Deep Learning-Based Anomaly Detection to Classify Inaccurate Data and Damaged Condition of a Cable-Stayed Bridge

HYESOOK SON\textsuperscript{1}, YUN JANG\textsuperscript{1}, (Member, IEEE), SEUNG-EOCK KIM\textsuperscript{2}, DONGJOO KIM\textsuperscript{2}, AND JONG-WOONG PARK\textsuperscript{3}

\textsuperscript{1}Department of Computer Engineering and Convergence Engineering for Intelligent Drone, Sejong University, Seoul 05006, South Korea
\textsuperscript{2}Department of Civil and Environmental Engineering, Sejong University, Seoul 05006, South Korea
\textsuperscript{3}School of Civil and Environmental Engineering, Urban Design and Studies, Chung-Ang University, Seoul 06974, South Korea

Corresponding author: Yun Jang (jangy@sejong.edu)

This work was supported in part by the Basic Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (MSIT) under Grant 2019R1A4A1021702, and in part by the Institute of Information & Communications Technology Promotion (IITP) funded by the Korea Government, MSIT (Development of Big Data and AI Based Energy New Industry Type Distributed Resource Brokerage System), under Grant 2019-0-00374.

\textbf{ABSTRACT} Cables of cable-stayed bridges are gradually damaged by weather conditions, vehicle loads, and corrosion of materials. Stayed cables are an essential factor closely related to the stability of a cable-stayed bridge. Damaged cables might lead to the bridge collapse due to tension capacity lost. Therefore, it is necessary to develop structural health monitoring (SHM) techniques that check the cable conditions. Besides, the sensor network system development has contributed to the state analysis, such as damage detection and structural deformation, by allowing us to collect large-scale SHM data. However, the collected SHM data might include abnormal data due to device malfunctioning or unexpected environmental inconstancies. Furthermore, since data anomalies interfere with accurate structural evaluation, we need to identify anomalies and treat them appropriately in the data preprocessing stage. However, the cause of anomalies may be either temporary errors or actual structural deformation. Anomalies caused by structural damage or sensor device failure are informative data that must not be replaced or deleted. In this paper, we distinguish between anomalies as inaccurate data and anomalies related to the state of structures or sensor devices and propose a framework to identify each of them. We train a Long Short Term Memory (LSTM) network based Encoder-Decoder architecture that processes multivariate time series and learn temporal correlation. The trained LSTM network discovers anomalies by calculating anomaly scores. We determine the anomalies emerging intermittently as errors and correct the erroneous data. If the anomalies persist, we recognize the data as generated by bridge damage or sensor device failure. We evaluate the proposed technique with cable tension data from an actual cable-stayed bridge.

\textbf{INDEX TERMS} SHM, LSTM, anomaly detection, time series, deep learning.

\section*{I. INTRODUCTION}

The importance of structural health monitoring (SHM) is increasing to ensure the safety and durability of large structures such as tunnels, buildings, infrastructure, and bridges. The key to SHM is to detect damages within structures and avoid socio-economic losses such as infrastructure collapse. In cable-stayed bridges, the structures tend to be damaged and even collapsed due to various causes such as weather conditions, vehicle loads, and material corrosion. Stayed cables are an essential part of the cable-stayed bridge, which significantly influences the structural integrity as it transmits the load of the bridge deck to the pylons. Damaged cables may deteriorate the bridge condition due to the loss of load-carrying capacity [1]. Therefore, a strategy is needed to evaluate and analyze the cable condition accurately. The development of the sensor network system allows us to accumulate a vast amount of SHM data. Also, large-scale time-series data accumulated over time from the SHM system serves as the basis for structural assessment and is broadly
used in SHM technology. However, the SHM data contains anomalies, which are inaccurate values recorded for various reasons. In particular, anomalies caused by system malfunctions, data transmission problems, harsh environments, or sensor faults might cause distorted structural analysis and interfere with damage detection, which leads to inevitable data cleaning [2]. We can detect anomalies during the data preprocessing stage and manage detected anomalies by deleting or replacing them. However, some anomalies can be indispensably valuable data. For example, due to structural changes, sensor devices failure, environmental changes such as earthquakes, or material damage, the SHM data may be collected differently from the patterns when the structure and sensor devices are in a healthy state. Anomalies due to structural damage or sensor device failure are essential data that should not be deleted. However, there is no clear criterion to distinguish whether anomalies are caused by data collection system malfunction or structural damage. Therefore, detecting anomalies and appropriate handling is still a significant problem.

Since it is challenging to collect labeled data, and it is very inefficient to label all anomalous data manually, data preprocessing such as anomaly detection is an expensive task that requires high-level expertise and consumes time and labor. Moreover, it is complicated to detect or identify anomalies in time series data because there are fewer abnormal data than normal data. In particular, a multivariate time series such as data from an SHM system having many sensors increases the complexity of modeling for anomaly detection. Many researchers have proposed model-based SHM techniques that evaluate the structure state and diagnose anomalies using statistical models and machine learning to resolve these challenges.

