Spiking neuron network for image segmentation

Boudjelal Meftah, M. Debakla, M. Zaagane, A. Benyettou, Olivier Lezoray

HAL Id: hal-00329522

https://hal.archives-ouvertes.fr/hal-00329522

Submitted on 11 Oct 2008

Spiking Neuron Network for Image Segmentation

B. Meftah, M. Debakla, M.Zaagane
Laboratoire de Recherche sur les Systèmes Biologiques et la Géomatique (LRSBG) équipe EDTEC, Centre Universitaire Mustapha Stambouli, Mascara, Algérie

A. Benyettou
Laboratoire SIMPA, BP 1505, Département d'Informatique, Faculté des Sciences, USTO-Mohamed Boudiaf, Oran, Algérie

O. Lezoray
Université de Caen, GREYC UMR CNRS 6072, 6 Bd. Marchal Juin, F-14050, Caen, France

Abstract

The process of segmenting images is one of the most critical ones in automatic image analysis whose goal can be regarded as to find what objects are presented in images. Artificial neural networks have been well developed. First two generations of neural networks have a lot of successful applications. Spiking Neuron Networks (SNNs) are often referred to as the 3rd generation of neural networks which have potential to solve problems related to biological stimuli. They derive their strength and interest from an accurate modeling of synaptic interactions between neurons, taking into account the time of spike emission. SNNs overcome the computational power of neural networks made of threshold or sigmoidal units. Based on dynamic event-driven processing, they open up new horizons for developing models with an exponential capacity of memorizing and a strong ability to fast adaptation. Moreover, SNNs add a new dimension, the temporal axis, to the representation capacity and the processing abilities of neural networks. In this paper, we present how SNN can be applied with efficacy in image segmentation.

Key words: Classification, Segmentation, Learning, Microscopic cellular images, Segmentation, Spiking Neuron Network.

1. Introduction

Image segmentation consists of subdividing an image into its constituent parts and extracting these parts of interest. A large number of segmentation algorithms have been developed since the middle of 1960's [1], and this number continually increases at a fast rate.

Simple and popular methods are threshold-based and process histogram characteristics of the pixel intensities of the image. Of course, thresholding has many limitations: the transition between objects and background has to be distinct and the result does not guarantee closed object contours, often requiring substantial post-processing.

Region-based methods have also been developed; they exploit similarity in intensity, gradient, or variance of neighboring pixels. Watersheds methods can be included in this category. The problem with these methods is that they do not employ any shape information of the image, which can be useful in the presence of noise.

Meanwhile, artificial neural networks are already becoming a fairly renowned technique within computer science. Since 1997, Maass [2,3] has quoted that computation and learning has to proceed quite differently in SNNs. He proposes to classify neural networks as follows:

- 1st generation: Networks based on McCulloch and Pitts' neurons as computational units, i.e. threshold gates, with only digital outputs (e.g.perceptrons, Hopfield network, Boltzmann machine, multilayer perceptrons with threshold units).

- 2nd generation: Networks based on computational units that apply an activation function with a continuous set of possible output values, such as sigmoid or polynomial or exponential functions (e.g. MLP, RBF networks). The real-valued outputs of such networks can be interpreted as firing rates of natural neurons.

- 3rd generation of neural network models: Networks which employ spiking neurons as computational units, taking into account the precise firing times of neurons for information coding.

The use of spiking neurons promises high relevance for biological systems and, furthermore, might be more flexible for computer vision applications [4].

Spiking neurons with leaky integrator synapses have been used to model image segmentation and binding by synchronization and desynchronization of neuronal group activity. The model, which is called RFSLISSOM, integrates the spiking leaky integrator model with the RF-LISSOM structure, modeling self-organization and

functional dynamics of the visual cortex at a more accurate level than earlier models [5].

Many of the existing segmentation techniques, such as supervised segmentation use a lot of parameters which are difficult to tune to obtain segmentation where the image has been partitioned into homogeneously colored regions. In this paper, a spiking neural network approach is used to segment images with unsupervised learning.

The paper is organized as follows: in the first Section, related works present in the literature of spiking neural network (SNNs). The second Section is the central part of the paper and is devoted to the description of the SNN segmentation method and its main features. The results and discussions of the experimental activity are reported in the third Section. Last Section concludes.

2. Spiking Neural Network

2.1 Biological background

Neurons are remarkable among the cells of the body in their ability to propagate signals rapidly over large distances. They do this by generating characteristic electrical pulses called action potentials, or more simply spikes that can travel down nerve fibers.

