A Principal Component Analysis (PCA)-Deep Neural Network Model for Automated Anomaly Detection Using Metro Turnout Data

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

***Abstract*— Transportation engineering is being transformed with the advent of data driven statistical models. Additionally, industry 4.0 has changed anomaly detection and predictive maintenance using data from sensors and monitoring systems. Urban metro systems form the backbone of public transportation in many cities, ensuring timely and safe movement of millions of passengers. Among the key infrastructural components of a metro network are turnouts, also known as railway switches, which allow trains to move from one track to another. The performance and reliability of turnouts directly affect the efficiency and safety of the rail system. However, due to their mechanical and electrical complexity, turnouts are prone to failures and malfunctions. Automated anomaly detection using metro turnout data has emerged as a critical approach to address these challenges, enabling proactive monitoring and predictive maintenance. This paper presents a combination of Principal Component Analysis (PCA) and Deep Neural Network based statistical model for automated anomaly detection using metro turnout data. It has been shown that the proposed work attains lower error percentage compared to existing research frameworks in the domain.**

Keywords*— Transportation Engineering, Statistical Modelling, Metro Turnout Imbalanced Datasets, Principal Component Analysis (PCA), Deep Neural Network, Error Percentage.

I. INTRODUCTION

One of the most profound impacts of Industry 4.0 on transportation engineering is the digitalization of infrastructure and operations. Smart sensors, communication networks, and IoT devices allow for real-time monitoring of roads, railways, airports, and shipping ports [1]. These connected systems can detect congestion, monitor wear and tear, and provide continuous feedback to operators [2]. Digital twins—virtual replicas of transportation infrastructure—enable engineers to simulate traffic scenarios, predict maintenance needs, and optimize resource allocation, leading to more efficient and reliable transport systems [3].

Fig.1 Concept of Industry 4.0

Industry 4.0, often referred to as the fourth industrial revolution, represents the convergence of digital technologies with traditional industries through automation, connectivity, and intelligent systems [5]. While it was initially associated with manufacturing, its principles are increasingly being applied to transportation engineering. Transportation systems worldwide are under pressure to become more efficient, sustainable, and resilient [6] By leveraging technologies such as the Internet of Things (IoT), artificial intelligence (AI), big data analytics, and cyber-physical systems, Industry 4.0 is transforming transportation engineering into a smarter, safer, and more adaptive domain [7].

II. APPLI CATIONS OF INDUSTRY 4.0 IN METRO TRANSPORT.

Transportation infrastructure requires continuous upkeep, and Industry 4.0 offers powerful tools for predictive maintenance [8]. By analyzing sensor data from bridges, tracks, runways, and vehicles, engineers can detect early signs of wear, cracks, or mechanical failures [9]. Predictive maintenance reduces unexpected breakdowns, minimizes downtime, and lowers repair costs by addressing problems before they escalate. In addition, asset management systems powered by big data provide transportation agencies with comprehensive insights into the lifecycle of infrastructure components, enabling smarter investment and resource allocation. Despite its potential, implementing Industry 4.0 in transportation engineering comes with challenges [10]. High costs of infrastructure modernization, cybersecurity risks, and interoperability issues between legacy systems and new technologies can hinder large-scale deployment. Moreover, workforce adaptation is essential—transportation engineers must acquire new digital and data-driven skills. Ethical and legal concerns regarding autonomous vehicles and data privacy also pose significant barriers. Addressing these challenges requires collaborative efforts from governments, industries, and academia to create robust frameworks for safe and equitable adoption [11].

Metro railway systems have become an integral part of urban transportation networks, offering fast and efficient mobility for millions of passengers daily [12]. At the heart of these systems lie critical infrastructure components such as turnouts or switches, which enable trains to change tracks and navigate complex routes. Ensuring the reliability and safety of these turnouts is paramount . Traditionally, inspection and maintenance of turnouts have relied on manual methods or rule-based systems, which are not only labor-intensive but also prone to human error. With the increasing complexity and usage frequency of metro systems, there is a growing need for intelligent, automated, and scalable monitoring solutions [13]

Metro Railway Turnouts refer to the critical components in urban rail transit (URT) systems that enable trains to switch tracks safely and efficiently. These are part of the Switch & Crossing (S&C) systems, which include rails, actuators, and switch machines [14].

