Anomaly Detection using Autoencoder with Gramian Angular Summation Field in Multivariate Time Series Data

Umaporn Yokkampon[1], Abbe Mowshowitz[2], Sakmongkon Chumkamon[1], Eiji Hayashi[1]

[1]Graduate School of Computer Science and Systems Engineering, Kyushu Institute of Technology,
680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan
[2]Department of Computer Science, The City College of New York,
160 Convent Avenue, New York, NY 10031, USA
E-mail: may@mmcs.mse.kyutech.ac.jp, amowshowitz@ccny.cuny.edu,
m-san@mmcs.mse.kyutech.ac.jp, haya@mse.kyutech.ac.jp
www.kyutech.ac.jp

Abstract

Uncertainty is ubiquitous in data and constitutes a challenge in real-life data analysis applications. To deal with this challenge, we propose a novel method for detecting anomalies in time series data based on the Autoencoder method, which encodes a multivariate time series as images by means of the Gramian Angular Summation Field (GASF). Multivariate time series data is represented as 2D image data to enhance the performance of anomaly detection. The proposed method is validated with four time-series data sets. Experimental results show that our proposed method can improve validity and accuracy on all criteria. Therefore, effective anomaly detection in multivariate time series data can be achieved by combining the methods of Autoencoder and Gramian Angular Summation Field.

Keywords: Anomaly detection, Factory automation, Autoencoder, Multivariate time series.

1. Introduction

Anomaly detection involves discovering unusual patterns in time series data. Such detection is an important and challenging task, widely applied in various fields, such as credit card processing, medical diagnosis, sensor network operations, intrusion detection, and other areas.[1] Many algorithms have been developed to detect anomalies, but most of them are limited to analysis of univariate time series.

Recently, deep learning techniques have been applied to anomaly detection, the most common being Principal Component Analysis (PCA), which is a linear dimensionality reduction technique. PCA projects high-dimensional time series into a low-dimensional sequence. However, PCA is not flexible and cannot perform non-linear operations. A newer dimensionality reduction method is autoencoder[2,3], which has become a popular method for anomaly detection. Autoencoder is used to perform dimension reduction by stacking up layers to form a deep autoencoder. The hidden units are expected to extract features that represent the data faithfully by reducing the number of units in the hidden layer. Autoencoder judges whether something is abnormal by considering the difference between encoded data and the original.[4,5] Moreover, the autoencoder can perform both linear and non-linear operations.[6]

In this paper, we propose an anomaly detection method based on the autoencoder which encodes the multivariate time series data as 2D images designed to enhance performance. Our proposed method is validated by comparing the encoded results with the original data

using four standard datasets. Four validation criteria have been used, namely, precision, recall, F1-score, and g-mean.

2. Proposed Method

This section introduces our proposed method for improving anomaly detection performance, and discusses the datasets and evaluation metrics used in this paper. Fig. 1 displays the overall structure of the proposed method. First, four time series data sets are converted into 2D images using Gramian Angular Summation Field (GASF), and then input into the autoencoder to identify anomalies by constructing the encoder and decoder. The results are then obtained from the autoencoder reconstruction values.
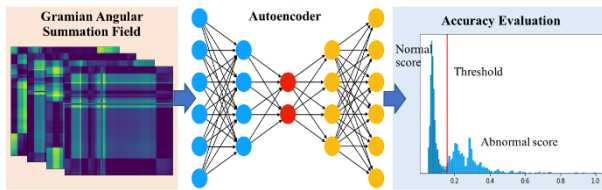

Fig. 1. Concept of the proposed method.

2.1. *Autoencoder*

Autoencoder is an unsupervised artificial neural network, which can leverage neural networks for the task of representation learning[7]. Autoencoder learns how to compress and encode data efficiently and then to reconstruct the data from the reduced encoded representation to one that is as close to the original input as possible. It consists of an encoder, a latent variable, and a decoder.

