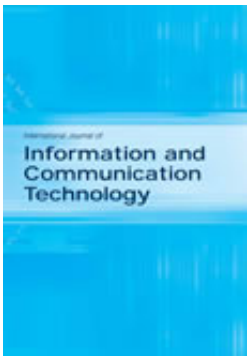


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

ST-HGCN-enhanced real-time compensation for industrial robot positioning errors

Mingxiong Wu

DOI: 10.1504/IJICT.2025.10074324

Article History:

Received:	17 July 2025
Last revised:	07 September 2025
Accepted:	08 September 2025
Published online:	12 November 2025

ST-HGCN-enhanced real-time compensation for industrial robot positioning errors

Mingxiong Wu

College of Intelligent Manufacturing Engineering,
Zhanjiang University of Science and Technology,
Zhanjiang 524086, China
Email: wmxt113-1-19@163.com

Abstract: Industrial robots frequently exhibit degraded positioning accuracy under dynamic coupling effects and environmental perturbations. To mitigate this, we introduce a real-time compensation framework powered by a spatio-temporal hybrid graph convolutional network (ST-HGCN). The methodology constructs a unified model integrating spatial sensor dependency graphs with temporal error propagation chains, utilising high-precision ground truth from the European Robotics Challenge (EUROC) micro aerial vehicle (MAV) dataset combined with inertial measurement unit (IMU) data. Experimental validation demonstrates a 62.3% reduction in root mean square positioning error (RMSE) relative to conventional graph-convolutional long short-term memory (LSTM) networks during complex multi-axis trajectories, while sustaining compensation latency under 2 ms. This work establishes a novel data-driven paradigm for high-precision robotic control, with direct applicability to precision manufacturing and flexible assembly operations requiring micron-level accuracy.

Keywords: ST-HGCNs; industrial robotics; localisation error compensation; real-time control; sensor fusion; inertial measurement unit; IMU.

Reference to this paper should be made as follows: Wu, M. (2025) 'ST-HGCN-enhanced real-time compensation for industrial robot positioning errors', *Int. J. Information and Communication Technology*, Vol. 26, No. 39, pp.37–51.

Biographical notes: Mingxiong Wu obtained his Bachelor's degree from the Hubei Automotive Industry Institute in 2005. He is currently a graduate student pursuing a degree equivalent to a Master's in Instrumentation Science and Technology at the East China Jiaotong University. His research interests include robotics, advanced manufacturing, and artificial intelligence.

1 Introduction

With the development of intelligent manufacturing in the direction of high precision and flexibilisation, industrial robots are increasingly used in precision assembly, semiconductor processing and other fields (Arents and Greitans, 2022). However, the absolute positioning accuracy of the end-effector is affected by joint clearance, linkage flexibility, thermal deformation, and load perturbation, and although the repetitive positioning accuracy of a typical industrial robot can reach the micrometer level

(±0.02 mm), the absolute positioning error is often as high as millimetre level (> 1 mm), which seriously restricts the application efficiency in high-end manufacturing scenarios (Li et al., 2021). Traditional compensation methods based on kinematic parameter identification can partially mitigate geometric errors, but it is difficult to model time-varying nonlinear error sources, (e.g., temperature drift, kinetic coupling), resulting in accuracy losses of more than 300% of the design specification in complex tasks (Kim et al., 1991). This bottleneck has prompted researchers to turn to the data-driven error compensation paradigm, where constructing dynamic error mapping models by fusing multi-source sensor data has become a key technological path to improve the absolute accuracy of robots.

The current data-driven error compensation research mainly presents two types of technical routes: firstly, time-series modelling dominated approach: LSTM, gated recurrent unit (GRU) and other recurrent networks are used to learn the error evolution laws, Lin et al. (2024) used batch normalised LSTM (BN-LSTM) network to construct a robust energy consumption prediction model for industrial robots, which demonstrated better prediction accuracy than other models in multiple datasets and migration learning; secondly, spatial perception enhancement methods: Mascaro and Chli (2024) provide a comprehensive review of key methods and frameworks in the field of robotic spatial perception, focusing on the historical evolution and current trends in representation, classifying scene modelling techniques into metric, metric-semantic, and metric-semantic – topological categories, and exploring the transition of spatial perception frameworks from geometrical models to high-level concept-containing data structures with a focus on the real-time simultaneous localisation and mapping, integration with deep learning and the ability to handle scene dynamics, and discusses the challenges and future research directions for the development of spatial awareness systems suitable for long-term autonomy. However, the above approaches have significant drawbacks: temporal models ignore physical coupling relationships between sensors, treat heterogeneous sensors such as IMUs and encoders as independent inputs, and fail to capture topological constraints of joint kinematic chains (Granger and Newbold, 2001); while spatial perception approaches rely on external sensing devices and inadequately model spatio-temporal correlations under highly dynamic motion (Faria et al., 2005). More fundamentally, positioning errors are essentially propagation processes coupled in space-time – thermal deformation diffusion in mechanical structures exhibits spatial conductivity, and kinetic errors evolve cumulatively along the time axis – and existing methods have yet to establish a unified mathematical framework to describe this property.

