Pyramidal Neural Networks with Variable Receptive Fields Designed by Genetic Algorithms

Alessandra M. Soares
University of Pernambuco, Brazil
ams3@ecomp.poli.br

Bruno J.T. Fernandes
University of Pernambuco, Brazil
bjtf@ecomp.poli.br

Carmelo J.A.Bastos-Filho
University of Pernambuco, Brazil
carmelofilho@ecomp.poli.br

ABSTRACT

Pyramidal Neural Networks (PNN) are computational systems inspired in the concept of receptive fields from the human visual system. In the original approach, the size of the receptive field within the same 2D layer is constant. However, their size is variable in the human visual system. This paper proposes a PNN with variable receptive fields, which might be determined by a Genetic Algorithm, called Variable Pyramidal Neural Network with Genetic Algorithms ($VPNN$-GA). We observed from preliminary experiments aiming at detecting faces in images that our approach can achieve better classification rates than the original.

CCS Concepts

•Computing methodologies → Search methodologies; Object recognition; Neural networks;

Keywords

Pyramidal Neural Network; Genetic Algorithm; Receptive Field

1. INTRODUCTION

Pyramidal Neural Network (*PyraNet*) [5] and its generalization, Lateral Inhibition Pyramidal Neural Network (*LIPNet*) [2], are Artificial Neural Networks (ANN) deployed to computer vision applications. They include implicit feature extraction avoiding the need for prior knowledge to predefine shapes or geometric relationships regarding the patterns to be recognized. This is an interesting characteristic for a pattern recognition system since the feature extraction procedure generally depends on the problem and, consequently, requires a huge effort from the specialist for each application. Both networks are based on the concept of receptive fields of the human visual system. In the human eye, photoreceptor cells are situated at the posterior area, called retina, and produce electrical impulses when stimulated by light. Each sensory neuron is responsible for capturing the impulses from a specific receptive region of the retina, so that the neuron activates when a proper stimulus is identified.

One of the limitations presented in the PNN models described in the literature regards on the use of regular spaced equal sized receptive fields. This is not the case for the human eye, where the sizes of the receptive regions can vary depending on the quantity of information to be processed [3]. Since we believe variable receptive fields can help to represent features of the patterns to be recognized and considering the biological inspiration for PNNs, this paper proposes a Variable Pyramidal Neural Network ($VPNN$), which is a PNN with variable receptive fields. In the proposal introduced in this paper, the size of each receptive field is determined by a Genetic Algorithm.

GECCO '15 July 11-15, 2015, Madrid, Spain

ACM ISBN 978-1-4503-3488-4/15/07.

DOI: http://dx.doi.org/10.1145/2739482.2764684

2. VPNN DESIGNED BY GA

$VPNN$ is a supervised multilayer neural network composed by two types of layers: two-dimensional (2D), which are responsible for feature extraction and data reduction, and are formed by a bidimensional matrix of neurons; one-dimensional (1D), which are composed by feedforward layers deployed for classification. The main goals of VPNN are to reduce the error in the classification, which is evaluated as the difference between the desired and the obtained output.

In this architecture, an image is received as an input of the first 2D layer, which is located in the base of the pyramid. The sequential layers are organized in a cascade building a pyramid-like structure: each neuron that belongs to the upper-layer is connected to a specific rectangular region of neurons in the former layer (i.e., the receptive field). The output of lower-layer neurons are used as input of upper-layer ones. Between the last 2D layer and the first 1D layer, the neurons are reorganized from a matricial arrangement to a linear one. The other 1D layers are connected exactly as a Multi-Layer Perceptron neural network. The outputs of the last 1D layer represent the outputs of the VPNN.

For 2D layers, let $R^l_{u,v}$ be the rectangular receptive field of neuron in position (u, v) from layer l. Each receptive field is defined by a center $(cx^l_{u,v}, cy^l_{u,v})$ located at layer $l-1$ and two sizes $rx^l_{u,v}$ and $ry^l_{u,v}$ (one for each coordinate). Mathematically, the receptive field is defined by the expression $R^l_{u,v} = \{(i,j) | cx^l_{u,v} - rx^l_{u,v} \leq i \leq cx^l_{u,v} + rx^l_{u,v}; cy^l_{u,v} - ry^l_{u,v} \leq j \leq cy^l_{u,v} + ry^l_{u,v}\}$.

In 2014, Soares, Fernandes and Bastos-Filho [7] proposed a procedure to optimize the $LIPNet$ parameters using Particle Swarm Optimization. For the studied case with 26 receptive fields for a regular $LIPNet$ [7], one must define the width rx and the height ry for each one of the receptive fields, resulting in 52 parameters to be determined. This optimization can be considered inappropriate to be defined empirically. Besides, the input image size applied in [7] have low resolution compared to others benchmarks and an increase would imply in even more parameters to be determined. These reasons stimulated the application of an optimization heuristic to determine the size of receptive fields.

