Chaotic Data Classification by Using Attractors Including Time Information

Aika Ohno, Yoko Uwate and Yoshifumi Nishio
Dept. of Electrical and Electronic Engineering of Tokushima University
2-1 Minami-Josanjima, Tokushima-shi, Tokushima-ken 770-8506, Japan
Email: {a-ohno, uwate, nishio}@ee.tokushima-u.ac.jp

***Abstract*—Chaotic data can be seen in various kinds of data such as biological signal. It is important to analyze chaotic data and know slight difference of chaotic parameters with better accuracy. As a method of chaos theory, time series data can be embedded into attractors which can be shown as images. In addition, recent years, neural network is used for image classification of various kinds of data.**

In this study, we classify chaotic time series data by using embedded attractors images and residual networks which is one kind of neural networks, and we propose embedded attractors that include time information for dataset. Then, we discuss the simulation results by using confusion matrixes.

I. Introduction

Recently, neural network has been got a lot of attention, and it achieves significant results especially on image recognition. It is used for various kinds of tasks such as image recognition, prediction, segmentation and natural language processing. Also, it is applied for a lot of kinds of data and fields not only academic fields but also industrial and medical fields. Image recognition by neural network is surpass human error rate [1]. Thereby, neural networks can recognize images even human cannot classify. Residual Networks (ResNet) is one of the most popular networks for image classifications [2]. ResNet is the winner network of ILSVRC 2015, and it is the networks that enable to have deep layered networks without vanishing gradient problem.

Chaotic characteristics can be observed in various kinds of time series data such as biological signals and stock price fluctuations, and chaotic theory is applied for these data analysis. Chaos means the signals that seems to be disorder, but actually it is deterministic. It has features sensitivity of initial conditions and non-periodicity. Analytics of pulse wave is one of the application example of chaos theory. Some studies show different feelings have different statement of chaos [3]. Then, proper recognition of the differences of chaos parameters enables to estimate their bodies and feeling conditions with better accuracies. Therefore, it is important to recognize chaotic features and classify depends on the differences of chaotic statements. However, as chaos seems to be disorder, it is difficult to recognize parameters differences by only seeing time series data. Therefore, Takens' embedding theorem is used, and some chaos characteristics that are difficult to be seen in time series data can be shown [4].

Classification of chaotic time series data by using reconstructed attractors with neural networks has been studied [5]. However, reconstructed attractors would be lost the time information. From that perspective, we investigate the effect of using attractor that include time information for dataset. By proposing this method, our study object is finding effective preprocessing for chaotic time series data for neural networks.

II. Chaos Theory and Embedded Attractors

By using Takens' embedding theorem, time series data $x(t)$ can be embedded into n-dimensional vector $v(t)$ with time delay τ as shown in Eq. (1).

$$v(t) = (x(t), x(t+\tau), x(t+2\tau), ..., x(t+(n-1)\tau)) \quad (1)$$

1-dimension (1D) time series data is converted into 3-dimensions (3D) time delay coordinate system, and it is shown in Fig. 1. This 3D object is called attractor.

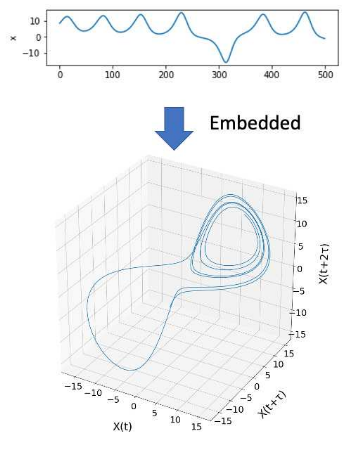

Fig. 1. Reconstructed 3D attractor from time series data.

III. Residual Networks

ResNet is commonly used for the classification of image data. In this study, ResNet50 that has 50 layers is used. Figure 2 shows the structure of ResNet50, and it is pretrained by ImageNet dataset. To make the figure simple, the repetitions of batch normalization and activation between convolution layers are omitted. Dotted lines indicate the process to increase dimensions with a stride of 2.

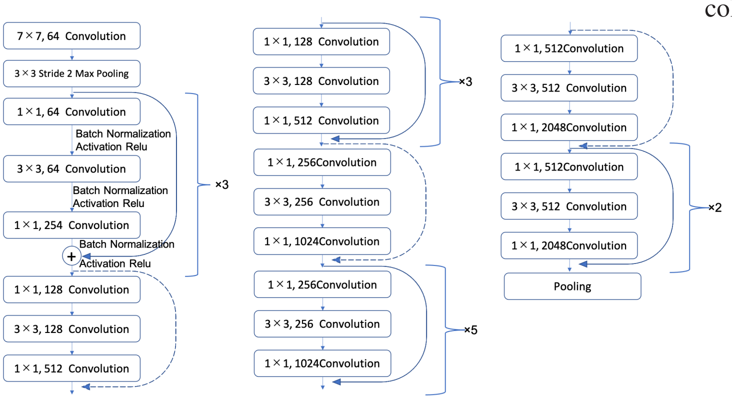

Fig. 2. The structure of ResNet50.