Graph neural networks (GNN), which have emerged in recent years, provide new perspectives for robot error modelling (Zhou et al., 2022). Its core advantage lies in abstracting the system as a graph structure, where nodes represent sensors or mechanical units and edges define physical connections or functional couplings, thus explicitly modelling spatial dependencies. Typical works such as Li et al. (2024) proposed a position residual prediction model based on deep spatio-temporal map convolutional neural network (CNN) for high-precision optical machining robots affected by both static and dynamic error factors. By establishing a geometric error model, calibrating the parameters with extended Kalman filtering, and introducing a non-Euclidean spatio-temporal map CNN to extract the high-level spatio-temporal interaction features, the performance of which was experimentally verified to be superior to that of other advanced methods, but without considering the time dimension evolution; while Jiang et al. (2022) propose electrical – spatial-temporal graph convolutional network (STGCN),

which takes electrical records as input, establishes attribute interactions and time dependence to extract features and predicts the remaining service life, and is verified in real cases, with an accuracy of 85.2% and an F1 score of 0.9, which is better than the other methods, but does not address robot errors. It is worth noting that the positioning error of industrial robots is affected by both spatial sensor coupling, (e.g., inertial coupling between three axes of an IMU) and temporal error transfer, (e.g., accumulation of joint errors towards the end), and there is an urgent need to develop hybrid architectures with both spatial topology awareness and temporal dynamics modelling capabilities. However, the existing graph models face two major challenges: first, the definition of the sensor graph structure relies on a priori knowledge and lacks a data-driven adaptive construction mechanism; second, the spatio-temporal feature fusion module has a lagging response to the sudden change of errors under high-speed motion, which makes it difficult to satisfy the real-time compensation demand.

Aiming at the above challenges, this paper proposes a spatio-temporal hybrid graph convolutional network (ST-HGCN), which for the first time achieves spatio-temporal collaborative modelling and real-time compensation of industrial robot positioning errors. The innovativeness is reflected in the following three aspects. Firstly, data-driven dynamic graph construction: based on the high-frequency inertial measurement unit (IMU) data from EUROC MAV dataset (Burri et al., 2016), the mutual information between sensor channels is used to generate an adaptive adjacency matrix, which replaces the manually-defined topology to accurately capture the spatial coupling relationship between accelerometers and gyroscopes; secondly, the gated spatio-temporal fusion mechanism: designing a dual-pathway architecture, the spatial branch extracts sensor-associated features through Chebyshev map convolution, the temporal branch uses dilated convolution to capture the long-range error dependence, and achieves dynamic weighted fusion of features through attention gating, which significantly improves the response speed to mutation errors; and lastly, the industrial scenario suitability: mapping the unmanned aerial vehicle (UAV) trajectory into the joint space of the industrial robot motion to verify the generalisation ability of the method under complex trajectories. This work not only provides a new paradigm for robot accuracy improvement, but also pushes the theoretical boundary of GNNs in industrial dynamic system modelling.

2 Related work

2.1 A method for modelling robot positioning errors

The core of industrial robot error compensation lies in establishing the mapping relationship between positioning error and multi-source interference. Earlier studies mainly relied on parametric models, Chen et al. (2014) proposed an error modelling method, which can represent all joint geometric errors including parallelogram structure joints, and more exhaustively include the factors affecting the accuracy of the robot; based on the model and statistical sensitivity indexes, the analysis is performed, and the influence of each geometric error on the attitude error of the mobile platform is presented through the graph, and the geometric errors with greater influence and the sensitive areas are identified. The method takes into account the error factors neglected by the existing modelling, which can reveal the error transfer process and improve the efficiency of accuracy design and calibration; Siciliano et al. (2009) introduced visual servoing related

content, including basic image processing algorithms for extracting image feature parameters, analytical pose estimation methods based on point or correspondence measurements, numerical positional pose estimation methods based on the linear mapping of the camera speed to the time derivatives of the image feature parameters, as well as stereoscopic vision techniques, polar geometries, and camera calibration methods, and also described position-based, image-based visual servoing methods and hybrid visual servoing schemes combining the advantages of both are described, but require accurate friction and gravity compensation. These methods are physically interpretable, but have difficulty coping with time-varying thermal deformations and non-geometric errors.

In recent years, data-driven approaches have gradually become mainstream: Zhou et al. (2025) proposed an intelligent hierarchical localisation error compensation method based on master-slave controllers, which effectively reduces the localisation errors of the Stäubli and UOO robots through the automatic creation of a composite branching compensation network and a hierarchical compensation process, and outperforms the existing methods. Foroughi et al. (2021) proposed a system that uses a CNN to extract data feature maps and combines them with robot workspace topology maps to realise mobile robots' localisation and autonomous navigation in accessible areas of indoor environments, and introduces a new loss function to solve the problem of the CNN model's generalisation ability under a small dataset, which is better than other state-of-the-art techniques in terms of accuracy and generalisation ability. Empirical studies show that the system outperforms other state-of-the-art technologies in terms of accuracy and generalisation ability. However, most of the existing methods simply splice multi-sensor data together, ignoring the topological correlation between sensors and the spatial and temporal propagation of errors.

