A Review On Statistical Models For Anomaly Detection Using Metro Turnout Data

Ramavtar Singh Sikarwar[1], Dr. Sunil Sugandhi [2]
[1, 2] Dept of Civil Engineering
[1, 2] JIT, Borawan, India

Abstract- *Metro railway systems are among the most important modes of urban transportation, providing safe, reliable, and efficient mobility for millions of passengers every day. One of the most critical components of a metro railway network is the turnout or switch system, which enables trains to change tracks and facilitates smooth traffic management. Since turnouts are subjected to continuous mechanical stress and environmental variations, they are vulnerable to wear, degradation, and unexpected failures. Therefore, timely fault detection in metro turnout data is essential to ensure operational safety, minimize service disruptions, and reduce maintenance costs.A turnout system consists of several mechanical and electrical components, including switch rails, point machines, locking mechanisms, motors, and sensors. During operation, these components generate various forms of data such as current signals, voltage measurements, vibration signatures, temperature readings, and switching times. The analysis of these data enables maintenance engineers to assess the health condition of the turnout and identify abnormalities that may indicate the onset of faults. Effective fault detection systems can provide early warnings and prevent catastrophic failures. This paper presents a review of existing statistical models in the domain of research.*

Keywords: Statistical Models, Metro Turnout, Imbalanced Datasets, Machine Learning, Classificaiton Error.

I. INTRODUCTION

Traditional fault detection methods primarily rely on manual inspections and rule-based diagnostic techniques. These approaches are labor-intensive, time-consuming, and often incapable of identifying incipient faults [1]. In many cases, maintenance activities are performed periodically rather than based on the actual condition of the turnout components. Such preventive maintenance strategies may either lead to unnecessary maintenance expenses or fail to prevent sudden failures. Consequently, there is a growing need for intelligent and automated fault detection mechanisms [2].

With the advancement of sensor technologies and data acquisition systems, large amounts of turnout operational data can now be collected continuously. This development has encouraged the adoption of condition monitoring and predictive maintenance approaches. By analyzing historical and real-time data, engineers can determine the degradation trends of turnout components and estimate their remaining useful life. These data-driven approaches significantly improve maintenance planning and enhance system reliability [3].

The traffic planning of railways primarily relies on controlling turnouts. Turnouts are essential elements of railway systems because they allow track switching and enable trains to travel on different routes. However, these turnouts may encounter failures due to various factors, disrupting the transportation process. Therefore, the proper functioning of turnouts is crucial to maintaining the safety and efficient operation of trains, as well as to enhancing transportation efficiency [4].

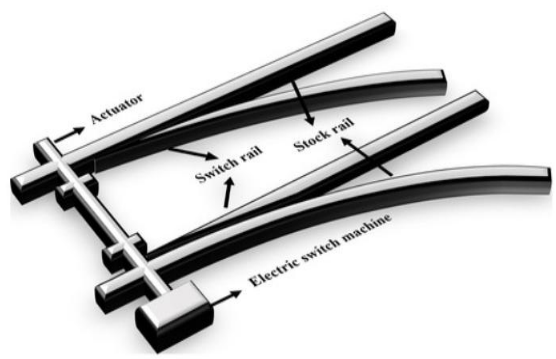

Fig. Turnout Structure

Figure 1 depicts the turnout schematic, which has the stock rails, switch rails, and switch machines.

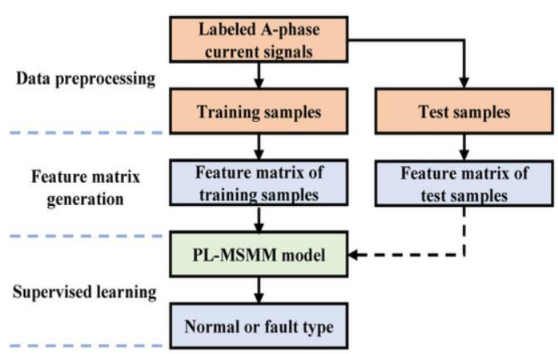

Fig.2Generic Model for Fault Detection from Metro Turnout Data

Figure 2 depicts the Generic Model for Fault Detection from Metro Turnout Data. The intelligent maintenance of railway equipment has garnered increasing attention as a way to enhance sustainable transportation and manufacturing. As an essential topic in prognostics and health management (PHM), fault diagnosis can help reduce the workload for inspectors and enhance the efficiency of traditional regular inspections [5].

