Proceedings of the 11th World Congress on New Technologies (NewTech'25)
Paris, France - August, 2025
Paper No. ICCEIA 162
DOI: 10.11159/icceia25.162

Understanding Steel Bridge Corrosion through Model Interpretation on Nationwide Inspection Data

Yongpeng Zhao[1], Kouichi Takeya[1], Yuichi Ito[1], Eiichi Sasaki[1]
[1] Institute of Science Tokyo
2-12-1 Ookayama, Meguro-ku, Tokyo, Japan
zhao.y.0782@m.isct.ac.jp, takeya.k.aa@m.titech.ac.jp, ito.y.ca@m.titech.ac.jp, sasaki.e.55bb@m.isct.ac.jp

Abstract - Considering the increasing number of aging infrastructure and the declining availability of skilled technicians in Japan, there is a growing demand for more efficient maintenance strategies. Leveraging information from existing inspection data presents a promising approach to address this challenge. This study proposes a novel model-interpretation based framework that integrates multiple databases to quantitatively evaluate the effects of internal structural conditions and environmental factors on the corrosion of steel bridge main girders. By enhancing the interpretability of predictive models, the proposed framework provides actionable insights to support targeted data-driven maintenance planning. The proposed approach shows potential to be broadly applicable for the maintenance of various civil engineering structures, contributing to the development of more efficient inspection and maintenance programs, which can rapidly adapt to changing environmental conditions.

***Keywords*:** corrosion, steel bridge, deterioration model, machine learning, model interpretation.

1. Introduction

Road infrastructure built in Japan during the post-war economic boom is aging while the maintenance workforce continues to decline [1] [2]. In response, standardized inspections were mandated in 2014, resulting in a nationwide data platform xROAD (the nationally unified data platform) to be constructed. This platform remains under construction and is based on the digital road map database (DRM-DB), which links structural data, ETC2.0, and other databases, and promotes the development of databases for each road facility. Integrating these platforms with publicly available environmental and bridge data is expected to help promote the development of data-based technologies that can be used in various fields beyond maintenance and management.

Machine learning (ML) offers promising tools to extract insights from inspection data. However, while complex models often improve prediction accuracy, their lack of interpretability limits practical use. Interpretation techniques help demonstrate key decision factors in these complex models [3], improving transparency and applicability for infrastructure management.

Several studies have explored bridge deterioration using bridge inspection data from various methods. Minami et al. [4] identified coastline proximity and material of the structure as key factors in Ishikawa Prefecture. Okazaki et al. [5] emphasized the role of age in crack propagation for concrete girders. Miao [6] applied artificial neural networks (ANN) and sensitivity analysis to predict deterioration with approximately 65% accuracy. Santos et al. [7] used Markov and ANN models for inspection interval optimization on over 10,000 Brazilian bridges. Igarashi and Abe [8] combined ensemble learning with interpretability techniques to assess crack severity but noted limitations in classification-based approaches. Saito [9] used multiple regression on public data to evaluate geographic and weather-related factors, highlighting the need for bridge-type-specific analysis. While these studies have contributed valuable insights, several limitations remain. Firstly, many approaches prioritized predictive accuracy at the cost of interpretability or relied on limited regional datasets, hindering their generalization ability and practical use for maintenance planning. Secondly, few studies have focused on the corrosion of steel bridge main girders, despite their critical role in structural safety.

To address these gaps, this study has developed a model-interpretation based framework combining bridge inspection records with publicly available structural and environmental data. By focusing on steel bridge main girders and applying

DOI: 10.11159/icceia25.162

easily interpretable machine learning techniques, this research aims to quantify the influence of key factors on corrosion progression and provide actionable insights for targeted, data-driven maintenance planning.

2. METHODOLOGY

2.1. Corrosion mechanism

Environmental factors and bridge design influence the development of corrosion. Steel corrodes if water and oxygen are present, while salts and acids further accelerate this process. Corrosion reaction rates accelerate with rising temperatures, especially in environments contaminated by sea salt in coastal areas and de-icing salts in mountainous regions. Structural differences also affect corrosion rates, especially in vulnerable areas, including girder ends and member crossings. Direct sunlight accelerates paint deterioration and increases corrosion risk in the areas exposed. The following data collection is based on the analysis of factors affecting corrosion.

