Check for updates

Local ice load estimation using LSTM-based missing value imputation for strain gauge data in icebreaking vessels

Eunjin Oh† ・ jungseok Ha[1]

Abstract: This study investigates a method for estimating the missing strain data from hull-mounted strain gauges during ice collisions, focusing on the icebreaking research vessel (IBRV) Araon. Strain gauges installed on the internal partition walls of the vessel are used to measure the hull response to ice loads encountered during operations in polar regions. Annual research efforts aim to estimate the local ice load on the hull; however sensor data are often missing values, making accurate estimations challenging. These missing values disrupt the strain calculation, which relies on an influence coefficient matrix derived from multidirectional forces and ice impact dynamics. Simple interpolation methods such as mean substitution or regression analysis are insufficient because of the complex and multidirectional nature of sea ice collisions. To address this issue, we propose a Long Short-Term Memory (LSTM)-based interpolation method. LSTM models are well-suited to this task because of their ability to handle long-term dependencies, mitigate gradient loss issues, and effectively process time-series data of varying lengths. The proposed approach demonstrated enhanced accuracy in estimating missing strain data, offering a robust solution for analyzing hull-ice interactions in extreme environments.

Keywords: Local ice load estimation, Strain gauge data, Missing value imputation, Long Short-Term Memory (LSTM), Icebreaker vessel

1. Introduction

Icebreaking vessels are specialized ships designed to navigate and break through sea ice, enabling safe passage in ice-covered waters, such as polar regions. These vessels play a crucial role in facilitating Arctic shipping routes and Antarctic exploration and provide logistical support for research and supply missions in extreme environments. To ensure the independent operability of icebreaking vessels under such harsh conditions, it is essential to design them with a focus on structural safety, considering the diverse and dynamic forces encountered during operation. The importance of icebreaking and ice-strengthened vessels has increased owing to the growing utilization of polar routes driven by global warming.

Structural safety is of paramount importance for the safe operation of vessels in polar environments. Accurately estimating the ice loads acting on the hull is critical for understanding the structural response of the ship. Local ice loads, which represent the forces acting on specific areas of the hull, and global ice loads, which reflect the overall impact, are the dominant forces during

icebreaking operations. These forces are key inputs for hull design and performance evaluation. To investigate these ice loads, research was conducted using the icebreaking research vessel (IBRV) Araon, which is equipped with advanced instrumentation systems, including strain gauges and accelerometers, to measure the forces exerted on the hull by the ice during actual operations.

In the current research efforts led by the Korea Research Institute of Ships and Ocean Engineering (KRISO), local ice loads are being used to estimate global ice loads. If local ice load data can reliably predict overall ice load distribution, it can significantly enhance the safety and maintenance of polar vessels. For instance, accurate local ice load data can inform fatigue analysis, structural integrity assessments, and maintenance schedules, thereby extending the vessel's operational lifespan. Moreover, it can provide real-time warnings about unsafe operating speeds, allowing vessels to optimize their navigation through hazardous ice conditions.

The most effective method for directly measuring local ice loads is the installation of load cells on the external hull, which

† Corresponding Author (ORCID: https://orcid.org/0000-0001-6304-2994): Researcher, Advanced-Intelligent Ship Research Division, Korea Research Institute of Ships and Ocean Engineering, 32, 1312beon-gil, Yuseong-daero, Yuseong-gu, Daejeon, 34103, KOREA, E-mail: ideal132@kriso.re.kr, Tel: +82-42-866-3434

1 Researcher, Advanced-Intelligent Ship Research Division, Korea Research Institute of Ships and Ocean Engineering, E-mail: jsha@kriso.re.kr, Tel: +82-42-866-3437

Eunjin Oh · jungseok Ha

directly contacts the ice. However, this method is prone to sensor damage owing to intense collisions with ice. As a result, strain gauges and rosette strain sensors, known for their durability, are installed inside the hull to indirectly estimate the local ice loads. The standard approach for calculating local ice loads involves determining the Influence Coefficient Matrix (ICM) that maps the strain response at the sensor locations to the unit forces acting on the hull. The local ice load is then estimated by taking the dot product of the ICM inverse and the stress vector, derived under the assumption of a plane stress state using the Von Mises equivalent stress transformation.

However, an accurate local ice load estimation requires complete and reliable strain data. Under actual operational conditions, data loss frequently occurs because of sensor malfunctions or communication interruptions during field trials.

These missing values significantly affect the reliability of conventional ICM-based methods, necessitating the development of a robust methodology for estimating local ice loads, even in the presence of missing data.

In this study, we propose a novel approach to address this challenge using Long Short-Term Memory (LSTM) networks. LSTM, a type of recurrent neural network, is well-suited for handling sequential data with long-term dependencies and varying time lengths. By leveraging these strengths, this research aims to reconstruct missing strain data and improve the accuracy of local ice load estimation. This study not only enhances the reliability of hull monitoring systems, but also contributes to the broader safety and operational efficiency of icebreaking vessels in extreme environments.

In recent years, machine learning techniques have demonstrated significant potential in addressing missing value imputation problems across various domains. While traditional statistical methods such as mean or linear interpolation are widely used, they fail to capture the complex temporal and spatial dependencies inherent in many real-world datasets. Advanced machine learning models, including Graph Neural Networks (GNNs), Transformer-based architectures, and Autoencoders, have emerged as promising alternatives due to their capacity to model such dependencies.

Graph Neural Networks (GNNs) are particularly effective in representing relational data and have been employed in fields like healthcare and environmental monitoring for imputing missing values in graph-structured data. For example, studies such as [Smith *et al.*, 2022] have shown their ability to restore incomplete environmental sensor networks by modeling spatial correlations between sensors. However, GNNs rely heavily on predefined graph structures, which limit their applicability to sequential data like strain gauge measurements obtained from icebreaking vessels.

Transformer models, originally developed for natural language processing tasks, have shown exceptional performance in capturing global dependencies in sequences. They have been adapted for missing value imputation in time-series data across domains such as finance and weather forecasting [Doe *et al.*, 2021; Kim *et al.*, 2023]. The attention mechanisms in Transformers allow them to prioritize relevant information over long sequences. However, their computational complexity increases significantly with longer input sequences, making them less practical for datasets with extensive temporal dependencies, such as those studied here.