II. APPLICATIONS OF INDUSTRY 4.0 IN METRO TRANSPORT

Statistical machine learning techniques have emerged as powerful tools for fault detection in metro turnout systems. Algorithms such as Support Vector Machines, Random Forests, Decision Trees, K-Nearest Neighbors, and Gradient Boosting can learn patterns associated with normal and faulty operating conditions. These methods are capable of handling large datasets and identifying hidden relationships among multiple variables. Machine learning-based approaches provide higher accuracy and faster fault classification compared with conventional methods [6].

Deep learning techniques have further enhanced the capability of fault diagnosis systems. Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, Autoencoders, and Transformer architectures can automatically extract complex features from raw sensor signals without extensive manual feature engineering [7]. These models are particularly effective in capturing temporal dependencies and nonlinear relationships in turnout data, leading to improved anomaly detection and fault classification performance.Signal processing methods also play a crucial role in metro turnout fault detection. Techniques such as wavelet transforms, Fourier transforms, empirical mode decomposition, and current curve analysis are widely employed to extract meaningful information from electrical and vibration signals. These methods help identify subtle variations in signal characteristics that may indicate mechanical wear, motor degradation, or abnormal switching behavior [8]..

Recent research has focused on integrating artificial intelligence with Internet of Things (IoT) technologies and cloud computing platforms for intelligent turnout monitoring. Smart sensors and wireless communication systems enable continuous collection and transmission of operational data to centralized monitoring systems. The application of big data analytics and edge computing facilitates real-time fault detection and supports predictive maintenance strategies. Such intelligent monitoring frameworks contribute to improved reliability and reduced downtime in metro operations [9].

Despite significant advancements, several challenges remain in fault detection for metro turnout systems. These include the scarcity of labeled fault data, class imbalance problems, noisy sensor measurements, and the difficulty of interpreting complex deep learning models. Furthermore, variations in environmental conditions and operational characteristics can affect the performance of diagnostic algorithms. Addressing these challenges requires the development of robust, explainable, and adaptive fault detection techniques [10]..

This imbalance also complicates evaluation, since metrics like accuracy become misleading for minority class detection.

III. EXISTING WORK IN THE DOMAIN

This section presents the summary of literature review in the domain:

Reetz et al.[11]proposed an expert system-based framework for diagnosing faults in railway point machines using Bayesian networks. Their approach utilized probabilistic reasoning to model the dependencies among various fault conditions and system components. The Bayesian network enabled efficient handling of uncertainty and incomplete information, thereby improving the reliability of fault identification and supporting predictive maintenance strategies for railway infrastructure.

Huang et al. [12] investigated turnout fault diagnosis through the application of the DT7 model combined with template matching techniques. Their methodology focused on analyzing characteristic signatures obtained from turnout operation data and comparing them with predefined fault templates. The study demonstrated that template matching could effectively distinguish between normal and abnormal operating conditions, providing a practical solution for turnout maintenance.

Bian et al. [13]introduced a degradation state mining and identification method for railway point machines. By analyzing degradation patterns from operational data, the authors established different health states of the turnout system and proposed mechanisms for identifying the progression of deterioration. Their work highlighted the importance of condition-based maintenance and emphasized the need for continuous monitoring to enhance system reliability.

Haro et al. [14] employed variational autoencoders for sensor fault detection. The proposed deep learning approach learned latent representations of sensor behavior and identified anomalies through reconstruction errors. The results demonstrated the capability of unsupervised learning methods in detecting failures without requiring extensive labeled datasets, making the technique suitable for complex railway monitoring environments.

Tang et al. [15] presented a comprehensive review of artificial intelligence applications in railway transportation. Their survey covered machine learning, deep learning, computer vision, and predictive analytics techniques employed in railway monitoring and fault diagnosis. The study emphasized the growing significance of intelligent maintenance systems and identified several research challenges, including data imbalance, interpretability, and real-time implementation.

Fu et al. [16] developed a Convolution-Transformer-based model for railway turnout fault diagnosis. The hybrid architecture combined convolutional neural networks for local feature extraction with Transformer networks for capturing long-range dependencies in sequential data. Experimental results showed improved diagnostic accuracy and robustness compared with conventional deep learning methods, demonstrating the potential of attention mechanisms in railway applications.

Wang [16] explored image-based inspection techniques for detecting turnout geometry defects and structural abnormalities. Their work utilized computer vision algorithms to process visual information acquired from cameras and image sensors. The study revealed that image-based approaches provide non-contact and automated inspection capabilities, reducing human intervention and enhancing maintenance efficiency.

Akula et al. [17] investigated acoustic signal-based damage detection in railway tracks using signal processing and pattern recognition techniques. By analyzing sound signatures generated during train operations, the authors identified structural anomalies and defects with considerable accuracy. Their approach demonstrated the feasibility of utilizing acoustic sensing as an economical and non-invasive solution for railway condition monitoring.

