Nuclear Engineering and Technology 55 (2023) 506–515

Contents lists available at ScienceDirect

Nuclear Engineering and Technology

journal homepage: www.elsevier.com/locate/net

Original Article

Proposal of a new method for learning of diesel generator sounds and detecting abnormal sounds using an unsupervised deep learning algorithm

Hweon-Ki Jo, Song-Hyun Kim, Chang-Lak Kim*

KEPCO International Nuclear Graduate School, 658-91 Heamaji-ro, Seoseong-myeon, Ulju-gun, Ulsan, 45014, South Korea

A R T I C L E I N F O

Article history:
Received 12 June 2022
Received in revised form
13 October 2022
Accepted 15 October 2022
Available online 21 October 2022

Keywords:
Unsupervised deep learning algorithm
Diesel generator sounds
Detecting abnormal sounds
Diagnosing the fault

A B S T R A C T

This study is to find a method to learn engine sound after the start-up of a diesel generator installed in nuclear power plant with an unsupervised deep learning algorithm (CNN autoencoder) and a new method to predict the failure of a diesel generator using it. In order to learn the sound of a diesel generator with a deep learning algorithm, sound data recorded before and after the start-up of two diesel generators was used. The sound data of 20 min and 2 h were cut into 7 s, and the split sound was converted into a spectrogram image. 1200 and 7200 spectrogram images were created from sound data of 20 min and 2 h, respectively. Using two different deep learning algorithms (CNN autoencoder and binary classification), it was investigated whether the diesel generator post-start sounds were learned as normal. It was possible to accurately determine the post-start sounds as normal and the pre-start sounds as abnormal. It was also confirmed that the deep learning algorithm could detect the virtual abnormal sounds created by mixing the unusual sounds with the post-start sounds. This study showed that the unsupervised anomaly detection algorithm has a good accuracy increased about 3% with comparing to the binary classification algorithm.

1. Introduction

The diesel generator installed in the nuclear power plant is automatically started in the event of an external power loss during reactor output operation, and functions to supply power to safety-related facilities to safely stop the reactor. In addition, during decommissioning of the nuclear power plant, it is automatically started in case of loss of external power, and it still plays an important role because it supplies power to the cooling facility for cooling the spent fuel. In all nuclear power plants, at least two diesel generators are installed per reactor unit, and in case of an emergency, even if one unit fails due to some unexpected cause, the remaining unit is started. Diesel generators are designed with a multi-defense concept to adequately respond to power plant emergencies. Since the diesel generator is an important facility not only during operation of a nuclear power plant but also during decommissioning, periodic diagnosis is essential.

Currently, the failure detection method for diesel generators in nuclear power plants was developed by the Korea Hydro & Nuclear Power Research Institute for about 7 years from 2006 to 2013. According to related references, the diesel engine condition diagnosis system consists of measurement sensors, data acquisition system, and main controller. The pressure sensor measures the explosion pressure inside the engine, and the vibration sensor and ultrasonic sensor monitor the operation status of the valves. The temperature sensor measures the temperature of the exhaust gas, and the tach sensor detects the measurement time of all measurement signals. The data acquisition system supplies power to the installed sensors, amplifies and filters the measured signals, and converts analog-to-digital signals. The main controller manages the test data received from the data acquisition system and the operating variables of each diesel engine and auxiliary systems as a database using the engine condition diagnosis program, and evaluates the status of the diesel engine and auxiliary systems based on these data [1–3].

Recently, AI technology is rapidly developing, and methods for detecting machine failures using deep learning algorithms based on sound are being actively studied. A commonly used method is to convert a sound signal recorded for each machine state into a spectrogram image and learn it with Convolution Neural Network

* Corresponding author.
E-mail address: clkim@kings.ac.kr (C.-L. Kim).

https://doi.org/10.1016/j.net.2022.10.019

(CNN) [4–6] or Deep Neural Network (DNN) [7] algorithm to diagnose failures with increasing accuracies between 2% and 30%. In addition, there is a method for diagnosing a failure using CNN [8,9], DNN [10], or Long Short-Term Memory (LSTM) [11] algorithm by converting the sound signal into a spectrum where the characteristics of the frequency domain can be easily seen. And there is a method that uses both sound and vibration signals as input signals and uses the DNN [12] or CNN [13] algorithm. Finally, there is also a way to use the CNN algorithm by receiving sound from several microphones and composing the input with multiple channels [14]. In this study, machine failure was diagnosed using commonly used spectrogram images and algorithms. However, there are limitations for directly applying these techniques into emergence diesel generators (EDGs) for nuclear power plants: 1) the diagnostic performance was not enough verified for high-powered diesel generators such as EDG; 2) the abnormal datasets for the machine learning are not enough in EDGs of nuclear power plants.

Therefore, in this study, the diagnostic performances of an unsupervised deep learning algorithm (CNN autoencoder) are evaluated and compared with the method of general binary classification by deep learning technique. First, the two sounds generated when the two large-powered diesel generators were started were prepared. By learning the sounds of two diesel generators with two deep learning algorithms, it was investigated whether the sound during operation (normal) and the sound before starting (abnormal) could be distinguished. In addition, five unusual sounds (pistol, cannon, siren, car crash, and explosion) were mixed with the diesel generator post-start sounds to create virtual abnormal sounds. It was confirmed whether these virtual abnormal sounds could be detected by a deep learning algorithm. These virtual abnormal sounds can be assumed to be abnormal engine sounds or abnormal sounds generated in the field while the diesel generator is running. It was checked whether these virtual abnormal sounds could be detected as abnormal with the learned deep learning algorithm.

