**Enhancing Fault Identification in Chemical Plants: A Multimodal Approach Combining CNN and Continuous Wavelet Transform**

Chinatsu Ukawaa, Yoshiyuki Yamashitaa\*

aTokyo University of Agriculture and Technology, 2-24-16 Naka-cho, Koganei, Tokyo 184-8588, JAPAN

\*yama\_pse@cc.tuat.ac.jp

Abstract

Fault detection and identification is a crucial task to ensure safe and stable operation of chemical plants. This study introduces a novel multimodal model that incorporates a continuous wavelet transform (CWT) along with two distinct convolutional neural networks (CNN): a two-dimensional CNN (2DCNN) and a three-dimensional CNN (3DCNN).  Utilizing a time-shifting window, multivariate time series data from chemical plants are segmented, with the resulting segments combined into two-dimensional input data. Prior to the combination, the data were transformed into scalograms using CWT. The multimodal model processes these 2D and 3D input data, producing outputs that indicate the occurrence of faults in the chemical plant processes. We applied the proposed method to the Tennessee Eastman process (TEP) dataset. The proposed method showed superior accuracy in fault identification compared to existing methods.

**Keywords**: Fault Identification, Continuous Wavelet Transform, Three-dimensional convolutional neural network

* 1. Introduction

Fault detection and identification is one of the most important steps in chemical process monitoring to ensure safe, stable, and efficient operation. Up to now, numerous studies have been conducted due to the aforementioned fact that fault detection and identification are essential for chemical process monitoring. In recent years, a variety of deep learning methods have been applied to this field due to the exponential growth of computer technology. There are a lot of types of deep learning methods. For instance, CNN, deep belief network, autoencoder, long short-term memory, and recurrent neural network (Chen et al., 2022).

In our previous work, we introduced the fault detection and identification model using CWT and 3DCNN (Ukawa et al., 2023). Although the model showed good performance on about half of the process conditions of the TEP datasets, the accuracy of the normal operating condition and some of the faulty condition showed lower accuracy than other comparative models. We had assumed that this was due to the lack of information about the absolute values of the chemical process data, which occurs when CWT is used to pre-process the original data. Therefore, in this paper, we propose a novel multimodal model using CWT and two types of CNN: 2DCNN and 3DCNN, to retain and effectively use the absolute values of chemical process data.

* 1. Method
		1. CWT

CWT is one of the wavelet transforms which is used for signal and image processing. The wavelet function is shown in equation 1. The parameter *a* represents the scale index which determines the center frequency of the function ψ (a-1(t-b)). The parameter *b* indicates the time shifting.

CWT generates scalograms from original signal data. Scalograms are the plot of the CWT coefficient and show the feature on the time-frequency domain. The application of spectrograms has been reported in various fields.

* + 1. 2DCNN and 3DCNN

The CNN has been widely used for various types of research. Especially, 2DCNN has gotten attention because of its high ability in the fields of image recognition, action recognition, medical imaging, and machinery. It is also used for research on process systems engineering and demonstrated good performance (Zhu et al., 2019) (Wu et al., 2018) (Chen et al., 2022).

3DCNN has garnered attention in the fields of action recognition (Liu et al., 2019) and 3D image analysis (Riahi et al., 2022). It is the logical extension of 2DCNN. The 2DCNN uses 2D convolution filter to extract the features of the 2D images. This process will be conducted for each channel of the input images. On the other hand, 3DCNN uses 3D filters and extracts the features of all channels or all dimensions simultaneously. Due to this characteristic, 3DCNNs enable feature extraction from data while preserving relationships across channels or dimensions. In this study, 3DCNN was used to deal with 3 dimensions: time, frequency, and process variables.

* + 1. Proposed method

The approach involves the following steps: initially, the process data is divided by applying a moving time window to each process variable. This segmentation process results in 2D input data derived from segments of each process variable, which is subsequently prepared for the 2DCNN model. These segments transform into spectrograms through CWT, capturing the unique features of chemical process data in the time-frequency domain. After this preprocessing, 3D input data is generated for the 3DCNN. A multimodal model is then trained, integrating both 2D and 3D data, with the output representing various process states, encompassing normal operation and fault conditions.

* 1. Case study

The proposed model was evaluated using TEP datasets introduced by Rieth et al. (2017). This process consists of five main units: a reactor, a condenser, a stripper, a separator, and a compressor. There are 52 process variables including 41 measured variables and 11 manipulated variables. In this study, we used 33 variables including 22 measured variables and all manipulated variables.

