SSL-GAT: A self-supervised learning-based graph attention network for agricultural machinery trajectory operation mode identification

Weixin Zhai[1†], Zekun Lyu[1†], Zhichao Li[1], Jiawen Pan[1,2,3,4*], Defang Li[5], Deming Sun[5], Usmon Mamur Mahmadyorzoda[6], Nozim Numon Alizoda[6]

(1. *College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China*;
2. *College of Computer and Information Sciences, Fujian Agriculture and Forestry University, Fuzhou 350002, China*;
3. *Center for Artificial Intelligence in Agriculture, Fujian Agriculture and Forestry University, Fuzhou 350002, China*;
4. *Department of Geography, Ghent University, Krijgslaan 281 S8, 9000, Ghent, Belgium*;
5. *Weichai Lovol Intelligent Agricultural Technology CO., LTD, Weifang 261206, China*;
6. *Tajik Agrarian University named after Shirinsho Shotemur, Dushanbe 734003, Republic of Tajikistan*)

Abstract: In the agricultural domain, agricultural machinery trajectory operation mode identification is essential for spatiotemporal trajectory processing. Its goal is to classify machinery trajectories into road travel or field operations by extracting latent spatiotemporal features. However, conventional models lack effective feature enhancement and ignore the varying importance of trajectory points, thereby weakening feature representation and reducing identification accuracy. To overcome these challenges, a self-supervised learning-based GAT is proposed for identifying agricultural machinery trajectory operation modes. First, to enhance trajectory feature representation, a multi-dimensional feature enhancement module is introduced based on statistical methods. To mitigate the impact of redundant features on model performance, a bidirectional feature fusion module is subsequently proposed that captures both interpoint and intrapoint dependencies, thereby enhancing spatiotemporal representations and suppressing irrelevant information. Next, to capture the importance of trajectory points, a graph attention network (GAT) with a masked attention mechanism is introduced. Finally, to reduce the dependence on labeled data and improve the model's feature learning capability, self-supervised learning is used as a pretraining step for the GAT. To evaluate SSL-GAT, experiments are conducted on two real-world paddy and wheat harvester datasets. For the paddy dataset, SSL-GAT reaches 95.92% accuracy and a 92.42% F1-score, exceeding those of GAN-BiLSTM by 4.67% and 5.24%, respectively. On the wheat dataset, it achieves 93.92% accuracy and a 90.72% F1-score, with gains of 5.58% and 4.49% over GAN-BiLSTM. These results collectively demonstrate that our SSL-GAT model achieves superior performance, establishing it as a new state-of-the-art model in agricultural machinery trajectory operation mode identification.
Keywords: agricultural machinery trajectory operation mode identification, multi-dimensional feature enhancement module, bidirectional feature fusion module, attention mechanism, self-supervised learning
DOI: 10.25165/j.ijabe.20261902.9966

Citation: Zhai W X, Lyu Z, Li Z C, Pan J W, Li D F, Sun D M, et al. SSL-GAT: A self-supervised learning-based graph attention network for agricultural machinery trajectory operation mode identification. Int J Agric & Biol Eng, 2026; 19(2): 282–293.

1 Introduction

The identification of the agricultural machinery trajectory operation mode is an essential task in trajectory data analysis. It extracts spatiotemporal features from GNSS/Beidou trajectories to distinguish field points from road points[1-4]. As shown in Figure 1, the goal is to classify raw trajectory points into the two operation modes. Since agricultural machinery trajectories are real-time spatiotemporal sequences collected by onboard terminals[5-9], accurate mode identification supports multiple downstream applications, including field area estimation[10], agricultural route planning[11-13], cross-field operation detection[14,15], and operational efficiency evaluation[16,17]. It also benefits large-scale smart agriculture analysis, particularly in terms of agricultural machinery management and cost analysis[18-20], as well as AI- and big data-driven precision agriculture applications[21-24]. It further provides reliable support for constructing rural road networks that are often underrepresented in mainstream mapping platforms[25-28].

Existing studies mainly use traditional machine learning or deep learning. Decision trees[29,30] rely on simple motion features and overlook richer spatiotemporal patterns. DBSCAN[31] requires careful hyperparameter tuning and is sensitive to noise. GCN-based methods[32] use graph convolution to aggregate neighbor information but apply fixed weights and rely on Laplacian normalization, which may weaken discrimination ability. DR-XGBoost[33] improves

Received date: 2025-06-16 **Accepted date:** 2026-01-03
Biographies: Weixin Zhai, PhD, Associate Professor, research interest: big data of spatio temporal, big data mining of agricultural machinery, cartography and geographic information system, Email: zhaiweixin@cau.edu.cn; **Zekun Lyu**, Master, research interest: artifical intelligence and generative AI, Email: zklv030@163.com; **Zhichao Li**, Master, research interest: big data mining of agricultural machinery, Email: 15610535583@163.com; **Defang Li**, Bachelor, Senior Engineer, Email: lidefang@lovol.com; **Deming Sun**, Bachelor, Senior Engineer, Email: sundeming@lovol.com; **Usmon Mamur Mahmadyorzoda**, PhD, Professor, research interest: precision agriculture, Email: rectortau31@mail.ru; **Nozim Numon Alizoda**, PhD, Professor, research interest: precision agriculture, Email: nnozim@gmail.com.
†The authors contributed equally to this work
***Corresponding author: Jiawen Pan**, Associate Professor, research interest: data mining, computer vision. College of Computer and Information Sciences, Fujian Agriculture and Forestry University, Fuzhou 350002, China. Tel: +86-15588880367, Email: panjiawen@fafu.edu.cn.

feature selection but depends on extensive grid search. GAN-BiLSTM[34] captures bidirectional temporal dependencies yet fails to model intra-point feature correlations.

Overall, existing methods have two main limitations. First, they do not fully capture spatiotemporal dependencies among trajectory points. Critical points such as turns or stops strongly indicate operational mode, but many models treat all points equally. Second, intra-point feature interactions are often ignored, even though combinations of timestamp, location, speed, and direction provide more reliable cues—for example, high speed with stable direction often indicates road travel, whereas lower speed with frequent direction changes indicates field operations. Real-world GNSS data further introduces noise from multipath effects, atmospheric delays, satellite clock bias, and machinery vibration, increasing the difficulty of distinguishing field and road patterns and placing higher demands on model robustness.

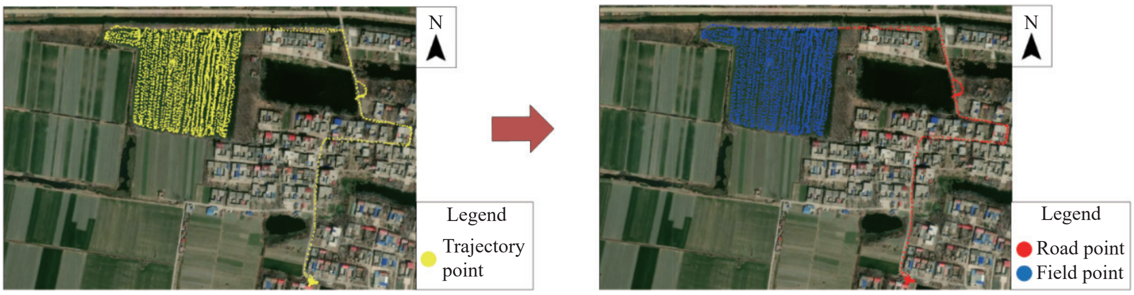

Figure 1 Process of agricultural machinery trajectory operation mode identification

To address these challenges, a self-supervised learning-based graph attention network (SSL-GAT) is proposed for trajectory operation mode identification. First, a multi-dimensional feature enhancement module extracts instant features that describe the machinery state at a specific moment and temporal features that summarize motion trends over time using statistical and sliding-window techniques. Second, a bidirectional feature fusion module performs inter-point and intra-point fusion to reduce redundancy and highlight informative features. Third, a graph attention network is constructed using temporal edges, spatial edges, and self-loops, and masked attention is used to adaptively weight neighbor information, reducing the influence of noisy points. Finally, a masked autoencoder-based self-supervised module reconstructs partially masked trajectory features using neighbor context, enabling the model to learn discriminative representations without relying heavily on labeled data. The main contributions of this study are:

1) A multi-dimensional feature enhancement module that captures both instant and temporal spatiotemporal patterns in agricultural machinery trajectories.

2) A bidirectional feature fusion module integrating inter-point and intra-point information to refine representations and suppress redundancy.

3) A graph attention mechanism that adaptively adjusts neighbor weights, improving the robustness and accuracy of spatiotemporal feature extraction.

4) A self-supervised learning module enabling automatic learning of discriminative representations from raw trajectories with reduced labeling effort.

5) Comprehensive experiments on a real-world joint grain harvester dataset from the Key Laboratory of Agricultural Machinery Monitoring and Big Data Application, Ministry of Agriculture and Rural Affairs, showing that SSL-GAT achieves competitive performance compared with state-of-the-art models.

