A deep neural network for broken rail prediction using multi-source data

Dr. Xin Wang & Dr. Xiang Liu, & Mulin Wan
Department of Civil and Environmental Engineering, Rutgers, The State University of New Jersey, Piscataway, New Jersey, USA

IHHA 2025
13TH INTERNATIONAL HEAVY HAUL ASSOCIATION CONFERENCE 2025

November 17-21, 2025 | The Broadmoor, Colorado Springs, CO, USA

Introduction

Broken rails are the main factor leading to freight-train derailments in terms of both the number of trains derailed and the number of cars derailed, which significantly jeopardizes the safety and efficiency of rail transportation. According to the Federal Railroad Administration (FRA) accident records, derailment takes up a significant portion of the average annual financial loss in the railroad network. Therefore, the prediction of broken rail risk is of great interest and practical value, which can assist railroad staff in preventive maintenance that aims to reduce risk and service disruptions by informing optimal planning of inspection and maintenance activities.

This study develops a spatial-temporal neural network model based on ResNet-Transformer architecture to predict the occurrence of broken rails one year in advance. The railroad data for this research includes infrastructure data, operational data, condition-related data, and maintenance activities. The proposed spatial-temporal model not only aids in proactive maintenance to improve the safety and reliability of rail transportation but also contributes to more strategic capital planning in the railroad industry.

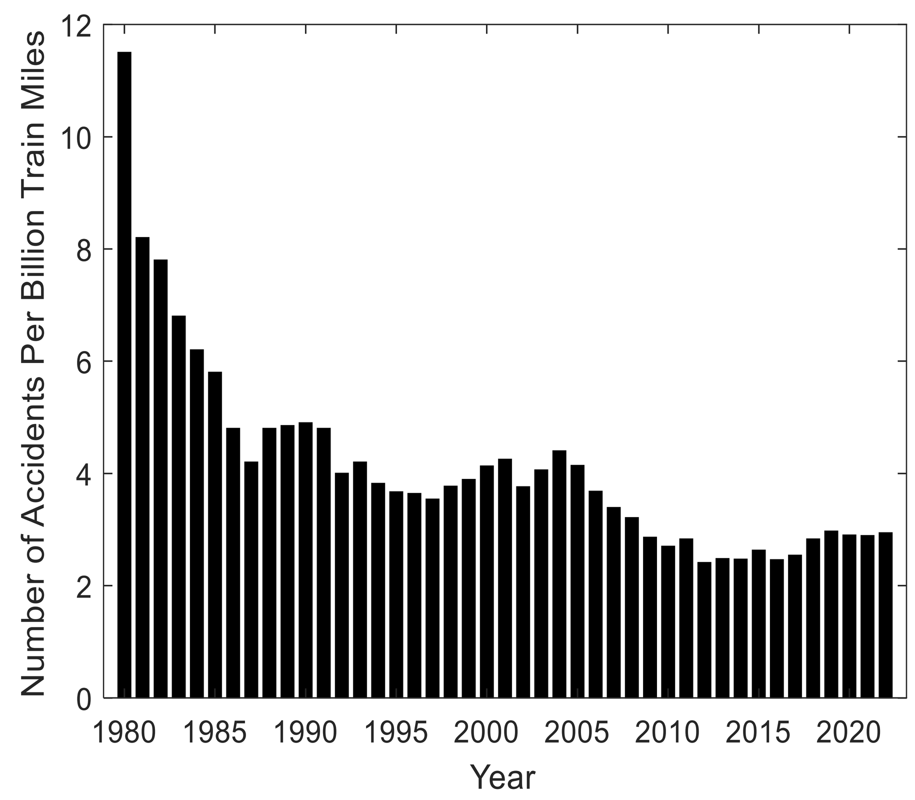

Figure 1. Accident Rate by Years

Methodology

This methodological framework mainly contains two parts: data preprocessing and model development. Data preprocessing generates the model's input and corresponding output. It utilizes 2D and 3D data structures as inputs of the model to offer a detailed representation of the spatial correlation and temporal dependencies in the data, which considers the impact of the adjacent track segments and historical railroad conditions. The prediction target is defined as the broken rail occurrence of the central track segment by the time of next year.