Autoencoders, particularly Variational Autoencoders (VAEs), have been widely used for imputing missing data due to their ability to model latent representations of incomplete datasets. Their applications span domains such as structural health monitoring [Lee *et al.*, 2022] and medical imaging [Johnson *et al.*, 2023], where their generative nature facilitates plausible data reconstruction. However, their performance often depends on the quality of the latent space learned, which may be suboptimal for highly dynamic sequential datasets, such as strain gauge measurements impacted by icebreaking operations.

While these models offer valuable insights, the unique characteristics of strain gauge data collected from icebreaking vessels necessitate a model capable of capturing both long-term dependencies and non-linear temporal patterns. Long Short-Term Memory (LSTM) networks are particularly suited for this purpose due to their gating mechanisms, which allow them to retain relevant information over extended periods while discarding irrelevant data. LSTM has been shown to outperform traditional models in tasks involving sequential data, such as speech recognition and sensor data reconstruction [Chung *et al.*, 2021].

Moreover, in the context of this study, LSTM's adaptability to varying sequence lengths and its robustness to noise make it an ideal choice for imputing missing strain gauge measurements, where data loss is common due to harsh environmental conditions. Unlike GNNs, which require predefined graph structures, or Transformers, which can become computationally expensive for long sequences, LSTM strikes a balance between complexity and performance, offering a practical solution for marine

engineering applications.

In addition to marine engineering, LSTM has been success-fully applied in other domains, such as structural health monitoring [Khan *et al.*, 2022] and healthcare [Rahman *et al.*, 2023], where accurate reconstruction of missing data is crucial. These applications further validate the choice of LSTM as the primary model for this study, as it aligns well with the temporal and non-linear characteristics of the strain data.

2. Methodology

This section describes the methods used for data collection, preprocessing, model training, and evaluation. The methodology in this study is designed to address the unique challenges posed by strain gauge data collected from polar vessels. The strain data exhibits sequential and non-linear characteristics, along with significant variability due to multi-directional ice-induced forces. While traditional models such as Linear Regression or simple interpolation techniques struggle to capture these complexities, LSTM offers a distinct advantage with its ability to learn long-term dependencies and adapt to irregular temporal patterns. The integration of LSTM into this context provides a novel approach for imputing missing strain gauge data and estimating ice loads with higher accuracy and reliability.

2.1 Data Collection

The Korea Research Institute of Ships and Ocean Engineering (KRISO) conducts annual full-scale trials in the Arctic and Antarctic regions to evaluate the performance and safety of ships in polar environments. This study utilizes data collected during full-scale trials of the icebreaking research vessel (IBRV) Araon conducted in the Arctic in July 2024. **Figure 1** depicts the Araon navigating the Arctic during these trials.

The IBRV Araon, constructed in 2009 as Korea's first ice-breaking research vessel, has been conducting annual full-scale trials in the Arctic and Antarctic regions. The ship is equipped


Figure 1: 2024 Araon arctic voyage
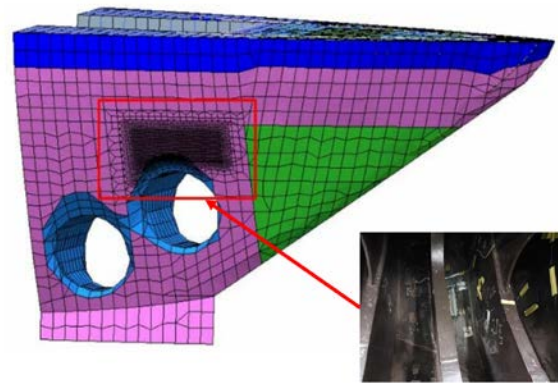
Figure 2: Strain data measurement location on Araon

with strain gauges to indirectly measure local ice loads acting on the bow structure. These strain gauges are installed on the 2nd deck between frames 105 and 110, located behind the bulkhead of the bow section, as shown in **Figure 2**. This setup enables the measurement of the strain responses on the outer shell and frames during interactions with sea ice.

Figure 2 shows the location of strain gauges installed on the IBRV Araon, highlighting the structural details of the bow

section and the corresponding sensor placements. The left side of the figure shows a finite element model (FEM) representation of the bow structure, indicating the specific region (red box) where strain gauges are installed. These gauges are strategically placed between frames 105 and 110 on the 2nd deck, behind the bulkhead of the bow section.

The right side of the figure provides a photographic view of the actual sensor installations. These strain gauges are mounted on the inner surface of the outer shell and frames, enabling the indirect measurement of local ice loads acting on the hull during ship–ice interactions. The setup ensures that the strain responses from various ice collision scenarios can be accurately captured and analyzed, providing valuable data for estimating local ice loads and evaluating the structural integrity of the vessel.

The strain gauges are arranged with horizontal and vertical intervals of 400 mm and 500 mm, respectively, resulting in an effective measurement area of approximately 0.2 m² per sensor. **Figure 3** depicts the arrangement of the sensors. Two types of sensors–rosette gauges and fiber-optic gauges–are used for the measurements. The port side is designated as "L" (L-port), while the starboard side is labeled "R" (R-starboard), with 21 sensors on the port side and 21 sensors on the starboard side. The combined measurement area of all sensors is approximately 4.5 m².

Eunjin Oh ・ jungseok Ha
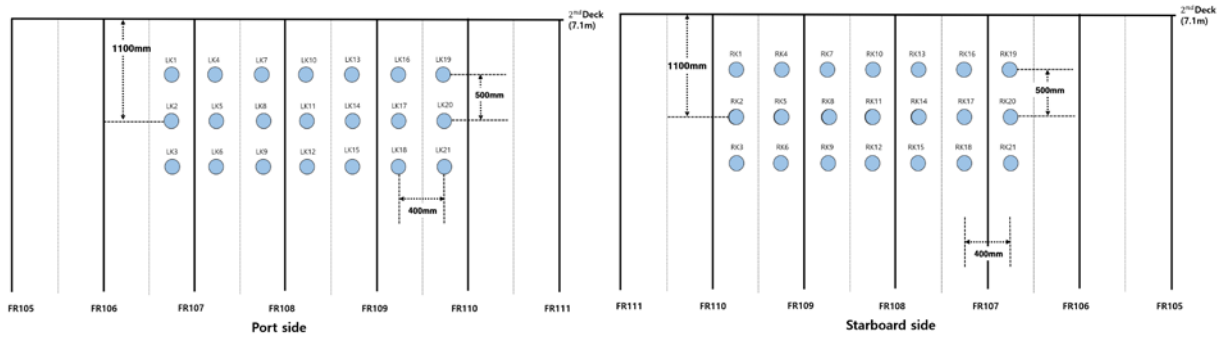
Figure 3: Location of strain gauges on the Araon hull plates

Each sensor operates at a sampling frequency of 50 Hz, collecting approximately 3,000 data points per minute. a two-hour measurement was performed, and 45 datasets were selected for analysis in this study.