Since the size of receptive fields are discrete variables, we decided to apply an evolutionary computation approach which presents good results for this type of variable. This class of algorithms is widely used for optimization in complex problems and the evolution through the generations happens because of GA operators: crossover, mutation and selection. Problem solutions are represented by genes.

In VPNN-GA, the individual is composed by heights and widths of rectangular receptive fields. The fitness evaluation happens in three steps: initialization, training and testing. The first step consists of randomly generating the initial weights of the network. The randomness of this process might produce an instability in the optimization that must be avoided. In order to overcome this problem, one needs to evaluate the fitness more than once. Therefore, each gene is evaluated q_{fit} times and the average Area Under the Curve (AUC)[1] is regarded as the individual's fitness. If q_{fit} is too small, the optimization process can be unstable; otherwise, experiments require high computational cost, since more initialization, training and testing procedures are needed. In the training step, the algorithm applied to update the weights is the same specified by *PyraNet* and *LIPNet* sources, which is the Resilient Propagation [6]. Finally, the last step is to test images not used in training, in order to check the generalization capability of the network. By the end of this procedure, the AUC is returned. AUC is widely used for assessing pattern recognition algorithms and it is related to the true positive and false positives rates.

3. EXPERIMENTS AND RESULTS

For the experiments, we extracted images from Center for Biological and Computational Learning benchmark [4], which is already divided in two sets: training and testing.

Regarding the GA operators, we considered the convergence as the criterion to choose the algorithms. The reason is GA itself is known for its slow convergence and, on the other hand, the fitness' computational cost is already high. Then, the intention is to balance quality of the optimization process and the computational cost.

The network output layer contains two neurons: the first output neuron gives the probability of the assessed image to be a face; and the second one, to be a non-face. The network settings defined according to [7] produces 52 genes for each individual. Due to the variability of the size of receptive fields, not all genes are effectively considered by the network since some neurons may be ignored in the process. Consequently, some image regions might not be assessed and the model tends to focus on the most important regions.

In fact, results showed receptive fields identified by the model emphasize important regions of face elements, such as the contrast between nose, cheeks and eyes. However, some receptive fields highlighted local regions, which are not important to the recognition process. This means the optimization process can be susceptible to be trapped in local minima in some cases. This can be related to the GA operators, since the intention of speeding up the optimization might cause a premature convergence of the algorithm.

We compared our approach in terms of classification rates with *PyraNet* and *LIPNet*. The settings defined for *PyraNet* and *LIPNet* are extracted from Fernandes *et al.* [2] and Soares *et al.* [7], respectively. Results showed VPNN-GA achieved a median classification rate of 89%, which is the best one (86% for *PyraNet* and 83.5% for *LIPNet*). We performed the signed-rank Wilcoxon statistical test, which confirmed our hypothesis for $p = 5\%$.

4. CONCLUSIONS

This paper proposed a VPNN with variable receptive fields. This approach is more consistent with the real human visual system than previous networks, such as *PyraNet* and *LIPNet*. In order to optimize the VPNN receptive fields, we used a GA resulting in a hybrid model, called VPNN-GA. We assessed our proposal in a face recognition problem and we reached better classification rates than the previous pyramidal networks. Moreover, regions with more information were emphasized by the receptive fields defined by VPNN-GA. For future works, other algorithms to define the receptive fields might be developed. Besides, we plan to apply our proposal to different benchmarks and problems.

5. ACKNOWLEDGMENTS

This research was partially supported by Brazilian agencies Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento Pessoal de Nível Superior (CAPES) and Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE).

6. REFERENCES

[1] A. P. Bradley. The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7):1145–1159, 1997.

[2] B. J. T. Fernandes, G. D. C. Cavalcanti, and T. Ren. Lateral inhibition pyramidal neural network for image classification. *IEEE Trans. Cybernetics*, 43:2082–2092, 2013.

[3] C. D. Gilbert and T. N. Wiesel. Receptive field dynamics in adult primary visual cortex. *Nature*, 356(6365):150–152, 1992.

[4] B. Heisele, T. Poggio, and M. Pontil. Face detection in still gray images. In *A.I. memo 1687, Center for Biological and Computational Learning*, 2000.

[5] S. L. Phung and A. Bouzerdoum. A pyramidal neural network for visual pattern recognition. *IEEE Trans. Neural Networks*, 18(2):329–343, 2007.

[6] M. Riedmiller and H. Braun. A direct adaptive method for faster backpropagation learning: The rprop algorithm. In *In Proc. IEEE Int. Conf. Neural Networks.*, pages 586–591, 1993.

[7] A. M. Soares, B. J. T. Fernandes, and C. J. A. Bastos-Filho. Lateral inhibition pyramidal neural networks designed by particle swarm optimization. In S. Wermter, C. Weber, W. Duch, T. Honkela, P. Koprinkova-Hristova, S. Magg, G. Palm, and A. Villa, editors, *Artificial Neural Networks and Machine Learning ICANN 2014*, volume 8681 of *Lecture Notes in Computer Science*, pages 667–674. Springer International Publishing, 2014.