2.2. Data collection and integration

In this study, multiple publicly accessible datasets were integrated to obtain comprehensive information on bridge characteristics, environmental conditions, and bridge maintenance history. The primary source of inspection data is the national road facility inspection database, part of the xROAD platform, which has been publicly available since mid-2022. This centralized database promotes technological advancement, efficient maintenance planning, and academic research by providing access to periodic inspection records. It includes detailed damage assessments categorized, as well as records of various major bridge components, along with traffic volumes and large vehicle mixing rates.

Additional structural and maintenance details, such as bridge types and painting history, were obtained from the MICHI (Bridge management chart and road management data), which provides access via APIs and offers downloadable data on bridge specifications, repairs, and painting records.

Meteorological conditions were represented using data from AMeDAS (The Automated Meteorological Data Acquisition System), operated by the Japan Meteorological Agency (JMA). Two datasets were used: the nearest weather station's long-term average values and high-resolution mesh annual normal values interpolated from AMeDAS data. These datasets provided localized climate information for all bridges in the dataset based on their geographic location.

To further capture environmental exposure, two geographic factors—elevation and distance from the shoreline—were incorporated. Elevation data was retrieved using the data from the Geospatial Information Authority service of Japan (GSI), while offshore distances were calculated using shoreline data and bridge coordinates.

2.3. Data Preprocessing

2.3.1. Deterioration Indicator

The inspection records used for this study document bridge damage taken at three different levels: full bridge, per span, and component element. Among these, component level damage can be assessed using a quantitative criterion, whereas evaluations at the span and bridge levels rely more heavily on the experience and subjective judgment of an inspector.

Indicators of existing bridge health, such as soundness level, provide a broad view into overall bridge condition but fail to capture the severity of specific deterioration types in individual components. However, the available data that details the damage level of individual component elements is very fragmented and inconvenient to use. This hinders the development of targeted maintenance strategies, as understanding specific deterioration is essential for prioritizing interventions and allocating resources efficiently. This study has considered several conventional bridge health indicators, including the maximum damage level, soundness level and the damage level development rate. However, these indicators often lack sensitivity to damage distribution or require subjective interpretation by inspectors. A new indicator, the damage score per unit area (S_{aa}), has therefore been developed in this study to overcome these shortcomings.

This novel indicator S_{aa} introduced above, converts component-level damage assessments into numeric scores using the classification scheme shown in Table 1 and normalizes the total damage score by the area (length × width) of the bridge. This allows consistent comparisons to be made across structures of varying size. S_{aa} targets specific types of deterioration for specific components, offering a fine-tuned health assessment at component level. Compared with traditional soundness levels, S_{aa} has three key advantages: (1) it enables detailed and component-specific maintenance scheduling, (2) it offers continuous values instead of broad categories, allowing better distinction between damage severity, and (3) it minimizes reliance on subjective inspector judgment.

Proceedings of the 11th World Congress on New Technologies (NewTech'25)
Paris, France - August, 2025
Paper No. ICCEIA 162
DOI: 10.11159/icceia25.162

Table 1: Damage Level to Score. The first three lines are the individual grades and their definitions, and the last line is the score corresponding to a particular damage level.

Damage Level	a	b	c	d	e
Depth of Damage	No damage	Shallow	Shallow	Deep	Deep
Area of damage	No damage	Small	Large	Small	Large
Score	0	1	2	3	4

Deep: Significant expansion or obvious reduction in thickness is visible on the surface of the steel.
Shallow: Rust is superficial, and no significant reduction in thickness is visible.
Large: Rust is observed in the entire area, or multiple rusting areas are spreading.
Small: The damaged area is small and localized.

2.3.2. Data cleaning and filtering

The original dataset consisted of 32,174 inspection records of both steel and concrete bridges from two rounds of nationwide surveys conducted between 2014 and 2018, and 2019 to 2023. Among these records, complete information from both inspections was available for 12,279 bridges, and these were merged into a single dataset. As this study focuses on the corrosion behaviour of steel bridge main girders and to reduce noise, the samples were limited to those from steel bridges. Among 3,981 steel bridge samples, bridges with only simple girder structures (n = 2,005) were selected as these structures are more uniform in behaviour and easier to compare. Since coating materials affect the development of corrosion, only samples (n = 1,365) with clearly recorded painting histories were included. Finally, to ensure consistency in the coating conditions across samples and to avoid the influence of mixed coatings, samples with only phthalic resin coating (n = 375) were selected as this was the most common coating type in the dataset. Fig.1 shows the distribution of the selected bridges.