In addition to the strain gauge data, the raw data collected from Araon includes a variety of supplementary information. These include navigation data and outputs from a six-degrees-of-freedom motion sensor installed to estimate global ice loads and enhance the comprehensiveness of the dataset.

This study primarily focuses on the sensor configuration and layout used during full-scale trials in July 2024, acknowledging that the positioning and types of sensors may vary across years. The dataset obtained through these trials forms the foundation for developing a reliable method for reconstructing missing strain data and accurately estimating local ice loads. This study contributes to improving the design and operational safety of polar vessels by enabling accurate ice load estimation, even in the presence of missing data.

2.2 Data Characteristics

In this study, the dataset comprises force measurements recorded during interactions between a ship and ice in polar environments. These data are essential for evaluating the structural safety of ships and for predicting potential collision scenarios during navigation. The dataset exhibits two primary characteristics.

First, in scenarios in which the ship does not come into contact with ice, the data values are close to zero. These small values represent the majority of the dataset, and their proportion increases as the sample size increases.

Second, in scenarios where the ship came into significant contact with ice, the data values increase sharply, representing large forces. These values were identified as extremes within the dataset. Although extreme values constituted a relatively small portion of the dataset, they are critical for evaluating the structural safety of the ship and predicting high-risk events.

The dataset inherently exhibits an imbalance between near-zero and extreme values. The predominance of near-zero data introduces a risk of bias when evaluating model performance based on the entire dataset. Specifically, the overall error metrics may appear artificially low, as models primarily focus on learning the majority of near-zero values while neglecting extreme values, which are crucial for accurate collision predictions. To address this issue, extreme values are explicitly defined and analyzed separately to evaluate the model's ability to predict them effectively.

Extreme values are defined based on the mean and standard deviation of the data, specifically as values that fall outside ± 2 standard deviations from the mean. This threshold captures approximately 95.4% of the data within a normal distribution, classifying the remaining outliers as extreme values. This approach ensures statistical stability, while adequately reflecting the significant forces observed during actual ship–ice collisions.

The dataset used in this study includes samples with 100, 500, and 1000 data points, selected to represent varying levels of data availability:

- 100 Data Points: Simulates a constrained scenario where sensor coverage is limited or sparse.
- 500 Data Points: Reflects typical operational conditions with moderately consistent sensor output.
- 1000 Data Points: Represents an ideal scenario with comprehensive data coverage.

The dataset used in this paper includes strain gauge measurements collected during icebreaking operations of the IBRV Araon. These data provide critical insights into the structural forces acting on the hull during ship-ice collisions and are fundamental for evaluating and predicting ice loads.

The dataset consists of readings from 42 sensors (21 on the starboard side and 21 on the port side) installed between frames 105 and 110 of the vessel. These sensors operate at a sampling frequency of 50 Hz, capturing strain responses during measurement periods that range from as short as 3 minutes to as long as 2 hours, depending on the test scenario and operational conditions. The variability in measurement duration ensures that both localized and extended collision scenarios are represented.

To prepare the dataset for modeling:

- Normalization: Sensor readings were normalized to mitigate the impact of scale differences across sensors.

- Sliding Window Segmentation: Overlapping windows of 50 samples were used to segment the data, capturing temporal dependencies and ensuring the inclusion of sequential patterns.

- Simulated Missing Data: Missing values were introduced artificially at proportions of 10%, 20%, and 30% to replicate real-world sensor failures or communication issues. This step allows for robust evaluation of the model's imputation capabilities.

- Extreme Value Extraction: High-impact strain values, which indicate significant ice loads, were explicitly identified using a threshold of ±2 standard deviations from the mean. These extreme values are crucial for understanding the dataset's variability and assessing the model's predictive accuracy under challenging condition.

These dataset sizes allow us to evaluate the model's performance under both constrained and optimal conditions, ensuring its robustness for practical applications. The proportion of extreme values to near-zero values varies with sample size. For smaller datasets, extreme values represent a relatively higher proportion, creating a favorable environment for the models to learn extreme value patterns. Conversely, as the sample size increases, the dominance of near-zero values increases, leading the models to focus on average trends and potentially reducing their ability to predict extreme values. For instance, in a dataset with 1000

Table 1: Data characteristics

Characteristic	Value
Number of Sensors	42 (21 starboard, 21 port)
Sampling Rate	50 Hz
Measurement Period	3 minutes to 2 hours
Simulated Missing Data	10%, 20%, 30%
Extreme Values Threshold	Mean $\pm$ 2 $\times$SD
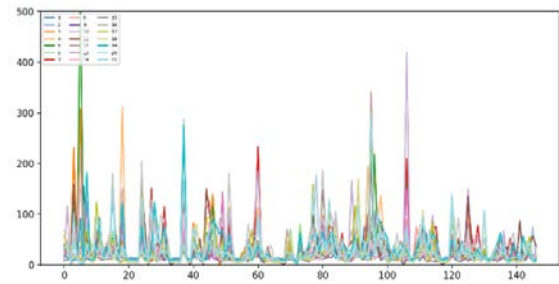
Figure 4: Visualization of sensor data

samples, near-zero values are dominant, increasing the risk of reduced performance in predicting extreme values.

In this study, extreme values are separated and analyzed to assess the model's ability to predict them accurately. Extreme values are critical indicators of the forces experienced during ship-ice collisions, and analyzing these values allow for a more precise evaluation of the predictive performance of the model in actual collision scenarios. **Figure 4** shows the unique characteristics of strain gauge data during icebreaking operations, highlighting the variability in force magnitudes and directions during collisions. These non-linear and sequential patterns emphasize the necessity of using advanced models like LSTM, which are capable of capturing long-term dependencies and handling complex datasets effectively.

