2020 RISP International Workshop on Nonlinear Circuits,
Communications and Signal Processing (NCSP 2020)
Honolulu, Hawaii, USA, February 28 - March 2, 2020

Effective Data Augmentation Method with Sequence Numbers from Chaos Phenomenon for Convolution Neural Network Learning

Yuji Yamauchi, Yuichi Miyata, Yoko Uwate and Yoshifumi Nishio

Dept. of Electrical and Electronic Engineering, Tokushima University
2-1 Minami-Josanjima, Tokushima 770-8506, Japan
Email: {yamauchi, y.miyata, uwate, nishio}@ee.tokushima-u.ac.jp

Abstract

Convolution Neural Network (CNN) exhibits excellent performance in image recognition, and is widely used as a base network for solving various tasks such as segmentation and various object detection. However, there are still many fields where it is difficult to collect enough data to learn CNN. An effective technique for learning with such a small amount of data is data augmentation.

In this study, image classification is performed for fields where it is difficult to collect sufficient data. At this time, we propose to perform image processing is performed using a sequence obtained from chaos phenomenon to augment the data. The proposed method is evaluated by comparing the learning accuracy of the model trained with the augmented data after image processing using random numbers and the test accuracy of image classification based on the model.

1. Introduction

A neural network that can acquire the features of an image group by learning a large amount of labeled learning data, and can cope with unknown data. Among them Convolutional Neural Network (CNN) is consisted of "convolution layers" that extracts features such as edge and a "pooling layers" that provides robustness to the extracted features to a convention neural network in Fig. 1 [1]. The Convolution layer can maintain spatial information. The output obtained by processing the feature points with the filter from the original image becomes one of the units of the next layer. The pooling layer is usually applied after the Convolution layer. To transform the input data into a more manageable form, the information is compressed and down-sampled. There are effects such as being robust against minute position changes, suppressing over-learning to some extent, and reducing the calculation cost. The units in each layer of the fully connected layer are connected to all units in the next layer. It is often used before the output layer, and this part is an identification part for classifying into the prediction result by the combination of detected features. The output value in the output layer is the probability that the category is predicted. The term "robust" as used herein means that the image position is not affected. This model exhibits excellent performance mainly in image recognition, and is also widely used as a base network for solving various tasks such as segmentation and various object detection.

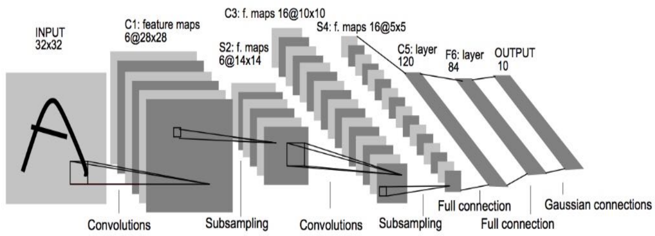


Figure 1: LeNet in CNN with typical structure.

Machine learning has greatly developed various fields such as natural language analysis [2], sound recognition, reinforcement learning, and so on by the recent development of CNN. For instance, application such as detecting and labeling a feature included in an image, or converting human voice into text can be cited. In the fashion industry, the Electronic Commerce (EC) market has expanded due to computer learn trends from a large amount of data accumulated every day, and analyzing product images from each person's purchase history to find products that suit those people. However, EC in a state where it is impossible to add value to vintage clothing, such as materials and manufacturing techniques that are rare now, and values that arise from the state changes over time. For this reason, vintage secondhand clothing cannot enter the EC market so much, and there is little amount of learning data is needed for CNN to learn. When learning such a small amount of data, there are three methods of using high quality data, augmentation, and transfer learning. Therefore, in this study, we focus on the augmentation of data and aim to improve the learning accuracy and image classification test accuracy by the new augmentation method of a small amount data of image data.

