

Anomaly Detection for Vietnam Railway Using Unsupervised Learning Based on IoT Device Data Monitoring Railway Level Crossings

Ngo Anh Tuan[1], Nguyen Nhu Hai Linh[2], Pham Van Hiep[2], Vu Thu Diep[3], and Phan Duy Hung[1(✉)]

[1] FPT University, Hanoi, Vietnam
tuan23mse13141@fsb.edu.vn, hungpd2@fe.edu.vn
[2] Hanoi Railway Signal and Telecom, JSC, Hanoi, Vietnam
{linh.nnh,hiep.phv}@hasitec.vn
[3] HaNoi University of Science and Technology, Hanoi, Vietnam
diep.vuthu@hust.edu.vn

Abstract. In Vietnam, the railway industry has always been of interest to society, with its scale constantly expanding and upgrading. According to statistics, there are approximately 10,000 railway level crossings across the nation. At railway level crossings, system of protective equipment is deployed to ensure traffic safety. The operational data of these protective devices is collected and transmitted to the data center for monitoring and management. We developed and evaluated machine learning (ML) models to detect potential anomalies in crossing operations and identify potential safety risks. The system analyzes real-time sensor data including magnetic sensors, barrier status monitors, power supply indicators, and environmental sensors from multiple crossing locations. By using one class support vector machine (OCSVM) models to exploit the collected data, with the dual objectives of early detection of potential risks as well as the stable operation of the equipment we achieved 96% accuracy. This research proposes a methodology for improving the quality of management as well as the application of science and technology in the digital transformation and modernization of the railway industry in Vietnam.

Keywords: Railway Level Crossing · Railway Safety · Unsupervised Learning · SVM One Class · Anomaly Detection

1 Introduction

1.1 Problem and Motivation

At railway level crossings in Vietnam, a comprehensive system of devices and sensors is deployed to ensure safety operations. The core infrastructure includes six magnetic sensors strategically positioned along the tracks, these magnetic sensors function by counting the pulses generated each time a train wheel axle passes over the sensor location,

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2026
O. Gervasi et al. (Eds.): ICCSA 2025 Workshops, LNCS 15886, pp. 3–15, 2026.
https://doi.org/10.1007/978-3-031-97576-9_1

with two sensors placed one to five kilometers before the crossing point for early train detection, and two sensors at the intersection to confirm train presence. The system also incorporates barrier gates equipped with traffic light signals, powered by both AC and DC supplies to ensure operational reliability. The IoT infrastructure continuously collects various parameters from these devices (Fig. 1).

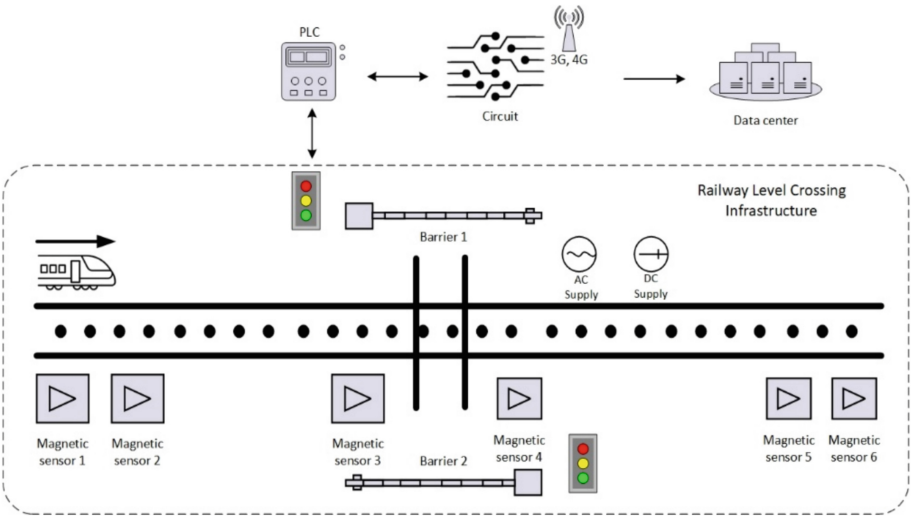

Fig. 1. System Architecture

The crossing protection devices are connected to a PLC (Programmable Logic Controller), which controls the operation of the entire infrastructure. Sensor data is sent to the PLC, and circuits collect it through indirect interfaces with the PLC. The processed data is then transmitted to the server through a Wi-Fi module, which connects to 3G/4G access points for reliable data transmission to centralized data centers.

