Transactions of the Korean Nuclear Society Virtual Autumn Meeting
October 21-22, 2021

Convolutional Neural Network using Plot Image Data for Abnormal Scenario Diagnosis in Nuclear Power Plants

Sang Won Oh, Sang Hyun Lee, Ji Hun Park, and Man Gyun Na*
Dept. of Nuclear Engineering, Chosun Univ., 309 Pilmun-daero, Dong-gu, Gwangju, 61452
**Corresponding author: magyna@chosun.ac.kr*

1. Introduction

The safety of nuclear power plants (NPPs) can be threatened because of a variety of causes. The NPP operators should take proper action after identifying the cause of an accident. However, the human errors of the operators that have to identify the accident in a limited situation (i.e., limited time and many surveillance elements) can increase.

Recently, many studies have been conducted to assist operators using artificial intelligence (AI) aiming to reduce human errors. Since the NPP data have time-series characteristics, it has a high performance by applying time-series data-based methods (recurrent neural network, long short-term memory, gated recurrent unit, etc.). In addition, the application of explainable AI (XAI) to ensure the AI reliability in various fields is being conducted. However, the application of XAI is more actively performed on image data compared to time-series data.

In this paper, we have conducted basic research for a real-time image-based accident identification system. Here, this study focused on the abnormal situation of the NPPs. Specifically, this study was carried out in the following steps; 1) image conversion of time-series data, 2) converted image data-based AI application, and 3) application of XAI.

First, in researches related to image conversion of time-series data, the Gramian angular field (GAF) method is mainly utilized [1]. The GAF method is a technique that converts the one-dimensional time-series data into a two-dimensional matrix by maintaining the time relationship between variables. However, there are several problems to apply to this study; 1) the conversion time of image data is not suitable for the real-time system, and 2) the converted image is difficult to understand intuitively. Therefore, a trend-based plot image conversion method that can be suitable for real-time systems and extract an intuitive image, is utilized.

Second, a convolutional neural network (CNN) [2], which is mainly utilized in AI studies using image data, has been adopted as an application model.

Finally, the application of gradient-weighted class activation mapping (Grad-CAM), which is an XAI method in the CNN model is considered. By applying XAI method, the operators can intuitively verify the focusing point of AI in the plot image. Therefore, it is expected to enhance the AI reliability by confirming the focusing point of the AI.

2. Method

2.1 Convolutional Neural Network

The structure of CNN is divided into a part that extracts features of an image and a part that classifies the class. The feature extraction area consists of single or multiple convolution and pooling layers. At the end of the CNN, a fully connected (FC) layer is constructed which is used to classify the flattened data. Fig. 1 shows the CNN structure.

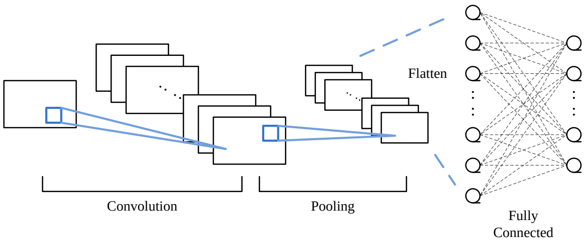

Fig. 1. Overview of the CNN structure.

In the convolution layer, a filter called a kernel moves at a specified interval (i.e., stride value); and a feature map is formed through the convolution of the kernel and the input data. Here, convolution is to multiply the input data by the corresponding elements of the kernel and then calculate the sum of the multiplied values (refer to Fig. 2). Finally, the feature map formed through the convolution represents the characteristics of the image.

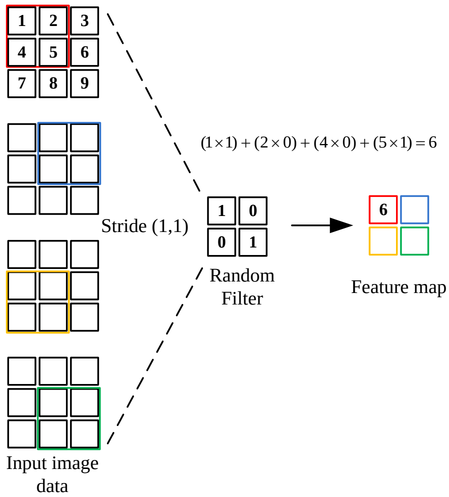

Fig. 2. An illustration of convolution calculation procedure and feature map extraction process.

In the pooling layer, the specific data of the feature map is emphasized, and the size of the data is reduced to lower the computational complexity. The pooling layer is divided into max pooling and average pooling depending on whether the maximum or average value is collected in the feature map. Since max pooling is mainly used, max pooling was used in this study. The last layer of CNN, the FC layer, is used to derive the output.