2. Method

2.1. Overview of proposed diagnosis system

As shown in Fig. 1, diagnosis of the existing diesel generator uses sensors such as pressure sensor, vibration sensor, ultrasonic sensor, temperature sensor, and tach sensor to monitor the operating status of the diesel engine and auxiliary equipment, and to find faults using a diagnostic program. This method has the advantage of being able to detect component or system failure by observing whether the value of the operating variable detected by a specific sensor is out of the normal range or shows a specific waveform in the graph. However, this method has strong dependency on the sensors in the diesel generators in analyzing many driving variables and related graphs, and thus, some advanced technologies should be introduced for independently verifying the performance of EDG.

The diesel generator failure diagnosis method proposed in this paper is a method of diagnosing the failure by analyzing the sound generated during operation of the diesel generator with an unsupervised deep learning algorithm. As a deep learning algorithm, it uses an autoencoder algorithm, which is frequently used for failure detection in fault detections [15]. This approach stores the sound generated from before starting to after stopping the diesel generator through a microphone, converts it into a spectrogram image, uses it as input data, and learns with a deep learning algorithm. It was checked whether a failure was detected by inputting virtual abnormal sounds that could occur during operation of a diesel generator into the learned deep learning algorithm. This method is a new method that can detect failure independently of the existing failure detection method. Using this method, it is possible to detect or predict the failure of a diesel generator only with sound, and when used with the existing failure method, more precise failure diagnosis is possible.

2.2. Overview of deep learning algorithm

The CNN autoencoder consists of an encoder that compresses the input X to a low dimension and a decoder that restores it back to its original size as shown in Fig. 2. The autoencoder performs learning so that the difference between the input image X and the output image X' (restored X) becomes smaller, and in the process, the features of the input data are extracted. Autoencoder is an unsupervised learning method that performs learning only with samples corresponding to normal sounds. When normal samples are used as both the input and the output for training the autoencoder-based neural network, the difference between input and output is small, and when abnormal samples are input, the difference between input and output is large. Here, the difference between the input and the output is called the reconstruction error. After training the autoencoder, a threshold value is set to distinguish between normal and abnormal samples. If the reconstruction error of the input sample is greater than the threshold value, it is judged as abnormal sound [15].

The structure of the CNN autoencoder used in this study is shown as given in Table 1. As the input of the autoencoder, the original images are resized to (96, 96, 4) [16,17]. So the size of the

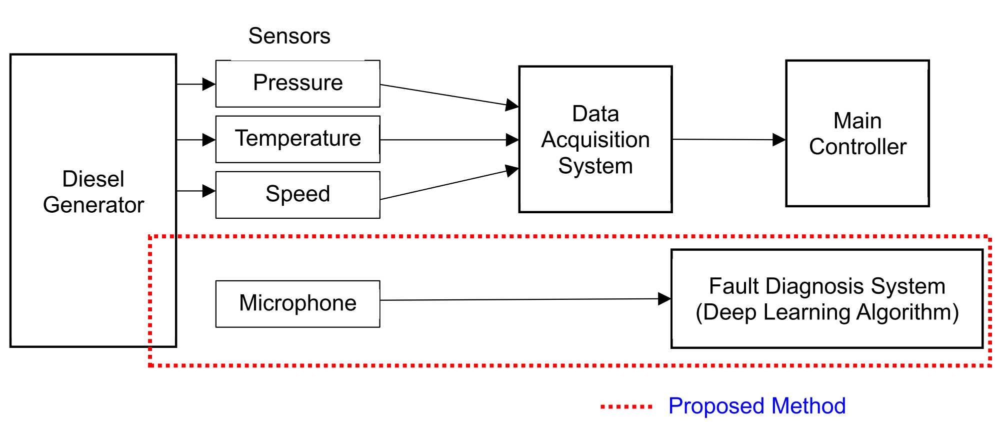


Fig. 1. Proposed fault diagnosis method of diesel generator.

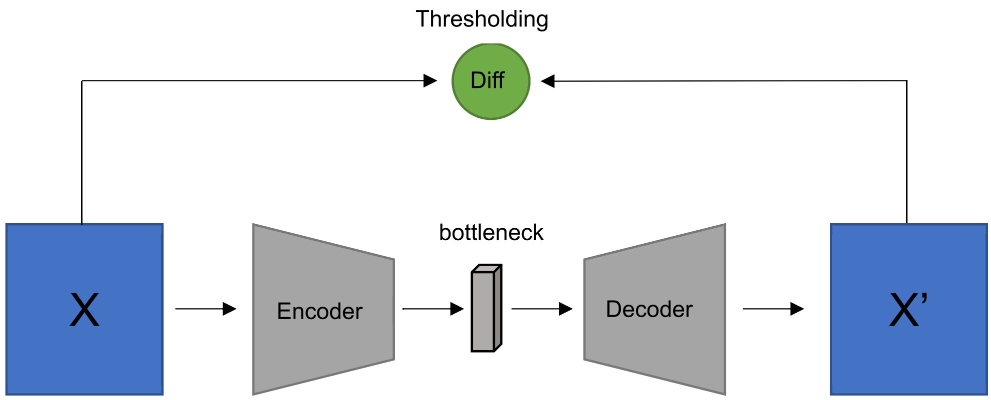


Fig. 2. Schematic of CNN autoencoder.

