Article

Application of Variational Graph Autoencoder in Traction Control of Energy-Saving Driving for High-Speed Train

Weigang Ma, Jing Wang *, Chaohui Zhang *, Qiao Jia, Lei Zhu ⓘ, Wenjiang Ji ⓘ and Zhoukai Wang

Faculty of Computer Science and Engineering, Xi'an University of Technology, Xi'an 710048, China; mwg_markey@xaut.edu.cn (W.M.); 2191221087@stu.xaut.edu.cn (Q.J.); leizhu@xaut.edu.cn (L.Z.); wjj@xaut.edu.cn (W.J.); zkwang@xaut.edu.cn (Z.W.)
* Correspondence: 2221220003@stu.xaut.edu.cn (J.W.); 2231221144@stu.xaut.edu.cn (C.Z.);
  Tel.: +86-19834405728 (J.W.); +86-18793810835 (C.Z.)

Abstract: In a high-speed rail system, the driver repeatedly adjusts the train's speed and traction while driving, causing a high level of energy consumption. This also leads to the instability of the train's operation, affecting passengers' experiences and the operational efficiency of the system. To solve this problem, we propose a variational graph auto-encoder (VGAE) model using a neural network to learn the posterior distribution. This model can effectively capture the correlation between the components of a high-speed rail system and simulate drivers' operating state accurately. The specific traction control is divided into two parts. The first part employs an algorithm based on the K-Nearest Neighbors (KNN) algorithm and undersampling to address the negative impact of imbalanced quantities in the training dataset. The second part utilizes a variational graph autoencoder to derive the initial traction control of drivers, thereby predicting the energy performance of the drivers' operation. An 83,786 m long high-speed train driving section is used as an example for verification. By using a confusion matrix for our comparative analysis, it was concluded that the energy consumption is approximately 18.78% less than that of manual traction control. This shows the potential and effect of the variational graph autoencoder model for optimizing energy consumption in high-speed rail systems.

Keywords: high-speed train; variational graph auto-encoders (VGAE); energy saving; traction control; link prediction

1. Introduction

The traction control of high-speed trains is a factor critical to ensuring their safety, punctuality, and reduced energy consumption. However, traditional traction control methods often make it challenging to achieve the optimal traction distribution in the face of complex dynamic characteristics and environmental conditions during the operation of high-speed trains. Especially in regards to the issue of energy consumption, drivers often need to make a difficult trade-off between traction and speed, resulting in an unnecessary waste of energy and reduced operating efficiency.

Saving energy during the operation of trains is essentially an optimization control problem and has attracted the attention of many scholars. Control strategies based on dynamic models are more common than others, but such methods face challenges in accurately describing the dynamic characteristics of train operation [1]. Scholars have applied modern intelligent control methods to the field of train operational control; thus, a series of methods using deep learning have been developed. In terms of the driving strategy of high-speed railway trains, Lei et al. optimized the genetic algorithm based on the speed limit, slope of the actual route, and the constraints of the vehicle. Although it has an obvious energy-saving effect, it has a high degree of computational complexity and strong data dependence, and it is difficult to adjust its parameters [2]. Ning et al. proposed a high-speed

railway train track optimization model considering continuous train constraints but could not guarantee the global optimal solution [3]. Cao et al. proposed a method based on mixed integer linear programming to optimize the trajectory of high-speed trains, but the process of transforming the problem into a mathematical model is more complicated [4]. Based on the Soft Actor Critic method, Su et al. proposed a method to optimize the train driving strategy, but it is sensitive to the initial strategy, which may cause the algorithm to fall into the local optimal solution or the convergence speed to be slow [5]. Zhu et al. transformed the inverse problem of training energy-saving control into a decision-making machine learning algorithm solved by a finite Markov process through deep reinforcement. They adopted the algorithm of the Soft Actor Critic method to determine the optimal driving strategy, but selecting the appropriate hyperparameters often requires a lot of experiments and tuning [6]. Zhang et al. proposed a real-time energy-saving optimization method of timetable based on an improved differential evolution algorithm and fast iteration method, but the effect of reducing energy consumption is not obvious [7]. Ying et al. optimized the curve of energy saved during train operation by quadratic constraint linear programming, but it was limited to only the neutral zone system [8]. Havaei et al. proposed a new intelligent proportional–integral–differential controller to optimize the velocity trajectory. However, when there is integral saturation in the system, the performance of the controller will decrease [9].

A high-speed railway system usually contains complex nodes and edges, and the interaction and influence of these are complex to model by traditional methods. The GNN (graph neural network) can capture important information in the network through mechanisms such as convolution, attention addition, and information dissemination, thus providing more accurate energy consumption prediction and energy-saving optimization [10]. The VGAE developed from the variational auto-encoder (VAE) is a graph neural network that uses latent variable learning to approximate the Gaussian distribution and can effectively solve the link prediction problem [11–13]. Nowadays, an algorithm based on graph neural networks is mainly applied in the field of transportation to solve the problem of traffic flow prediction in urban traffic networks and is rarely applied to issues related to high-speed trains [14–16].

The VGAE can effectively learn the relationship between nodes in complex systems, which is crucial for modeling the interaction between components in high-speed rail systems. By modeling the high-speed rail system as a graph, the VGAE can learn the low-dimensional representation of nodes to encode and represent the system state effectively. In addition, the VGAE can also expand the training data and improve the generalization ability of the model by generating synthetic samples. Based on previous research, this paper applies the link prediction method of the VGAE to the optimal control problem of trains. To solve this problem, the implicit graph structure is established, and the train traction problem is abstracted into a multi-classification problem. The VGAE is employed to derive an intelligent traction control, which is continuously adjusted to calculate the traction's energy consumption.

2. The Formation of Energy-Saving Problem of High-Speed Trains

To address the energy-saving control problem of high-speed trains, we aim to reduce energy consumption, improve traction efficiency, and reduce operating costs during train operation. The multiple aspects covered by this question are shown in Table 1. This paper applies advanced traction control and braking strategies to model the traction process, accurately and dynamically adjusting the strategy to adapt to different operating conditions. The force analysis of the train operation is shown in Figure 1. The train is regarded as a single-particle model, and it is subjected to five forces during its operation: train gravity G, support force N, traction force F, running resistance W, and braking force B. Because the traction and braking force are two different working conditions of a train power system, traction F and braking force B cannot exist simultaneously.

Table 1. Classification of energy-saving strategies for high-speed trains.

| Challenge	Description
Braking strategy	Minimize energy consumption during deceleration and parking
Traction control	Through intelligent traction control, the train can use energy more efficiently
Energy recovery	The energy released by the train braking is re-injected into the power supply system
Power system	Optimize the power system of the train and improve the efficiency of energy utilization
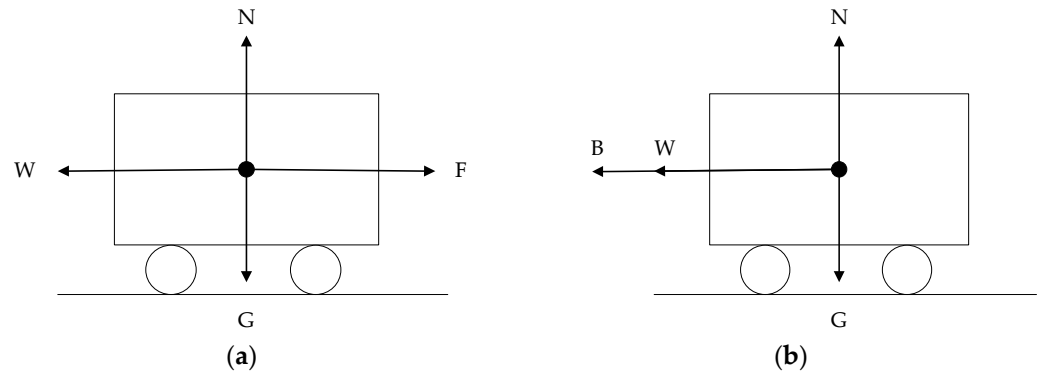
Figure 1. Force analysis of straight line. (**a**) Acceleration or constant speed; (**b**) Braking.

