optimization
PID Control Parameters	Proportional Gain (K_p)	1.5	2.1	Ziegler-Nichols Tuning
PID Control Parameters	Integral Gain (K_i)	0.2	0.5	Robustness Optimization
PID Control Parameters	Derivative Gain (K_d)	0.3	0.8	Adaptive Filtering Adjustment

The local planning adopts the MPC (Model Predictive Control) framework, the rolling window length is set to 15 steps, and the optimizer is called every 50ms to solve the minimum cost path [24]. In order to ensure the real-time performance and dynamic adaptability of the system, a feedback closed-loop control process with a cycle of 50ms is designed. The following operations are completed in each cycle:

(1) Incremental update of graph structure

Input the latest lidar and IMU data into Graphormer, and only update the affected node embedding through the Edge Gate mechanism;

(2) Risk distribution reassessment:

Based on the updated graph embedding H, the MLP classifier re-predicts the terrain risk heat map and passes it to the path planning module;

(3) PID instruction generation:

Based on the error between the planned path and the current posture, the PID control amount is calculated:

$$\tau = K_p e + K_i \int e\,dt + K_d \frac{de}{dt} \tag{11}$$

where $K_p = 2.1$, $K_i = 0.5$, and $K_d = 0.8$ are the parameters after adjustment, and the error e includes position deviation and attitude angle deviation. The final control command is sent to the chassis driver through digital-to-analog conversion to achieve attitude adjustment.

III. EXPERIMENTAL VERIFICATION AND PERFORMANCE EVALUATION

NODE DISPLACEMENT ERROR EVALUATION

The node positioning error assessment is based on the three-dimensional coordinate residuals of the equipment geometric center and the terrain feature points. The experiment uses a Leica TS16 total station to perform high-precision coordinate calibration on the geometric centers of key substation equipment (main transformer, disconnector, and lightning arrester) as the true value of the equipment node. The true value of the terrain feature point is extracted from the digital elevation model (DEM) generated by the fusion of drone oblique photography and ground laser scanning, and is aligned with the SLAM coordinate system through a six-parameter affine transformation [25]. After the sensor raw data is Kalman filtered to eliminate IMU motion distortion, the improved Iterative Closest Point (ICP) algorithm is used to complete multi-source point cloud registration [26].

In the process of constructing the equipment-topographic map structure, the equipment node is initialized by the center coordinates of the power equipment marked in the CAD drawing, and the terrain feature points are sampled and generated by surface fitting the DEM using the Moving Least Squares (MLS) method. The coordinates can be converted into the local coordinate system within the station. The optimal matching between graph nodes and true value points is achieved through the Hungarian algorithm. The cost matrix is defined as the square of the Euclidean distance. The positioning error can be expressed as follows according to the node configuration strategy:

$$C_{ij} = \left\| p_i^{\text{graph}} - p_j^{\text{truth}} \right\|_2^2, \quad \text{MSE} = \frac{1}{N} \sum_{i=1}^{N} w_i \left\| p_i^{\text{graph}} - p_i^{\text{truth}} \right\|_2^2 \tag{12}$$

where p_i^{graph} is the coordinate of the graph node, and p_j^{truth} is the coordinate of the corresponding true value point. The matching success threshold is set to 3 times the standard deviation of the laser radar ranging noise; N is the total number of valid matching nodes. The error is decomposed into equipment nodes and terrain nodes, and their axial errors ϵ_x, ϵ_y, ϵ_z and radial errors $\epsilon_r = \sqrt{\epsilon_x^2 + \epsilon_y^2 + \epsilon_z^2}$ are calculated respectively. The error spatial autocorrelation is analyzed using the semivariogram function:

$$\gamma(h) = \frac{1}{2M} \sum_{i=1}^{M} [\epsilon(s_i) - \epsilon(s_i + h)]^2 \tag{13}$$

where h is the lag distance and M is the number of node pairs with an interval of h. The spatial adjacency threshold of the optimized graph construction is adjusted from a fixed value to an adaptive radius $r = \mu_d + 2\sigma_d(\mu_d = 2.3\,\text{m}, \sigma_d = 0.7\,\text{m})$.

Aiming at the asynchrony problem between the lidar and the visual odometer, an error state Kalman filter (ESKF) with timestamp alignment is designed. The Jacobian matrix of the observation model introduces the Lie group perturbation term:

$$H_k = \left. \frac{\partial h\left(\exp(\xi^\wedge) T_k\right)}{\partial \xi} \right|_{\xi=0} \tag{14}$$

The distribution characteristics of node positioning error and multimodal data synchronization delay in equipment-topography map structure modeling are shown in Figure 4. After collecting N = 200 groups of sample points, each sample corresponds to the positioning error of a map node (unit: cm) and the synchronization delay of a multimodal data collection (unit: ms). The positioning error is shown in Figure 4 (a) and follows a normal distribution, with a mean of 3.14 cm and a standard deviation of 1.48 cm. The error is mainly concentrated in the 2-4 cm range, indicating that the improved ICP (Iterative Closest Point) registration algorithm and Kalman filter effectively suppress SLAM drift. The synchronization delay is concentrated in the [30, 50] ms range, reflecting that the asynchronous error of the lidar, IMU and visual sensor after the timestamp alignment is effectively controlled.