III. STATISTI CAL MODELS FOR METRO TURNOUT MONITORING

Metro turnouts, also known as railway switches, are critical components that guide trains from one track to another. Their reliable functioning directly impacts operational safety and service continuity in metro networks. Due to mechanical stresses, electrical wear, and environmental influences, turnouts are prone to anomalies that can lead to costly disruptions. To detect early signs of deterioration, engineers often rely on statistical models, which analyze operational data such as switching current, actuation time, vibration levels, and temperature. These models provide systematic methods to identify abnormal patterns, support decision-making, and reduce the risk of failures [15].

Descriptive Statistical Models: The simplest statistical models for turnout monitoring involve descriptive measures such as mean, variance, standard deviation, and range of key parameters. For instance, tracking the average switching current over time helps establish a baseline for normal operation. Though basic, descriptive models offer intuitive insights and are useful for establishing initial benchmarks in turnout monitoring [16].

Control Chart Models: Control charts, widely used in quality control, are a common tool in turnout monitoring. Shewhart control charts, for example, monitor whether parameters like switching time or vibration remain within statistically acceptable limits. Cumulative Sum (CUSUM) and Exponentially Weighted Moving Average (EWMA) charts are particularly effective for detecting small or gradual shifts in turnout performance. These methods help engineers distinguish between natural variability and genuine anomalies, thus improving reliability in maintenance decisions [17].

Regression Models: Regression analysis provides another layer of insight into turnout behavior. Linear and multiple regression models are used to establish relationships between turnout performance variables and influencing factors such as temperature, train load, or usage frequency [18]. For example, regression can predict how actuation current changes with seasonal temperature variations. Logistic regression is also applied to classify turnout states into healthy, degraded, or faulty conditions. Such models are valuable for predictive maintenance strategies, as they can quantify the impact of external conditions on turnout reliability [19]

Time-Series Models: Since turnout data is collected sequentially over time, time-series models like Autoregressive Integrated Moving Average (ARIMA) and its variants are particularly relevant. These models capture trends, seasonality, and autocorrelations in turnout parameters, allowing for accurate forecasting of future conditions. For example, ARIMA can predict when switching current is likely to exceed safe thresholds, enabling preemptive interventions. Time-series models are well suited for continuous monitoring systems that require near-real-time decision support [20]

Probabilistic and Survival Models: Probabilistic models, including Bayesian approaches and survival analysis, are also used in turnout monitoring. Bayesian inference allows for incorporating prior knowledge about turnout performance and updating failure probabilities as new data becomes available. Survival models, such as the Weibull distribution, estimate the expected lifetime of turnout components under varying operating conditions. These models support long-term asset management and help optimize maintenance schedules based on component reliability [21].

Benefits of Statistical Models: Statistical models provide a systematic, data-driven approach to turnout monitoring, offering several advantages. They are relatively easy to implement, interpretable for engineers, and capable of detecting both sudden anomalies and gradual wear. By quantifying variability and forecasting potential issues, these models reduce unplanned failures, improve safety, and lower maintenance costs. Moreover, statistical methods form the foundation for more advanced machine learning and AI techniques, acting as a bridge between traditional monitoring and intelligent predictive systems [22].

IV. EXISTING CHALLENGS AND CLASS IMBALANCE

Despite their usefulness, statistical models face limitations in complex turnout environments. They often assume stationarity or linearity, which may not hold true for real-world turnout data influenced by nonlinear interactions and noise. Control charts and regression models may generate false alarms under fluctuating environmental conditions. Time-series models, while powerful, require large amounts of historical data and careful parameter tuning. Therefore, statistical models are most effective when combined with engineering expertise or integrated with modern data-driven approaches such as machine learning [23].

Automated anomaly detection and fault classification methods are increasingly applied to turnout monitoring. In most real-world datasets, healthy turnout conditions dominate, while faulty or anomalous conditions are relatively rare. This imbalance affects the performance of statistical and machine learning models, often leading to biased outcomes [24].