Table 1
Structure of CNN autoencoder.

Input	The image of the (96, 96, 4) shape
Encoder	Conv2D (16, (3, 3), padding = 'same', activation = 'relu', input_shape = (96, 96, 4)) MaxPooling2D (pool_size = (4, 4), padding = 'same') Conv2D (8, (3, 3), activation = 'relu', padding = 'same') MaxPooling2D (pool_size = (4, 4), padding = 'same') Conv2D (3, (3, 3), activation = 'relu', padding = 'same') MaxPooling2D (pool_size = (2, 2), padding = 'same')
Decoder	Conv2D (3, (3, 3), activation = 'relu', padding = 'same') UpSampling2D ((2, 2)) Conv2D (8, (3, 3), activation = 'relu', padding = 'same') UpSampling2D ((4, 4)) Conv2D (16, (3, 3), activation = 'relu', padding = 'same') UpSampling2D ((4, 4)) Conv2D(4, (3, 3), activation = 'sigmoid', padding = 'same')
Learning	optimizer = 'adam', loss = 'mean_squared_error'
Output	The reconstructed image of the (96, 96, 4) shape

original image is reduced to the size of the input image using the resize() function. The encoder that compresses the input image consists of three-step convolution and pooling layers, and compresses the input image of size (96, 96, 4) to size (3, 3, 3). The decoder consists of three-step convolution and upsampling layers, and restores data compressed to the size (3, 3, 3) to the original size (96, 96, 4) [16,17].

For the comparison of the proposed diagnostic method, as shown in Fig. 3, the CNN binary classification algorithm is selected as a supervised learning method that classifies two images with different features as 0 or 1, respectively. In this study, the abnormal sample is set to 0 and the normal sample is set to 1. Therefore, if the output of any sample input to the learned binary classification algorithm is lower than 0.5, it is classified as abnormal, and if it is greater than 0.5, it is classified as normal [18].

As shown in Table 2, the CNN binary classification algorithm extracts the features through three-step convolution and pooling layers, and then classifies two types of data through a fully connected layer. In the convolutional layer, 'relu' was used as the activation function, and in the fully connected layer, droupout(0.5) and two activation functions('relu' and 'sigmoid') were used. The size of the original image was changed to the size of (96, 96, 4) or (150, 150, 4) using the resize() function and used as an input [7].

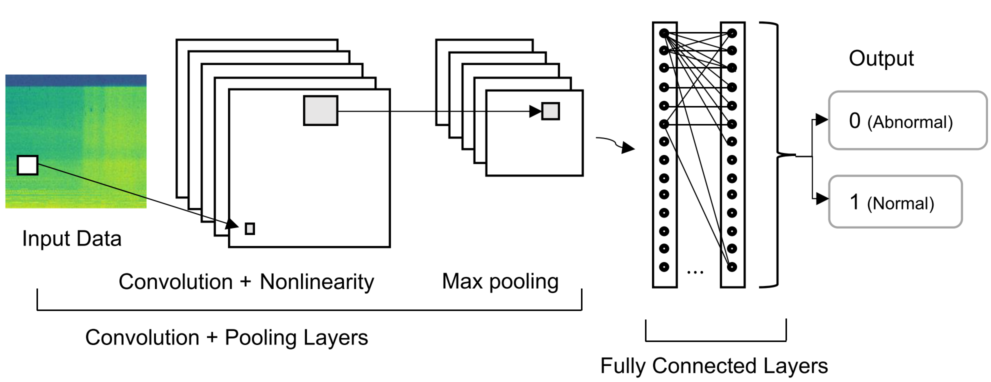


Fig. 3. Schematic of CNN binary classification.

Table 2
Structure of CNN binary classification.

Input	The image of the (96, 96, 4) shape
Body	Conv2D (32, (3, 3), input_shape = (96, 96, 4)) Activation('relu') MaxPooling2D (pool_size = (2, 2)) Conv2D (32, (3, 3)) Activation('relu') MaxPooling2D (pool_size = (2, 2)) Conv2D (64, (3, 3)) Activation('relu') MaxPooling2D (pool_size = (2, 2)) Flatten () Dense (64) Activation('relu') Dropout (0.5) Dense (1) Activation('sigmoid')
Learning	optimizer = 'adam', loss = 'binary_crossentropy', metrics = ['accuracy']
Output	0–1

2.3. Input data and preprocessing

2.3.1. Samples of sounds before and after the diesel generator start-up

To learn the sound of a diesel generator with a deep learning algorithm, samples were made by cutting the recorded sound into 7-s-long sounds as shown in Table 3. In this study, the sounds of two diesel generators were used, one was a 20-min recording and the other was a 2-h recording. Materials recorded in stereo format must be converted into mono format WAV files in advance. In the case of a 20-min sound, 7-s-long samples were made at 1-s intervals to make a total of 1200 samples (mono-type WAV files) as shown in Table 3. A total of 7200 samples (mono-type WAV files) were made for the 2-h sound. However, in order to learn with a deep learning algorithm, sound must be converted into an image. In this study, the method of converting sound into spectrogram image was used. Spectrogram is one of the methods of analyzing sound. The x-axis represents time, the y-axis represents frequency, and the z-axis represents amplitude (indicated by color) [3,4]. Here, the x and y axes are linear scales, and the z axes are log scales. In addition, when creating a spectrogram image, Fourier transform is performed. At this time, the window length is 20 msec, the overlap length is 5 msec, and the window shape is 'Hann'. The spectrogram image is saved as a PNG file with a size of 465x442 $\times$ 32b (sampling rate 44.1 kHz) or 466x481 $\times$ 32b (sampling rate 48 kHz). This study was conducted using Keras module in Windows 10(64-bit), Python ver.3.8.8(64-bit), and GPU environment (NVIDIA GeForce GTX 1660 Ti, CUDA ver.11.4, cuDNN ver.8.2.1).