Neurons are highly specialized for generating electrical signals in response to chemical and other inputs, and transmitting them to other cells. Some important morphological specializations are the dendrites that receive inputs from other neurons and the axon that carries the neuronal output to other cells. The elaborate branching structure of the dendritic tree allows a neuron to receive inputs from many other neurons through synaptic connections [6].

The membrane potential $U_j(t)$ of a postsynaptic neuron N_j varies continuously through time (cf. Figure 1). Each action potential, or spike, emitted by a presynaptic neuron connected to generates a weighted PostSynaptic Potential (PSP) which is function of time.

If the W_{ij} synaptic weight is excitatory, the EPSP is positive: Sharp increasing of the potential $U_j(t)$ and then smoothly decreasing back to null influence. If W_{ij} is inhibitory then the IPSP is negative: Sharp decreasing $U_j(t)$ and then smoothly increasing. At each time, the value of $U_j(t)$ results from the addition of the still active PSPs variations. Whenever the potential $U_j(t)$ reaches the threshold value ϑ of N_j, the neuron fires or emits a spike, that corresponds to a sudden and very high increase of $U_j(t)$, followed by a strong depreciation and a smooth return to the resting potential U_0 [7].
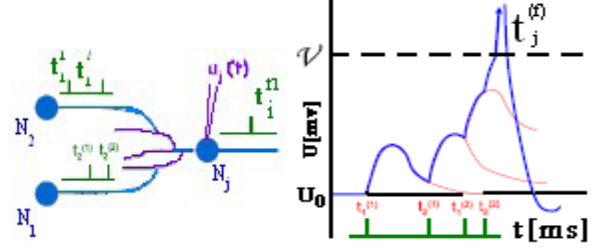
Figure 1. Emission of spike

2.2 Models of spiking neurons

Since the works of Santiago Ramon y Cajal and Camillo Golgi, a vast number of theoretical neuron models have been created, with a modern phase beginning with the work of Hodgkin and Huxley [8].

We divide the spiking neuron models into three main classes, namely threshold-fire, conductance based and compartmental models. Because of the nature of this paper we will only cover the class of threshold-fire and specially spike response model (SRM).

The SRM as defined by Gerstner [9] is simple to understand and to implement. The model expresses the membrane potential u at time t as an integral over the past, including a model of refractoriness.

Let $F_i = \left\{ t_i^f ; 1 \leq f \leq n \right\}$ denote the set of all firing times of neuron N_i and $\Gamma_j = \{i; N_i \text{ is presynaptic to } N_j\}$ denote the set of all presynaptic neuron to N_j.

The state $U_j(t)$ of neuron N_j at time t is given by:

$$u_j^{(t)} = \sum_{t_j^{(f)} \in F_j} \eta_j \left(t - t_j^{(f)} \right)$$
$$+ \sum_{i \in \Gamma_j} \sum_{t_i^{(f)} \in F_i} w_{ij} \epsilon_{ij} \left(t - t_j^{(f)} - d^k \right) \quad (1)$$

η_j models the potential reset after a spike emission, w_{ij} describes the response to presynaptic spikes. For the kernel functions, a choice of usual expressions is given by:

$$\eta_j(s) = -\vartheta e^{\left(-\frac{s}{\tau} \right)} H(s) \quad (2)$$

where H is the Heaviside function, ϑ is the threshold and τ a time constant defining the decay of the PSP. The function $\epsilon(t)$ is an α-function as:

$$\epsilon(t) = \frac{t}{\tau} e^{\left(1 - \frac{t}{\tau} \right)} \text{ for } t > 0$$
$$\text{else } \epsilon(t) = 0 \quad (3)$$

3. Segmentation Using Spiking Neural Network

However, before building a SNN, we have to explore three important issues: network architecture, information encoding and learning method. After that we will use the SNN to segment images.

3.1 Network architecture

The network architecture consists of a fully connected feedforward network of spiking neurons with connections implemented as multiple delayed synaptic terminals (cf. Figure 2).
The network consists of an input layer, a hidden layer, and an output layer. The first layer is composed of three inputs neurons (RGB values) of pixel. Each node in the hidden layer has a localized activation $\Phi^n = \Phi(\|X - C_n\|, \sigma_n)$ where $\Phi^n(.)$ is a radial basis function (RBF) localized around C_n with the degree of localization parameterized by σ_n.

