

Article

Investigation of Issues in Data Anomaly Detection Using Deep-Learning- and Rule-Based Classifications for Long-Term Vibration Measurements

Imdad Ullah Khan ^{1,2}, Seunghoo Jeong ³  and Sung-Han Sim ^{2,*} 

¹ Department of Civil, Architectural and Environmental System Engineering, Sungkyunkwan University, Suwon 16419, Republic of Korea

² Department of Global Smart City, Sungkyunkwan University, Suwon 16419, Republic of Korea

³ Advanced Railroad Civil Engineering Division, Korea Railway Research Institute, Uiwang 16105, Republic of Korea

* Correspondence: ssim@skku.edu

Abstract: Structural health monitoring (SHM) systems are widely used for civil infrastructure monitoring. Data acquired from the SHM systems play an important role in assessing structural integrity and determining further maintenance activities. Considering that sensors in the SHM systems are installed in a harsh environment for long-term measurements, some sensors can malfunction and produce faulty data. As a large amount of measured data is often desired to be automatically processed and can adversely affect structural assessments, identifying such abnormal data is important. This paper provides critical investigations of the automated detection of data anomalies using existing deep-learning-based classification in conjunction with a simple rule-based approach. The issues investigated in this study include (1) the presence of ambiguous data that cannot be categorized as an anomaly class in the literature, (2) information loss during the conversion of time-series data into images for the deep-learning-based approach, and (3) additional issues, such as misclassification by trained models and requirements of the threshold section in the rule-based approach. The results of these key investigations can be utilized to develop an effective anomaly detection process.

Keywords: structural health monitoring; anomaly detection; rule-based approach; deep learning



Citation: Khan, I.U.; Jeong, S.; Sim, S.-H. Investigation of Issues in Data Anomaly Detection Using Deep-Learning- and Rule-Based Classifications for Long-Term Vibration Measurements. *Appl. Sci.* **2024**, *14*, 5476. <https://doi.org/10.3390/app14135476>

Academic Editor: Yang Xu

Received: 22 May 2024

Revised: 19 June 2024

Accepted: 19 June 2024

Published: 24 June 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the recent development of advanced sensing techniques, structural health monitoring (SHM) is considered an effective maintenance tool for civil engineering structures [1–3]. SHM systems were deployed in several full-scale structures for long-term monitoring [2,4–6]. Monitoring data from SHM systems can be used to evaluate structural performance, predict the future behavior of structures, detect damage, and assess the condition of structures [7–12]. However, the sensors in SHM systems often malfunction, resulting in various types of data anomalies. Identifying and removing anomalies before data processing are crucial for proper SHM. Thus, effective approaches for handling such anomalies in the measured data are essential.

Data-driven approaches, including statistical approaches, have been studied for identifying anomalies in SHM data [13–17]. A method using principal component analysis was proposed for detecting, isolating, and reconstructing faulty sensor data [18]. A Mahalanobis-distance-based algorithm was used to detect outliers in the time-series data obtained from different structures in the laboratory [19]. Based on the Gaussian process approach, the sensor anomalies were identified using the residual between the measured and expected responses [20]. Although these statistics-based approaches can

achieve reasonably high performance, they are intrinsically limited to data with a single type of data anomaly.

As a result, deep learning was introduced to resolve problems associated with statistics-based identification. Recently, deep learning has been widely used for the automatic analysis of time-series data, such as prediction [21], classification [22,23], and anomaly detection [24–30]. One of the earliest approaches employing deep learning was the development of a deep neural network in which the inputs were images converted from measured time-history data [25]. For better classification performance, an approach using a convolutional neural network (CNN) was proposed with enhanced input information, combining images of time- and frequency-domain data [26]. Furthermore, multichannel images consisting of the time-domain response, spectrogram, and probability density function information were proposed to further increase the information in the input images for accurate anomaly detection [27]. Although data enhancement in the images used for training the neural networks was shown to improve the classification performance, the methods employing converted images have critical issues, such as ambiguous labeling of training images, imbalanced data for each class, and limited performance due to low-resolution converted images.

Therefore, investigating effective and practical anomaly detection using deep learning is necessary. Most importantly, the measured time-history data often have an intrinsic ambiguity that cannot classify a dataset into either normal or abnormal cases. For example, the “Trend” category [26] indicates a data anomaly with an overall increasing or decreasing trend; knowing that no dataset has exactly a zero mean, data labeling implies ambiguity that makes the deep learning approaches difficult. Furthermore, the converted images used as input to the CNN typically have a low resolution compared to the original time histories, which can lead to similar effects on the quantization errors. These issues should be appropriately considered before applying image-classification-based deep-learning approaches for detecting data anomalies.

This paper presents a comparative study of two automated data anomaly detection schemes of CNN- and rule-based classifications to provide in-depth discussions regarding the important issues described previously. Herein, rule-based classification is selected because it has advantages in exploring the issues in that (1) conversion to images is unnecessary, (2) class ambiguity can be better investigated and handled, and (3) classification performance is less sensitive to false detections. This paper first defines rules for data anomaly classes typically used in the literature and prepares labeled datasets from actual long-term SHM data. Subsequently, CNN-based and rule-based classification approaches are compared to identify the characteristics and limitations of these methods for data anomaly detection.

2. Background: CNN-Based Classification

Among the state-of-the-art approaches for anomaly detection described in the Introduction, CNN-based classification using converted images [26] was selected for comparison. This approach involves converting both time- and frequency-domain data into images that are, in turn, used as the input to a CNN model for classifying multi-class data anomalies. A total of seven classes are considered: *Missing* (most or all of the data are missing), *Minor* (the vibration response oscillates with a small amplitude), *Outlier* (one or more outliers appear in data), *Square* (the vibration response oscillates abnormally over a range of the accelerometer), *Trend* (the response is non-stationary with a monotonous trend), *Drift* (the response is non-stationary with random drift), and *Normal*. This method, validated using acceleration data measured from a cable-stayed bridge in China, exhibited 94% accuracy.

In the data preprocessing step, the measured time-history data and their Fourier transforms are converted into images to prepare the input to the CNN. Image conversion includes plotting the time and frequency responses without axes and storing them as images. The single-channel images of the time- and frequency-domain data were stacked

together to form a dual-channel image in which the time and frequency responses were stored in the red and green channels, respectively. The size of each dual-channel image was 100×100 pixels. Only the acceleration response of the bridge was considered for anomaly detection [26].

The neural network used for data anomaly detection includes typical convolutional layers. The architecture of the CNN is shown in Figure 1. The first hidden layer is a feature map with a size of 60×60 , generated after implementing the convolutional layer with 20 filters of a size of 41×41 and a stride of one. The third layer is of size 30×30 , generated after implementing a max pooling layer with a pooling size of 2×2 and a stride of two.

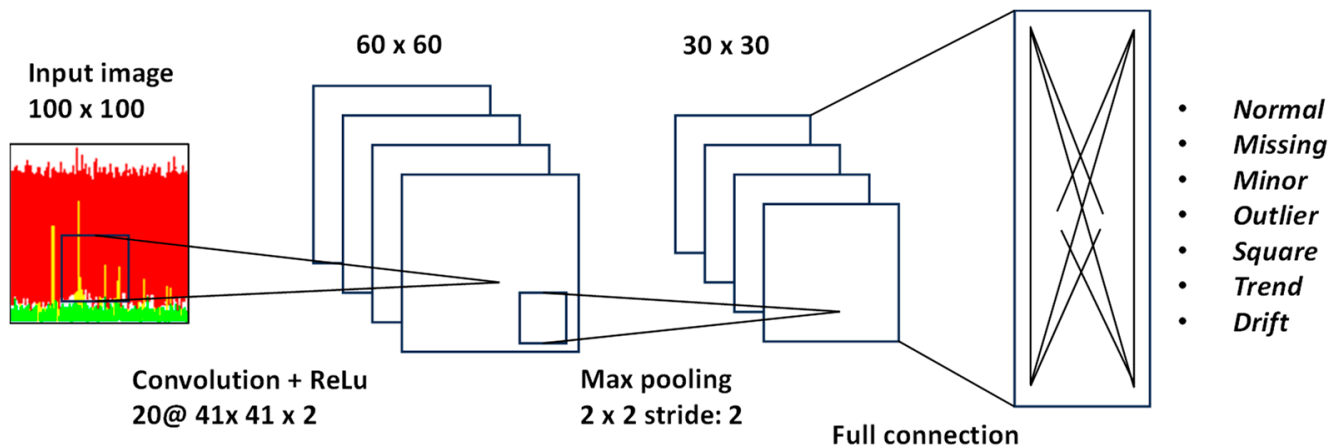


Figure 1. Detailed CNN architecture [26].

Although the CNN-based approach has shown potential for data anomaly detection, several important issues have been identified, as discussed previously.

- Conversion of measured data into images: low-resolution input images have a similar effect to quantization errors;
- Intrinsic class ambiguity in measured data: the presence of data that cannot be classified into a specific class;
- False detections: classes with similar shapes are falsely detected.