2. Date augmentation

Date augmentation is a technique that increases the amount of data by converting the original data. Conversion of images include increasing noise and adjusting contrast and brightness, and so on. For example, when you rotate a plant image, you can recognize that it is an image of a plant with the same name. In other words, it is not necessary to change the label for the image. In addition, there are actually photographs from various angles, which helps to increase robustness. The quality of images is not constant, such as shooting from various angles or shadows of other objects, so an approach that improves robustness is important [3].

3. Logistic map

Determinism is that "giving an initial value inevitably gives future behavior". By combining non-linearity and deterministic rules, a phenomenon in which future behavior cannot be assumed even with a very small initial value error is a chaos phenomenon.

The logistic map [4] is one of the simplest systems that generate chaos, and shown by the following one-variable quadratic difference equation. x represents a variable, and α represents a branch parameter.

$$x_{n+1} = \alpha x_n(1 - x_n). \tag{1}$$

Where n is the number of steps (n=0,1,2...). It means the development of discrete time and returns a real value between 0 and 1 by giving an initial value.

The change of the value of x_n depends on what value is given to the parameter α. The value of α is taken between 2 and 4. When the parameter α is between 2 and 3, x_n converges to a fixed point. When parameter α exceeds 3, the convergence destination of x_n is not a fixed point but a periodic solution of period 2. If parameter a is further increased, the period of the periodic solution at the convergence destination doubles to 4, 8, and 16, and finally a chaotic time change appears in Fig. 2.

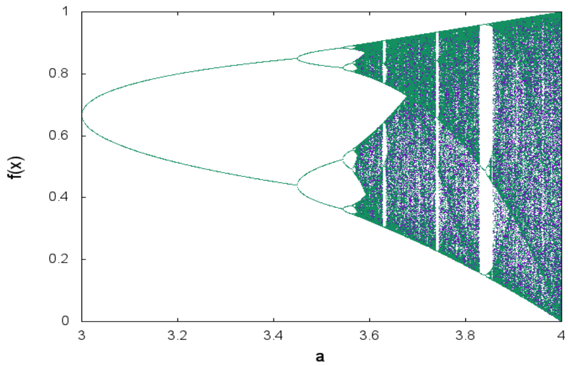
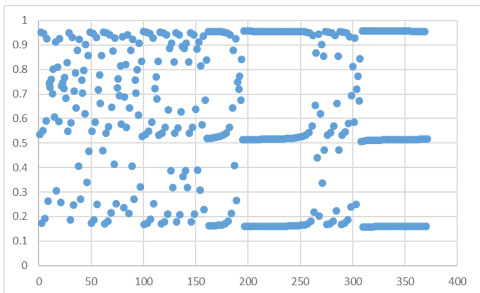
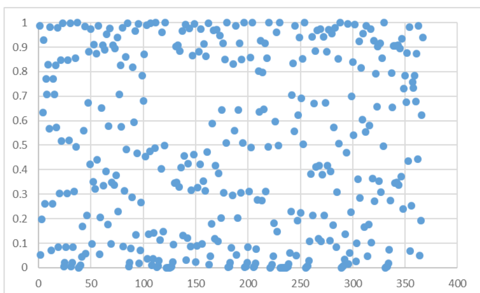


Figure 2: Bifurcation diagram of logistic map.

- Intermittent chaos (α =3.83)
  The time domain in which behaves irregularly and the time domain in which periodic behavior of 3 period occur are alternately generated in Fig. 3(a).
- Pure chaos (α =4.00)
  The logistic map has all natural number k periodic orbits, but all of them are unstable periodic orbits, and the orbits keep going around the interval [0,1] without asymptoticing to the periodic orbits.
  The sequence of pure chaos arranged in chronological order is determined according deterministic regular rules, unlike completely random numbers in Fig. 3(b).



(a) Intermittent chaos (α =3.83)



(b) Pure chaos (α =4.00)

Figure 3: Time series of intermittent and pure chaos.