Some researchers have extracted Damaged Sensitive Features (DSFs) to detect structure damages using time series models such as Autoregressive Moving Average Model (ARMA) [3], [4] and AutoRegressive Moving Average models with eXogenous inputs (ARMAX) [5]. Outliers could be either a single data that does not match the data set or a cluster [6]. The clustering, one of the outlier estimation methods in wireless sensor networks, has the advantage of not requiring prior knowledge about distribution in multivariate data [7]. The clustering-based outlier detection involves creating a grouped data set and determining outlier clusters [8]. For example, we can classify outliers by calculating anomaly scores based on the distances in the k-means clustering method [9]. Alamdarly et al. [10] propose a modified k-means clustering to identify the difference between the normal and distorted signals using the Spectral Moments (SMs). Diez et al. [11] apply the clustering technique to classify damaged group substructures or joints for bridges.

Deep learning models are tools for training complex non-linear correlations within datasets. When sufficient data is available for the network training, the deep learning models perform best in various tasks such as nonlinear feature extraction, classification, regression, and anomaly detection. Deep learning is also widely applied to detect structural damages. Pathirage et al. [12] train a deep autoencoder (DAE) to predict the steel frame structure state with modal information such as frequencies and mode shapes. They generate training data with an updated finite element model and predict the stiffness parameters representing the impairment state by supervised learning. They show that the proposed method enables efficient dimension reduction and accurate damage prediction. We can also detect outliers even in time-series data using deep learning models. Malhotra et al. [13] propose a Long Short-Term Memory Networks (LSTM) based Encoder-Decoder architecture for Anomaly Detection (EncDec-AD) that detects anomalies in multi-sensor time-series such as data collected from sensors mounted on mechanical devices. EncDec-AD is trained to reconstruct target sequence data identical to the input sequence data. Since they train the EncDec-AD with a ‘normal’ time-series that do not contain outliers, they detect sections with high reconstruction errors as anomalous sequences. However, they only detect when an outlier occurs but do not identify which sensor causes the outlier. Zang et al. [14] propose a Multi-Scale Convolutional Recurrent Encoder-Decoder (MSCRED) that diagnoses outliers in multivariate time-series data. They create a signature matrix representing the inter-correlations of two time-series pairs as input data for MSCRED. MSCRED consists of a convolutional encoder that captures a spatial pattern of signature matrices, an attention-based convolutional LSTM [15] that catches temporal dependencies, and a convolutional decoder that reconstructs the same signature matrix as the input data. The trained MSCRED detects in which time steps it contains outliers. Given the detection results, they rank the anomaly scores of each time series using the prediction error. Furthermore, they identify specific ranking series as sensors causing outliers. As presented in previous studies [13], [14], we can train a deep learning model and identify data with high prediction errors as anomalies.

In SHM, machine learning techniques have been proposed to detect anomalous data due to sensor system malfunction in time series data. Bao et al. [16] propose a two-step outlier detection method that converts time series data into images and trains deep learning classifiers. They divide the acceleration data of the long-span bridge into sections and plot each section to create an image. Moreover, they break the anomalies pattern into six categories: missing, minor, outlier, square, trend, drift, and manually label the visualized categorized data patterns. Then, they train the deep learning classifier to classify multi-pattern anomalies by the graphical characteristics of images. To further complement the study of Bao et al. [16], Tang et al. [17] converts raw time-series data into time response domain and frequency response domain by Fast Fourier Transform (FFT), and visualize both domains to make the time-domain characteristics abundant. Besides, they secure the ratio of each anomaly type to be the same by random selection to obtain a balanced training set. Furthermore, instead of the autoencoder neural network proposed by Bao et al. [16], Tang et al. [17] have a convolutional neural
network (CNN) learn representations of anomalies features. However, to train a classifier with supervised learning, labels for outliers are necessary. However, these labels are difficult to obtain, and all of them must be manually sorted if there is no label. Mao et al. [18] point out the class imbalance and incompleteness of anomalous patterns of the dataset as a difficulty of supervised methods for anomaly detection. They propose an unsupervised learning technique that combines the generative adversarial networks (GAN) and autoencoders (AE). They convert acceleration data collected from 14 vertical accelerometers installed on a long-span bridge into Gramian Angular Field images for network training. They improve the accuracy by transferring the structure and weights of the generator of the GANs to the AE decoder. Anomaly detection method from the viewpoint of safety analysis for structural condition evaluation and damage detection has also been proposed. Ni et al. [19] introduce a two-step framework consisting of anomaly detection and data compression. They first design a one-dimensional CNN and classify abnormal data from acceleration data of steel box girder suspension bridge in China. They note that abnormal data should be stored without compression since abnormal data is related to the structure state. Moreover, they compress the autoencoder network using only normal data and show that data can be recovered with high accuracy. Sarmadi and Karamodin [20] propose an anomaly detection method using adaptive Mahalanobis-squared distance and one-class kNN rule to detect early damage in various environments. Nguyen and Goulet [21] define anomaly as a change in structure behavior. They propose a combination of Bayesian Dynamic Linear Models and the Switching Kalman Filter and consider both a prior probability of an anomaly and transition probabilities between a normal and an abnormal state. The proposed technique is semisupervised learning and does not require normal and abnormal labels. They apply the proposed technique to the horizontal displacement data collected from the dam to identify anomalies caused by refection work. Gu et al. [22] propose a multilayer artificial neural network (ANN) that detects structural damages through unsupervised learning. To eliminate the effect of temperature on fluctuations in frequencies, they utilize the frequencies and temperature data into the ANN input data. Target data has the same frequencies as the input data, and they calculate the Euclidean distance between the ANN output and the target data for the anomaly detection. The ANN is trained only with data extracted from the undamaged structure. Therefore, when the Euclidean distance increases, it is possible to distinguish data from the damaged structure. In this way, it is possible to train a deep learning model for anomaly detection with an unsupervised approach without anomaly labels. Currently existing techniques do not distinguish between anomalies caused by temporary error and anomalies induced by structural damage or sensor malfunctioning. In this paper, we propose a framework for processing anomalies generated by two causes by training a deep learning model with multiple signals. As illustrated in Figure 1, the proposed method consists of anomaly detections in two stages. We train the LSTM with the normal data, which can process multivariate time series and learn temporal correlation. (a) is the first process to detect a damaged condition and inaccurate data. In the first anomaly detection step, the trained LSTM network computes anomaly scores of the test dataset containing abnormal data. We detect when anomalies occur based on the anomaly score and test the anomaly persistence. If anomalies continuously persist, we conclude that the data must be investigated since the structure condition is abnormal or the sensor device is defective. However, the proposed method does not distinguish whether the continuous outlier is caused by the bridge condition or the sensor device problem, which is a challenging subject to be investigated in the future. For temporarily fallacious errors, we treat them as intermittent outliers and we judge intermittent outliers as inaccurate data that interferes with the structure state analysis. We apply the interpolation to replace the detected inaccurate data and proceed into
the second anomaly detection step, as shown in Figure 1 (b). In the second outlier detection step, we check whether the inaccurate data is appropriately managed. If inaccurate data is correctly classified and replaced with appropriate values in the first outlier detection step, most outliers must have been eliminated in the second outlier detection step.