This cleaning and filtering process was carried out to reduce data noise while ensuring the selected dataset retains structural diversity and sufficient generality for steel girder bridges.

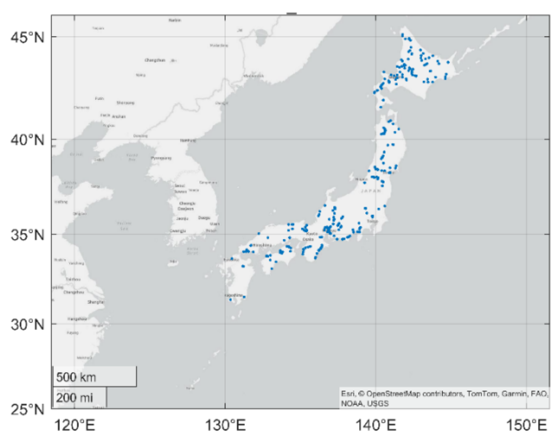

Fig.1: The distribution of the final 375 samples.

2.3.3. Normalization

The values of the considered features were transformed into a comparable scale to enhance the efficacy of model training. Based on assessments of skewness, two normalization methods were employed: log scaling and z-score scaling. Log scaling was applied when the skewness exceeded one, whereas z-score scaling was adopted when skewness was less than or equal to one. Since some features may possess values less than or equal to zero, which cannot be subjected directly to log scaling, a shifting process was executed before scaling. For a given feature x, if its minimum value, denoted as min(x), is less than 1, all feature values were adjusted by adding 1 - min(x) to ensure that the transformed values exceeded zero.

2.3.4. Dimensionality reduction

Proceedings of the 11th World Congress on New Technologies (NewTech'25)
Paris, France - August, 2025
Paper No. ICCEIA 162
DOI: 10.11159/icceia25.162

Principal Component Analysis (PCA) is a dimension reduction technique that can mitigate the impact of feature correlation on models. By identifying orthogonal principal components (PCs) and linear combinations of the original features PCA transforms high-dimensional data into a lower-dimensional space while retaining maximum variance. However, as it involves some loss of original data information, model interpretability is reduced. Therefore, PCA is applied only to variables with significant correlation.

2.4. Model Selection and Evaluation

Considering the small dataset size, this study avoided overly complex models. This study focuses on commonly used nonlinear algorithms that are effective in capturing nonlinear patterns and feature interactions, namely Support Vector Machine (SVM), Artificial Neural Networks (ANNs), Random Forest (RF), and Gaussian Process Regression (GPR). Hyperparameter optimization enhances performance, which mean squared error (MSE) evaluates. Grid Search is applied to both ANN and RF, while Random Search is used for GPR and SVM to balance search efficiency with model complexity. A 5-fold cross-validation was conducted for the accurate assessment of the performance of the model.

2.5. Model Interpretation

Model interpretation is a critical tool to help clarify the underlying mechanisms of machine learning models, and facilitate prediction validation, model refinement, and extraction of actionable insights. While interpretable models such as decision trees permit direct analysis of feature contributions, more complex models require advanced interpretability techniques. Among these techniques, Shapley Additive Explanations (SHAP) [10], proposed in 2017 and grounded in game-theoretic Shapley values, has gained prominence for its model-agnostic nature and capacity to provide local and global interpretative insights. Through interpretation algorism, it has become possible to quantify feature importance, capture feature interactions, and understand their influence on model outputs.

3. Results and Discussion

3.1. Dataset

There were 375 samples retained after cleaning and filtering, where all feature values were normalized to an approximate range. Correlation analysis revealed strong interdependencies among meteorology-related features (Fig. 2). Features that exhibited correlation coefficients ≥0.7 with at least two other features were selected for PCA. PC1 and PC2 are the first and second principal components derived through PCA applied to highly correlated meteorological features. These 2 Principal components explained 94.57% of the total variance, effectively capturing most of the information in the original six features. Accordingly, the original variables were replaced by PC1 and PC2 in subsequent analyses.

