Detecting Railway Track Irregularities Using Conformal Prediction

Andreas Plesner[1] and Allan P. Engsig-Karup[2] and Hans True[2]

[1] ETH Zurich, Switzerland,
aplesner@ethz.ch,
[2] Technical University of Denmark (DTU), Denmark,
{apek,htru}@dtu.dk

Abstract. This study addresses the challenge of assessing railway track irregularities using convolutional neural networks (CNNs) and conformal prediction techniques. Using high-fidelity sensor data from high-speed trains, the study proposes a CNN model that outperforms state-of-the-art results, achieving a mean unsigned error of 0.31 mm on the test set. Incorporating conformal prediction with the CV-minmax method, the model delivers prediction intervals with 97.18% coverage, averaging 2.33 mm in width, ensuring reliable uncertainty estimation. The model also exhibits impressive computational efficiency, processing data at a rate suitable for real-time applications, with the capacity to evaluate over 2,000 kilometers of track data per hour. These advances demonstrate the potential of the model for practical implementation in continuous monitoring systems, providing a contribution to the field of predictive maintenance within the railway industry.

Keywords: Railway track integrity, convolutional neural networks, conformal prediction, predictive maintenance, sensor data analysis, machine learning

1 Introduction

In the evolving landscape of transportation, railways play a pivotal role, offering a blend of efficiency, reliability, and environmental sustainability. As rail networks burgeon, paralleled by an upsurge in speed and passenger expectations, the imperatives of track safety and maintenance have ascended to the forefront of railway operations. The integrity of railway tracks, susceptible to irregularities due to wear and external forces, directly influences the safety, comfort, and operational efficiency of rail services. Traditional track inspection methodologies, although precise, grapple with limitations such as high operational costs, limited coverage, and high latency between inspections. The advent of machine learning and sensor technology indicates a transformative approach that allows continuous, real-time monitoring of track conditions by collecting data from in-service railway vehicles to generate predictive machine learning models.

This study delves into applying convolutional neural networks (CNNs) and conformal prediction methods to preemptively identify track irregularities from dynamic responses of in-service railway vehicles. The research is based on the use of high-fidelity sensor data from a high-speed train, embodying a shift from conventional reactive maintenance strategies to a predictive maintenance paradigm. The fusion of CNNs with conformal prediction offers a robust way to quantify the uncertainty of predictions, improving the reliability of the predictive framework. This integration not only showcases the potential of deep learning to decipher complex patterns from high-dimensional data but also underscores the importance of conformal prediction in providing robust, uncertainty-aware inferences.

Focusing on the prowess of the convolutional neural network, the study highlights the architecture, training, and optimization decisions that underpin the successful application of CNNs to the task at hand. CNNs are used because of their ability to handle spatial hierarchies in data, making them especially suitable for analyzing the nuanced dynamics captured by onboard sensors. The investigation extends to the realm of conformal prediction, highlighting its utility in giving prediction intervals that encapsulate the expected deviations with a quantifiable confidence level. The results derived from this application of CNNs and conformal prediction not only demonstrate a marked advancement in the accuracy and reliability of track irregularity prediction but also show the way for operationalizing these insights in real-world railway maintenance operations.

The prediction errors of the models, i.e., the difference between the predictions and the actual values, will have to be low enough so that it can be determined if the operating limits are exceeded. The EN:13848-5 standard [8] is used to establish a benchmark. This document contains operating limits for the track measurements at various speeds. The strictest limits are deviations of 1 mm for 100-meter running means and standard deviations. Based on this, a 0.1 mm benchmark will be chosen for the mean unsigned error, ME. Furthermore, to ensure that the predictions can be safely used to assess operating limits, further benchmarks are established to say that the maximum of the unsigned errors is below 0.5 mm. Since the limit values in [8] are only specified to the nearest millimeter, a maximum unsigned error of less than 0.5 mm would imply that all running means are within 0.5 mm of the actual value.

Kawasaki and Youcef-Toumi [11] give error ranges that would accept maximum unsigned errors of less than 4 mm and an ME of around 1 mm, while Hao et al [9] set a benchmark of 0.25 mm and 0.45 mm of the mean unsigned error in the wavebands [3 m, 42 m] and [42 m, 120 m], respectively. These are less strict than our benchmarks mentioned above. We therefore set additional benchmarks of an ME of 0.35 mm (the mean of 0.25 mm and 0.45 mm) and a maximum unsigned error of 4 mm and call these the “satisfying” levels.