II. PROPOSED METHODOLOGY

In this section, we present our proposed method to detect anomalies for data preprocessing and anomalies for structural condition evaluation. Figure 1 illustrates the framework of the proposed method. We detect anomalies twice with a deep learning model trained with the multivariate time series tension data. In the first anomaly detection step, the deep learning model is employed to determine when the structure damages occur and discover when random inaccurate anomalies appear. In the second anomaly detection step, we apply the deep learning model again with the preprocessed time series to check if inaccurate outliers are handled properly. In this section, we investigate Long Short Term Memory (LSTM) network-based Encoder-Decoder for anomaly detection.

A. LONG SHORT TERM MEMORY NETWORKS BASED ENCODER-DECODER SCHEME FOR ANOMALY DETECTION

Long Short Term Memory (LSTM) networks [23] are recurrent neural networks (RNNs) that are widely adopted for time-series data training by catching complex temporal correlations through adjusting the information flow of sequence data, such as power plant data, multi-sensor engine data, space shuttle data, etc. Figure 2 present the LSTM structure. The input data \( x_t \) is the \( t \)-th \( m \)-dimensional vector of the input sequence. The LSTM computes the input gate \( i_t \), forget gate \( f_t \), and output gate \( o_t \) to update the cell state \( c_t \) and the hidden state \( h_t \). The equation of the LSTM is as follows.

\[
i_t = \sigma(W_{ix}x_t + W_{ih}h_{t-1} + b_i) \quad (1)
\]

\[
f_t = \sigma(W_{fx}x_t + W_{fh}h_{t-1} + b_f) \quad (2)
\]

\[
o_t = \sigma(W_{ox}x_t + W_{oh}h_{t-1} + b_o) \quad (3)
\]

\[
c_t = f_t \odot c_{t-1} + i_t \odot \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (4)
\]

\[
h_t = o_t \odot \tanh(c_t) \quad (5)
\]

where \( \odot \) is Hadamard product, \( \sigma \) is sigmoid function, and \( \tanh \) is the hyperbolic tangent. \( W \) and \( b \) represent the weight parameter and the bias respectively.

Sutskever et al. [24] proposes an Encoder-Decoder network composed of LSTM layers and shows that the long-term dependency can be trained. Figure 3 shows the Encoder-Decoder structure. The input data is a time-series with length \( L \), \( X = [x_1, x_2, \ldots, x_L] \), which is also the target data. When \( X \) is applied to the encoder, the hidden state of LSTM, \( H^E_L = [h^E_1, h^E_2, \ldots, h^E_L] \), is updated. Here, the dimension of \( h^E_t \) is the number of LSTM units. The encoder maps the input sequence to a hidden state, which is a vector of fixed dimensions. For the information transfer from the encoder to the decoder, the decoder hidden state is initialized to the last hidden state of the encoder. We have hidden states, \( H^D = \{h^D_1, h^D_2, \ldots, h^D_L\} \), and use the \( H^D_0 \) to obtain the prediction, \( X' = [x'_1, x'_2, \ldots, x'_L] \), of the target sequence. Therefore, the decoder plays the role of reconstructing the vector represented by the encoder as a sequence again. The Encoder-Decoder structure is employed for tasks in various fields such as machine translation, speech recognition, and multiple predictions as it can take a sequence and reconstruct a sequence. Malhotra et al. [13] propose an outlier detection technique using an LSTM based Encoder-Decoder network. They train LSTM network so that the encoder receives the sequence and reconstructs the input sequence in the decoder. Since only normal sequences are used for the LSTM network training, they remark that the reconstruction error will be higher than that of the normal sequences given anomalous sequences. Therefore, they calculate the anomaly score with the network error and detect the segments containing outliers.