The high-speed train adopts a composite braking system with electric braking as the main method and air braking as the auxiliary method. The most commonly used braking method prioritizes non-wearing electric braking. In contrast, air braking or composite braking is used in emergency braking situations and is difficult to describe in specific application scenarios. The basic braking control modes include regular, emergency, anti-skid, and parking braking [17]. Compared with pure mechanical braking, the driving strategy combining regenerative and mechanical braking saves more energy [18]. The brake handling level contains different braking modes, which can meet the requirements of high-speed trains under different operating conditions. These include common brake levels (1A, 1B, 2, 3, 4, 5, 6, 7, 8), OC, REL, and EB. In the same operating environment, different braking operations consume varying amounts of traction electrical energy [19].

Without considering the influence of other conditions, the average increase in traction energy consumption per second for each brake lever position is shown in Table 2. REL corresponds to the traction position, and its traction energy consumption is greater than that of other operations. In the event of an emergency, the brake lever needs to be placed in the "EB" position. However, if the train is stopped on a track segment where evacuation is not feasible, the driver can place the brake lever in the "OC" position to cancel the excessive braking force. Therefore, compared to common braking positions, it results in higher energy consumption, which is consistent with the statistical results in Table 2.

Table 2. Brake handling level statistics.

Brake Handling Level	OC			REL			EB		
Brake instructions	Ignore passengers and activate the emergency brake position			Running brake position			Emergency brake		
Average energy consumption increase (kwh/s)	0.1226			0.4718			0.1403		

Brake handling level	1A	1B	2	3	4	5	6	7	8
Average energy consumption increase (kwh/s)	0.08	0.0017	0.0008	0.001	0.0015	0.0029	0.0039	0.0051	0.0115

By improving the operation mode of high-speed trains, such as by optimizing the acceleration and deceleration strategies of vehicles, improving the speed control of trains

on different road sections, and optimizing the parking strategies of trains, the traction energy consumption during train operation can be effectively reduced [20,21]. The optimal traction control is studied to determine the application time of the brakes and the type of operating brake handling. The braking problem can be transformed into the following equivalent decision problem.

$$y_B = q(S_t) \tag{1}$$

$$S_{nt} = yB(S_t) \tag{2}$$

In Equations (1) and (2) above, S_t represents the running state of the train at decision time t, which is comprehensively reflected by the running environment, the attributes of the train, and the running state. $q(.)$ maps the train state to the braking operation that the drivers may perform; and y_B represents the specific braking operation, and its value is a specific number of discrete quantities, namely $y_B \epsilon \{OC, REL, 1A, 1B, 2, 3, 4, 5, 6, 7, 8, EB\}$. Therefore, the process of obtaining an intelligent braking application strategy by learning from actual driving data can be considered as solving a multi-classification problem. y_B operates for the drivers to change the current running state from S_t to S_{nt}.

The energy saving can be determined by whether the resultant force Δ of the train is greater than 0. The energy-saving auxiliary labels defined for each sample are shown in (3):

$$L_e = \begin{cases} a, \Delta < 0 \\ b, \Delta = 0 \\ c, \Delta > 0 \end{cases} \tag{3}$$

where, when $\Delta < 0$, the direction of the resultant force on the train is consistent with the running direction. In this case, most of the trains are idle, and the energy consumption does not increase much. When $\Delta = 0$, the resultant force on the train in the horizontal direction is 0. In this case, most trains are in the cruise state, and the energy consumption does not increase much. When $\Delta > 0$, the direction of the resultant force on the train is opposite to the running direction. In this case, most of the trains are running with an increased level of energy consumption.

3. Energy-Saving Traction Control Model Based on VGAE

The proposed energy-saving traction control model based on the VGAE is divided into two parts: one constructs the equilibrium training set, and the other constructs the implicit graph structure according to the problem and uses the VGAE to solve the problem. The overall process is shown in Figure 2 as follows:

The class of data with the least number of brake handling operations is defined as P0, and the other classes are defined as $Pi, i = 1, 2, ... , w - 1$.

3.1. Constructing the Equilibrium Training Set

Like most multi-classification studies, the braking decision problem of high-speed trains includes the phenomenon of sample imbalance; that is, the number of samples varies greatly between categories, and it is difficult to extract effective information from classes with too few samples.

The training set without any preprocessing leads to poor performance in multi-class classification results, with unsatisfactory levels of model accuracy, robustness, and other performance indicators. Moreover, the classification error of the minority class is more serious than that of the majority class [22]. The phenomenon of data imbalance is often ignored in the braking problem. While solving the air-braking decision problem of heavy-haul trains, the literature [23] mentions that the information embodied in the few data sets that are ignored may be related to some special conditions encountered by the train.
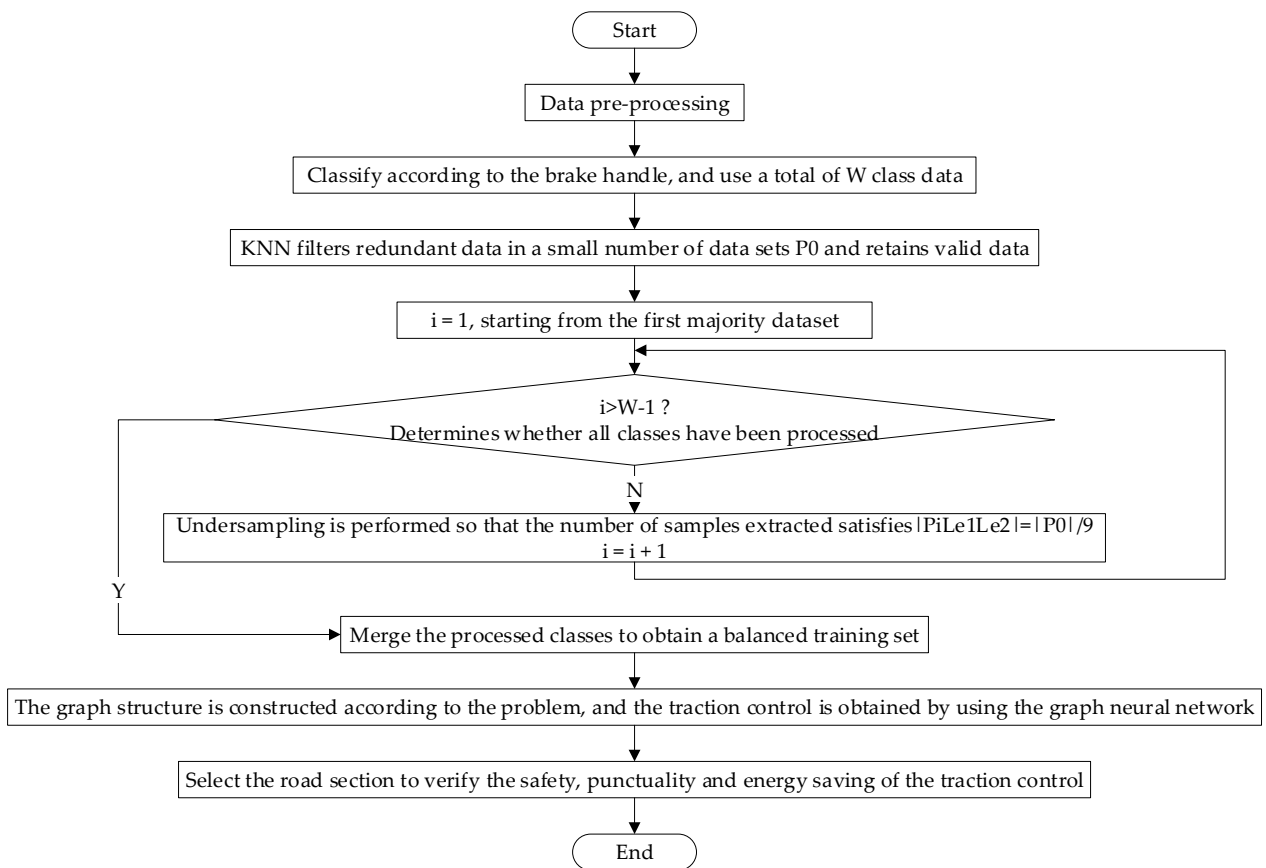
```
                          ┌─────────┐
                          │  Start  │
                          └────┬────┘
                               ▼
                     ┌───────────────────┐
                     │ Data pre-processing│
                     └─────────┬─────────┘
                               ▼
        ┌──────────────────────────────────────────────┐
        │ Classify according to the brake handle, and   │
        │ use a total of W class data                    │
        └──────────────────┬───────────────────────────┘
                           ▼
    ┌───────────────────────────────────────────────────────┐
    │ KNN filters redundant data in a small number of data   │
    │ sets P0 and retains valid data                          │
    └───────────────────────┬───────────────────────────────┘
                            ▼
        ┌──────────────────────────────────────────┐
        │ i = 1, starting from the first majority    │
        │ dataset                                     │
        └──────────────────┬─────────────────────────┘
                           ▼
                   ◇ i>W-1 ? ◇
       Determines whether all classes have been processed
```