To address these issues, this study introduces a rule-based classification method that does not require data-to-image conversion, allows the investigation of class ambiguity, and is less sensitive to false detections. The rule-based classification is described in detail in the following section by defining the classes of data anomalies [26].

3. Rule-Based Classification

This study introduces a rule-based classification and compares it with the CNN-based classification to investigate the issues of intrinsic class ambiguity and information loss in data-to-image conversion. The overall process of the rule-based classification is shown in Figure 2. Each data anomaly class is defined by specific rules and algorithms that minimize the need for extensive data preprocessing, such as converting data into images. The measured time-history data were scrutinized for each class and classified into a single anomalous class upon detection. If no anomalous classes were found, the data were considered *Normal*. Note that the time-history data can have multiple classes of data anomalies, except for *Normal*. Herein, we define the data anomaly classes and provide rules for determining classes.

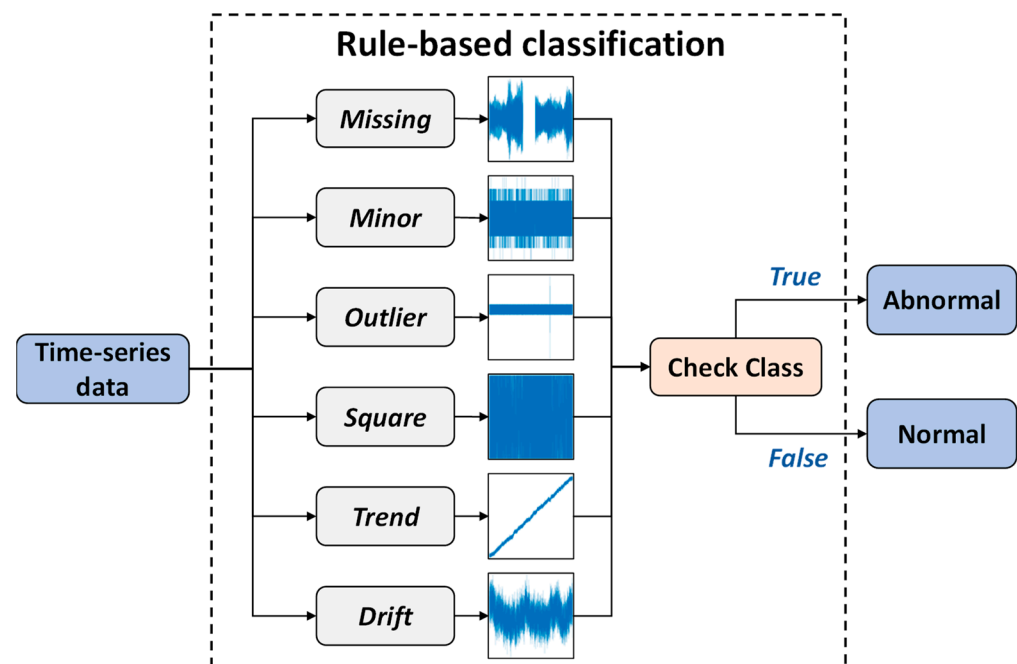


Figure 2. Flowchart of the rule-based classification for data anomalies.

3.1. Definitions of Data Anomaly Classes

This study considered the same anomaly classes used previously [26] for classifying the measured time-history data, with the rules defined for each class based on prior descriptions [26]. The definitions in this section involve the necessary threshold values that cannot have universal values. The specific values used in this study are presented in the next section, along with the actual measured data.

3.1.1. Missing Class

Missing is when the measured data are not complete. Examples of the *Missing* class are shown in Figure 3. A simple approach comparing the number of data points in the measured data with the desired length can be used to determine the *Missing* class.

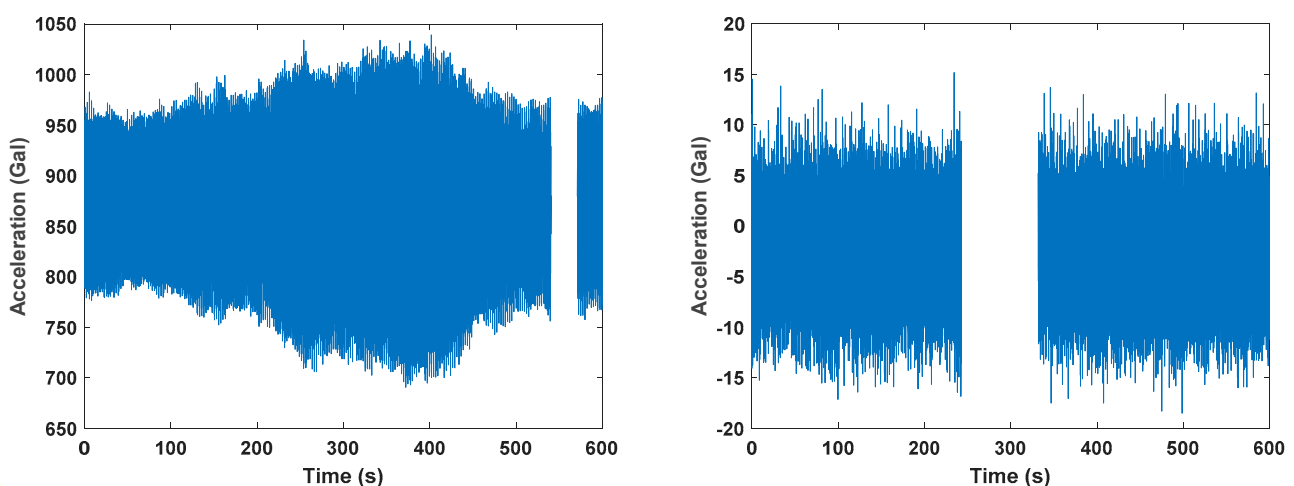


Figure 3. Examples of the *Missing* class.

3.1.2. Minor Class

The *Minor* class indicates data with clear quantization errors that typically occur when the response level is low compared with the resolution of the analog-to-digital converter (ADC), as shown in Figure 4. To detect the *Minor* class, the number of unique values

in the measured structural response was counted. If the number was less than a pre-defined threshold, the data were classified as a *Minor* class. The threshold can be selected considering that a 24-bit ADC is common in the market.

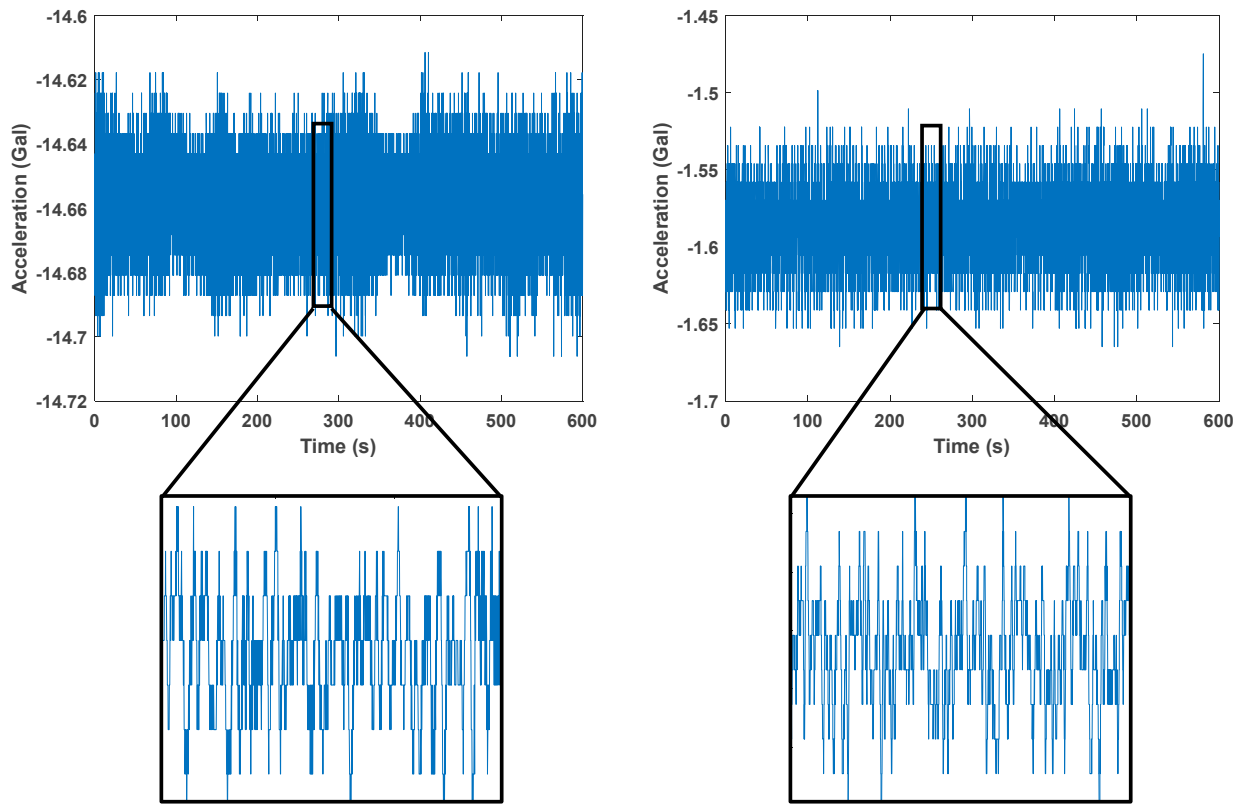


Figure 4. Examples of the *Minor* class.