2.3.2. Samples of virtual abnormal sounds

The sound samples after the diesel generator start-up and 5 unusual sounds were mixed to make 10 abnormal sound samples for each unusual sound. By mixing the normal sound with the unusual sound, it can be assumed that the engine sound is abnormal during the operation of the diesel generator or the abnormal sound occurs in the field. As shown in Table 4, new abnormal sound samples were created by mixing the normal sounds of diesel generators A and B and 5 unusual sounds.

3. Result

3.1. Training result of CNN autoencoder algorithm

In diesel generator A, out of 1200 samples, 600 for train, 150 for test, and the remaining 450 were used for verification after completion of training. In diesel generator B, 3600 of the total 7200 samples were used for train, 600 for test, and 2560 for verification. After the sensitivity studies that the cost functions are converged at eacg specific epoch, autoencoder was trained with batch size = 30, epoch = 20 in diesel generator A, and batch size = 50, epoch = 15 in generator B. As shown in Table 5, the largest reconstruction errors in generators A and B were 0.002078 and 0.001792, and the threshold values were set to 0.0021 and 0.0018, respectively, as values larger than those reconstruction errors. Therefore, normal and abnormal samples can be detected based on these threshold values. The reconstruction error of the normal sample is less than or equal to the threshold value, and the reconstruction error of the abnormal sample is greater than the threshold value.

- Normal sample (normal sound): Reconstruction error $\leq$ Threshold value
- Abnormal sample (abnormal sound): Reconstruction error > Threshold value

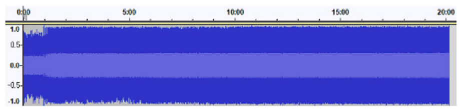
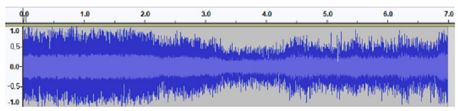
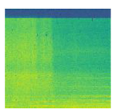
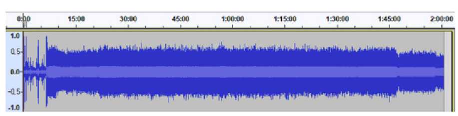
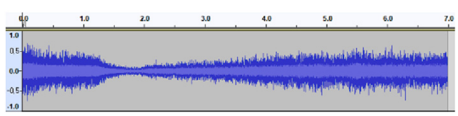
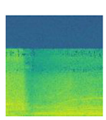
3.1.1. Verification for sounds before and after starting point

The starting point of diesel generator A is 58 s, and the starting point of diesel generator B is 420 s. Therefore, how the learned autoencoder judges the samples of the sounds before and after the starting point of the diesel generators A and B were recorded as shown in Table 6. In the case of diesel generator A, sample numbers 0 to 60 and 117 were determined to be abnormal, and the remaining samples were determined to be normal. Since the starting time is 58 s, it can be considered appropriate to judge the samples 59–60 as abnormal, but it can be seen as a singularity because the sample 117 was expected to be judged as normal. In the case of diesel generator B, samples 0 to 421 were determined to be abnormal, and the rest were determined to be normal. Since the starting time is 420 s, the results determined as abnormal up to sample 421 can be considered appropriate.

3.1.2. Verification for sounds after start-up

Table 7 shows how the learned autoencoder judges the unlearned sounds after the diesel generator startup. In the case of diesel generator A, since the reconstruction error was lower than the threshold value for 210 samples corresponding to the sound after start-up, all were determined to be normal. In the case of diesel generator B, as shown in Table 8, the reconstruction error was higher than the threshold value for 6 out of 2560 samples of the sound after start-up, so it was determined to be abnormal.

Table 3
Input data generation for diesel generator sound learning.

	Waveform or Spectrogram	Quantity	File Type	Length or Size
Diesel Generator A				
Overall Sound		1	Mono-WAV	20 min
Split Sound		1200	Mono-WAV	7 s
Input Image		1200	PNG	465 × 442 × 32b
Diesel Generator B				
Overall Sound		1	Mono-WAV	20 min
Split Sound		7200	Mono-WAV	7 s
Input Image		7200	PNG	466 × 481 × 32b

3.2. *Training result of CNN binary classification algorithm*

In order to learn with the binary classification algorithm, it is also necessary to learn samples for abnormal sounds. Therefore, in diesel generator A, out of 1200 samples, 630 samples for train (600 normal samples and 30 abnormal samples), 210 samples for test (200 normal samples and 10 abnormal samples), and 360 samples were used for verification. In diesel generator B, 3900 samples (3600 normal samples and 300 abnormal samples) for train out of 7200 samples, 650 samples (600 normal samples and 50 abnormal samples) for test, and 2650 samples were used for verification. The binary classification algorithm was trained with batch size = 30, epoch = 20 in diesel generator A, and batch size = 50, epoch = 15 in generator B.