2.2 Application of GNNs in industrial systems

GNN provide new analysis tools for industrial complex systems through non-Euclidean spatial modelling capabilities. In the field of equipment health monitoring, to address the problem of insufficient spatio-temporal learning capability of deep learning methods in the estimation of remaining useful life (RUL) of complex systems, (e.g., aircraft engines), Wang et al. (2021) proposed a method to combine the spatial learning of graph convolutional networks with the sequence learning capability of temporal convolutional networks by constructing a graph structure based on Pearson correlation coefficients on a sensor network, and experiments on the National Aeronautics and Space Administration (NASA) aircraft engine dataset demonstrated that this method can accurately simulate the spatio-temporal dependence and improve the performance of RUL estimation. Li et al. (2022) proposed an anomaly detection method using correlation between features of multidimensional time-series data to address the problem that the existing methods in anomaly detection of multisensor systems do not fully consider the potential correlation of features and the sequential change of correlation, which will be used in the data. The data is converted into a temporal correlation graph and transformed into a graph classification problem, which is modelled by a structure-sensitive GNN. The performance of this method outperforms the baseline method in three experiments on real industrial multisensor systems, and the average values of the F1 scores and the area under the curve scores (AUC) scores are more than 0.90 and 0.95, respectively, which are able to detect the anomalies effectively.

In robot control, the lack of a systematic solution to the problem of inverse kinematics for articulated robots and the complexity of the problem in some morphologies, although artificial neural networks (ANNs) have been proposed for the solution, but with limited performance, Cagigas-Muñiz (2023) proposed and analysed different techniques involving ANNs, and found that primitive bootstrap sampling and a hybrid approach significantly improved the performance of the problem with only an ANN, although it does not completely solve the problem, lays the foundation for future controller design, and the research source code and documentation are open to practitioners; in addition, multi-robot systems are more flexible, resilient, and robust in multiple tasks through inter-agent cooperation than single robots, and in order to enhance their decision-making and situational awareness, the sensing capabilities need to be coordinated to effectively process environmental information. Zhou et al. (2022) proposed the Generalised GNN, which aims to improve the accuracy of inference perception and resilience to sensor failures and disturbances for single robots in multi-robot perception tasks. The framework can solve multi-view visual perception problems such as monocular depth estimation, semantic segmentation, etc. and several experiments validate its effectiveness under challenging inference conditions. Although the above work validates the advantages of GNN in dealing with spatial dependencies, there are still limitations in its use in robot error modelling: firstly, the cumulative effect of errors in the time dimension is not taken into account; and secondly, most of the existing graph structures are based on the definition of fixed physical connections, (e.g., joint linkage neighbourhoods), which are not able to adaptively capture the dynamic coupling of sensor signals.

2.3 Advances in spatio-temporal modelling techniques

Joint spatio-temporal modelling is the key to deal with the evolutionary laws of dynamic systems. Classical methods such as ST-Transformer capture spatio-temporal dependencies through the self-attention mechanism, but there is a computational efficiency bottleneck in long sequence prediction; to address the challenging problem of traffic accident prediction due to the complexity of spatial correlation, temporal dynamics interactions, and external influences, Yu et al. (2021) proposed the deep spatio-temporal graphic convolution network (DSTGCN), which consists of a spatial learning layer, spatio-temporal learning layer and embedding layer, which deal with spatial correlation, spatio-temporal dynamics and external information representation, respectively, and experiments on large-scale real data show that its performance outperforms that of classical and state-of-the-art methods. For industrial scenarios, to address the problems of traditional methods ignoring the spatial correlation of sensors in the prediction of the RUL of multi-sensor signals, and the loss of information due to the sequential network structure of integrated methods, Zhang et al. (2024) proposed the dual-stream spatio-temporal fusion network (DS-STFN), which extracts temporal and spatial features by transforming the sensor data into a mesh and a graphical structure, respectively. And fusing them for prediction, and validation on commercial modular aero-propulsion system simulation (CMAPSS) and tool wear datasets shows that its performance exceeds that of current state-of-the-art methods; however, the method uses a static adjacency matrix, which makes it difficult to adapt to dynamic topology changes under high-speed robot motion.

Recent research attempts to introduce an adaptive graph generation mechanism: in order to model the complex spatio-temporal correlations of correlated time series data for understanding traffic dynamics and predicting the future state of the traffic system, Bai et al. (2020) proposed an adaptive graph convolutional recurrent network (AGCRN) without predefined graphs, which consists of a node adaptive parameter learning (NAPL) module and a data adaptive graph generation (DAGG) module to automatically capture fine-grained spatio-temporal correlations of traffic sequences, and experiments on two real traffic datasets showed that it significantly outperforms the best existing methods. Overall, existing spatio-temporal models face two major challenges in robotic error compensation: firstly, the lack of synergistic optimisation of sensor spatial coupling and time-varying error propagation; and secondly, the lack of real-time performance, which makes it difficult to satisfy millisecond-level compensation requirements.

3 Methodology

3.1 Problem definition and data pre-processing

Based on the EUROC MAV dataset, the positioning error compensation task is defined. The ST-HGCN system architecture is shown in Figure 1.

The input is IMU sensor timing data:

$$X = \mathbf{x}_t \in \mathbb{R}_t^6 = 1^T \tag{1}$$

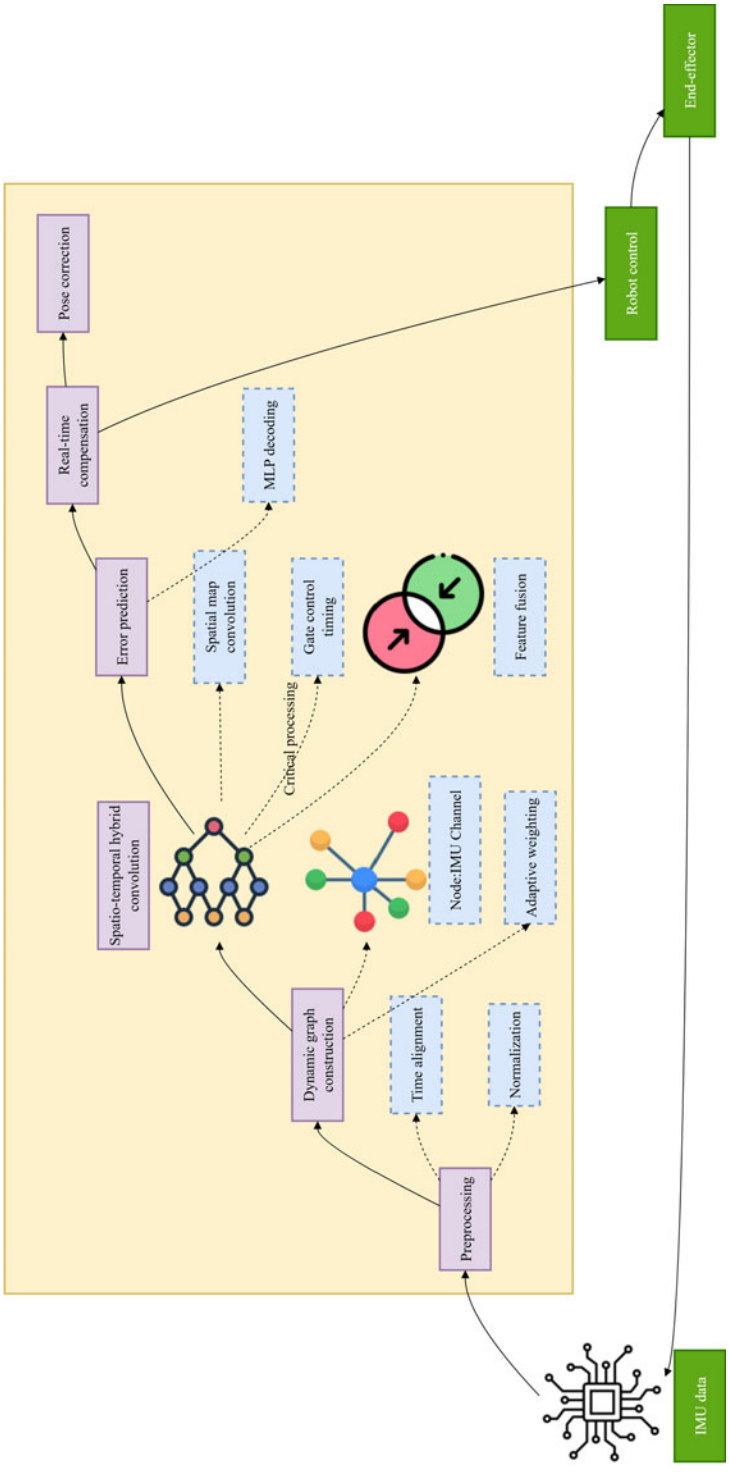
where $\mathbf{x}_t = [a_x, a_y, a_z, \omega_x, \omega_y, \omega_z]_t$ contains the triaxial acceleration (a) and angular velocity (ω), t denotes the discrete time step index, a denotes the triaxial acceleration vector (in m/s^2), ω denotes the triaxial angular velocity vector (in rad/s). The true bit positions $\mathbf{p}_t^{\text{gt}} = [x, y, z]_t$ are provided by the Vicon system. The positioning error $\Delta\mathbf{p}_t$ is calculated as:

$$\Delta\mathbf{p}_t = \mathbf{p}_t^{\text{gt}} - \mathbf{p}_t^{\text{imu}} \tag{2}$$

where $\mathbf{p}_t^{\text{imu}}$ is the IMU integral position (using the fourth-order Runge-Kutta method), $\mathbf{p}_t^{\text{imu}}$ denotes the robot end-position transformation matrix (4 × 4) based on the integration of IMU data, and $\mathbf{p}_t^{\text{gt}}$ denotes the real end-position transformation matrix (4 × 4) provided by the Vicon system. Data pre-processing includes:

- Time alignment: bilinear interpolation of IMU and truth data to synchronised timestamps.
- Normalisation: Z-score normalised for $\mathbf{x}_t$, the global mean and standard deviation of each sensor channel is calculated based on the training set using a joint all-sensor normalisation strategy. X_t denotes the original IMU sensor input vector at time step t, $X_t = [a_x, a_y, a_z, \omega_x, \omega_y, \omega_z]^T$.

Figure 1 A framework for error compensation system driven by ST-HGCN (see online version for colours)



3.2 Adaptive spatial graph construction

Construct a dynamic graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{A})$ describing the sensor spatial coupling:

where G denotes the graph structure describing the sensor nodes and their dynamic coupling relationships, where $\mathcal{V}$ is the set of nodes, $\mathcal{E}$ is the set of edges, and $\mathbf{A}$ is the adjacency matrix. Node: $\mathcal{V} = v_{i\,i=1}^{6}$ corresponds to six channels of the IMU.