Comparison of the two sets of data shows that the discreteness of the positioning error is low (the standard deviation is only 1.48cm), achieving the optimal node-truth matching strategy. The positioning error mainly originates from the mapping and positioning system based on LiDAR and SLAM fusion, rather than GPS or pure visual odometer. The synchronization delay fluctuation is relatively significant (the right-skewed distribution tail extends to 70ms), which is related to the state update mechanism of the ESKF filter. Although the Lie group perturbation term can compensate for the IMU bias and spatial discretization errors. The difference in computational load between visual feature extraction and lidar point cloud processing leads to the accumulation of some data frame delays, which is manifested as a tail shift of the delay distribution, verifying the performance difference of the system in spatial modeling and time synchronization.

PATH SUCCESS RATE

This experiment uses the control variable method to design 10 typical substation scenarios, and verifies the obstacle avoidance and target reachability of the method in complex terrain through quantitative comparison. The scenarios include 5 types of terrain complexity (flat ground/gentle slope/step/gully/loose ground) and 2 types of obstacle density (low density: 5/100m2; high density: 15/100m2). Each type of scenario runs 30 benchmark tests independently, and the number of times the robot successfully reaches the target point without colliding or getting stuck in obstacles is counted, and the path success rate is calculated. The path planning module implements global path search based on the improved A* algorithm. Its cost function comprehensively considers the path length d_{path} and the terrain risk probability p_{risk}, and is defined as follows:

$$C = \alpha d_{\text{path}} + \beta \int p_{\text{risk}}(x)\, dx \tag{15}$$

where α and β are weight coefficients. The optimal configuration is determined through ablation experiments. The model predictive control framework is used to perform local path rolling optimization to improve realtime performance. In each control cycle, the local path trajectory is dynamically adjusted according to the current chassis posture and the latest terrain perception results.

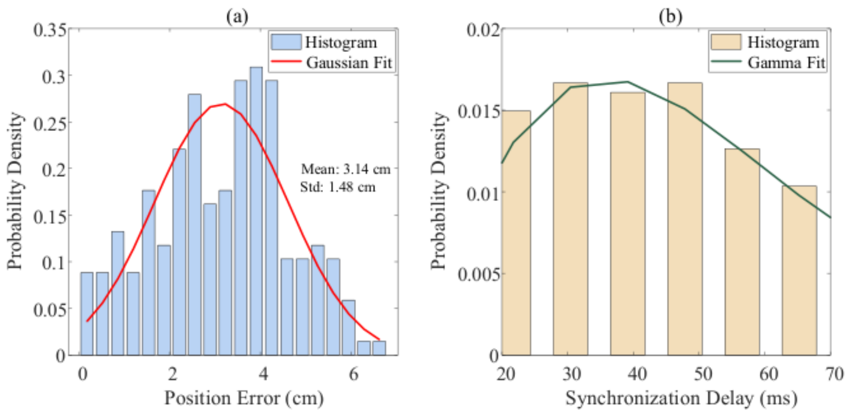

FIGURE 4. Sample error/delay distribution histogram: (a) Node Position Error Distribution; (b) Multimodal Data Sync Delay Distribution

To verify the effect of improving the path success rate, a comparative experiment was designed, using the traditional A* algorithm, the rapidly-exploring random tree search (RRT), the graph convolutional network driven planning (GCN), and the graph attention network optimization scheme (GAT) for path planning. The experimental data of the path success rate under different methods are shown in Table 2.

The results show that the traditional A* frequently freezes or rolls over in complex terrain because it does not consider terrain risks, and the average path success rate is 78.3%. RRT has an unstable path due to random sampling, and the success rate is 71.1%. Although GCN can guarantee the shortest path, it lacks the ability to avoid high-risk areas, and the success rate is 85.0%. GAT can guarantee the success rate in less complex environments, but the overall success rate is 86.4%. This method combines graph structure modeling with the risk map output by multimodal fusion perception, which significantly improves the obstacle avoidance ability. The path success rate reaches 91.6%, which is 13.3 percentage points higher than the traditional A*. It should be noted that the 'simulated terrain collapse' in the experiment is a predefined, controllable scenario synthesized from CAD and DEM. The generalization limitation primarily refers to the insufficient adaptability to unknown collapse patterns beyond the training data distribution, such as sudden foundation collapse or unstructured rock piles.