2 Related work

Vehicle dynamics Traditionally, the integrity assessment of railway infrastructure has relied heavily on periodic inspections using specialized measurement vehicles, a process that, while accurate, suffers from limitations such as high costs, limited coverage, and the potential for subjective error. Ravitharan [18] highlight the operational benefits of proactive maintenance strategies, advocating for continuously monitoring track conditions using in-service railway vehicles, a concept explored over the past two decades [11, 26]. Lee et al [12] underscores the direct correlation between vehicle dynamics and track conditions, laying the foundation for the use of vehicle dynamics as a means of assessing track quality.

Classic Methods Prior works have largely focused on classical mathematical analysis tools, such as Kalman filters, system identification techniques, digital and analog processing, and other signal processing methods [3, 5, 11, 12, 15, 16, 22, 25]. These works have prioritized interpretable models over complex data-driven solutions. These traditional methods often encounter mathematical difficulties, such as the issue of double integration of accelerations to obtain positions, which complicates their application in real-world scenarios [26].

Deep Learning Recent advances in machine learning, particularly in the application of convolutional neural networks (CNNs), present promising alternatives to traditional methods. Data-driven machine learning models have begun to shift the paradigm in various domains, demonstrating superior performance in fields such as image analysis [1, 17]. In the context of monitoring the condition of the railway track, initiatives have explored the use of cameras on board and binary classification techniques to differentiate between good and bad track conditions [7, 14, 21, 27]. However, the adoption of machine learning in this domain is not without its challenges. Despite their promise, these approaches face their own set of limitations, including computational demands and the lack of severity assessment in track irregularities [14].

Research by Hao et al [9] presents a notable advancement, which showcases the potential of deep learning approaches to predict vertical track irregularities with a high degree of precision. However, this method does not address lateral irregularities and is based on simulated data, which may not fully capture the complexity of real-world track conditions. Similarly, the use of autoencoders to compress irregularity data presents innovative solutions but is again limited to simulated environments and specific types of irregularities [13].

Exploring data-driven methods for road quality monitoring has also yielded encouraging results, suggesting that similar approaches could be beneficial for the maintenance of railway tracks [23].

3 Methodology and Data

This section outlines the methodology employed to predict railway track irregularities using Convolutional Neural Networks (CNNs) complemented by conformal prediction techniques to estimate the uncertainties of these predictions. The section also includes a bit of background for these methods, how the CNN is designed specifically for the task at hand, and training the model. In addition, we will look at the data used for this project.

Label	Unit	Description	Notes
Position	km	Position along the track	
Lateral left D1	mm	Lateral irregularities of the left and right rail in the D1 wavelength domain	D1 is the first frequency band with wavelengths in [3 m, 25 m]
Lateral right D1	mm		
Vertical left D1	mm	Vertical irregularities of the left and right rails in the D1 wavelength domain	
Vertical right D1	mm		
Lateral left D2	mm	Lateral irregularities of the left and right rail in the D2 wavelength domain	D2 is the second frequency band with wavelengths in [25 m, 70 m]
Lateral right D2	mm		
Vertical left D2	mm	Vertical irregularities of the left and right rails in the D2 wavelength domain	
Vertical right D2	mm		
Lateral left D3	mm	Lateral irregularities of the left and right rail in the D3 wavelength domain	D3 is the third frequency band with wavelengths in [70 m, 200 m]
Lateral right D3	mm		
Vertical left D3	mm	Vertical irregularities of the left and right rails in the D3 wavelength domain	
Vertical right D3	mm		

Table 1. Features in a sample of geometry dataset – With labels used in this project.

3.1 Data Collection and Preprocessing

The data used in this study comprise high-fidelity sensor readings from a high-speed train, capturing various dynamic responses under operating conditions. The preprocessing steps involved the removal of outliers and normalization and segmentation to ensure compatibility with the CNN architecture. This preprocessing facilitated the transformation of raw sensor data into a structured format conducive to machine learning models.

Data are collected using multiple accelerometers located at various points on the railway vehicle; see Fig. 1 for locations. The input data consist of time series with measurements from each accelerometer. The output data consist of the irregularities of the track in the lateral and vertical directions for the left and right rails. These have been split into three frequency domains D1, D2, and D3 with wavelengths of [3 m, 25 m], [25 m, 70 m], and [70 m, 200 m], respectively, thus giving 12 output series. An outlier analysis found that five of the sensors had regions where they were faulty; in these regions, the faulty data were zeroed. Data were collected with a sampling frequency of ≈ 1000 Hz, and this was interpolated to have a constant sample spacing of 0.167 m. The train did not drive at a constant speed, so the spacing was irregular in the positional domain before interpolation. The features in the track geometry (output) data