Class imbalance arises because turnouts operate normally for the majority of their lifecycle, and only a small fraction of data corresponds to anomalies such as mechanical wear, obstruction, or actuator failures. For example, 95–98% of collected records may indicate normal operation, while less than 2–5% may reflect fault states [25].

This unequal distribution of classes creates challenges in training predictive models, as they become biased toward the majority (healthy) class, often ignoring the minority (faulty) class that is actually of greater importance for safety and maintenance. In statistical models, imbalance reduces the effectiveness of threshold-based methods, as thresholds may be overly influenced by dominant healthy data. Similarly, in machine learning models, classifiers like logistic regression, decision trees, or support vector machines tend to become biased toward the majority class. Deep learning models face overfitting risks, as they may memorize majority class features while underrepresenting the minority class. This imbalance also complicates evaluation, since metrics like accuracy become misleading for minority class detection [26].

IV. PROPOSED ALGORITHM

This work proposes the amalgamation of two statistical models:

1. Principal Component Analysis (PCA)
2. Deep Neural Networks

Each of them is explained next:

Principal Component Analysis (PCA): It is a widely used dimensionality reduction technique that plays a crucial role in handling large-scale and complex datasets. In metro turnout systems, which are critical components of railway infrastructure, monitoring and analyzing data is essential for ensuring safety, reliability, and efficient operation. Turnouts generate high-dimensional data from various sensors measuring parameters such as vibration, temperature, current, and displacement. Analyzing this high-dimensional data can be challenging, and PCA offers a powerful tool to simplify the process while preserving important information. Thus applying the PCA would yield in a reduced data vector for training given by [27]:

$$[X]_n \xrightarrow[PCA]{} [X]_{n-k} \quad (1)$$

Here,

X is the original data vector

N is the dimension of the original data vector

K is the dimensional reduction factor

n-k is the reduced dimensions of the data vector after the application of PCA.

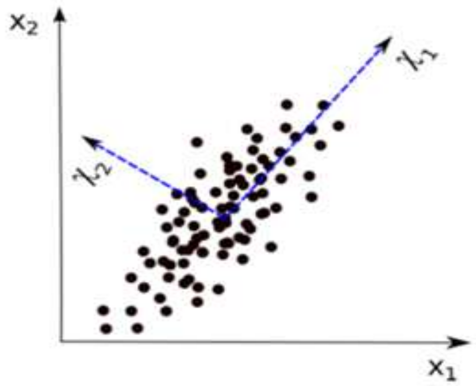

Fig.2 Concept of PCA Vectors

It can be observed that the PCA vectors map the data points onto the orthogonal plane to minimize the correlation and maximize the variance. As metro turnouts are subjected to continuous mechanical stress and environmental conditions, leading to diverse patterns in sensor data. PCA helps in identifying the most significant features by transforming correlated variables into a smaller set of uncorrelated principal components.

These components capture the maximum variance in the dataset, making it easier to detect underlying patterns. By reducing redundancy, PCA allows engineers to focus on the most informative variables, improving the efficiency of monitoring and fault detection in metro turnouts.

Deep Neural Networks: The deep neural network model in this case is the BayesNet with penalty based regularization. It is an improved version of the conventional Naïve Bayes. The gradient is considered as the objective function to be reduced in each iteration. A probabilistic classification using the Bayes theorem of conditional probability is given by:

$$P\left(\frac{H}{X}\right) = \frac{P\left(\frac{X}{H}\right)P(H)}{P(X)} \quad (2)$$

Here,

Posterior Probability [P (H/X)] is the probability of occurrence of event H when X has already occurred
Prior Probability [P (H)] is the individual probability of event H
X is termed as the tuple and H is is termed as the hypothesis.
Here, [P (H/X)] denotes the probability of occurrence of event X when H has already occurred.