B. DATA PREPROCESSING AND DAMAGE DETECTION USING ANOMALY SCORES

As shown in Figure 1, we first train the LSTM network with the normal tension data. Then, we compute the anomaly score by the trained network with the test dataset, including abnormal tension data. We utilize the anomaly score to discover...
when anomalies occur, which is called anomaly detection. The anomaly score $a_i$ at point $i$ for anomaly detection is defined as the log of the probability density function of the multivariate normal distribution as follows

$$a_i = -\log \left( \frac{1}{\sqrt{(2\pi)^d \det \Sigma}} \times \exp \left( -\frac{1}{2} (x_i - \mu)^T \Sigma^{-1} (x_i - \mu) \right) \right),$$

(6)

where $e_i$ is the reconstruction error, $|x_i - x'|$, and $m$ is the dimension of $X$. $\mu$ and $\Sigma$ are the parameters of the normal distribution, $N(\mu, \Sigma)$. We approximate $\mu$ and $\Sigma$ with the maximum likelihood estimation using randomly sampled data from the training set. The anomaly is the point, $\{ i | a_i > \gamma \}$, where the anomaly score $a_i$ is greater than the threshold $\gamma$. Then, we check whether anomalies are transient or continuous and presume the transient anomalies as inaccurate data.

We consider the data with continuous anomalies as abnormal signals caused by the bridge structure damage or defective sensor devices. To process the inaccurate data, we classify which cable tension causes the anomalies within the data that contains transient anomalies, which is called anomaly identification. We identify the cable with the largest prediction error as the anomaly at the point $i$ where the anomaly score is greater than the threshold $[14]$. Therefore, we pay attention to the tension data of the cable whose reconstruction error $e_i$ is the maximum. We replace the abnormal tension data of the associated cable with interpolated values to obtain the preprocessed time series data. We manage the inaccurate data utilizing the spline interpolation. We again calculate the anomaly score using the LSTM network with the preprocessed time series data. Finally, we evaluate the suitability of interpolation by checking how many outliers have been removed in the preprocessed data.

### III. EVALUATION OF THE PROPOSED ANOMALY DETECTION METHOD

In this section, we analyze the performance of the proposed anomaly detection method utilizing tension data from the actual cable-stayed bridge. In the first anomaly detection step, we present that the proposed framework successfully evaluates the bridge condition by the damage detection. Then, we examine the inaccurate data detection and anomaly identification accuracy. In the second step, we confirm that inaccurate data is appropriately managed with the preprocessed time series data.

### A. DATA DESCRIPTION

This study utilizes the data provided by the International Project Competition for Structural Health Monitoring (IPC-SHM2020). Cable tension was monitored at a double tower and double cable-plane cable-stayed bridge in China. This cable-stayed bridge consists of 168 cables, and each cable is given a name according to the rules. The cables are numbered as 1-21 toward the direction of the bank and river from the tower. N or S is assigned depending on whether it is North or South, and A or J is assigned according to whether it is on the bank or river side. Moreover, S or X is assigned according to whether it is up or downstream. The tension data were collected from the 14 cables, SJS08 to SJS14 and SJX08 to SJX14. The data were captured for ten days, 2006-05-13 to 2006-05-19, 2007-12-14, 2009-05-05, and 2011-11-01. The data sampling frequency was 2 Hz, and 172,800 tension data records were stored per day from one cable.

Table 1 shows a summary of statistics for the tension data of each cable. We observe that the tension data of the SJS13 cable contains negative numbers. Also, the maximum value of the SJS13 data is substantially compared to the maximum value of other cables. Since the first quartile, median, and third quartile of SJS13 are not very different from other cables, we assume that the SJS13 data contains a few outliers, whose values are either too large or too small. We also suspect the sensor is defective in the 2011-11-01 data, where the tension of the SJS13 cable is much smaller than the others. There is a technique to detrend the data with domain knowledge. Li et al. [25] evaluate the bridge state after excluding the effects of external load, sensor error, and zero-shift of sensors from the data decomposed.

| cable  | min   | Q1    | median | mean  | Q3    | max  | std  |
|--------|-------|-------|--------|-------|-------|------|------|
| SJS08  | 2119.005 | 2181.019 | 2185.262 | 2183.009 | 2191.802 | 2404.355 | 23.827 |
| SJS09  | 2307.982 | 2386.797 | 2391.532 | 2394.551 | 2402.281 | 2715.739 | 38.86  |
| SJS10  | 2396.246 | 2467.282 | 2474.061 | 2473.07  | 2481.431 | 2697.429 | 23.216 |
| SJS11  | 1991.049 | 2433.727 | 2473.528 | 2466.52  | 2533.458 | 2540.727 | 2545.953 |
| SJS12  | 2466.52  | 2533.458 | 2540.727 | 2545.953 | 2552.87  | 2805.937 | 24.205 |
| SJS13  | -13657.9 | 2830.562 | 2838.355 | 2837.763 | 2846.436 | 18372.76 | 68.837 |
| SJS14  | 2907.944 | 2980.916 | 2990.826 | 2990.452 | 2999.101 | 3218.309 | 23.904 |
| SJX08  | 68.427  | 2206.922 | 2277.202 | 2046.559 | 2282.758 | 2456.852 | 660.242 |
| SJX09  | 2071.368 | 2147.378 | 2351.79  | 2290.97  | 2337.832 | 2534.484 | 108.17 |
| SJX10  | 2352.811 | 2387.504 | 2444.887 | 2427.888 | 2447.813 | 2540.863 | 33.637 |
| SJX11  | 1826.704 | 2403.631 | 2523.213 | 2417.343 | 2478.397 | 2487.014 | 2487.468 |
| SJX12  | 2417.343 | 2478.397 | 2487.014 | 2487.468 | 2494.771 | 2773.543 | 21.998 |
| SJX13  | 1071.588 | 2839.545 | 2857.981 | 2679.686 | 2864.968 | 3075.099 | 536.118 |
| SJX14  | 2799.36  | 2974.378 | 2988.092 | 2978.416 | 3004.666 | 3224.864 | 52.035 |

