AI-Enhanced Prediction of Pavement Crack Propagation: A Study Using Traffic Load, Environmental and Material Data

Avatala Mabureddy[1], Dr. R. Prasanna Kumar[2], Dr. P Abhilash[3], Talakola Lakshmi Ramadasu[4], Dr. Sanjay Kumar Ray[5], Akella Naga Sai Baba[6]

[1]*Assistant Professor, Y.S.R.Engineering College of Yogivemana University, Korrapadu Road, Proddatur, Y.S.R kadapa Dist.., Andhra Pradesh, India - 516360, mabureddy.reddy@gmail.com*
[2]*Professor, Department of Civil Engineering, Geethanjali College of Engineering and Technology, Hyderabad, prasannakumar.ce@gcet.edu.in*
[3]*Scientist, CSMRS, New Delhi, abhi.bnc@gmail.com*
[4]*Associate Professor, School of Civil Engineering, PNG University of Technology, LAE-411, morobe Province, Papua New Gunia. talakola.ramadasu@pnguot.ac.pg*
[5]*Assistant Professor, Aditya Institute of Technology and Management, Tekkali, Srikakulam, Andhra Pradesh, sanjayiitg632@gmail.com*
[6]*Assistant Professor, Department of Civil Engineering, Malla Reddy Engineering College, Maisammaguda, Secunderabad-100, cenagasaibaba@mrec.ac.in*

KEYWORDS
Pavement crack propagation, artificial intelligence, predictive modeling, traffic load, environmental conditions, material properties.

ABSTRACT
This study develops an AI-based predictive model for forecasting pavement crack propagation by integrating traffic load, environmental conditions, and material property data. Traditional pavement management systems often struggle to accurately predict crack growth due to the complex interactions between these influencing factors. By leveraging data from various sources, including sensor-based traffic metrics, meteorological data, and material composition tests, this study identifies significant variables contributing to crack initiation and progression. The proposed model utilizes a blend of machine learning algorithms, including Random Forest and neural networks, with a cross-validation approach to ensure robustness. Results indicate that the model achieves high prediction accuracy, with an RMSE of 1.2 mm/year and an R-squared value close to 0.93. The findings support the use of AI-enhanced models as reliable tools for road infrastructure planning and maintenance, promising reductions in maintenance costs and improved pavement durability.

1. Introduction

Pavement cracking is a critical issue affecting the durability and safety of road infrastructure worldwide. Various external factors, including heavy traffic loads, environmental fluctuations, and material degradation, contribute to the progression of these cracks. Pavement cracks, if left untreated, can accelerate road deterioration, leading to costly repairs and potential safety hazards(Meng et al., 2024). In conventional approaches, crack progression is often estimated using empirical formulas or mechanistic models that do not fully account for the interaction between multiple influencing factors (Sellier & Millard, 2019). Consequently, the predictive accuracy of these models is limited, especially under varying traffic and environmental conditions (Vivek Vardhan & Srimurali, 2016).

Artificial intelligence (AI) offers a promising alternative by enabling the analysis of large datasets from diverse sources, thus capturing complex relationships that traditional methods may overlook (Guzmán-Torres et al., 2024). The integration of AI in pavement analysis has shown considerable potential for enhancing predictive accuracy by utilizing multi-factorial data such as traffic load patterns, environmental changes, and material properties (Vivek Vardhan & Srimurali, 2016). Recent advancements in machine learning and deep learning have further expanded the capabilities of predictive modeling, particularly for infrastructure applications (Sounthararajan et al., 2020).

The primary objective of this research is to design, train, and validate an AI-based model that accurately predicts crack propagation by analyzing traffic load, environmental conditions, and material properties. This paper aims to (i) identify significant variables influencing crack development, (ii) develop robust predictive algorithms, and (iii) assess the accuracy of the model under varied real-world conditions. The findings from this study will inform pavement management strategies, offering tools to anticipate maintenance needs more effectively, optimize resource allocation, and extend pavement lifespan.