Each node is associated with a conditional probability distribution that quantifies the effect of its parents in the graph. Bayes Nets provide a structured way to model joint probability distributions, allowing for efficient inference and learning. They are particularly useful in domains where relationships among variables are complex and uncertain, such as metro turnout. The probability function can be computed using [28]:

$$P\left(\frac{X}{X_{i,}k_1,k_2,M}\right) = \frac{P\left(\frac{X_i}{X,k_2,M}\right)P\left(\frac{X_i}{k_1,M}\right)}{P\left(\frac{X}{k_1,k_2,M}\right)} \quad (3)$$

Here,
P denotes probability
X_i denotes the set of weight and bias
X denotes the training data set
M denotes the network architecture in terms of the hidden layers and neurons
k_1 and k_2 are the regularization parameters for the network

Incorporating prior distributions over the parameters or network structures, guiding the learning process towards more plausible models. Priors can reflect domain knowledge or be designed to favor simpler models, thereby enhancing generalization.

Generally, the term $\rho = \frac{k_1}{k_2}$ is called the regularization ratio. The regularization parameter is adopted in this case to limit the variations in the weights by introducing a penalty factor to the learning algorithm's cost function or objective function $\boldsymbol{J}$. The regularization is different from early stopping or convergence in the sense that the earlier truncates the iterations prior to convergence to a minimum value of $\boldsymbol{J}$ whereas the latter tries to restrict the values of weights and number of parameters by modifying the cost function. Thus, regularization allows a much steeper decrease in the cost function and eventually lesser values as compared to early stopping. This significantly helps to reduce the time complexity of the algorithm.

Algorithm:

The training algorithm adopted in this work is given by:

Step.1: Initialize weights (w) randomly.

Step.2: Fix the maximum number of iterations (n) and compute $\rho = \frac{k_1}{k_2}$

Step.3: Update weights using gradient descent with an aim to minimize the objective function J given by:

$$J = \frac{1}{m}\sum_{i=1}^{m}(v_i - v'_i)^2 \quad (4)$$

Step.4: Compute the Jacobian Matrix $\boldsymbol{J}$ given by:

$$J = \begin{bmatrix} \frac{\partial^2 e_1}{\partial w_1^2} & \cdots & \frac{\partial^2 e_1}{\partial w_m^2} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 e_n}{\partial w_1^2} & \cdots & \frac{\partial^2 e_n}{\partial w_m^2} \end{bmatrix} \quad (5)$$

Here,
The error for iteration 'i' designated by e_i is computed as:

$$e_i = (y_i - y'_i) \quad (6)$$

Here
y_i is the actual value
y'_i is the predicted value

Step.5: Iterate steps (1-4) till the cost function J stabilizes or the maximum number of iterations set in step 2 are reached, whichever occurs earlier.

Regularization enhances the robustness and generalizability of Bayesian Networks by preventing overfitting. By constraining the model complexity, regularization techniques ensure that the learned network captures the essential dependencies among variables without being influenced by noise. This leads to improved predictive performance on new data and more reliable inferences. Additionally, regularization facilitates the interpretation of the network by avoiding unnecessarily complex structures, making it easier to understand and communicate the relationships among variables.

Performance Metrics:

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

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

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

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

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

V. EXPERIMENTAL RESULTS

This section presents the experimental results. The dataset consists of 15169480 data points collected at 1Hz and is described by 15 features from 7 analogue (1-7) and 8 digital (8-15) sensors:
TP2 (bar) – the measure of the pressure on the compressor.
TP3 (bar) – the measure of the pressure generated at the pneumatic panel.
H1 (bar) – the measure of the pressure generated due to pressure drop when the discharge of the cyclonic separator filter occurs.
DV pressure (bar) – the measure of the pressure drop generated when the towers discharge air dryers; a zero reading indicates that the compressor is operating under load.
Reservoirs (bar) – the measure of the downstream pressure of the reservoirs, which should be close to the pneumatic panel pressure (TP3).
Motor Current (A) – the measure of the current of one phase of the three-phase motor; it presents values close to 0A - when

it turns off, 4A - when working offloaded, 7A - when working under load, and 9A - when it starts working.

Oil Temperature (ºC) – the measure of the oil temperature on the compressor.

COMP - the electrical signal of the air intake valve on the compressor; it is active when there is no air intake, indicating that the compressor is either turned off or operating in an offloaded state.

DV electric – the electrical signal that controls the compressor outlet valve; it is active when the compressor is functioning under load and inactive when the compressor is either off or operating in an offloaded state.