1. http://sstl.cee.illinois.edu/ipc-shm2020/
FIGURE 4. Tension data of 14 cables on (a) 2006-05-15, (b) 2006-05-18, (c) 2007-12-14, (d) 2009-05-05, and (e) 2011-11-01. (a) and (b) are normal data, and (c) and (d) contain inaccurate data. (e) is the data with the bridge damage. In (c) and (d), the red boxes indicate the enlarged outlier data segment. We see that ten outliers occurred consecutively, which means that the outliers persist for 5 seconds.

with a low-pass filter. However, the proposed technique is a data-based outlier detection that does not require domain knowledge. In the future, we plan to investigate LSTM results with the detrended tension data. Figure 4 shows the tension data on (a) 2006-05-15, (b) 2006-05-18, (c) 2007-12-14, (d) 2009-05-05, and (e) 2011-11-01. Unlike the 7-day data in 2006, anomalies such as Figure 4 (c) and (d) appear on 2007-12-14 and 2009-05-05. All of these anomalies are found in the data from the cable SJS13. In the case of the SJS13 data on 2007-12-14, negative values are found at 21,600, 25,200, 46,800, 54,000, 61,200, 79,200 seconds, all of which are accurately divided by 3,600 seconds. Since there is a negative value exactly every hour, we judge the data as inaccurate data generated by an unusual circumstance. We manually label inaccurate data on 2007-12-14 and 2009-05-05 to evaluate the proposed method. Our first evaluation aims to detect inaccurate data and identify cables with abnormal tension values, such as SJS13. Meanwhile, one of 14 cables is known as damaged on 2011-11-01. We verify whether our deep learning model can classify the data on 2011-11-01 as a damaged condition in the first anomaly detection step, as described in Section II. Although it is presumed that there is a sensor anomaly in the 2011-11-01 data, our model does not distinguish whether the cause of the persistent outlier is a sensor failure or damage to the bridge. Therefore, the 2011-11-01 data is labeled as damaged.

B. TRAINING LSTM NETWORK
Normal data are utilized for training a deep learning model in the studies [13], [14], [22], and the test set includes both normal and abnormal data. In this study, we use the normal data in 2006-05-13 to 2006-05-16 as a training set. We randomly sample 25% of the training set to estimate \( \mu \) and \( \Sigma \) in Equation (6). The data on 2006-05-17 and 2007-12-14 are set to the validation set, which includes both inaccurate data and normal data to determine the anomaly score threshold. In the validation set, we set the threshold \( \gamma \) to the larger score between the average anomaly score in the normal data and the average anomaly score in the abnormal data. We use the data on 2006-05-18, 2006-05-19, 2009-05-05, and 2011-11-01 as a test set to contain all of the normal data, inaccurate data, and damaged bridge data. Then, there are 691,200 training set, 345,600 validation set, and 691,200 test set. Moreover, in order to input each dataset into the LSTM network, we divide the data for ten days into 120 length intervals. Therefore, the dataset consists of sequences with 14-dimension with length 120. Also, we scale the data in the range between 0 and 1 with the Min-Max normalization.

As shown in Figure 3, we configure the encoder and decoder as two LSTM layers each and set the dimension of hidden states to 16. We train the model to minimize the mean absolute error of the network using the ADAM optimizer [26] with a learning rate of 0.0001. We set the number of epochs to 100 and the batch size to 128. We implement the network using the Pytorch library [27].

C. RESULTS
We determine the data as an anomaly when the anomaly score is greater than the threshold \( \gamma \) set to 4,084.859, which is the
average of the anomaly scores for abnormal data on 2007-12-14 in the validation set. Figure 5 presents anomaly scores for 10 days including normal data in (a) to (g), inaccurate data in (h) and (i), and damaged condition data in (j). As shown in Figure 5 (a) to (g), the anomaly scores of the training set, validation set, and test set are less than threshold $\gamma$. However, in (h) and (i), we observe that the anomaly scores are relatively low most of the time but significantly high at some times. We notice that in (i), the anomaly scores exceed the threshold $\gamma$ only at the sudden high anomaly scores. Since the abnormality and normality are repeated in the data on 2007-12-14 and 2009-05-05, we assume that anomalies identified by our deep learning model occur intermittently. As mentioned earlier, data on 2007-12-14 and 2009-05-05 are labeled as inaccurate data in Section III-A. We, therefore, verify that it is possible to classify inaccurate data from predicting intermittent anomalies. Similarly, in (j), intermittently high anomaly scores and relatively low anomaly scores are seen, but overall anomaly scores are further increased. In (j), the minimum anomaly score is 731,282.984, and the anomaly scores for the damaged condition data are always greater than the threshold $\gamma$. The data in 2011 contain signals for both damaged data and inaccurate data. The entire data on 2011-11-01 are labeled as abnormal data. We observe a series of anomalies in the data on 2011-11-01 when the bridge is damaged. From Figure 5, which shows the anomalies detected in the first step, the anomalies in the damaged condition are persistent.