Figure 2. The overall process of energy−saving driving strategy.

Therefore, this paper needs to overcome the problem of data imbalance in the training samples so that the number of samples in each case is as balanced as possible. The implementation steps of this module are divided into two parts.

3.1.1. Eliminating Redundant Noise Data Based on KNN

The redundant data generated by high-speed train operations includes repeated data and some useless surplus data. Noise data are generated by incorrect operation when drivers are nervous. This information has a greater impact on a small number of data sets [24], so this part only focuses on a small number of data sets. Considering the classification performance and computational cost, the redundant and noisy data are filtered by the KNN-based method. An equal number of samples were randomly selected from small data sets and big data sets as known categories, and samples from the small data sets were chosen successively as unknown samples. Equation (4) was used to calculate the distance between unknown samples and known samples:

$$d_{x_i x_j} = \sqrt{\sum_{r=1}^{n} \left(x_i^r - x_j^r \right)^2} \tag{4}$$

Wherein $d_{x_i x_j}$ represents the Euclidean distance between the unknown sample x_i and the known sample x_j, and r represents the rth attribute.

We found the k samples with a high degree of similarity to the unknown samples and determined the category of the unknown samples as the category with the most occurrence times among the k samples. The selection of the k value was determined by the five-fold cross-validation method, and the k value with the higher average accuracy was selected. After determining k, we compared the actual value and the predicted value in the original

high-speed train driving data for consistency. If they were not consistent, the data were removed from a small number of data sets by filtering.

3.1.2. The Undersampling Algorithm to Solve the Imbalance Problem

After obtaining the filtered data set, the undersampling algorithm is used to randomly select samples from big data sets to construct a training set with a balanced sample number so that the number of samples extracted from each class satisfies Equation (5) as follows:

$$|PiL_{e1}\ L_{e2}| = |P0|/|L_e|^2 \tag{5}$$

The difference between this study and other undersampling algorithms is that each class of big data sets satisfies $|Pi| = |P0|$; that is, under the premise of ensuring that the number of each braking operation in the training set is the same, the number of energy-saving state transitions in each class is also as consistent as possible. We suppose there is no mapping of a certain relation under a certain operation or the number of samples is less than the specified number of samples extracted. In that case, we should increase the number of samples in the other operation cases to ensure the balance of the training sets.

3.2. Solution of Variational Graph Auto-Encoder Strategy

By constructing the graph structure of a high-speed rail system, the VGAE strategy represents the complex relationship between the braking operation and the train attributes in the operation process as a network of nodes and edges. It uses the neural network to learn the posterior distribution and the potential representation of the nodes to design the corresponding optimization algorithm or decision strategy so that the train drivers can flexibly adjust the braking intensity and timing of the train according to the current operational situation and environmental information to achieve the optimal level of energy consumption.

3.2.1. Diagram Construction

According to the description of the problem in Section 2, the braking operation and running condition of a high-speed train are modeled as graph G = (V, E, F), as shown in Figure 3:
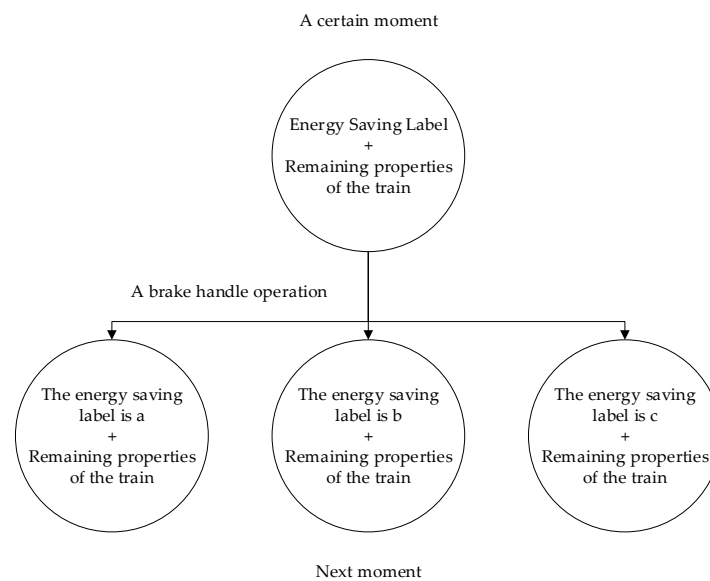

Figure 3. Simplified abstract diagram of high-speed train braking problem.

Where node V represents the set of running states of high-speed trains, and node set $F = \{L_e, f_2, \ldots, f_m\}$ contains attributes that reflect the running state of high-speed trains,

in which the number of attributes is m, including the energy-saving label L_e. The edge set E contains the direction information, and the node representing S_t points to the node representing S_{nt}, indicating that the specific braking operation y_B executed at a certain time makes the train's running state change from S_t to S_{nt}.

The problem of this study is transformed into a link prediction problem of a directed graph, which predicts the possibility of a certain type of link between nodes. It is similar to the description of the possibility of synergistic lethality in the medical field [25]; that is, according to the attributes of high-speed operations and their energy-saving labels, the most likely braking operation is speculated.