TOWERS – the electrical signal that defines the tower responsible for drying the air and the tower responsible for draining the humidity removed from the air; when not active, it indicates that tower one is functioning; when active, it indicates that tower two is in operation.

MPG – the electrical signal responsible for starting the compressor under load by activating the intake valve when the pressure in the air production unit (APU) falls below 8.2 bar; it activates the COMP sensor, which assumes the same behaviour as the MPG sensor.

LPS – the electrical signal that detects and activates when the pressure drops below 7 bars.

Pressure Switch - the electrical signal that detects the discharge in the air-drying towers.

Oil Level – the electrical signal that detects the oil level on the compressor; it is active when the oil is below the expected values.

Caudal Impulse – the electrical signal that counts the pulse outputs generated by the absolute amount of air flowing from the APU to the reservoirs.

The target variable in this case is the motor current. The proposed model tries to map the relation among the input variables **(X)** and the dependent or target variable **(Y).** Accurate mapping of the variables X and Y would results in lower error rates and high resultant accuracy.

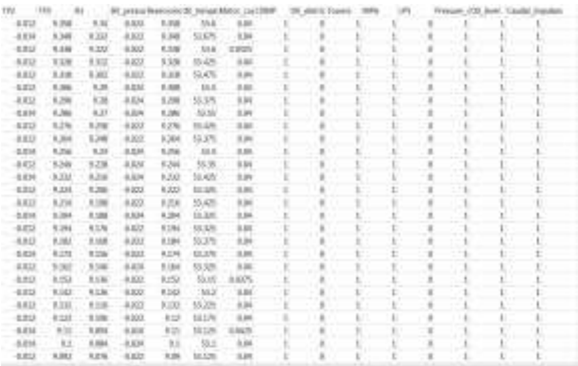

Fig.3 Raw Data

The figure above shows the raw data used in the study.

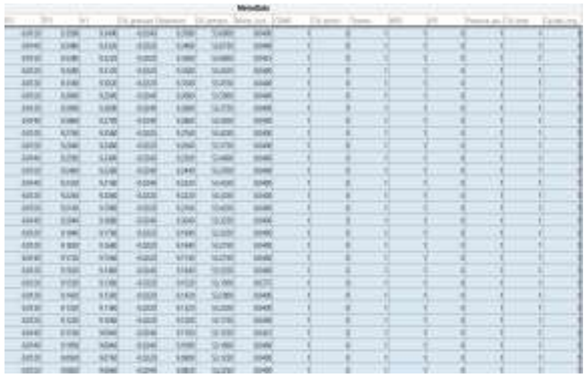

Fig.4 Importing raw data to MATLAB workspace

Figure above shows importing of the raw data to MATLAB workspace.

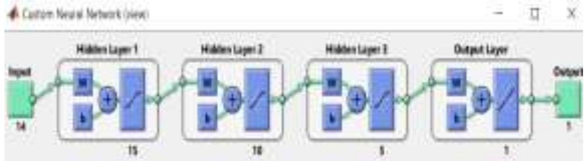

Fig.5 Network Visualization

Figure above shows the designed deep neural network with total of 5 layers, which is a shallow deep net. The hidden layer configuration has been taken as 15-10-5.

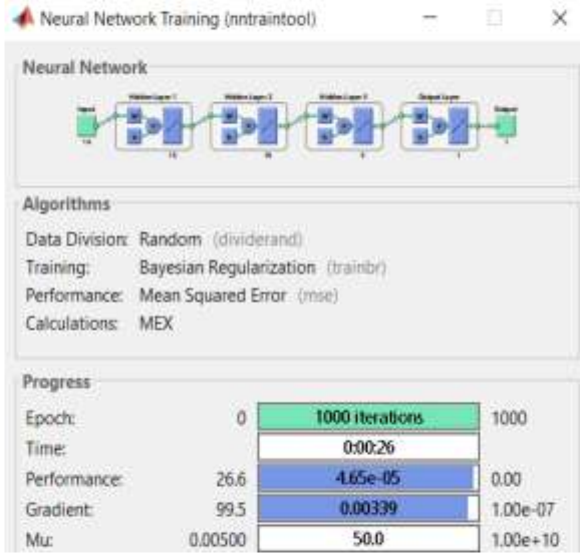

Fig.6 Network Visualization

Figure above shows the training of the neural network which trains in 1000 iterations and 26s. The values of the gradient and learning rate of the model can also be observed.