Figure 6 shows the prediction performance for the first anomaly detection. (a) presents the receiver operating characteristic (ROC) curve, and the area under the curve (AUC) is 0.9905. The ROC curve is employed for the quantitative evaluation of the model by plotting the FP rate against the TP rate for several thresholds. Since the AUC is close to 1, we can tell that the proposed anomaly detection method is very efficient. (b) reveals the confusion matrix when the
FIGURE 7. (a) is raw tension data and (b) is the preprocessed data after the first anomaly detection step. The red box is the zoomed-in part of the SJS13 cable.

TABLE 2. Classified results for anomaly identification.

| No | SJS08 | SJS09 | SJS10 | SJS11 | SJS12 | SJS13 | SJS14 | SX08 | SX09 | SX10 | SX11 | SX12 | SX13 | SX14 |
|----|-------|-------|-------|-------|-------|-------|-------|------|------|------|------|------|------|------|
| 0  | 0     | 0     | 0     | 0     | 0     | 261   | 0     | 0    | 27   | 0    | 0    | 0    | 0    |

threshold $\gamma$ is determined using the validation set. (c) presents a part of the actual outliers and detected outliers in the data of the cable SJS13. As shown in (c), only one outlier sample is not correctly classified. TN and TP are the numbers of samples for which normal data and anomalies are correctly classified, respectively. FP is the number of samples for which the detection model incorrectly classifies normal data as anomalies. FN is the number of samples for which the detection model does not classify abnormal data as anomalies. Precision, recall, accuracy, and F1 scores are computed as follows.

\[
\text{precision} = \frac{TP}{TP + FP} \tag{7}
\]

\[
\text{recall} = \frac{TP}{TP + FN} \tag{8}
\]

\[
\text{accuracy} = \frac{TP + TN}{2TP + TN + FN + FP} \tag{9}
\]

\[
F1 = \frac{2TP}{2TP + FP + FN} \tag{10}
\]

In the anomaly detection results, precision, recall, accuracy, and F1 scores are 0.9568, 0.9201, 0.9998, and 0.9381, respectively, and all are above 0.9. For the anomaly identification, we distinguish the cable with the highest anomaly score in the interval where the anomaly detection prediction is the same as the actual label. Table 2 shows the number of samples identified as data causing anomaly for each cable. As discussed in Section III-A, the cable containing inaccurate data is always SJS13. To check the classification accuracy for anomaly identification, we classify the cables only for the actual anomaly labels. Note that we do not reflect the predicted labels. As seen in Table 2, most samples on SJS13 are classified as inaccurate data and the accuracy is 0.90625.

We have preprocessed the raw tension data by substituting the data predicted as anomalies with the interpolated values. In Figure 7, (a) and (b) are the tension data before and after the preprocessing, respectively. The sharp increase or decrease in (a) does not appear in (b). From Figure 7, it is visually confirmed that most outliers have disappeared in the preprocessed data. The preprocessed data is applied in the second anomaly detection step to check whether the outliers have been appropriately handled. If inaccurate data occur frequently, the quality of the data deteriorates. Also, it is not easy to analyze the structure state properly. Hence, it is necessary to make sure that inaccurate data must be suitably treated and replaced. The anomaly scores are computed with the LSTM network in the second anomaly detection step after replacing the outliers with the interpolated values. Similar to the first step, anomalies are classified with anomaly scores greater than the threshold $\gamma$. In Figure 8, (a) and (b) show 2009-05-05 data classified as outliers in the first anomaly detection step and the second anomaly detection step, respectively. The anomaly score in (b) has been generally reduced, and the number of outliers has been significantly reduced. To be precise, the number of outliers classified before and after the data preprocessing is 301 and 35, respectively. Outliers still occur after the preprocessing step. The problem could be flaws in the misclassified data and interpolation in the first outlier detection stage. Also, the years in which the training set and the test set were collected are different. The training set contains data collected only in 2006, and the test set contains data in 2006, 2009, and 2011. Different data collecting years could make accurate outlier classification more complicated. Although there are errors, we notice that most outliers are removed compared to the first step.

Figure 9 (a) and (b) present the number of aggregated anomalies by 5 seconds within the data on 2009-05-05 in the first anomaly step and the second anomaly step, respectively. The red dot indicates the point where all of ten tension data samples for 5 seconds are classified as outliers. From the data...
on 2009-05-05, we recognize data samples where the outliers continue to occur for 5 seconds. As confirmed in Figure 8, most outliers have been deleted after the data preprocessing, and accordingly, most outliers lasting for 5 seconds disappear. In particular, the number of red dots representing the data samples where the outliers persist for 5 seconds decreases from 20 to 1. Since the data collection period was 24 hours, we could easily separate the days containing inaccurate data from the days when the bridge was damaged. We classify data on 2011-11-01 containing continuous outliers for 24 hours as the damaged state in the first outlier detection step. Moreover, we classify the data on 2009-05-05 as inaccurate data rather than classifying it as damaged. However, depending on the structure domain knowledge, the judgment on the outlier duration in the damaged state can be very diverse. Setting a threshold for the length of successive outliers will be a challenging task. Training data containing two labels may be required to distinguish between inaccurate data and damage conditions accurately. Alternatively, we can use the outliers detected in the second anomaly detection step. In this study,
outliers lasting about 5 seconds could be resolved simply by the interpolation. However, even though the interpolation error is low, if there are many data samples whose outliers are not well processed in the second step, it may be wrong to classify the data as incorrect data. We will work on classifying the two labels in the future.