IV. Dataset

As a chaotic data, Lorenz data which is famous chaotic equations is used. These equations are shown the following Eq. (2).

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

In this study, as the parameters, $\sigma = 10$ and $\beta = 8/3$ are used, and the dataset has three classes $\rho = 27, 28, 29$. The data have different initial values. X-axis data is used, and the length of each time series data is 500 time steps from time step 1000 to 1500. These are embedded with time delay $\tau = 10$.

Train data and test data are divided randomly. Train data has 1800 data and test data has 1200 data in total of 3 classes, and each train and test data have the same ratio of 3 classes. Each image size is 224×224.

V. Proposed Method

By using Takens' embedding theorem, time information is lost. In our proposed method, time information will be added on reconstructed attractors.

3rd dimension's information $x(t + 2\tau)$ is replaced by time information t. Equation (3) shows the vector when time series data is embedded into 3-dimension vector.

$$v(t) = (x(t), x(t + \tau), x(t + 2\tau)) \tag{3}$$

Equation (4) shows the vector when includes time information that we propose.

$$v(t) = (x(t), x(t + \tau), t) \tag{4}$$

Both of the 3rd dimensions of the vectors are expressed by coloring of plot as shown in Fig 4. Figure 3 shows normal 3D plot attractor with color plotting, and Fig. 4 shows proposed 3D plot attractor that includes time information. For easier understanding, axes and labels are shown in Fig. 3 and 4, but actual images of dataset does not include axes and labels and color bars. Dataset images have only attractors.

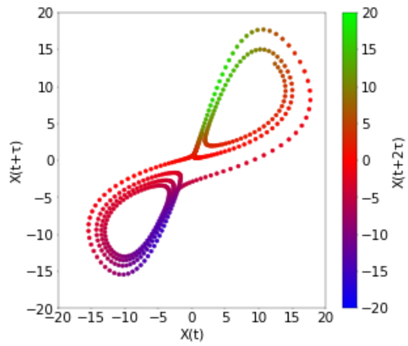

Fig. 3. The reconstructed 3D plot attractor.

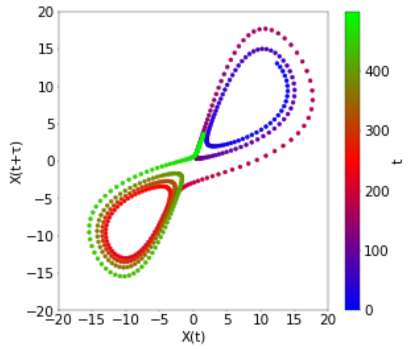

Fig. 4. The reconstructed 3D plot attractor includes time information.

Figure 5 and Fig. 6 shows the another 3D line attractors. It is drawn in 3D phase space which is different from aforementioned 3D plot attractors that express 3rd dimension by coloring of plots. Figure 5 shows normal 3D line attractor, and Fig. 6 shows proposed 3D line attractor includes time information.

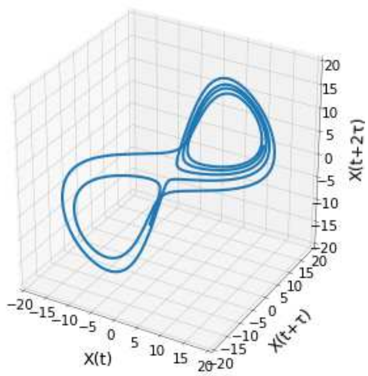

Fig. 5. The reconstructed 3D line attractor.

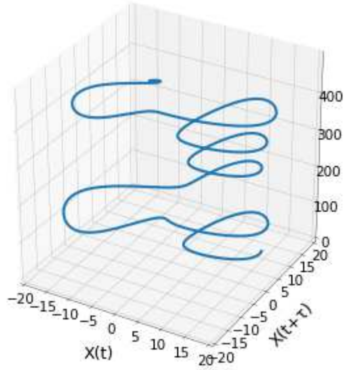

Fig. 6. The reconstructed 3D line attractor includes time information.

VI. Simulation Results

Batch size is 16, and reduce learning rate method is used. Initial learning rate is 0.0001.

Table 1 shows the simulation results of 3D plot attractor dataset, these are the average score of 3 times simulations.

