The Effect of Preprocessing with Gabor Filters on Image Classification Using CNNs

Akito Morita
Graduate School of Information Science and Technology, Osaka Institute of Technology,
1-79-1 Kitayama, Hirakata, Osaka 573-0196, Japan

Hirotsugu Okuno
Faculty of Information Science and Technology, Osaka Institute of Technology,
1-79-1 Kitayama, Hirakata, Osaka 573-0196, Japan
E-mail: hirotsugu.okuno@oit.ac.jp
www.oit.ac.jp

Abstract

In image classification tasks, preprocessing of input images is one of the promising approaches for improving the performance. In this study, we investigated the effect of neuro-inspired preprocessing, such as Gabor filtering. We compared the averaged classification accuracy of multiple CNNs with the following three types of preprocessing: no preprocessing, Gabor filtering, and calculation of the difference between two Gabor filtered signals in the opposite color channels. The results showed that Gabor filtering increased the classification accuracy.

Keywords: convolutional neural network, Gabor filter, image classification, bio-inspired

1. Introduction

Convolutional neural networks (CNNs) are one of the most useful components in the field of image classification. In the CNNs that receive RGB images, features required for classification are extracted in the hidden layers. For example, Krizhevsky et al. showed that some of the kernels of the first layer in their CNN were selective to a particular image feature, such as spatial frequency, orientation, and color.[1] Using an input image in which features are extracted in advance could improve the accuracy and shorten the time required for learning. Actually, several models that combine CNNs and feature extraction with Gabor filters have been proposed.[2,3]

On the other hand, in the mammalian visual nervous systems, by which CNNs were inspired, the spatial characteristics of a type of neurons found in the primary visual cortex are modeled as Gabor filters.[4] Applying the processing in the early visual cortex to preprocessing of CNNs is expected to improve the performance of CNNs because the sophisticated visual functions in the brain, such as classification, are achieved by using the signals preprocessed by the primary visual cortex, which is the front end of the visual signal processing in the brain.

The purpose of this study is to investigate the effects of neuro-inspired preprocessing, such as Gabor filtering and calculation of the difference between opposite color signals, on CNNs. We compared the classification performance of CNNs with the following three types preprocessing: no preprocessing, Gabor filtering, and calculation of the difference between two Gabor filtered signals in the opposite color channels. In addition, in order to remove the influence of the topology of a particular CNN, we evaluated the performance in terms of the average value of the results of 50 CNNs whose parameters were determined randomly.

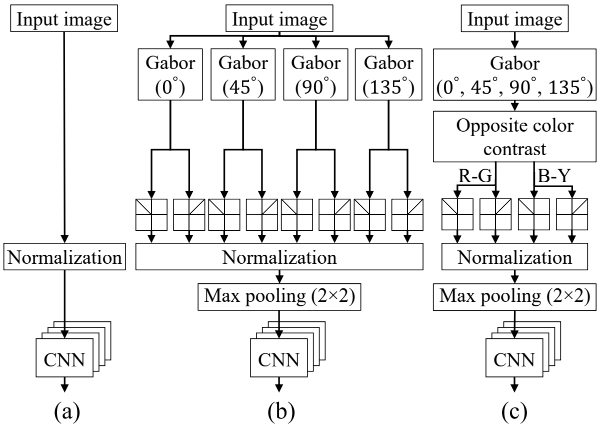

Fig. 1. Processing flow of classification. (a) Processing flow without preprocessing. (b) Processing flow with Gabor filtering. Four types of Gabor filters that enhance the following orientations were used: 0, 45, 90, and 135 degrees. The rectangles above the normalization represents rectification. (c) Processing flow with the calculation of the difference between two Gabor filtered signals in the opposite color channels. The processing to achieve the opposite color contrast is applied to the Gabor filtered signals of each orientation separately.

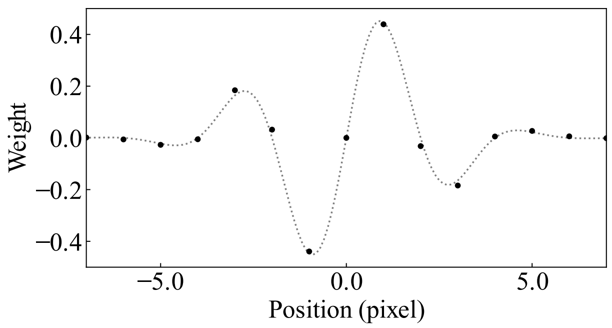

Fig. 2. Weights of the Gabor filter used in this study at $y = 0$. The black dots represent the weights of the Gabor filter, and the dotted line plots the fitted curve of the Gabor function.

2. Image classification algorithm

2.1. *Processing flow of classification*

Fig.1 shows the processing flow diagrams with three different preprocessing: no preprocessing (Fig. 1(a)), Gabor filtering (Fig.1 (b)), and calculation of the difference between two Gabor filtered signals in the opposite color channels (Fig.1 (c)). The input images are composed of RGB three channels.