Choosing $\Phi(Z, \sigma) = e^{-\frac{z^2}{2\sigma^2}}$ gives the Gaussian RBF. This layer transforms real values to temporal values. The activations of all hidden nodes are weighted and sent to the output layer. Instead of a single synapse, with its specific delay and weight, this synapse model consists of many sub-synapses, each one with its own weight and delay d^k.
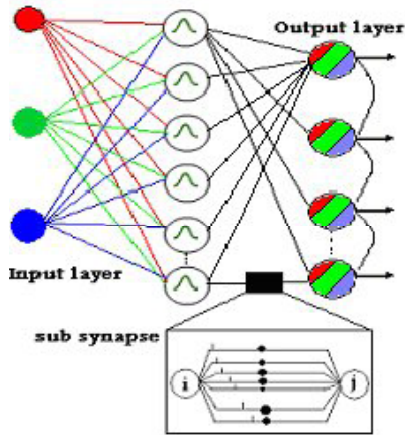
Figure 2. Network architecture

3.2 Information encoding

The first question that arises when dealing with spiking neurons is how neurons encode information in their spike trains, since we are especially interested in a method to translate an analog value into spikes. We distinguish essentially three different approaches [9] in a very rough categorization:

- Rate coding: the information is encoded in the firing rate of the neurons.

- Temporal coding: the information is encoded by the timing of the spikes.

- Population coding: information is encoded by the activity of different pools (populations) of neurons, where a neuron may participate of several pools.

We have used the temporal encoding proposed by Bohte et al. in [10]. By this method, the input variables are encoded with graded and overlapping activation functions, modeled as local receptive fields.
Each neuron of entry is modeled by a local receiving field. A receiving field is a Gaussian function. Each receiving field i have a center C_i given by the equation (4) and a width σ_i given by the equation (5) such as: m is number of receptive fields in each population and $\gamma = 1.5$.

$$c_i = \frac{i - 1.5}{m - 2} \qquad (4)$$

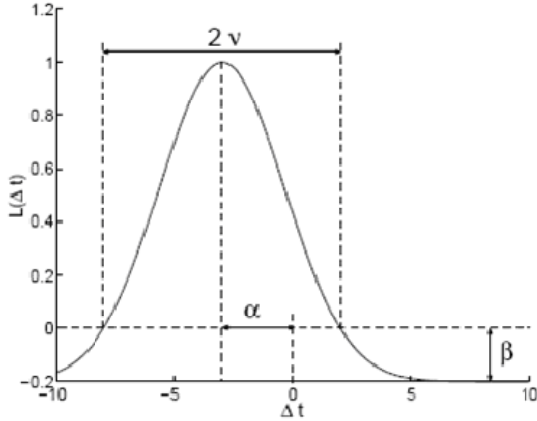$$\sigma_i = \frac{1}{\gamma(m - 2)} \qquad (5)$$

3.3 Learning method

The approach presented here implements the Hebbian reinforcement learning method through a winner-takes-all algorithm [11]. For unsupervised learning, a Winner-Take-All learning rule modifies the weights between the input neurons and the neuron first to fire in the output layer using a time-variant of Hebbian learning: if the start of a PSP at a synapse slightly precedes a spike in the output neuron, the weight of this synapse is increased, as it had significant influence on the spike-time via a relatively large contribution to the membrane potential. Earlier and later synapses are decreased in weight, reflecting their lesser impact on the output neuron's spike time. For a weight with delay d^k from neuron i to neuron j we use:

$$\Delta w_{ij} = \eta L(\Delta t_{ij}) \qquad (6)$$

And

$$L(\Delta t) = (1 + \beta)e^{\frac{(\Delta t - \alpha)^2}{2(k-1)}} - \beta \qquad (7)$$

with $\kappa = 1 - \frac{\vartheta^2}{2ln\frac{\beta}{1+\beta}}$

Figure 3. Gaussian learning function

The learning window is defined by the following parameters:

- v :this parameter, determines the width of the learning window where it crosses the zero line and affects the range of Δt_{ij}, inside which the weights are increased.

- Inside the neighborhood the weights are increased, otherwise they are decreased.

- β: this parameter determines the amount by which the weights will be reduced and corresponds to the part of the curve laying outside the neighborhood and bellow the zero line.

- α: because of the time constant τ of the EPSP, a neuron i firing exactly with j does not contribute to the firing of j, so the learning window must be shifted slightly to consider this time interval and to avoid reinforcing synapses that do not stimulate j.