3.2.2. Variational Graph Autoencoder Solution

The VGAE is an unsupervised learning framework that can utilize neural networks to learn posterior distributions and obtain interpretable latent representations of undirected graphs.

Compared with other tasks that can make link predictions [26], the VGAE model can naturally integrate the characteristics of nodes, so the model has achieved good results in the link prediction task. The basic structure of the VGAE defined in this study is shown in Figure 4:
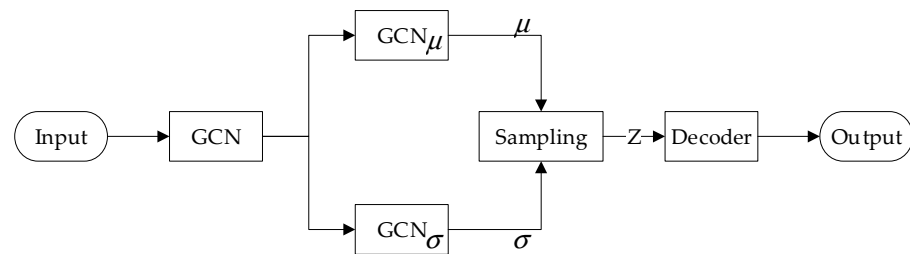

Figure 4. The basic structure of variational graph auto-encoders.

In the figure, the first layer is the shared GCN, and the second layer is composed of parallel $GCN\mu$ and $GCN\sigma$. The brake diagram has f nodes, where each node contains m train operation attributes, and all the attributes are expressed as the attribute information characteristic matrix X of $f \times m$. The node represented by the current train state S_t has a connection with all the nodes represented by the next train state S_{nt}, indicating that the drivers have performed the operation, and the relationship between the nodes with the connection is 1; otherwise, the edge relation value is 0, and all the edge relations form an adjacency matrix A of $f \times f$. The feature matrix X and the adjacency matrix A of the graph under each perspective are input into the graph autoencoder of each perspective, respectively. The mean and variance of the target distribution are calculated, and the Gaussian distribution of the target is obtained by using the graph convolution network, as shown in Equations (6) and (7).

$$\mu = GCN_\mu(X, A) \tag{6}$$

$$log\, \sigma = GCN_\sigma(X, A) \tag{7}$$

The mean vector μ and the covariance matrix σ are learned by the GCN, where the first layer parameter W_0 in GCN_μ and GCN_σ is shared, and the second layer parameter W_1 is not shared. The mean and variance are sampled by reparameterization, and the posterior distribution of the input is determined by the mean and variance. Then, the potential vector Z is obtained, as shown in (8) and (9).

$$q(Z|X, A) = \prod_{i=1}^{N} q(z_i | X, A) \tag{8}$$

$$q(z_i|X,\ A) = N\left(z_i\ \middle|\mu_i,\ diag\left(\sigma_i^2\right)\right) \tag{9}$$

Here, z_i is the element in the latent vector Z, and $q\ (Z\mid X,\ A)$ is the obtained posterior distribution. The adjacency matrix is reconstructed using the inner product of hidden variables, and the reconstruction is realized by calculating the probability of edges between points. Finally, a reconstruction diagram of the energy-saving traction for high-speed train driving is obtained as shown in (10) and (11).

$$p(A|Z) = \prod_{i=1}^{N}\prod_{j=1}^{N} p\left(A_{ij}|z_i,\ z_j\right) \tag{10}$$

$$p\left(A_{ij} = 1\middle|z_i,\ z_j\right) = \sigma\left(z_i^T\ z_j\right) \tag{11}$$

$p(A\mid Z)$ is the process of calculating probability; $p\left(A_{ij} = 1\middle|z_i,\ z_j\right)$ takes the inner product of latent variables sampled from the distribution obtained by the encoder as the decoder, where A_{ij} is an element of A, and $\sigma\ (.)$ is the Sigmoid activation function. The loss function needs to determine the similarity between the reconstructed graph and the original graph, as well as between the distribution calculated by the GCN and the standard Gaussian distribution, which are measured by the cross entropy and *KL* divergence, respectively, as shown in (12):

$$\mathcal{L} = E_{q(Z\mid X,\ A)}[\ log\ p\ (A\mid Z)]KL[q(Z\mid X,\ A)|\ |p(Z)] \tag{12}$$

4. Simulation Process and Results of Train Energy-Saving Optimization

In the previous section, an energy-saving traction control model based on the VGAE was proposed. This section describes our supervised training and reconstruction of train traction based on real data of high-speed train operation, a simulation of this, and a comparison of the results. Through the analysis of the algorithm and its energy-saving performance, the effectiveness of the proposed model is verified.

4.1. Simulation Experiment Settings

The simulation experiment is carried out on the basis of the proposed scheme. Firstly, it is necessary to select and set the experimental circuit, including the sample size and circuit information of the brake handling. In addition, various parameters need to be set, such as the sample size of the training set, the number of iterations, and the embedding dimension. These parameter settings will directly affect the results and reliability of the simulation experiment.

4.1.1. Experimental Circuit

The original data are the actual data of more than 190 thousand Tangshan Rail Car Company high-speed trains running from 4 December to 22 December 2012. The unprocessed data include 202 attributes, which can be divided into train attributes, operational characteristics, and the predicted values of braking operations. The statistics of the level of brake handling during the operation of the high-speed trains are shown in Table 3.

Table 3. Statistics of the sample size of brake handling.

Handling Level	OC	REL	1A	1B	2	3
Sample size	71	174,626	4672	4526	5619	3966
Proportion (%)	0.036	88.958	2.380	2.306	2.862	2.020
Handling level	**4**	**5**	**6**	**7**	**8**	**EB**
Sample size	2065	367	143	144	100	3
Proportion (%)	1.052	0.187	0.073	0.073	0.051	0.002

During the operation of high-speed trains, the number of samples at each brake handling level varies greatly, and there are too few cases in which the brake handling level is "EB". This study focuses on the normal driving process. Therefore, a few data sets in this study are defined as cases in which the brake handling level is at "OC". The line containing all the braking level operations is used to verify whether the optimal control method proposed in this paper is effective. The specific line data are shown in Table 4, and the statistics of the brake handling levels in the simulation experiment data are shown in Table 5.

Table 4. Specific line data.

	Verified Line
Running distance	83,786 m
Running time	23 min 54 s
Maximum speed limit	300 km/h
Operating energy consumption	1747.175 kwh

Table 5. Statistics of brake handling level in simulation experiment data.

Handle Level	OC	REL	1A	1B	2	3
Sample size	2	1029	109	92	93	30
Proportion (%)	0.0014	72.06	7.63	6.44	6.51	2.1
Handle Level	**4**	**5**	**6**	**7**	**8**	**EB**
Sample size	40	23	1	7	2	0
Proportion (%)	2.8	1.61	7.0028	0.0049	0.0014	0

4.1.2. Parameter Setting

The number of balanced training sets extracted from big datasets and small datasets is 30. The k value determined by cross-validation is four, and we use the European distance function. The number of valid minority datasets obtained by filtering is 65. The VAGE model uses a two-layer GCN as the encoder and uses the Adam optimizer to train 300 iterations at a learning rate of 0.01. The embedding dimension of the VAGE is 32.