Figure 5 compares the path length deviation (PLD) of different algorithms in 10 scenarios, reflecting the degree of deviation between the actual path and the optimal path, and the rate of risk-area pass (RRP), reflecting the proportion of crossing high-risk terrain. Taking scenario 4 (gully terrain) as an example, the path length deviation of the method used in this paper is 6.8%, which is significantly lower than 19.1% of RRT and 10.3% of traditional A*. The risk area pass rate is controlled at 6.3%, which is better than 8.0% of GCN and 7.7% of GAT. In the comprehensive comparison of 10 scenarios, the average path length deviation of the proposed method in the 10 scenarios is stable at 5.72%, and the average risk area pass rate is 5.17%. The average path length deviation of the traditional A* method is 8.9%, and the average risk area pass rate of RRT is as high as 14.06%, which proves that the method used in this paper has comprehensive advantages in path efficiency and safety.

ATTITUDE ANGLE FLUCTUATION RANGE

The experiment uses a six-axis IMU (sampling rate 200Hz, angle random walk noise $\leq 0.15°/\sqrt{\text{hr}}$) to synchronously collect Euler angle data during chassis movement. The Kalman filter observation noise $R = \text{diag}(0.1, 0.1, 0.5)$ is used to separate the gravity acceleration and motion acceleration components to obtain high-precision attitude angle θ (pitch angle) and ϕ (roll angle) sequences. The data collection cycle covers 10 typical terrain scenarios, and 30 repeated path planning experiments are performed for each scenario to record the dynamic response of the attitude angle during the complete movement process.

The calculation window is a 5-second transient response segment when the chassis passes through a characteristic terrain area (such as a slope top and a gully) to evaluate the attitude stability. The characteristic terrain boundary is identified through the terrain segmentation algorithm, and the attitude angle data processed by low-pass filtering in the characteristic interval is intercepted, and the standard deviation after mean normalization is calculated. The experimental comparison group includes traditional PID control, LQR (Linear Quadratic Regulator) control and pure reinforcement learning control strategy without the graph neural network (Reinforcement Learning without Graph Neural Network, RL without GNN). The LQR controller employed in the comparison is based on a linearized chassis dynamics model, with states including position, velocity, and attitude angles. The RL strategy without GNN adopts the same PPO framework as in this paper, but its states only include IMU attitude angles and chassis velocity, without incorporating graph embedding information. The attitude

TABLE 2. Path success rate (%)

Terrain Type	Obstacle Density	Proposed Method	A*	RRT	GCN	GAT
Flat Ground	Low	96.7	88.3	82.5	91.2	93.5
Gentle Slope	High	90.0	76.7	68.3	82.5	83.7
Steps	Low	93.3	81.7	75.0	87.3	89.2
Gullies	High	88.3	73.5	65.2	80.4	81.6
Loose Ground	Low	93.0	82.1	77.6	88.5	89.7
Mixed Terrain (Flat + Slope)	High	91.7	78.9	70.5	85.6	86.4
Mixed Terrain (Steps + Gullies)	Low	90.3	79.2	72.1	84.9	86.1
Simulated Terrain Collapse	Medium	86.5	66.9	59.3	77.6	78.9
Dynamic Obstacle Replanning	Dynamic	92.1	74.8	67.0	83.3	84.5
Multi-Target Sequential Task	High	94.0	80.5	73.9	89.1	90.4
Mean	–	91.6	78.3	71.1	85.0	86.4