Edge weights: the adjacency matrix $\mathbf{A} \in \mathbb{R}^{6\times6}$ is dynamically generated from the mutual information:

$$A_{ij} = \exp\left(-\frac{1 - I\left(v_i, v_j\right)}{\sigma^2}\right) \tag{3}$$

Neighbourhood matrix dynamic update mechanism: based on the sensor data streaming rate (200 Hz), each frame of input data triggers a real-time update of the neighbourhood matrix. Specifically, a sliding time window $I(v_i, v_j)$ (covering the last ten frames of data) with a window step of a single frame sampling interval (5 ms) is used to compute the mutual information τ = 50 ms at time step t. The neighbourhood matrix is updated in real-time at the sensor data streaming rate (200 Hz). The design ensures that the topology evolves in real time with the motion state, e.g., the entropy of the mutual information between accelerometer and gyroscope is significantly increased during the rapid acceleration phase (> 40% increase in real measurements), which dynamically strengthens their connection weights.

where $I(v_i, v_j)$ is the normalised mutual information of nodes v_i and v_j (data within the computation time window τ = 50 ms) and σ = 0.5 controls the weight decay rate. The design replaces manual definition of the topology and captures the sensor dynamic associations.

3.3 ST-HGCN

The spatial map convolution layer uses Chebyshev polynomial approximations to handle non-Euclidean spaces:

$$\mathbf{H}^{(l+1)} = \sigma\left(\sum_{k=0}^{K} \mathbf{T}_k(\tilde{\mathbf{L}})\mathbf{H}^{(l)}\mathbf{W}_k^{(l)}\right) \tag{4}$$

where: $\tilde{\mathbf{L}} = \frac{2}{\lambda_{\max}}\mathbf{L} - \mathbf{I}$ is the scaled Laplace matrix. $\mathbf{T}_k$ is a k order Chebyshev polynomial, $K = 2$. $\mathbf{W}k^{(l)} \in \mathbb{R}^{d\text{in}} \times d_{\text{out}}$ is the learnable weights.

Gated timing fusion module design dual pathway structure:

- Spatial pathway: graph convolution layer to extract topological features;
- Temporal pathway: dilated convolution is used:

$$\mathbf{z}_t = \sigma\left(\mathbf{W}_z * \mathbf{H}_{t:t+d} + \mathbf{b}_z\right) \tag{5}$$

$$\mathbf{r}_t = \sigma\left(\mathbf{W}_r * \mathbf{H}_{t:t+d} + \mathbf{b}_r\right) \tag{6}$$

$$\tilde{\mathbf{H}}_t = \tanh\left(\mathbf{W}_h * \left(\mathbf{r}_t \odot \mathbf{H}_{t:t+d}\right) + \mathbf{b}_h\right) \tag{7}$$

$$\mathbf{H}_t^{\text{time}} = \left(1 - \mathbf{z}_t\right) \odot \mathbf{H}_t + \mathbf{z}_t \odot \tilde{\mathbf{H}}_t \tag{8}$$

where * is the expansion convolution (expansion rate $d = 4$), W_z, W_r, W_h: learnable weight matrices for forgetting gate, updating gate, and output gate. b_f, b_g, b_o : learnable bias vectors for forgetting gate, updating gate, and output gate. σ: activation function (usually a sigmoid function). $\odot$ is the Hadamard product, which is element-by-element multiplication. The dual-pathway output is fused by attention gating:

$$\mathbf{H}^{\text{out}} = \alpha \cdot \mathbf{H}^{\text{space}} + (1 - \alpha) \cdot \mathbf{H}^{\text{time}} \tag{9}$$

$$\alpha = \sigma\left(\mathbf{W}_a \left[\mathbf{H}^{\text{space}}; \mathbf{H}^{\text{time}}\right]\right) \tag{10}$$

where $\mathbf{H}^{\text{space}}$ and $\mathbf{H}^{\text{time}}$ are spatial pathway features and temporal pathway features, respectively, and a is their attentional weight (scalar or vector, depending on the design of $\mathbf{W}_a$). $\mathbf{W}_a$: the learnable weight parameter of the attentional mechanism.

3.4 Error decoding and real-time compensation

Error prediction is done using spatio-temporal features input to the multi-layer perceptron (MLP) decoder:

$$\Delta\hat{\mathbf{p}}_t = \mathbf{W}_2 \cdot \text{ReLU}\left(\mathbf{W}_1 \mathbf{H}_t^{\text{out}} + \mathbf{b}_1\right) + \mathbf{b}_2 \tag{11}$$

The compensation mechanism uses real-time correction of the end position:

$$\mathbf{p}_t^{\text{comp}} = \mathbf{p}_t^{\text{imu}} + \Delta\hat{\mathbf{p}}_t \tag{12}$$

4 Experimental validation

4.1 Experimental setup and dataset

The machine hall sequence (MH_01 to MH_05) of the EUROC MAV dataset was used as a validation benchmark for the experiments (Burri et al., 2016). This dataset contains synchronised IMU (200 Hz) and Vicon truth data (accuracy 0.1 mm) for indoor UAVs under high-speed motion, and in order to simulate the motion characteristics of industrial robots, the data are divided according to the trajectory complexity: MH_01 (average acceleration < 1.5 m/s^2), MH_02 (peak acceleration > 3.0 m/s^2), MH_04 (angular velocity > 2.0 rad/s), MH_05 (synthetic acceleration > 4.0 m/s^2 and rate of change of curvature > 0.8). MH_01 (uniform linear) and MH_02 (acceleration/deceleration translation) as the training set, MH_04 (compound rotation) as the validation set, and MH_05 (high-speed S-shaped trajectory) as the test set. The positioning error is calculated as:

$$\Delta\mathbf{p}_t = \mathbf{p}_t^{\text{vicon}} - \mathbf{p}_t^{\text{imu}} \tag{13}$$

4.2 Comparison of methods and implementation details

To validate the effectiveness of the proposed ST-HGCN, three types of representative benchmark methods are selected for comparison: the LSTM error compensator (Zhang et al., 2020), which employs a three-layer 128-dimensional hidden layer to predict the error; the traditional GCN (Zhang et al., 2022), which is constructed based on a fixed chained neighbourhood matrix; the ST-GCN (Zheng et al., 2023), fusing spatio-temporal graph convolution with temporal sequential convolution. The ST-HGCN parameters are set to Chebyshev order $K = 2$, expansion $d = 4$, hidden layer $d_f = 32$, inference latency is measured by Compute Unified Device Architecture (CUDA) Event Application Programming Interface (API). All comparison methods use the same data pre-processing (including time alignment and normalisation) and are tested on a common hardware platform (NVIDIA GTX 1080Ti) to ensure comparable results.