2 Materials and methods

2.1 Dataset

200 real agricultural machinery trajectories were selected, comprising a total of 904 839 points, provided by the Key Laboratory of Agricultural Machinery Monitoring and Big Data Applications, Ministry of Agriculture and Rural Affairs, China. During data collection, the Beidou/GNSS terminal installed on agricultural machinery accurately records real-time trajectory data. This data was transmitted via mobile communication networks to IoT platforms, which then immediately sent it to the key laboratory's Beidou-based big data platform for agricultural machinery trajectories. These trajectories were collected in June and September 2021, covering six provincial-level administrative regions in China: Shanxi, Shaanxi, Hebei, Xinjiang, Shandong, and Jiangsu. To evaluate model performance, two types of trajectories were used as the experimental dataset: wheat harvester trajectories and paddy harvester trajectories. These two types differ significantly in the number of trajectories, operation duration, spatial distribution, and sampling frequency. Each trajectory point of the two types includes five attributes: direction (°), timestamp (YYYY/MM/DD hh:mm:ss), velocity (m/s), geospatial coordinates (longitude and latitude based on the WGS84 coordinate system), and a semantic label obtained through manual annotation. Due to the use of two different GNSS receivers, HOMER3S and Model 781 100, the paddy and wheat trajectories were collected at different sampling intervals of 5 s and 30 s. Detailed information is listed in Table 1.

Table 1 Detailed information for two types of trajectory

Parameters	Paddy harvester trajectory dataset	Wheat harvester trajectory dataset
No. of covered provinces	5	4
No. of trajectories	100	100
Total number of points	109 389	795 450
Acquisition time	2021.9-2021.10	2021.6
Acquisition interval	30 s	5 s

2.2 Overview

To comprehensively extract the latent spatiotemporal information, a self-supervised graph attention network was proposed for agricultural machinery trajectory operation mode identification. The pipeline of our model is shown in Figure 2 and consists of the following key components:

Trajectory points preprocessing: The preprocessing method consists of removing drift, overlapping, and stationary trajectory points (Section 2.3).

Multi-dimensional feature enhancement: This feature enhancement method involves two key components: instant feature and temporal feature. Instant feature can capture instantaneous features of agricultural machinery's motion, and temporal feature can capture the movement trends of agricultural machinery over a period of time (Section 2.4).

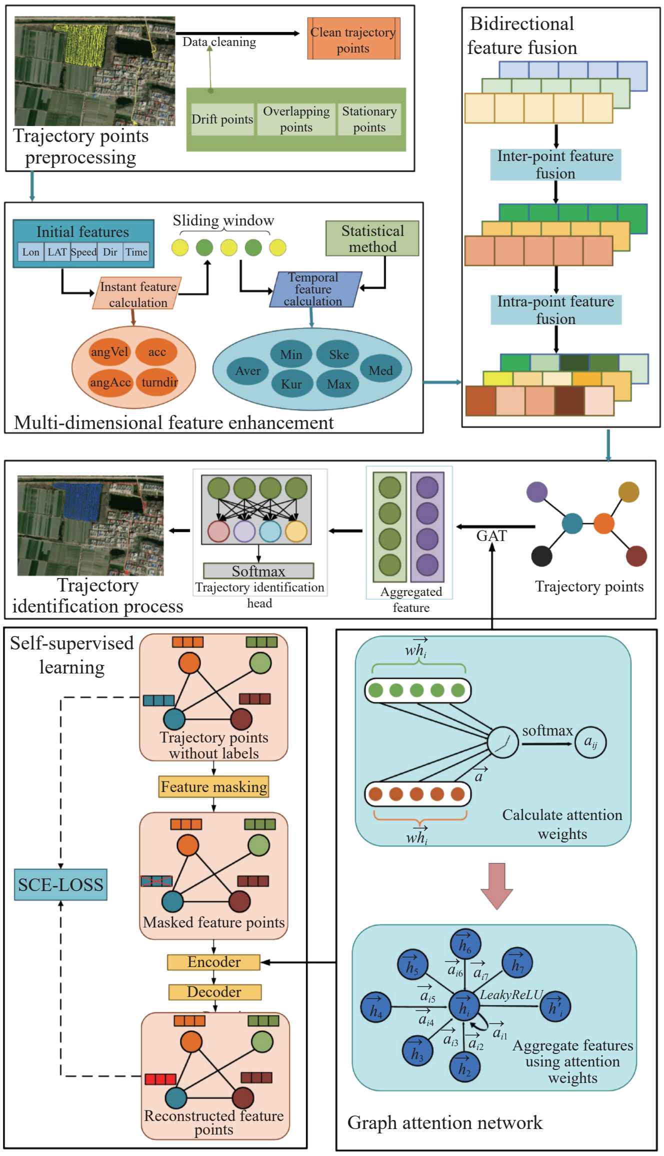

Figure 2 Pipeline of SSL-GAT model

Bidirectional feature fusion: This method captures dependencies between trajectory points and refines feature information through linear transformations across both inter-point and intra-point dimensions (Section 2.5).

Graph attention network: This network captures the relationships between trajectory points by using an attention mechanism to weigh the significance of neighboring nodes in the graph. The attention mechanism further helps the model focus on more important trajectory points, thereby reducing noise and improving prediction accuracy (Section 2.6).

Trajectory identification head: A feed-forward network was introduced that assigns semantic labels to individual trajectory points by leveraging learned feature representations (Section 2.7).

Self-supervised learning: The self-supervised learning method was used to utilize raw data without requiring additional manual annotations by reconstructing masked trajectory features, which helps in learning more robust representations under noisy and complex trajectory points (Section 2.8).

2.3 Trajectory points preprocessing

To ensure the accuracy of trajectory points, three key preprocessing steps were performed: handling drifting points, removing halt points, and eliminating duplicate points. First, drifting points caused by abnormal speeds or invalid sampling intervals were filtered by setting reasonable speed thresholds (2-30 m/s) and discarding points with overly long time gaps. Second, halt points, typically with zero velocity, were removed to avoid noise caused by GNSS drift during machine stops. Third, duplicate points resulting from measurement errors or signal fluctuations were eliminated by retaining only the first occurrence.

2.4 Multi-dimensional feature enhancement

To effectively capture the latent spatiotemporal information of agricultural machinery trajectories, a multi-dimensional feature enhancement is proposed, which consists of two key components: the instant feature and the temporal feature. Instant features describe point-wise motion states, whereas temporal features summarize movement trends within a sliding window.

2.4.1 Instant feature calculation

A trajectory $R=\{P_1, P_2, P_3, \ldots, P_n\}$ was defined as a sequence of trajectory points, where P_i represents the i-th point and n is the total number of points in the trajectory. Each point P_i is represented by a vector of five initial features: $P_i=\{\text{longitude}_i, \text{latitude}_i, \text{speed}_i, \text{direction}_i, \text{time}_i\}$, where longitude$_i$ and latitude$_i$ correspond to the geographic coordinates of the i-th point, and speed$_i$, directions$_i$, and time$_i$ denote the speed, direction, and timestamp, respectively. Based on the trajectory, several instant features were derived by applying kinematic formulas to compute the speed, angle difference, acceleration, angular velocity, and angular acceleration at each trajectory point. The speed was directly obtained from the original trajectory points, while the angle difference was obtained by calculating the change in direction between two consecutive points. The acceleration was computed by analyzing the change in speed between consecutive points over the time interval. The angular velocity was computed by measuring the change in direction between adjacent points, divided by the time difference. The angular acceleration was computed by the rate of change of angular velocity over the time interval. The formulas for these calculations are presented in Equations (1)-(4).

$$td_i = |d_{i+1} - d_i| \tag{1}$$

$$\text{acc}_i = \frac{(s_{i+1} - s_i)}{(t_{i+1} - t_i)} \tag{2}$$

$$\text{angVel}_i = \frac{(tr_{i+1} - tr_i)}{(t_{i+1} - t_i)} \tag{3}$$

$$\text{angAcc}_i = \frac{(\text{angVel}_{i+1} - \text{angVel}_i)}{(t_{i+1} - t_i)} \tag{4}$$

where, t_i, d_i, td_i, s_i, acc$_i$, angVel$_i$, and angAcc$_i$ represent the timestamp (YYYY/MM/DD hh:mm:ss), direction (°), turning angle (°), speed (m/s), acceleration (m/s^2), angular velocity [(°)/s], and angular acceleration [(°)/s^2] of p_i trajectory point, respectively.