2.2 Grad-CAM

Class activation map (CAM) is a methodology to find factors that greatly influenced the classification results in images by putting global average pooling (GAP) in the last convolution layer of the CNN [3]. In detail, CAM calculates the average value for each element through GAP and then utilizes the weight associated with the value and classification result. However, GAP is utilized at the latter part of CNN, which causes limitations in model construction.

To solve this problem, Grad-CAM has been proposed [4]. Grad-CAM improved the problem by finding the gradient of the weights connected to the feature (i.e., element). That is, the main elements can be found without modifying the model structure. Finally, the main elements from the input image are output as a heatmap, and the result is emphasized.

3. Experiment

3.1 Data Collection

Data were collected using a compact nuclear simulator (CNS). The CNS was developed with reference to the design of Westinghouse 3 loop, 993MWe pressurized water reactor Kori NPP units 3 & 4 for the purpose of basic system education [5]. In data collection, 5 scenarios related to the pressurizer (PZR), which were easy to identify based on the same variables, were selected. The selected scenarios are the PZR level channel failure "high", the PZR pressure-operated relief valve (PORV) open, PZR safety valve failure, PZR spray valve failure "open" and normal scenario. Data on 2,222 variables of the CNS were collected from the selected scenarios.

3.2 Data Pre-Processing

In the data preprocessing, 6 variables with high correlation concerning PZR were selected among 2,222 variables of the collected data. The selected variables are shown in the Table I.

Table I: Selected input variables

No.	Variable name	Description
1.	PPRZ	PZR pressure
2.	QPRZH	Proportional heater fractional power
3.	UPRZ	PZR temperature
5.	WPORV	PZR PORV flow rate to pressurizer relief tank (PRT)
6.	ZPRZUN	PZR uncompensated level

For the image format, a subplot format was used so that 6 selected variables could be included in one image. The image data are collected using a trend-based plot image conversion method (refer to Fig. 3). That is, the plot image data for each scenario was plotted by overlapping every 30 (or 60, 90) seconds from the start (e.g., 1 to 30sec, 1 to 60sec, 1 to 90sec, etc.).

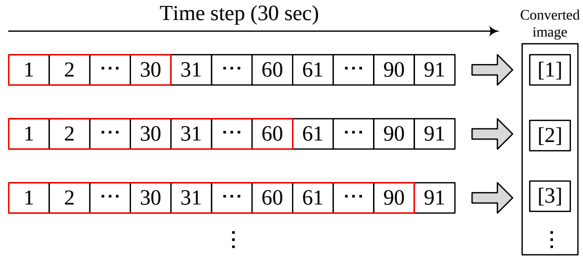

Fig. 3. Trend-based plot image conversion method.

In an extracted image, each variable was distinguished by fixing the position and color of the plot line. Fig. 4 represents the subplot image of input data for trend-based plot image conversion. In order to exclude unnecessary information in image recognition when training the AI model, the axis and legends in input data were excluded from the image.

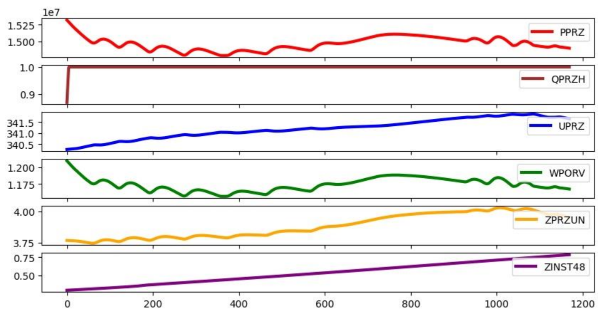

Fig. 4. Subplot image of PZR PORV open scenario.

4. Diagnosis Result of CNN Model

When training AI, the larger the number of data and the more balanced the data distribution, the better the performance. However, as can be observed in Table II, the data imbalance problem occurred due to the small amount of normal data. Therefore, the under-sampling method was used to prevent data imbalance problem.

Under-sampling is the operation of balancing data by reducing the number of large data set [6]. Therefore, the number of the training data was adjusted to 150 according to the normal scenario with the least data. However, 150 training data is regarded as a small number for general deep learning. Nevertheless, this study has proposed a classification of 5 simple scenarios as basic research; that is, a simple classification problem that can be classified even with a small amount of data. Data information used for training and testing is presented in Table II.