In the binary classification algorithm, abnormal samples were trained as 0 and normal samples as 1. Therefore, if the output of the input sample was less than 0.5, it was determined as abnormal, and if it was greater than or equal to 0.5, it was determined as normal. How the CNN binary classification algorithm judges the samples used for training is shown in Table 9. In the case of diesel generator A, it was confirmed that most of the 840 samples used for train and test were accurately determined as normal or abnormal. However, only sample 117 was erroneously determined to be abnormal. In the case of diesel generator B, it was confirmed that 4550 samples were accurately determined as normal or abnormal.

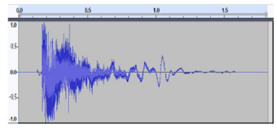
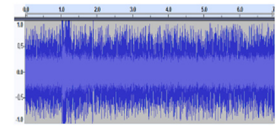
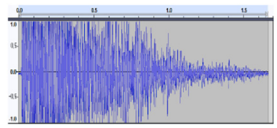
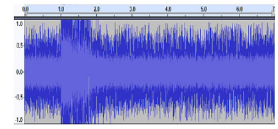
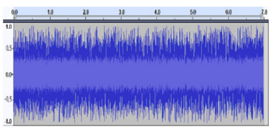
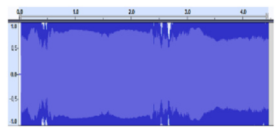
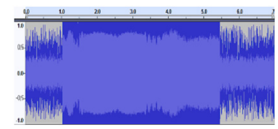
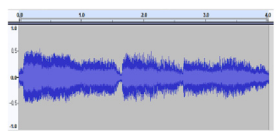
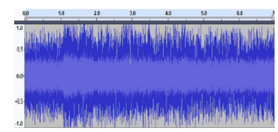
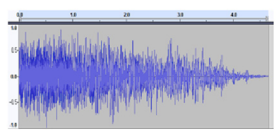
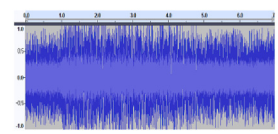
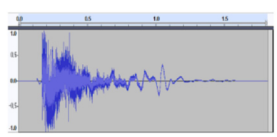
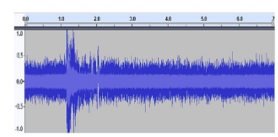
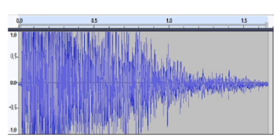
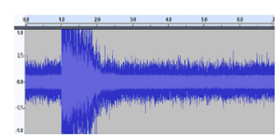
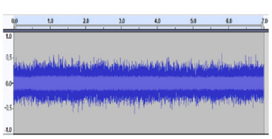
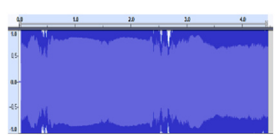
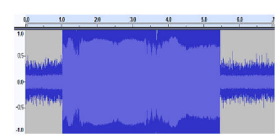
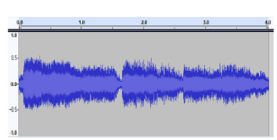
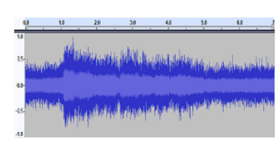
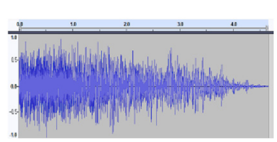
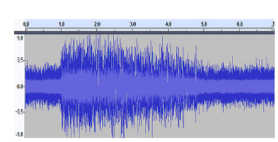
3.2.1. *Verification for sounds before and after starting point*

How the learned binary classification algorithm judges the sounds before and after the diesel generator start-up was confirmed as shown in Table 10. The starting point of diesel generator A was 58 s, and it was confirmed that it was determined as normal from sample 52. However, the autoencoder determined to be normal from sample 61. It was confirmed that the starting point of diesel generator B was 420 s, and it was determined as normal from sample 420. In this case, the autoencoder determined that it was normal from sample 422.

3.2.2. *Verification for sounds after start-up*

It was confirmed how the learned binary classification algorithm judges the unlearned sounds after the start-up of the diesel generator. As shown in Table 11, in the case of diesel generator A, 320 normal samples were accurately determined as normal. In the

Table 4
Virtual abnormal sounds for diesel generator A and B sound.

Normal Sound	Unusual Sound		Virtual Sound
Diesel Generator A			
		Pistol	
		Cannon	
		Siren	
		Car Crash	
		Explosion	
Diesel Generator B			
		Pistol	
		Cannon	
		Siren	
		Car Crash	
		Explosion	

case of diesel generator B, it was confirmed that only 14 samples out of 2560 samples were incorrectly judged as abnormal. And the output values of 14 samples that were judged incorrectly are shown in Table 12. Here, the peculiar thing is that the 14 samples judged erroneously by this algorithm include all 6 samples erroneously judged by the autoencoder.

Table 5
Results of samples for train & test and settings of threshold value.

Diesel Generator	Category	Input		Output (Recon. Error)		Threshold Value
		Sample No.	Quantity	Min	Max	
A	Train	240–839	600	0.000936	0.002078	0.0021
	Test	900–1049	150	0.000996	0.001382	
B	Train	440–4039	3600	0.000781	0.001792	0.0018
	Test	4040–4239 5000–5199 6000–6199	600	0.000789	0.001404	

Table 6
Results of samples before and after the diesel generator start-up.