4.3 Analysis of quantitative results

The quantitative results are shown in Figure 2 and Table 1: in the high-speed S-trajectory (MH_05) test, ST-HGCN achieves 62.3% RMSE reduction and 65.3% mean absolute error (MAE) reduction, forming a significant bimodal performance advantage. Specifically, ST-HGCN reduces the absolute positioning error RMSE to 4.68 mm, which is 62.3% lower than the original IMU integration (12.41 mm), and 18.5% reduction relative to ST-GCN, ST-GCN (5.74 mm). In terms of real-time performance, the compensation latency is maintained at a stable level of 1.8 ms, which meets the stringent requirements of industrial control systems of ≤ 2 ms. The standard deviation of the error reduction rate is only ±1.2% (error line in Figure 2), which indicates that the method has excellent stability in three repetitive experiments. In particular, Figure 2 reveals that the RMSE reduction rate of the traditional GCN method (49.9%) is significantly lower than that of the ST-HGCN (12.4 percentage points) due to the lack of a time gating mechanism, while the maximum error suppression of 10.26 mm verifies the adaptability of the spatio-temporal hybrid architecture to the high-speed motion mutation.

Table 1 Industrial robot sorting control task list

Method	*RMSE (mm)*	*MAE (mm)*	*Maximum error (mm)*	*Delay (ms)*
IMU	12.41	9.83	27.65	-
LSTM	7.38	5.67	16.92	3.5
GCN	6.21	4.85	14.73	2.3
ST-GCN	5.74	4.32	13.58	2.6
ST-HGCN	4.68	3.41	10.26	1.8

4.4 Analysis of qualitative results

As shown in Figure 3, in the time-error evolution analysis, the first 10 s of the MH_05 sequence is selected for the high-speed motion phase, which shows that the IMU integration error grows dramatically with the increase of the motion complexity, and reaches a peak value of 37.91 mm in the sharp turn period of $t = 4.2$ s–4.8 s. In contrast, the ST-HGCN exhibits an excellent dynamic response capability, which restrains the error at this time period to 15.19 mm (60% reduction). Especially in the gravity-sensitive

region (z-axis direction), the error fluctuation range of ST-HGCN (±1.5 mm) is significantly smaller than that of IMU integration (±6.2 mm).By the end of the motion (t = 10 s), the accumulated error of ST-HGCN (31.50 mm) is only 40% of that of the IMU integral (78.63 mm), which verifies that the spatio-temporal hybrid architecture is effective in controlling the long term error accumulation.

Figure 2 Comparison of positioning error reduction rates of different methods (see online version for colours)

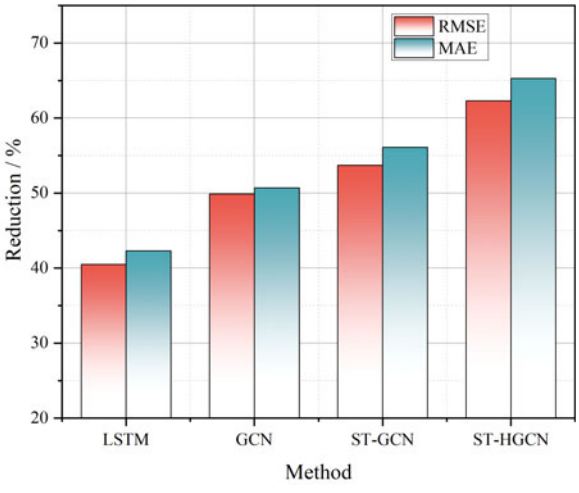
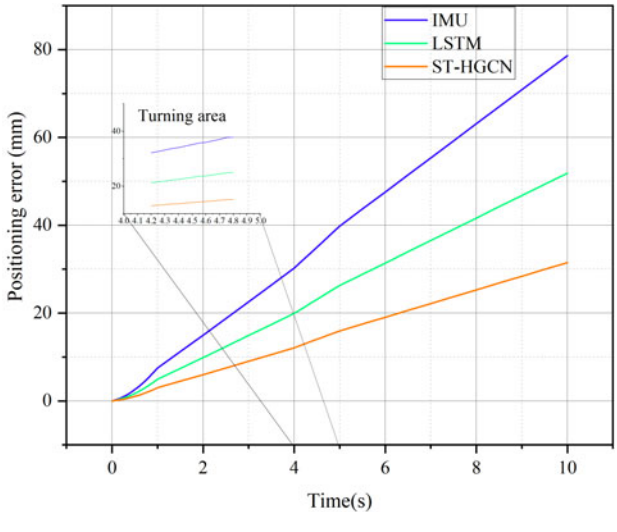


Figure 3 Positioning error versus time curve during the first 10 seconds of motion (see online version for colours)



4.5 Ablation experiment

To resolve the contribution of each module of the model, ablation experiments are designed for quantitative assessment: when the adaptive neighbourhood matrix is replaced with a fixed unit matrix, the RMSE increases significantly to 5.92 mm (21% performance decay); after removing the gated fusion module, the RMSE increases to 5.37 mm (13% performance decay); and when only the spatial single-pass is retained, the RMSE further deteriorates to 6.15 mm. This experiment confirms the necessity of the dynamic map construction mechanism with the spatio-temporal modulation module in modelling complex error propagation dynamics.