**IV. CONCLUSION**

In this paper, we introduced an anomaly detection method based on deep learning to examine outliers in SHM data. We designed a multilayered LSTM and applied multi-sensor time-series data from multiple cables into the network. We classified abnormal data by calculating anomaly scores with the reconstruction errors through the LSTM network trained only with normal data. We verified the network performance trained by the unsupervised learning method using tension data measured on 14 cables of an actual cable-stayed bridge. In the first anomaly detection step, we checked the persistence of outliers to discover damaged conditions and inaccurate data and confirmed that the proposed method achieved high accuracy. If data in 2011, when the bridge was actually damaged, was applied into the network, all data samples were classified as outliers. In this way, it is possible to identify continuous outliers as alarms in the SHM system for the structural condition evaluation. Meanwhile, we considered intermittent outliers as inaccurate data. We searched for the cable locations that caused outliers to manage inaccurate data. We applied the interpolation method to replace inaccurate data, and then in the second anomaly detection step, we checked whether the outliers of the interpolated data were processed correctly. Most of the inaccurate data were converted to normal data in the second step, despite obstacles such as interpolation error and different collection periods. However, the proposed method does not differentiate the exact cause of outliers. The causes of outliers include structural damage, system malfunctions, data transmission problems, harsh environments, and sensor device failure. In the future, we will investigate the true source of outliers by capturing data trends and patterns that vary depending on the cause of outliers and the persistence of outliers. Moreover, we will study the outlier persistence further. We have shown that we can classify the damage status and inaccurate data based on outlier persistence. Since the length of continuous outliers is either around 24 hours or 5 seconds, we could easily separate the outlier intervals in our dataset. However, the degree of outlier persistence can vary depending on the dataset and the cause of outliers. Therefore, we will examine the degree of persistence and the data pattern while identifying the cause of the outliers.

**REFERENCES**

[1] H. Jo, S.-H. Sim, K. A. Mechitov, R. Kim, J. Li, P. Moinzadeh, B. F. Spencer, Jr., J. W. Park, S. Cho, H.-J. Jung, C.-B. Yun, J. A. Rice, and T. Nagayama, “Hybrid wireless smart sensor network for full-scale structural health monitoring of a cable-stayed bridge,” Proc. SPIE, vol. 7981, Apr. 2011, Art. no. 79810F.

[2] Y. Bao, Z. Chen, S. Wei, Y. Xu, Z. Tang, and H. Li, “The state of the art of data science and engineering in structural health monitoring,” Engineering, vol. 5, no. 2, pp. 234–242, 2019.

[3] K. K. Nair, A. S. Kiremidjian, and K. H. Law, “Time series-based damage detection and localization algorithm with application to the ASCE benchmark structure,” J. Sound Vib., vol. 291, nos. 1–2, pp. 349–368, Mar. 2006.

[4] L. Yu and J.-C. Lin, “Cloud computing-based time series analysis for structural damage detection,” J. Eng. Mech., vol. 143, no. 1, Jan. 2017, Art. no. C4015002.

[5] Q. Mei and M. Gül, “Novel sensor clustering–based approach for simultaneous detection of stiffness and mass changes using output-only data,” J. Struct. Eng., vol. 141, no. 10, Oct. 2015, Art. no. 04014237.

[6] L. Duan, L. Xu, Y. Liu, and J. Lee, “Cluster-based outlier detection,” Ann. Oper. Res., vol. 168, no. 1, pp. 151–168, 2009.

[7] Y. Zhang, N. Meratnia, and P. Havinga, “Outlier detection techniques for wireless sensor networks: A survey,” IEEE Commun. Surveys Tuts., vol. 12, no. 2, pp. 159–170, 2nd Quart., 2010.

[8] S.-Y. Jiang and Q.-B. An, “Clustering-based outlier detection method,” in Proc. 5th Int. Conf. Fuzzy Syst. Knowl. Discovery, vol. 2, Oct. 2008, pp. 429–433.

[9] R. Pamula, J. K. Deka, and S. Nandi, “An outlier detection method based on clustering,” in Proc. 2nd Int. Conf. Emerg. Appl. Inf. Technol., Feb. 2011, pp. 253–256.

[10] M. M. Alamdari, T. Rakotoariveloh, and N. L. D. Khoa, “A spectral-based clustering for structural health monitoring of the Sydney Harbour Bridge,” Mech. Syst. Signal Process., vol. 87, pp. 384–400, Mar. 2017.

[11] A. Diez, N. L. D. Khoa, M. M. Alamdari, Y. Wang, F. Chen, and P. Runcie, “A clustering approach for structural health monitoring on bridges,” J. Civil Struct. Health Monit., vol. 6, no. 3, pp. 429–445, Jul. 2016.

[12] C. S. N. Pathirage, J. Li, L. Li, H. Hao, and W.-Q. Liu, “Application of deep autoencoder model for structural condition monitoring,” J. Syst. Eng. Electron., vol. 29, no. 4, pp. 873–880, 2018.