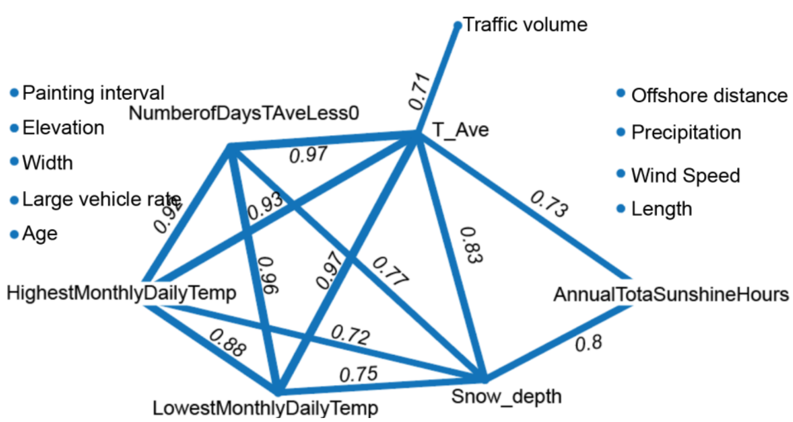

Fig.2: Feature Correlation Network. Features are connected when their correlation coefficient is greater than 0.7.

The target variable S_{aa}, introduced in this study, quantifies the severity of corrosion damage by normalizing damage scores across bridge area. S_{aa} provides a continuous and component-specific measure of deterioration, enhancing objectivity compared with traditional categorical indicators like soundness level.

DOI: 10.11159/icceia25.162

Although the final analytical set only included 375 bridges, which may appear limited, this sample was intentionally selected to maintain uniformity across structure and coating and thus improve internal validity. A five-fold cross-validation was utilized during the model training step to reduce the risk of overfitting due to limited sample size.

Table 2: Summary statistics of input features used for model training

	Before Normalization						After Normalization				
Features	Min	Median	Max	Mean	Std	Skewness	Min	Median	Max	Mean	Std
Length	7.4	30.0	521.1	46.0	51.9	4.6	2.0	3.4	6.3	3.5	0.7
Width	4.1	10.3	51.5	11.4	4.6	3.7	1.4	2.3	3.9	2.4	0.3
Age	19.0	46.0	82.0	45.3	9.0	0.0	-2.9	0.1	4.1	0.0	1.0
Traffic volume	35	8448	89697	13910	13911.0	1.8	3.6	9.0	11.4	9.0	1.2
Large vehicle rate	2.5	19.1	69.4	20.3	9.3	0.7	-1.9	-0.1	5.3	0.0	1.0
Offshore distance	0.1	24.4	106.9	28.9	24.4	1.0	-1.2	-0.2	3.2	0.0	1.0
Elevation	0.0	77.0	1053.1	149.0	189.1	2.2	0.0	4.4	7.0	4.1	1.5
Precipitation	778.9	1600.2	3469.0	1734.3	574.8	0.8	-1.7	-0.2	3.0	0.0	1.0
Wind Speed	0.0	2.0	4.9	2.0	0.8	0.2	-2.4	0.0	3.5	0.0	1.0
Painting interval	3.0	20.0	50.0	20.4	7.5	0.3	-2.3	0.0	4.0	0.0	1.0
PC1							-3.9	-0.4	4.8	0.0	2.7
PC2							-2.3	-0.2	2.6	0.0	0.9
S_{aa}	0.0	0.1	0.8	0.1	0.1	2.7	0.0	0.1	0.8	0.1	0.1

3.2. Model

Table 3 shows the minimum cross-validation MSE and hyperparameters, where GPR was identified as the most suitable approach. The model for interpretation was trained using the complete dataset and specified hyperparameters, achieving high accuracy (MSE=1.69×10^{-6}, MAE=4.25×10^{-4}, R^2=0.99). Its performance based on/considering the full dataset exceeded that of cross-validation, raising concerns of overfitting, which limits accurate predictions on unseen data. However, since hyperparameter tuning guided by cross-validation and evaluation occurred on the test set, the performance difference is likely due to insufficient data rather than overfitting. To improve model performance, it is important to ensure an adequate sample size, minimize conditional biases during data screening, and perform careful hyperparameter tuning within the training process.

Table 3: Minimum MSE (CV) results from hyperparameter tuning.

Model	MSE(CV)	Hyperparameters	
ANN	0.00656	Activation	"sigmoid"
		Layer size	[60]
		Lambda	7.7e-4
RF	0.00637	Min Leaf Size	3
		Max Num Splits	155
GPR	0.00606	Sigma	0.005301
		Basis Function	"linear"
		Kernel Function	"ARDRationalQuadratic"
SVM	0.00671	Epsilon	0.466
		Kernal Founction	"gaussian"
		Box Constraint	638.07

3.3. Interpretation

Table 4 presents Spearman's correlation coefficients between each predictor feature and its Shapley value. Age, elevation, and length show the highest correlations.