3.1.3. Outlier Class

The identification of the *Outlier* class involves data points that deviate significantly from the rest of the measured data. Examples of the *Outlier* class are shown in Figure 5. A statistical approach can be used to determine the *Outlier* class. The initial step was to detrend the data to eliminate any underlying trends (e.g., constant, linear, or higher-order). Subsequently, all the peaks in the measured data were obtained. Three conditions were considered for classifying the time-history data to be classified as *Outlier*:

- The magnitude of the maximum peak ($\hat{\Downarrow}_p$) divided by the standard deviation of the time history (σ_t) is greater than a pre-defined threshold;

$$r_{outlier,1} = \hat{\Downarrow}_p / \sigma_t \quad (1)$$

- The magnitude of the maximum peak divided by the mean of the time history (μ_t) is greater than a pre-defined threshold;

$$r_{outlier,2} = \hat{\Downarrow}_p / \mu_t \quad (2)$$

- A data segment was selected right after the maximum peak. The ratio of standard deviations of the data segment (σ_d) and standard deviation of the raw time history (σ_t) is less than a pre-defined threshold.

$$r_{outlier,3} = \sigma_d / \sigma_t \quad (3)$$

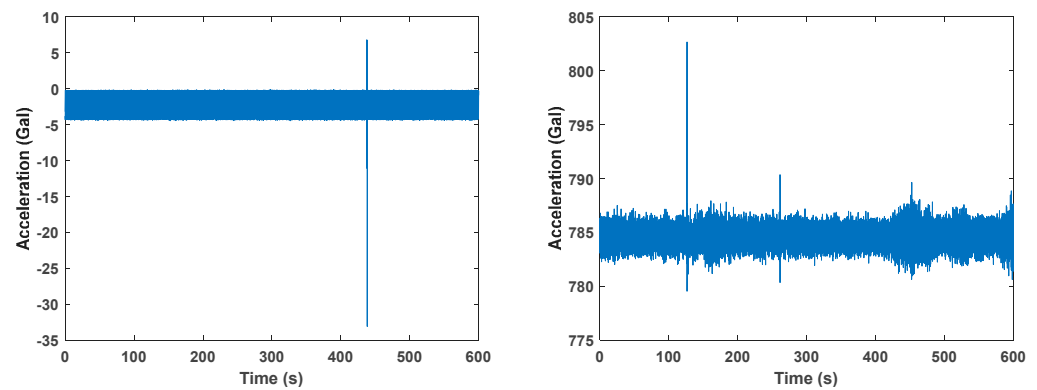


Figure 5. Examples of the *Outlier* class.

The third condition is necessary for differentiating the impulse response from an outlier in the time-history data.

3.1.4. Square Class

The *Square* class occurs when the vibration response oscillates abnormally over the measurement range. Thus, most data points were approximately at the maximum or minimum amplitudes, resulting in a square shape, as shown in Figure 6a. Consequently, the data histogram has two distinct peaks corresponding to the maximum and minimum amplitudes, as shown in Figure 6b. The characteristics of the *Square* class can be utilized for classification. Consider the following criteria for determining the *Square* class:

- The two maximum numbers of data points in the histogram (maximum bins) divided by the total number of data points (t_1) are greater than a pre-defined threshold;
- Bins with larger or smaller acceleration values than those of the maximum bins do not possess data points (t_2) that exceed a pre-defined threshold.

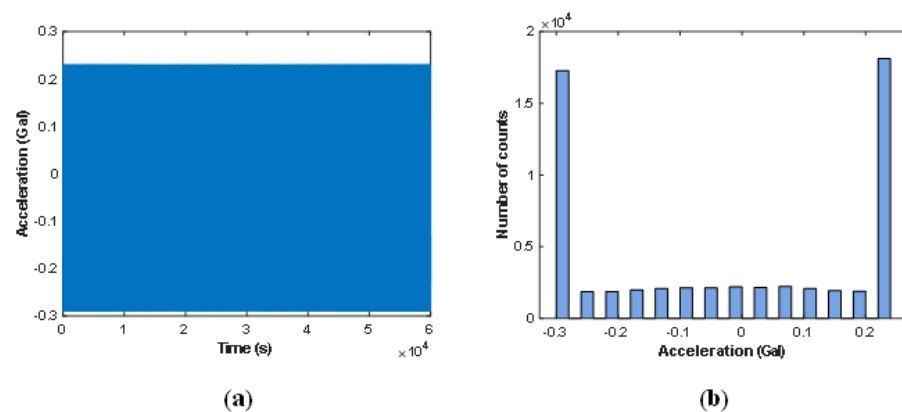


Figure 6. Example of the *Square* class: (a) time history and (b) histogram plot.

The second criterion was added to ensure the histogram has two distinct peaks on the far left and right sides, as shown in Figure 6b.

3.1.5. Trend Class

A vibration response of the *Trend* class is non-stationary, having a consistent and monotonous trend, as shown in Figure 7. The *Trend* class was determined by examining the slope and vibration amplitude of the signal. If the ratio of the slope to amplitude was greater than a pre-defined threshold, the signal can be classified as *Trend*. The standard deviation of the signal after linear detrending was used as the amplitude.

$$r_{trend} = \frac{s}{A} \quad (4)$$

where s and A denote the slope and vibration amplitude of the signal, respectively.

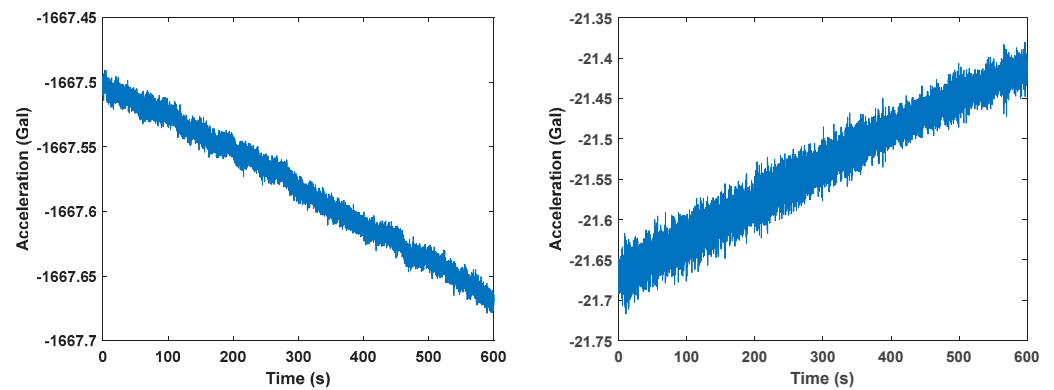


Figure 7. Examples of the *Trend* class.

3.1.6. Drift Class

The *Drift* class is characterized by nonstationary vibrations with random drift, as illustrated in Figure 8. To identify the *Drift* class, the measured data were partitioned into equally sized segments. Here, we define the following index to determine the *Drift* class:

$$d_{drift} = \frac{\sigma_{\mu_s}}{\mu_{\sigma_s}} \quad (5)$$

where σ_{μ_s} is the standard deviation of the mean of the segments and μ_{σ_s} is the mean of the standard deviation of the segment. If the index d_{drift} is greater than a pre-defined threshold, the signal is classified as *Drift*. σ_{μ_s} indicates the level of variation in the segmented signals. By dividing σ_{μ_s} by μ_{σ_s} which is the overall amplitude of the original signal, the index d_{drift} can be used as a measure of the random drift in a signal.