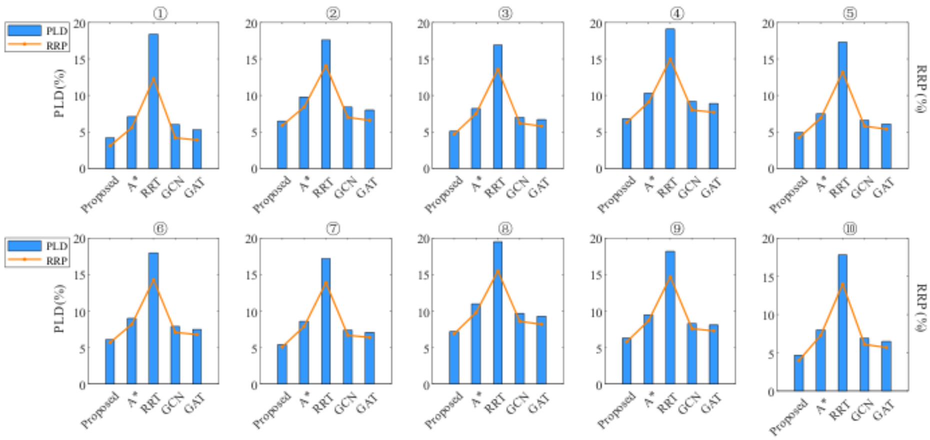

FIGURE 5. Path length deviation/risk area pass rate indicator evaluation

stability quantification uses two indicators including the standard deviation σ_θ of the pitch angle sequence and the standard deviation σ_ϕ of the roll angle sequence. In order to study the correlation between attitude fluctuation and control strategy, a triple analysis mechanism is established: 1) Time domain statistical analysis: the box plot distribution of σ_θ and σ_ϕ in each scenario is calculated as shown in Figure 6, outliers are identified and the Pearson correlation coefficient between terrain complexity (quantified by slope variance σ^2_{slope}) and attitude fluctuation is analyzed as shown in Table 3. 2) Frequency domain energy distribution analysis: The attitude angle energy density spectrum is extracted through continuous wavelet transform to locate the main oscillation frequency components. The attitude angle energy density spectrum curve is shown in Figure 7; 3) Control input correlation analysis: The cross-correlation function between the attitude fluctuation peak and the motor torque $\Delta\tau$ is calculated (maximum lag step size 100ms) to verify the response delay of dynamic adjustment.

Figure 6 shows the box plot distribution of the pitch angle standard deviation σ_θ and the roll angle standard deviation σ_ϕ under the control of PID, LQR, RL without GNN and the proposed method in 10 typical terrain scenes. The vertical axis represents 10 terrain scenes of different complexity, and the horizontal axis is the attitude angle fluctuation value. The fluctuation distribution of the pitch angle and roll angle is shown in each scene. As the complexity of the terrain increases, the posture fluctuation of the traditional control method generally increases, especially in complex terrains such as steps and gullies. The Graphormer-based control method proposed in this paper shows a smaller and more stable fluctuation range in most scenarios. The Pearson correlation coefficient in Table 3 further quantifies the relationship between posture fluctuation and terrain complexity. In complex terrain with steps and gullies, the correlation coefficient between σ_θ and σ^2_{slope} under PID control is as high as 0.85, indicating that it is highly sensitive to terrain disturbances. However, in the proposed method, this coefficient is significantly reduced to 0.51, indicating that the model effectively reduces the dependence of attitude response on terrain changes.

Figure 7 shows the frequency domain characteristics of attitude angle fluctuations under different control strategies, including the power spectral density (PSD) curve and the energy proportion of the 4-8 Hz frequency band. In Figure 7(a), the horizontal axis is the frequency range (0-15 Hz) and the vertical axis is the energy density value (g2/Hz). As shown in Figure 7(a), in the traditional PID, LQR and RL without GNN methods, the attitude fluctuation has obvious resonance peaks in the 4-8 Hz frequency band,

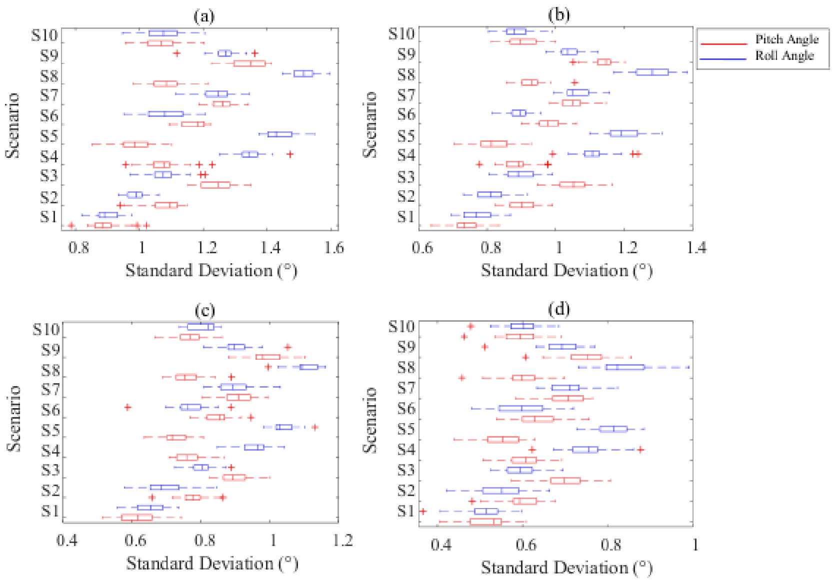

FIGURE 6. Standard deviation of pitch angle sequence σ_θ and standard deviation of roll angle sequence σ_ϕ in each scenario: (a) Traditional PID control; (b) LQR control; (c) RL without GNN; (d) Method in this paper

