Biological Time Series Analysis by Neural Network Using Attractor Features

Ryosuke Shimizu, Tomiyuki Furugaki , Yoko Uwate and Yoshifumi Nishio
Dept. of Electrical and Electronic Engineering, Tokushima University
2-1 Minami-Josanjima, Tokushima 770–8506, Japan
Email:{shimizu, furugaki, uwate, nishio}@ee.tokushima-u.ac.jp

***Abstract*—Time series classification is an important and challenging problem in data analysis. Recently, time series analysis using neural networks has attracted much attention. However, the analysis of time series data with complex oscillations is difficult. Therefore, it is important to search for effective features of the data. In this study, we transform the dimensionality of the data and search for features suitable for NN classification. In this study, we investigate the effect of dimensionality reduction methods on accuracy and the characteristics of the data.**

I. Introduction

In recent years, there has been a great deal of research on the analysis of time series data. This real-world data has various characteristics, such as not only periodic oscillation but also random oscillation. Chaos theory deals with data that obey deterministic laws. However, in order to observe chaos, it is necessary to transform the data. One of the methods is the time-delay coordinate system, which is a kind of chaos theory [1]. Using this method, one dimensional data (1d data) can be transformed into multi-dimensional data called attractors. Many phenomena in the real world are represented by attractors, which can be treated as mathematical models. Attractors are also applied to systems, physical phenomena, and economic phenomena in the living body [2]. In recent years, there is a lot of research on time series data analysis using neural networks (NN) such as 1-dimensional convolutional neural networks (1d-CNN) and recurrent neural networks (RNN). The advantage of NN is that they can capture features by themselves. However, even with these models, it is difficult to learn numerical and time-series features of time series data with very irregular oscillations. Therefore, data preprocessing is important in NN. In this study, to solve this problem, the data is extended to 3D data using a time-delay coordinate system and various dimensional reduction methods to compress the data into 1D data. The objective is to find features that can be easily learned by NN from time series data with irregular oscillations.

II. Convolutional Neural Network

Convolutional Neural Network (CNN) is one of the NN and is mainly used in image recognition [3]. Based on the hierarchical processing mechanism of information in visual cortical pathways, Fukushima and Miyake proposed a self-organizing hierarchical network called “ neocognitron ”, the predecessor of CNN, in 1982 [4]. However, CNN are required high computational power. To compensate for these shortcomings, in 2015, Kiranyaz proposed the first compact and adaptive 1d-CNN that works directly with patient-specific ECG data [5]. Today, 1d-CNN have quickly achieved state-of-the-art performance levels in several applications such as biomedical data classification and early diagnosis, structural health monitoring, anomaly detection and identification in power electronics and electrical [6]-[8]. CNN is a good method for learning features in a local frequency band. The CNN consists of an input layer, an output layer, and several hidden layers in between them. Usually, it consists of a convolutional layer, a pooling layer, and a fully connected layer (Fig. 1).

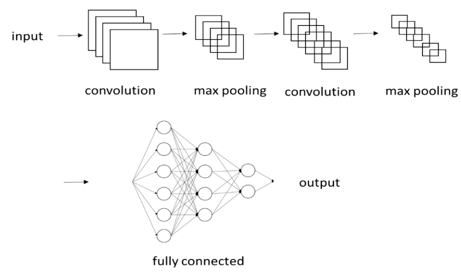

Fig. 1: An example of CNN structure.

III. Dataset

The data from the UCR repository [9] used in this study is an electromyogram (EMG) of hand movements. The free and repeated grasping of various objects necessary to move the hand was measured. The speed and force of the grip is intentionally left to the subject. To collect information on muscle activation, two forearm surface EMG electrodes (flexor, extensor, radial, long radial, and short radial muscles [10]) is fixed with a rubber band and a reference electrode was placed in the middle. For data collection, five healthy subjects (two males and three females) of approximately the same age (20 to 22 years old) are asked to repeatedly perform the following six movements considered to be basic hand movements. Five healthy subjects (two males and three females) of approximately the same age (20 to 22 years old) are asked to repeat the following three movements that are considered to be basic hand movements used in Fig. 2 [11]. Data (a) is Spherical for holding spherical tools. Data (b) is Tip for holding small tools. Data (c) is Palmar for grasping with palm. An example of EMG time series data is shown in (Fig. 3). The signals is band-pass filtered using a Butterworth Band Pass filter with low and high cutoff 15Hz and 500Hz respectively and a notch filter at 50Hz to eliminate line interference artifacts.

Fig. 2: Images of the three hand movements.