* 1. Result

Figure 1 The TEP flow sheet

* + 1. Evaluation metrics

We assessed the suggested approach and analyzed its effectiveness by examining the fault detection rate (FDR) and false positive rate (FPR). In this context, *TP* denotes true positive, while *FN* signifies false negative.

|  |
| --- |
|  |
| * + 1. Comparison of the application result
 |

Table 1 shows a comparison of model accuracies of 2DCNN, CWT-3DCNN, and Multimodal CWT-3DCNN. The accuracy average of CWT-3DCNN is lower than the 2DCNN model although it demonstrated better performance on half of the faults. On the other hand, multimodal CWT-3DCNN accomplished the best performance comparing the accuracy average. The accuracy of the normal operation condition is still not high, but it is improved compared with the previous CWT-3DCNN model. In the case When 2DCNN and CWT-3DCNN have scores of 0.8 or higher, the score of multimodal CWT-3DCNN becomes higher than that of 2DCNN and CWT-3DCNN. The average of FDR shows the best score among the three models. It indicates that using the time-frequency information with absolute values is possible to further improve and enhance the ability of the CNN model for fault detection and identification.



Figure 2 Training history of the accuracy and the loss

Table 1 Comparison of model accuracies

* 1. Discussion
		1. Model Training

Figure 2 shows the learning history of the models. CWT-3DCNN models can achieve higher accuracy than the 2DCNN model in exchange for the long training time. The proposed Multimodal model demonstrates the highest accuracy among the comparing models. CWT-3DCNN models are found to be less prone to overfitting even when trained for extended periods, as compared to other models.

* + 1. Batch size

Batch size affects the learning speed and model accuracy. We investigated the difference in the accuracy between the models which have different batch sizes. We selected 64, 128, 256, 512, and 1024 for batch size. Table 2 shows the result comparison. The best batch size is 512. The optimized batch size can contribute to shorten model training time.

Table 2 Result comparison of the models with different batch sizes.



* 1. Conclusion

In prior studies, the fault detection and identification model proposed using CWT and 3DCNN demonstrated notable effectiveness on benchmark process datasets. However, it failed to achieve a higher accuracy score compared to previous models, especially 2DCNN, particularly under normal operating conditions in TEP datasets. The preprocessing with CWT resulted in a lack of information from absolute values of chemical process data, as opposed to the abundance of time-frequency information.

Therefore, in this study, a model capable of handling both data types: time-frequency scalograms and absolute values of chemical processes is introduced. The incorporation of these data types is expected to enhance the model's performance in fault detection and identification. The proposed model is a multimodal CNN with 2D and 3D input layers, utilizing both 2DCNN and 3DCNN.

The model is evaluated on TEP datasets. The multimodal CWT-3DCNN achieves the best overall performance, surpassing the accuracy average of the other models although the model training time is longer than 2DCNN. Batch size is one of the important hyper- parameters of deep learning. We investigated the effects of the different batch sizes. In this case, the best batch size is 512. The bigger batch sizes can contribute to shortening model training time.

References

C. Ukawa, Y. Yamashita, Fault Detection and Diagnosis for Chemical Processes based on Deep Neural Networks with Continuous Wavelet Transform, Computer Aided Chemical Engineering, Elsevier (2023), <https://doi.org/10.1016/B978-0-443-15274-0.50267-5>

C. A. Rieth, B. D. Amsel, R. Tran, M. B. Cook, Additional Tennessee Eastman Process Simulation Data for Anomaly Detection Evaluation, Harvard Dataverse (2017), V1, <https://doi.org/10.7910/DVN/6C3JR1>

H. Chen, Jian. C, Z. Yang, W. Si, H. Cheng, Fault diagnosis of the dynamic chemical process based on the optimized CNN-LSTM network. ACS Omega (2022), <https://doi.org/10.1021/acsomega.2c04017>

G. Zhu, L. Zhang, P. Shen, J. Song, S. A. A. Shah, M. Bennamoun, Continuous Gesture Segmentation and Recognition Using 3DCNN and Convolutional LSTM, IEEE Trans. Multimedia (2019), <https://doi.org/10.1109/TMM.2018.2869278>

H. Wu, J. Zhao, Deep convolutional neural network model based chemical process fault diagnosis, Computers & Chemical Engineering (2018), <https://doi.org/10.1016/j.compchemeng.2018.04.009>

A. Riahi, O. Elharrouss, S. A. Maadeed, BEMD-3DCNN-based method for COVID-19 detection, Computers in Biology and Medicine (2022), <https://doi.org/10.1016/j.compbiomed.2021.105188>

Y. Liu, T. Zhang, Z. Li, 3DCNN-Based Real-Time Driver Fatigue Behavior Detection in Urban Rail Transit, IEEE Access (2019), <https://doi.org/10.1109/ACCESS.2019.2945136>

S. Shikhar, K. Kumar. ASL-3DCNN: American sign language recognition technique using 3-D convolutional neural networks, Multimedia Tools and Applications (2021), <https://doi.org/10.1007/s11042-021-10768-5>

S, Qian. *Introduction to time-frequency and wavelet transform.* 2002. Prentice Hall, New Jersey