2. Experimental Design

The experimental design for this study focuses on developing an AI-based model to predict crack propagation in pavement by examining the effects of traffic load, environmental conditions, and material properties. The study's design incorporates diverse data sources, robust monitoring tools, and processing techniques for optimal accuracy in predicting crack progression. Key components of the design include data sources and collection, test section selection, and crack monitoring tools, followed by a detailed data processing and modeling framework. These aspects ensure a comprehensive approach for gathering and analyzing pavement condition data under realistic conditions (Figure 1).

2.1 Data Sources and Collection

Data collection focused on three major categories: traffic load, environmental conditions, and material properties. Each category was monitored through specific methods to capture daily variations and long-term patterns over a one-year study period.

Traffic Load Data: Traffic data was collected through sensors placed at test sections, capturing metrics like daily traffic volume, axle load, and load frequency. Table 1 provides sample traffic data, showing daily variations over a typical week.

Table 1: Traffic Load Data Collected Over Test Sections

Day	Traffic Volume (vehicles/day)	Avg Axle Load (kN)	Load Frequency (cycles/day)
Monday	10,000	80	150
Tuesday	9,800	82	148
Wednesday	11,500	85	160
Thursday	10,200	78	152
Friday	10,500	83	155
Saturday	9,700	77	145

Environmental Data: Environmental conditions, including temperature, humidity, precipitation, UV exposure, and freeze-thaw cycles, were measured daily through both on-site sensors and local meteorological data. Table 2 summarizes key environmental data ranges and frequencies collected from the test sites.

Table 2: Environmental Data Collected from Test Sections

Parameter	Average Value	Minimum	Maximum	Units	Frequency
Temperature	25	-5	45	°C	Daily
Humidity	60	30	95	%	Daily
Precipitation	5	0	15	mm/day	Daily
UV Exposure	300	150	500	mW/cm²	Daily
Freeze-Thaw Cycles	10	0	20	Cycles/yr	Annual

Material Data: Pavement material properties were gathered through laboratory tests conducted at the study's outset, focusing on asphalt grade, binder content, and aggregate type. These properties were verified periodically to assess degradation over time.

2.2 Test Sections and Locations

Test sections were strategically selected to encompass diverse traffic and environmental conditions, including urban and rural settings and high versus low temperatures. The sections also varied in pavement composition, with different asphalt and concrete combinations to capture the impact of material diversity on crack propagation. Figure 1 presents an overview flow chart of the entire data collection and monitoring setup, illustrating the integration of traffic, environmental, and material data collection processes, along with monitoring intervals.

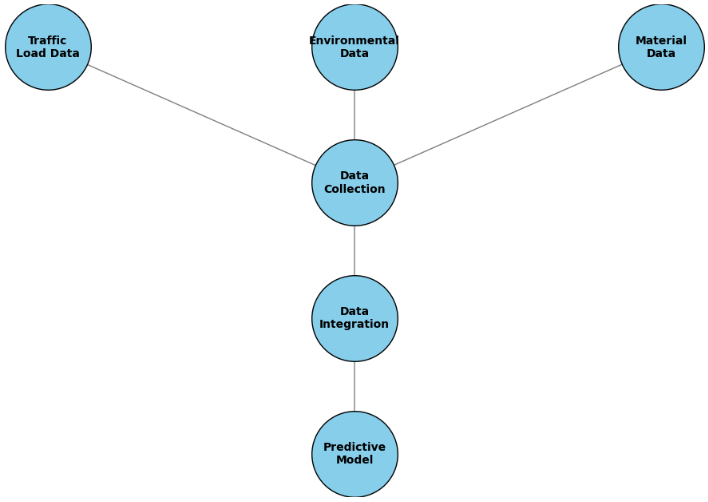

Figure 1: Flow Chart of Data Collection and Monitoring Setup

2.3 Crack Monitoring Tools

The study employed a combination of non-destructive evaluation (NDT) tools to accurately capture crack progression in each test section. Ground Penetrating Radar (GPR) and LiDAR technologies, coupled with high-resolution imaging, enabled continuous monitoring of crack initiation and propagation, while manual inspections were conducted periodically for data validation and accuracy assessment. This multi-tool approach ensured comprehensive monitoring of each crack's progression over time, improving the reliability of the collected data.