The encoder maps input data $x \in \mathbb{R}^d$ to a latent variable (code) $z \in \mathbb{R}^d$ and the decoder maps back from latent variable to input space. The autoencoder training procedure uses backpropagation to minimize the network's reconstruction loss. The loss function measures the differences between the original input and the consequent reconstruction. The loss function is defined as follows:

$$L(x, \hat{x}) = \| x - \hat{x} \|^2 \qquad (1)$$

After the training procedure, the autoencoder uses the reconstruction error as the anomaly score. The data with high anomaly scores are considered anomalies because only the normal data are used to train the autoencoder. The autoencoder will reconstruct normal data very well while failing to do so when confronted with anomalous data.

2.2. *Gramian Angular Summation Field (GASF)*

A Gramian Angular Field(GAF)[8] is an image obtained from a 1-dimensional time series, representing some temporal correlation between each time point. There are two methods available: Gramian Angular Summation Field(GASF) and Gramian Angular Difference Field. In this work, Gramian Angular Summation Field (GASF) is used to transform our time series data into 2D images, using a polar coordinate system.

In the Gramian matrix, each element is the cosine of the summation of pairwise temporal values. Given a time series $X = (x_1, x_2, ..., x_n)$ of n observations, we rescale X so that all values fall in the interval [0, 1]:

The signal is warped in the transform domain. After this, each time point in polar coordinates is compared with every other point for temporal correlation.

$$\tilde{X}_0^i = \frac{x_i - \min(X)}{\max(X) - \min(X)} \qquad (2)$$

At this point, we convert the obtained values into the polar coordinate system as follows:

$$\begin{cases} \phi = \arccos(\tilde{x}_i), & -1 \le \tilde{x}_i \le 1, \tilde{x}_i \in \tilde{X} \\ r = \dfrac{t_i}{N}, \ t_i \in N \end{cases} \qquad (3)$$

where t_i represents the time stamp and N is a constant factor used to regularize the polar coordinate system. After transforming the rescaled time series into the polar coordinate system, GASF can be written as follows:

$$\text{GASF} = \begin{bmatrix} \cos(\phi_1 + \phi_1) & \cdots & \cos(\phi_1 + \phi_n) \\ \cos(\phi_2 + \phi_1) & \cdots & \cos(\phi_2 + \phi_n) \\ \vdots & \ddots & \vdots \\ \cos(\phi_n + \phi_1) & \cdots & \cos(\phi_n + \phi_n) \end{bmatrix} \qquad (4)$$

$$\text{GASF} = \tilde{X}' \cdot \tilde{X} - \sqrt{I - \tilde{X}^2}' \cdot \sqrt{I - \tilde{X}^2} \qquad (5)$$

where I is the unit row vector.

3. Experiments

In this section, we first describe the details of the datasets[9] used in our experiments, and present the performance metrics for evaluating the performance of our system.

3.1. *Datasets*

To demonstrate the effectiveness of the proposed method, we conducted experiments on four datasets: Satellite, SonyAIBORobotSurface2, ItalyPowerDemand, and Wafer. Table 1. shows the details of the datasets. We use 80% of the normal data for training, and the remaining 20% of the normal data is used for testing purposes.

Table 1. Details of the datasets.

Datasets	Length	Number of instances	Anomaly Ratio
Satellite	36	6435	0.32
SonyAIBORobotSurface2	65	980	0.38
ItalyPowerDemand	24	1096	0.49
Wafer	152	7164	0.11

3.2. *Performance Evaluation*

To evaluate the detection performance of our proposed method, we employed four metrics, namely, Precision, Recall, F1-Score, and G-mean[10] which are defined as follows:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (6)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (7)$$

$$\text{F1-Score} = 2 \times \frac{\text{Precision} 識 ecall}{\text{Precision}+\text{Recall}} \quad (8)$$

$$G-mean = \sqrt{\frac{TP}{TP+FN} \times \frac{TN}{TN+FP}} \quad (9)$$

where *TP* is the correctly predicted anomaly, *FP* is the incorrectly predicted anomaly, *TN* is the correctly assigned normal, and *FN* is the incorrectly assigned normal.