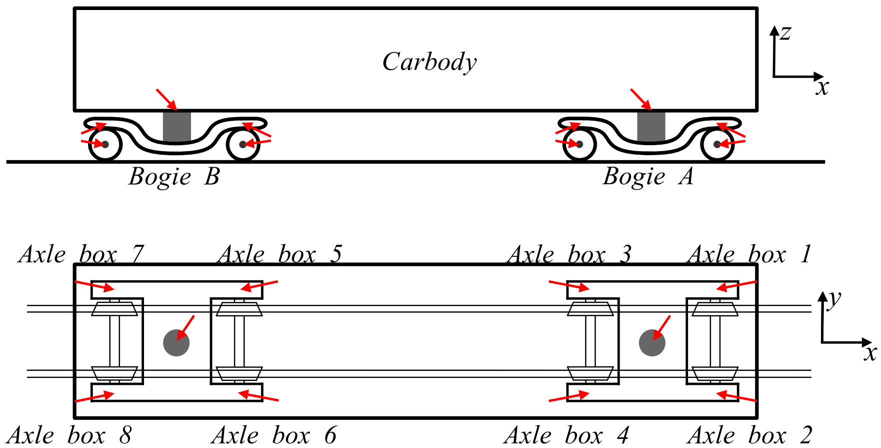

Fig. 1. Data measurement locations of the vehicle dynamics. The red arrows indicate the placement of the accelerometers on the axle boxes, bogies, and car body.

can be seen in Table 1 with a large table of all dynamics (input) features in Appendix Table 4.

We then split the data into training and testing regions. The data have 287,827 observations in total, and the 1st to 23,827th sample, the 94,001st to 117,827th sample, and the 188,001st to 211,827th sample are used as the test data. The training data consist of the 23,828th to 94,000th sample, the 117,828th to 188,000th sample, and the 211,827th to 281,827th sample. These regions have been shown in Fig. 2. The training data are further divided into training and validation segments by splitting it into 9 regions and using 1 for validation and the remaining 8 for training. Six of the nine regions are used for validation; a separate model is trained for each of the six validation regions, and the results are the mean across the six models.

3.2 Convolutional Neural Network (CNN) Architecture

The CNN architecture was designed to process time-series data, capturing spatial and temporal dependencies inherent in the train's dynamic responses. The model comprises multiple convolutional layers, each followed by pooling layers to reduce dimensionality and enhance feature extraction. Dropout layers were incorporated to mitigate overfitting, ensuring the model's generalizability across different track conditions. The model consists of batch normalization of the input and then 3 hidden CNN layers using batch normalization, the ELU activation function, and dropout of 60 %, with a final CNN layer to obtain the output [6, 10, 20]. The first convolutional layer uses very large kernels to ensure that features with 300 m wavelengths can be captured. This is relevant as the irregularities can exhibit wavelengths up to 200 m. A diagram of the final network has been shown in Fig. 4.

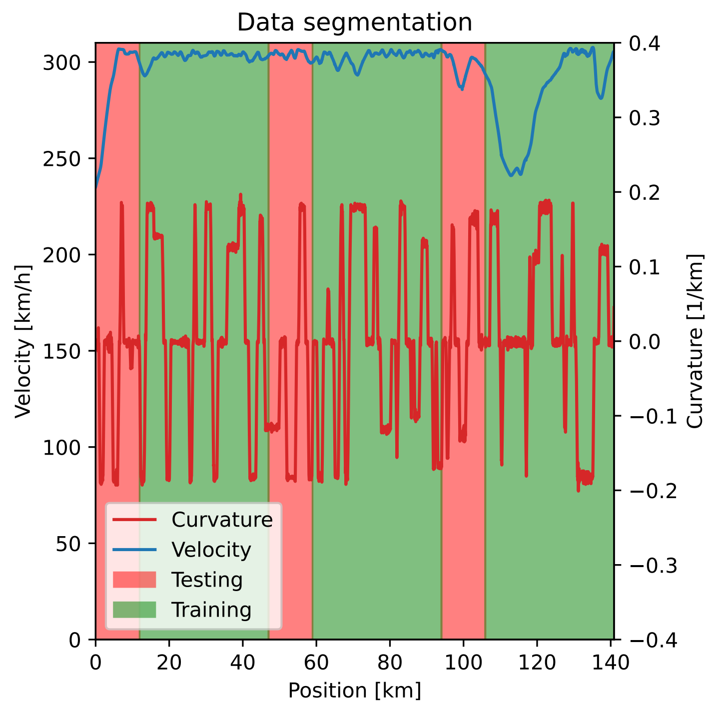

Fig. 2. Training and testing regions of the data. During the training of the models, the training data is divided into training and validation segments by splitting it into 9 regions and using 1 for validation and the remaining 8 for training.