2.4.2 Temporal feature calculation

Temporal features capture the spatiotemporal information of clusters of trajectory points that are geographically close. In this study, statistical methods and a sliding window approach were used to extract the temporal features of each trajectory point based on its neighboring points. Specifically, a sliding window moves through the set of trajectory points. For each point, the window covers a fixed number of adjacent points before and after it, based on timestamps, and computes statistical metrics such as the mean, skewness, and kurtosis of the points within the window as temporal features. Skewness and kurtosis are statistical measures that describe the shape of the probability distribution within the window. Skewness reflects the symmetry of the distribution, while kurtosis reflects the peakedness of the distribution. The formulas for calculating skewness and kurtosis are shown in Equations (5) and (6).

$$\mathrm{sk}_i = \frac{1}{n} \times \frac{\sum_{i=1}^{n} (x_i - \bar{x})^3}{s^3} \tag{5}$$

$$\mathrm{ku}_i = \frac{1}{n} \times \frac{\sum_{i=1}^{n} (x_i - \bar{x})^4}{s^4} \tag{6}$$

where, sk_i is the skewness coefficient of the P_i trajectory point, and ku_i is the kurtosis coefficient of the P_i trajectory point. n denotes the size of the sliding window, and x_i represents the value of the P_i point in the window. $\bar{x}$ represents the mean of all points within the window, and s represents their standard deviation. The complete process of multi-dimensional feature enhancement is shown in Figure 3.

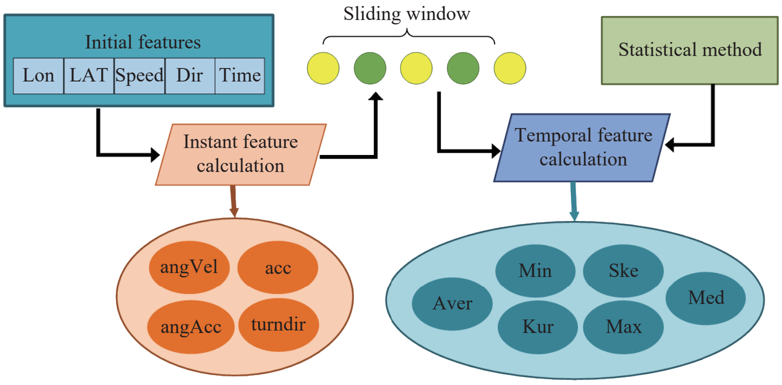

Figure 3 Complete process of multi-dimensional feature enhancement

2.5 Bidirectional feature fusion module

To enhance the model's ability to represent input features and reduce redundancy, a bidirectional feature fusion module was proposed that applies two sequential linear transformations to refine feature representations derived from trajectory points.

The first linear transformation performs inter-point feature fusion, while the second performs intra-point feature fusion[35]. These two fusion strategies jointly capture inter-point dependencies and refine intra-point feature representations. Specifically, given an input trajectory $X \in R^{P \times F}$, where P is the number of points, and F is the number of features within each point, the bidirectional feature fusion module first applies a linear transformation along the inter-point feature dimension:

$$X' = W_1 \cdot NormLayer(X) \tag{7}$$

where, X' denotes the intermediate representation after inter-point feature fusion. W_1 is a learnable weight matrix, and NormLayer is a normalization method along the last dimension of X. Then, another linear transformation is applied along the intra-point feature dimension:

$$X'' = W_2 \cdot \mathrm{NormLayer}(X') \tag{8}$$

where, X'' represents the final refined representation after intra-point feature fusion. W_2 assigns weights to different features within each point to reduce the influence of redundant features. The inter- and intra-point projection refines representations by capturing dependencies and suppressing redundant dimensions.

2.6 Graph attention network

In the agricultural machinery trajectory operation mode identification task, effectively capturing the spatiotemporal dependencies and varying importance between trajectory points is crucial for improving model performance. GCN proposed by Chen et al.[32] demonstrated the effectiveness of Graph Neural Networks (GNNs) in the trajectory operation mode identification task, showing their ability to learn the spatiotemporal dependencies of trajectories. However, since key trajectory points—such as junctions and turning points—carry more discriminative information for identification, the inability of GCN to capture their varying importance ultimately limits the model's performance. To overcome the limitation, a graph attention network was proposed for agricultural machinery trajectory operation mode identification.

2.6.1 Spatiotemporal graph construction

Trajectory points are typically stored in an unstructured manner without explicit relationships between them, which hinders the model's ability to exploit potential spatiotemporal dependencies. To address this problem, a spatiotemporal graph was constructed that maps trajectory points into a graph defined by spatiotemporal relationships. Additionally, the graph serves as an essential structure for extracting, aggregating, and propagating features within the graph attention network.

Graph $G = (V, E)$ was defined in this study, where V represents the set of GNSS trajectory points, and E represents the edges that encode spatiotemporal relationships between points. Two points are considered connected if a spatiotemporal relationship exists between them. To ensure that the graph model can effectively capture spatiotemporal information, three types of spatiotemporal edges are designed:

Self-loop edges: Each GNSS point is connected to itself via a self-loop edge. These edges are essential for enhancing feature learning and maintaining model stability, ensuring that the original features of each GNSS point are preserved during the message passing process in the GAT.

Temporal neighbor edges: Temporal neighbor edges connect a point to its four nearest neighbors based on time intervals. As illustrated in Figure 4, a field point p_i is connected to four points a_i, b_i, c_i, and d_i, where a_i is two time intervals after p_i, b_i is one time interval after p_i, c_i is one time interval before p_i, and d_i is two time intervals before p_i. These edges help capture temporal correlations in the trajectory points by linking trajectory points that are temporally close.

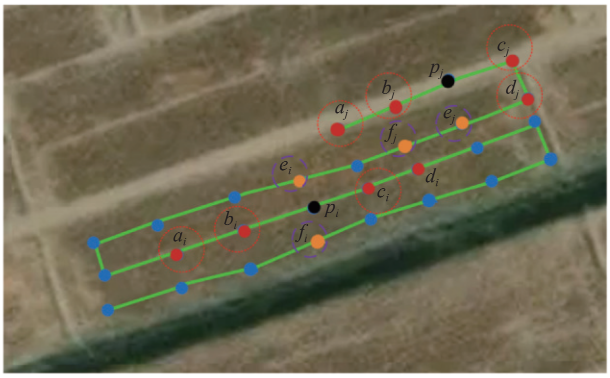

Figure 4 Spatiotemporal relationships between trajectory points

Spatial neighbor edges: Spatial neighbor edges connect each point to its two closest points in space (excluding those already connected by temporal neighbor edges). As shown in Figure 4, p_i is connected to its nearest spatial neighbors e_i and f_i. Through these spatial connections, the model captures the spatial information of the trajectory points. These three types of spatiotemporal edges precisely describe the temporal and spatial relationships between trajectory points. This design allows the model to aggregate information from neighboring points based on their spatial and temporal relevance, improving the accuracy of feature representation.

2.6.2 Graph attention network

To enhance the accuracy and flexibility of trajectory operation mode identification for agricultural machinery, a Graph Attention Network[36] model was proposed that builds upon the spatial-temporal graph structure introduced in Section 2.6.1. Unlike traditional graph-based approaches with manually defined connectivity, our model incorporates a learnable attention mechanism[36,37] to dynamically infer and utilize the underlying relationships among trajectory points. The visual depiction of the graph attention network is shown in Figure 5.

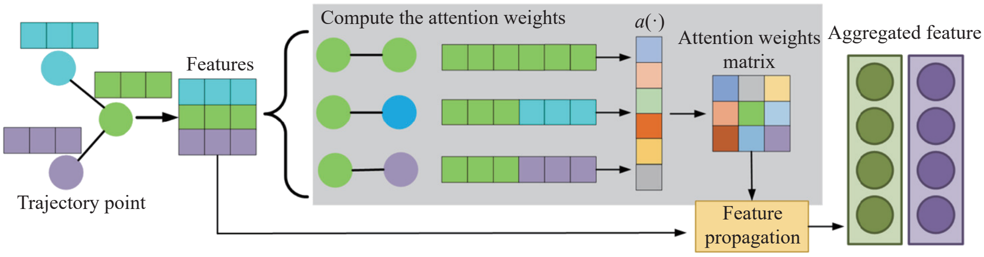

Figure 5 Illustration of the graph attention network

GAT employs attention-based message passing to assign distinct weights to neighbors and to focus on informative points. Specifically, we define a learnable attention matrix A, where $a_{ij} \in A$ quantifies the influence of node j on node i. The preliminary attention score e_{ij} between node i and its neighbor j is computed as:

$$e_{ij} = a\left(Wh_i \parallel Wh_j\right), \quad j \in N_i \tag{9}$$

where, h_i and h_j represent the input feature vectors of nodes i and j respectively, W is a shared learnable weight matrix, $\parallel$ denotes vector concatenation, and $a(\cdot)$ is a single-layer feedforward neural network that outputs a scalar attention score.

To ensure comparability among neighbors, the attention scores are normalized using the softmax function:

$$a_{ij} = \frac{\exp\left(\mathrm{LeakyReLU}(e_{ij})\right)}{\sum_{k \in N_i} \exp\left(\mathrm{LeakyReLU}(e_{ik})\right)}, \quad j \in N_i \tag{10}$$

where, a_{ij} represents the final attention weights from node j to node i. The updated feature representation of node i is then calculated as a weighted sum of the transformed features of its neighbors:

$$h_i' = \mathrm{LeakyReLU}\left(\sum_{j \in N_i} a_{ij} W h_j\right) \tag{11}$$

where, h_j denotes the input feature vector of node j, and h_i' represents the updated feature vector of node i after aggregating information from its neighbors using a_{ij} and a shared learnable weight matrix W.