Despite the centralized collection of sensor data from railway level crossings, its current application is limited to displaying information at monitoring centers, with incident response decisions still heavily dependent on human operators. This manual approach faces significant challenges as monitoring centers are responsible for supervising numerous level crossings simultaneously. The high volume of real-time data from thousands of crossing points creates a substantial cognitive load for operators, potentially leading to delayed response times and increased risk of oversight in critical situations.

The monitoring of level crossings faces numerous challenges due to the inability to comprehensively define all safety-affecting scenarios. The railway system extends throughout Vietnam with diverse operating conditions, and the message structure reveals over 40 different sensor parameters that must be monitored simultaneously. The combinations of these parameters create an extensive anomaly space, and many potentially dangerous situations have never occurred, making it impossible to pre-label them for traditional analysis methods.

Furthermore, the data exhibits complex temporal relationships. Train detection follows a specific sequence (approaching, passing through, and departing the crossing) that must trigger synchronized responses from barrier movements, light signals, and sound alerts. Any desynchronization in this sequence could indicate a potential safety risk that might be overlooked in manual monitoring.

We have recognized the potential to develop unsupervised learning models to detect these hidden risks. Unlike supervised learning methods that require extensively labeled training data, unsupervised learning can identify patterns and detect anomalies by leveraging the inherent structure of the data itself without explicit programming. This approach is particularly valuable for railway crossing safety where normal operational patterns can be learned, allowing deviations that may indicate potential safety issues to be automatically flagged for operator attention.

Therefore, there is a pressing need for automated systems that can not only collect data but also intelligently process and identify potential anomalies, supporting operators in making more timely and informed decisions.

1.2 Related Works

Statistics from 1990 to 2020 show that even countries with developed railway infrastructure such as China and Japan face significant challenges [1]. Both countries, along with many others [1, 2], continue to experience railway accidents due to environmental factors, poor management, and other challenges.

The complexity of railway safety management varies significantly across different economic contexts. Research on transport policy and management across low, middle, and high-income Asian countries demonstrated that while technological solutions are important, the effectiveness of safety measures heavily depends on the local context and management capabilities [3]. This is particularly relevant for Vietnam′s railway system, where the railway infrastructure is outdated compared to global standards.

In the context of railway management, proactive safety is always the first priority [4]. Traditional monitoring methods often fail to detect potential risks before they develop into accidents. Railway safety systems must integrate risk management with real-time monitoring capabilities, indicating that advanced technologies such as Machine Learning [5] can play a crucial role in enhancing safety measures. The integration of ML technologies offers promising solutions for enhancing railway safety systems through improved risk detection and management capabilities [6–8]. The deployment of IoT devices at these crossings generates an exceptionally large volume of data [9]. Machine learning excels at identifying subtle patterns in large, complex datasets that may not be apparent through traditional statistical analysis [10].

To build a machine learning model with high accuracy, data preprocessing and feature extraction play a crucial role, as sensors at railway level crossings typically have different measurement scales [11, 12]. Normalization methods such as min-max scaling or Z-score are particularly important for magnetic sensors used in train detection, where environmental conditions can significantly affect raw measurements, requiring normalization to maintain consistent detection thresholds across different locations and weather

conditions [13]. Additionally, dimensionality reduction helps significantly improve system performance. Methods such as Principal Component Analysis (PCA) or t-SNE are effective in reducing data complexity for railway systems with extensive data [14].

The One-Class Support Vector Machine (OCSVM) algorithm is particularly well suited to the problem of anomaly detection at railway crossings, as data in this area is often unlabeled. OCSVM is highly sensitive to rare and unusual data patterns [15], such as incidents with very low occurrence frequencies. This makes OCSVM particularly useful for safety monitoring at crossings, where severe anomalies may not have been present in historical data. However, the ability to detect these anomalies early is key to preventing accidents and ensuring safe operations.

2 Contribution

Our study proposes data filtering and normalization methods, thereby building a highly reliable training dataset that is suitable for the characteristics of the railway industry [26].