TABLE 3. Pearson correlation coefficient

Terrain Type	σ^2_{slope}	Metric	PID	LQR	RL without GNN	Proposed Method
Flat Ground	0.05	$\sigma_\theta - \sigma^2_{slope}$	0.23 ($p = 0.20$)	0.19 ($p = 0.30$)	0.17 ($p = 0.36$)	0.11 ($p = 0.53$)
Flat Ground	0.05	$\sigma_\phi - \sigma^2_{slope}$	0.18 ($p = 0.32$)	0.16 ($p = 0.38$)	0.14 ($p = 0.45$)	0.09 ($p = 0.61$)
Gentle Slope	0.22	$\sigma_\theta - \sigma^2_{slope}$	0.61 ($p = 0.001$)	0.53 ($p = 0.003$)	0.47 ($p = 0.008$)	0.29 ($p = 0.11$)
Gentle Slope	0.22	$\sigma_\phi - \sigma^2_{slope}$	0.57 ($p = 0.002$)	0.49 ($p = 0.006$)	0.43 ($p = 0.012$)	0.25 ($p = 0.17$)
Steps	0.35	$\sigma_\theta - \sigma^2_{slope}$	0.81 ($p < 0.001$)	0.74 ($p < 0.001$)	0.67 ($p = 0.001$)	0.42 ($p = 0.018$)
Steps	0.35	$\sigma_\phi - \sigma^2_{slope}$	0.79 ($p < 0.001$)	0.71 ($p < 0.001$)	0.63 ($p = 0.001$)	0.39 ($p = 0.027$)
Gullies	0.41	$\sigma_\theta - \sigma^2_{slope}$	0.76 ($p < 0.001$)	0.69 ($p < 0.001$)	0.61 ($p = 0.001$)	0.36 ($p = 0.041$)
Gullies	0.41	$\sigma_\phi - \sigma^2_{slope}$	0.83 ($p < 0.001$)	0.78 ($p < 0.001$)	0.72 ($p < 0.001$)	0.44 ($p = 0.013$)
Loose Ground	0.29	$\sigma_\theta - \sigma^2_{slope}$	0.69 ($p < 0.001$)	0.62 ($p = 0.001$)	0.56 ($p = 0.002$)	0.33 ($p = 0.065$)
Loose Ground	0.29	$\sigma_\phi - \sigma^2_{slope}$	0.72 ($p < 0.001$)	0.66 ($p = 0.001$)	0.61 ($p = 0.001$)	0.38 ($p = 0.039$)
Mixed Terrain (Flat + Slope)	0.27	$\sigma_\theta - \sigma^2_{slope}$	0.63 ($p = 0.001$)	0.57 ($p = 0.002$)	0.51 ($p = 0.004$)	0.28 ($p = 0.13$)
Mixed Terrain (Flat + Slope)	0.27	$\sigma_\phi - \sigma^2_{slope}$	0.58 ($p = 0.002$)	0.53 ($p = 0.004$)	0.48 ($p = 0.007$)	0.24 ($p = 0.19$)
Mixed Terrain (Steps + Gullies)	0.52	$\sigma_\theta - \sigma^2_{slope}$	0.85 ($p < 0.001$)	0.81 ($p < 0.001$)	0.76 ($p < 0.001$)	0.51 ($p = 0.003$)
Mixed Terrain (Steps + Gullies)	0.52	$\sigma_\phi - \sigma^2_{slope}$	0.87 ($p < 0.001$)	0.83 ($p < 0.001$)	0.79 ($p < 0.001$)	0.53 ($p = 0.002$)
Simulated Terrain Collapse	0.61	$\sigma_\theta - \sigma^2_{slope}$	0.82 ($p < 0.001$)	0.79 ($p < 0.001$)	0.73 ($p < 0.001$)	0.48 ($p = 0.006$)
Simulated Terrain Collapse	0.61	$\sigma_\phi - \sigma^2_{slope}$	0.89 ($p < 0.001$)	0.86 ($p < 0.001$)	0.81 ($p < 0.001$)	0.55 ($p = 0.002$)
Dynamic Obstacle Replanning	0.38	$\sigma_\theta - \sigma^2_{slope}$	0.73 ($p < 0.001$)	0.68 ($p = 0.001$)	0.63 ($p = 0.001$)	0.41 ($p = 0.023$)
Dynamic Obstacle Replanning	0.38	$\sigma_\phi - \sigma^2_{slope}$	0.71 ($p < 0.001$)	0.65 ($p = 0.001$)	0.60 ($p = 0.001$)	0.37 ($p = 0.034$)
Multi-Target Sequential Task	0.33	$\sigma_\theta - \sigma^2_{slope}$	0.67 ($p = 0.001$)	0.61 ($p = 0.002$)	0.58 ($p = 0.002$)	0.35 ($p = 0.047$)
Multi-Target Sequential Task	0.33	$\sigma_\phi - \sigma^2_{slope}$	0.64 ($p = 0.002$)	0.59 ($p = 0.003$)	0.55 ($p = 0.003$)	0.31 ($p = 0.078$)

indicating that the chassis system is susceptible to terrain disturbances in this frequency range and produces continuous oscillations. The Graphormer-based method proposed in this paper significantly reduces the energy density in this frequency band, showing a stronger attitude suppression capability. Figure 7(b) further quantifies this phenomenon. The statistical results show that the pitch angle and roll angle fluctuation energy in the 4~8Hz frequency band of the proposed method account for 14.0% and 15.5% respectively, which is more than 12.8% lower than that of other methods, indicating that it effectively suppresses the unstable response in the key frequency band.