4. Proposed method

In this study, the images are rotated when augmenting the data [5]. We use a sequence derived from the chaotic phenomenon and random number. The numbers in the sequence are from 0 to 1. These numbers are turned into angles. The image is rotated using the obtained angle in Fig.4.

The training data is a vintage t-shirt images, even among vintage clothing. As original data, 200 images of 70s, 80s and 90s vintage t-shirts are collected. This original data was augmented by 600 each using random numbers or the value obtained from chaos on the method that is the images are rotated, and a total data of 2400 images were obtained in Table 1.

(a) Original date (b) A data after processing

Figure 4: Image processing method.

Table 1: Learning data set.

Data set name	The number of data
Original data	200 images each (600 images in total)
Random	600 images each + original data (2400 images in total)
Intermittent chaos	600 images each + original data (2400 images in total)
Pure chaos	600 images each + original data (2400 images in total)

In this study the CNN structure is used in Fig. 5. The structure has two convolution layers, two pooling layers, and two layer fully connected layers. Processing in the convolution layer is performed by calculating an inner product for each filter, adding a bias, and applying an activation function.

The softmax function is used as the activation function. Softmax functions are commonly used in classification problems. It is a real number from 0 to 1 depending on the size of each input value, and the sum of output values is always 1. This means that the output of a softmax function can be treated as a "probability". The expression of the softmax function is expressed Eq. (2) in Fig. 6.

$$y_k = \frac{exp(a_k)}{\sum_{i=1}^{n} exp(a_i)} \tag{2}$$

In this study, we use a Spatial Pyramid Pooling (SPP) as a pooling layer. Max pooling is done in SPP which is the operation to select the largest one for a small area. In particular, SPP can obtain a fixed-length output regardless of the input image size. Furthermore, the test accuracy of iamge classification by using CNN of various structures can be improved by using SPP in the pooling layer. The convolution layer and Pooling layer function to detect features.

As with the Convolution layer, the output unit value of the fully connected layer is calculated by adding a bias to the inner product of each input unit value and the connection weight. In addition, dropout is applied to all bonding layers. At some update time, some nodes in the layer are disabled and CNN learning is performed. This improves generalization performance and avoids over-learning by forcing decrease in the degree of freedom of the network being learned [6].

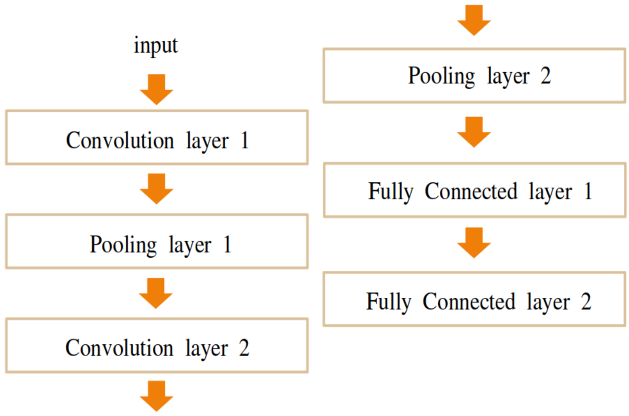


Figure 5: CNN structure used in this paper.

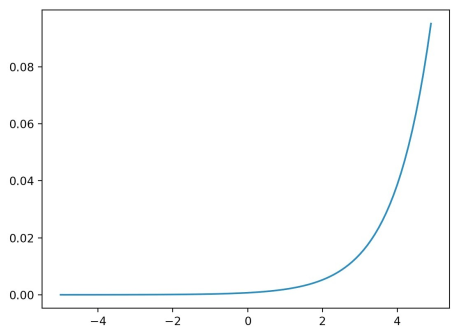


Figure 6: Softmax function.

CNN learns each data set. We prepare 75 images of 70s, 80s, and 90s vintage t-shirts as the test data different from the original data. The test data is input to each CNN, and the T-shirt of the input image is classified into 70s, 80s, and 90s. We compare and evaluate the learning accuracy and test accuracy of image classification by CNN trained on each data sets.