This research serves as a foundation for future approaches to processing data collected from devices at railway crossings via the Internet of Things (IoT) infrastructure. The proposed method demonstrates the ability to efficiently integrate and analyze multiple data streams, thereby supporting the early detection of potential safety-related risks and system operation anomalies.

This research offers valuable insights into the challenges and solutions associated with implementing smart monitoring systems in railway networks in developing countries where infrastructure and management capacity are limited. These insights are particularly important for countries facing similar issues in infrastructure development and modernization, providing a practical application framework that can be flexibly adapted to different operational contexts considering local technological and managerial constraints.

The remainder of this paper is organized as follows. Section 2 describes the methodology. Section 3 discusses experimental results. Finally, Sect. 4 provides conclusion and future works.

3 Methodology

Our method presents a comprehensive approach to detect anomalies at railway crossings. Raw data from IoT devices at these crossings is stored in our data center. The system processes this data through two parallel preprocessing threads:

- The first preprocessing thread prepares data specifically for the Modeling Module, where we employ a One-Class SVM machine learning model to establish reliable anomaly detection capabilities.
- The second preprocessing thread focuses on preparing data for real-time analysis, directly feeding into the Real-time Processing Module. This module leverages a sliding window technique to continuously sample incoming data, enabling immediate anomaly detection and prediction during railway crossing operations.

Our proposed overall architecture is illustrated as Fig. 2.

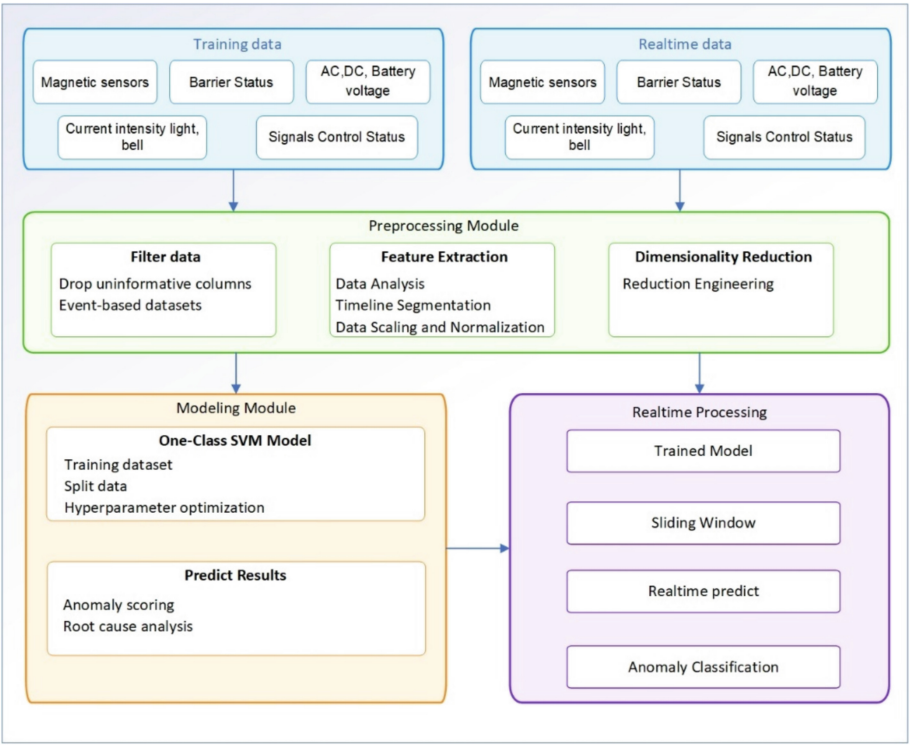

Fig. 2. Architecture overview

3.1 Preprocessing Module

Filter Data

Raw data will be filtered to remove unnecessary information fields. Each train passing through a level crossing will be a set containing valuable information collected from IoT devices over time. Information about these fields is described in Table 1.