The Gabor filters enhance edges with 0, 45, 90, and 135 degree orientations of each color channel. The component labeled "opposite color contrast" calculates the difference between the red and the green channels and the difference between the blue and the yellow (the average of red and green) channels, and this processing is applied to the Gabor filtered signals of each orientation separately. In addition, the outputs of Gabor filters and the outputs of the opposite color contrast processing are rectified as shown in Fig.1(b) and (c).

The average of outputs from 50 CNNs whose parameters were chosen randomly were used for evaluation; the parameters are the number of layers and neurons, and the size of kernels for convolution.

2.2. *Preprocessing*

2.2.1. *Gabor filtering*

A Gabor filter is a two dimensional spatial filter that enhances edges with a particular orientation, and was used to simulate the spatial characteristics of a simple cell, which is a well-studied neuron in the primary visual cortex.[4] In this study, we used Gabor filters that enhance edges with 0, 45, 90, and 135 degree orientations. The kernel of a Gabor filter that enhances edges with θ degree orientation is expressed as:

$$G(x, y) = A \cdot \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \times \cos\left(\frac{2\pi}{\lambda}(x\cos\theta + y\sin\theta) + \phi\right) \quad (1)$$

where x and y represent the coordinates in the kernel. The parameters were set as follows: $A = 0.506$, $\sigma = 2.0$, $\lambda = 4.0$, $\phi = \pi/2$. Fig. 2 shows the spatial characteristics at $y = 0$ of the filter whose $\theta = 0$.

The Gabor filtered images of RGB channels are hereinafter referred to as (G_R, G_G, G_B), respectively.

2.2.2. *Opposite color contrast*

Before calculating the difference between opposite color signals, the yellow channel is generated by

$$G_Y(x, y) = \frac{G_R(x, y) + G_G(x, y)}{2}, \quad (2)$$

where x and y represent the coordinates on the image.

The output signals of the opposite color contrast processing are given by $G_{R-G} = G_R - G_G$ and $G_{B-Y} = G_B - G_Y$. This processing is applied separately to each output image of the Gabor filter with four orientations.

2.3. *Convolutional neural network*

CNNs are neural networks (NNs) mainly composed of convolutional layers, fully connected layers, and pooling layers.

Fig.3 shows the architecture of the CNNs used in this study. The average of outputs from 50 CNNs whose parameters were chosen randomly were evaluated to remove the dependency on the topology of CNNs, and to investigate the effect of preprocessing alone. The following parameters were randomly selected: the number of layers and neurons, and the kernel size for convolution. Fig. 3(a) shows the entire network, whose components labeled "Randomly selected layers" were selected probabilistically from Fig. 3(b) and Fig. 3(c). The kernel size was randomly selected from {3, 5, 7}. The initial number of filters for convolutional layers was 32, and each time a layer was added, the number was doubled with a probability of 0.3. The number of neurons of the component labeled as "Fully connected layer" was chosen probabilistically from 2^n ($n \in \{5, 6, \ldots, 11\}$).

The rectified linear unit (ReLU) function[5] was used for the activation function of all layers except for the output layer, whose activation function was the softmax function.

The backpropagation algorithm and Adam[6] were used to train all weights. Adam is one of the most widely used optimization algorithms for NNs that changes the update amplitude of weights depending on the past updates of weights.

3. Experiments and results

3.1. *Experimental environment*

We implemented the methods described in Chapter 2 using python. The STL-10 dataset[8] was used for the evaluation. The STL-10 dataset is an image dataset consisting of 10 classes of images whose resolution is 96 × 96 pixels; each class has 500 training images and 800 test images. In order to perform the opposite color contrast processing, we removed grayscale images from the STL-10 dataset in this experiment.

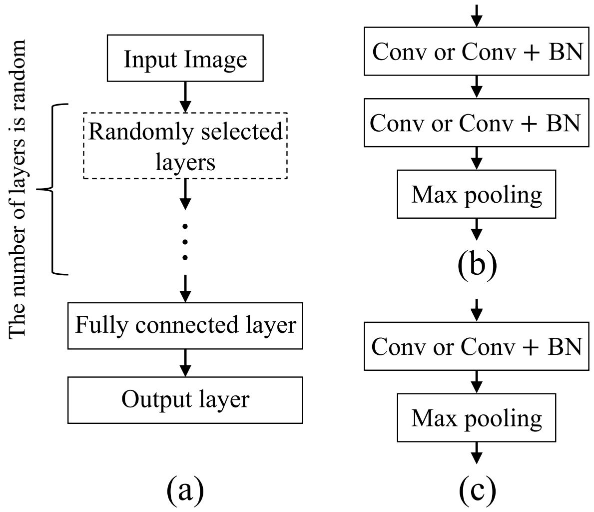

Fig. 3. Network architecture whose parameters are randomly selected. (a) Entire network architecture. Layers (b) or (c) are randomly embedded in the rectangle labeled "Randomly selected layers". The number of neurons of the component labeled "Fully connected layer" is random. (b), (c) Candidate layers to be embedded in the network. Conv represents the convolutional layer and BN represents the batch normalization layer[7]. In the component labeled "Conv or Conv + BN", only the convolutional layer or the combination of the convolutional layer and the batch normalization layer is selected probabilistically. The parameters in the layers are also random (See body text for detail).