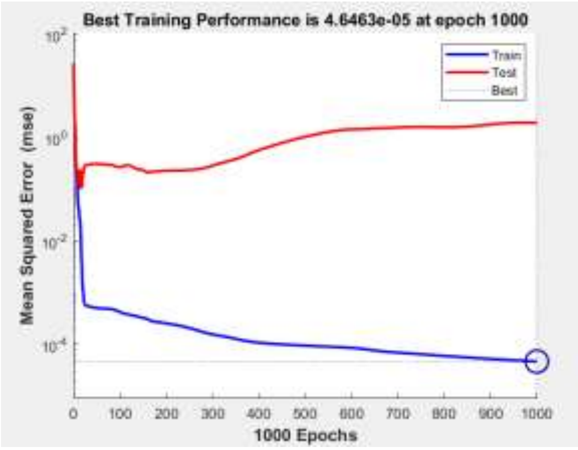

Fig.7 MSE to Convergence

It can be observed that the MSE is $4.5\ x\ 10^{-5}$ at convergence.

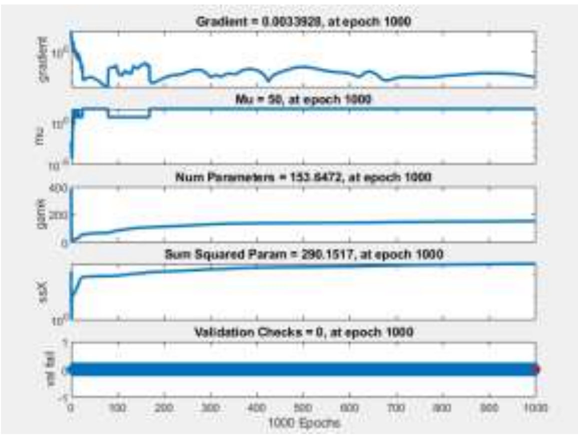

Fig.8 Training States

The figure above shows the training states or training parameters of the model.

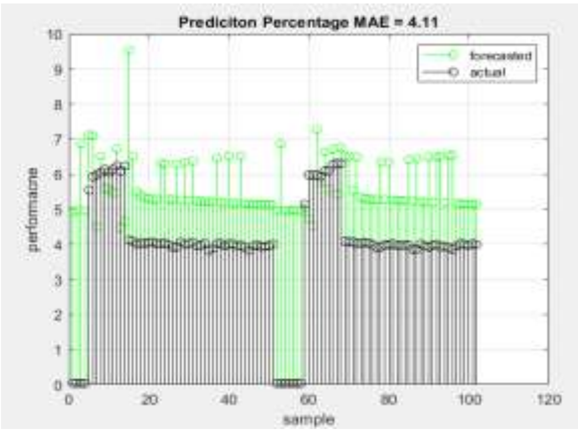

Fig.9 Percentage MAE obtained.

It can be observed that the proposed work attains an Percentage MAE or MAPE of 4.11 at convergence which depicts the accurate prediction capability of the proposed work.

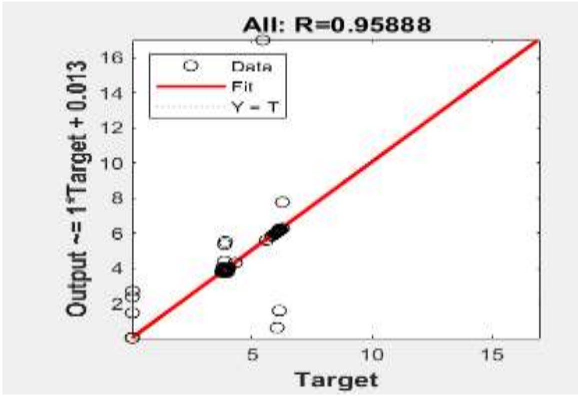

Fig.10 Overall Regression

The figure shows the overall regression (R^2) value of the model which is 0.95888.