Hyperparameters, including the learning rate, number of convolutional layers, kernel size, and dropout rate, were tuned using a combination of grid search and cross-validation to find the optimal model configuration by comparing the validation losses.

3.3 Conformal Prediction Framework

To quantify the uncertainty of CNN predictions, we applied conformal prediction methods. These methods use residuals from the training dataset to construct prediction intervals for new observations. Different variants of conformal prediction, Naïve, Holdout, and Cross-Validation, were evaluated to determine the most effective approach for this application. The Cross-Validation variant can further be split into three versions, CV, CV+, and CV-minmax [2, 4, 19, 24]. The best intervals were produced by CV+ and CV-minmax, so the results will include only these. These methods have assumption-free theoretical guarantees that the α level interval contains $> 1-\alpha$ of the samples [4]. For our results, we will use

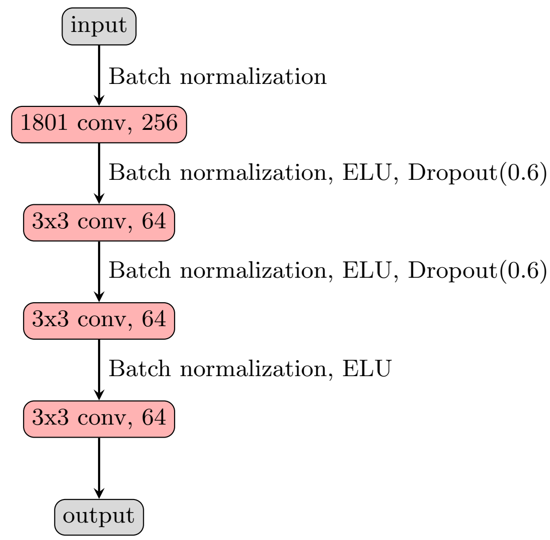

Fig. 3. Depiction of the CNN after tuning the hyperparameters.

$\alpha = 0.05$ intervals. The focus on CV+ and CV-minmax is due to the better theoretical guarantees of these methods [4].

3.4 Evaluation Metrics

The performance of the CNN model and the effectiveness of the conformal prediction intervals were evaluated using a few metrics. For CNN, the metrics were the mean and maximum of unsigned errors and the compute time. For conformal prediction, the focus was on the accuracy of the prediction intervals measured through the coverage probability (how often the true values were inside the interval) and the width of the interval assessed through the mean and maximum width.

4 Results

This section will present the results of this project for the best CNN model constructed, the use of conformal predictions, and, lastly, the compute time required to evaluate the model and produce prediction intervals.

4.1 CNN Predictions

The convolutional neural network (CNN) model showcased proficiency in predicting track irregularities from dynamic responses of in-service railway vehicles. The model, after rigorous tuning, achieved a satisfactory mean unsigned error (ME) by beating the “satisfying” benchmark for the ME. This significant achievement is depicted in Fig. 4, illustrating the training and validation mean errors across epochs, where the model's performance is notably highlighted by its capacity to maintain errors below the "satisfying" benchmark level.