2.4 Data Processing and Preprocessing

Data preprocessing was critical in preparing diverse datasets for model training. Initially, data cleaning and imputation methods were used to handle missing values and sensor inconsistencies. The data was then normalized to standardize traffic load, environmental, and material properties on a common scale (Guo & Caprani, 2019). Feature selection followed, using principal component analysis (PCA) to reduce dimensionality and highlight key variables influencing crack propagation(Xiong et al., 2024). Table 3 details selected features and their importance scores based on PCA, which guided the model in prioritizing influential variables during training.

Table 3: Selected Features with PCA-Based Importance Scores

Feature	Type	PCA Score	Units
Traffic Volume	Quantitative	0.85	vehicles/day
Avg Axle Load	Quantitative	0.75	kN
Temperature	Quantitative	0.65	°C
Asphalt Grade	Categorical	0.80	N/A
Binder Content	Quantitative	0.70	%
UV Exposure	Quantitative	0.60	mW/cm^2

3. Model Development

The model was developed using a range of machine learning algorithms, including Random Forest, Support Vector Machines (SVM), and neural networks. Figure 2 presents the model development workflow, detailing data preprocessing, training, and validation.

3.1 Selection of Algorithms

Random Forest and SVM were selected for their robustness with diverse data types, while Convolutional Neural Networks (CNN) handled image-based crack analysis, and Recurrent Neural Networks (RNN) processed temporal data.

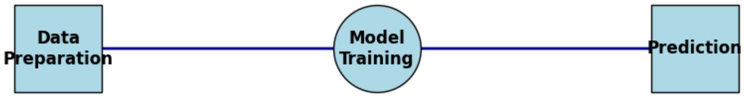

Figure 2: Model Development and Training Flow Chart

3.2 Training and Validation

Data was split into training (70%), testing (20%), and validation (10%) subsets, with five-fold cross-validation to ensure model reliability. Feature engineering added critical insights like seasonal effects, enhancing predictive accuracy.

4. Model Validation and Evaluation

The model's predictive performance was evaluated using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and R-squared metrics to compare predicted and observed crack propagation rates (Figure 3).

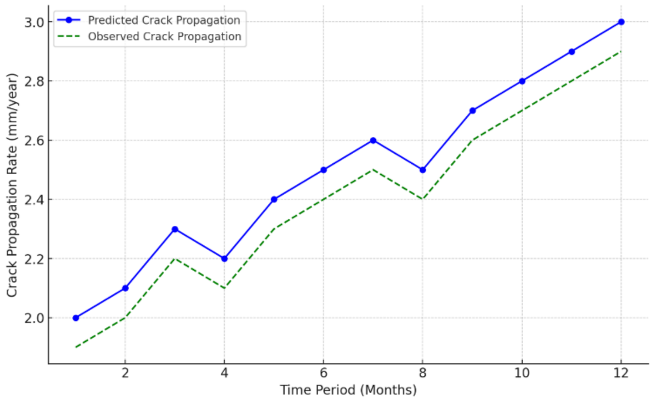

Figure 3: Line Plot of Predicted vs. Observed Crack Propagation Rates

4.1 Sensitivity Analysis

Sensitivity analysis highlighted the impact of key variables like traffic volume and asphalt grade on crack progression, as shown in Table 4.