DOI: 10.11159/icceia25.162

Shapley analysis results (Fig. 3) reveal that shorter and narrower bridges tend to have higher Shapley values, and this is likely due to the greater proportional influence of highly corrosive girder-end areas on main girder. The effect of corrosion diminishes with an increase in bridge size. Bridge age correlates positively with Shapley values, indicating an increased risk in deterioration over time. While proximity to the coastline does not directly influence Shapley values, low-elevation areas exhibit higher values. This reflects a moderate correlation between elevation and offshore distance ($\rho = 0.64$), where corrosion susceptibility increases in coastal environments. Wind speed is negatively associated with Shapley values, likely due to its role in moisture dissipation. These findings suggest that maintenance should prioritize older, shorter bridges in areas with low-elevation, low-wind, and high-precipitation. As shown in Fig. 4, the five most influential factors based on mean absolute Shapley values are PC2 (a dominant meteorological principal component), traffic volume, annual precipitation, bridge width, and bridge length.

Table 4: The correlation between feature value and its Shapley value. All correlation coefficients were statistically significant with p-values less than 0.05.

Feature	Correlation coefficient
Length	-0.75
Width	-0.52
Age	0.95
Traffic volume	-0.39
Large vehicle rate	0.28
Offshore distance	-0.45
Elevation	-0.81
Precipitation	0.54
Wind Speed	-0.58
Painting interval	0.48
PC1	0.74
PC2	0.57

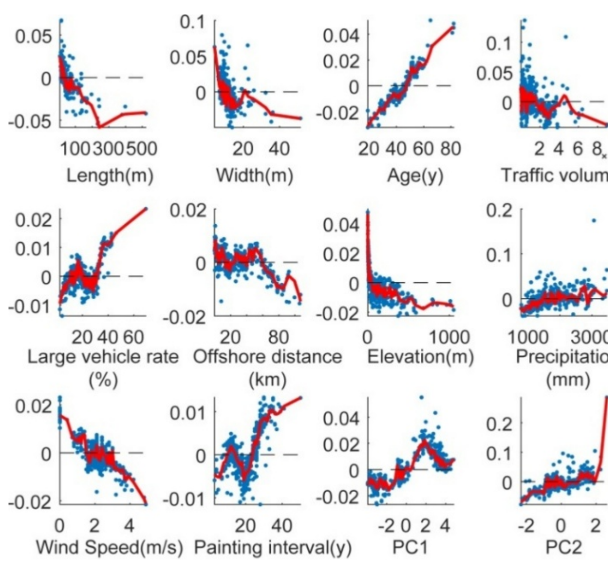

Fig. 3: Scatter diagrams of each feature value and its Shapley value and moving average curve. The y-axis indicates a feature's Shapley value.

In this analysis negative correlation was found between traffic volume and corrosion severity. Higher traffic volumes were expected to cause greater emissions of acidic gases, which accelerate corrosion. Another possible reason is the bridge's proximity to coastal areas with milder climates, which may also promote more aggressive corrosion. However, bridges with higher traffic volumes showed lower levels of corrosion. This suggests that heavily trafficked bridges may

DOI: 10.11159/icceia25.162

receive frequent inspections and maintenance, mitigating corrosion progression despite harsher conditions. Further investigation is needed to clarify the mechanisms and maintenance practices leading to this correlation.

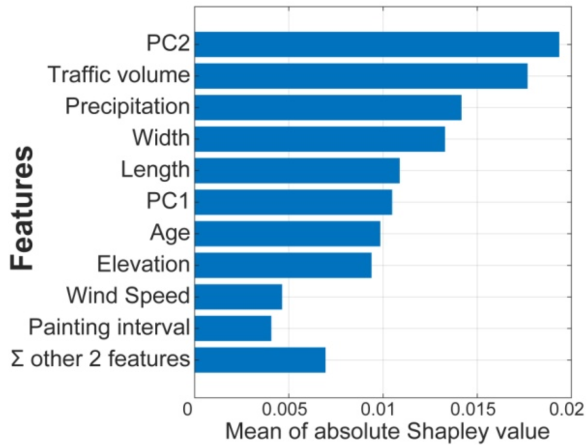

Fig.4: Feature importance rank based on average absolute Shapley value.