TABLE I
Simulation result of 3D plot attractor dataset

	Train accuracy	Test accuracy
3D plot attractors	0.9250	0.8547
Proposed method	0.9570	0.8845

The proposed method shows better accuracy than 3D plot attractors dataset in both of train and test accuracies.

Figure 7 shows the confusion matrix when 3D plot attractors dataset was used. Confusion matrix shows how many images are predicted which labels, and it shows also which actual labels.

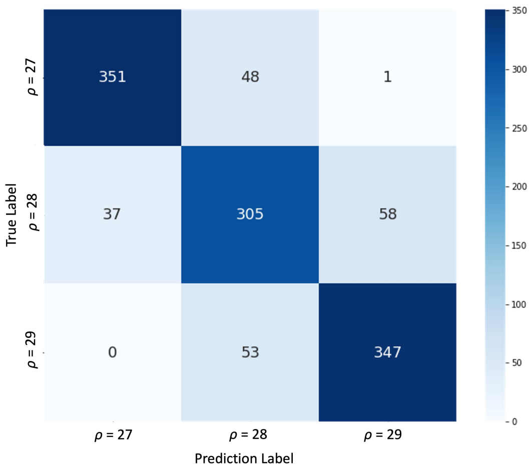

Fig. 7. The confusion matrix of 3D plot attractors.

From Fig. 7, there is more mis-labeled images on related $\rho = 28$. Categories of true $\rho = 28$ predicted $\rho = 29$ and true $\rho = 29$ predicted $\rho = 28$ has more mis-labeled than others.

Figure 8 shows the confusion matrix of proposed method 3D plot attractors dataset include time information.

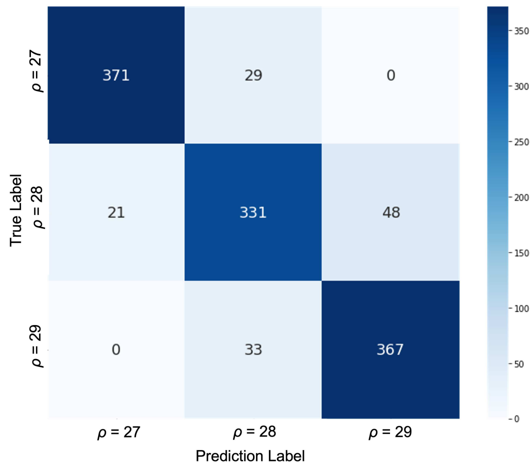

Fig. 8. The confusion matrix of 3D plot attractors include time information.

There is more mis-labeled images on related $\rho = 28$ which is the same tendency with 3D plot attractors dataset. The most frequent mis-labeled category is true $\rho = 28$ predicted $\rho = 29$.

Figure 9 shows differential matrix between 3D plot attractors dataset and proposed method. Differential matrix Cm is calculated by Eq. (5). $Cm1$ is the confusion matrix of 3D attractors dataset, and $Cm2$ is the confusion matrix of proposed 3D attractors include time information dataset.

$$Cm = Cm1 - Cm2 \tag{5}$$

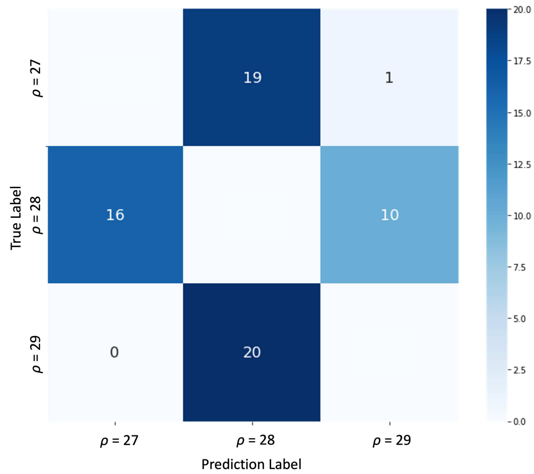

Fig. 9. The differential matrix of two kinds of 3D plot attractors dataset.

From Fig 9, the mis-labeled categories true label $\rho = 27$ predicted $\rho = 28$ and true label $\rho = 29$ predicted $\rho = 28$ have improved well compared with other mislabeled categories. On the other hand, true label $\rho = 28$ predicted $\rho = 29$ has less improvement than others, even though it has more mis-labeled images in both datasets.

Table 2 shows the simulation results of 3D line attractors dataset, these are the average score of 3 times simulations.

TABLE II
SIMULATION RESULT OF 3D LINE ATTRACTORS

	Train accuracy	Test accuracy
3D line attractors	0.9733	0.9292
Proposed method	0.9818	0.9739

Train accuracy of the proposed method is slightly better, and test accuracy of the proposed method shows better result than original 3D line attractors dataset.