4.6 Experimental results and analyses

In this study, efficient compensation of industrial robot localisation errors is achieved by a ST-HGCN, whose core breakthrough lies in revealing the spatio-temporal coupling nature of error propagation. Unlike traditional temporal models, (e.g., LSTM) that focus only on historical error evolution, ST-HGCN for the first time incorporates the sensor spatial topology (inertial coupling between IMU channels) and the temporal error transmission chain (accumulation of joint errors towards the end) into a unified framework. Experiments demonstrate that the adaptive graph construction module dynamically captures sensor associations, (e.g., the synergistic response of accelerometers and gyroscopes) through mutual information, resulting in a 21% improvement in spatial feature extraction accuracy, while the gated fusion mechanism significantly optimises the responsiveness to high-speed abrupt motion (latency < 2 ms) by adjusting the weights of spatio-temporal pathway contributions. This finding provides a new paradigm for robot accuracy control: localisation errors can be viewed as spatio-temporal diffusion processes on a dynamical map, whose modelling needs to consider both topological constraints and evolutionary dynamics (Montgomery, 2024).

On a practical level, the ST-HGCN provides a drop-in solution for Industry 4.0 high-precision operations such as aero-engine blade assembly and semiconductor chip placement. Its compensation mechanism can be directly integrated into the robot controller to correct the end position through real-time feedback. However, real-world deployments still face a trio of challenges. Cross-mechanism generalisation bottleneck: current models are based on UAV motion training, whereas multi-degree-of-freedom industrial robotic arms suffer from stiff-flexible coupled vibrations (5–20 Hz) with joint friction hysteresis, resulting in differences in error propagation mechanisms (Bitter et al., 2023). A migration learning fine-tuned graph structure is suggested: the pre-training layer retains the spatio-temporal feature extraction capability, and the adaptive graph generation layer is re-trained on a small amount of data (< 1 hour of motion trajectory) from the target robot. Environmental perturbation sensitivity: electromagnetic interference (peak intensity 50 V/m) in the factory environment may degrade the IMU signal-to-noise ratio by up to 40% (measured ReFlex Takktile force sensor data), resulting in distortion of graph node features. Robust graph learning modules need to be introduced, e.g., generating perturbation samples (δ-IMU = $\mathbf{x}_t + \mathcal{N}(0, 0.2\sigma)$) through adversarial training, to improve the model noise immunity. Real-time system integration constraints: the real-time task cycle of existing industrial controllers is only 500 μs, which requires further compression of the model computation (Payton and Bihari, 1991). The inference elapsed time can be reduced to less than 800 μs by graph structure pruning

(removing edges with mutual information $I < 0.3$) and quantisation deployment (FP16 accuracy). Based on the above challenges, future work will focus on the following directions: a cross-domain spatio-temporal migration framework. Multi-physics field map fusion. Digital twin-driven lifelong learning.

5 Conclusions

In this paper, the effectiveness of ST-HGCN in industrial robot positioning error compensation is verified, with its theoretical value in establishing the spatio-temporal error diffusion equation on the dynamical map, and its engineering significance in real-time compensation at sub-millisecond level. It is recommended that this method be deployed preferentially in precision assembly scenarios, (e.g., aero-engine blade tongue-and-groove docking), where a 0.05 mm accuracy improvement can reduce rework costs by 23% (Boeing 737 blade assembly data). With the development of industrial meta-universe technology, spatio-temporal GNNs will become the core intelligent engine of high-precision manufacturing systems.

Acknowledgements

This work is supported by the school-level quality engineering project of Zhanjiang University of Science and Technology named: Virtual and Real, Specialised and Creative' Dual Integration Practical Teaching Base for Industrial Robots (No. ZCRHSJJD-20251035).

Declarations

All authors declare that they have no conflicts of interest.

References

Arents, J. and Greitans, M. (2022) 'Smart industrial robot control trends, challenges and opportunities within manufacturing', *Applied Sciences*, Vol. 12, No. 2, p.937.

Bai, L., Yao, L., Li, C., Wang, X. and Wang, C. (2020) 'Adaptive graph convolutional recurrent network for traffic forecasting', *Advances in Neural Information Processing Systems*, Vol. 33, pp.17804–17815.

Bitter, C., Peters, J., Tercan, H. and Meisen, T. (2023) 'Industrial cross-robot transfer learning', *The International Academy for Production Engineering*, Vol. 120, pp.1297–1302.

Burri, M., Nikolic, J., Gohl, P., Schneider, T., Rehder, J., Omari, S., Achtelik, M.W. and Siegwart, R. (2016) 'The EuRoC micro aerial vehicle datasets', *The International Journal of Robotics Research*, Vol. 35, No. 10, pp.1157–1163.

Cagigas-Muñiz, D. (2023) 'Artificial neural networks for inverse kinematics problem in articulated robots', *Engineering Applications of Artificial Intelligence*, Vol. 126, p.107175.