Table. 1. Averaged accuracy (%) of the three methods for the test data.

no preprocessing	Gabor filtering	Gabor filtering and opposite color contrast
51.687	57.879	51.867

The results were evaluated in terms of averaged accuracy obtained from the 50 CNNs described in Section 2.3. The number of epochs for learning was 50, and the size of the mini-batch was 128.

3.2. *Results*

Table.1 shows the accuracy of the three methods with different preprocessing described in Section 2.1. The method using Gabor filtering obtained the highest accuracy, whereas no significant difference in accuracy was found between the method without preprocessing and that with both Gabor filtering and opposite color contrast processing.

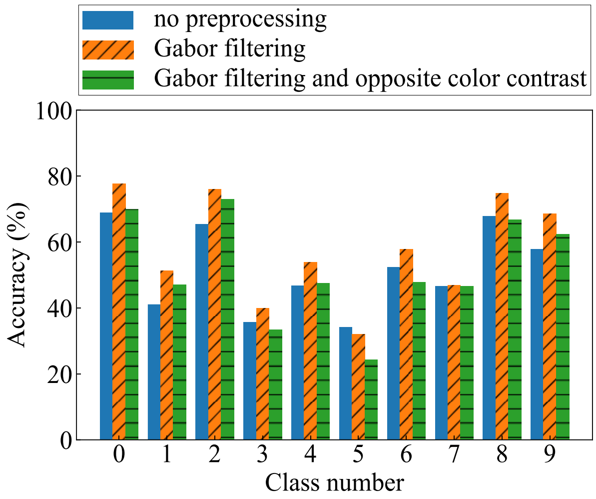

Fig. 4. Averaged accuracy of the three methods for each class.

Fig. 4 shows the accuracy for each class of the test data. As can be seen from Fig. 4, the method using Gabor filtering has the highest accuracy for all classes except for class 5.

4. Conclusion

In this study, we compared the classification accuracy of the average of multiple CNNs with the three types preprocessing to investigate the effect of neuro-inspired preprocessing. As a result, the highest accuracy was obtained by the method using Gabor filtering as a preprocessing, and no significant difference in accuracy was found between the method without preprocessing and that with both Gabor filtering and opposite color contrast processing. The results suggested that Gabor filtering used as a preprocessing for CNNs is an effective way to improve classification accuracy.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 19K12916.

References

1. A. Krizhevsky, I. Sutskever, G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Communications of the ACM, Vol.60, No.6, pp. 84-90, 2017.
2. M. M. Taghi Zadeh, M. Imani, B. Majidi, "Fast Facial emotion recognition Using Convolutional Neural Networks and Gabor Filters," 2019 5th Conference on Knowledge Based Engineering and Innovation, pp. 577-581, 2019.
3. S. Luan, C. Chen, B. Zhang, J. Han, J. Liu, "Gabor Convolutional Networks," IEEE Transactions on Image Processing, Vol. 27, No. 9, pp. 4357-4366, 2018.
4. S. Marcelja, "Mathematical description of the responses of simple cortical cells," Journal of the Optical Society of America, Vol. 70, No. 11 pp. 1297-1300, 1980.
5. V. Nair, G. E. Hinton, "Rectified Linear Units Improve Restricted Boltzmann Machines," Proceedings of the 27th International Conference on Machine Learning, pp. 807-814, 2010.
6. D. Kingma, J. Ba, "Adam: A Method for Stochastic Optimization," Proceedings of the 3rd International Conference for Learning Representations, 2015.
7. S. Ioffe, C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," Proceedings of the 32nd International Conference on Machine Learning, pp. 448-456, 2015.
8. A. Coates, H. Lee, A. Y. Ng, "An Analysis of Single-Layer Networks in Unsupervised Feature Learning," Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, pp. 215-223, 2011.

Authors Introduction

Mr. Akito Morita

He received his B.S. degree from the Department of Information Science and Technology, Osaka Institute of Technology, Japan in 2020. He is currently a Master's course student in Osaka Institute of Technology, Japan.

Dr. Hirotsugu Okuno

He received the Ph.D degree in electrical, electronic and information engineering from Osaka University, in 2008. He is currently an Associate Professor at the Faculty of Information Science and Technology, Osaka Institute of Technology. His research interests include visual information processing in the nervous system and their applications to robotics.