4. Results and Discussion

In this section, we evaluate the proposed method for detecting anomalies in a time series data set using the Autoencoder method. First, we transform all time series data sets into 2D images using GASF. Fig. 2 shows the GASF images example, in which the image represents some temporal correlation between each time point. The result is from the wafer dataset.
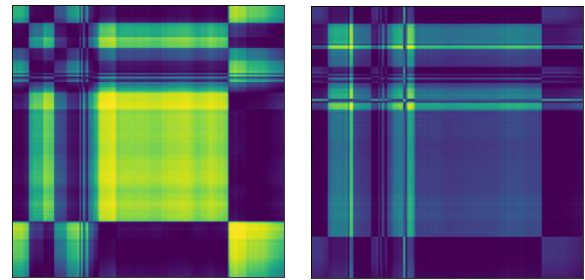

Fig. 2. The example of GASF images.

Our proposed method is trained with the normal samples, and then the model is verified using testing samples, including both normal and abnormal data. We conducted our experiment to improve the anomaly detection performance in four time series data sets using precision, recall, F1-score, and G-mean criteria. Table 2. shows the summary and comparison results of anomaly detection. The bold fonts in Table 2 indicate our proposed method could improve performance relative to relying on original data. For the Wafer dataset, the F1-Score value associated with the original data was 84.64%, whereas our method achieved 99.06%. In addition, the G-mean of the original data was only 51.61% compared with 99.48% for our proposed method. However, the SonyAIBORobotSurface2 and ItalyPowerDemand datasets of original data gave precision results better than our proposed method.

Table 2. Autoencoder compare of original data and our proposed method

| Datasets	Original Data				Our proposed method			
	Precision	Recall	F1-Score	G-mean	Precision	Recall	F1-Score	G-mean
Satellite	0.7075	0.6779	0.7761	0.5074	0.7369	**0.9989**	0.8481	0.8325
SonyAIBORobotSurface2	0.9043	0.7647	0.8287	0.7903	0.8891	**0.9982**	0.9405	0.9171
ItalyPowerDemand	0.7091	0.5166	0.5977	0.5263	0.6818	0.7813	0.7282	**0.8559**
Wafer	0.7349	0.9979	0.8464	0.5161	0.9820	**0.9992**	0.9906	0.9948
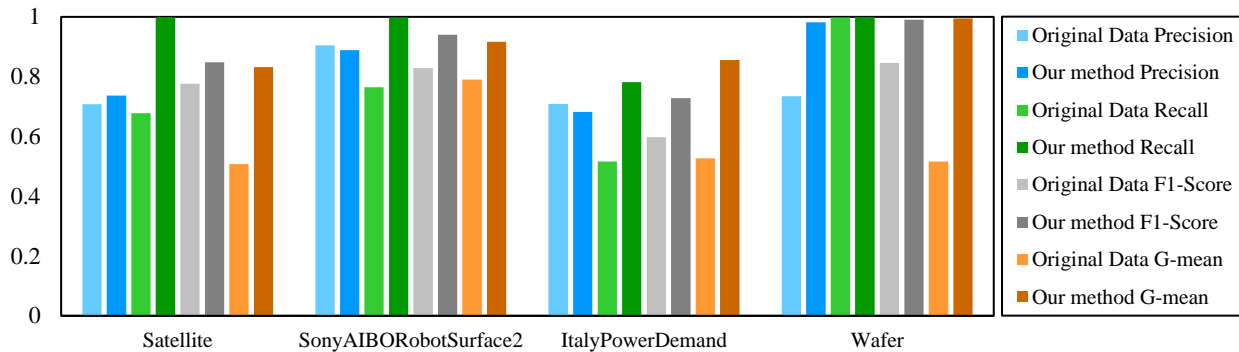
Fig. 3. All criteria comparisons of four datasets on Autoencoder

Fig. 3. shows the chart of all criteria comparisons of four datasets on Autoencoder. The light color line represents the original data, and the dark color line represents our proposed method. It is quite clear that our proposed method outperforms original data.