In order to further explore the dynamic response relationship between attitude fluctuation and chassis control input, this paper further carried out control input correlation analysis. Based on the cross-correlation function (CCF) of attitude fluctuation peak value and motor torque change $\Delta\tau$, and zero-mean processing, the cross-correlation function between attitude fluctuation and control input is constructed as follows:

$$R_{\bar{\theta},\bar{\Delta\tau}}(\tau_d) = \frac{1}{N}\sum_{t=1}^{N} \bar{\theta}(t) \cdot \bar{\Delta\tau}(t+\tau_d) \tag{17}$$

$$R_{\bar{\phi},\bar{\Delta\tau}}(\tau_d) = \frac{1}{N}\sum_{t=1}^{N} \bar{\phi}(t) \cdot \bar{\Delta\tau}(t+\tau_d) \tag{18}$$

where τ_d represents the lag step (unit: ms); N is the total number of sampling points; $\bar{\Delta\tau}$ is the change in motor output torque; $\bar{\theta}$ and $\bar{\phi}$ are normalized attitude fluctuation sequences.

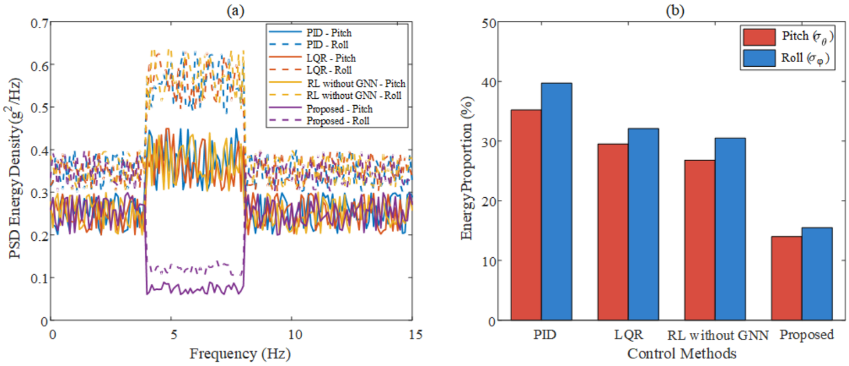

FIGURE 7. Frequency domain characteristics of attitude angle fluctuations: (a) Power Spectral Density of Attitude Fluctuations; (b) Energy Proportion in 4–8 Hz Band

By calculating the time lag position corresponding to the maximum correlation coefficient of $R_{\bar{\theta},\bar{\Delta\tau}}(\tau_d)$ and $R_{\bar{\phi},\bar{\Delta\tau}}(\tau_d)$ in the lag range of [0,100] ms, the time delay from the occurrence of attitude disturbance to the adjustment of control input can be obtained. The results in Figure 8 show that in the traditional PID control strategy, the maximum correlation between the posture fluctuation peak and the motor torque adjustment occurs at $\tau_d = 85$ ms, while in the proposed method it is significantly advanced to $\tau_d = 32$ ms, indicating that the model has better real-time response performance.