Feature Extraction

The data will be normalized and feature vectors will be extracted from the dataset of each train crossing event. The methods implement:

Min-max scale. For magnetic sensor data, which increases gradually over time and has unchanging peak values, using the maximum and minimum values method will significantly improve model performance [16, 17]. With threshold is the train detection threshold the transformation formula is:

$$\text{X_normalized} = \frac{\text{X} - \text{Threshold}}{\text{X_max} - \text{Threshold}}$$

Peak Difference Analysis (PDA). PDA is typically used when we have two or more variables/series that are closely correlated. In our system, sensor pairs are deployed consecutively to backup each other, so monitoring the ΔX at peak values is critically

Table 1. Description of IoT Data

No	Field Name	Description	Value Range/Unit
1	time_stamp	Timestamp	ddMMyyyy_HHmmss
2	X1, X2 … X6	Magnetic sensor values	Integer values
3	S5	PLC control signal	0: Not operational 1: Ready 2: Controlling
4	S6, S7	Control signal for opening/closing barrier 1	Same as above
5	S9, S10	Control signal for opening/closing barrier 2	Same as above
6	C7, C8	Open/closed status of barrier 1	0: Undetermined 1: Opened/Closed
7	C13, C14	Open/closed status of barrier 2	Same as above
8	AI1	AC power voltage	220 V
9	AC2	DC power voltage	24 V
10	I1	Warning light current intensity	Blink. 0 to 2000mA
11	I3, I4	Yellow lights current intensity	Same as above
12	I5, I6, I7, I8	Red lights current intensity	Same as above
13	I9	Bell current intensity	Always on

important. By extracting peak features and calculating the differences, we can highlight sudden deviations that models based on raw values might overlook:

$$\Delta X = \frac{\left|\max(X_norm_i) - \max(X_norm_{i-1})\right|}{|t_i - t_{i-1}|}$$

t_i and t_{i-1} are the timestamps when peaks occur at sensors i and i-1.

This is especially useful for detecting anomalies [18] related to irregular train speeds.

Temporal Feature Extraction. Identifying time markers to include in the feature vector is very significant [19]. For our problem, there is a strong relationship between sensor values and timestamps.

For example, when a train is detected entering the level crossing (X1 > Threshold), the time interval from when the barrier closure control signal is given until the barrier is completely closed has a major impact on safety. We will analyze and incorporate these time intervals Δt into the feature vector:

$$t_{close} = \text{time}(C_{closed}) - \text{time}(S_{signal_close})$$

Z-score. Is a powerful tool that can significantly improve the performance of models in anomaly detection tasks [20], especially when combined with min-max scaling, peak difference analysis, and Temporal Feature Extraction [21].

Apply Z-Score for each phase: train entering, train passing through, and train leaving the level crossing to enhance the sensitivity of anomaly detection when trains are running:

$$Z(X_{i,t}) = \frac{X_{i,t} - \mu_{i,t}}{\sigma_{i,t}}$$

- $X_{i,t}$ value of sensor X_i at time t in the train event
- $\mu_{i,t}$ $\sigma_{i,t}$ are mean value and standard deviation of sensor X_i at time t across multiple train events.

For example, to detect early mechanical degradation of the barrier system before failures occur, we can apply the formula.

$$Z(t_{close}) = \frac{t_{close} - \mu_{t_{close}}}{\sigma_{t_{close}}}$$

$\mu_{t_{close}}$ and $\sigma_{t_{close}}$ are the mean value and standard deviation of the barrier closing time.

Dimensionality Reduction
In this phase, after the data has been normalized using the aforementioned methods, we will proceed with dimensionality reduction to make the dataset more concise while still preserving the data characteristics. Common dimensionality reduction methods such as PCA (Principal Component Analysis), t-SNE (t-Distributed Stochastic Neighbor Embedding) or UMAP (Uniform Manifold Approximation and Projection) will be applied in this step

3.2 Modeling Module

At this phase, data is classified and anomaly levels are evaluated through unsupervised learning methods, where the One-Class SVM(OCSVM) model is selected to build the system. We choose OCSVM for anomaly detection in railway systems because:

- The data collected from railway systems are mainly normal operating cases, while anomalies are rare. OCSVM is specially designed for this imbalanced data, where normal data samples are the majority.
- One-class SVM can detect nonlinear relationships between features. This is important in this problem because the relationships between parameters such as magnetic sensor, Δt, and the states of the barriers are often nonlinear.

To implement One-Class SVM, the model is trained on a dataset representing normal operations to determine the boundary of standard behavior. The RBF (Radial Basis Function) kernel technique is used to better capture non-linear relationships between features. Hyperparameter tuning is applied to find optimal values for the hyperparameter pair (ν, γ), helping to balance between model sensitivity and specificity. Finally, a decision function is implemented to classify new observations, determining whether they belong to normal or anomalous behavior.