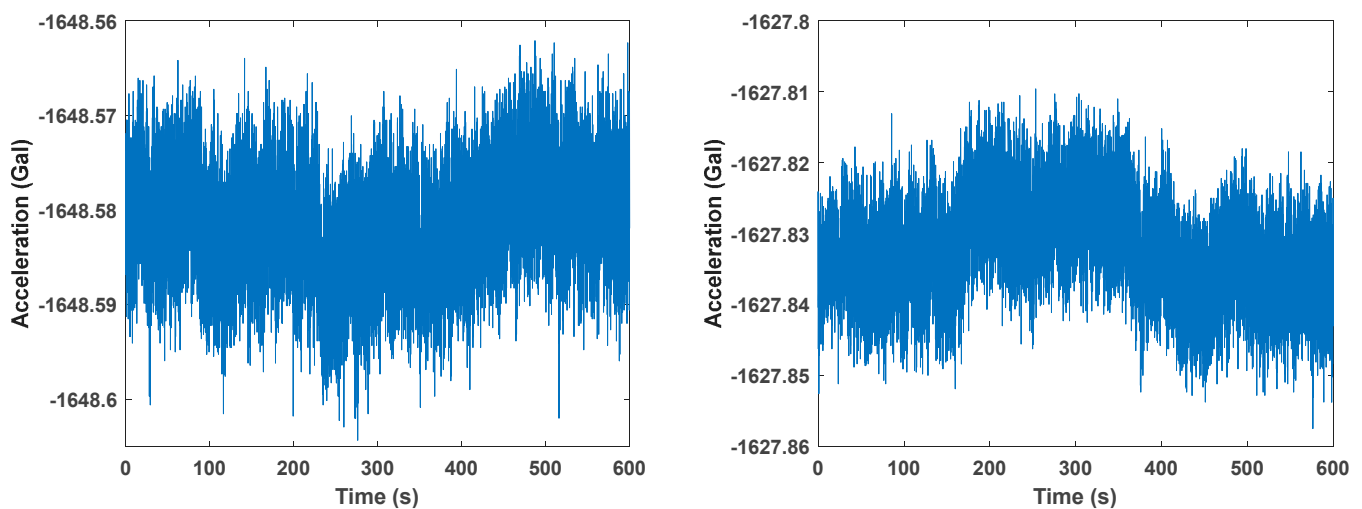


Figure 8. Examples of the *Drift* class.

3.1.7. Normal Class

If a measured signal is not classified into any of the anomaly classes, it is considered the *Normal* class.

This rule-based classification can be utilized to classify measured data into the above-defined classes. This classification approach is independent of data conversion into images and is less sensitive to false detections. The rule-based classification with actual measured data is discussed in detail in Section 4.

3.2. Issue in the Definitions of Anomaly Classes

Although proper definitions for classifying the measured data into various classes are provided in Section 3.1, ambiguity exists in the data anomaly classes for a certain dataset. For example, the *Trend* and *Drift* classes have intrinsic ambiguities because the threshold values for r_{trend} and d_{drift} cannot be selected perfectly, causing difficulties in labeling data for CNN-based classifications. Further discussion on this issue is provided in a later section.

4. Data Classification Results

4.1. Data Source

The measured time-history data from the Hwatae Bridge, a cable-stayed bridge located in the city of Yeosu, Republic of Korea, were used for the analysis of data anomaly detection. This bridge was constructed in 2015 with two 130-meter-high pylons and a 500-meter span between the two pylons. The data consisted of acceleration, displacement, and wind data. Only the acceleration data of the bridge were used for anomaly detection. To collect the acceleration data, 36 channels were placed at various parts of the bridge, such as piers, decks, pylons, and cables. The sensor layout of the 36 accelerometers on the bridge is shown in Figure 9. Detailed information on the positions of these channels is presented in Table 1. Acceleration data were recorded every 10 min at a sampling frequency of 100 Hz.

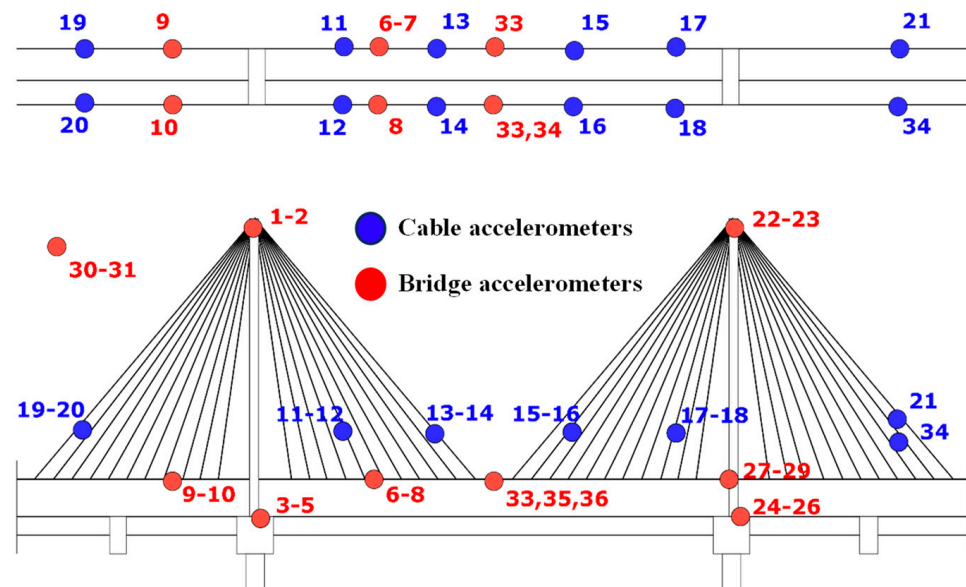


Figure 9. Schematic of accelerometers installed on the Hwatae bridge.

Table 1. Information about selected channels for anomaly detection.

Channels	Position
1, 2	Top of the left pylon
3–5	Foundation of the left pylon
6–8	Girder at $\frac{1}{4}$ span
9, 10	Girder at $\frac{1}{2}$ span
11–21, 34	Cables
22, 23	Top of the right pylon
24–26	Foundation of the right pylon
27–29	Girder end
30–32	Free field
33, 35, 36	Girder at $\frac{1}{2}$ span

Considering that the data for the *Minor* and *Square* classes do not exist in the measured data, the data for these two classes were artificially produced. The data for the *Minor* class

were generated by quantizing random signals of the original acceleration data. Similarly, the *Square* class was artificially prepared using the square wave phenomenon. A total of 5105 *Minor* and 5000 *Square* class samples were generated. The artificially generated *Minor* and *Square* class examples are shown in Figure 10.

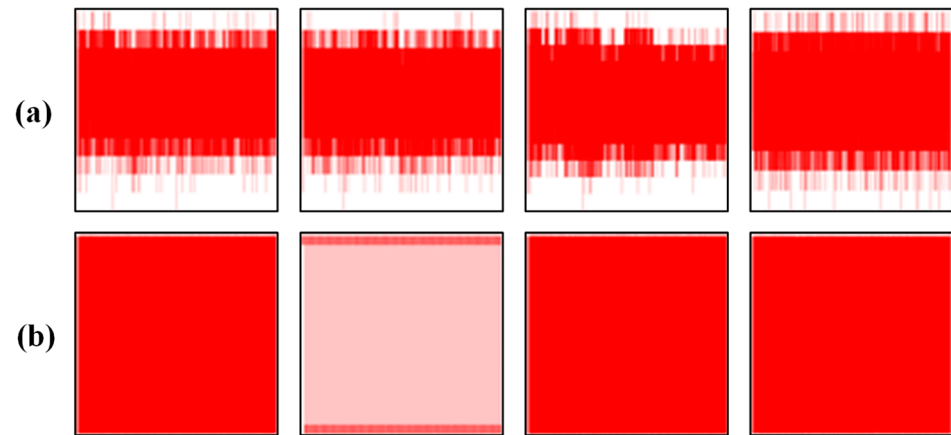


Figure 10. Examples of artificially generated classes: (a) *Minor* and (b) *Square*.

Rule- and CNN-based methods were applied to the data measured from the bridge to investigate their applicability and performance. First, a rule-based method was applied to the measured data, which allowed us to identify class ambiguity in the data. Based on these results, labeled data were prepared for the CNN-based method. Furthermore, the CNN-based classification results were obtained for the labeled dataset. Given that data labeling is conducted using the rule-based classification, in this study, we do not intend to compare the accuracy of the two methods but to investigate their characteristics and associated issues. An in-depth discussion of the classification results is presented in the next section.

4.2. Results of the Rule-Based Classification

The rule-based classification was applied to the acceleration data from the Hwatae Bridge along with added artificial data to investigate its applicability. The threshold values used in the rule-based classification are listed in Table 2. The selection of the specific values is discussed later in this section. Acceleration data were obtained from December 2020 to May 2021 for all 36 channels. During the classification, *Normal*, *Missing*, *Outlier*, *Trend*, and *Drift* classes were identified in the actual acceleration data. Examples of these classes representing 10 min acceleration data are shown in Figure 11.

Table 2. Threshold values for the data anomaly classes.