Ou et al. [18] proposed early hybrid detection techniques based on current curves and waveform analysis for railway turnouts. Their methodology combined electrical current analysis with signal waveform characteristics to detect abnormalities in point machine operations. The study established the effectiveness of hybrid diagnostic techniques in providing early warning indicators and minimizing unexpected failures.

Yang et al. [19] examined deep learning approaches for anomaly detection in turnout monitoring systems. The authors reviewed various neural network architectures, including convolutional and recurrent networks, and discussed their ability to automatically learn discriminative features from large datasets. Their findings indicated that deep learning models significantly improve fault detection performance and offer promising solutions for intelligent railway maintenance systems.

Ou et al. [20] proposed an online classification framework for fault diagnosis of railway turnouts using sensor data acquired during operation. The developed method enabled real-time fault classification and continuous monitoring of turnout conditions. Experimental results confirmed that the approach provided rapid and accurate fault detection, thereby enhancing the safety and operational reliability of railway systems.

IV. RESEARCH GAP

The research gap in the domain is presented here:

Although significant progress has been achieved in the field of fault detection and condition monitoring of railway and metro turnouts, several limitations and research challenges still exist. These gaps provide opportunities for the development of more efficient and intelligent diagnostic systems [21].

1. Limited Availability of Fault Data

Most existing studies are based on small datasets or laboratory-generated fault scenarios. Real-world metro turnout failures occur infrequently, resulting in highly imbalanced datasets and limited availability of labeled fault samples. Consequently, the generalization capability of many machine learning and deep learning models remains inadequate.

2. Dependence on Single-Modal Data

Several approaches rely only on a single source of information, such as current signals, vibration signals, acoustic data, or images. However, turnout failures are complex phenomena involving mechanical, electrical, and environmental factors. The integration of multiple sensor modalities has not been sufficiently explored, limiting the comprehensiveness of existing fault diagnosis systems.

3. Insufficient Real-Time Implementation

Although many algorithms demonstrate high accuracy in offline experiments, their deployment in real-time metro environments remains limited. The computational complexity of deep learning models and the need for low-latency decision-making pose significant challenges for practical implementation.

4. Lack of Explainability

Advanced deep learning architectures, including CNNs, LSTMs, Autoencoders, and Transformers, often function as "black boxes." Maintenance engineers require interpretable diagnostic information to understand the root causes of failures. Existing studies provide limited attention to explainable artificial intelligence (XAI) techniques for turnout fault diagnosis.

5. Inadequate Prediction of Fault Progression

Most research focuses on fault detection and classification after abnormalities occur. Comparatively fewer studies address degradation modeling, remaining useful life estimation, and early prediction of fault progression. Predictive maintenance frameworks for metro turnouts remain underdeveloped.

6. Class Imbalance Problem

Fault events constitute only a small proportion of turnout operational data. Conventional machine learning models tend to be biased toward the majority class, resulting in poor sensitivity to rare fault conditions. More effective imbalance handling techniques and anomaly detection frameworks are required.

7. Limited Use of Unsupervised and Semi-Supervised Learning

Most existing methods rely on supervised learning, which requires extensive labeled datasets. Since obtaining fault labels is expensive and time-consuming, there is a need for unsupervised, semi-supervised, and self-supervised learning techniques capable of learning from large amounts of unlabeled turnout data.

8. Lack of Robustness Under Changing Operating Conditions

Metro turnouts operate under varying environmental conditions, train loads, speeds, and weather conditions. Many existing models are developed for specific datasets and may experience performance degradation when applied to different operational scenarios. More adaptive and transferable models are required.

9. Limited Integration with IoT and Digital Twins

Although IoT-based monitoring and digital twin technologies are gaining attention, their application in metro turnout fault detection is still in its early stages. Research on combining real-time sensor networks, cloud computing, and digital twins for intelligent maintenance remains insufficient.

10. Scarcity of Hybrid Intelligent Models

Most studies employ individual machine learning or deep learning techniques. The potential benefits of hybrid approaches combining signal processing, Bayesian methods, optimization algorithms, Transformers, and deep neural networks have not been fully exploited. Such hybrid frameworks could improve both accuracy and robustness.