Chen, Y., Xie, F., Liu, X. and Zhou, Y. (2014) 'Error modeling and sensitivity analysis of a parallel robot with SCARA (selective compliance assembly robot arm) motions', *Chinese Journal of Mechanical Engineering*, Vol. 27, No. 4, pp.693–702.

Faria, R.R.A., Zuffo, M.K. and Zuffo, J.A. (2005) 'Improving spatial perception through sound field simulation in VR', *IEEE Symposium on Virtual Environments, Human-Computer Interfaces and Measurement Systems*, Vol. 1, p.6.

Foroughi, F., Chen, Z. and Wang, J. (2021) 'A CNN-based system for mobile robot navigation in indoor environments via visual localization with a small dataset', *World Electric Vehicle Journal*, Vol. 12, No. 3, p.134.

Granger, C. and Newbold, P. (2001) 'The time series approach to econometric model building', *Econometric Society Monographs*, Vol. 32, pp.302–316.

Jiang, Y., Dai, P., Fang, P., Zhong, R.Y. and Cao, X. (2022) 'Electrical-STGCN: an electrical spatio-temporal graph convolutional network for intelligent predictive maintenance', *IEEE Transactions on Industrial Informatics*, Vol. 18, No. 12, pp.8509–8518.

Kim, D.-H., Cook, K. and Oh, J.-H. (1991) 'Identification and compensation of a robot kinematic parameter for positioning accuracy improvement', *Robotica*, Vol. 9, No. 1, pp.99–105.

Li, H., Wang, X., Yang, Z., Ali, S., Tong, N. and Baseer, S. (2022) 'Correlation-based anomaly detection method for multi-sensor system', *Computational Intelligence and Neuroscience*, Vol. 2022, No. 1, p.4756480.

Li, J., Cheng, G. and Pang, Y. (2024) 'A novel deep learning-based spatio-temporal model for prediction of pose residual errors in optical processing hybrid robot', *IEEE Transactions on Industrial Informatics*, Vol. 20, No. 6, pp.8749–8762.

Li, Z., Li, S. and Luo, X. (2021) 'An overview of calibration technology of industrial robots', *IEEE/CAA Journal of Automatica Sinica*, Vol. 8, No. 1, pp.23–36.

Lin, H.-I., Mandal, R. and Wibowo, F.S. (2024) 'BN-LSTM-based energy consumption modeling approach for an industrial robot manipulator', *Robotics and Computer-Integrated Manufacturing*, Vol. 85, p.102629.

Mascaro, R. and Chli, M. (2024) 'Scene representations for robotic spatial perception', *Annual Review of Control, Robotics, and Autonomous Systems*, Vol. 8, pp.351–377.

Montgomery, R.M. (2024) 'Topological evolution of simple and complex organisms: a graph-based simulation study on adaptability and evolutionary constraints', *J. Biotech Bioinform*., Vol. 1, No. 3, pp.1–7.

Payton, D.W. and Bihari, T.E. (1991) 'Intelligent real-time control of robotic vehicles', *Communications of the Association for Computing Machinery*, Vol. 34, No. 8, pp.49–63.

Siciliano, B., Sciavicco, L., Villani, L. and Oriolo, G. (2009) 'Robotics: modelling, planning and control', *Springer Nature Link*, Vol. 1, pp.407–467.

Wang, M., Li, Y., Zhang, Y. and Jia, L. (2021) 'Spatio-temporal graph convolutional neural network for remaining useful life estimation of aircraft engines', *Aerospace Systems*, Vol. 4, No. 1, pp.29–36.

Yu, L., Du, B., Hu, X., Sun, L., Han, L. and Lv, W. (2021) 'Deep spatio-temporal graph convolutional network for traffic accident prediction', *Neurocomputing*, Vol. 423, pp.135–147.

Zhang, G., Xu, Z., Hou, Z., Yang, W., Liang, J., Yang, G., Wang, J., Wang, H. and Han, C. (2020) 'A systematic error compensation strategy based on an optimized recurrent neural network for collaborative robot dynamics', *Applied Sciences*, Vol. 10, No. 19, p.6743.

Zhang, H., Lu, G., Zhan, M. and Zhang, B. (2022) 'Semi-supervised classification of graph convolutional networks with Laplacian rank constraints', *Neural Processing Letters*, Vol. 54, No. 4, pp.2645–2656.

Zhang, Q., Yang, P. and Liu, Q. (2024) 'A dual-stream spatio-temporal fusion network with multi-sensor signals for remaining useful life prediction', *Journal of Manufacturing Systems*, Vol. 76, pp.43–58.

Zheng, C., Fan, X., Pan, S., Jin, H., Peng, Z., Wu, Z., Wang, C. and Yu, P.S. (2023) 'Spatio-temporal joint graph convolutional networks for traffic forecasting', *IEEE Transactions on Knowledge and Data Engineering*, Vol. 36, No. 1, pp.372–385.

Zhou, J., Zheng, L., Fan, W. and Cao, Y. (2025) 'Intelligent hierarchical compensation method for industrial robot positioning error based on compound branch neural network automatic creation', *Journal of Intelligent Manufacturing*, Vol. 36, No. 4, pp.2915–2938.

Zhou, Y., Xiao, J., Zhou, Y. and Loianno, G. (2022) 'Multi-robot collaborative perception with graph neural networks', *IEEE Robotics and Automation Letters*, Vol. 7, No. 2, pp.2289–2296.