Class	Criteria	Threshold
<i>Missing</i>	Existence of a missing data point	Not applicable
<i>Minor</i>	Quantization lines in data	64
<i>Outlier</i>	$r_{outlier,1}$	30
	$r_{outlier,2}$	3
	$r_{outlier,3}$	3
<i>Square</i>	t_1	0.05
	t_2	0.05
<i>Trend</i>	r_{trend}	0.00035
<i>Drift</i>	d_{drift}	0.3

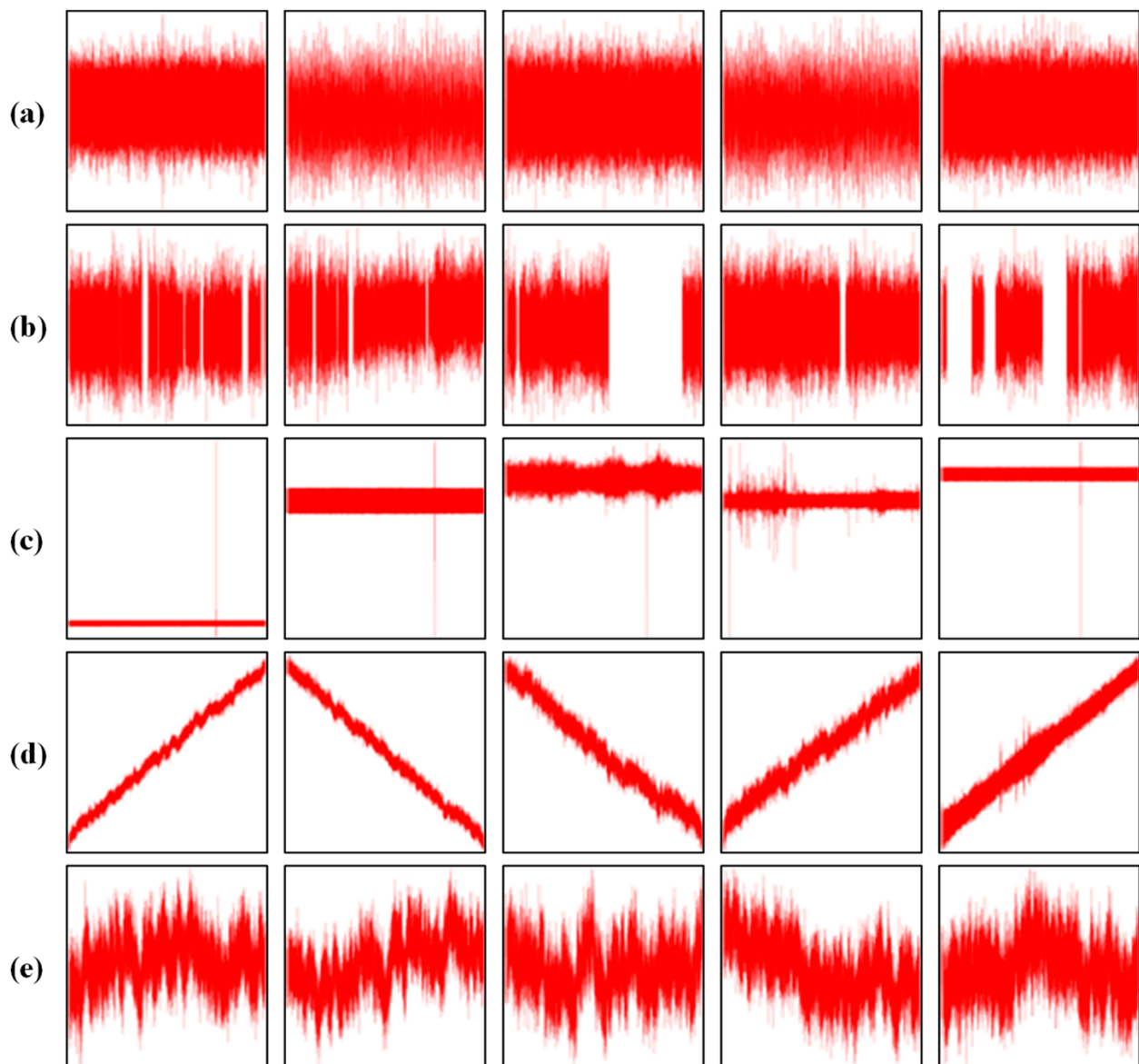


Figure 11. Example data for anomaly classes: (a) *Normal*, (b) *Missing*, (c) *Outlier*, (d) *Trend*, and (e) *Drift*.

A critical issue in applying a rule-based method is that the measured data have an intrinsic ambiguity in the anomaly classes. In particular, certain data are difficult to identify as the *Trend*, *Drift*, or *Normal* class because the data do not exhibit distinctive features for classification. To better describe the phenomenon, the index defined for the *Trend* class, r_{trend} , for all acceleration data is plotted in Figure 12a, showing an almost continuous distribution without clear patterns that can be used for classifying the *Trend* class. Thus, we selected two threshold values for r_{trend} to separate data clearly for both *Trend* and *Normal* classes, such that data above the red line (i.e., the first threshold of $r_{trend} = 0.00035$) and below the green line (i.e., the second threshold of $r_{trend} = 0.00005$) in Figure 12b, c are observed as obvious *Trend* and *Normal* classes, respectively. The data between the red and green lines are defined as *Trend-ambiguous*. A procedure similar to that for the d_{drift} distribution shown in Figure 13a was used for the *Drift* class. Time history examples for the *Drift* and *Normal* classes above and below the green lines are shown in Figures 13b and 13c, respectively. The data between the red and green lines correspond to *Drift-ambiguous*.

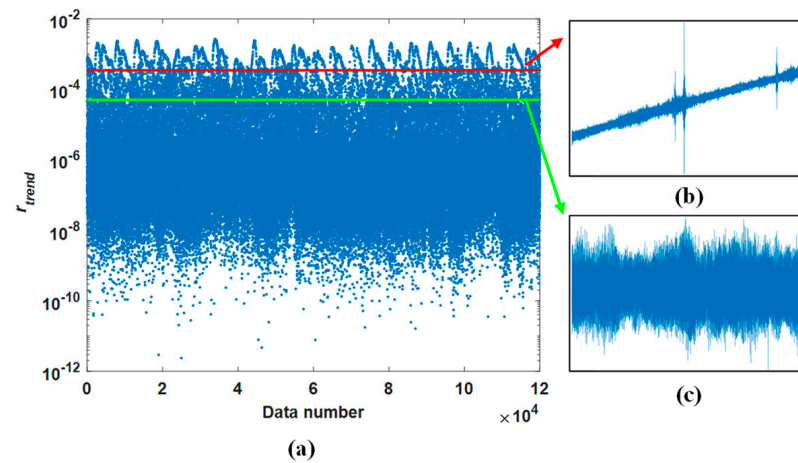


Figure 12. (a) *Trend* class classification index distribution, (b) time history example right above the red line, and (c) time history example right below the green line.

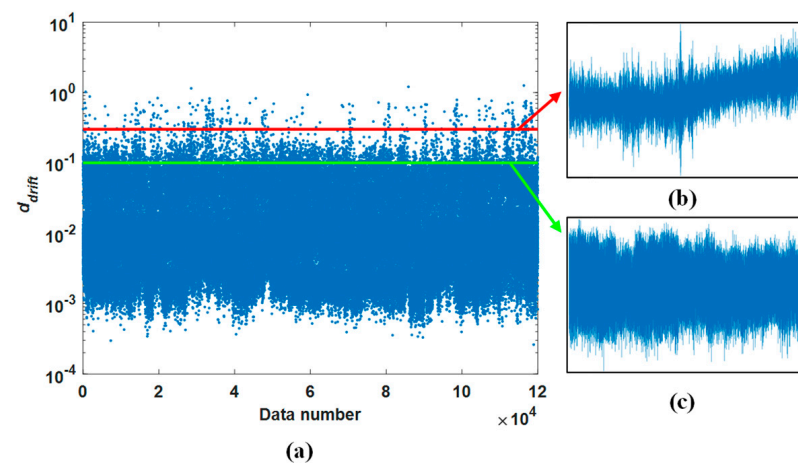


Figure 13. (a) *Drift* class classification index distribution. (b) Time history example right above the red line and (c) time history example right below the green line.

Examples of both *Trend-ambiguous* and *Drift-ambiguous* classes are shown in Figure 14. One can clearly see the difficulty of considering it as a perfect *Normal*, *Trend*, or *Drift* class. In total, 7216 *Trend-ambiguous* samples and 3489 *Drift-ambiguous* samples were selected from the measured acceleration data with pre-defined thresholds. Using rule-based classification, the total numbers of samples for each class and their proportions with respect to the total number of samples labeled in the database are listed in Table 3.