To evaluate the model, the first step is to extract normal operational data from various level crossings for training. Next, a synthetic anomaly dataset is generated with varying

anomaly occurrence rates to assess the model's detection performance. Performance is then measured using precision, recall, F1-score, and AUC-ROC. Finally, a hyperparameter sensitivity analysis is conducted to ensure the model's stability and generalization capability

3.3 Realtime Processing

To deploy and apply trained models in production system, there is a huge challenge: there will be many incidents that are only detected when the train reaches a specific time point in the train running event chain. For each stage, we can detect different abnormalities. To solve this problem, we have implemented a method that combines Sliding Window and Dynamic Template.

Sliding Windows
Sliding Window technique plays an important role in detecting abnormalities by analyzing multivariate data series from real stages for the railway system, through analyzing multivariate data series from IoT points. This method is based on a number of defined technical characteristics, allowing to track statistical characteristics and correlations between sensors in each specific stage of the train passing event.

We can evaluate the similarity between windows and the standard template [22] using the formula:

$$\text{Eros(A, B, w)} = \sum_{i=1}^{n} w_i |\langle a_i, b_i \rangle| = \sum_{i=1}^{n} |\cos \theta_i|$$

$\langle a_i, b_i \rangle$ is the inner product between a_i and b_i,
w_i is a weight vector which is based on the eigenvalues of dataset.
$\cos \theta_i$ is the angle between a_i and b_i.

Dynamic Template
Applying the Dynamic Template allows early detection of anomalies without waiting for the event to complete in order to issue an alert [23]. Real-time data of train crossing events are continuously updated, while the standard template automatically adjusts when gradual changes occur. As a result, the system does not need to store the entire historical data, helping to optimize computational resources.

4 Experiments and Results

4.1 Data Collection

Our data was collected from the Hanoi-Hai Phong railway route, one of the busiest railway corridors in northern Vietnam. We selected 8 consecutive level crossings with the most stable operations to collect data over a 5-day period (from March 1 to March 5, 2025). These crossings were chosen because they have reliable infrastructure, creating favorable conditions for the research process.

For each day, the raw data is stored as a text file. After filtering the data, our dataset is transformed into CSV files, with a summary presented in Table 2.

Table 2. Summary of dataset

Total CSV files	40
Total train sessions	404
Total data rows	30561
Rows per event	75.46

4.2 Experiments

Since performance evaluation in unsupervised learning is crucial, all our collected data consists of train events operating under normal conditions (without incidents). This creates a major challenge known as class imbalance [24] in machine learning. To address the lack of incident data samples, we used historical error data, combined with simulated scenarios developed by railway experts, and integrate these samples into the original dataset. Table 3 summarizes the types of anomalies we simulated and the implementation methods used to generate these anomalies in our dataset.

Table 3. Anomaly types and implementation methods

Anomaly Type	Implementation Method
Barrier Malfunction	Extend closing time
Signal Control Delay	Increased time intervals
Power Instability	Voltage fluctuation
Sensor Detection Failure	Set sensor pair value

Barrier Malfunction errors are defined by the following process: find the first data point that satisfies the condition $C7 = 0$ and $C8 = 1$, then mark the next 10 consecutive records with values $C7 = 0$ and $C8 = 0$. The barrier closure time will be extended by the duration corresponding to the next 10 records, which results in a Barrier Malfunction.

The time from when the number of pulses counted by X1 or X6 exceeds the detection threshold (usually 7 pulses) until $S7 = 1$ and $S10 = 1$ is defined as the system reaction time. To generate the Signal Control Delay error, we update the value at the first time $S7 = 1$ and the next two records to 0, because this time interval is usually very short.

AC power typically has a threshold of 220V and DC power with a threshold of 24V usually operates very stably with a 5% error margin. We modify a few random records with a deviation higher than 10% to create the Power Instability error.

The magnetic sensors count pulses generated each time a train wheel axle passes, and at peak operation, the variation between sensors typically remains within an expected range of values. To simulate a Sensor Detection Failure scenario for a train event, we will set all X1 values after the pulse count reaches 30 to a fixed value of 30. Figure 3 illustrates the variation in magnetic sensor values for a normal train event and a simulated faulty train event.

