Artificial Vision by Deep CNN Neocognitron

Kunihiko Fukushima

Fuzzy Logic Systems Institute, Fukuoka, Japan

E-mail: fukushima@m.ieice.org

Recently, deep convolutional neural networks (deep CNN) have become very popular in the field of visual pattern recognition. The neocognitron [1], which was first proposed by Fukushima (1979), is a network classified to this category. Its architecture was suggested by neurophysiological findings on the visual systems of mammals. It is a hierarchical multi-layered network. It acquires the ability to recognize visual patterns robustly through learning. Although the neocognitron has a long history, improvements of the network are still continuing. This talk discusses the recent neocognitron, focusing on differences from the conventional deep CNN.

For training intermediate layers of the neocognitron, the learning rule called AiS (Add-if-Silent) is used [2]. Under the AiS rule, a new cell is generated and added to the network if all postsynaptic cells are silent in spite of non-silent presynaptic cells. The generated cell learns the activity of the presynaptic cells in one-shot. Once a cell is generated, its input connections do not change any more. Thus the training process is very simple and does not require time-consuming repetitive calculation.

In the deepest layer, a method called IntVec (Interpolating-Vector) is used for classifying input patterns based on the features extracted by the intermediate layers [3][4]. For the recognition by the IntVec, we search, in the multi-dimensional feature space, the nearest plane or line that is made of a trio or pair of reference vectors. Computer simulation shows that recognition error can be made much smaller by the IntVec than by the WTA (Winner-Take-All) or even by the SVM (support vector machine).

For training the deepest layer, a supervised learning rule called mWTA (margined Winner-Take-All) is used [5]. Every time when a training pattern is presented during the learning, if the result of classification by the WTA is an error, a new cell is generated in the deepest layer. Here we put a certain amount of margin to the WTA. In other words, only during the learning, a certain amount of handicap is given to cells of classes other than that of the training vector, and the winner is chosen under this handicap. By introducing the margin to the WTA, we can generate a compact set of cells, with which a high recognition rate can be obtained with a small computational cost.

Some other functions of the visual system can also be realized by networks extended from the neocognitron, for example, mechanism of selective attention, recognition and completion of partly occluded patterns, restoring occluded contours, and analysis of optic flow.

References

[1] Fukushima K.: "Artificial vision by multi-layered neural networks: neocognitron and its advances", Neural Networks, Vol. 37, (2013), pp 103-119.

[2] Fukushima K.: "Training multi-layered neural network neocognitron", Neural Networks, Vol. 40, (2013), pp 18-31.

[3] Fukushima K.: "Interpolating vectors for robust pattern recognition", Neural Networks, Vol. 40, No. 8, (2007), pp 904-916.

[4] Fukushima K. and Shouno H.: "Deep convolutional network neocognitron: improved interpolating-vector", IJCNN 2015, Killarney, Ireland, (2015), pp 1603-1610.

[5] Fukushima K.: "Margined winner-take-all: new learning rule for pattern recognition", IJCNN 2016, Vancouver, Canada, N-16157, (2016), pp 977-984.