5. Simulation result

Figure 7 shows the learning accuracy and step on CNN. All the augmented datasets show better results than the original data set. The learning accuracy when learning the data augmented using the sequence obtained from the pure chaos reach 100% learning accuracy faster about 20 steps than the learning accuracy when learning the data augmented using random numbers.

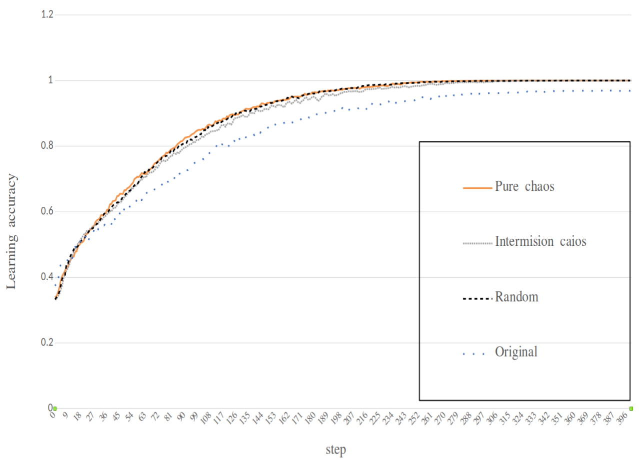


Figure 7: The accuracy of image classification.

Table 2 shows the test accuracy of image classification when CNN learns each data set. The CNN that learns the augmented data using the sequence obtained from pure chaos shows the highest test accuracy.

Table 2: Test accuracy.

	Test accuracy (%)
Original data	48.8637
Random	52.2727
Intermittent chaos (α =3.83)	51.1364
Pure chaos (α =4)	56.8182

However, for the both learning rate and test accuracy, the method using intermittent chaos show a lower result than that using random numbers.

We consider that the differences between the properties of pure chaos and random is related. Pure chaos is a non-periodicity that is not completely random like random numbers and that is not periodic, further has a characteristic that is never takes the same value.

6. Conclusion

In this study, we used data from an area where it is difficult to collect enough data for CNN learning.. The data is augmented by image processing using sequences derived from chaotic phenomena. In addition, image classification was performed by letting CNN learn the augmented data. The proposed method was evaluated by comparing the learning accuracy of the CNN trained with the data augmented by image processing using random numbers and the accuracy of image classification by using the CNN.

For the both learning accuracy and test accuracy, the data augmented used the sequence obtained from the pure chaos showed better results than the the augmented data using random numbers. Therefore, it is effective to combine a sequence obtained from chaos for data expansion. We consider that the important thing here is pure chaos is a non-periodicity that is not completely random and not periodic, further has a characteristic that is never takes the same value.

References

[1] Yann LeCun, Leon Botton, Yoshua Bengio and Patrick Haffner, "Gradient-Based Leaning Applied to Document Recognition," PROC. OF THE IEEE, 1998.

[2] Collobert, Ronan, Weston and Jason, "A Unified Architecture for Natural Language Processing: Deep Neural Networks with Multitask Learning," Proceedings of the 25th International Conference on Machine Learning. ICML '08,2008.

[3] Luke Taylor and Geoff Nitschke, "Improving Deep Learning using Generic Data Augmentation," Submitted on 20 Aug 2017.

[4] Robert M. May, "Simple mathematical models with very complicated dynamics," Nature, vol.261, 459-467 1976.

[5] Yu-Dong Zhang, Zhengchao Dong, Xianqing Chen, Wenjuan Jia, Sidan Du, Khan Muhammad and Shui-Hua Wang, "Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation," Multimed Tools Appl 2019.

[6] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever and Ruslan Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," The Journal of Machine Learning Research, Volume 15 Issue 1, January 2014.