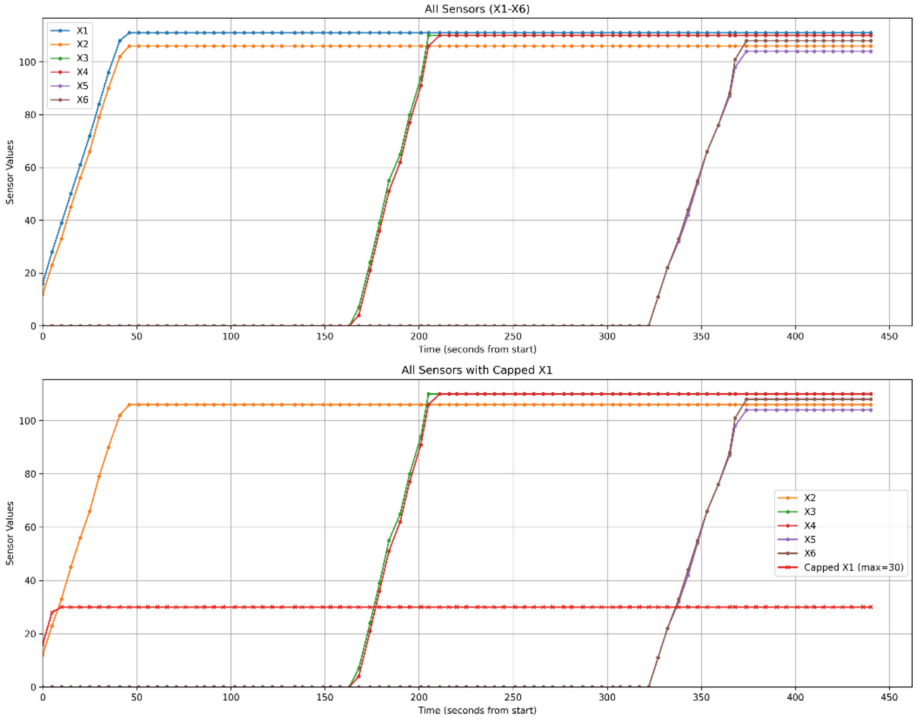

Fig. 3. Magnetic sensor values for a normal train event and a simulated faulty train event

4.3 Results

The characteristics of the dataset were collected under normal conditions, including 400 train events as described in Table 2. We divided the data into a training set and a testing set at a 50% ratio, with each set containing approximately 200 train events. From the training set, we proceeded to train the OCSVM model.

To evaluate the model based on metrics such as Accuracy, Precision, Recall, and F1-Score, our approach involves mixing different error ratios with our testing dataset, introducing anomalies at various proportions: 20% (40 samples), 10% (20 samples), 5% (10 samples), and 2% (4 samples) in a dataset of 200 samples. These ratios help evaluate our model′s performance under both high anomaly density scenarios and more realistic conditions where anomalies are relatively rare. Testing with such varied ratios helps us gain deeper insights into how the model responds in different scenarios, while also helping identify the sweet spot [25] the point at which the model maintains effectiveness despite very low anomaly rates. This is particularly important in railway safety applications, where the cost of missing an incident (false negative) can be extremely high. Table 4 describes the metrics of the OCSVM model with default parameters $\nu = 0.5$, $\gamma = \text{scale}$, kernel = rbf.

To increase the model′s performance, we performed hyperparameter tuning as shown in Table 5. We achieved very good results during experimentation. Table 6 describes the results after tuning with $\nu = 0.01$, $\gamma = \text{scale}$, kernel = linear.

Table 4. OCSVM default hyper params metrics

Sample error	Accuracy	Precision	Recall	F1-Score	ROC AUC
40	0.7865	0.8974	0.8642	0.8805	0.6173
20	0.7765	0.9429	0.8148	0.8742	0.5185
10	0.7952	0.9706	0.8148	0.8859	0.5185
4	0.8293	0.9855	0.8395	0.9067	0.5309

Table 5. Tuning OCSVM hyper params

Parameter	Tuning Range
nu	0.01, 0.05, 0.1, 0.2
gamma	scale, auto, 0.1, 0.01
kernel	rbf, linear

Table 6. Best metrics of OCSVM model

Sample error	Accuracy	Precision	Recall	F1-Score	ROC AUC
40	0.8989	0.9091	0.9877	0.9467	0.8519
20	0.9294	0.9518	0.9753	0.9634	0.8519
10	0.9398	0.9750	0.9630	0.9689	0.7160
4	0.9634	0.9875	0.9753	0.9814	0.8519

5 Conclusion and Future Works

This research has successfully developed an anomaly detection subsystem for railway crossings in Vietnam, applying unsupervised learning on IoT sensor data. The implementation of unsupervised learning not only helps reduce manual classification efforts but also enhances detection efficiency based on large-scale datasets.