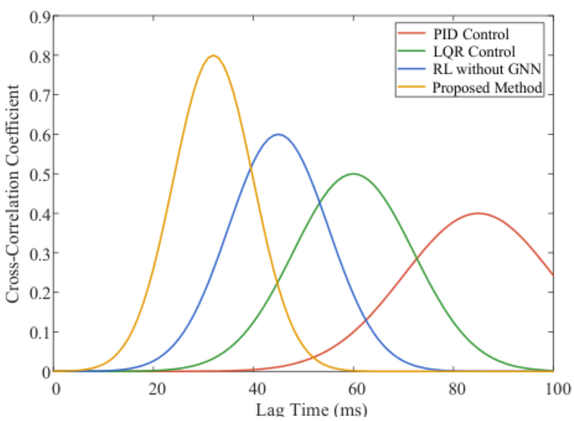

FIGURE 8. Hysteresis curve

IV. CONCLUSION

The Graphormer-based substation equipment-topographic map collaborative modeling and chassis control framework proposed in this paper achieves stable navigation in complex environments through a three-stage algorithm design, providing a robust solution for ensuring the continuous monitoring and power stability vital for sensitive industrial operations, such as those within the manufacturing sector. This paper combines LiDAR and CAD topology to construct a device-terrain map structure, and uses centrality coding and spatial coding to quantify the importance and positional relationship of nodes. The graph embedding vector is dynamically updated through a multi-head attention mechanism to provide global terrain perception for the control module; cross-modal attention is used to fuse IMU, vision and LiDAR features, and reinforcement learning is combined to generate a multi-modal driven chassis control strategy. Finally, closed-loop optimization is achieved through incremental graph update and MPC rolling planning. Experimental results show that the average path success rate of this method in 10 types of terrain is 13.3% higher than that of traditional A*, and the fluctuation of posture angle is reduced by more than 12.8% compared with PID, LQR, and RL without GNN. However, the generalization ability of the current model in more complex terrains (such as collapsed areas) is still limited, and the attention mechanism of Graphormer is time-consuming to calculate, and it is necessary to further introduce a sparse strategy to optimize efficiency. Future work can be extended to dynamic obstacle prediction and multi-robot collaborative navigation, and explore its versatility in industrial scenarios other than power equipment inspection, such as leveraging this framework for autonomous material handling and site monitoring within complex, large-scale manufacturing facilities.

AUTHOR CONTRIBUTIONS

Conceptualization –Jian Duan, Yinghui Lu, Jianxun Zhao and Yongchao Zhu; methodology – Jian Duan, Yinghui Lu, Jianxun Zhao and Yongchao Zhu; investigation – Jian Duan, Yinghui Lu and Jianxun Zhao ; writing-original draft preparation – Jian Duan, Yinghui Lu, Jianxun Zhao and Yongchao Zhu. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research was supported by Science and Technology Project funded by State Grid Hebei Electric Power Co., LTD. (kj2023-090- Intelligent spraying robot development).

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] Z. Qin, Z. D. Xu, Q. C. Sun, and P. Balamurugan, “Investigation of intelligent substation inspection robot by using mobile data,” International journal of humanoid robotics, vol. 20, no. 02n03, Art. no. 2240003, 2023, doi: 10.1142/S0219843622400035.

[2] L. Dong, N. Chen, J. Liang, T. Li, Z. Yan, and B. Zhang, “A review of indoor-orbital electrical inspection robots in substations,” Industrial Robot: the international journal of robotics research and application, vol. 50, no. 2, pp. 337-352, 2022, doi: 10.1108/IR-06-2022-0162.

[3] J. Ren, T. Wu, X. Zhou, C. Yang, J. Sun, M. Li, et al., “SLAM, path planning algorithm and application research of an indoor substation wheeled robot navigation system,” Electronics, vol. 11, no. 12, pp. 1838, 2022, doi: 10.3390/electronics11121838.

[4] Y. Yan, W. Jiang, Z. Luo, J. Zhang, and W. Liu, “System optimization and robustness stability control for GIS inspection robot in complex microgrid networks,” Complexity, vol. 2021, no. 1, Art. no. 6691905, 2021, doi: 10.1155/2021/6691905.

[5] D. Luo, X. Huang, Y. Huang, M. Miao, and X. Gao, “Optimal Trajectory Planning for Wheeled Robots (OTPWR): A Globally and Dynamically Optimal Trajectory Planning Method for Wheeled Mobile Robots,” Machines, vol. 12, no. 10, pp. 668, 2024, doi: 10.3390/machines12100668.

[6] Z. Liu, J. Wang, Z. Hui, B. Zhao, and F. Zhang, “Research on positioning and mapping algorithm of sliding window optimization for substation monitoring robot,” Energy Reports, vol. 9, no. 10, pp. 898-908, 2023, doi: 10.1016/j.egyr.2023.05.166.

[7] L. Yang, C. He, Y. Yao, Y. Zhang, and J. Guo, “Research on a detection robot for power equipment status in hazardous scenarios,” Journal of Computational Methods in Sciences and Engineering, vol. 25, no. 1, pp. 223-242, 2025, doi: 10.1177/14727978241299592.

[8] C. Wang, Z. Wang, H. Hu, and L. Li, “Innovative design and kinematic characteristics analysis of floating mobile chassis of inspection robot,” Machines, vol. 11, no. 1, pp. 24, 2022, doi: 10.3390/machines11010024.

[9] C. Yao and S. Liu, “Design of a Control System for an Indoor Substation Patrol Robot with Lifting Platform,” Journal of East China University of Science and Technology, vol. 47, no. 1, pp. 116-122, 2021, doi: 10.14135/j.cnki.1006-3080.20191203005.

[10] Y. Ren, N. Yan, X. Yu, F. Tang, Q. Tang, Y. Wang, et al., “On automatic camera shooting systems via PTZ control and DNN-based visual sensing,” Intelligent Service Robotics, vol. 16, no. 3, pp. 265-285, 2023, doi: 10.1007/s11370-023-00462-w.

[11] D. Yan, H. Cao, T. Wang, R. Chen, and S. Xue, “Graph-based knowledge acquisition with convolutional networks for distribution network patrol robots,” IEEE Transactions on Artificial Intelligence, vol. 2, no. 5, pp. 384-393, 2021, doi: 10.1109/TAI.2021.3087116.