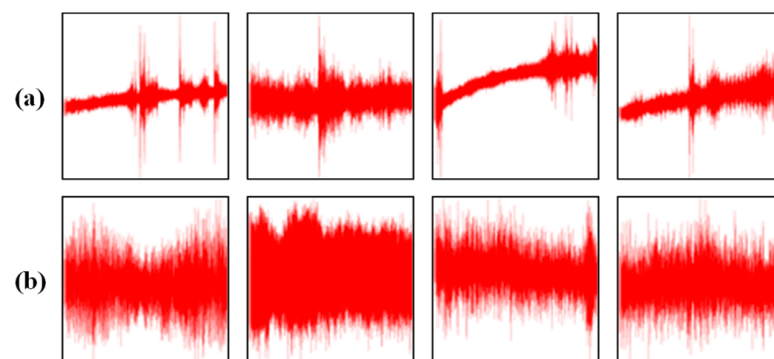


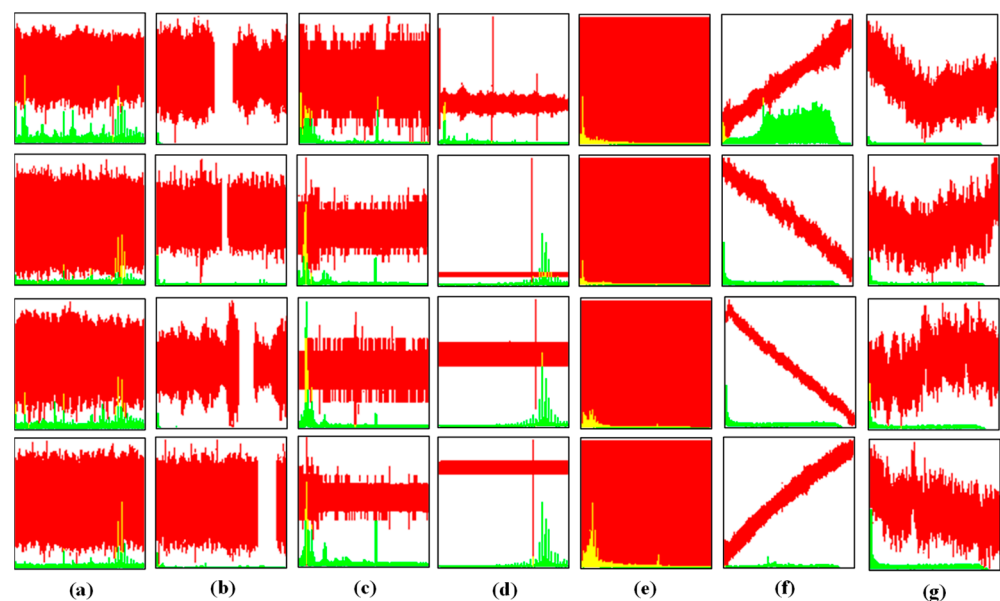
Figure 14. Examples of (a) *Trend-ambiguous* class and (b) *Drift-ambiguous* class.

Table 3. Total samples and percentage of each class anomaly in database.

Number	Class	Quantity	Ratio (%)
1	<i>Normal</i>	20,000	34.05
2	<i>Missing</i>	5105	8.69
3	<i>Minor</i>	5091	8.67
4	<i>Outlier</i>	4536	7.72
5	<i>Square</i>	5000	8.51
6	<i>Trend</i>	5883	10.01
7	<i>Drift</i>	2425	4.13
8	<i>Trend-ambiguous</i>	7216	12.28
9	<i>Drift-ambiguous</i>	3489	5.94
Total		58,745	100

4.3. Results of the CNN-Based Classification

A CNN-based classification method was applied to the prepared labeled datasets. The labeled data must be converted into images for the CNN-based classification. The converted dual-channel images contain time and frequency information in the red and green channels, respectively. Examples of the dual-channel images for all seven classes are shown in Figure 15.

**Figure 15.** Examples of dual-channel images: (a) *Normal*, (b) *Missing*, (c) *Minor*, (d) *Outlier*, (e) *Square*, (f) *Trend*, and (g) *Drift*.

The training dataset was prepared by randomly selecting data samples from the labeled dataset. *Trend-ambiguous* and *Drift-ambiguous* classes were not included in the training and test datasets to ensure consistency with a previous CNN-based classification study [26]. The details of both the training and test datasets are listed in Table 4. A balanced dataset with the same amount of data for each class was used to train the CNN model because an imbalanced dataset usually shows poor performance in training the network. A total of 1000 random samples from each class were selected from the database to train the CNN model. The training dataset for the CNN model was further divided into 70% training and 30% validation datasets. The remaining data samples from the database were used as datasets to test the accuracy of the trained CNN model.

Table 4. Training and test dataset information for the database.

Labels	1	2	3	4	5	6	7	Total
Class	<i>Normal</i>	<i>Missing</i>	<i>Minor</i>	<i>Outlier</i>	<i>Square</i>	<i>Trend</i>	<i>Drift</i>	
Quantity	20,000	5105	5091	4536	5000	5883	2425	48,040
Train	1000	1000	1000	1000	1000	1000	1000	7000
Test	19,000	4105	4091	3536	4000	4883	1425	41,040

The trained CNN model was then used to test for data anomalies in the remaining test dataset of the database. The results for the test dataset are shown in Figure 16. The overall accuracy of the test dataset was 93.3%. For the *Normal* and *Minor* classes, the recall values were 88.6% and 92.8%, respectively. Except for the *Normal* and *Minor* classes, all classes exhibited >95% recall values. The *Minor* and *Drift* classes showed 83.3% and 56.8% precision values, respectively, which were less than the precision values of the other classes. The results are similar to those of a previous study [26], proving that the CNN-based classification was appropriately implemented. Nonetheless, certain critical issues such as intrinsic ambiguity in the anomaly class and low resolution in the converted images must be investigated to determine the applicability of the data anomaly detection process. Further discussion of these issues is presented in the next section of this paper.

Confusion Matrix								
Actual \ Predicted	1	2	3	4	5	6	7	
	16828 41.0%	377 0.9%	747 1.8%	25 0.1%	6 0.0%	71 0.2%	946 2.3%	88.6% 11.4%
	102 0.2%	3946 9.6%	6 0.0%	5 0.0%	2 0.0%	11 0.0%	33 0.1%	96.1% 3.9%
	263 0.6%	8 0.0%	3797 9.3%	2 0.0%	0 0.0%	1 0.0%	20 0.0%	92.8% 7.2%
	11 0.0%	1 0.0%	6 0.0%	3504 8.5%	1 0.0%	5 0.0%	8 0.0%	99.1% 0.9%
	0 0.0%	3 0.0%	1 0.0%	0 0.0%	3996 9.7%	0 0.0%	0 0.0%	99.9% 0.1%
	8 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	4846 11.8%	29 0.1%	99.2% 0.8%
	30 0.1%	20 0.0%	1 0.0%	3 0.0%	0 0.0%	7 0.0%	1364 3.3%	95.7% 4.3%
	97.6% 2.4%	90.6% 9.4%	83.3% 16.7%	99.0% 1.0%	99.8% 0.2%	98.1% 1.9%	56.8% 43.2%	93.3% 6.7%

Figure 16. Confusion matrix for the unseen test dataset.

5. Issues in Data Anomaly Classification

5.1. Ambiguous Data Investigation

Ambiguous classes were investigated based on the results of the CNN-based classification. Data samples from both *Trend-ambiguous* and *Drift-ambiguous* were tested using the trained CNN model. Note that this analysis does not evaluate the performance of the CNN-based classification but discusses the effects of intrinsic ambiguity in the measured data. The classification results for *Trend-ambiguous* and *Drift-ambiguous* conditions are shown in Figures 17 and 18, respectively. In the case of *Trend-ambiguous*, the data were mostly classified into *Normal*, *Trend*, and *Drift* classes. Similarly, in the case of *Drift-ambiguous*, most of the data were classified into *Normal* and *Drift* classes, whereas few were classified into

Trend and *Minor* classes by the trained CNN. These ambiguous classes not only make data labeling difficult but also considerably degrade the classification accuracy of CNN-based methods. Class ambiguity also applies to the rule-based classification. Perfect thresholds that can distinguish ambiguous classes from *Normal* classes do not exist.

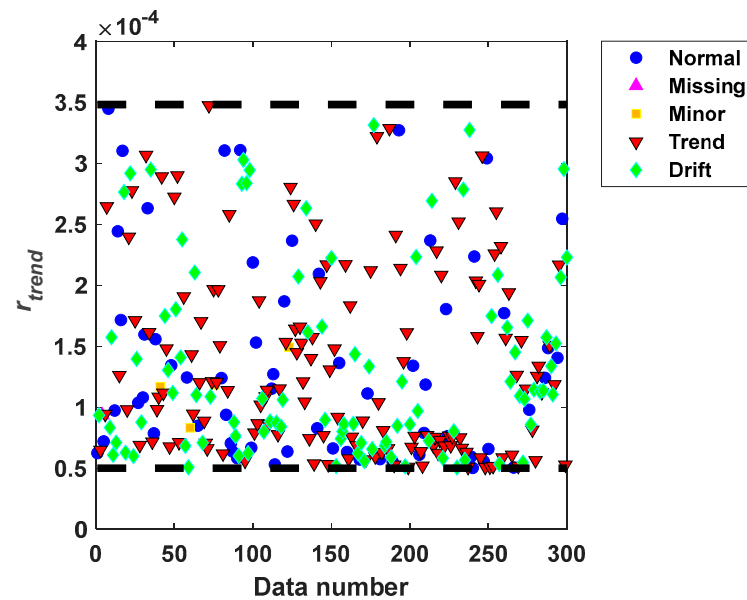


Figure 17. CNN-based classification results for *Trend-ambiguous* data samples.