This attention-based aggregation allows the model to dynamically focus on the most informative neighbors while preserving the local structural information of each node.

The graph attention network employs a dynamic adjacency matrix generation mechanism based on pairwise feature similarity, replacing traditional fixed or manually defined graph structures. Specifically, the influence of edges is determined by the attention weights a_{ij}, which are learned during training. This adaptive construction enables the graph topology to reflect complex spatial and temporal dependencies embedded in the trajectory. Consequently, the model gains greater flexibility and expressiveness.

In addition, to enhance computational efficiency and model scalability, two optimization strategies are implemented. First, sparse attention was applied where only relevant neighboring nodes with significant attention weights participate in aggregation, thereby reducing computational overhead. Second, a weight-sharing strategy was employed that assumes symmetric attention coefficients (i.e., e_{ij}=e_{ji}) in certain scenarios, which simplifies parameter learning and improves training stability.

2.7 Trajectory identification head

In this study, each node was encoded as a feature representation, which is fed into a fully connected layer followed by a softmax function to obtain the predicted probability result over the two classes, as Equation (12).

$$Y_i = \mathrm{softmax}(h_i' \omega + b) \tag{12}$$

where, ω and b are learnable parameters of the network, and h_i' is the feature representation for node i. The class with the highest probability in Y_i is selected as the prediction result for node i. Repeating this process for all nodes yields the final identification of the agricultural machinery trajectory operation modes.

2.8 Self-supervised learning

To enhance the model's capability in learning expressive and discriminative feature representations, this study proposes a self-supervised learning[38,39] that does not rely on manually annotated labels. The visual depiction of the self-supervised learning is shown in Figure 6. The core idea is to train the model to reconstruct the masked trajectory features, thereby capturing the underlying structure and dependencies within the points. The proposed method comprises two major components:

1) Feature masking: A subset of trajectory point features is randomly selected and replaced with zero vectors, simulating missing data and encouraging the model to learn dependencies between points.

2) Graph-based feature reconstruction: A graph encoder-decoder architecture is employed to recover the masked features based on the relational structure of the trajectory. Specifically, the encoder utilizes graph attention network in Section 2.6.2 to project the masked input features X^* into a latent representation H, formulated as:

$$H = f(G, X^*) \tag{13}$$

where $G=(V,E)$ denotes the graph structure with nodes E representing trajectory points and edges V encoding spatial or temporal relationships; f denotes graph attention network that encodes the input features X^* into latent representations H. Subsequently, the latent feature representations H are passed into a standard transformer decoder module, which is directly adopted without any architectural modifications, formulated as:

$$143\tilde{X} = P(H) \tag{14}$$

The decoder comprises multiple layers of multi-head self-attention and position-wise feed-forward networks, consistent with the original transformer design. Positional encodings are added to the input sequence to preserve the spatial or temporal dependencies among trajectory points. The output of the decoder $\tilde{X}$, which has the same dimensionality as the original feature matrix X, allows for direct feature-level reconstruction. Layer normalization and residual connections are incorporated following the standard configuration to enhance model stability and learning efficiency.

To optimize the reconstruction, a self-supervised contrastive embedding loss (SCE-Loss) is introduced, defined as:

$$146L_{\text{sce}} = \sum_{i \in M} \left(1 - \left\langle \tilde{X}_i, X_i \right\rangle\right)^{\alpha} \tag{15}$$

where, $\tilde{X}_i$ and X_i represent the normalized reconstructed and ground-truth feature vectors for the masked nodes i, and α is a tunable hyperparameter controlling the sensitivity of the loss.

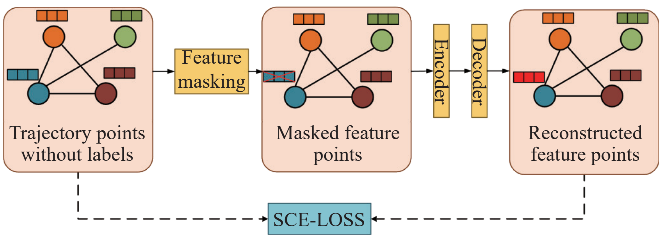

Figure 6 Illustration of the self-supervised learning

After the self-supervised pretraining phase, the learned representations can be effectively fine-tuned within the GAT framework for downstream tasks such as field-road identification. This method not only enhances the robustness of trajectory feature learning but also improves model generalization under conditions of noise and trajectory incompleteness, while significantly reducing dependence on labeled points.

2.9 Performance evaluation metrics

To evaluate the performance of our model, this study employed Precision, Recall, F1-score, and Accuracy as assessment metrics. Specifically, Precision (Equation (16)) quantifies the proportion of correctly identified positive trajectory points among points predicted as positive. Recall (Equation (17)) measures the proportion of correctly identified positive trajectory points relative to the total number of actual positive points. The F1-score (Equation (18)), defined as the harmonic mean of precision and recall, provides a balanced evaluation of these two metrics. Accuracy (Equation (19)) represents the overall proportion of correctly identified trajectory points across the entire dataset. The formulas are shown in Equations (16)-(19). These metrics were computed based on true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). In the context of agricultural machinery trajectory operation mode identification tasks, considering the “field” class as positive and the “road” class as negative, TP denotes the number of trajectory points correctly classified as “field”, TN denotes the number of trajectory points correctly classified as “road”, FP denotes the number of “road” points misclassified as “field”, and FN denotes the number of “field” points misclassified as “road”. Conversely, when the “road” class is treated as positive and the “field” class as negative, the definitions of TP, TN, FP, and FN are reversed accordingly. To comprehensively assess the model’s performance, Precision, Recall, and F1-score were first computed separately for both classes (“field” and “road”) and subsequently averaged to obtain the mean Precision, mean Recall, and mean F1-score. Among these metrics, Accuracy and mean F1-score serve as overarching indicators of overall model performance.

$$\text{Precision} = \frac{\text{TP}}{\text{TP}+\text{FP}} \tag{16}$$

$$\text{Recall} = \frac{\text{TP}}{\text{Tp}+\text{FN}} \tag{17}$$

$$\text{F1}-\text{score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{18}$$

$$\text{Accuracy} = \frac{\text{TP}+\text{TN}}{\text{Tp}+\text{TN}+\text{FP}+\text{FN}} \tag{19}$$

3 Results and discussion

3.1 Experimental settings

Baseline models: To evaluate the effectiveness of our proposed method, several baselines were selected for the agricultural machinery trajectory operation mode identification task. GAN-BiLSTM[34] represents the current state-of-the-art model for this task. In addition, DirDist-DBSCAN[31], Decision Tree[30], Random Forest[29], Graph Convolutional Network[32], and DR-XGBoost[33] were introduced as comparative baseline methods.

Model implementation: All models were implemented using Python 3.9 and PyTorch 2.6.0. A total of 200 trajectory samples were divided into training and testing sets with a ratio of 8:2. All baseline models were configured according to the settings reported in their original papers. For the proposed SSL-GAT model, two graph attention layers were used with feature dimensions of 43 and 108, respectively. During training, the learning rate was set to 0.001. The model adopts the Adam optimizer[40] with a weight decay of 0.0005.

3.2 Method comparison and discussion

All models are trained on the entire dataset and compared the performance of SSL-GAT with the six baseline methods. Tables 2 and 3 show the performance results on the paddy harvester and wheat harvester trajectory datasets, respectively.