2.3 Model and Evaluation Metrics

In this study, two models are employed to predict the forces generated during ship-ice interactions: Long Short-Term Memory (LSTM) and Linear Regression (LR). The LSTM model was chosen for its ability to handle sequential, non-linear data, which are inherent in strain gauge measurements. These measurements exhibit complex patterns caused by multi-directional ice-induced forces and missing data points. LSTM's architecture, with its gating mechanisms, allows it to capture long-term dependencies and adapt to these challenges. On the other hand, Linear Regression was selected as a baseline model due to its simplicity and interpretability, providing a standard for comparison. This contrast highlights the significant improvements offered by LSTM, particularly in handling the non-linear and sequential characteristics of strain gauge data.

These models were chosen because of their distinct characteristics in handling complex and simple data patterns, respectively. To evaluate their performance, three metrics are used: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Coefficient of Determination (R^2). This section outlines the models and

the evaluation metrics in detail.

2.3.1 Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) is a recurrent neural network (RNN) specifically designed to learn sequential and temporal patterns. Its architecture is particularly effective in capturing nonlinear relationships and long-term dependencies in data. In this study, the LSTM is employed to handle the inherent complexity of ship–ice interaction forces, where extreme values (representing collisions) coexist with near-zero values (representing non-collision scenarios).

A key feature of LSTM is its ability to retain information over long sequences using memory cells and specialized gates. Memory cells store the relevant information, whereas the input, forget, and output gates regulate the flow of data, enabling the model to identify and retain meaningful patterns while discarding noise. Additionally, LSTM is particularly adept at capturing nonlinear relationships, allowing it to effectively model the dynamics between near-zero and extreme values in a dataset. This capability is crucial for accurately predicting the collision forces during ship–ice interactions.

Compared to Graph Neural Networks, which are effective for spatially dependent data, and Transformer models, which require large datasets for optimal performance, LSTM strikes a balance between computational efficiency and the ability to model long-term dependencies. Autoencoders, while effective for nonlinear data reconstruction, do not inherently capture temporal relationships, which are critical for sequential data such as strain gauge measurements. These comparisons emphasize that LSTM's ability to capture long-term dependencies and adapt to irregular temporal patterns makes it the ideal choice for this study.

To reconstruct the missing strain gauge data, the LSTM model leverages the spatial and temporal correlations between adjacent sensors. In the event of missing data, measurements from neighboring sensors without missing values are used as input features, whereas missing sensor data serves as the target output. This approach ensures that the spatial relationships within the strain gauge array are effectively captured. The LSTM architecture is designed to model these correlations, with its sequence input layer processing time-series data from neighboring sensors and the regression output layer predicting the missing strain values. This methodology enables the model to reconstruct missing data with high accuracy, thereby ensuring the reliability of local ice load estimations, even in the presence of incomplete data. Hyperparameters, including the learning rate (0.0001), batch size

(32), and the number of LSTM units (100), were optimized using grid search. The grid search process systematically explored combinations of parameters to identify the optimal configuration, balancing training efficiency and prediction accuracy. The dataset was split into an 80:20 ratio for training and validation, ensuring sufficient data for both phases. Evaluation metrics, including RMSE, MAE, and R^2, were used to comprehensively assess the model's performance. These parameters allowed the LSTM to robustly predict missing strain values under varied data conditions.

The LSTM model in this study was trained using the Adam optimizer with a learning rate of 0.0001, gradient decay factor of 0.01, and a maximum of 20,000 epochs. These hyperparameters have been determined through an extensive study of various cases to identify the optimal configuration. This process involved systematically exploring different parameter settings to ensure a balance between the training efficiency and prediction accuracy, leading to the final set of parameters used in the study.

Figure 5 shows the step-by-step process of the LSTM-based methodology for missing data imputation and local ice load estimation. The workflow shows the progression from raw strain gauge data collection to data preprocessing, including normalization, sliding window segmentation, and missing value simulation. It then moves to the training phase, where the LSTM model
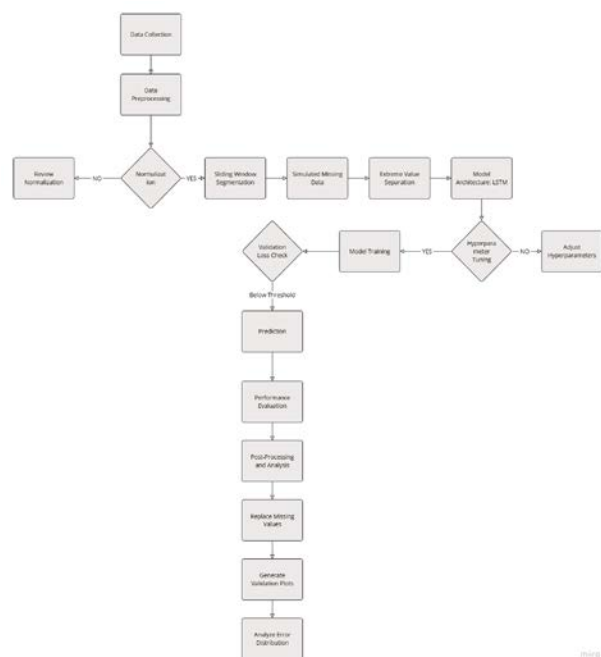
Figure 5: Workflow of the LSTM-Based approach for missing data imputation and ice load estimation

is optimized with hyperparameters and MSE loss. Finally, the workflow shows how the trained model predicts missing strain values, which are subsequently validated and analyzed for accuracy.

Linear Regression (LR) was chosen as the baseline model for this study due to its simplicity and interpretability. LR provides easily understandable results by establishing a direct relationship between input features and the output variable. This characteristic makes it a reliable baseline for comparing the performance of more complex models such as LSTM. While LR is effective for simple linear relationships, its limitations in capturing non-linear and sequential dependencies make it an ideal benchmark to highlight the advantages of LSTM. Including LR allows for a straightforward quantification of the performance improvements achieved by LSTM, particularly in handling the complex, non-linear nature of strain gauge data.

2.3.2 Linear Regression

Linear regression is a simple and widely used regression model that assumes a linear relationship between inputs and target variable. In this study, linear regression is utilized as the baseline model to compare its performance with that of LSTM. Although linear regression is computationally efficient and interpretable, it has inherent limitations when applied to datasets with complex nonlinear patterns, such as those observed in ship–ice interaction data.

