Transactions of the Korean Nuclear Society Spring Meeting
Jeju, Korea, May 9-10, 2024

Multi-channel Signal Loss Restoration based on Autoencoder

Jaejun Lee [a], Hogeon Seo [a,b], Yonggyun Yu [a,b*]
[a]*Korea Atomic Energy Research Institute, 111, Daedeok-daero 989 beon-gil , Yuseong-gu , Daejeon, 34057, Korea*
[b]*Korea National University of Science & Technology, 217, Gajeong-ro, Yuseong-gu , Daejeon, 34113, Korea*
[*]*Corresponding author: ygyu@kaeri.re.kr*

***Keywords :** multi-channel signal, signal loss, signal restoration, autoencdoer

1. Introduction

In facilities such as nuclear power plants, efficiency and reliability are critical factors in successful operations. In particular, data collected by sensors, such as acceleration, power load, and temperature, provides crucial information for monitoring and controlling operational processes. However, during the data collection process, issues such as sensor failure or transmission errors can cause data to be lost or distorted. This can negatively impact operational processes and threat the structural safety and prognostics of a facility or system [1-2]. Therefore, it is necessary to develop effective recovery algorithms that maintain the reliability and consistency of data.

In this study, an autoencoder is utilized to restore the lost signal. This is expected to minimize the loss of information in the data, contributing to a better understanding of the operational status of the facility and improving the reliability of the system.

2. Proposed Method

This study utilizes a subset of the Electrical Transformer Temperature (ETT) dataset, which is used to predict oil temperature based on the power load of an electrical transformer. This is a dataset consisting of 69,680 data points collected at 15-minute intervals over a two-year period. Each data point consists of six power load characteristic values and a target value, the oil temperature; thus, the total number of data attributes is seven.

2.1 Data Preprocessing

Since the values of each attribute in the data have different ranges and magnitudes, we applied min-max scaling to ensure that all attributes are scaled on the same scale and to facilitate comparisons between attributes. This ensures that the data has values from 0 to 1 and that the model can be trained faster and more reliably.

The training data for training the Autoencoder model used 60% (41808×7) of the total data, and the validation and test data used 20% (13936×7) each. In addition, the test data was set to zero instead of its original value, assuming that one attribute value was randomly lost, as shown in Fig. 1.

After the training and validation process uses normally collected signal data with no lost attributes as input to train the model to restore the data, the trained model was tested using data with one randomly lost attribute as input to restore the signal data.

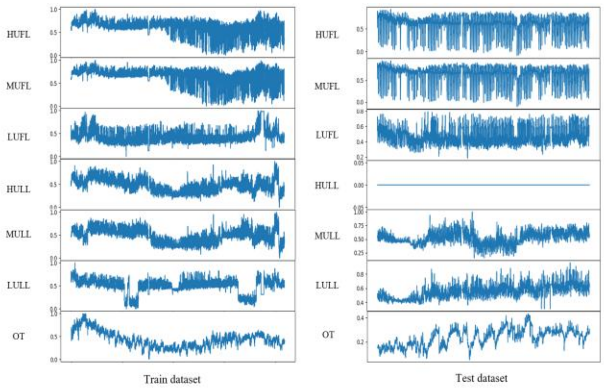

Fig. 1. Schematic of preprocessed data.

2.2 Autoencoder Model

The autoencoder model consists of an encoder that extracts meaningful features based on input data and maps them to a latent vector z and a decoder that reconstructs the data based on the latent vector z. Autoencoders, especially those based on unsupervised learning methods, can operate effectively even when data is poorly labeled and have been used for anomaly detection [3], time series forecasting [4], and more. For this reason, we applied autoencoders as a method to restore lost signal data.

The detailed configurations of the encoder and decoder of the proposed autoencoder are shown in Table I and Table II. Except for the latent vector z and the layer that outputs the restored data, we use batch normalization (BN) and rectified linear unit (ReLU) as the activation function.

The hyperparameters for training the model were set to an epoch number of 100, a batch size of 32, and a learning rate of 5e-4. We used root mean squared error (RMSE) as the loss function to measure the error between the input and reconstructed data. Also, we set an early stop to prevent the model from overfitting and shorten the learning time.