Table 2 Performance comparison of the seven methods on the paddy harvester trajectory dataset

Method	Accuracy/ %	Precision/ %	Recall/ %	F1-score/%	Time/ s	Parameter
DirDist-DBSCAN	79.40	70.41	76.26	73.19	20.3	-
DT	76.42	68.13	74.02	70.94	5.27	-
RF	80.24	77.40	78.03	77.66	4.51	-
GCN	85.57	78.32	83.86	80.99	577.00	0.1601 M
DR-XGBoost	87.58	83.90	84.26	84.01	15.4	-
GAN-BiLSTM	91.25	84.11	90.60	87.18	792.00	1.188 M
SSL-GAT	95.92	90.47	94.58	92.42	451.00	1.270 M

Table 3 Performance comparison of the seven methods on the wheat harvester trajectory dataset

Method	Accuracy/ %	Precision/ %	Recall/ %	F1-score/%	Time/ s	Parameter
DirDist-DBSCAN	72.45	66.42	71.38	68.77	35.30	-
DT	77.36	68.12	75.23	71.49	13.20	-
RF	76.89	72.75	74.24	73.44	11.90	-
GCN	81.74	77.61	78.08	77.83	914.00	0.1601M
DR-XGBoost	83.47	79.60	80.73	80.13	42.70	-
GAN-BiLSTM	88.34	84.42	88.17	86.23	1329.00	1.188M
SSL-GAT	93.92	88.06	93.58	90.72	802.00	1.270M

On the paddy harvester trajectory dataset (Table 2), SSL-GAT achieved the highest F1-score of 92.42%, significantly outperforming all other methods. Compared to the second-best model, GAN-BiLSTM, SSL-GAT improved the F1-score by 5.24%. Table 3 presents the results on the wheat harvester trajectory dataset, where SSL-GAT again achieved the best performance with an F1-score of 90.72%. It surpassed GAN-BiLSTM by 4.49%, and exceeded DR-XGBoost, GCN, RF, DT, and DirDist-DBSCAN by 10.59%, 12.89%, 17.28%, 19.23%, and 21.95%, respectively. These results confirm the superior robustness and accuracy of SSL-GAT. Tables 4 and 5 list the performance of all seven models on recognizing the "field" and "road" points in the paddy and wheat datasets. On the paddy harvester trajectory dataset, SSL-GAT achieved an F1-score of 94.91% for the "field" class, improving upon DirDist-DBSCAN, DT, RF, GCN, DR-XGBoost, and GAN-BiLSTM by 18.37%, 13.29%, 8.09%, 6.48%, 4.09%, and 3.40%, respectively. For the "road" class, it achieved an F1-score of 89.93%, outperforming the same six methods by 20.08%, 29.68%, 21.44%, 16.37%, 12.76%, and 7.08%. On the wheat harvester trajectory dataset, SSL-GAT achieved F1-scores of 93.56% for "field" and 87.88% for "road", exceeding the six comparative methods by 19.00% and 24.90%, 12.72% and 25.73%, 13.50% and 21.07%, 8.93% and 16.84%, 7.02% and 14.16%, and 3.92% and 5.06%, respectively.

Table 4 Performance comparison of the seven methods on the "field" and "road" classes on the paddy harvester trajectory dataset

Method	Field			Road		
	Precision/ %	Recall/ %	F1-score/%	Precision/ %	Recall/ %	F1-score/%
DirDist-DBSCAN	72.37	81.22	76.54	68.45	71.31	69.85
DT	79.24	84.16	81.62	57.02	63.89	60.25
RF	88.60	85.12	86.82	66.21	70.95	68.49
GCN	85.74	91.31	88.43	70.91	76.42	73.56
DR-XGBoost	93.31	88.47	90.82	74.50	80.05	77.17
GAN-BiLSTM	90.28	92.79	91.51	77.94	88.42	82.84
SSL-GAT	95.01	94.82	94.91	85.93	94.34	89.93

Table 5 Performance comparison of the seven methods on the "field" and "road" classes on the wheat harvester trajectory dataset

Method	Field			Road		
	Precision/ %	Recall/ %	F1-score/%	Precision/ %	Recall/ %	F1-score/%
DirDist-DBSCAN	70.48	79.15	74.56	62.37	63.62	62.98
DT	77.34	84.69	80.84	58.91	65.77	62.15
RF	81.12	79.04	80.06	64.39	69.44	66.81
GCN	85.24	84.03	84.63	69.98	72.14	71.04
DR-XGBoost	87.45	85.66	86.54	71.76	75.8	73.72
GAN-BiLSTM	89.01	90.28	89.64	79.83	86.06	82.82
SSL-GAT	91.88	95.31	93.56	84.25	91.85	87.88

Figures 7 and 8 show the prediction results of all seven models, alongside the ground truth. Red dots represent points classified as "field", while green dots represent those classified as "road". As shown in the figures, SSL-GAT produced results closest to the ground truth, demonstrating the most accurate trajectory recognition.

Across both the paddy and wheat harvester trajectory datasets, SSL-GAT needed more training time than the traditional machine learning models did, but it was still much faster than the other deep learning methods. SSL-GAT took 451 s and 802 s on the two datasets. In comparison, the traditional models, DT, RF, and DR-XGBoost, finished training within only a few to several dozen seconds. This difference is reasonable because SSL-GAT uses graph attention operations, which naturally require more computations than simple tree-based models do.

GCN and GAN-BiLSTM performed better than the traditional models in terms of recognition ability, but their training times were much longer, 577 s/914 s for GCN and 792 s/1329 s for GAN-BiLSTM. In contrast, SSL-GAT not only achieved higher recognition performance than both GCN and GAN-BiLSTM but also required noticeably less training time. This shows that SSL-GAT has a clear advantage in both accuracy and efficiency within deep learning methods.

In terms of model complexity, SSL-GAT contains 1.270M parameters, which is substantially larger than those of GCN (0.1601M) and GAN-BiLSTM (1.188M). Because traditional machine-learning models (e.g., DT, RF, and DR-XGBoost) differ fundamentally from deep learning models in structure, their parameter counts are not directly comparable and are therefore not included. The increased parameter count of SSL-GAT mainly comes from the multi-head graph attention layers, which introduce more learnable weight matrices than standard graph convolutions or recurrent units do. Although SSL-GAT has the highest number of parameters among the deep-learning baselines, it still achieves a shorter training time than GCN and GAN-BiLSTM, indicating that the proposed architecture maintains good computational efficiency despite its larger model capacity.

Overall, even though SSL-GAT takes more time than the traditional models do, its great improvement in recognition ability fully justifies the extra training time. Considering both accuracy and efficiency, SSL-GAT provides the best overall balance and is the most practical and reliable choice among all the methods.

The reasons for SSL-GAT's superior performance are as follows:

DirDist-DBSCAN is a density-based clustering algorithm that only uses spatial coordinates, failing to capture spatiotemporal patterns. Its reliance on spatial proximity often causes

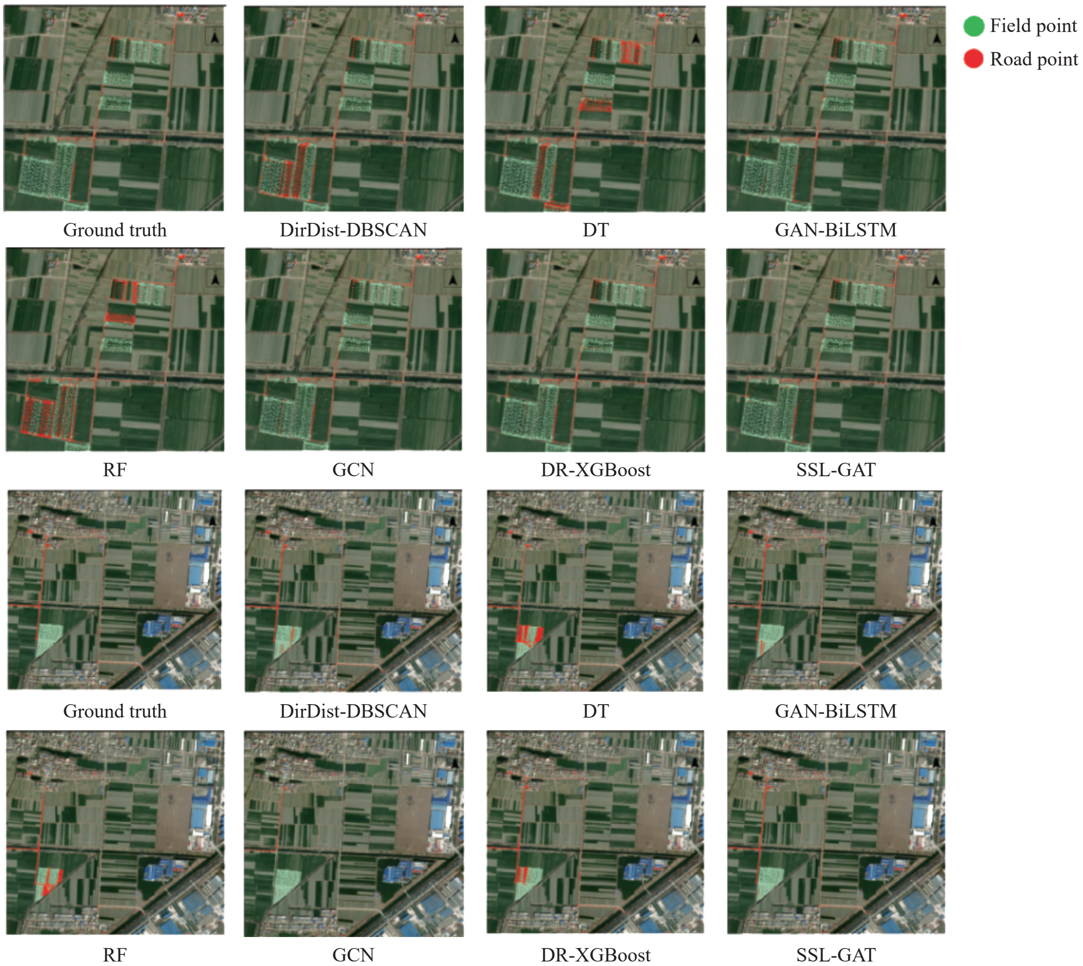

Figure 7 Identification results for two scenarios on the wheat harvester dataset

misclassification at field-road boundaries.