4.2. Simulation Process

Firstly, the static attributes that cannot describe the sample distribution are eliminated. The chi-square test is used to reduce the dimension of feature selection and eliminate the attributes that are not related to the classification problem. The average interpolation method is used for filling, and attributes such as train operation status and arrival distance are increased. Finally, each sample contains 51 attribute values. The first part of the strategy is used to obtain a balanced training set, and the attributes of these samples are extracted at the next moment according to the time stamp to illustrate the changes in the train operation caused by the braking level. In order to eliminate the influence of the graph direction on model training, the braking problem is divided into three perspectives according to the value of the energy-saving label l_{e1} under the current high-speed train S_t of the training set, as shown in Figure 5.

The center of each figure represents the node under the current status S_t of the high-speed train operation, and the possible situation for the next status S_{nt} is around the center node. The figure is based on the true statistics of all possible situations in the training set. Multiple VAGE models are used to reconstruct the braking relationship diagram of high-speed trains from each perspective, and supervised training is performed in combination with the braking value of the training set. In the process of verification, the energy-saving braking label at each moment is applied to the reconstruction map from the corresponding perspective in turn, and the edge with the highest probability of occurrence is selected as

the most likely operation of the current driver. The predicted value constitutes an initial strategy for train operation.
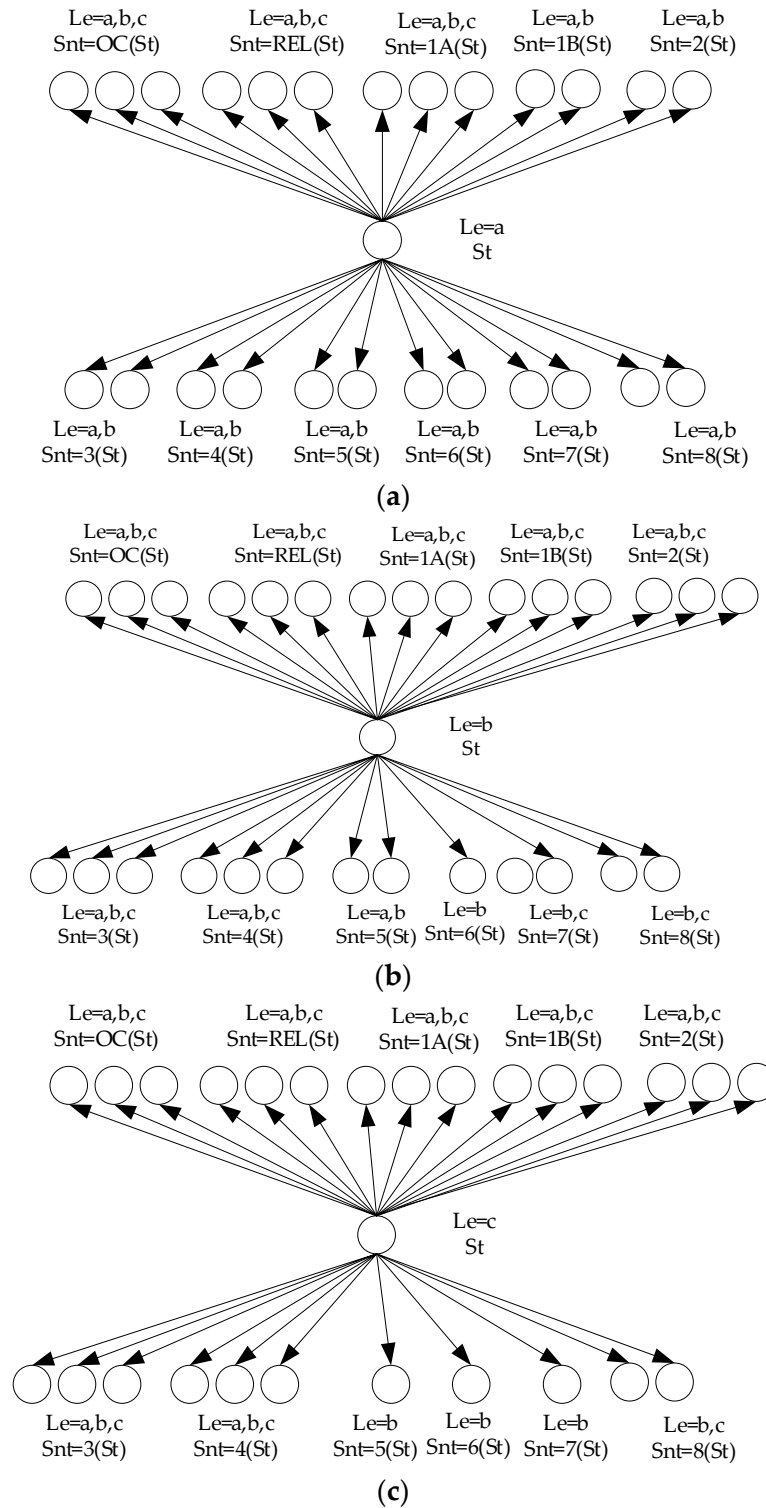

Figure 5. The structure of the specific diagram is constructed. (**a**) All braking conditions with current energy-saving label a. (**b**) All braking conditions with current energy-saving label b. (**c**) All braking conditions with current energy-saving label c.

4.3. Analysis of Simulation Results

Our analysis of the simulation results covers many aspects. The first is the evaluation of the performance of the algorithm, which includes a comprehensive evaluation of the performance of the proposed method in the experimental environment. The second is the evaluation of its energy-saving performance. Compared with the traditional method, a reduction in energy consumption and improvement in system energy efficiency are achieved. Finally, the proposed method is compared with other existing energy-saving methods to evaluate its advantages and disadvantages in practical applications so as to verify its effectiveness and practicability.

4.3.1. Evaluation of Algorithm Performance

In this paper, the AUC and loss are used to measure the performance of the model, and their changes with the number of iterations are shown in Figure 6 as follows:
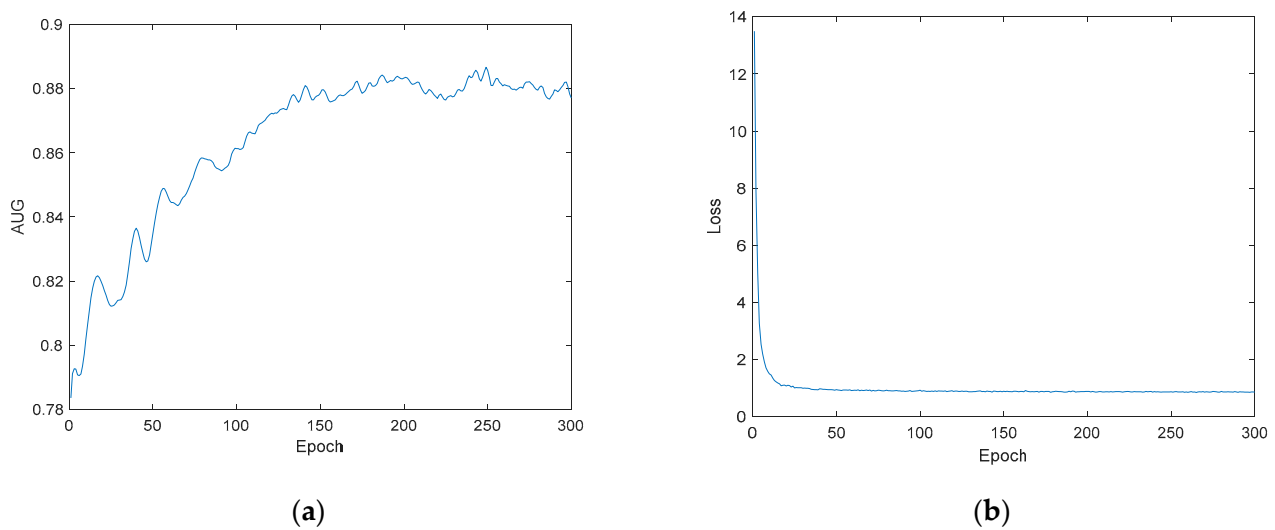

(**a**)   (**b**)

Figure 6. Comparison of AUC and loss of VGAE algorithm. (**a**) Changes in AUC. (**b**) Changes in loss.