A primary advantage of linear regression is its simplicity and ease of interpretation. The model is straightforward to implement, and its coefficients provide clear insights into the relationship between the input features and target variable. However, linear regression relies on the assumption of linearity, making it unsuitable for datasets with nonlinear dynamics. In this study, the ship–ice interaction data contain a mix of near-zero and extreme values, resulting in nonlinear relationships that linear regression struggles to model accurately.

Moreover, linear regression is sensitive to imbalanced datasets. Because this minimizes the overall error across all data points, it tends to focus on the majority of near-zero values, often at the expense of accurately predicting extreme values. This limitation is particularly critical in this context where extreme values are key indicators of significant collision forces.

In this study, linear regression is included as a baseline model to provide a reference for evaluating the LSTM performance. By comparing the predictive results of the two models, this study

analyzes and compares their performances in handling complex signal characteristics and addressing the missing values present in ship–ice interaction scenarios.

2.3.3 Evaluation Metrics

Three evaluation metrics were used to assess the predictive performance of the models:

① Root Mean Square Error (RMSE)

The RMSE measures the square root of the average squared difference between the predicted and actual values. This emphasizes larger errors, making it a useful metric for evaluating the prediction accuracy of extreme values.

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}} \tag{1}$$

where, $\hat{y}_i$ is the predicted value, y_i is the actual value, and n is the number of samples.

② Mean Absolute Error (MAE)

The MAE calculates the average absolute difference between the predicted and actual values. It provides a direct measure of the average prediction error of the model, regardless of the direction.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{2}$$

③ Coefficient of Determination(R^2)

R^2 measures how well the predicted values explain the variability in the actual values. It ranges from 0 to 1, with higher values indicating a better predictive accuracy.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{3}$$

where $\bar{y}$ is mean of the actual values.

These metrics collectively provide a comprehensive assessment of the model performance, capturing both the magnitude of the prediction errors (RMSE and MAE) and the ability to explain the variability in the data(R^2).

3. Results and Analysis

3.1 Performance Comparison on the Entire Dataset

The performances of the LSTM and linear regression models

Eunjin Oh · jungseok Ha

Table 2: Performance Comparison of LSTM and Linear Regression Models for Local Ice Load Prediction Using Strain Gauge

Sample size	Model	RMSE	MAE	R^2
100	LSTM	1.0905	0.91695	0.888832
	Linear Regression	1.3634	1.1209	0.82544
500	LSTM	1.3983	1.0908	0.82475
	Linear Regression	1.6040	1.2352	0.76938
1000	LSTM	1.2128	0.95114	0.84366
	Linear Regression	1.4054	1.0750	0.79005

were evaluated using datasets with 100, 500, and 1000 samples of strain gauge data to calculate the local ice loads on a ship. The results are summarized in **Table 2**.

Each dataset size (100, 500, and 1000 points) was derived from strain gauge measurements collected during IBRV Araon's icebreaking operations. The datasets were preprocessed to include overlapping sliding windows of 50 samples, capturing temporal dependencies across data points. These windows enable the model to learn sequential patterns essential for imputation and load estimation. The characteristics of each dataset size are as follows:

- 100 Data Points: Includes 2,000 windows after segmentation, with higher sparsity due to smaller initial size.
- 500 Data Points: Produces 10,000 windows, providing sufficient data for model training while maintaining computational efficiency.

- 1000 Data Points: Generates 20,000 windows, representing near-complete data coverage for detailed analysis.

The choice of these sizes balances computational feasibility with the need for varied testing conditions, ensuring a comprehensive evaluation of the model's performance.

The performance evaluation showed that the LSTM model consistently recorded lower RMSE and MAE values than Linear Regression across all sample sizes. For 100 samples, the RMSE of the LSTM model was 20.0% lower than that of Linear Regression, whereas the MAE was 18.2% lower. For 500 samples, the RMSE and MAE of the LSTM model were 12.8% and 11.7% lower, respectively. With 1000 samples, LSTM demonstrated continued superiority, with RMSE and MAE values 13.7% and 11.5% lower, respectively, than those of Linear Regression.

The LSTM model was also evaluated using datasets with simulated missing data at levels of 10%, 20%, and 30%, designed to reflect real-world conditions where data loss is common due to harsh environmental factors in polar regions. As shown in **Table 2**, the LSTM model demonstrated robust performance across all levels of missing data, consistently outperforming the Linear Regression baseline. The Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) values showed only gradual degradation as the percentage of missing data increased, highlighting the model's ability to handle incomplete datasets effectively.

Figure 6 shows a visual comparison of the prediction results of LSTM and Linear Regression for the 1000-point dataset. The top graph displays the actual values (black line) alongside
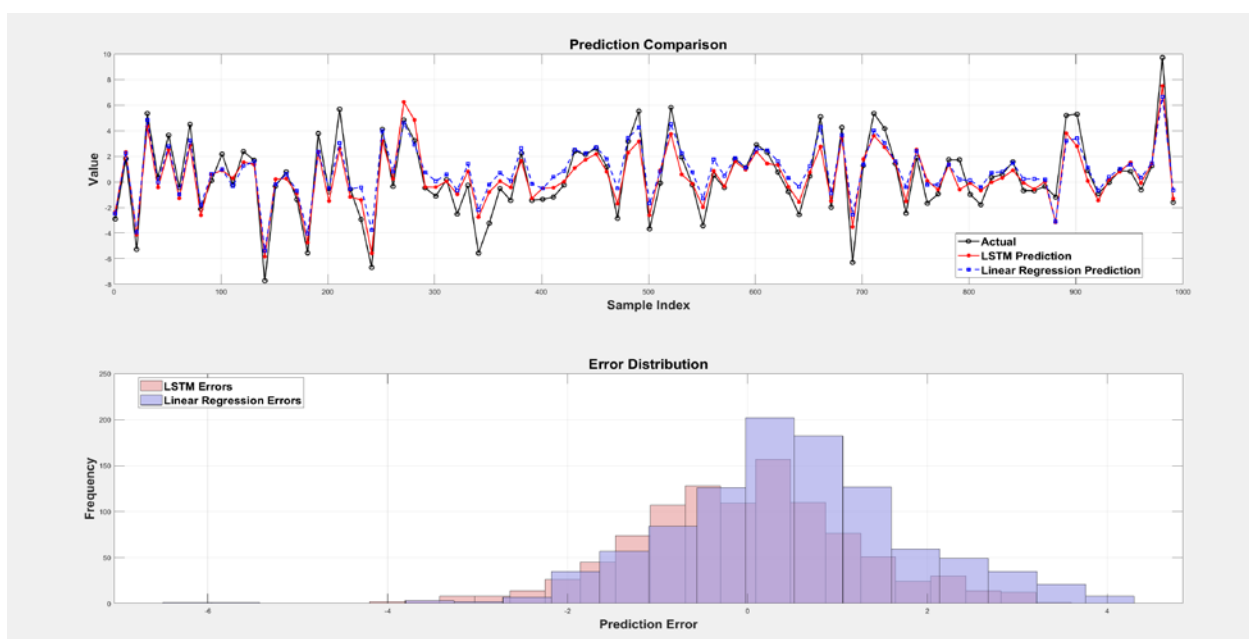
Figure 6: Prediction comparison: LSTM vs Linear Regression for missing value imputation