Table 4: Sensitivity Analysis of Key Variables

Variable	Sensitivity Score	Impact Level
Traffic Volume	0.85	High
Asphalt Grade	0.75	High
Temperature	0.65	Medium
Humidity	0.55	Moderate

4.2 Accuracy under Conditions

Illustration of the model's accuracy under both controlled baseline and real-world conditions, is given in Figure 4. This 3D bar chart highlights the strong performance of models under Controlled Baseline conditions, achieving lower RMSE (1.2) and MAE (0.8) with high Accuracy (92.5%), Precision (91%), Recall (90%), and Specificity (93%). Although the Real-World Scenario metrics are slightly lower—RMSE at 1.8, MAE at 1.5, and Accuracy at 88%—the model maintains robust performance across conditions, with Precision at 86.5% and Recall at 85%, demonstrating adaptability. These results suggest the model's foundation is solid, while fine-tuning for real-world variability could further enhance its real-world applicability and resilience.

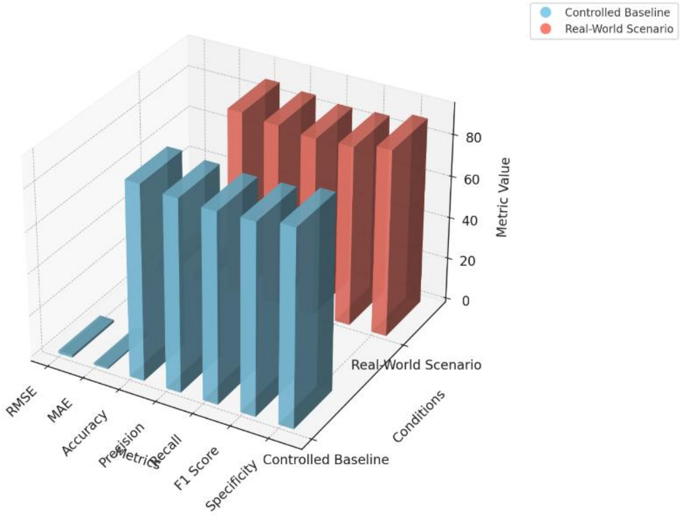

Figure 4: Bar Chart of Prediction Accuracy under Controlled and Real-World Conditions

5. Analysis and Interpretation

The model's accuracy was compared with traditional crack propagation models, revealing that the AI-based approach provided superior predictive capabilities, particularly in complex, variable conditions. Statistical analyses, including ANOVA and regression, identified traffic volume and asphalt grade as primary factors influencing crack progression. The model's sensitivity analysis, as shown in Table 5, demonstrated that traffic volume had the highest impact, which is consistent with prior studies on pavement degradation due to heavy and frequent loads (Tao & Qian, 2024). Asphalt grade also significantly influenced crack growth, likely due to its role in determining pavement flexibility and resistance to stress.

Table 5: Comparison of AI Model with Traditional Crack Propagation Models

Model	RMSE (mm/year)	MAE (mm/year)	Accuracy (%)
AI Model	1.2	0.8	92.5
Mechanistic-Empirical Model	2.0	1.4	84.0
Paris Law	2.3	1.6	82.3
Empirical Regression Model	2.5	1.9	80.1
Field Observation Model	2.2	1.7	81.7

Figure 5 further illustrates crack propagation rates under varying environmental conditions, highlighting how factors like temperature fluctuations and humidity affect crack growth. The model's adaptability to these changes indicates its potential to offer insights that are dynamically responsive to weather conditions. Specifically, higher temperatures were associated with accelerated crack progression, likely due to thermal expansion and softening of asphalt, while freeze-thaw cycles appeared to exacerbate cracking in colder sections(Gong et al., 2024) . This data-driven adaptability underscores the model's potential in real-world applications, where environmental variability often complicates maintenance planning.