Subsequently, in the model development phase, ResNet is applied to extract spatial information from the raw data while the Transformer captures the temporal correlations and sequential patterns in the latent representations.

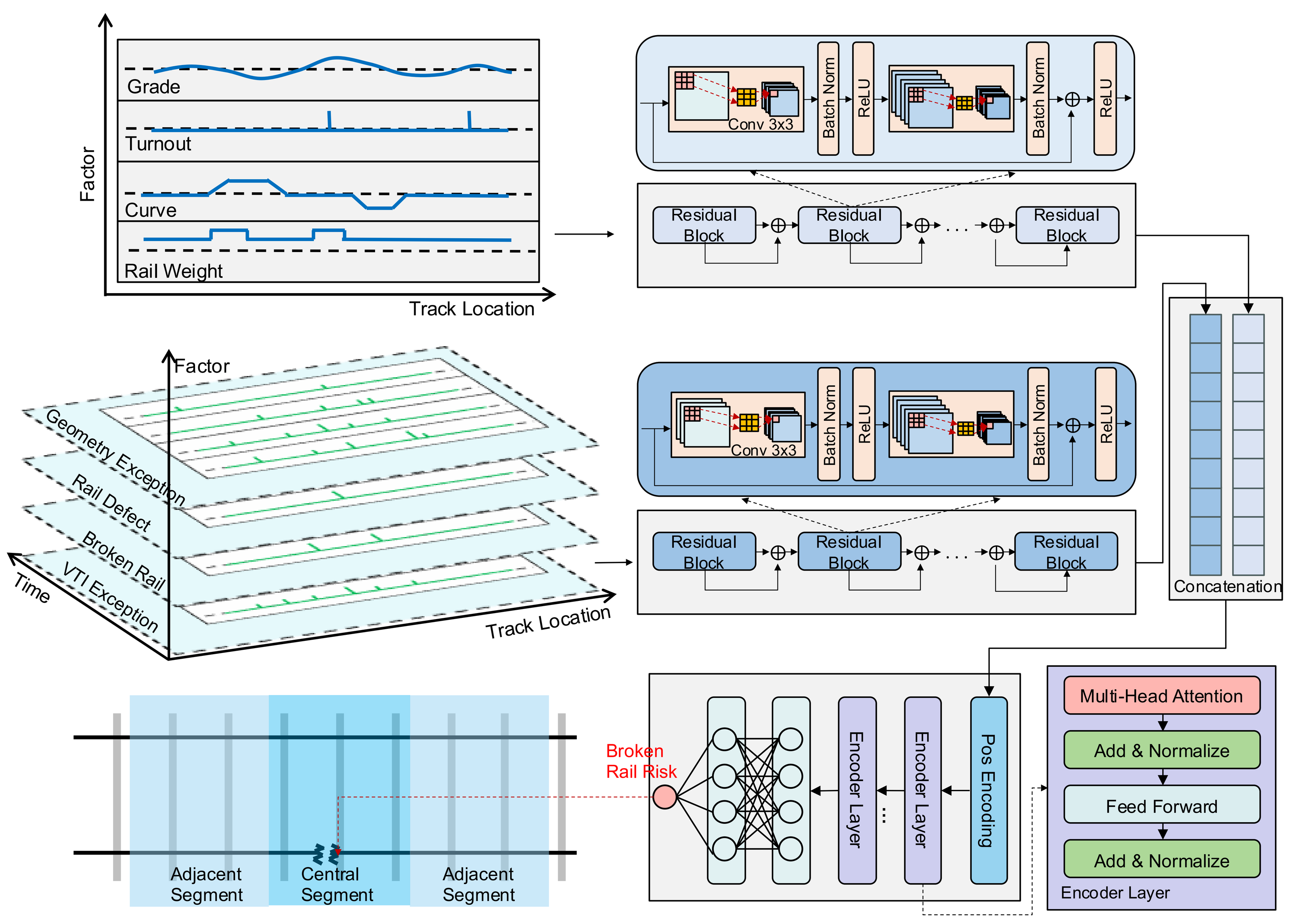

Figure 2. Methodological Framework for Broken Rail Prediction

Results

The studied railroad data covers about 20,000 miles of track spanning nine years, from 2013 to 2021. It can be classified into four categories: operational data, infrastructure data, condition-related data, and maintenance activities. The experiment results indicate that AUC values of the proposed model for the training, validation, and test set are 0.84, 0.81, and 0.81, respectively, demonstrating that the model has a relatively good performance and generalizes reasonably well to the validation and test set. The proposed model outperforms traditional machine learning approaches such as XGBoost, especially in identifying high-risk segments. When screening 10% of the highest-risk rail segments, the model can capture 41.6% of broken rails, compared to only 33.1% detected by XGBoost and 38.0% detected by ResNet-only model.

Additionally, the ablation experiment was conducted to evaluate the contributions of key components of our model, specifically focusing on the impact of the ResNet architecture. The result indicates that the ResNet-only structure can capture 38.0% of broken rails when 10% of the network is inspected. While the ResNet architecture excels at extracting features from spatial data, it may not fully leverage the temporal dependencies inherent in the combined 2D and 3D input formats.