Performance Metrics:

The training is stopped based on the mean square error or mse given by [22]:

$$mse = \frac{\sum_{i=1}^{n} e_i^2}{n} \quad \textbf{(7)}$$

The final computation of the performance metric is the mean absolute percentage error given by:

$$MAPE = \frac{100}{M}\sum_{i=1}^{N}\frac{E-E_i|}{i} \quad \textbf{(8)}$$

Here,

n is the number of errors
i is the iteration number
E is the actual value
E_i is the predicted value

V. CONCLUSION

It can be concluded that fault detection in metro turnout data is a crucial aspect of ensuring the safety and efficiency of urban railway systems. The transition from traditional maintenance practices to intelligent data-driven approaches has significantly improved the ability to identify faults at an early stage. Machine learning, deep learning, signal processing, and IoT-based technologies have emerged as promising solutions for developing reliable and automated fault diagnosis systems. As metro networks continue to expand worldwide, advanced fault detection methods will play an increasingly important role in achieving safe, sustainable, and cost-effective railway operations. The review of the state of the art methods in the domain of research pave the way for further research in the domain.

REFERENCES

[1] Bian, C., Yang, S., Huang, S., Xu, Q., Liu, J., & Zio, E. (2018). Degradation detection method for railway point machines. Engineering Failure Analysis. (Conference/Preprint).

[2] Ou, D., Xue, R., & Cui, K. (2019). A data-driven fault diagnosis method for railway turnouts. Transportation Research Record: Journal of the Transportation Research Board, 2673(4), 448–457.

[3] Dong, O., et al. (2018). Hybrid fault diagnosis of railway switches based on current curve analysis and DTW. (Conference paper).

[4] Guo, Z., Wan, Y., & Ye, H. (2020). An unsupervised fault-detection method for railway turnouts. IEEE Transactions on Instrumentation and Measurement, 69(11), 8881–8901.

[5] Li, M., Hei, X., Ji, W., Zhu, L., Wang, Y., & Qiu, Y. (2022). A fault-diagnosis method for railway turnout systems based on improved autoencoder and data augmentation. Sensors, 22(23), 9438.

[6] Hu, X., et al. (2023). Fault detection for point machines: A review, challenges, and perspectives. Machines, 12(10), 391.

[7] Barkhordari, P., Galeazzi, R., & Blanke, M. (2021). Monitoring of railpad long-term condition in turnouts using extreme value distributions. arXiv preprint.

[8] Zhou, L., et al. (2024). A research paradigm on AE-based switch rail condition monitoring at turnout areas. Journal/Proceedings (2024).

[9] Chen, C., Xu, Z., Mei, M., Huang, K., & Lo, S. M. (2024). Fault diagnosis scheme for railway switch machine using multi-sensor fusion tensor machine. Computational Materials Continua, 79(3), 4533–4549.

[10] Li, Y., et al. (2024). A fault diagnosis method for turnout switch machines based on sound signals: ERS (EMD+ReliefF+SVM). Electronics, 13(23), 4839.

[11] Reetz, S., et al. (2024). Expert system based fault diagnosis for railway point machines: A Bayesian network approach. Proceedings/Journal (2024).

[12] Huang, S., et al. (2019). Turnout fault diagnosis using DT7 and template matching, dvances in Traffic Safety Methodologies and Technologies, Wiley.

[13] Bian, C., Yang, S., & Zio, E. (2019). Degradation state mining and identification for railway point machines. Reliability Engineering & System Safety, 188, 432–443.

[14] Moreno Haro, L. M., et al. (2025). Failure detection in sensors via variational autoencoders (related technique). Sensors (2025).

[15] Tang, R., et al. (2022). A literature review of AI applications in railway transport (includes turnout monitoring). Transportation Research Part C (2022).

[16] Fu, Y., Yang, Z., Li, L., & Liu, J. (2024). Convolution-Transformer based fault diagnosis for railway turnouts. Engineering Research Express, 2025 – IOP Science

[17] Wang, H., & Zhang, Y. (2019). Image-based inspection methods for turnout geometry and defects. Measurement, 2025 – Elsevier, 247 (116737).

[18] Akula, P., et al. (2024). Acoustic signal-based detection of damage in railway tracks, 2024 IEEE Conference on Engineering Informatics (ICEI), Melbourne, Australia, 2024, 1-13.

[19] Ou, D., et al. (2018). Early hybrid detection methods for turnout current curves and waveform analysis, Applied Mechanics and Materials, 2013 - Trans Tech, 427–429, 1022–1027.

[20] 28. Yang, X., et al. (2022). Deep learning approaches for anomaly detection in turnout monitoring, Neural Computing Springer, 35, 4073–4097

[21] D Ou, Y Ji, L Zhang, H Liu, (2020), An online classification method for fault diagnosis of railway turnouts, Sensors 2020, MDPI, 20(16), 4627.

[22] Ding, X., He, Q., Shao, Y., & Huang, W. (2019). Transient feature extraction based on time–frequency manifold image synthesis for machinery fault diagnosis. IEEE Transactions on Instrumentation and Measurement, 68(11), 4242–4252.