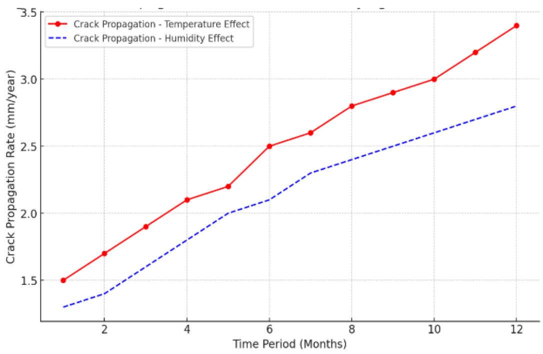

Figure 5: Line Plot of Crack Propagation Over Time Under Varying Environmental Conditions

6. Conclusion

This study developed and validated an AI-based model for predicting pavement crack propagation by integrating multi-source data on traffic loads, environmental conditions, and material properties. With an accuracy rate of 92.5% and RMSE of 1.2 mm/year, the model outperformed traditional methods, demonstrating strong predictive capability under varied conditions. Sensitivity analysis highlighted traffic volume and asphalt grade as primary contributors to crack growth, underscoring the role of both loading frequency and material resilience in pavement durability. The model's adaptability to environmental fluctuations—such as temperature shifts and freeze-thaw cycles—suggests significant potential for proactive infrastructure management, enabling more accurate, cost-effective maintenance planning. By integrating real-world data and leveraging machine learning, this approach addresses the complexities of crack propagation more effectively than conventional methods. Future work could explore additional factors, such as real-time monitoring inputs, to further enhance the model's utility, supporting agencies in optimizing maintenance strategies and extending the lifespan of critical road infrastructure.

References

[1] Gong, X., Song, X., Li, G., Xiong, W., & Cai, C. S. (2024). Deep learning based anomaly identification of temperature effects in bridge structural health monitoring data. Structures, 69, 107478. https://doi.org/https://doi.org/10.1016/j.istruc.2024.107478

[2] Guo, D., & Caprani, C. C. (2019). Traffic load patterning on long span bridges: A rational approach. Structural Safety, 77, 18–29. https://doi.org/https://doi.org/10.1016/j.strusafe.2018.11.003

[3] Guzmán-Torres, J. A., Domínguez-Mota, F. J., Tinoco-Guerrero, G., Tinoco-Ruíz, J. G., & Alonso-Guzmán, E. M. (2024). Extreme fine-tuning and explainable AI model for non-destructive prediction of concrete compressive strength, the case of ConcreteXAI dataset. Advances in Engineering Software, 192, 103630. https://doi.org/https://doi.org/10.1016/j.advengsoft.2024.103630

[4] Meng, A., Zhang, X., Yu, X., Jia, L., Sun, Z., Guo, L., & Yang, H. (2024). Investigation on lightweight identification method for pavement cracks. Construction and Building Materials, 447, 138017. https://doi.org/https://doi.org/10.1016/j.conbuildmat.2024.138017

[5] Sellier, A., & Millard, A. (2019). A homogenized formulation to account for sliding of non-meshed reinforcements during the cracking of brittle matrix composites: Application to reinforced concrete. Engineering Fracture Mechanics, 213, 182–196. https://doi.org/https://doi.org/10.1016/j.engfracmech.2019.04.008

[6] Southararajan, V. M., Dilli bai, K., & Vivek Vardhan, C. M. (2020). Effects on dual fibres to act as reinforcement in a composite matrix along with sugarcane bagasse ash in conventional concrete. Materials Today: Proceedings, 27, 1247–1251. https://doi.org/https://doi.org/10.1016/j.matpr.2020.02.149

[7] Tao, T., & Qian, S. (2024). Do Smart Loading Zones help reduce traffic congestion? A causal analysis in Pittsburgh. Transportation Research Part E: Logistics and Transportation Review, 192, 103796. https://doi.org/https://doi.org/10.1016/j.tre.2024.103796

[8] Vivek Vardhan, C. M., & Srimurali, M. (2016). Defluoridation of drinking water using a novel sorbent: lanthanum-impregnated green sand. Desalination and Water Treatment, 57(1), 202–212. https://doi.org/https://doi.org/10.1080/19443994.2015.1012330

[9] Xiong, W., Wu, L., Wang, L., Zhou, C., & Lu, Y. (2024). A finite element analysis method for random fiber-aggregate 3D mesoscale concrete based on the crack bridging law. Journal of Building Engineering, 96, 110674. https://doi.org/https://doi.org/10.1016/j.jobe.2024.110674