Decision Tree and Random Forest rely on handcrafted spatial features but overlook temporal information, which limits their ability to model the dynamic behavior of agricultural machinery.

GCN, DR-XGBoost, and GAN-BiLSTM incorporate spatiotemporal correlations to varying degrees but still exhibit significant limitations. GCN performs static message passing without considering the importance of individual trajectory points. DR-XGBoost uses limited features such as speed and acceleration and lacks deeper feature representations. GAN-BiLSTM, though combining adversarial learning and bidirectional sequence modeling, primarily relies on basic statistics such as mean and standard deviation, ignoring richer descriptors like skewness, kurtosis, and trend features. Additionally, none of these models account for the redundancy introduced by enhanced features.

In contrast, SSL-GAT extracts comprehensive feature representations from individual trajectory points using a statistical enhancement module, suppresses redundant information via bidirectional fusion, and dynamically weights neighbors through the attention mechanism to emphasize key points. Furthermore, the use of large-scale unlabeled trajectories in the self-supervised pretraining phase helps the model learn intrinsic feature representations, thereby improving generalization. Together, these innovations enable SSL-GAT to outperform the baseline models with greater adaptability and robustness.

3.3 Ablation experiments

To verify the contribution of each core module, ablation experiments were conducted on both the paddy and wheat datasets. The impact of the multi-dimensional feature enhancement module (MFE), bidirectional feature fusion module (BFM), self-supervised learning (SSL), and graph attention mechanism (ATT) was evaluated. The overall results are shown in Tables 6 and 7.

Table 6 Performance comparison of the ablation experiments on the paddy harvester trajectory dataset (w/o means without)

Method	Accuracy/%	Precision/%	Recall/%	F1-score/%
SSL-GAT	95.92	90.47	94.58	92.42
SSL-GAT w/o MFE	92.65	85.60	90.36	87.80
SSL-GAT w/o BFM	87.12	84.62	86.75	85.55
SSL-GAT w/o SSL	88.75	82.67	85.69	83.99
SSL-GAT w/o ATT	90.77	85.29	88.89	86.94

Table 7 Performance comparison of the ablation experiments on the wheat harvester trajectory dataset

Method	Accuracy/%	Precision/%	Recall/%	F1-score/%
SSL-GAT	93.92	88.06	93.58	90.72
SSL-GAT w/o MFE	94.78	84.38	91.05	87.55
SSL-GAT w/o BFM	91.16	83.87	88.25	85.99
SSL-GAT w/o SSL	86.45	80.95	85.38	83.07
SSL-GAT w/o ATT	92.01	85.12	90.48	87.71

First, the MFE module was removed, and the results show that the F1-score decreased by 4.62% on the paddy harvester trajectory dataset (from 92.42% to 87.80%) and by 3.17% on the wheat harvester trajectory dataset (from 90.72% to 87.55%). This demonstrates that the MFE module effectively extracts comprehensive features and enhances the model's recognition capability.

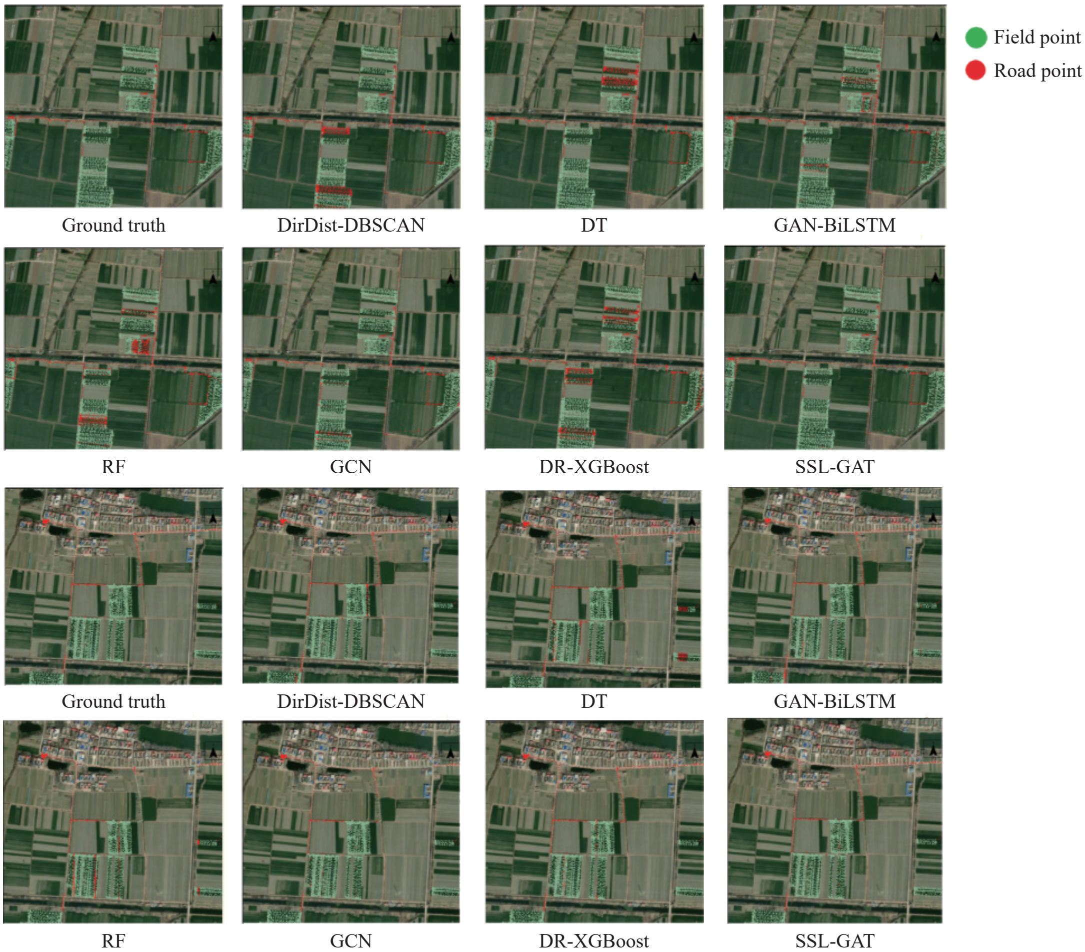

Figure 8 Identification results for two scenarios on the paddy harvester dataset

Second, the BFM module was removed, and the results show that the F1-score dropped by 6.87% (to 85.55%) on the paddy harvester dataset and by 4.73% (to 85.99%) on the wheat harvester trajectory dataset. This highlights the BFM's importance in enhancing identification of both "field" and "road" points in agricultural machinery trajectories.

Next, the attention mechanism was removed, and SSL-GCN was constructed to examine the impact of attention-based dynamic aggregation. The F1-score dropped by 5.48% on the paddy harvester trajectory dataset and by 3.01% on the wheat harvester trajectory dataset. Based on Tables 8 and 9, the F1-scores for the "field" class decreased by 3.22% and 2.48%, while those for the "road" class dropped by 7.74% and 3.53%, respectively. These results demonstrate the critical role of the attention mechanism, which adaptively assigns weights to neighbors, ultimately improving flexibility and robustness in graph-based feature propagation.

Table 8 Performance comparison of the ablation experiments on the "field" and "road" classes for the paddy harvester trajectory dataset

Method	Field			Road		
	Precision/%	Recall/%	F1-score/%	Precision/%	Recall/%	F1-score/%
SSL-GAT	95.01	94.82	94.91	85.93	94.34	89.93
SSL-GAT w/o MFE	92.58	91.29	91.93	78.63	89.43	83.68
SSL-GAT w/o BFM	93.04	88.81	90.87	76.21	84.7	80.23
SSL-GAT w/o SSL	90.33	86.05	88.13	75.02	85.34	79.84
SSL-GAT w/o ATT	92.94	90.48	91.69	77.65	87.31	82.19

Table 9 Performance comparison of the ablation experiments on the "field" and "road" classes for the wheat harvester trajectory dataset

Method	Field			Road		
	Precision/%	Recall/%	F1-score/%	Precision/%	Recall/%	F1-score/%
SSL-GAT	91.88	95.31	93.56	84.25	91.85	87.88
SSL-GAT w/o MFE	89.14	92.67	90.87	79.62	88.21	84.24
SSL-GAT w/o BFM	87.53	90.39	88.93	80.21	86.11	83.05
SSL-GAT w/o SSL	85.12	86.28	85.69	76.78	84.49	80.45
SSL-GAT w/o ATT	88.72	93.58	91.08	81.53	87.38	84.35

Finally, the self-supervised learning module was removed. This led to an F1-score reduction of 8.43% (from 92.42% to 83.99%) on the paddy harvester trajectory dataset and 7.65% (from 90.72% to 83.07%) on the wheat harvester trajectory dataset. This confirms the critical role of SSL in learning meaningful representations from unlabeled data.