This work establishes a foundation for future intelligent monitoring systems that can be adapted to different operational contexts while maintaining high reliability and performance. Our dataset, extracted in CSV format, can serve as a valuable reference resource for the community [26], supporting more effective research and development of models. This work can be applied to machine learning and data mining problems with tabular data [27–30].

In the future, there is significant potential for expanding this research. For instance, the system can be upgraded to monitor detailed train journeys as they pass through crossings. Additionally, after the preliminary classification, integrating another machine learning model to specifically identify each type of incident will help obtain accurately

labeled data. This data, in turn, can be used to improve and retrain the original model, enhancing the reliability and efficiency of the entire system.

References

1. Cao, Y., An, Y., Su, S., Xie, G., Sun, Y.: A statistical study of railway safety in China and Japan 1990–2020. Accid. Anal. Prev. **175**, 106764 (2022)
2. Kyriakidis, M., Hirsch, R., Majumdar, A.: Metro railway safety: an analysis of accident precursors. Saf. Sci. **50**(7), 1535–1548 (2012)
3. Shah, S.A.R., Ahmad, N., Shen, Y., Pirdavani, A., Basheer, M.A., Brijs, T.: Road safety risk assessment: an analysis of transport policy and management for low, middle, and high-income Asian countries. Sustainability **10**(2), 389 (2018)
4. Bulakh, M., Okorokov, A., Baranovskyi, D.: Risk System and Railway Safety. IOP Conf. Ser. Earth Environ. Sci. **666**, 042074 (2021)
5. Alawad, H., Kaewunruen, S., An, M.: Learning from accidents: machine learning for safety at railway stations. IEEE Access **8**, 633–648 (2020)
6. Li, H., et al.: Improving rail network velocity: a machine learning approach to predictive maintenance. Transp. Res. Part C: Emerg. Technol. **45**, 17–26 (2014)
7. Thaduri, A., Galar, D., Kumar, U.: Railway assets: a potential domain for big data analytics. Procedia Comput. Sci. **53**, 457–467 (2015)
8. Paltrinieri, N., Comfort, L., Reniers, G.: Learning about risk: machine learning for risk assessment. Saf. Sci. **118**, 475–486 (2019)
9. Singh, P., Elmi, Z., Meriga, V.K., Pasha, J., Dulebenets, M.A.: Internet of things for sustainable railway transportation: past, present, and future. Cleaner Logistics Supply Chain **4**, 100065 (2022). https://doi.org/10.1016/j.clscn.2022.100065
10. Siddiqui, H.U.R., Saleem, A.A., Raza, M.A., Zafar, K., Munir, K., Dudley, S.: IoT based railway track faults detection and localization using acoustic analysis. IEEE Access **10**, 106520–106533 (2022). https://doi.org/10.1109/ACCESS.2022.3210326
11. Islam, U., et al.: A novel anomaly detection system on the internet of railways using extended neural networks. Electronics **11**(18), 2813 (2022). https://doi.org/10.3390/electronics11182813
12. Inan, M.S.K., Liao, K., Shen, H., Jayaraman, P.P., Georgakopoulos, D., Tang, M.J.: DeepHeteroIoT: deep local and global learning over heterogeneous IoT sensor data. proceedings of the 20th EAI international conference on mobile and ubiquitous systems: computing, networking and services (2023)
13. Kumar, N., Krause, L., Wondrak, T., Eckert, S., Eckert, K., Gumhold, S.: Robust reconstruction of the void fraction from noisy magnetic flux density using invertible neural networks. Sensors **24**(4), 1213 (2024). https://doi.org/10.3390/s24041213
14. Lasisi, A., Attoh-Okine, N.: Hybrid rail track quality analysis using nonlinear dimension reduction technique with machine learning. Can. J. Civ. Eng. **48**(12), 1713–1723 (2021). https://doi.org/10.1139/cjce-2019-0832
15. Yang, K., Kpotufe, S., Feamster, N.: An efficient one-class SVM for anomaly detection in the internet of things. arXiv. https://arxiv.org/abs/2104.11146 (2021)
16. Mejri, N., Lopez-Fuentes, L., Roy, K., Chernakov, P., Ghorbel, E., Aouada, D.: Unsupervised anomaly detection in time-series: an extensive evaluation and analysis of state-of-the-art methods. Expert Syst. Appl. **256**, 124922 (2024)
17. Zhang, M., Xu, B., Wang, D.: An anomaly detection model for network intrusions using one-class SVM and scaling strategy. In: Guo, S., Liao, X., Liu, F., Zhu, Y. (eds.) Collaborative Computing: Networking, Applications, and Worksharing, pp. 280–291. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-28910-6_24