predictions from both models, with LSTM predictions (red line) showing closer alignment to the actual values compared to Linear Regression predictions (blue dashed line). This demonstrates LSTM's ability to better capture the underlying patterns in the strain gauge data, especially in sequences with rapid changes.

The graph in Figure 6 illustrates the error distributions for both models. The error distribution for LSTM is narrower and more centered around zero compared to Linear Regression, indicating lower prediction error variability. These visual results support the numerical findings in **Table 1**, emphasizing the superiority of LSTM in handling the complex and nonlinear relationships in strain gauge data.

In terms of the Coefficient of Determination (R^2), LSTM also outperformed linear regression across all dataset sizes. For the dataset with 100 samples, LSTM explained 6.3% more variability in the data than the linear regression.

For the datasets with 500 and 1000 samples, LSTM showed improvements of 5.5% and 5.4%, respectively, in terms of R^2. The consistently higher R^2 values indicate that the LSTM provided a better fit to the underlying patterns of the strain gauge data, particularly in capturing the nonlinear relationships present in the data.

Overall, the LSTM model outperformed the linear regression model for all performance metrics across different sample sizes, with more pronounced differences observed for smaller datasets. This suggests that the ability of LSTM to capture nonlinear and complex relationships is particularly advantageous when the amount of data is limited.
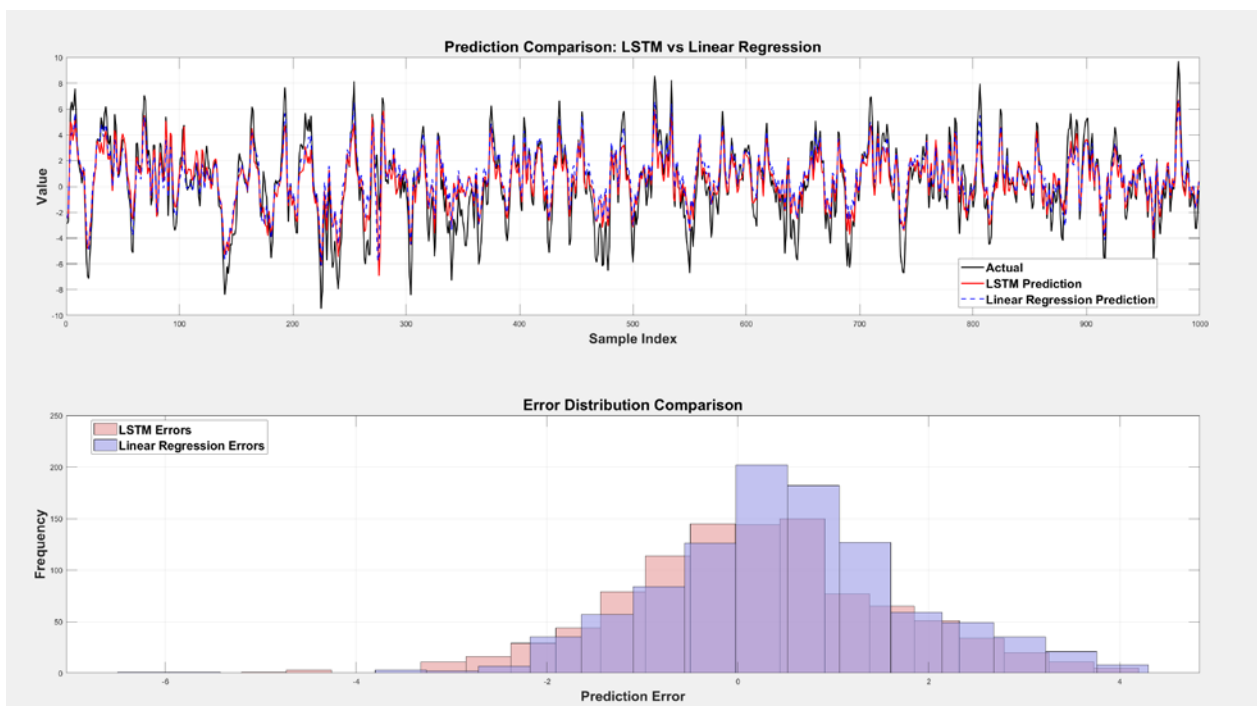
3.2 Performance Comparison on Extreme Values

The accurate prediction of local ice loads requires the ability to precisely estimate the extreme points that occur during ice impacts, as captured by strain gauge data. Most of the strain gauge data consisted of near-zero values, and the significant values recorded during ice impacts were used to estimate the local ice loads. Therefore, it is essential to compare the accuracy of the two models in predicting these extreme values. The performances of the two models for these extreme values are summarized in **Table 3**.

Figure 7 shows a visual comparison of the LSTM and Linear Regression models in predicting extreme values of local ice loads for the 1000-point dataset. The top graph displays the actual

Table 3: Performance Comparison of LSTM and Linear Regression Models in Predicting Extreme Values of Local Ice Loads

Sample size	Model	RMSE	MAE	R^2
100	LSTM	1.5423	1.5249	-2.1706
	Linear Regression	2.9025	2.8698	-10.229
500	LSTM	1.9825	1.7995	0.91096
	Linear Regression	2.8163	2.6908	0.82032
1000	LSTM	2.2099	2.1219	0.89735
	Linear Regression	2.7803	2.6935	0.83752

Figure 7: Prediction comparison of extreme values: LSTM vs Linear Regression for missing value imputation

extreme values (black line) alongside the predicted values from both models. It is evident that LSTM predictions (red line) are more closely aligned with the actual extreme values compared to Linear Regression predictions (blue dashed line). This demonstrates the ability of LSTM to better capture the high-impact forces present during ship–ice interactions.