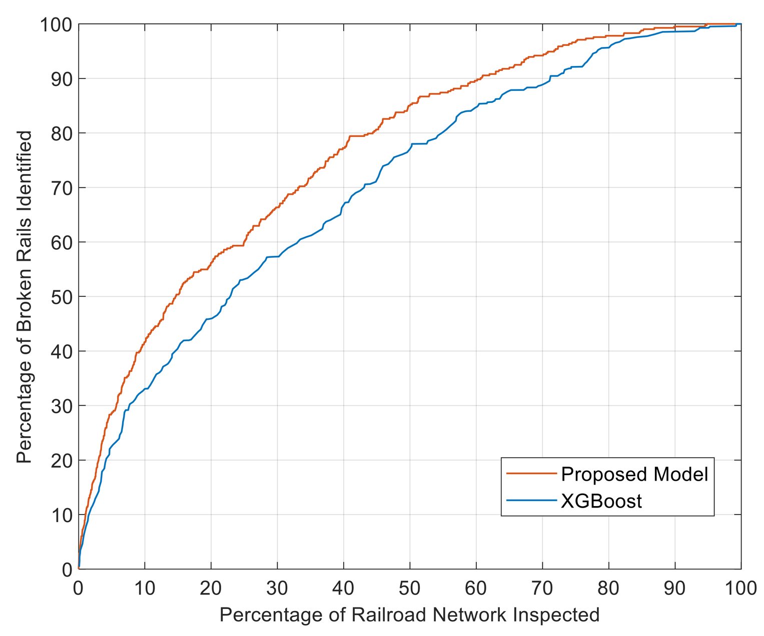

Figure 3. Performance of the Proposed Model and XGBoost

Figure 4. Predicted Top 10% Risk of Broken Rails over Studied Railroad Network

Conclusions

1). Development of a Pruned ResNet-Transformer Model: This study introduces a novel spatial-temporal prediction model for broken rails that integrates a pruned ResNet architecture with Transformer.

2). Enhanced Spatial and Temporal Feature Integration: This study utilizes microscale track segments to capture detailed spatial variations in rail properties (e.g., curvature and turnout presence). Additionally, the proposed approach incorporates time-series changes of track to consider the temporal variations that influence broken rail occurrences.

3). Practical Implications for Railroad Management: The proposed model provides a robust tool for railroad operators and maintenance planners to forecast broken rails more accurately.

This work is based on Xin Wang, Junyan Dai, & Xiang Liu (2025). A spatial-temporal neural network based on ResNet-Transformer for predicting railroad broken rails. Advanced Engineering Informatics, 65, 103126

For further information, please contact:
Dr. Xiang Liu, Rutgers University, xiang.liu@rutgers.edu

RUTGERS THE STATE UNIVERSITY OF NEW JERSEY

INTERNATIONAL HEAVY HAUL ASSOCIATION