A

Advanced Cellular Neural Networks Image Processing

J. Álvaro Fernández
University of Extremadura, Badajoz, Spain

INTRODUCTION

Since its introduction to the research community in 1988, the Cellular Neural Network (CNN) (Chua & Yang, 1988) paradigm has become a fruitful soil for engineers and physicists, producing over 1,000 published scientific papers and books in less than 20 years (Chua & Roska, 2002), mostly related to Digital Image Processing (DIP). This Artificial Neural Network (ANN) offers a remarkable ability of integrating complex computing processes into compact, real-time programmable analogic VLSI circuits as the ACE16k (Rodríguez *et al.*, 2004) and, more recently, into FPGA devices (Perko *et al.*, 2000).

CNN is the core of the revolutionary Analogic Cellular Computer (Roska *et al.*, 1999), a programmable system based on the so-called CNN Universal Machine (CNN-UM) (Roska & Chua, 1993). Analogic CNN computers mimic the anatomy and physiology of many sensory and processing biological organs (Chua & Roska, 2002).

This article continues the review started in this Encyclopaedia under the title *Basic Cellular Neural Network Image Processing*.

BACKGROUND

The standard CNN architecture consists of an $M \times N$ rectangular array of cells $C(i,j)$ with Cartesian coordinates (i,j), $i = 1, 2, \ldots, M$, $j = 1, 2, \ldots, N$. Each cell or neuron $C(i,j)$ is bounded to a sphere of influence $S_r(i,j)$ of positive integer radius r, defined by:

$$S_r(i,j) = \left\{ C(k,l) \middle| \max_{1 \le k \le M,\, 1 \le l \le N} \left\{ |k-i|, |l-j| \right\} \le r \right\} \tag{1}$$

This set is referred as a $(2r+1) \times (2r+1)$ neighbourhood. The parameter r controls the connectivity of a cell. When $r > N/2$ and $M = N$, a fully connected CNN is obtained, a case that corresponds to the classic Hopfield ANN model.

The state equation of any cell $C(i,j)$ in the $M \times N$ array structure of the standard CNN may be described by:

$$C\frac{dz_{ij}(t)}{dt} = -\frac{1}{R} z_{ij}(t) + \sum_{C(k,l) \in S_r(i,j)} \left[A(i,j;k,l) \cdot y_{kl}(t) + B(i,j;k,l) \cdot x_{kl} \right] + I_{ij} \tag{2}$$

where C and R are values that control the transient response of the neuron circuit (just like an RC filter), I is generally a constant value that biases the state matrix $Z = \{z_{ij}\}$, and S_r is the local neighbourhood defined in (**1**), which controls the influence of the input data $X = \{x_{ij}\}$ and the network output $Y = \{y_{ij}\}$ for time t.

This means that both input and output planes interact with the state of a cell through the definition of a set of real-valued weights, $A(i, j; k, l)$ and $B(i, j; k, l)$, whose size is determined by r. The cloning templates A and B are called the feedback and feed-forward operators, respectively.

An isotropic CNN is typically defined with constant values for r, I, A and B, implying that for an input image X, a neuron $C(i, j)$ is provided for each pixel (i, j), with constant weighted circuits defined by the feedback and feed-forward templates A and B. The neuron state value z_{ij} is adjusted with the bias parameter I, and passed as input to an output function of the form:

$$y_{ij} = \frac{1}{2}\left(\left| z_{ij}(t) + 1 \right| - \left| z_{ij}(t) - 1 \right| \right) \tag{3}$$

The vast majority of the templates defined in the CNN-UM template compendium of (Chua & Roska, 2002) are based on this isotropic scheme, using $r = 1$ and binary images in the input plane. If no feedback (i.e. $A = 0$) is used, then the CNN behaves as a convolution network, using B as a spatial filter, I as a threshold and the piecewise linear output (**3**) as a limiter. Thus, virtually any spatial filter from DIP theory can be implemented on such a feed-forward CNN, ensuring binary output stability via the definition of a central feedback absolute value greater than 1.

ADVANCED CNN IMAGE PROCESSING

In this section, a description of more complex CNN models is performed in order to provide a deeper insight into CNN design, including multi-layer structures and nonlinear templates, and also to illustrate its powerful DIP capabilities.

Nonlinear Templates

A problem often addressed in DIP edge detection is the robustness against noise (Jain, 1989). In this sense, the EDGE CNN detector for grey-scale images given by

$$A = 2,\ B_{EDGE} = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix},\ I = \text{-}0.5 \qquad (4)$$

is a typical example of a weak-against-noise filter, as a result of fixed linear feed-forward template combined with excitatory feedback. One way to provide the detector with more robustness against noise is via the definition of a nonlinear B template of the form:

$$B_{CONTOUR} = \begin{bmatrix} b & b & b \\ b & 0 & b \\ b & b & b \end{bmatrix} \text{ where } b = \begin{cases} 0.5 & \left|x_{ij} - x_{kl}\right| > th \\ -1 & \left|x_{ij} - x_{kl}\right| \le th \end{cases} \qquad (5)$$

This nonlinear template actually defines different coefficients for the surrounding pixels prior to perform the spatial filtering of the input image X. Thus, a CNN defined with nonlinear templates is generally dependent of X, and can not be treated as an isotropic model.

Just two values for the surrounding coefficients of B are allowed: one excitatory for greater than a threshold th luminance differences with the central pixel (i.e. edge pixels), and the other inhibitory, doubled in absolute value, for similar pixels, where th is usually set around 0.5. The feedback template $A = 2$ remains unchanged, but the value for the bias I must be chosen from the following analysis:

For a given state z_{ij} element, the contribution w_{ij} of the feed-forward nonlinear filter of (**5**) may be expressed as:

$$\begin{aligned} w_{ij} &= -1.0 \cdot p_s + 0.5 \cdot p_e \\ &= -\left(8 - p_e\right) + 0.5 \cdot p_e \\ &= -8 + 1.5 \cdot p_e \end{aligned} \qquad (6)$$

where p_s is the number of similar pixels in the 3 × 3 neighbourhood and p_e the rest of edge pixels. E.g. if the central pixel has 8 edge neighbours, w_{ij} = 12 – 8 = 4, whereas if all its neighbours are similar to it, then w_{ij} = –8. Thus, a pixel will be selected as edge depending on the number of its edge neighbours, providing the possibility of noise reduction. For instance, edge detection for pixels with at least 3 edge neighbours forces that $I \in (4, 5)$.

The main result is that the inclusion of nonlinearities in the definition of B coefficients and, by extension, the pixel-wise definition of the main CNN parameters gives rise to more powerful and complex DIP filters (Chua & Roska, 1993).

Morphologic Operators

Mathematical Morphology is an important contributor to the DIP field. In the classic approach, every morphologic operator is based on a series of simple concepts from Set Theory. Moreover, all of them can be divided into combinations of two basic operators: erosion and dilation (Serra, 1982). Both operators take two pieces of data as input: the binary input image and the so-called structuring element, which is usually represented by a 3×3 template.

A pixel belongs to an object if it is active (i.e. its value is 1 or black), whereas the rest of pixels are classified as background, zero-valued elements. Basic morphologic operators are defined using only object pixels, marked as 1 in the structuring element. If a pixel is not used in the match, it is left blank. Both dilation and erosion operators may be defined by the structuring elements