Time Series Analysis with Three Types of Noise-Mixing Effects by Neural Network

Takuya Nakamura, Yoshifumi Nishio and Yoko Uwate
Dept. of Electrical and Electric Engineering, Tokushima University
2-1 Minami-Josanjima, Tokushima, 770-5806, Japan
e-mail:{takuya, nishio, uwate}@ee.tokushima-u.ac.jp

Abstract— **One-dimensional convolutional neural networks are often used in time series analysis. However, when noise is mixed into the data, incorrect recognition is performed and discrimination is affected. Therefore, in this study, three types of color noise are mixed into the chaotic data and compared with the case where an autocorrelation function is used.**

Keywords; neural networks; time series analysis;

I. INTRODUCTION

Time series data can be found in a variety of phenomena such as temperature and seismic waves. Many studies have been conducted to analyze the characteristics of those time series. As a method to observe those characteristics, many time series have been converted into data and time series analysis using neural networks (NN) such as recurrent neural networks (RNN) and one-dimensional convolutional neural networks (1D-CNN) have been studied in recent years. However, the data often contains information that is unnecessary for analysis, called noise. The more of it there is, the more wrong perceptions are made, which hinders the analysis. In this study, three types of noise are mixed into time-series data, and the classification accuracy of neural networks is verified.

II. ONE-DIMENSIONAL CONVOLUTIONAL NEURAL NETWORK

Convolutional Neural Networks (CNN) is a type of NN with an additional convolutional and pooling layer. CNN is used in natural language processing and image recognition. In addition, they have a wide range of applications in the medical and automotive fields. However, CNN has mainly used for image recognition, they are computationally expensive and time consuming. Therefore, one-dimensional CNN (1D-CNN) can be used to reduce the computational cost. In addition, features can be automatically extracted through learning. As a model for the NN, I use a one-dimensional residual network (1d-ResNet). This is the model proposed by He at Facebook AI Research in 2015. Using this model prevents the gradient loss problem, which prevents learning from progressing in 1d-ResNet.

III. PROPOSED METHODS

In this study, the following method is used to confirm the accuracy of the classification of noise in the data. Noises are added to the original data and compare the accuracy between the original
data case and the case using Autocorrelation Function (ACF).

Step 1. The data to be trained is replaced with ACF from the original data.
Step 2. ACF is trained on the 1D-CNN.
Step 3. ACF is trained on the 1d-ResNet.

A. Autocorrelation Function

The autocorrelation function (ACF) is defined as the product of time t and data shifted by *k* from *t* in time series data. The correlation between the current data and the data shifted by *k* in the past is examined. ACF measures the relationships between a lagged version of itself over successive time intervals. ACF is expressed by the following Eq. (1).

$$\frac{\sum_{i=k+1}^{n}(x_t-\bar{x})(x_{t-k}-\bar{x})}{\sum_{i=1}^{n}(x_t-\bar{x})^2} \tag{1}$$

n is the sampling number and *k* is the time lag. Figure 1 shows the example of the original data. Figure 2 shows autocorrelation of the original data.

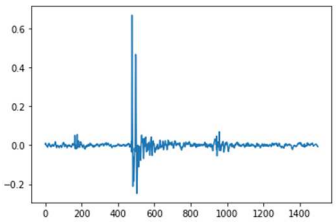

Figure 1. Original Data.

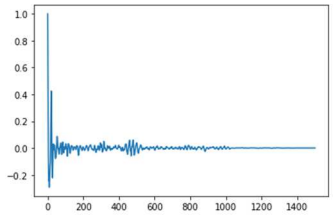

Figure 2. ACF of Figure 1.

B. Noise

In the field of time series analysis, noise refers to not necessary data that is not subject to analysis. The greater the amount of noise, the less accurate the analysis becomes. In this case, we used three types of noise: white noise, pink noise, and colored noise (red, blue and violet). These noises are added to the training data to train the 1d-ResNet.

IV. Verification Simulation Model

In this study, two types of classification accuracy are tested using original data for training data and using an autocorrelation function, and original data is used for test data. A dropout layer is also used to prevent over-training.

The 1d-ResNet is used as the classification model 1d-resnet is used to solve the gradient loss problem, which is a problem of multilayer networks, by learning residuals. Figure 3 shows the structure of residual learning. Convolution(1, 128) means convolution layer with 1 filter and 128 channel.

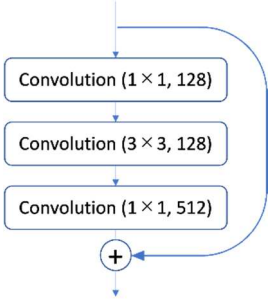

Figure 3. Structure of 1d-resnet

V. Dataset

The data used in this study are chaotic data generated by the Lorenz equation. These equations are shown the following Eq. (2).

$$\begin{cases} \dfrac{dx}{dt} = \sigma(y - x) \\ \dfrac{dy}{dt} = x(\sigma - z) \\ \dfrac{dz}{dt} = xy - \beta z \end{cases} \tag{2}$$

In this study, as the parameters, σ=10, β=8/3 are used, and the dataset has three classes that have different ρ, and recognize the parameters differences by using 1d-resnet. Table I shows the number of the train data and test data.

Table I. The Nunber of The Train Data and Test Data

Data	*ACF*	*Original*
Train Data	1800	1800
Test Data	1200	1200

VI. Simulation Result

Tables II and III show the classification accuracy of the ACF and 0.0 and 0.5 original data verified for three different noise rates. These tables show that the classification accuracy is higher when ACF is used for white and pink noise.

Table II. Test Accuracy of White Noise

Noise Rate	*Test Accuracy (Original) [%]*	*Test Accuracy (ACF) [%]*
0.0	65.35	65.91
0.5	63.37	64.98

Table III. Test Accuracy of Pink Noise

Noise Rate	*Test Accuracy (Original) [%]*	*Test Accuracy (ACF) [%]*
0.0	60.92	64.87
0.5	59.92	61.34

Tables IV, V, and VI show that when ACF is used for color noise, the classification accuracy for the three color noise patterns often decreases.

Table IV. Test Accuracy of Red Noise

Noise Rate	*Test Accuracy (Original) [%]*	*Test Accuracy (ACF) [%]*
0.0	66.57	65.17
0.5	62.30	60.20

Table V. Test Accuracy of Blue Noise

Noise Rate	*Test Accuracy (Original) [%]*	*Test Accuracy (ACF) [%]*
0.0	66.47	64.27
0.5	65.62	56.73

Table VI. Test Accuracy of Violet Noise

Noise Rate	*Test Accuracy (Original) [%]*	*Test Accuracy (ACF) [%]*
0.0	63.73	66.10
0.5	64.00	60.27

VII. Conclusion

In this study, three types of noise are mixed into the chaotic data. The results showed that ACF is effective in increasing the classification accuracy for white and pink noises. However, it is found to have effect on the three colored noise patterns. As a future work, we would like to search effective methods for colored noise.

References

[1] K. He, X. Zhang, S. Ren, J. Sun, “Deep residual learning for image recognition”, In Proceedings of the IEEE. Conference on computer vision and pattern recognition, pp.770-778, 2016.

[2] M. Matsugu, K. Mori, Y. Mitari and Y. Kaneda, “Subject independent facial expression recognition with robust face detection using a convolutional neural network”, Neural networks 16, vol.5, pp.555-559, 200