18. Lee, B.S., Kaufmann, J.C., Rizzo, D.M., Haq, I.U.: peak anomaly detection from environmental sensor-generated watershed time series data. In: Lossio-Ventura, J.A., Valverde-Rebaza, J., Díaz, E., Alatrista-Salas, H. (eds.), Information Management and Big Data, pp. 156–170. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-35445-8_11
19. Zhao, P., Chang, X., Wang, M.: A novel multivariate time-series anomaly detection approach using an unsupervised deep neural network. IEEE Access **9**, 109025–109041 (2021). https://doi.org/10.1109/ACCESS.2021.3101844
20. Tutivén, C., Vidal, Y., Insuasty, A., Campoverde-Vilela, L., Achicanoy, W.: Early fault diagnosis strategy for WT main bearings based on SCADA data and one-class SVM. Energies **15**(12), 4381 (2022). https://doi.org/10.3390/en15124381
21. Salam, A., et al.: Securing smart manufacturing by integrating anomaly detection with zero-knowledge proofs. IEEE Access **12**, 36346–36360 (2024)
22. Zhang, M., Xu, B., Gong, J.: An anomaly detection model based on one-class SVM to detect network intrusions. In: Proceedings of the 11th International Conference on Mobile Ad-hoc and Sensor Networks (MSN), pp. 102–107. IEEE (2015)
23. Lee, S., Park, D.: A real-time abnormal beat detection method using a template cluster for the ECG diagnosis of IoT devices. HCIS **11**(04) (2021). https://doi.org/10.22967/HCIS.2021.11.004
24. Hosseini, S.M., Shafique, A., Babaie, M., Tizhoosh, H.R.: Class-imbalanced unsupervised and semi-supervised domain adaptation for histopathology images. In: Proceedings of the 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pp. 1–7. IEEE (2023)
25. Rufino, V.Q., et al.: Improving predictability of user-affecting metrics to support anomaly detection in cloud services. IEEE Access **8**, 198152–198167 (2020)
26. Tuan, N.A.: Vietnamese railway crossing IoT sensor data. Kaggle (2023). https://www.kaggle.com/datasets/ngtuan12/vietnamese-railway-crossing-iot-sensor-data
27. Hung, P.D.: Detection of central sleep apnea based on a single-lead ECG. In: Proceedings of the 5th International Conference on Bioinformatics Research and Applications (ICBRA '18). Association for Computing Machinery, New York, NY, USA, pp. 78–83 (2018)
28. Hung, P.D.: Central sleep apnea detection using an accelerometer. In: Proceedings of the 1st International Conference on Control and Computer Vision (ICCCV '18). Association for Computing Machinery, New York, NY, USA, pp. 106–111 (2018)
29. Dat, D.Q., Hung, P.D.: Improvement for time series clustering with the deep learning approach. In: Luo, Y. (eds.) Cooperative Design, Visualization, and Engineering. CDVE 2021. Lecture Notes in Computer Science(), vol. 12983. Springer, Cham (2021)
30. Minh, N.Q., Hung, P.D.: the system for detecting vietnamese mispronunciation. In: Dang, T.K., Küng, J., Chung, T.M., Takizawa, M. (eds.) Future Data and Security Engineering. Big Data, Security and Privacy, Smart City and Industry 4.0 Applications. FDSE 2021. Communications in Computer and Information Science, vol. 1500. Springer, Singapore (2021)