The AUC value in the figure tends to be 0.9; the loss value gradually decreases with the increase in training rounds and finally stabilizes. This shows the validity of the model.

In addition, in order to further compare and illustrate our model, the confusion matrix is defined to describe the classification effects of multi-classification models, and P_i, R_i, and F_i are used to represent the accuracy, recall, and overall performance of model classification, respectively, as shown in (13)–(15):

$$P_i = \frac{n_{ii}}{\sum_{j=1}^{m} n_{ji}} \tag{13}$$

$$R_i = \frac{n_{ii}}{\sum_{j=1}^{m} n_{ij}} \tag{14}$$

$$F_i = \frac{2P_i R_i}{(P_i + R_i)} \tag{15}$$

Wherein n_{ii} is the number of correctly predicted samples in the class, and n_{ij} is the number of samples that incorrectly classify brake operation i as brake operation j. The results of P_i, R_i, and F_i are shown in Tables 6–8.

Table 6. Comparison of P_i calculation results.

	OC	REL	1A	1B	2	3	4	5	6	7	8
VGAE	0.008	0.96	0.17	0.39	0.63	0.95	0.24	0.65	0	0.09	0
NBM	0.01	0.98	0.54	0.31	1	0.33	0.23	1	0	0.08	0
BP	0.005	0.95	0	0.46	0.4	0.92	0.95	0.45	0	0.05	0

Table 7. Comparison of R_i calculation results.

	OC	REL	1A	1B	2	3	4	5	6	7	8
VGAE	0.5	0.88	0.15	0.34	0.15	0.6	0.53	0.25	0	0.39	0
NBM	0.5	0.86	0.28	0.53	0.22	0.33	0.8	0.17	0	0.29	0
BP	1	0.69	0	0.49	0.04	0.4	0.5	0.87	0	0.43	0

Table 8. Comparison of F_i calculation results.

	OC	REL	1A	1B	2	3	4	5	6	7	8
VGAE	0.01	0.92	0.16	0.36	0.24	0.74	0.33	0.36	0	0.15	0
NBM	0.02	0.91	0.37	0.39	0.36	0.33	0.36	0.29	0	0.13	0
BP	0.009	0.79	0	0.47	0.07	0.55	0.65	0.59	0	0.09	0

From the above table, it can be seen that the classification effect of the VGAE for traction control in this paper is not much different from that of the naive Bayesian and BP neural networks. The classification effect of REL braking is better than that of others because the number of samples of other classes in the validation section is too small. If the number of samples of each class in the validation set can reach a certain value, the accuracy, recall, and overall evaluation will be greatly improved. In addition, this study does not have requirements that are too strict on the accuracy of classification. It is reasonable that the predicted braking operation is inconsistent with the original operation, as we focus on determining the amount of energy saved.

4.3.2. Energy-Saving Performance Evaluation

The running curve of the final running strategy is shown in Figure 7. The simulated high-speed train reaches the speed limit when it goes from the starting point to more than 20,000 m, continues to cruise to more than 70,000 m at a speed close to the speed limit, and then begins to slow down. Compared with the original real running curve, the optimized traction control makes the change in speed smoother and ensures that the train runs within the speed limit range, which ensures safety. Based on the original manual driving data, the time error of two minutes is acceptable. The line data after the implementation of the strategy are shown in Table 9.

Table 9. Simulation circuit data.

	Travel Time	Travel Distance
Original run policy	23 min 54 s	83,786 m
Optimized operation strategy	22 min 30 s	83,786 m

In the table, the running time of the optimized driving scheme obtained in this study is 1 min 24 s faster than that of the original scheme. This can be explained by the fact that the strategy in this study achieves the punctuality and accurate parking of trains on the basis of safe driving, which are the preconditions for verifying energy saving.
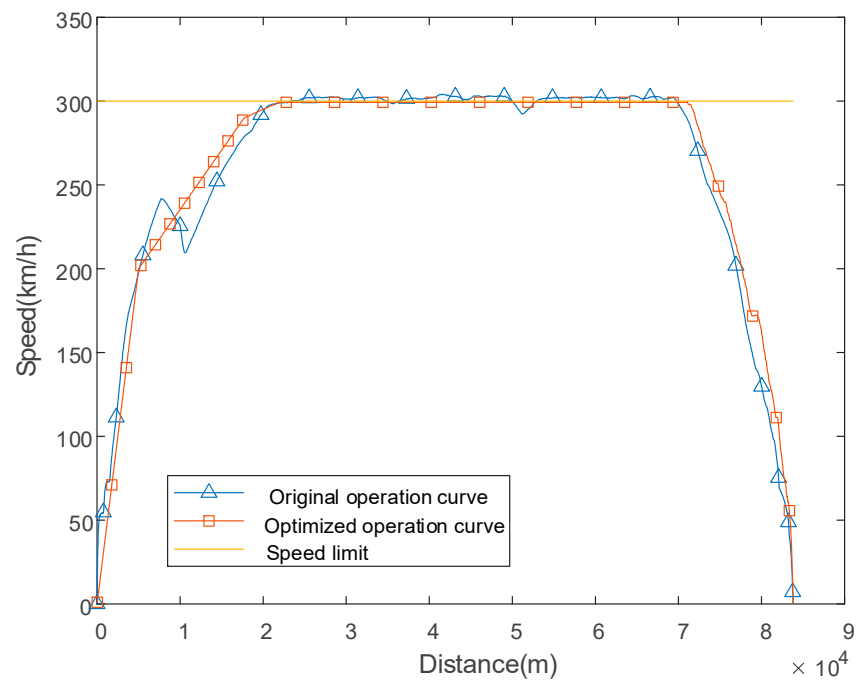
Figure 7. Running simulation.