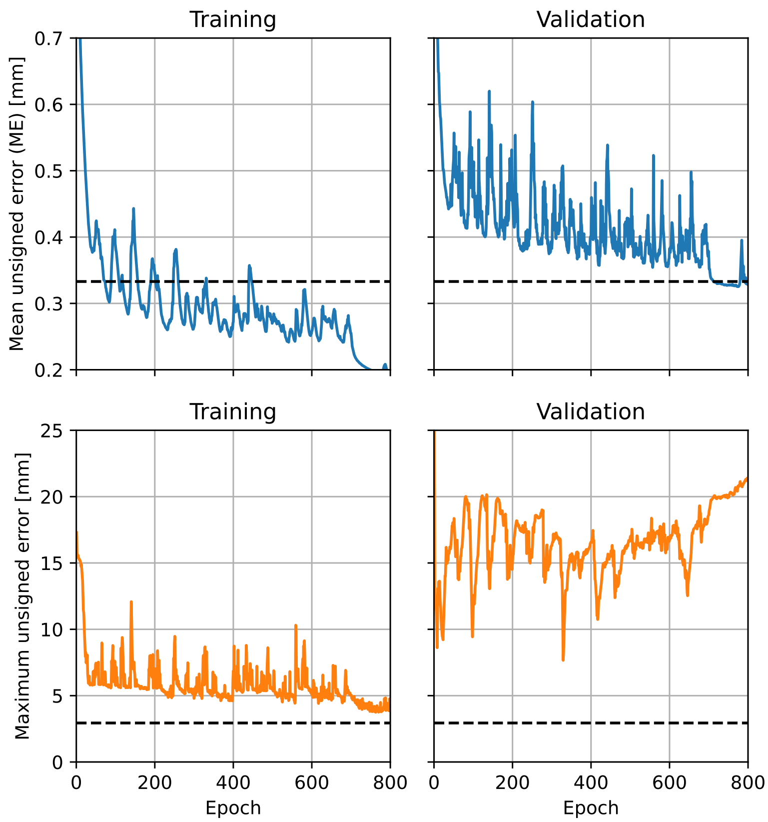

Fig. 4. The training and validation mean and maximum unsigned error during training for the best performing CNN model. The black dashed lines are the "satisfying" benchmark levels. We can see that the model gets a satisfactory mean unsigned error (ME), but the maximum is still off.

Architecture and hyperparameter optimization played a pivotal role in enhancing the model's accuracy. The final CNN model utilized a sophisticated arrangement of convolutional layers coupled with dropout regularization and batch normalization techniques. These elements collectively contributed to a robust model capable of discerning the intricate patterns associated with track irregularities from the vast and complex data derived from railway dynamics.

An extensive error analysis was conducted to dive into the predictive capabilities of the model and areas of improvement. This analysis was crucial in understanding the nuances of the model's performance, including the instances where it deviated from expected outcomes. Despite achieving high accuracy, the model faced challenges with maximum errors, especially in the validation data, prompting a detailed examination of error characteristics to identify potential model enhancements. The result of this analysis showed that the model makes the largest errors in regions with faulty sensor data. This is highlighted in Table 2, which shows key statistics for the model evaluated on the test data. The test data did not contain faulty sensor data. From the table, we see that the model also gets a satisfactory mean, but not a satisfactory maximum, on the test data. However, the aggregated maximum unsigned error is much smaller in the test data compared to the validation data errors seen in Fig. 4. Additionally, the model beats the state-of-the-art results from [9] for short wavelengths and the aggregated mean unsigned errors (ME). However, the model falls short of the 0.1 and 0.5 mm benchmarks for, respectively, the mean and maximum of unsigned errors.

Thus, to further improve the model, the focus should be on the data used to train the model. This might then eliminate the issues with missing sensor data.

	Mean [mm]	Maximum [mm]
Lateral left D1	0.14	2.54
Lateral right D1	0.13	2.62
Vertical left D1	0.27	2.23
Vertical right D1	0.28	2.56
Lateral left D2	0.20	3.55
Lateral right D2	0.18	3.49
Vertical left D2	0.30	2.28
Vertical right D2	0.31	2.63
Lateral left D3	0.35	**7.33**
Lateral right D3	0.33	**7.61**
Vertical left D3	**0.63**	**6.58**
Vertical right D3	**0.64**	**6.19**
Aggregates	0.31	**7.61**

Table 2. Mean and maximum of the unsigned test errors for each of the 12 output features. Values highlighted in red and in bold are those that exceed satisfactory levels. Recall that D1, D2, and D3 correspond to wavelengths of [3 m, 25 m], [25 m, 70 m], and [70 m, 200 m], respectively. From this, the mean unsigned errors for the three wavelength regions are 0.205 mm, 0.248 mm, and 0.486 mm, respectively.

4.2 Conformal Prediction

Integration of conformal prediction methods notably enhanced the predictive capabilities of the CNN model. The CV+ and CV-minmax methods were used to calculate the prediction intervals for the test data, which would improve the confidence in the model outputs for new predictions. The predictions are made as a mean aggregate of the six model instances trained for each validation segment. As depicted in Fig. 5, the CV+ and CV-minmax methods achieved high true value coverage rates, illustrating their effectiveness in encompassing data variability. The CV+ intervals are slightly narrower than those for CV-minmax but are overall very similar.

Table 3 presents the aggregate statistics for the $\alpha = 0.05$ intervals. It reveals that although CV-minmax offers higher coverage at 97.18 % compared to CV+'s 95.76%, it produces wider intervals on average (2.33 mm for CV-minmax versus 1.78 mm for CV+). Notably, as will be shown later, the CV-minmax method's prediction intervals are computed faster than those of CV+, an advantage for real-time applications.