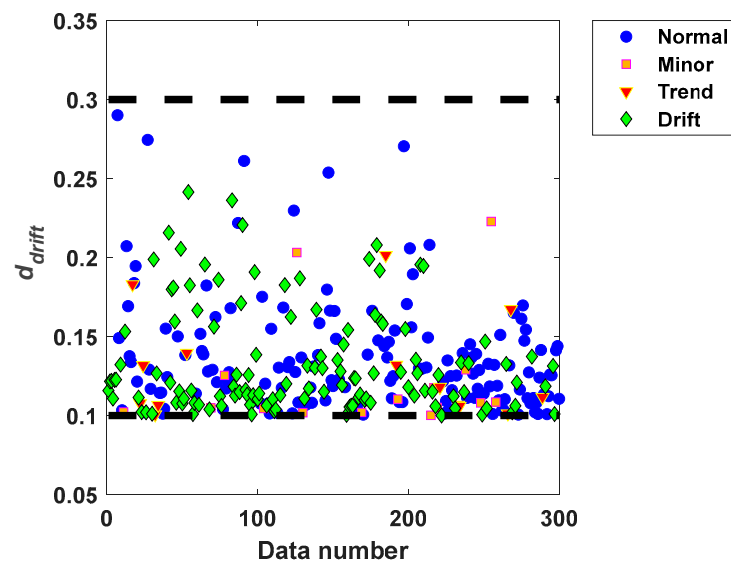


Figure 18. CNN-based classification results for *Drift-ambiguous* data samples.

5.2. Image-Resolution-Related Issue

The conversion of time-series data into images can impair data anomaly classification because the information in the data is partially lost during the conversion process. The converted image, which was used as the input of the CNN model, could not contain the full information of the measured dynamic signal. The image size used in this study was only 100×100 pixels, whereas the acceleration data contained 60,000 data points digitized with a 24-bit analog-to-digital converter (ADC). Indeed, lost information can lead to the misclassification of relevant data.

One of the adverse effects of image conversion is the loss of information on the amplitude axis, similar to quantization error. A vertical length of 100 pixels, which is less than 2^7 , is equivalent to applying a 6- or 7-bit ADC to the original acceleration signal.

The resulting images after the low-bit ADC are often seen as data of the *Minor* class, the main reason for which is the quantization phenomenon. Figure 19 shows examples of misclassification (i.e., data in the *Normal* class were classified as *Minor*).

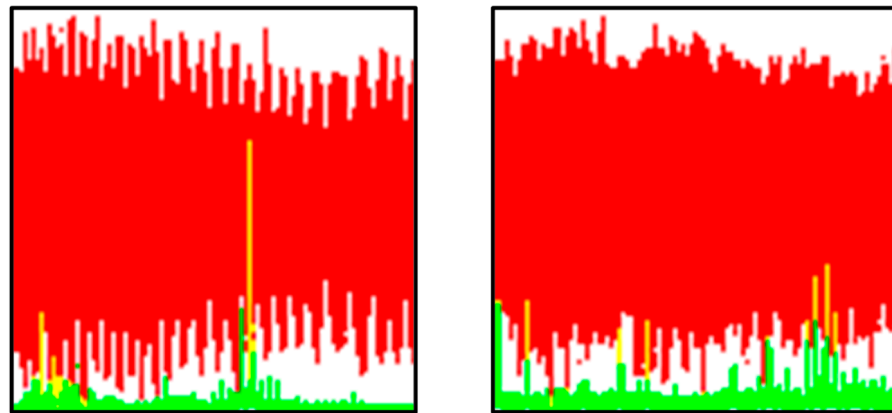


Figure 19. Misclassification of the *Normal* class into the *Minor* class.

Image conversion also causes information loss on the time axis, which behaves like resampling. In this dataset, one pixel corresponded to 600 data points. Thus, fewer than 600 missing data points could not be included in the converted image. For example, the images at the top of Figure 20 exhibit missing data, whereas the converted images shown at the bottom are classified as *Normal*.

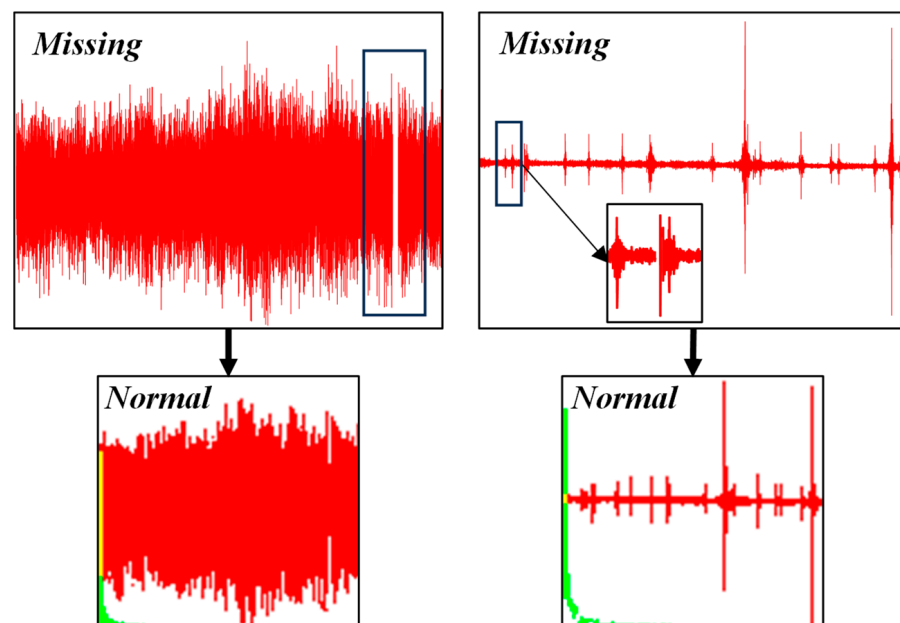


Figure 20. Misclassification of the *Missing* class due to low-resolution images.

5.3. Supplementary Issues

Other issues in the data anomaly detection process are discussed in this section. Although the accuracy of the CNN-based classification was high, false detections were observed, as shown in Figure 16. Because the converted images were used as inputs to the CNN, data with similar shapes can be misclassified, as shown in Figure 21. For example, an image of impulse responses belonging to the *Normal* class (see Figure 21a) can be classified as *Outlier* by the CNN-based method.

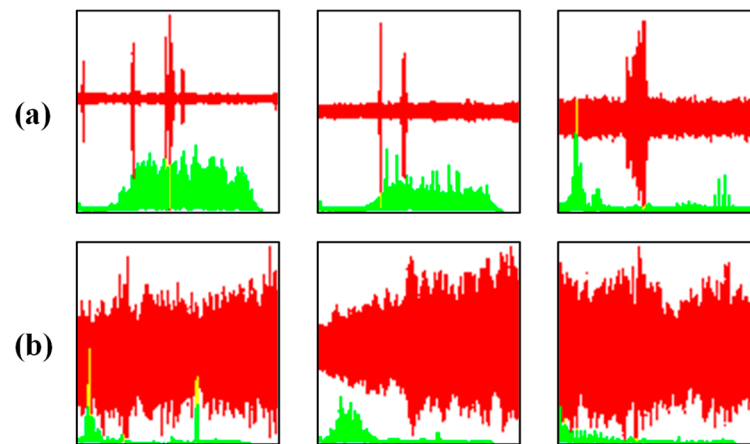


Figure 21. Misclassification of the *Normal* class into the (a) *Outlier* and (b) *Drift* classes.

One of the most salient drawbacks of the rule-based classification is that an appropriate selection of threshold values is necessary for each SHM dataset. Determining the threshold values defined in this study is a subjective task that cannot be generalized. In addition, as the threshold values are manually selected before the actual application, the automation of data anomaly detection becomes difficult.

6. Conclusions

The critical issues in the data anomaly detection were investigated with the rule-based and the CNN-based approaches. The rules for the seven classes of data anomaly used in [26] were defined for the rule-based classification. As the two classes *Trend* and *Drift* were not seen to be strictly separated from *Normal*, *Trend-ambiguous* and *Drift-ambiguous* were further defined. The acceleration data obtained from the Hwatae Bridge were labeled using the defined rules and were used to train the CNN model for classification. The following issues in the data anomaly classification were identified in the process of applying rule-based and CNN-based classifications:

- (1) Measured data have intrinsic ambiguity, as we defined two ambiguous classes. Thus, both determining a perfect rule that can separate the data without ambiguity and labeling such data for training the CNN model were seen to be impossible.
- (2) Important information can be lost during data conversion into images, as this conversion generates non-original data anomalies due to quantization and resampling phenomena.
- (3) The CNN model misclassified normal data as an incorrect anomaly class when the data had a feature in which a similar shape was shown in the anomaly class.
- (4) Rule-based classification requires definitions of data anomalies and the selection of the associated threshold values as prior knowledge. This requirement is the most important reason for developing the CNN-based classification, while the deep learning-based approach used to date also has clear drawbacks as discussed in this paper.