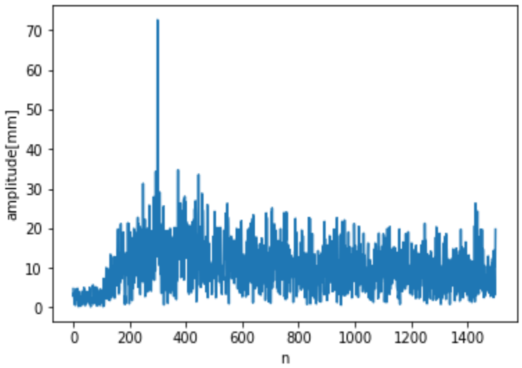

Fig. 3: Electromyogram of hand movement.

IV. Proposed Method

In this study, we propose a method of feature translation for using multidimensional features of time series data. Figure 4 shows the flow of the proposed method of this study.

1) The data is transformed into three dimensional data using time delay coordinate system by Takens' Embedding Theorem.

2) The expanded data is compressed using various dimensionality reduction methods. This allows us to extract the necessary features.

(a) Attractor construction

Time-delayed coordinate systems are commonly known as dimensional dilation methods by Takens ' Embedding Theorem. It is a method of dilation of dimensions. Let the value of the data at a certain time be $x(n)$. Further, let the value of time delay be τ. This system is represented by Eq. (1) and is shown in Fig. 5.

$$f(x) = [x(n),\ x(n+\tau),\ x(n+2\tau)\ ...] \tag{1}$$

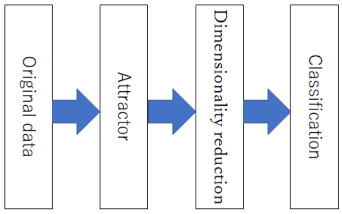

Fig. 4: Proposed method.

In this study, we extended the data to three dimensions.

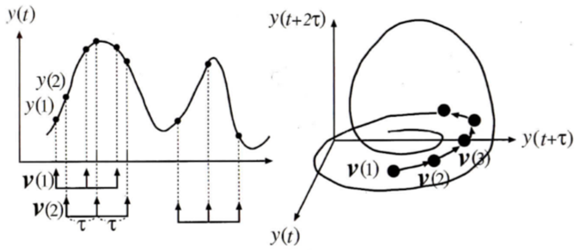

Fig. 5: Time delay coordinate system.

(b) Dimensionality reduction

Dimensionality reduction refers to the reduction of high-dimensional feature vectors to low-dimensional vectors while retaining the distribution information of the original dataset. Dimensionality reduction methods can be broadly divided into linear and non-linear methods. Linear methods are good at representing "global" data structures in low dimensions, while non-linear methods are good at representing "local" structures. In this study, two linear methods and one nonlinear method are used. The procedure of three methods is shown in the following steps.

Principal Component Analysis (PCA)

1) Determine the axis that can maximize the variance
2) Determine the axis of the second that is orthogonal to the first
3) Get the subspace as eigenvectors by singular value decomposition
4) Compute lower-dimensional embedding.

PCA is one of the most widely used low-dimensional embedding methods. As the name implies, it is an algorithm that searches for the important element principal component in the observed data.

Non-negative Matric Factorization (NMF)

1) Prepare a non-negative matrix.
2) Decompose it into the inner product of two non-negative basis matrices and a non-negative weight matrix.

3) Use only the non-negative basis matrix to compress the dimensionality.
4) Compute lower-dimensional embedding.

NMF is a mathematically very simple formulation of decomposing a matrix consisting only of non-negative values, yet it can automatically enumerate frequently occurring patterns in sound, image documents, etc.

Local Linear Embedding (LLE)

1) Select K nearest neighbors from data points
2) Consider a linear combination in a neighborhood of K
3) Approximate a point with a linear combination
4) Compute lower-dimensional embedding.

Linear modeling of local relationships can be performed while maintaining the relationships in the original data.

V. Simulation Model

In this study, the 1d-CNN is used for classification model. Figure 6 shows the structure of the 1d-CNN.

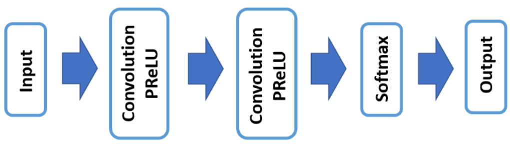

Fig. 6: 1d-CNN structure used in this study.

The 1d-CNN consists of dozens of layers, where each layer is trained to detect different features of the 1d-data. For each training data, multiple filters with different resolutions are applied, and the output of each convolutional data is used as input for the next layer. With repeated training, the filters increase the level of feature complexity. The convolution of a 1d-CNN is expressed by Eq. (2). n is the length of the convolution filter. k is the number of the filter. x is the input image data. In this study, the number of the first filter is 32 and the number of the second filter is 64.

$$f(x) = \sum_{t=0}^{n-1} w_t^{(k)} x_{j+t} + b^{(k)} \tag{2}$$

The pooling layer is removed for speedup. An activation function, Parametric Rectified Linear Unit (PReLU), is added after the convolutional layer to prevent overtraining [12]. The negative gradient a $(0 < a < 1)$ is adjusted by network learning. The PReLU function is expressed in Eq. (3).