Table I: The Configuration of The Encoder

Layer composition	Number of input features	Number of output features
Linear Layer + BN	7	64
Linear Layer + BN	64	32
Linear Layer	32	16

Table II: The Configuration of The Decoder

Layer composition	Number of input features	Number of output features
Linear Layer + BN	16	32
Linear Layer + BN	32	64
Linear Layer	64	7

2.3 Restoration Result

The reconstruction results for the test data in Fig. 1 are shown in Fig. 2, based on an autoencoder model trained only on normal and intact data. Although there was one lost attribute out of seven features, the restoration resulted in a close approximation to the values and trends of the attribute in the original data. However, the restoration was not complete due to the lost attribute, which affected the other normal attributes. In particular, the distorted attribute has a strong correlation with the lost attribute, suggesting that the autoencoder model may have been affected in the process of restoring other attributes.

A comparison of the RMSE loss before and after restoration based on normal test data when each feature was lost in the data is shown in Fig. 3. Overall, when each attribute was restored by the autoencoder model, lower losses were abtained, showing improvement.

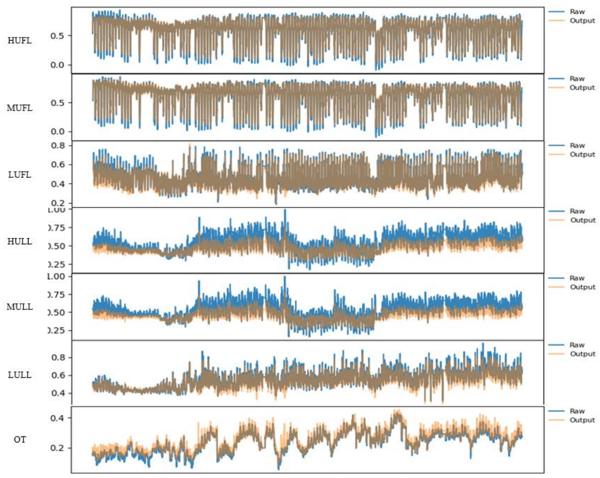

Fig. 2. Results of restoring the autoencoder model.

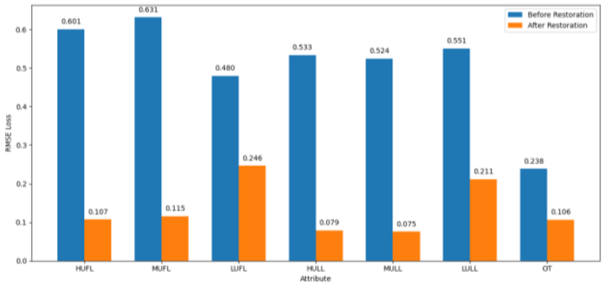

Fig. 3. Comparison of RMSE loss before and after restoration.

3. Conclusions

In this study, we propose a method to restore signal data with lost features through an autoencoder model trained only on normal and intact data. This suggests that autoencoders can be an effective tool for restoring signal data with lost features and contribute to the compensation of the information loss of the data. The results show that lower losses were obtained when each attribute was restored by the autoencoder model, supporting that the proposed method contributes to improving the resilience to sensor failures. We plan to conduct future research by considering sequences rather than data points or utilizing datasets that contain attributes that were lost during model training to obtain more complete restoration results.

ACKNOWLEDGMENTS

This work was supported in part by Korea Atomic Energy Research Institute R&D Program under Grant KAERI-524540-24 and in part by the National Research Foundation of Korea under Grant RS-2023-00253853 funded by the Korean government (Ministry of Science and ICT).

REFERENCES

[1] J. Ma, J. Jiang, Applications of fault detection and diagnosis methods in nuclear power plants: A review, Progress in Nuclear Energy, Vol. 53, No. 3, pp. 255-266, 2011.

[2] H. Kim, M. G. Na, G. Heo, Application of monitoring, diagnosis, and prognosis in thermal performance analysis for nuclear power plants, Nuclear Engineering and Technology, Vol. 46, No. 6, pp. 737-752, 2014.

[3] Z. Chen, C. K. Yeo, B. S. Lee, C. T. Lau, Autoencoder-based Network Anomaly Detection, 2018 Wireless Telecommunications Symposium (WTS), pp. 1-5, 2018.

[4] A. Sagheer, M. Kotb, Unsupervised Pre-training of a Deep LSTM-based Stacked Autoencoder for Multivariate Time Series Forecasting Problems, Scientific Reports, Vol. 9, No. 1, pp. 19038, 2019.