In conclusion, the MFE, BFM, SSL modules, and graph attention mechanism are all essential components of the proposed model. Together, they collaboratively enhance the accuracy and robustness of the agricultural machinery trajectory operation mode identification task.

4 Method analysis and discussion

4.1 Multi-dimensional feature enhancement module

The Multi-dimensional Feature Enhancement (MFE) module plays a vital role in improving feature representation. Most existing methods rely solely on raw spatiotemporal features, which limits their capacity to extract rich latent motion information. In contrast, the proposed MFE module leverages both instant features and temporal features derived through statistical methods to enrich the representation of trajectory points.

To further illustrate the contribution of the Multi-dimensional Feature Enhancement (MFE) module, case-based analysis is conducted on both paddy and wheat harvester trajectories. For paddy harvesters, the low sampling frequency (30 s) often causes the raw trajectory to miss fine-grained turning behaviors and short-duration acceleration–deceleration patterns. After the MFE module is incorporated, the temporal features extracted by sliding windows effectively compensate for missing local dynamics. In particular, the skewness and kurtosis of the direction changes helped the model distinguish field turning segments that were previously misidentified as road segments in the "w/o MFE" variant.

For wheat harvesters, which have a higher sampling frequency (5 s), the instant features (acceleration and angular velocity) provide richer short-term motion information. These features allowed the model to more accurately identify boundary points between the field and the road, especially at field entrances where the machinery rapidly decelerates before turning. Without MFE, these segments were frequently misclassified as "road", explaining the performance drop shown in Tables 6-9.

Ablation experiments (Tables 6 and 7) confirm the effectiveness of MFE, showing 4.62% and 3.17% improvements in F1 scores on the paddy and wheat datasets, respectively, demonstrating that the enhanced features significantly strengthen the model's ability to distinguish different operation modes. While some previous models, such as GAN-BiLSTM and DR-XGBoost, incorporate feature augmentation, they lack explicit modeling of temporal trends. By capturing both instantaneous and temporal characteristics, the MFE module provides an essential foundation for accurate and robust operation mode identification.

4.2 Bidirectional feature fusion module

After feature enhancement, trajectory points may still contain redundant information as well as noisy points caused by measurement errors. To address this issue, the bidirectional feature fusion module (BFM) is introduced to enhance trajectory representation by fusing features across both inter-point and intra-point dimensions. BFM performs feature fusion, enabling the model to learn latent dependencies among motion features within each point while suppressing the influence of redundant points.

In practical scenarios, BFM has clear advantages. For paddy harvester data, GNSS noise often appears near field edges where signals are obstructed by vegetation or terrain; without BFM, these noisy clusters are prone to be misclassified as field–road transitions owing to abrupt deviations in direction or velocity. With BFM, spatial and temporal neighborhood information is integrated, enabling the model to correct these distortions and maintain stable predictions. For wheat harvester trajectories with denser sampling, BFM mitigates redundancy by suppressing repetitive high-frequency points along straight road segments, allowing the model to emphasize meaningful motion variations instead of being influenced by redundant patterns.

The effectiveness of BFM is clearly demonstrated in the ablation results (Tables 6-9). On the paddy dataset, the overall F1 score decreases from 92.42% to 85.55% when BFM is removed (Table 6), with "field" and "road" class F1 scores dropping from 94.91% to 90.87% and from 89.93% to 80.23%, respectively (Table 8). On the wheat dataset, the overall F1 score drops from 90.72% to 85.99% (Table 7), with corresponding class-level decreases from 93.56% to 88.93% and from 87.88% to 83.05% (Table 9). These results confirm that BFM not only suppresses noise and redundancy but also enhances the model's ability to recognize challenging segments such as turning areas, field entrances, and field–road transition points. Overall, its bidirectional fusion mechanism provides essential robustness for accurately identifying agricultural machinery operation modes under complex real-world conditions.

4.3 Graph attention network

To capture the varying importance of different trajectory points, the SSL-GAT model adopts a Graph Attention Network (GAT) as its backbone, enabling dynamic weighting of spatiotemporal relationships. The GAT component operates in two key steps: first, a spatiotemporal graph is constructed from trajectory points using three types of edges: temporal-connection edges that link points close in time, spatial-connection edges that connect points close in space, and self-loop edges that preserve node-specific information. This structure encodes complex trajectory dependencies and transforms raw trajectory sequences into a graph representation. Second, GAT performs dynamic feature aggregation through an attention mechanism that adaptively adjusts node attention weights based on implicit relationships among trajectory points. Important points receive higher weights, while irrelevant or noisy ones are downweighted, thereby improving the stability and robustness of feature extraction.

The importance of attention is clearly demonstrated in the ablation results (Tables 6-9). On the paddy harvester trajectory dataset, removing attention reduces the overall F1-score from 92.42% to 86.94%, and the "field" and "road" class F1 scores from 94.91% to 91.69% and from 89.93% to 82.19%, respectively. On the wheat dataset, the overall F1-score decreases from 90.72% to 87.71%, with class-level declines from 93.56% to 91.08% ("field") and from 87.88% to 84.35% ("road"). These reductions show that attention is essential for effectively weighting neighboring trajectory points.

In practical agricultural scenarios, the GAT plays a critical role through the attention mechanism in the following typical situations. First, at field–road boundary regions, trajectory points often contain noise or mixed movement patterns. With attention, the model tends to rely more on stable neighbors, reducing boundary misclassification. Without attention, uniform aggregation increases errors, particularly for the "road" class. Second, for turning points, the motion changes sharply. The attention mechanism highlights these points, while removing attention causes them to be averaged with nearby stable points, making turning segments harder to identify. Third, in long straight road segments, occasional GNSS jitter may resemble field-like fluctuations. The attention mechanism suppresses such outliers, whereas the non-attention model is more easily influenced by these noisy points.

Overall, the attention mechanism enables the model to concentrate on the most informative trajectory points and reduces the impact of noise, which is especially important in boundary regions, turning segments, and long straight road segments.

4.4 Self-supervised learning

Due to variations in operating environments, agricultural trajectories often exhibit highly complex and inconsistent spatiotemporal patterns. To address these challenges, a self-supervised learning (SSL) composed of feature masking and graph-based feature reconstruction is introduced, enabling the model to learn structural information from unlabeled trajectory data. As shown in Tables 6 and 7, removing the SSL module decreases the F1-score from 92.42% to 83.99% on the paddy dataset and from

90.72% to 83.07% on the wheat dataset, demonstrating that SSL substantially enhances representation quality.

The effectiveness of SSL is evident in several typical agricultural scenarios. First, for low sampling-rate trajectories where crucial motion transitions between adjacent points may be missing, SSL enables the model to infer these patterns by learning consistent temporal structures during pretraining. Without SSL, the model must rely solely on sparse local observations, resulting in unstable predictions. Second, trajectory patterns differ significantly across regions due to factors such as crop type, field size, and operator habits. SSL pretraining allows the model to learn generalizable structural properties shared across datasets, thereby reducing the impact of distribution shifts. Models trained without SSL show reduced adaptability to unfamiliar motion characteristics. Third, during field operations with irregular or noisy behavior, such as frequent steering adjustments or stop-and-go movements, SSL helps the model separate meaningful trends from noise by learning spatiotemporal correlations. When the SSL is removed, misclassification rates increase noticeably, especially in complex field segments.

Overall, these case analyses show that SSL improves model robustness in three challenging situations: sparse sampling, regional variability, and noisy or irregular field operations. This explains the substantial decline in performance when SSL is removed and confirms that SSL is an indispensable component of the proposed model.

5 Conclusions

Agricultural machinery trajectory operation mode identification, which aims to distinguish field points from road points, is fundamental to downstream trajectory analysis tasks. To address the limitations of existing methods in extracting spatiotemporal features, this study proposes a self-supervised learning-based graph attention network (SSL-GAT). Through the combination of the modules, SSL-GAT effectively captures complex spatiotemporal relationships, reduces noise and redundant features, and enhances the contribution of key trajectory points, thereby improving the accuracy and robustness of operation mode identification.

Experimental results on paddy and wheat harvester datasets demonstrate the effectiveness of the proposed method. On the paddy dataset, SSL-GAT achieves F1 scores of 94.91% and 89.93% for the "field" and "road" categories, respectively; on the wheat dataset, it achieves F1 scores of 93.56% and 87.88%. These results show that SSL-GAT consistently outperforms existing state-of-the-art models.

Despite these advantages, the model remains constrained by the limited number of available features, which affects its identification capability. Future work will explore more advanced feature enhancement strategies to further strengthen spatiotemporal representation learning. Additionally, since model performance improves with higher trajectory sampling frequencies, we plan to investigate methods for increasing the density of trajectory point generation to further enhance recognition performance.