$$f(x) = \begin{cases} x & (x \geq 0) \\ ax & (x < 0) \end{cases} \tag{3}$$

The classification probability is derived by the softmax function expressed in Eq. (4). p is the probability of becoming class j, x is the output of the NN, and n is the total number of discriminant classes. In this study, the total number C of discriminant classes is 3. For all classification, $f(x)$ satisfies $0 < f(x) < 1$.

$$f(x) = \frac{exp(x_j)}{\sum_{i=0}^{n} exp(x_i)} \tag{4}$$

Table I shows the number of the EMG time series data in this study. Table II shows the batch size, epochs, and learning rate in this study. Batch size is the amount of data processed at once. The epoch is the number of the learnings. The learning rate is the degree of learning progress. By setting as shown in Table II , the training accuracies reach 100%.

TABLE I: The number of the data.

	data1	data2	data3
train data	250	250	250
test data	50	50	50

TABLE II: Learning parameters.

	Batch size	Learning rate	Epoch
1D-CNN	42	1e-3	25

VI. Simulation results

We investigate the average of 10 times of test accuracy. Table III shows the test accuracy of the original data expanded to three dimensions and compressed by the proposed method. The test accuracy of the data expanded to three dimensions and compressed by NMF is the best. It can be seen that the other test accuracies are also better than the accuracy of the original data. Even if the original data is extended to more than three dimensions, the original characteristics can be maintained.

TABLE III: Test accuracy of the proposed method.

	test accuracy
original	0.888
PCA	0.934
NMF	0.965
LLE	0.912

Considering the characteristics of each method, it can be considered that the features are coherent and decomposing them into matrices made it easier to grasp the features. In addition, when comparing the nonlinear LLE method and the linear method, the linear method has better accuracy, which suggests that the structure of the transformed data is a global manifold. The data in this study consisted of non-negative data only, and the features are properly extracted in dimensionality reduction using NMF.

VII. Conclusion

In this study, we extended time series data to 3D space using time-delay embedding method and compressed the data using various reduction methods. we checked how the accuracy of time series classification by 1d-CNN changes by extending 1D data to 3D data. The results show that the accuracy of the test using the proposed method is better than that of the test using the conventional method. In addition, we found that there is an important relationship between the test accuracy and the characteristics of the data structure. Using this relationship, we search for the best method in various data. We will also use models such as RNN that are specialized for analyzing the flow of time series, or construct optimal extended dimensions.

References

[1] F.Takens, Detectingstrange attractors intur bulance,LectureNotes in Mathematics, Vol.898(Springer-Verlag. 1981) 366-381.

[2] Christian Oestreicher,"A history of chaos theory", doi: 10.31887/DCNS.2007.9.3/coestreicher

[3] 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, 2003.

[4] K. Fukushima, S. Miyake, Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position, Pattern Recognit. 15 (1982) 455469. doi 10.1016/0031-3203(82)90024-3

[5] S. Kiranyaz, T. Ince, R. Hamila, M. Gabbouj, Convolutional Neural Networks for patient specific ECG classification, in: Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS, 2015. doi:10.1109/EMBC.2015.7318926

[6] Shraddha R. Bajare, Vaishali V. Ingale, "ECG Based Biometric for Human Identification using Convolutional Neural Network", Computing Communication and Networking Technologies (ICCCNT) 2019 10th International Conference on, pp. 1-7, 2019.

[7] Pragya Sharma, Swet Chandan, B. P. Agrawal, "Vibration Signal-based Diagnosis of Defect Embedded in Outer Race of Ball Bearing using 1-D CNN", Computational Performance Evaluation (ComPE) 2020 International Conference on, pp. 531-536, 2020.

[8] Imene Mitiche, Alan Nesbitt, Stephen Conner, Philip Boreham, Gordon Morison, "1D-CNN based real-time fault detection system for power asset diagnostics", Generation Transmission Distribution IET, vol. 14, no. 24, pp. 5766-5773, 2020.

[9] Yanping Chen, Eamonn Keogh, Bing Hu, Nurjahan Begum, Anthony Bagnall, Abdullah Mueen and Gustavo Batista (2015). The UCR Time Series Classification Archive.

[10] E. F. Delagi, A. O. Perotto, J. Iazzetti, and D. Morrison, Anatomical guide for the Electromyographer: The Limbs and Trunk, Charles C Thomas Pub Ltd, 5ht ed., 2011

[11] T. S. Saponas, D. S. Tan, D. Morris, R. Balakrishnan, J. Turner, and J A. Landay, Enabling Always-Available Input with Muscle-Computer Interfaces´ in 22nd annual ACM symposium on User interface software and technology, Association for Computing Machinery, New York, USA, pp 167-176, 2009

[12] K. He, X. Zhang, S. Ren and J. Sun, "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification ", IEEE International Conference on Computer Vision and Pattern Recognition, arXiv:1502.01852, 6 Feb 2015.