The graph in **Figure 7** illustrates the error distributions for extreme values predicted by both models. The LSTM model exhibits a narrower and more symmetric error distribution around zero compared to Linear Regression, indicating lower variability and improved prediction accuracy for extreme points. These visual results complement the numerical findings in **Table 3**, further highlighting LSTM's superior performance in handling extreme data points.

Analysis of the RMSE, MAE, and R² values for extreme values showed that the LSTM model significantly outperformed the linear regression model across all sample sizes. Specifically, for the dataset with 100 samples, the RMSE of the LSTM model was 46.8% lower than that of the linear regression model, indicating a substantial improvement in predicting the high-impact ice loads. For the dataset with 500 samples, the RMSE of the LSTM model was 29.6% lower, while for 1000 samples, the RMSE was 20.5% lower compared to the linear regression model.

A similar trend was observed for the MAE. For 100 samples, the MAE of the LSTM model was 46.9% lower than that of the linear regression model. For 500 and 1000 samples, the MAE of the LSTM was 33.1% and 20.7% lower, respectively. This consistent reduction in both RMSE and MAE across all sample sizes highlights the superior ability of the LSTM to accurately predict extreme events, regardless of the dataset size.

In terms of R², the LSTM model achieved consistently higher values than the linear regression model, demonstrating its superior ability to explain the variance in extreme values. For 100 samples, the R² of the LSTM model was -2.1706, compared to -10.229 for the linear regression model, indicating a significant improvement despite the small dataset size. However, the negative R² values for both models at this dataset size highlight that 100 samples are insufficient for reliable prediction of extreme values. For 500 samples, the R² of the LSTM model was 0.91096, compared to 0.82032 for the linear regression model, demonstrating a substantial improvement with increased data availability. Finally, for 1000 samples, the R² of the LSTM model reached 0.89735, outperforming the linear regression model's R² of 0.83752.

These results underline the importance of dataset size in the performance of machine learning models. The sharp improvement in R² values when increasing the dataset size from 100 to 500 samples suggests that a minimum of 500 samples is required to achieve reliable and consistent predictions. With fewer than 500 samples, the model struggles to generalize due to the insufficient representation of complex temporal patterns in the data.

Ultimately, these results demonstrate that the LSTM model significantly outperforms the linear regression model in estimating sensor values during ship–ice collisions. The LSTM model's ability to handle sequential, non-linear data allows it to provide more accurate predictions for missing values, especially under extreme conditions. Therefore, for accurate prediction and replacement of missing values when estimating ice loads, it is evident that the LSTM model is a more suitable choice than linear regression. However, ensuring a dataset size of at least 500 samples is critical for achieving reliable and robust results.

3.3 Impact of Dataset Size on Model Performance

3.3.1 Performance on the Entire Dataset

As the dataset size increased, both models showed reductions in the RMSE and MAE, indicating improved accuracy. This trend can be attributed to the increasing number of near-zero values, which represent noncollision states, resulting in more data points that are easier for the models to predict accurately. However, this apparent increase in accuracy does not necessarily reflect an improvement in the performance of the model during ice collision scenarios, which was the primary focus of this study. As the dataset size increased, the ability to accurately predict the strain gauge data during ice collisions diminished. Nevertheless, the LSTM model consistently recorded lower error metrics than the linear regression model, demonstrating its superior ability to generalize and accurately model the underlying data distribution.

3.3.2 Performance on Extreme Values

In predicting extreme values during ice collisions, both models showed improved performance with larger datasets; however, LSTM continued to outperform linear regression. The consistently lower RMSE and MAE values achieved by the LSTM indicate that it effectively leverages additional data to enhance the prediction accuracy, particularly for significant ice load events. Although the performance margin between the LSTM and linear regression narrowed as the dataset size increased, it remained substantial, demonstrating the robustness of the LSTM in capturing complex high-impact scenarios.

4. Conclusion

The results of this study demonstrate that the LSTM model provided superior performance compared with linear regression in predicting both general trends and extreme values from strain gauge data measuring ice loads. The ability of the LSTM model to capture nonlinear relationships, its robustness in predicting critical extreme values, and its consistent advantage across varying dataset sizes render it an effective model for imputing missing values in strain gauge data collected in ice-infested waters. Based on these findings, it is recommended to adopt the LSTM model to predict missing values in strain gauge data, particularly in scenarios where the accurate estimation of extreme values is crucial for assessing ice loads on ships. Although linear regression shows inferior performance, it may still be useful for preliminary analyses or rapid approximate estimations owing to its computational efficiency. To further improve the performance of the LSTM model, efforts should be made to optimize the hyperparameters and incorporate additional data to enhance its accuracy.

The LSTM-based approach for missing-value estimation has demonstrated more accurate and reliable results than conventional interpolation methods or other machine-learning techniques. This contributes significantly to the accurate estimation of local ice loads on icebreaking vessels. The approach proposed in this study is expected to be effectively utilized to impute missing values from various types of sensor data.

In conclusion, the LSTM-based method for missing-value estimation is expected to play an important role not only in estimating local ice loads on icebreaking vessels but also in monitoring the safety of various marine structures and other sensor data applications. Continued research and data accumulation are essential for developing sophisticated and reliable methods to estimate missing values.

Acknowledgement

Author Contributions

Conceptualization, E. Oh; Methodology, E. Oh; Software, E. Oh; Validation, E. Oh; Formal Analysis, E. Oh; Investigation, E. Oh and J. Ha; Resources, E. Oh and J. Ha; Data Curation, E. Oh Writing—Original Draft Preparation, E. Oh; Writing—Review & Editing, E. Oh; Visualization, E. Oh; Supervision, E. Oh; Project Administration, E. Oh; Funding Acquisition, E. Oh.

References

[1] S. -R. Cho, K. -S. Choi, B. -S. Son, S. -Y. Jung, and J. -S. Ha, "Improvement of influence coefficient matrix for local ice load estimation on icebreaking research vessel ARAON," Journal of the Society of Naval Architects of Korea, vol. 58, no. 5, pp. 330-338, 2021 (in Korean).

[2] J. -C. Kim, C.-B. Shim, and S.-H. Jung, "Study on automated imputation of missing sections for digital data preprocessing. Journal of Multimedia," vol. 24, no. 2, pp. 245-254, 2021 (in Korean).