Table II: Number of the data for each scenario

Class	Scenario name	No. of image data	No. of train data	No. of test data
Ab20-04	PZR level channel failure “high”	555	150	405
Ab21-12	PZR PORV open	969	150	819
Ab19-02	PZR safety valve failure	805	150	655
Ab21-11	PZR spray valve failure “open”	1,536	150	1386
Normal	Normal	199	150	49

Fig. 5 shows the model test results through the confusion matrix. The confusion matrix is a table that shows how well the predicted class matches the true class. The training was performed on a total of 750 train data, and as a result of 3,314 test data, all data scenarios were classified with 100% accuracy. Accuracy is expressed as the following Eq. (1).

$$\text{Accuracy} = \frac{\text{number of test data predicted correctly}}{\text{number of total test data}} \tag{1}$$

$$= \frac{49+655+405+1386+819}{49+655+405+1386+819} = 1$$

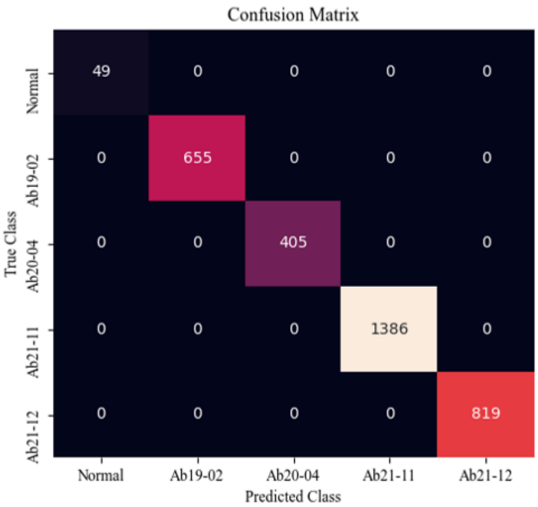

Fig. 5. A confusion matrix representing the test result of a trained CNN model.

Fig. 6 is an interpretation result to which the Grad-CAM method is applied to intuitively understand the diagnosis result of AI. In Fig. 6, the part highlighted in red is the UPRZ variable part, indicating that the increase in the PZR temperature became the evidence for the diagnosis result of AI.

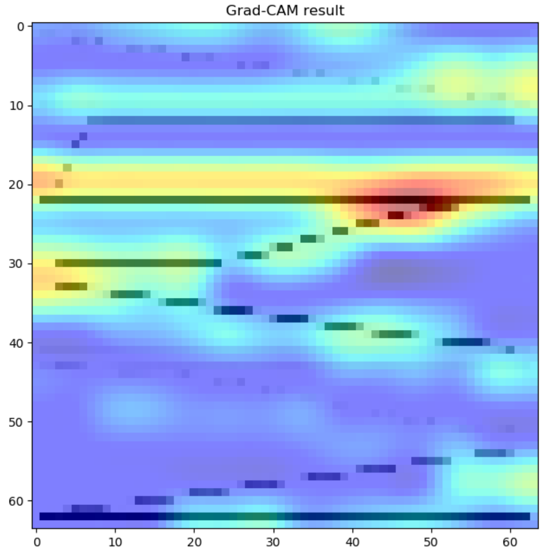

Fig. 6. Grad-CAM interpretation result when the diagnosed scenario is Ab21-12.

5. Conclusions

In this study, image processing, model application, and XAI application for abnormal situations identification were performed based on image data. Multivariate time-series data were converted to plot image data, and the converted data was applied to the AI model as input data. The model performance showed 100% accuracy. Additionally, the Grad-CAM method was applied to confirm the diagnosis results of reliable AI. From these results, it is meaningful to conduct additional research on various image-based AI methods.

This study was performed only on a simple case (i.e., classification of 5 scenarios) as a part of basic research. In future works, we plan to extract images with complex problems (i.e., classification problem with more than 5 scenarios and various input variables) and use different models to identify different kinds of abnormal scenarios. Through such image-based AI research, more intuitive image data can be used, and it is expected that the XAI will be applied actively.

REFERENCES

[1] Z. Wang, T. Oates, Imaging Time-Series to Improve Classification and Imputation, Proceedings of the Twenty-Fourth International Joint Conference on Artificial

Intelligence (IJCAI-2015), July. 25-31 2015. Buenos Aires, Argentina.
[2] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai, T. Chen, Recent Advances in Convolutional Neural Networks, Pattern Recognition, Vol. 77, pp. 354-377, 2018.
[3] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, A. Torralba, Learning Deep Features for Discriminative Localization, In Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, pp. 2921-2929, 2016.
[4] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization, Proceedings of the IEEE International Conference on Computer Vision, pp. 618-626, 2017.
[5] K.H. Cha, J.C. Park, K.C. Kwon, G.Y. Lee, Reuse of Compact Nuclear Simulator Software under PC with Linux, Proceedings of the Korean Nuclear Society Virtual Autumn Meeting, Oct. 26-27, 2000, Daejeon, Korea.
[6] S. J. Yen, Y.S. Lee, Cluster-Based Under-Sampling Approaches for Imbalanced Data Distributions, Expert Systems with Applications, Vol. 36, pp. 5718-5727, 2009