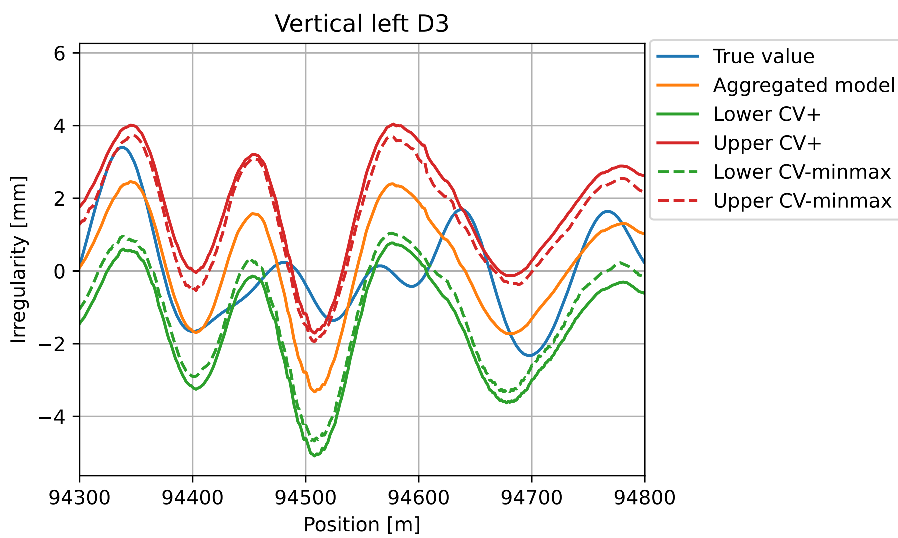

Fig. 5. Comparison of prediction intervals from CV+ and CV-minmax methods against the true values of track irregularities for the vertical D3 (long wavelengths) irregularities of the left rail. We see that the intervals often capture the true value, but there are instances, where they fail.

4.3 Compute Time

A crucial aspect of the CNN model's design was its ability to process and make predictions at a rate that exceeds the operational speeds of high-speed railway vehicles, which can reach speeds over 300 km per hour. The CNN model could process 35.30 km of test data in 60.72 seconds using a GTX 970. Thus, the model can process track data at a rate of 2093 km per hour. Meanwhile, for conformal

Measure	CV+	CV-minmax
True value coverage (%)	95.76	97.18
Average interval width (mm)	1.78	2.33
Maximum interval width (mm)	4.54	5.25

Table 3. Aggregate statistics for conformal prediction intervals using the CV+ and CV-minmax methods. CV+ produces narrower intervals, but they have a slightly lower coverage.

predictions, the CV-minmax and CV+ methods require 0.12 seconds and 12 minutes, respectively, to process the test data. This means they can process track data at rates of over 1,000,000 km and 176.5 km per hour, respectively.

Thus, the CV+ method cannot process sufficiently fast. However, the CNN model and CV-minmax demonstrated exceptional efficiency, capable of evaluating substantial lengths of track data within a constrained timeframe, thereby ensuring its applicability in real-time monitoring systems. We propose a CNN-based model with uncertainty quantified via conformal predictions as a solution for continuous real-time monitoring of railway track conditions.

5 Conclusion

This research ventured into the domain of using data-driven machine learning methods, with a focus on convolutional neural networks (CNNs) and conformal prediction, to predict railway track irregularities from the observed dynamics of in-service railway vehicles. The core achievement was the development of a predictive model that not only delivered satisfactory accuracy in detecting track irregularities but also incorporated conformal prediction to estimate the uncertainty of these predictions reliably. Satisfactory results are set as a mean unsigned error (ME) of 0.35 mm based on state-of-the-art results from related work. Our model has a mean unsigned error of 0.31 mm on the test set, thus improving the state-of-the-art results of [9]. Interestingly, the conformal prediction methodology achieved a high coverage of 97.18 % of the true values, with prediction intervals of an average width of 2.33 mm, thus ensuring a robust and reliable predictive framework.

However, it was noted that, while the prediction coverage was impressively high, the width of the intervals, though relatively small, indicates room for optimization to refine the precision further. These intervals were derived using the CV-minmax method, highlighting the potential for real-time application of this approach, given its ability to evaluate more than 1M km of track data per hour. Additionally, the CNN model could process track data at a rate of more than 2,000 km per hour. This efficiency underscores the feasibility of deploying this methodology in real-world settings, where it can serve as a cornerstone for continuous real-time monitoring of railway track conditions using in-service high-speed vehicles.

The journey to improve the accuracy and reliability of track irregularity detection through machine learning is far from over. Future endeavors can pivot around several key areas to push the boundaries of current achievements. Primarily, addressing the identified data issues will be crucial. This includes refining sensor data quality by removing or correcting data from faulty sensors and handling outliers more effectively. The model could, for instance, be made more robust so that it can allow for faulty sensors.

Further exploration of vehicle modeling offers a promising avenue for advancement. Transitioning the codebase to Julia has opened up new possibilities for using scientific computing methods. For example, delving into the domain of scientific machine learning, specifically through the lens of Neural Ordinary Differential Equations (NODEs), presents an exciting frontier. This approach could fundamentally change the way we model vehicle dynamics by integrating data-driven insights directly into the differential equations governing these dynamics.