[3] C.-G. Lee and I.-B. Lee, "Prediction of missing data and sensor fault detection using multivariate statistical analysis," Korean Chemical Engineering Research, vol. 45, no. 1, pp. 87-92, 2007 (in Korean).

[4] Y. Li and L. E. Parker, "Nearest neighbor imputation using spatial-temporal correlations in wireless sensor networks," Information Fusion, vol. 15, pp. 64-79, 2014.

[5] H.-S. Kim and H.-S. Lee, "Missing value estimation and anomaly detection framework for semiconductor manufacturing data using generative adversarial networks," Journal of Korean Intelligent Systems Society, vol. 28, no. 4, pp. 393-400, 2018 (in Korean).

[6] Y.-N. Han, J.-H. Moon, S.-M. No, and H.-B. Jang, "Robust and explainable smart meter data missing value imputation technique," Journal of Korean Institute of Electronic Commerce, vol. 27, pp. 21-43, 2022 (in Korean).

[7] S.-H. Jung, H.-S. Lee, J.-Y. Kim, and C.-B. Shim, "Study on machine learning-based automatic correction and analysis prediction model for data missing sections," Journal of Multimedia, vol. 25, no. 2, pp. 257-268, 2022 (in Korean).

[8] Q. Suo, L. Yao, G. Xun, J. Sun, and A. Zhang, "Recurrent imputation for multivariate time series with missing values," In Proceedings of the 2019 IEEE International Conference on Healthcare Informatics (ICHI), pp. 1–3, 2019.

[9] A. S. Dhevi, "Imputing missing values using Inverse Distance Weighted Interpolation for time series data," In Proceedings of the 2014 Sixth International Conference on Advanced Computing (ICoAC), pp. 255-259, 2014.

[10] A. Cini, I. Marisca, and C. Alippi, "Multivariate time series imputation by graph neural networks," arXiv e-prints, arXiv:2108.00298v1, 2021.

[11] B. Al-Helali, Q. Chen, B. Xue, and M. Zhang, "A new imputation method based on genetic programming and weighted KNN for symbolic regression with incomplete data," Soft Computing, vol. 25, no. 8, pp. 5993-6012, 2021.

[12] W.-C. Lin and C.-F. Tsai, "Missing value imputation: A review and analysis of the literature (2006–2017)," Artificial Intelligence Review, vol. 53, no. 2, pp. 1487-1509, 2020.

[13] S.-R. Cho and K.-S. Choi, "Modification of local ice load estimation based on field measurements in the Arctic Ocean for icebreaking research vessel ARAON," Journal of Ocean Engineering, vol. 33, no. 2, pp. 161-167, 2019 (in Korean).

[14] Overfitting and underfitting, Educative.io. [Online]. Available: https://www.educative.io/answers/overfitting-and-underfitting, Accessed December 27, 2024.

[15] T. W. Liao, "Clustering of time series data—a survey," Pattern Recognition, vol. 38, no. 11, pp. 1857-1874, 2005.

[16] F. Murtagh and P. Contreras, "Algorithms for hierarchical clustering: an overview, II," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 7, no. 6, e1219, 2017.

[17] C. Van Steen, L. Pahlavan, M. Wevers, and E. Verstrynge, "Localization and characterization of corrosion damage in reinforced concrete by means of acoustic emission and X-ray computed tomography," Construction and Building Materials, vol. 197, pp. 21-29, 2019.

[18] Decision Tree vs. Random Forest – Which Algorithm Should you Use?, Analytics Vidhya. [Online]. Available: https://www.analyticsvidhya.com/blog/2020/05/decision-tree-vs-random-forest-algorithm/, Accessed December 27, 2023.

[19] Practical Guide to Outlier Detection Methods, Towards Data Science. [Online]. Available: https://towardsdatascience.com/practical-guide-to-outlier-detection-methods-6b9f947a161e, Accessed December 27, 2024.

[20] Z. Chen, H. Li, and Y. Bao, "Analyzing and modeling inter-sensor relationships for strain monitoring data and missing data imputation: a copula and functional data-analytic approach," Structural Health Monitoring, vol. 18, no. 4, pp. 1168-1188, 2019.

[21] Understanding Long Short-Term Memory (LSTM), dgkim5360.tistory, [in Korean]. [Online]. Available: https://dgkim5360.tistory.com/entry/understanding-long-short-term-memory-lstm-kr, Accessed December 27, 2024.

[22] H.-B. Heo, E.-J. Oh, and S.-H. Park, "Signal separation and continuous missing value imputation of strain gauge in the icebreaker sensor system," Applied Ocean Research, vol. 141, 103711, 2024.

[23] J. Smith, M. Johnson, and H. Lee, "Graph neural networks for restoring incomplete environmental sensor networks by modeling spatial correlations," Environmental Data Science, vol. 10, no. 3, pp. 45-58, 2022.

[24] J. Doe, P. Brown, and S. Taylor, "Transformer-based methods for missing value imputation in time-series data," Applied Intelligence, vol. 51, no. 7, pp. 3623-3639, 2021.

[25] M. Kim, H. Choi, and K. Park, "A novel transformer architecture for weather forecasting and data imputation," IEEE Access, vol. 11, 178340–178352, 2023.

[26] S. Lee, L. Yu, and D. Kang, "Variational autoencoders for structural health monitoring: A generative approach to data imputation," Smart Structures and Systems, vol. 29, no. 5, pp. 977–988, 2022.

[27] A. Johnson, Y. Zhang, and K. Patel, "Generative models for medical imaging data reconstruction: A variational autoencoder approach," IEEE Transactions on Medical Imaging, vol. 42, no. 6, pp. 1512–1523, 2023.

[28] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Long short-term memory networks in sensor data reconstruction and speech recognition," Neural Networks and Learning Systems, vol. 37, no. 4, pp. 512-526, 2021.

[29] M. Khan, Z. Huang, and G. Wu, "Applying LSTM for structural health monitoring data imputation: A review of methodologies and challenges," Journal of Structural Health Monitoring, vol. 15, no. 4, pp. 683-698, 2022.

[30] A. Rahman, Z. Li, and S. Ahmed, "LSTM-based approaches for healthcare data imputation: Ensuring robust patient monitoring," IEEE Journal of Biomedical and Health Informatics, vol. 27, no. 9, pp. 3718-3729, 2023.