Diesel Generator A					Diesel Generator B				
Input			Output		Input			Output	
Sample No.	Start (mm:ss)	End (mm:ss)	Recon. Error	Result[a]	Sample No.	Start (mm:ss)	End (mm:ss)	Recon. Error	Result[a]
0	00:00	00:07	0.009501	F	0	00:00	00:07	0.008694	F
1	00:01	00:08	0.003808	F	1	00:01	00:08	0.007535	F
2	00:02	00:09	0.005388	F	2	00:02	00:09	0.006767	F
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
57	00:57	01:04	0.010120	F	415	06:55	07:02	0.008301	F
58	00:58	01:05	0.009177	F	416	06:56	07:03	0.007876	F
59	00:59	01:06	0.006413	F	417	06:57	07:04	0.006343	F
60	00:60	01:07	0.002901	F	418	06:58	07:05	0.004614	F
61	00:61	01:08	0.001551	T	419	06:59	07:06	0.003460	F
62	00:62	01:09	0.001477	T	420	07:00	07:07	0.002629	F
63	00:63	01:10	0.001867	T	421	07:01	07:08	0.002056	F
⋮	⋮	⋮	⋮	⋮	422	07:02	07:09	0.001517	T
116	01:56	02:03	0.001641	T	423	07:03	07:10	0.000961	T
117	01:57	02:04	0.002345	F	424	07:04	07:11	0.000899	T
118	01:58	02:05	0.001436	T	425	07:05	07:12	0.000948	T
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
239	03:59	04:06	0.001355	T	439	07:20	07:27	0.000981	T

[a] Result: T – Normal, F – Abnormal.

Table 7
Results of samples after the diesel generator start-up.

Diesel Generator	Input		Output (Reconstruction Error)	
	Sample No.	Quantity	Min	Max
A	840–899 1050–1199	210	0.000988	0.001590
B	4240–4999 5200–5999 6200–7199	2560	0.0007823	0.001772[a]

[a] Maximum value excluding 6 anomalous data.

3.3. Performance estimation of the anomaly detection algorithms

For verifying and comparing the performances of the algorithms, virtual abnormal sounds were first created by mixing a normal sound with an unusual sound, and investigated whether it can be detected when an abnormal sound occurs in the engine or when an unusual sound occurs in the field while driving a diesel generator. For example, 10 virtual abnormal sound samples were created by mixing the siren sound with 10 normal sound samples, and in the case of diesel generator B, the autoencoder and binary classification algorithm correctly determined all 10 as abnormal as shown in Table 13. Overall, in the case of diesel generator A, the average accuracy of detecting these virtual abnormal sounds was low, around 50%. On the other hand, diesel generator B showed high accuracy of over 90%. In the case of diesel generator B, the accuracy was slightly lower only for the sound of a car crash, and 100% of the other abnormal sounds were detected. Analysis showed that the sound for the diesel generator A relatively had large noises comparing to the sound of pistol, car crash and explosion;

Table 8
Falsely judged samples by CNN autoencoder.

Diesel Generator B				
Input			Output	
Sample No.	Start (mm:ss)	End (mm:ss)	Recon. Error	Result[a]
7117	01:58:37	01:58:44	0.001842	F
7118	01:58:38	01:58:45	0.001843	F
7119	01:58:39	01:58:46	0.001858	F
7121	01:58:41	01:58:48	0.001850	F
7122	01:58:42	01:58:49	0.001851	F
7152	01:59:12	01:59:19	0.001874	F

[a] Result: T – Normal, F – Abnormal.

Table 9
Results of samples for train & test (CNN binary classification).

Diesel Generator	Category		Input		Output	
			Sample No.	Quantity	Min	Max
A	Train	Abnormal	0–29	30	0.000065	0.022862
		Normal	80–679	600	0.972010[a]	0.999999
	Test	Abnormal	30–39	10	0.004263	0.038201
		Normal	700–739 800–839 900–939 1000–1039 1100–1139	200	0.999989	1.0
B	Train	Abnormal	0–299	300	7.243176e-09	2.401789e-05
		Normal	440–4039	3600	0.970257[b]	1.0
	Test	Abnormal	300–349	50	5.2437e-09	2.743405e-07
		Normal	4040–4239 5000–5199 6000–6199	600	1.0	1.0

[a] Disel Generator A: sample117.png = 0.4527759.
[b] Disel Generator B: sample551.png = 0.873443, sample553.png = 0.827125.

Table 10
Results of samples before and after the diesel generator start-up (CNN binary Class.).

Diesel Generator A					Diesel Generator B				
Input			Output		Input			Output	
Sample No.	Start (mm:ss)	End (mm:ss)	Recon. Error	Result[a]	Sample No.	Start (mm:ss)	End (mm:ss)	Recon. Error	Result[a]
40	00:40	00:47	0.014404	F	350	03:50	03:57	1.876e-08	F
⋮	⋮	⋮	⋮	⋮	351	03:51	03:58	1.264e-08	F
49	00:49	00:56	0.010443	F	352	03:52	03:59	1.077e-08	F
50	00:50	00:57	0.033435	F	⋮	⋮	⋮	⋮	⋮
51	00:51	00:58	0.033540	F	415	06:55	07:02	6.645e-07	F
52	00:52	00:59	0.905535	T	416	06:56	07:03	9.433e-07	F
53	00:53	01:00	0.999721	T	417	06:57	07:04	6.669e-06	F
54	00:54	01:01	0.999996	T	418	06:58	07:05	4.932e-05	F
55	00:55	01:02	1.0	T	419	06:59	07:06	0.000893	F
56	00:56	01:03	1.0	T	420	07:00	07:07	0.999997	T
57	00:57	01:04	1.0	T	421	07:01	07:08	0.943151	T
58	00:58	01:05	1.0	T	422	07:02	07:09	0.992148	T
59	00:59	01:06	1.0	T	423	07:03	07:10	0.999998	T
60	01:00	01:07	0.999996	T	424	07:04	07:11	0.999989	T
61	01:01	01:08	0.999985	T	425	07:05	07:12	0.999913	T
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
79	01:19	01:26	0.999623	T	439	07:20	07:27	1.0	T

[a] Result: T – Normal (≧0.5), F – Abnormal (<0.5).