We tried using transfer learning to simulate vehicle dynamics and pre-training the model on these simulated data. However, this did not produce the expected benefits in this study, suggesting a potential misalignment in data formatting or a lack of representation in the ODE system. Future research could aim to refine these aspects, potentially leading to breakthroughs in model performance and generalizability. Similarly, future work could try using physics-informed neural networks (PINNs) to enhance the model by incorporating physical laws directly into the learning process.

In sum, the groundwork laid by this project not only contributes to the current body of knowledge but also charted a course for future research to explore uncharted territories in railway track maintenance and safety through the lens of advanced machine learning techniques.

A Appendix

ID	Label	Unit	Description
0	Position	km	Position along the track
1	Velocity	km/h	Velocity of the railway vehicle
2	AccB1Y	m/s^2	Lateral acceleration of axle box 1
3	AccB1Z	m/s^2	Vertical acceleration of axle box 1
4	AccCR1Y	m/s^2	Lateral acceleration of bogie A at axle box 1
5	AccCR1Z	m/s^2	Vertical acceleration of bogie A at axle box 1
6	AccB2Y	m/s^2	Lateral acceleration of axle box 2
7	AccB2Z	m/s^2	Vertical acceleration of axle box 2
8	AccCR2Y	m/s^2	Lateral acceleration of bogie A at axle box 2
9	AccCR2Z	m/s^2	Vertical acceleration of bogie A at axle box 2
10	AccB3Y	m/s^2	Lateral acceleration of axle box 3
11	AccB3Z	m/s^2	Vertical acceleration of axle box 3
12	AccCR3Y	m/s^2	Lateral acceleration of bogie A at axle box 3
13	AccCR3Z	m/s^2	Vertical acceleration of bogie A at axle box 3
14	AccB4Y	m/s^2	Lateral acceleration of axle box 4
15	AccB4Z	m/s^2	Vertical acceleration of axle box 4
16	AccCR4Y	m/s^2	Lateral acceleration of bogie A at axle box 4
17	AccCR4Z	m/s^2	Vertical acceleration of bogie A at axle box 4
18	AccB5Y	m/s^2	Lateral acceleration of axle box 5
19	AccB5Z	m/s^2	Vertical acceleration of axle box 5
20	AccCR5Y	m/s^2	Lateral acceleration of bogie B at axle box 5
21	AccCR5Z	m/s^2	Vertical acceleration of bogie B at axle box 5
22	AccB6Y	m/s^2	Lateral acceleration of axle box 6
23	AccB6Z	m/s^2	Vertical acceleration of axle box 6
24	AccCR6Y	m/s^2	Lateral acceleration of bogie B at axle box 6
25	AccCR6Z	m/s^2	Vertical acceleration of bogie B at axle box 6
26	AccB7Y	m/s^2	Lateral acceleration of axle box 7
27	AccB7Z	m/s^2	Vertical acceleration of axle box 7
28	AccCR7Y	m/s^2	Lateral acceleration of bogie B at axle box 7
29	AccCR7Z	m/s^2	Vertical acceleration of bogie B at axle box 7
30	AccB8Y	m/s^2	Lateral acceleration of axle box 8
31	AccB8Z	m/s^2	Vertical acceleration of axle box 8
32	AccCR8Y	m/s^2	Lateral acceleration of bogie B at axle box 8
33	AccCR8Z	m/s^2	Vertical acceleration of bogie B at axle box 8
34	AccCSAY	m/s^2	Lateral acceleration of car body at bogie A
35	AccCSAZ	m/s^2	Vertical acceleration of car body at bogie A
36	AccCSBY	m/s^2	Lateral acceleration of car body at bogie B
37	AccCSBZ	m/s^2	Vertical acceleration of car body at bogie B
38	Curvatura	1/m	Curvature of the circle the track is forming

Table 4. Features in a sample of dynamics dataset – With provided labels and labels used in this project.