Table 1 Summary of Results

S.No	Parameter	Value
1.	Dataset Parameters	15
2.	Model	PCA- Deep Neural Network Hybrid
3.	Hidden Layer Configuration	15-10-5
4.	Algorithm	Bayesian Regularization
5.	Iterations	1000
6.	MSE at convergence	$4.5\ x\ 10^{-5}$
7.	Gradient at convergence	0.00339
8.	**Percentage MAE (Proposed Work)**	**4.11 (PCA + Bayesian Deep Neural Network)**
9.	**Percentage MAE (Previous Work, Chen et al., [29])**	**8% (Convolutional Auto-encoder based Neural Network)**

The approach attains higher classification accuracy compared to baseline approaches [29].

CONCLUSION: Metro railway systems are critical infrastructures that demand high levels of safety, reliability, and efficiency. One of the most vulnerable components of this system is the turnout, which enables trains to switch tracks. Faults or anomalies in turnout systems can lead to severe disruptions or even accidents. To enhance operational safety, the adoption of intelligent data-driven methods has become increasingly important. The proposed approach integrates Principal Component Analysis (PCA) with Deep Neural Networks (DNNs), which allows both

dimensionality reduction and robust anomaly detection in complex turnout datasets. The PCA–DNN model represents a powerful approach for automated anomaly detection in metro turnout systems. By combining dimensionality reduction and deep learning, it provides an effective solution to handle complex, high-dimensional data and accurately detect anomalies. This would allows safer and swifter operation of the metro systems. The proposed work attains lower error percentage of 4.11 compared to existing work 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. Su, X., et al. (2024). Research on ultrasonic guided wave-based high-speed detection for turnout switch rail fractures. Mechanical Systems and Signal Processing (2024).
12. Sun, Y., et al. (2023). Condition monitoring and fault diagnosis strategy of railway switch machines. Transportation Science and Engineering (2023).
13. Del Río, A. M., et al. (2024). Pattern recognition based on statistical methods combined with DTW for turnout current curve analysis. Expert Systems with Applications (2024).
14. Li, W., et al. (2024). A semi-supervised adaptive matrix machine approach for switch machine fault diagnosis. Sensors/MDPI (2024).
15. Reetz, S., et al. (2024). Expert system based fault diagnosis for railway point machines: A Bayesian network approach. Proceedings/Journal (2024).
16. Huang, S., et al. (2019). Turnout fault diagnosis using DT7 and template matching, dvances in Traffic Safety Methodologies and Technologies, Wiley.
17. Bian, C., Yang, S., & Zio, E. (2019). Degradation state mining and identification for railway point machines. Reliability Engineering & System Safety, 188, 432–443.
18. Moreno Haro, L. M., et al. (2025). Failure detection in sensors via variational autoencoders (related technique). Sensors (2025).
19. Tang, R., et al. (2022). A literature review of AI applications in railway transport (includes turnout monitoring). Transportation Research Part C (2022).
20. Fu, Y., Yang, Z., Li, L., & Liu, J. (2024). Convolution-Transformer based fault diagnosis for railway turnouts. Engineering Research Express, 2025 – IOP Science
21. Wang, H., & Zhang, Y. (2019). Image-based inspection methods for turnout geometry and defects. Measurement, 2025 – Elsevier, 247 (116737).
22. 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.
23. 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.
24. 28. Yang, X., et al. (2022). Deep learning approaches for anomaly detection in turnout monitoring, Neural Computing Springer, 35, 4073–4097
25. 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.
26. 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.
27. Liu J and Rao B (2019), "Sparse Bayesian Learning for Robust PCA: Algorithms and Analyses," in IEEE Transactions on Signal Processing, 67(22),, 5837-5849,
28. Pereira A, Antoni J, Leclere Q (2019), Empirical Bayesian regularization of the inverse acoustic problem, Applied Acoustics, Elsevier, 97, 11-19.
29. Chen, C., Li, X., Huang, K., Xu, Z., & Mei, M. (2023). A convolutional autoencoder based fault detection method for metro railway turnout. Computer Modeling in Engineering & Sciences, 136(1), 471–485.