The traction energy consumption of the train is a piecewise function, and the calculation needs to be segmented. Combined with whether the brake handling is switched and the speed changes, the energy consumption of different segments is counted, and the curve of the total energy consumption with the driving distance based on the strategy proposed in this paper is obtained, as shown in Figure 8.
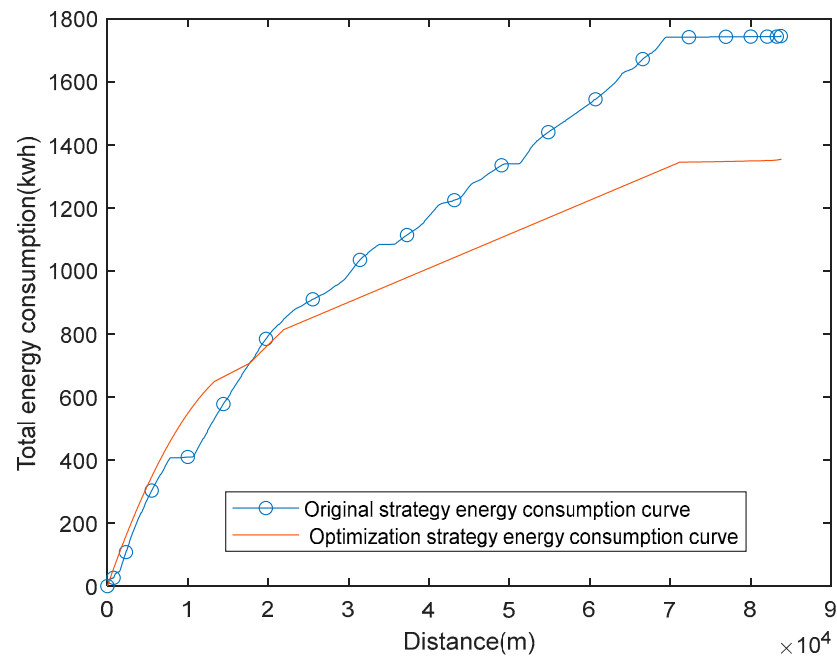

Figure 8. Relationship between total energy consumption and running distance comparison.

After the high-speed train goes 30,000 m in the diagram, the energy saving of this research strategy begins to show. For each interval of 10,000 m, the cumulative energy consumption of the original and the optimized strategy is shown in Table 10.

Table 10. Cumulative energy consumption comparison.

Travel Distance	Original Policy	Optimization Strategy	Energy Saving	Energy Saving Percentage
30,000 m	989.41 kwh	962.08 kwh	27.33 kwh	2.76%
40,000 m	1173.89 kwh	1069.72 kwh	104.17 kwh	8.87%
50,000 m	1339.81 kwh	1178.25 kwh	161.56 kwh	12.06%
60,000 m	1530.78 kwh	1285.89 kwh	244.89 kwh	15.99%
70,000 m	1714.92 kwh	1393.53 kwh	321.39 kwh	18.75%
83,786 m	1747.18 kwh	1418.98 kwh	328.2 kwh	18.78%

In Table 10, with the increase in driving distance, the energy-saving percentage in this study increases continuously, saving about 18.78% of the traction energy for the whole journey. In the normal driving environment, compared with the manual driving strategy, the optimized driving strategy in this study shows that some energy is saved.

4.3.3. Performance Comparison Test with Other Methods

In order to comprehensively evaluate the effect of the high-speed rail energy-saving strategy proposed in this paper, based on existing data, comparative experiments are designed to discover the performance differences between the VAGE method and the Jaccard and Adamic–Adar methods in depth. The energy-saving percentage of the traction energy consumption with the same distance interval under the VAGE, Adamic–Adar, and Jaccard methods compared with the original strategy is shown in Table 11.

Table 11. Energy saving percentage under different methods.

Distance	Original	VGAE		Adamic–Adar		Jaccard	
	Consumption	Consumption	Energy-Saving	Consumption	Energy-Saving	Consumption	Energy-Saving
30,000 m	989.4 kwh	984.13 kwh	0.5%	1078.39 kwh	—	1092.78 kwh	—
40,000 m	1173.89 kwh	1092.67 kwh	6.92%	1230.19 kwh	—	1200.42 kwh	—
50,000 m	1339.8 kwh	1201.27 kwh	10.34%	1337.83 kwh	0.15%	1308.06 kwh	2.37%
60,000 m	1530.78 kwh	1308.84 kwh	14.5%	1445.47 kwh	5.57%	1415.69 kwh	7.5%
70,000 m	1714.92 kwh	1417.38 kwh	17.35%	1553.11 kwh	9.44%	1511.43 kwh	11.87%
83,786 m	1747.18 kwh	1438.26 kwh	17.68%	1591.89 kwh	8.89%	1519.44 kwh	13.03%

The VGAE algorithm saves about 17.68% of the traction power throughout the simulated driving route and shows energy savings earlier than the other algorithms. A comparison of the speed and travel time between the proposed method and other existing methods is shown in Table 12.

Table 12. The comparison of speed and travel time.

Method	Figure Correlation (× Represents That the Method Doesn't Use It, and √ Represents the Opposite)	Using GNN (× Represents That the Method Doesn't Uses It, and √ Represents the Opposite)	Travel Time	Maximum Speed
Original	×	×	23 min 54 s	304.09 km/h
Adamic-Adar	√	×	23 min 58 s	299.42 km/h
Jaccard	√	×	24 min 30 s	299.82 km/h
VGAE	√	√	22 min 41 s	299.75 km/h

The maximum speed of the VGAE method does not exceed 300 km per hour, and the travel time is shorter than that of the other methods. The experimental results can verify the effectiveness of using a multi-view variational graph autoencoder to solve the

energy-saving problem of high-speed trains, that is, using less energy in a shorter period of travel.

5. Conclusions

This paper studies the energy consumption algorithm of high-speed train operations and proposes a graph optimization algorithm based on the VGAE, which can fully consider the complex relationship between nodes in the operation of trains. Through a simulation and comparative tests, under the premise of ensuring safe, on-time, and accurate parking, the control method in this study shows increased energy savings.

In the process of solving the driving strategy, this study balanced the sample size of each class of the training set and considered some special cases. A algorithm based on the KNN and undersampling is used to balance the negative impact of the number imbalance between the data sets, and the braking problem of the high-speed train is transformed into the link prediction problem of the implicit graph. The variational graph autoencoder is used to solve the initial strategy of train braking, and the time of intermediate trips is adjusted to ensure that the driving distance is consistent with the length of the route. The actual feasible braking application time and brake handling operation are obtained. However, the training data are based on relatively safe driving conditions, which makes the actual driving situation, especially in the face of a complex or dangerous driving environment, too idealistic. Therefore, for future research, we suggest inputting more train operation factors and real-time data into the VGAE model to obtain a more optimized strategy.

In detail, in order to further evaluate the proximity between research and practical applications, the following aspects need to be considered. First of all, the simulation data used in this study may not be able to fully simulate the various complex situations and changes during the operation of actual trains. The actual train operation environment may be affected by many factors, such as the weather, road conditions, traffic conditions, etc. These factors may have an important impact on the energy consumption of trains and the driving strategy of their drivers. Secondly, our research focuses on the safety, punctuality, and accuracy of trains, but in practical applications, safety is always the primary consideration. Therefore, any control method that optimizes energy consumption must ensure that the safety of the train is not sacrificed. Finally, the training data in the study are relatively restricted, which may limit the generalizability of the algorithm to actual driving environments. Therefore, incorporating more actual operating data and factors into the model should be the focus of future research to ensure the effectiveness and reliability of the algorithm in practical applications.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

Symbol	Symbolic Meaning	
S_t	Train operation state at decision moment t	S is state T is moment
$q(.)$	Map the train status to the braking operation that the drivers may perform	q is a mapping operation The input in the function is S_t
y_B	Specific brake operation level	y is the target forecast label B is brake
$y_B(.)$	Actions imposed by the drivers	The input in the function is S_t
S_{nt}	Train operation status at the next moment of t	nt is the next moment
l_e	Energy-saving label	l is label e is energy-saving
$P0$	Small datasets	P is a set
Pi	Big Datasets	P is a set, Class i, i = 1, 2,..., 11
$P_i l_{e1} l_{e2}$	When the current energy-saving label is l_{e1}, the drivers apply Class i brake handling to change the energy-saving label to the data set of l_{e2}	$l_{e1}, l_{e2} \in S_{la}$
$d_{x_i x_j}$	Similarity judgment of two nodes	d is the distance; subscripts represent the ith sample and the jth sample
$GCN(.)$	Graph convolution function	
W_0	Weight parameters to be learned	
W_1	Weight parameters to be learned	
A	Adjacency matrix	
$\check{A}$	Symmetric normalized matrix	
X	Node characteristic matrix	
D	Degree matrix	
$q(Z \mid X, A)$	Distributed calculation	
$p(A \mid Z)$	Probability calculation	
L	Value of the loss	