Both the rule-based and CNN-based classifications have advantages and disadvantages with respect to their applicability. The rule-based classification requires prior knowledge necessary to determine threshold values for the data anomalies sought. On other hand the CNN-based classification is free of such definitions of data anomalies; however, information loss due to the conversion of data into images can be an issue. Those characteristics of the rule-based and CNN-based methods discussed in this study result in differences in applicability, depending on target SHM system.

Author Contributions: Conceptualization, S.-H.S.; methodology, I.U.K. and S.-H.S.; software, I.U.K., S.J. and S.-H.S.; validation, I.U.K. and S.J.; investigation, S.J.; writing—original draft, I.U.K.; writing—review and editing, S.J. and S.-H.S.; supervision, S.-H.S.; funding acquisition, S.-H.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Sungkyunkwan University and the BK21 FOUR (Graduate School Innovation) funded by the Ministry of Education (MOE, Korea) and National Research Foundation of Korea (NRF), as well as by a grant (RS-2020-KA156887) from the Smart Construction Technology Development Program funded by the Ministry of Land, Infrastructure, and Transport of the Korean Government.

Data Availability Statement: The data presented in this study is available on request it from the Korea Authority of Land and Infrastructure Safety.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kralovec, C.; Schagerl, M. Review of structural health monitoring methods regarding a multi-sensor approach for damage assessment of metal and composite structures. *Sensors* **2020**, *20*, 826. [\[CrossRef\]](#)
2. Bao, Y.; Chen, Z.; Wei, S.; Xu, Y.; Tang, Z.; Li, H. The State of the Art of Data Science and Engineering in Structural Health Monitoring. *Engineering* **2019**, *5*, 234–242. [\[CrossRef\]](#)
3. Rocha, H.; Semprinoschnig, C.; Nunes, J.P. Sensors for process and structural health monitoring of aerospace composites: A review. *Eng. Struct.* **2021**, *237*, 112231. [\[CrossRef\]](#)
4. Zhang, F.; Norouzi, M.; Hunt, V.J.; Helmicki, A. Structural health monitoring system for Ironton-Russell Bridge, Ohio and Kentucky: Phase 1. Substructure construction. *Transp. Res. Rec.* **2015**, *2504*, 159–167. [\[CrossRef\]](#)
5. Comisu, C.C.; Taranu, N.; Boaca, G.; Scutaru, M.C. Structural health monitoring system of bridges. In *Procedia Engineering*; Elsevier Ltd.: Amsterdam, The Netherlands, 2017; pp. 2054–2059. [\[CrossRef\]](#)
6. Koo, K.Y.; Brownjohn, J.M.W.; List, D.I.; Cole, R. Structural health monitoring of the Tamar suspension bridge. *Struct. Control Health Monit.* **2013**, *20*, 609–625. [\[CrossRef\]](#)
7. Cho, S.; Yun, C.B.; Sim, S.H. Displacement estimation of bridge structures using data fusion of acceleration and strain measurement incorporating finite element model. *Smart Struct. Syst.* **2015**, *15*, 645–663. [\[CrossRef\]](#)
8. Neves, A.C.; González, I.; Leander, J.; Karoumi, R. Structural health monitoring of bridges: A model-free ANN-based approach to damage detection. *J. Civ. Struct. Health Monit.* **2017**, *7*, 689–702. [\[CrossRef\]](#)
9. Arjomandi, K.; Araki, Y.; MacDonald, T. Application of a hybrid structural health monitoring approach for condition assessment of cable-stayed bridges. *J. Civ. Struct. Health Monit.* **2019**, *9*, 217–231. [\[CrossRef\]](#)
10. Jeong, S.; Kim, H.; Lee, J.; Sim, S.H. Automated wireless monitoring system for cable tension forces using deep learning. *Struct. Health Monit.* **2021**, *20*, 1805–1821. [\[CrossRef\]](#)
11. Corbally, R.; Malekjafarian, A. Detecting changes in the structural behaviour of a laboratory bridge model using the contact-point response of a passing vehicle. *J. Struct. Integr. Maint.* **2023**, *8*, 226–238. [\[CrossRef\]](#)
12. McCrum, D.P.; Wang, S.; O'Brien, E.J. Monitoring the health of bridges using accelerations from a fleet of vehicles without knowing individual axle weights. *J. Struct. Integr. Maint.* **2023**, *8*, 249–259. [\[CrossRef\]](#)
13. Xu, B.; He, J.; Masri, S.F. Data-Based Model-Free Hysteretic Restoring Force and Mass Identification for Dynamic Systems. *Comput.-Aided Civ. Infrastruct. Eng.* **2015**, *30*, 2–18. [\[CrossRef\]](#)
14. Al-Amri, R.; Murugesan, R.K.; Man, M.; Abdulateef, A.F.; Al-Sharafi, M.A.; Alkahtani, A.A. A review of machine learning and deep learning techniques for anomaly detection in iot data. *Appl. Sci.* **2021**, *11*, 5320. [\[CrossRef\]](#)
15. Bao, Y.; Li, H. Machine learning paradigm for structural health monitoring. *Struct. Health Monit.* **2021**, *20*, 1353–1372. [\[CrossRef\]](#)
16. Ni, F.T.; Zhang, J.; Noori, M.N. Deep learning for data anomaly detection and data compression of a long-span suspension bridge. *Comput.-Aided Civ. Infrastruct. Eng.* **2020**, *35*, 685–700. [\[CrossRef\]](#)
17. Son, H.; Jang, Y.; Kim, S.E.; Kim, D.; Park, J.W. Deep Learning-Based Anomaly Detection to Classify Inaccurate Data and Damaged Condition of a Cable-Stayed Bridge. *IEEE Access* **2021**, *9*, 124549–124559. [\[CrossRef\]](#)
18. Kerschen, G.; De Boe, P.; Golinval, J.C.; Worden, K. Sensor validation using principal component analysis. *Smart Mater. Struct.* **2005**, *14*, 36–42. [\[CrossRef\]](#)
19. Gul, M.; Catbas, F.N. Statistical pattern recognition for Structural Health Monitoring using time series modeling: Theory and experimental verifications. *Mech. Syst. Signal Process* **2009**, *23*, 2192–2204. [\[CrossRef\]](#)
20. Abdelghani, M.; Friswell, M.I. Sensor validation for structural systems with additive sensor faults. *Struct. Health Monit.* **2004**, *3*, 265–275. [\[CrossRef\]](#)
21. Han, Z.; Zhao, J.; Leung, H.; Ma, K.F.; Wang, W. A Review of Deep Learning Models for Time Series Prediction. *IEEE Sens. J.* **2019**, *21*, 7833–7848. [\[CrossRef\]](#)
22. Kim, H.; Ahn, E.; Shin, M.; Sim, S.H. Crack and Noncrack Classification from Concrete Surface Images Using Machine Learning. *Struct. Health Monit.* **2019**, *18*, 725–738. [\[CrossRef\]](#)
23. Fawaz, H.I.; Forestier, G.; Weber, J.; Idoumghar, L.; Muller, P.A. Deep learning for time series classification: A review. *Data Min. Knowl. Discov.* **2019**, *33*, 917–963. [\[CrossRef\]](#)
24. Lee, K.; Jeong, S.; Sim, S.H.; Shin, D.H. A novelty detection approach for tendons of prestressed concrete bridges based on a convolutional autoencoder and acceleration data. *Sensors* **2019**, *19*, 1633. [\[CrossRef\]](#)

25. Bao, Y.; Tang, Z.; Li, H.; Zhang, Y. Computer vision and deep learning-based data anomaly detection method for structural health monitoring. *Struct. Health Monit.* **2019**, *18*, 401–421. [[CrossRef](#)]
26. Tang, Z.; Chen, Z.; Bao, Y.; Li, H. Convolutional neural network-based data anomaly detection method using multiple information for structural health monitoring. *Struct. Control Health Monit.* **2019**, *26*, e2296. [[CrossRef](#)]
27. Shajihan, S.A.V.; Wang, S.; Zhai, G.; Spencer, B.F. CNN based data anomaly detection using multi-channel imagery for structural health monitoring. *Smart Struct. Syst.* **2022**, *29*, 181–193. [[CrossRef](#)]
28. Mao, J.; Wang, H.; Spencer, B.F. Toward data anomaly detection for automated structural health monitoring: Exploiting generative adversarial nets and autoencoders. *Struct. Health Monit.* **2021**, *20*, 1609–1626. [[CrossRef](#)]
29. Arul, M.; Kareem, A. Data Anomaly Detection for Structural Health Monitoring of Bridges Using Shapelet Transform. *arXiv* **2020**, arXiv:2009.00470.
30. Zhang, Y.; Lei, Y. Data anomaly detection of bridge structures using convolutional neural network based on structural vibration signals. *Symmetry* **2021**, *13*, 1186. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.