The results highlight that environmental conditions and traffic volume play predominant roles in bridge deterioration. Due to the small sample size in this study, the generalization ability of these results is open to discussion.

4. Conclusion

This study proposed a data-driven analytical framework to identify and interpret key factors contributing to the corrosion of steel bridge main girders, by integrating xROAD with publicly available environment databases. By constructing a corrosion severity indicator and applying an interpretation algorithm on machine learning models, the analysis in this study has revealed environmental variables and PC2 play a key role in corrosion progression. The proposed framework has demonstrated predictive performance and high interpretability, making it a valuable tool for developing targeted maintenance strategies and supporting decision-making in infrastructure asset management.

Nonetheless, several limitations remain as although the selected dataset was thoroughly filtered for consistency, it was small and remained spatially imbalanced. This could have introduced biases and constrained the generalization ability of the model. In addition, key data sources such as inspection imagery, structural details, and objective damage metrics were not incorporated, limiting the model's capacity to capture complex deterioration mechanisms.

Future work could focus on expanding the dataset by integrating richer and more diverse data set, including detailed inspection records, structural health monitoring data, and 3D point cloud models. Incorporating GIS-based spatial analysis and real-time monitoring techniques may further enhance the proposed framework's accuracy, robustness, and scalability, ultimately supporting more proactive and cost-effective bridge maintenance.

Acknowledgements

The author gratefully acknowledges the Subcommittee on Practical Research of AI and Data Science, the Committee on Structural Engineering, Japan Society of Civil Engineers, for providing access for this bridge inspection and associated management database. Financial support from the WISE-SSS Program of the Academy for Super Smart Society, Institute of Science Tokyo, is also sincerely appreciated.

References

[1] MLIT, "Current Situation and Issues in the Construction Industry," 2023. [Online]. Available: https://www.mlit.go.jp/policy/shingikai/content/001610913.pdf
[2] MLIT, "Percentage of social infrastructure that is more than 50 years old after construction," Accessed: Jan. 01, 2025. [Online]. Available: https://www.mlit.go.jp/sogoseisaku/maintenance/_pdf/50year_percentage.pdf

DOI: 10.11159/icceia25.162

[3] A. Adadi and M. Berrada, “Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI),” IEEE Access, vol. 6, pp. 52138–52160, 2018.
[4] Takahiro MINAMI, Makoto FUJIU, Shoichiro NAKAYAMA, Jyunichi TAKAYAMA, and Yasuo CHIKATA, “ Analysis of Relationship between Soundness of Bridges and Natural Environments,” Journal of Japan Society of Civil Engineers, Ser. D3 (Infrastructure Planning and Management), vol. 72, no. 5, Art. no. 5, 2016.
[5] Yuriko OKAZAKI, Shinichiro OKAZAKI, Pang-jo CHUN, Shingo ASAMOTO, and Kazuaki OKUBO, “Crack Propergation Model of Concrete Bridges in Shikoku Based Data Driven Approach,” Journal of Japan Society of Civil Engineers, Ser. F4 (Construction and Management), vol. 74, no. 2, p. I_107-I_118, 2018.
[6] P. Miao, “Prediction-Based Maintenance of Existing Bridges Using Neural Network and Sensitivity Analysis,” Advances in Civil Engineering, vol. 2021, pp. 1–17, 2021.
[7] A. F. Santos, M. S. Bonatte, H. S. Sousa, T. N. Bittencourt, and J. C. Matos, “Improvement of the Inspection Interval of Highway Bridges through Predictive Models of Deterioration,” Buildings, vol. 12, no. 2, Art. no. 2, Jan. 2022.
[8] Kouji IGARASHI, and Kazuhisa ABE, “Examination of Crack Damage Prediction in Concrete Bridges using Machine Learning Algorithms,” Artificial Intelligence and Data Science, vol. 4, no. 4, pp. 1–15, 2023.
[9] Junpei Saito, “Factor Analysis of Geographical and Weather Conditions and Traffic Conditions Affecting the Deterioration Rate of Road Bridges in Japan,” Journal of Structural Engineering A, vol. 70, pp. 126–140, 2024.
[10] S. M. Lundberg and S.-I. Lee, “A Unified Approach to Interpreting Model Predictions,” Advances in Neural Information Processing Systems, vol. 30, 2017.