4. Experimental Results and Discussion

We have chosen an image from Berkeley Segmentation Dataset and Benchmark [12] defined in pixel grid of 250x250 pixels (Fig 4).
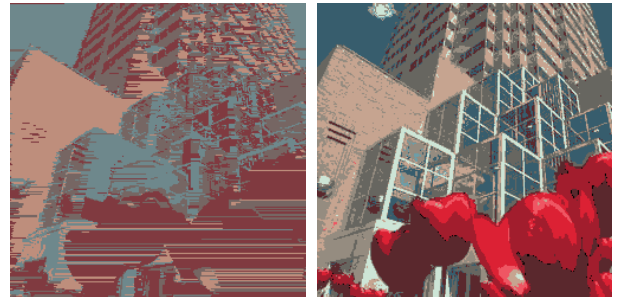

Fig. 4. Original image

To show the influence of the number of neurons at exit on the number of areas of the segmented image, we had fixed the number of sub-synapses at 14 between two neurons, the step of training to 0.35, the choice of the base of training is random starting from the image source of 5% and numbers of receiving fields with 18 (6 for each value of intensity) and we varied the number of classes at exit. The images obtained are shown in Figure 5:

Fig. 5. Segmented image with 5 and 10 classes

To show the influence of the number of sub-synapses on the number of areas of the segmented image we had fixed the number of area at exit at 10, the step of training to 0.35, the choice of the base of training is random starting from the image source of 5% and numbers of receiving fields with 18 (6 for each value of intensity) and we varied the number of subsynapses. The images obtained are shown in Figure 6:

Fig. 6. Segmented image with 4 and 14 subsynapses

To show the influence of the number of receptive fields on the number of classes of the segmented image we had fixed the number of area at exit at 10, the step of training

to 0.35, the choice of the base of training is random starting from the image source of 5%, the number of sub-synapses at 14 and we varied numbers of receiving fields. The images obtained are shown in Figure 7:

Fig. 7. Segmented image with 4 and 6 receptive fields for each value of intensity.

To show the influence of the percentage of simple training on the number of classes of the segmented image we had fixed the number of area at exit at 10, the step of training to 0.35, the number of sub-synapses at 14 and numbers of receiving fields with 18 (6 for each value of intensity) and we varied the number of percentage of simple training. The images obtained are shown in Figure 8:

Fig. 8. Segmented image with 5 and 20 % of simple training.

5. Conclusion

In this paper we applied spiking neural networks to image segmentation. At first, the network is build, a subset of the image pixel is taken to be learned by the network and finally the SNN processes the rest of the image to have as a result an important number of classes quantized the image.

References

[1] Y.J. Zhang, "An Overview of Image and Video Segmentation in the Last 40 Years", *Proceedings of the Sixth International Symposium on Signal Processing and Its Applications,* 2001, 148-151.

[2] W. Maass, "Networks of spiking neurons: The third generation of neural network models", *Neural Networks*, 1997, 10:1659-1671.

[3] W. Maass, "On the relevance of time in neural computation and learning", *Theoretical Computer Science,* 2001, 261:157-178 (extended version of ALT'97, in LNAI 1316:364-384).

[4]S. J. Thorpe, A. Delorme and R.VanRullen, "Spike-based strategies for rapid processing", *Neural Networks,* 2001, 14(6-7), 715-726.

[5] Y. Choe, R. Miikkulainen, "Contour Integration and Segmentation in A Self-organizing Map of Spiking Neurons", Biological *Cybernetics,* 2004, 75-88.

[6] P.Dayan, L.F. Abbott, "Computational and mathematical of modeling neural systems", *Theoretical neuroscience*, MIT Press, 2004.

[7] H.Paugam-Moisy, "Spiking neuron networks a survey", *IDIAP-RR,* 2006, 06-11.

[8] A.L.Hodgkin, A. F Huxley, "A quantitative description of ion currents and its applications to conduction and excitation in nerve membranes", *Journal of Physiology,* 1952, 117:500-544

[9] W.Gerstner, W.M.Kistler, "Spiking Neuron Models", *The Cambridge University Press*, Cambridge. 1st edition 2002.

[10] S.M.Bohte, H. La Poutre and J.N.Kok, "Unsupervised Segmentation with Spiking Neurons by Sparse Temporal Coding and Multi-Layer RBF Networks", *IEEE transactions on neural networks.* 2002, 13(2)

[11] M. Oster and S.C. Liu, "A winner-take-all spiking network with spiking inputs", *Proceedings of the 11th IEEE International Conference on Electronics, Circuits and Systems (ICECS 2004),* 2004, 11: 203-206

[12] D. Martin, C. Fowlkes, D. Tal and J. Malik, "A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics", *Proc. 8th Int'l Conf. Computer Vision,* 2001, 2:416-423.