5. Conclusion

In this paper, we have presented an anomaly detection method using an autoencoder to compare the original data represented as GASF images (2D images). Performance has been evaluated by the metrics Precision, Recall, F1-Score, and G-mean. The experiments on anomaly detection show that our proposed method could improve the accuracy of detecting anomalies for time series data. Therefore, anomaly detection in multivariate time series, judged by the criteria of precision, recall, F1-score, and G-mean, can be improved by using Gramian Angular Summation Field (GASF) image encoding.

Acknowledgements

References

1. V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey", *ACM computing surveys (CSUR)*, Vol. 41, No. 3, pp. 1-58, 2009.
2. A. Ng, "Sparse autoencoder", *CS294A Lecture notes*, Vol. 72, No. 2011, pp.1-19, 2011.
3. C. Zhou, and R.C. Paffenroth, "Anomaly detection with robust deep autoencoders", In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 665-674, August, 2017.
4. A. Borghesi, A. Bartolini, M. Lombardi, M. Milano, and L. Benini, "Anomaly detection using autoencoders in high performance computing systems", In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33, No. 01, pp. 9428-9433, July, 2019.
5. B. Zong, Q. Song, M.R. Min, W. Cheng, C. Lumezanu, D. Cho, and H. Chen, "Deep autoencoding gaussian mixture model for unsupervised anomaly detection", In *International conference on learning representations*, February, 2018.
6. M. Sakurada, and T. Yairi, "Anomaly detection using autoencoders with nonlinear dimensionality reduction", In *Proceedings of the MLSDA 2014 2nd workshop on machine learning for s ensory data analysis*, pp. 4-11, December, 2014.
7. Z. Chen, C.K. Yeo, B.S. Lee, and C.T. Lau, "Autoencoder-based network anomaly detection", In *Wireless Telecommunications Symposium (WTS)*, pp. 1-5, April, 2018.
8. Z. Wang, and T. Oates, "Imaging time-series to improve classification and imputation", In *Twenty-Fourth International Joint Conference on Artificial Intelligence*, June, 2015.
9. H.A. Dau, A. Bagnall, K. Kamgar, C.C.M. Yeh, Y. Zhu, S. Gharghabi, C.A. Ratanamahatana, and E. Keogh, 2019. "The UCR time series archive", *IEEE/CAA Journal of Automatica Sinica*, Vol. 6, No. 6, pp.1293-1305, 2019.
10. X. Hua, Y. Cheng, H. Wang, Y. Qin, and Y. Li, "Geometric means and medians with applications to target detection", *IET Signal Processing*, Vol. 11, No.6, pp.711-720, 2017.

Authors Introduction

Ms. Umaporn Yokkampon

Ms. Umaporn Yokkampon received Master of Science degree in Applied Statistics from King Mongkut's University of Technology North Bangkok in 2019. Presently she is a doctoral student in the Department of Computer Science and Systems Engineering at Kyushu Institute of Technology. Her research interests include data mining, anomaly detection and time series analysis.

Prof. Abbe Mowshowitz

Prof. Abbe Mowshowitz received the Ph.D. degree from University of Michigan in 1967. He has been professor of computer science at the City College of New York and member of the doctoral faculty at the Graduate Center of the City University of New York since 1984. His current research interests lie in two areas are organizational and managerial issues in computing, and network science. In addition to teaching and research, He has acted as consultant on the uses and impacts of information technology (especially computer networks) to a wide range of public and private organizations in North America and Europe.

Dr. Sakmongkon Chumkamon

Dr. Sakmongkon Chumkamon received Doctor of Engineering degree from Kyushu Institute of Technology in 2017. He was a postdoctoral researcher at Guangdong University of Technology in 2017-2019. Presently he is a postdoctoral researcher in Kyushu Institute of Technology since 2019. His research interests include factory automation robots and social robots.

Prof. Eiji Hayashi

Prof. Eiji Hayashi is a professor in the Department of Intelligent and Control Systems at Kyushu Institute of Technology. He received the Ph.D. (Dr. Eng.) degree from Waseda University in 1996. His research interests include Intelligent mechanics, Mechanical systems and Perceptual information processing. He is a member of The Institute of Electrical and Electronics Engineers (IEEE) and The Japan Society of Mechanical Engineers (JSME).