[12] A. Islam M, M. Ismail, R. Atat, and O. Boyaci, “Software-defined network-based proactive routing strategy in smart power grids using graph neural network and reinforcement learning,” E-Prime-Advances in Electrical Engineering, Electronics and Energy, vol. 5, no. 9, Art. no. 100187, 2023, doi: 10.1016/j.prime.2023.100187.

[13] J. Ruan, G. Liang, J. Zhao, H. Zhao, J. Qiu, F. Wen, et al., “Deep learning for cybersecurity in smart grids: Review and perspectives,” Energy Conversion and Economics, vol. 4, no. 4, pp. 233-251, 2023, doi: 10.1049/enc2.12091.

[14] Z. Zhao, B. Liu, Y. Zhai, W. Zhao, and P. Su, “Dual graph reasoning network for oil leakage segmentation in substation equipment,” IEEE Transactions on Instrumentation and Measurement, pp. (11); 73:1-15, 2023, doi: 10.1109/TIM.2023.3336439.

[15] H. Liu and W. Liu, “A Deep Learning-Based Fault Diagnosis Approach for Power System Equipment via Infrared Image Sensing,” Journal of Circuits, Systems and Computers, vol. 33, no. 11, Art. no. 2450154, 2024, doi: 10.1142/S0218126624501548.

[16] S. A. Kashinath, S. A. Mostafa, D. Lim, A. Mustapha, H. Hafit, and R. Darman, “A general framework of multiple coordinative data fusion modules for real-time and heterogeneous data sources,” Journal of Intelligent Systems, vol. 30, no. 1, pp. 947-965, 2021, doi: 10.1515/jisys-2021-0083.

[17] Q. Yang, J. Dong, M. Tan, J. Wang, D. Guo, H. Kang, et al., “A novel navigation assistant method for substation inspection robot based on multisensory information fusion,” Journal of advanced research, S2090-1232 (25) 00035-9, doi: 10.1016/j.jare.2025.01.016.

[18] S. Zheng, W. Du, J. Wang, G. Yang, and Y. Zhao, “Information Monitoring of Transmission Line Operating Environment based on Internet of Things Technology,” Scalable Computing: Practice and Experience, vol. 26, no. 2, pp. 699–707, 2025, doi: 10.12694/scpe.v26i2.3976.

[19] R. Chang, S. Zhou, Y. Zhang, N. Zhang, C. Zhou, and M. Li, “Research on insulator defect detection based on improved YOLOv7 and multi-UAV cooperative system,” Coatings, vol. 13, no. 5, pp. 880, 2023, doi: 10.3390/coatings13050880.

[20] M. M. Rahman, S. Akash A, H. Pirim, et al., “Enhancing Electrical Network Vulnerability Assessment with Machine Learning and Deep Learning Techniques,” Northeast Journal of Complex Systems (NEJCS), vol. 6, no. 1, pp. 2, 2024, doi: 10.22191/nejcs/vol6/iss1/2.

[21] J. Jia, Q. Li, C. Xu, L. Zhang, and S. Zhao, “Depth augmentation and cross-attention progressive fusion network for semantic segmentation using RGB-depth images of power equipment,” Journal of Electronic Imaging, vol. 34, no. 2, pp. 023064-023064, 2025, doi: 10.1117/1.JEI.34.2.023064.

[22] Z. Li, O. Yoshie, H. Wu, X. Mai, Y. Yang, and X. Qu, “Image Analysis Method of Substation Equipment Status Based on Cross-Modal Learning,” IEEJ Transactions on Electrical and Electronic Engineering, vol. 19, no. 9, pp. 1507-1521, 2024, doi: 10.1002/tee.24111.

[23] W. Zhang, W. Li, X. Zheng, and B. Sun, “Improved A* and DWA fusion algorithm based path planning for intelligent substation inspection robot,” Measurement and Control, vol. 1, no. 1, Art. no. 00202940251316687, 2025, doi: 10.1177/00202940251316687.

[24] D. Zuo, Y. Zhao, W. Zhang, and B. Sun, “Tube-MPC based Trajectory Tracking Control for Substation Inspection Robot,” Engineering Research Express, vol. 7, no. 015247, pp. 1-12, 2025, doi: 10.1088/2631-8695/adae57.

[25] P. Wang, C. Li, F. Cai, and L. Zheng, “An improved SLAM algorithm for substation inspection robot based on the fusion of IMU and visual information,” Energy Informatics, vol. 7, no. 1, pp. 86, 2024, doi: 10.1186/s42162-024-00390-8.

[26] Q. Yuan, Y. Luo, and H. S. Wang, “3D point cloud recognition of substation equipment based on plane detection,” Results in Engineering, vol. 15, no. 9, Art. no. 100545, 2022, doi: 10.1016/j.rineng.2022.100545.