Table 11
Results of samples after the diesel generator start-up.

Diesel Generator	Input		Output (0–1)	
	Sample No.	Quantity	Min	Max
A	680–699, 740–799 840–899, 940–999 1040–1099, 1140–1199	320	0.999979	1.0
B	4240–4999, 5200–5999, 6200–7199	2560	0.999599[a]	1.0

[a] Maximum value excluding 14 anomalous data.

therefore, sound-based diagnosis system should be well tested before the applications.

Also, the accuracy of 1200 samples for diesel generator A and 7200 samples for diesel generator B with considering the original data sizes recorded was compared as shown in Table 14 by autoencoder and binary classification algorithm. Here, it can be seen that from which sample is determined as normal near the starting point of a diesel generator is slightly different depending on the characteristics of the algorithm. A conservative approach was taken to calculate the accuracy. In the case of diesel generator A, the starting time is 58 s, based on the determination of normal from sample 58, and in the case of diesel generator B, the starting time is 420 s, and based on the determination of normal from sample 420 was set. When this standard is applied, the accuracy of the autoencoder of diesel generator A is 99.7% and the accuracy of the binary classification algorithm is 99.4%. In addition, the accuracy of the autoencoder of diesel generator B is 99.9% and the accuracy of the binary classification algorithm is 99.8%. The results showed that both algorithms predict well within 0.6% for distinguishing the sounds of starting and normal points; therefore, it is

Table 12
Falsely judged samples by CNN binary classification.

Diesel Generator B				
Input			Output	
Sample No.	Start (mm:ss)	End (mm:ss)	Value (0–1)	Result[a]
7116	01:58:36	01:58:43	4.849927e-05	F
7117	01:58:37	01:58:44	5.899933e-05	F
7118	01:58:38	01:58:45	1.624109e-05	F
7119	01:58:39	01:58:46	4.120367e-05	F
7120	01:58:40	01:58:47	0.000174	F
7121	01:58:41	01:58:48	1.679705e-05	F
7122	01:58:42	01:58:49	2.941326e-05	F
7148	01:59:08	01:59:15	0.002490	F
7149	01:59:09	01:59:16	0.001557	F
7150	01:59:10	01:59:17	0.047798	F
7151	01:59:11	01:59:18	0.000534	F
7152	01:59:12	01:59:19	0.000215	F
7153	01:59:13	01:59:20	0.061836	F
7154	01:59:14	01:59:21	0.000662	F

[a] Result: T – Normal (≧0.5), F – Abnormal (<0.5).

Table 13
Accuracy for virtual abnormal sound samples.

No.	Category	Diesel Generator A		Diesel Generator B	
		Autoencoder	Binary Class.	Autoencoder	Binary Class.
1	Pistol	30%	30%	100%	100%
2	Cannon	80%	70%	100%	100%
3	Siren	100%	70%	100%	100%
4	Car Crash	30%	30%	70%	80%
5	Explosion	40%	40%	100%	100%
Average Accuracy		56%	48%	94%	96%

expected that these algorithm can be utilized for checking the normal operation.

4. Conclusion

In this study, the autoencoder-based anomaly detection algorithm was introduced and compared with the binary classification method for verifying the diagnostic performance of sounds generated by diesel generators in nuclear power plant. The sounds of the diesel generator were recorded and converted for the machine learning, and the performances of the two algorithms were compared using sound samples which were not used in the machine learning. Based on the starting point, the samples for the sound before the start were judged as abnormal, and the samples for the sound after the start were almost exactly determined as normal. However, since one sample is 7 s long, samples near the starting point contain both normal and abnormal sound characteristics, so it can be seen that some errors occur. In addition, it was also confirmed that most of the remaining normal samples were determined to be normal. As a result, by learning the sound before and after start-up of the diesel generator with a deep learning algorithm, it was possible to accurately distinguish the sound after start-up as normal and the sound before start-up as abnormal.

Especially, it was confirmed that, using the virtual abnormal sound samples, it was possible to detect abnormal sounds of the engine during operation of the diesel generator or abnormal sounds generated in the surroundings while the engine was started. In particular, diesel generator A had a low detection accuracy of 50% for virtual abnormal samples, but diesel generator B showed a high accuracy of 90% or more. Using the results of this study, it is possible to predict engine failures by detecting engine abnormalities during operation of diesel generators in the field or by continuously monitoring the performance of diesel generators.

Additionally, the autoencoder sets a threshold value as a criterion for discriminating between normal and abnormal sounds, and the sensitivity of detecting abnormal sounds changes depending on how this value is set. That is, if the threshold value is set low, even a slight deviation from the normal sound can be detected as abnormal. Autoencoders are useful for detecting abnormal sounds that deviate from normal. Binary classification algorithm can find abnormal sounds by learning from samples of normal and abnormal sounds. However, as in the case of this diesel generator, when there is little data on abnormal sounds, there is a difficulty in performing supervised learning because the learning data is not sufficient. If data on various abnormal sounds can be obtained, it will be possible to classify abnormal sounds in detail. In other words, if there is data in which the diesel generator sound changes

Table 14
Accuracy for diesel generator sound data.