References

1. Julien, N.; Christian, S.; Christine, F.; Perrin, G. Optimization of train speed to limit energy consumption. *Veh. Syst. Dyn.* **2022**, *60*, 3540–3557.
2. Lei, Y.; Chen, Y. High-speed Railway Train Energy Driving Strategy Based on Improved Genetic Algorithm. In Proceedings of the 34th China Control and Decision-Making Conference, Hefei, China, 21–23 May 2022; p. 6. [CrossRef]
3. Ning, L.; Zhou, M.; Wu, W.; Zhang, Z.; Liu, C.; Dong, H. Train Trajectory Optimization for High-speed Railways under Constraints of Successive Trains. In Proceedings of the 2021 China Automation Congress (CAC), Beijing, China, 22–24 October 2021. [CrossRef]
4. Cao, Y.; Zhang, Z.; Cheng, F.; Su, S. Trajectory Optimization for High-Speed Trains via a Mixed Integer Linear Programming Approach. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 17666–17676. [CrossRef]
5. Su, S.; Zhu, Q.; Liu, J.; Tang, T.; Wei, Q.; Cao, Y. A data-driven iterative learning approach for optimizing the train control strategy. *IEEE Trans. Ind. Inform.* **2023**, *19*, 7885–7893. [CrossRef]
6. Zhu, Q.; Su, S.; Tang, T.; Xiao, X. Energy-efficient train control method based on soft actor-critic algorithm. In Proceedings of the 2021 IEEE Intelligent Transportation Systems Conference (ITSC), Indianapolis, IN, USA, 19–21 September 2021.
7. Zhang, L.; He, D.; He, Y.; Liu, B.; Chen, Y.; Shan, S. Real-time energy saving optimization method for urban rail transit train timetable under delay condition. *Energy* **2022**, *258*, 124853. [CrossRef]
8. Ying, P.; Zeng, X.; D'Ariano, A.; Pacciarelli, D.; Song, H.; Shen, T. Quadratically Constrained Linear Programming-based energy-efficient driving for High-speed Trains with neutral zone and time window. *Transp. Res. Part C Emerg. Technol.* **2023**, *154*, 104202. [CrossRef]

9. Havaei, P.; Sandidzadeh, M.A. Intelligent-PID controller design for speed track in automatic train operation system with heuristic algorithms. *J. Rail Transp. Plan. Manag.* **2021**, *22*, 100321. [CrossRef]
10. Wu, Z.; Pan, S.; Chen, F.; Long, G.; Zhang, C.; Philip, S.Y. A Comprehensive Survey on Graph Neural Networks. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *32*, 4–24. [CrossRef] [PubMed]
11. Luodi, X.; Huimin, H.; Qing, D. A Co-Embedding Model with Variational Auto-Encoder for Knowledge Graphs. *Appl. Sci.* **2022**, *12*, 715.
12. Nairouz, M.; Mohamed, B.; Riadh, K. A contrastive variational graph auto-encoder for node clustering. *Pattern Recognit.* **2024**, *149*, 149110209.
13. Zhou, F.; Yang, Q.; Zhong, T.; Chen, D.; Zhang, N. Variational Graph Neural Networks for Road Traffic Prediction in Intelligent Transportation Systems. *IEEE Trans. Ind. Inform.* **2021**, *17*, 2802–2812. [CrossRef]
14. Wentao, L.; Fan, L.; Luo, S.J.; Xiang, J.A. Multi-Site Anti-Interference Neural Network for ASD Classification. *Algorithms* **2023**, *16*, 315.
15. Xie, J.; Zhang, J.; Sun, K.; Ni, S.; Chen, D. Passenger and energy-saving oriented train timetable and stop plan synchronization optimization model. *Transp. Res. Part D* **2021**, *98*, 102975. [CrossRef]
16. Lv, M.; Hong, Z.; Chen, L.; Chen, T.; Zhu, T.; Ji, S. Temporal Multi-Graph Convolutional Network for Traffic Flow Prediction. *IEEE Trans. Intell. Transp. Syst.* **2020**, *22*, 3337–3348. [CrossRef]
17. Zhong, W.; Li, T.; Yuan, Q.; Xu, H. Cooperative optimal train operation algorithm for utilizing regenerative braking energy. *Appl. Math. Model.* **2024**, *127*, 172–192. [CrossRef]
18. Scheepmaker, G.M.; Goverde, R.M. Energy-efficient train control using nonlinear bounded regenerative braking. *Transp. Res. Part C Emerg. Technol.* **2020**, *121*, 102852. [CrossRef]
19. Atangulova, A.; Komshilov, K.; Sennikova, A.; Barbashov, N.; Shanygin, S.; Lobacheva, E. Method for Recovery of Vehicle Braking Energy with Electric Drive Powered by DC Network and Device for Its Implementation. *Transp. Res. Procedia* **2023**, *68*, 967–972. [CrossRef]
20. Liao, J.; Zhang, F.; Zhang, S.; Yang, G.; Gong, C. Energy-saving optimization strategy of multi-train metro timetable based on dual decision variables: A case study of Shanghai Metro line one. *J. Rail Transp. Plan. Manag.* **2021**, *17*, 100234. [CrossRef]
21. Mohamed, A.; Qiyuan, P. Ultra-Long-Distance High-Speed Railways Night Trains Operations: Differentiation Mode. *Transp. Res. Rec.* **2022**, *2676*, 366–370.
22. Liu, S.; Li, H.; Liu, Y.; Cheng, G.; Yang, G.; Wang, H.; Zheng, H.; Liang, D.; Zhu, Y. Highly accelerated MR parametric mapping by undersampling the k-space and reducing the contrast number simultaneously with deep learning. *Phys. Med. Biol.* **2022**, *67*, 185004. [CrossRef]
23. Wu, J.; Ling, L.; Zhou, K.; Zhou, K.; Zhang, E.; Wang, K.; Zhai, W. Coupler separation of slave locomotive in a 20,000-tonne combined heavy-haul train during air-braking release. *Veh. Syst. Dyn.* **2023**, *61*, 2761–2789. [CrossRef]
24. Collart-Dutilleul, S.; Bon, P.; Bougacha, R.; Laleauv, R. Engineering for Critical Systems: The Automatic Train Operation over European Train Control System for Freight Trains Use Case. *Int. J. Transp. Dev. Integr.* **2023**, *7*, 311–320. [CrossRef]
25. Hao, Z.; Wu, D.; Fang, Y.; Wu, M.; Cai, R.; Li, X. Collaborative lethal gene prediction algorithm based on supervised multi-view variational map autoencoder. *IEEE J. Biomed. Health Inform.* **2021**, *25*, 4041–4051. [CrossRef] [PubMed]
26. Jeongtae, S.; Dongsup, K. Applying network link prediction in drug discovery: An overview of the literature. *Expert Opin. Drug Discov.* **2023**, *19*, 43–56.