Acknowledgements

This research was financially supported by the National Key Research and Development Program of China (Grant No. 2025YFE0103600 and 2021YFB3901300), the National Natural Science Foundation of China (Grant No. 32301691), the China Scholarship Council (Grant No. 202306350108), and the 2115 Talent Development Program of China Agricultural University.

[References]

[1] Zhai W X, Xu Z, Pan J W, Guo Z, Wu C C. A general image classification model for agricultural machinery trajectory mode recognition. Computers and Electronics in Agriculture, 2024; 227: 109629.

[2] Wu C C, Li D, Zhang X Q, Pan J W, Quan L, Yang L L, et al. China's agricultural machinery operation big data system. Computers and Electronics in Agriculture, 2023; 205: 107594.

[3] Zhai W X, Xiong X Y, Mo G Z, Xiao Y Z, Wu C C, Xu Z, et al. A bagging-SVM field-road trajectory classification model based on feature enhancement. Computers and Electronics in Agriculture, 2024; 217: 108635.

[4] Pan J W, Wu C C, Zhai W X. A hybrid genetic slime mould algorithm for parameter optimization of field-road trajectory segmentation models. Information Processing in Agriculture, 2024; 11(4): 590–602.

[5] Zhai W X, Xu Z, Liu J M, Xiong X Y, Pan J W, Chung S, et al. Feature deformation network with multi-range feature enhancement for agricultural machinery operation mode identification. Int J Agric & Biol Eng, 2024; 17(4): 265–275.

[6] Pan J W, Guo Z, Wu C C, Zhai W X. Parameter optimization of the field-road trajectory segmentation model based on the chaos sensing slime mould algorithm. Soft Computing, 2024; 28(19): 11065–11132.

[7] Sun Q, Xia J H, Foster J, Falkmer T, Lee H. Pursuing precise vehicle movement trajectory in urban residential area using multi-GNSS RTK tracking. Transportation Research Procedia, 2017; 25: 2356–2372.

[8] Zhai W X, Kuang X R, Cheng X Y, Pan J W, Wu C C. Reconstruction of missing points in agricultural machinery trajectory based on bidirectional adjacent information. Computers and Electronics in Agriculture, 2024; 220: 108920.

[9] Zhai W X, Ni Z X, Xu Z, Pan J W, Wu C C. T2R-pix2pix: A method for constructing rural thematic road network based on pix2pix. Computers and Electronics in Agriculture, 2025; 230: 109911.

[10] Song X P, Potapov P V, Krylov A, King L, Di Bella C M, Hudson A, et al. National-scale soybean mapping and area estimation in the United States using medium resolution satellite imagery and field survey. Remote Sensing of Environment, 2017; 190: 383–395.

[11] Rodias E, Berruto R, Busato P, Bochtis D, Sørensen C G, Zhou K. Energy savings from optimised in-field route planning for agricultural machinery. Sustainability, 2017; 9(11): 1956.

[12] Liu Y, Shi L, Gao Y, Kou C, Yang S. Research on the optimized management of agricultural machinery allocation path based on teaching and learning optimization algorithm. Tehnički vjesnik - Technical Gazette, 2022; 29(2): 456–463.

[13] Li H C, Gao F, Zuo G C. Research on the agricultural machinery path tracking method based on deep reinforcement learning. Scientific Programming, 2022; 2022: 6385972.

[14] Liu L, Chen T, Yang S G, Wang X L. Analysis on the mode of trans-regional allocation of agricultural machinery. American Journal of Plant Sciences, 2020; 11(7): 1049.

[15] Li D, Liu X, Zhou K, Sun R Z, Wang C T, Zhai W X, et al. Discovering spatiotemporal characteristics of the trans-regional harvesting operation using big data of GNSS trajectories in China. Computers and Electronics in Agriculture, 2023; 211: 108003.

[16] Lacour S, Burgun C, Perilhon C, Descombes G, Doyen V. A model to assess tractor operational efficiency from bench test data. Journal of Terramechanics, 2014; 54: 1–18.

[17] Lee J W, Kim J S, Kim K U. Computer simulations to maximise fuel efficiency and work performance of agricultural tractors in rotovating and ploughing operations. Biosystems Engineering, 2016; 142: 1–11.

[18] Bochtis D D, Sørensen C G C, Busato P. Advances in agricultural machinery management: A review. Biosystems Engineering, 2014; 126: 69–81.

[19] Sopegno A, Calvo A, Berruto R, Busato P, Bocthis D. A web mobile application for agricultural machinery cost analysis. Computers and Electronics in Agriculture, 2016; 130: 158–168.

[20] Molari G, Mattetti M, Lenzini N, Fiorati S. An updated methodology to analyse the idling of agricultural tractors. Biosystems Engineering, 2019; 187: 160–170.

[21] Pagare V, Nandi S, Khare D K. Appraisal of optimum economic life for farm tractor: A case study. Economic Affairs, 2019; 64(1): 111–118.

[22] Bhat S A, Huang N F. Big data and AI revolution in precision agriculture: Survey and challenges. IEEE Access, 2021; 9: 110209–110222.
[23] Kamilaris A, Kartakoullis A, Prenafeta-Boldú F X. A review on the practice of big data analysis in agriculture. Computers and Electronics in Agriculture, 2017; 143: 23–37.
[24] Liakos K G, Busato P, Moshou D, Pearson S, Bochtis D. Machine learning in agriculture: A review. Sensors, 2018; 18(8): 2674.
[25] Pan C Y, Zhai W X. AMGFormer: A multilevel GNSS recording autoencoder for agricultural machinery trajectory operation mode identification. Computers and Electronics in Agriculture, 2025; 237: 110638.
[26] Zhai W X, Guo Z, Pan J W, Xu J W, Wu C C. Addressing local sparsity in massive agricultural machinery trajectories: A BiLSTM-GRU approach. Computers and Electronics in Agriculture, 2024; 226: 109376.
[27] Zhai W X, Zhang X Y, Liu J M, Pan J W, Wu C C. A dual-level interactive cascade network for agricultural machinery trajectory operation mode identification. Computers and Electronics in Agriculture, 2025; 238: 110788.
[28] Hao F Q, Zhao X Y, Bian C X, Kong H, Dong X J, Bai J Q. VEBiLSTM: A Neural network for field-road classification using enhanced spatiotemporal features. In: Neural Information Processing, 2025; 15291: 245–258.
[29] Poteko J, Eder D, Noack P O. Identifying operation modes of agricultural vehicles based on GNSS measurements. Computers and Electronics in Agriculture, 2021; 185: 106105.
[30] Kotsiantis S B. Decision trees: A recent overview. Artificial Intelligence Review, 2011; 39(4): 261–283.
[31] Chen Y, Zhang X Q, Wu C C, Li G Y. Field-road trajectory segmentation for agricultural machinery based on direction distribution. Computers and Electronics in Agriculture, 2021; 186: 106180.
[32] Chen Y, Li G Y, Zhang X Q, Jia J P, Zhou K, Wu C C. Identifying field and road modes of agricultural machinery based on GNSS recordings: A graph convolutional neural network approach. Computers and Electronics in Agriculture, 2022; 198: 107082.
[33] Xiao Y Z, Mo G Z, Xiong X Y, Pan J W, Hu B B, Wu C C, et al. DR-XGBoost: An XGBoost model for field-road segmentation based on dual feature extraction and recursive feature elimination. Int J Agric & Biol Eng, 2023; 16(3): 169–179.
[34] Zhai W X, Mo G Z, Xiao Y Z, Xiong X Y, Wu C C, Zhang X Q, et al. GAN-BiLSTM network for field-road classification on imbalanced GNSS recordings. Computers and Electronics in Agriculture, 2024; 216: 108457.
[35] Tolstikhin I, Houlsby N, Kolesnikov A, Beyer L. MLP-Mixer: An all-MLP architecture for vision. 2021; arXiv e-press, arXiv: 2105.01601. doi: 10.48550/arXiv.2105.01601.
[36] Veličković P, Cucurull G, Casanova A, Romero A, Liò P, Bengio Y. Graph Attention Networks. 2018; arXiv e-press, arXiv: 1710.10903. doi: 10.48550/arXiv.1710.10903.
[37] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, et al. Attention is all you need. Advances in Neural Information Processing Systems, 2017; 30. Available: https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html. Accessed on [2025-12-03].
[38] Liu X, Zhang F J, Hou Z Y, Mian L, Wang Z Y, Zhang J, et al. Self-supervised learning: Generative or contrastive. IEEE Transactions on Knowledge and Data Engineering, 2021; 35(1): 857–876.
[39] Liu Y X, Jin M, Pan S R, Zhou C, Zheng Y, Xia F, et al. Graph self-supervised learning: A survey. IEEE Transactions on Knowledge and Data Engineering, 2023; 35(6): 5879–5900.
[40] Kingma D P, Ba J. Adam: A method for stochastic optimization. 2017; arXiv e-press, arXiv: 1412.6980. doi: 10.48550/arXiv.1412.6980.