Diesel Generator	Algorithm	Total Number of Samples	Erroneous Samples		Accuracy	
			No.	Quantity	Calculation	%
A	Autoencoder	1200	58–60, 117	4	1196/1200	99.7
	Binary Class.		52–57, 117	7	1193/1200	99.4
B	Autoencoder	7200	420, 421 7117–7119 7121, 7122 7152	8	7192/7200	99.9
	Binary Class.		7116–7122 7148–7154	14	7186/7200	99.8

according to various causes, it is possible to classify various abnormal sounds by changing to a multi-classification algorithm instead of a binary classification. Analysis showed that the unsupervised anomaly detection algorithm has a good accuracy compared with the general classification algorithm. In addition, as considering the applicability that there is not enough datasets on the anomaly detection in nuclear power plants, it could be a good option for diagnosing the fault in EDGs.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by the Korea Institute of Energy Technology Evaluation and Planning [KETEP Project No. 20204010600130] and the Ministry of Trade, Industry & Energy (MOTIE) of the Republic of Korea.

References

[1] K.H. Choi, S.G. Lee, A study on the power plant application of engine condition diagnosis technology for diesel generator, J. Korean Soc. Power Syst. Eng. 17 (2013) 17–22.
[2] S.G. Lee, Development status of emergency diesel generator status diagnosis technology for nuclear power plants, J. Electr. World (2014) 82–88. Special And+.
[3] S.G. Lee, D.W. Kim, Development and study of engine condition diagnosis program for emergency diesel generator for power generation, J. Korean Soc. Power Syst. Eng. 19 (2015) 67–72.
[4] K.W. Kang, K.M. Lee, CNN-Based automatic machine Fault Diagnosis method using spectrogram images, J. Instit. Converg. Signal Process 21 (2020) 121–126.
[5] S.H. Lee, J.S. Kim, B.B. Shin, CNN based actuator fault cause classification system using noise, in: Proceedings of the Winter Conference of the Korean Society of Computer and Information Engineers 26, 2018, pp. 7–8.
[6] T. Tran, J. Lundgren, Drill fault diagnosis based on the scalogram and mel spectrogram of sound signals using artificial intelligence, IEEE Access 8 (2020) 203655–203666, https://doi.org/10.1109/ACCESS.2020.3036769.
[7] H. Liu, L. Li, J. Ma, Rolling Bearing Fault Diagnosis Based on STFT-Deep Learning and Sound Signals, Hindawi Publishing Corporation Shock and Vibration, 2016, 6127479, https://doi.org/10.1155/2016/6127479, 2016.
[8] K.M. Lee, A Machine Fault Diagnosis System Based on Convolutional Neural Network Using Acoustic Power Spectral Density, Pukyong National University, 2019. https://repository.pknu.ac.kr: 8443/handle/2021.oak/23346.
[9] H. Nakamura, K. Asano, S. Usuda, Y. Mizuno, A diagnosis method of bearing and stator fault in motor using rotating sound based on deep learning, Energies 14 (2021) 1319, https://doi.org/10.3390/en14051319.
[10] C. Vununu, A Sound Diagnosis System with Deep Learning for Machine Fault Detection, Pukyong National University, 2017. https://repository.pknu.ac.kr: 8443/handle/2021.oak/14281.
[11] D. Li, B. Li, C. Wang, P. Cheng, B. Jiao, A fault diagnosis equipment of motor bearing based on sound signal and CNN, J. Phys.: Conf. Ser. (2010), 012159, https://doi.org/10.1088/1742-6596/2010/1/012159.
[12] S.H. Lee, B.B. Sin, Y.J. Kim, J.S. Kim, Deep-learning based PHM embedded system using noise · vibration, in: Proceedings of the Summer Conference of the Korean Society of Computer and Information Engineers 25, 2017, pp. 9–10.
[13] S. Zhao, E. Wang, J. Hao, Fault diagnosis method for energy storage mechanism of high voltage circuit breaker based on CNN characteristic matrix constructed by sound-vibration signal, J. Vibroeng. 21 (2019) 1665–1678, https://doi.org/10.21595/jve.2019.20781.
[14] Y. Yao, H. Wang, S. Li, Z. Liu, G. Gui, Y. Dan, J. Hu, End-to-end convolutional neural network model for gear fault diagnosis based on sound signals, Appl. Sci. 8 (9) (2018) 1584, https://doi.org/10.3390/app8091584.
[15] L. Weng, From Autoencoder to Beta-VAE, GITHUB, 2018. https://lilianweng.github.io/lil-log/2018/08/12/from-autoencoder-to-beta-vae.html.
[16] S. Jin, Introduction to the spectrum spectrogram, J. Kor. Soc. Laryngol. Phoniatr. Logop. 19 (2008) 101–106. http://jkslp.org/journal/view.php?number=1878.
[17] S.A. Broughton, K. Bryan, Discrete Fourier Analysis and Wavelets: Applications to Signal and Image Processing, second ed., Wiley, 2018, ISBN 978-1-119-25824-7.
[18] F. Chollet, Deep Learning with Python, Manning, 2017, ISBN 9781617294433.