Figure 10 shows the confusion matrix when 3D line attractors dataset was used.

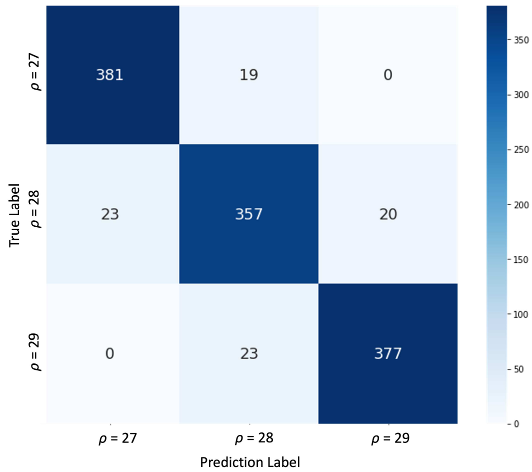

Fig. 10. The confusion matrix of 3D line attractors.

There is more mis-labeled images on $\rho = 28$ as well as 3D plot attractors dataset. 4 categories related on $\rho = 28$ have almost same amount of mis-labeled images.

Figure 11 shows the confusion matrix when 3D attractors line that include time information dataset was used.

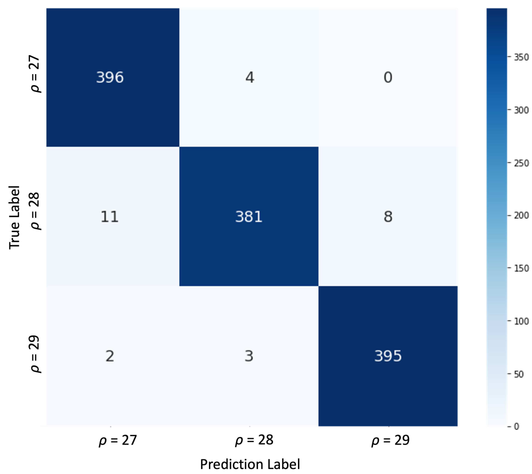

Fig. 11. The confusion matrix of 3D line attractors include time information.

Figure 12 shows differential matrix between 3D line attractors dataset and proposed method. As same as the aforementioned differential matrix, it is calculated by Eq. (5).

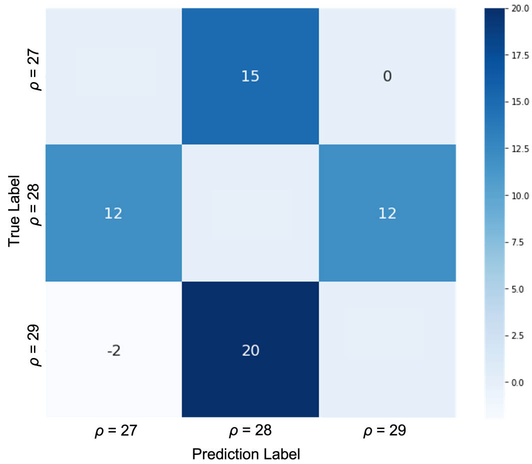

Fig. 12. The differential matrix of two kinds of 3D plot attractors dataset.

From Fig 12, the mis-labeled categories true label $\rho = 27$ predicted $\rho = 28$ and true label $\rho = 29$ predicted $\rho = 28$ have improved well compared with other mislabeled categories. This is the same tendency with 3D plot attractors dataset.

VII. CONCLUSIONS

In this study, we classified chaotic time series data Lorenz data by ResNet, and we proposed the method that use 3D attractors includes time information. Our proposed method showed better accuracy than normal 3D embedded attractors on both way of 3D plot attractors and 3D line attractors. Then, we discussed the simulation result by using confusion matrixes, and it is found that the mis-labeled categories true label $\rho = 27$ predicted $\rho = 28$ and true label $\rho = 29$ predicted $\rho = 28$ have improved well by our proposed method. For our future works, we would like to study other preprocessing method to improve accuracy.

REFERENCES

[1] He, Kaiming, et al. "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification." Proceedings of the IEEE international conference on computer vision. 2015.

[2] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[3] Oyama-Higa, Mayumi, et al. "Development of a Self-Check System for Mental Health Using a Pulse Wave Mouse." ICSOFT (ISDM/EHST/DC). 2007.

[4] K. Aihara, T. Ikeguchi, Y. Yamada and M. Komuro, "Fundamentals and applications of chaotic time series analysis", Industrial book, p.23, 2000.

[5] Chen, Yen-Lin, et al. "High-Dimensional Phase Space Reconstruction with a Convolutional Neural Network for Structural Health Monitoring." Sensors 21.10 (2021): 3514.