Bibliography

[1] Aslam Y, N S (2019) A review of deep learning approaches for image analysis. In: 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT), pp 709–714, https://doi.org/10.1109/ICSSIT46314.2019.8987922

[2] Balasubramanian V, Ho SS, Vovk V (2014) Conformal Prediction for Reliable Machine Learning: Theory, Adaptations and Applications. Elsevier Inc., https://doi.org/10.1016/C2012-0-00234-7

[3] Balouchi F, Bevan A, Formston R (2021) Development of railway track condition monitoring from multi-train in-service vehicles. Vehicle System Dynamics 59(9):1397–1417

[4] Barber RF, Candès EJ, Ramdas A, Tibshirani RJ (2021) Predictive inference with the jackknife+. Annals of Statistics 49(1):486–507, https://doi.org/10.1214/20-AOS1965

[5] Chudzikiewicz A, Bogacz R, Kostrzewski M, Konowrocki R (2017) Condition monitoring of railway track systems by using acceleration signals on wheelset axle-boxes. Transport (Vilnius, Lithuania) 33(2):555–566

[6] Clevert DA, Unterthiner T, Hochreiter S (2020) Fast and accurate deep network learning by exponential linear units (elus). arxiv 2015. arXiv preprint arXiv:151107289

[7] De Rosa A, Kulkarni R, Qazizadeh A, Berg M, Di Gialleonardo E, Facchinetti A, Bruni S (2021) Monitoring of lateral and cross level track geometry irregularities through onboard vehicle dynamics measurements using machine learning classification algorithms. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit 235(1):107–120

[8] En C (2017) 13848-5, railway applications–track–track geometry quality–part 5: geometric quality levels–plain line, switches and crossings. European Committee for Standardization, Brussels

[9] Hao X, Yang J, Yang F, Sun X, Hou Y, Wang J (2023) Track geometry estimation from vehicle–body acceleration for high-speed railway using deep learning technique. Vehicle system dynamics 61(1):239–259

[10] Ioffe S, Szegedy C (2015) Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: International conference on machine learning, pmlr, pp 448–456

[11] Kawasaki J, Youcef-Toumi K (2002) Estimation of rail irregularities. Proceedings of the American Control Conference 5:3650–3660, https://doi.org/10.1109/acc.2002.1024495

[12] Lee JS, Choi S, Kim SS, Kim YG, Kim SW, Park C (2012) Waveband analysis of track irregularities in high-speed railway from on-board acceleration measurement. Journal of solid mechanics and materials engineering 6(6):750–759

[13] Li C, He Q, Wang P (2021) Estimation of railway track longitudinal irregularity using vehicle response with information compression and bayesian deep learning. Computer-aided civil and infrastructure engineering 37(10):1260–1276

[14] Mittal S, Rao D (2017) Vision based railway track monitoring using deep learning. arXiv preprint arXiv:171106423

[15] Muñoz S, Ros J, Urda P, Escalona JL (2021) Estimation of lateral track irregularity through kalman filtering techniques. IEEE Access 9:60,010–60,025

[16] Naganuma Y, Kobayashi M, Okumura T (2010) Inertial measurement processing techniques for track condition monitoring on shinkansen commercial trains. Journal of mechanical systems for transportation and logistics 3(1):315–325

[17] O'Mahony N, Campbell S, Carvalho A, Harapanahalli S, Hernandez GV, Krpalkova L, Riordan D, Walsh J (2020) Deep learning vs. traditional computer vision. In: Advances in Computer Vision: Proceedings of the 2019 Computer Vision Conference (CVC), Volume 1 1, Springer, pp 128–144

[18] Ravitharan R (2019) Safer rail operations: Reactive to proactive maintenance using state-of-the-art automated in-service vehicle-track condition monitoring. 2018 International Conference on Intelligent Rail Transportation, Icirt 2018 p 8641587, https://doi.org/10.1109/ICIRT.2018.8641587

[19] Shafer G, Vovk V (2008) A tutorial on conformal prediction. Journal of Machine Learning Research 9:371–421

[20] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R (2014) Dropout: a simple way to prevent neural networks from overfitting. The journal of machine learning research 15(1):1929–1958

[21] Tsunashima H (2019) Condition monitoring of railway tracks from car-body vibration using a machine learning technique. Applied Sciences 9(13):2734

[22] Tsunashima H, Hirose R (2022) Condition monitoring of railway track from car-body vibration using time–frequency analysis. Vehicle System Dynamics 60(4):1170–1187

[23] Varona B, Monteserin A, Teyseyre A (2020) A deep learning approach to automatic road surface monitoring and pothole detection. Personal and Ubiquitous Computing 24(4):519–534

[24] Vovk V, Gammerman A, Shafer G (2005) Algorithmic learning in a random world. Springer US, https://doi.org/10.1007/b106715

[25] Wei X, Liu F, Jia L (2016) Urban rail track condition monitoring based on in-service vehicle acceleration measurements. Measurement 80:217–228

[26] Weston P, Roberts C, Yeo G, Stewart E (2015) Perspectives on railway track geometry condition monitoring from in-service railway vehicles. Vehicle system dynamics 53(7):1063–1091

[27] Yang C, Sun Y, Ladubec C, Liu Y (2020) Developing machine learning-based models for railway inspection. Applied Sciences 11(1):13