[13] P. Malhotra, A. Ramakrishnan, G. Anand, L. Vig, P. Agarwal, and G. M. Shroff, “LSTM-based encoder-decoder for multi-sensor anomaly detection,” CoRR, vol. abs/1607.00148, pp. 1–5, Jul. 2016. [Online]. Available: http://arxiv.org/abs/1607.00148
[14] C. Zhang, D. Song, Y. Chen, X. Feng, C. Lumezanu, W. Cheng, J. Ni, B. Zong, H. Chen, and N. V. Chawla, “A deep neural network for unsupervised anomaly detection and diagnosis in multivariate time series data,” in Proc. AAAI Conf. Artif. Intell., vol. 33, Palo Alto, CA, USA: AAAI Press, 2019, pp. 1409–1416.

[15] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-C. Woo, “Convolutional LSTM network: A machine learning approach for precipitation nowcasting,” in Proc. Adv. Neural Inf. Process. Syst., vol. 28, 2015, pp. 802–810.

[16] Y. Bao, Z. Tang, H. Li, and Y. Zhang, “Computer vision and deep learning–based data anomaly detection method for structural health monitoring,” Struct. Health Monit., vol. 18, no. 2, pp. 401–421, 2019.

[17] Z. Tang, Z. Chen, Y. Bao, and H. Li, “Convolutional neural network-based data anomaly detection method using multiple information for structural health monitoring,” Struct. Control Health Monit., vol. 26, no. 1, p. e2296, Jan. 2019.

[18] J. Guo, M. Gul, and X. Wu, “Damage detection under varying temperature using artificial neural networks,” Struct. Control Health Monit., vol. 20, no. 4, pp. 1609–1626, 2020.

[19] F. Ni, J. Zhang, and M. N. Noori, “Deep learning for data anomaly detection and data compression of a long-span suspension bridge,” Comput.-Aided Civil Infrastruct. Eng., vol. 35, no. 7, pp. 685–700, Jul. 2020.

[20] H. Sarmadi and A. Karamodin, “A novel anomaly detection method based on adaptive Mahalanobis-squared distance and one-class kNN rule for structural health monitoring under environmental effects,” Mech. Syst. Signal Process., vol. 140, Jun. 2020, Art. no. 106495.

[21] L. H. Nguyen and J.-A. Goulet, “Anomaly detection with the switching Kalman filter for structural health monitoring,” Struct. Control Health Monit., vol. 25, no. 4, p. e2136, Apr. 2018.

[22] J. Guo, M. Gul, and X. Wu, “Damage detection under varying temperature using artificial neural networks,” Struct. Control Health Monit., vol. 24, no. 11, p. e1998, Nov. 2017.

[23] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997.

[24] I. Sutskever, O. Vinyals, and Q. V. Le, “Sequence to sequence learning,” in Proc. Adv. Neural Inf. Process. Syst., 2014, pp. 3104–3112.

[25] S. Li, S. Wei, Y. Bao, and H. Li, “Condition assessment of cables by pattern recognition of vehicle-induced cable tension ratio,” Eng. Struct., vol. 155, pp. 1–15, Jan. 2018.

[26] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” in Proc. 3rd Int. Conf. Learn. Represent. (ICLR), Y. Bengio and Y. LeCun, Eds., San Diego, CA, USA, May 2015, pp. 1–15. [Online]. Available: http://arxiv.org/abs/1412.6980

[27] A. Paszke et al., “Pytorch: An imperative style, high-performance deep learning library,” in Advances in Neural Information Processing Systems, vol. 32, H. Wallach, H. Larochelle, A. Beygelzimer, F. d’Alché-Buc, E. Fox, and R. Garnett, Eds. Red Hook, NY, USA: Curran Associates, 2019, pp. 8024–8035. [Online]. Available: http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf

HYEYJONG PARK receives the B.S. degree in civil engineering from Hanyang University, South Korea, in 2008, and the M.S. and Ph.D. degrees in civil engineering from KAIST, South Korea, in 2009 and 2013, respectively. He currently works as an Assistant Professor with Chung-Ang University, South Korea. His research interests include wireless smart sensor networks for SHM, and sensor fusion algorithms.

YUN JANG (Member, IEEE) received the bachelor’s degree in electrical engineering from Seoul National University, South Korea, in 2000, and the master’s and Ph.D. degrees in electrical and computer engineering from Purdue University, in 2002 and 2007, respectively. From 2007 to 2011, he was a Postdoctoral Researcher with CSCS and ETH Zürich, Switzerland. He is currently an Associate Professor of computer engineering with Sejong University, Seoul, South Korea. His research interests include machine learning, interactive visualization, and data analytics.

SEUNG-EOCK KIM received the Ph.D. degree from Purdue University, in 1996. Since 1996, he has been a Professor with Sejong University, South Korea. He has his expertise in the development and evaluation of ultimate strength of structures using nonlinear inelastic analysis, LRFD design of steel and composite structures, AI application to structure analysis.

DONGJOO KIM received the Ph.D. degree from the University of Michigan, Ann Arbor, MI, USA, in 2009. Since 2009, he has been an Assistant Professor and an Associate Professor with Sejong University, South Korea. He has his expertise in the development and evaluation of high performance cement based construction materials with high tensile strength, ductility and energy absorption capacity in addition to self-sensing or self-healing capability.

H. Son et al.: Deep Learning-Based Anomaly Detection to Classify Inaccurate Data and Damaged Condition