Basic Cellular Neural Networks Image Processing

J. Álvaro Fernández
University of Extremadura, Badajoz, Spain

INTRODUCTION

Since its seminal publication in 1988, the Cellular Neural Network (CNN) (Chua & Yang, 1988) paradigm have attracted research community's attention, mainly because of its ability for integrating complex computing processes into compact, real-time programmable analogic VLSI circuits (Rodríguez *et al.*, 2004).

Unlike cellular automata, the CNN model hosts nonlinear processors which, from analogic array inputs, in continuous time, generate analogic array outputs using a simple, repetitive scheme controlled by just a few real-valued parameters. CNN is the core of the revolutionary Analogic Cellular Computer, a programmable system whose structure is the so-called CNN Universal Machine (CNN-UM) (Roska & Chua, 1993). Analogic CNN computers mimic the anatomy and physiology of many sensory and processing organs with the additional capability of data and program storing (Chua & Roska, 2002).

This article reviews the main features of this Artificial Neural Network (ANN) model and focuses on its outstanding and more exploited engineering application: Digital Image Processing (DIP).

BACKGROUND

In the following paragraphs, a definition of the parameters and structure of the CNN is performed in order to clarify the practical usage of the model in DIP.

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 connected neighbourhood or sphere of influence $S_r(i,j)$ of positive integer radius r, which is the set of all neighbouring cells satisfying the following property:

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

This set is sometimes referred as a $(2r+1) \times (2r+1)$ neighbourhood, e.g., for a 3×3 neighbourhood, r should be 1. Thus, the parameter r controls the connectivity of a cell, i.e. the number of active synapses that connects the cell with its immediate neighbours.

When $r > N/2$ and $M = N$, a fully connected CNN is obtained, where every neuron is connected to every other cell in the network and $S_r(i,j)$ is the entire array. This extreme case corresponds to the classic Hopfield ANN model (Chua & Roska, 2002).

The state equation of any cell $C(i,j)$ in the $M \times N$ array structure of the standard CNN may be described mathematically 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, typically set to unity for the sake of simplicity), I is generally a constant value that biases or thresholds the state matrix $Z = \{z_{ij}\}$, and S_r is the local neighbourhood of cell $C(i, j)$ 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 the neighbourhood radius r. The matrices or cloning templates A and B are called the feedback and feed-forward (or control) operators, respectively.

A standard CNN is typically defined with constant values for r, I, A and B, thus implying that for a fixed input image X, a neuron $C(i, j)$ is provided for each

B

pixel (i, j), with constant weighted circuits defined by the feedback template A that connects the cell with the output plane Y, and by the control template B, which connects the neuron to the neighbouring pixels of input $x_{ij} \in X$. The value of the neuron state z_{ij} is then adjusted with the bias parameter I, and passed as input to a piecewise-linear function in order to determine the output value y_{ij}. This function may be expressed as

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

In the Image Processing context, a grey-scale image input X can be represented pixel-wise using a linear map between a pixel value (e.g. a 8-bit integer luminance matrix with 256 grey-scale levels) and the CNN input interval [–1, +1], where the lower limit is used to implement full luminance (i.e. white) and the upper for black pixels (Chua & Yang, 1988).

BASIC CNN IMAGE PROCESSING

The main application of the CNN model, due to its convolution-like scheme, has been DIP modelling and design. In the next subsections a number of basic DIP approaches are introduced, underlining the importance of the network parameters by giving illustrative examples of application. Starting from the standard model described in the previous section, the definition of the standard isotropic CNN follows. Then, an example of application in logic DIP processing is performed in order to introduce the nonlinear effects that implies the using a non-zero feedback template.

The Isotropic CNN Model

For a still image, X will be invariant with time, and for video, $X = X(t)$. In the most general case, r, A, B and I may vary with position and time, and the cloning templates are defined as nonlinear, with the possibility of integrating inhibitory signals for the state matrix and even nonlinear templates that interact with mixed input-output-state data (Chua & Roska, 2002).

These possible extensions raise the definition of a special (and simpler) class of CNN, called isotropic or space-invariant, in which r, A, B and I are fixed for the whole network and where linear synaptic operators are utilized.

In other words,

$$\sum_{C(k,l)\in S_r(i,j)} A(i,j;k,l)\cdot y_{kl} = \sum_{|k-i|\le r}\sum_{|l-j|\le r} A(i-k,j-l)\cdot y_{kl}$$

$$\sum_{C(k,l)\in S_r(i,j)} B(i,j;k,l)\cdot x_{kl} = \sum_{|k-i|\le r}\sum_{|l-j|\le r} B(i-k,j-l)\cdot x_{kl}$$

and $I_{ij} = I$. (4)

The vast majority of the templates defined in the template compendium of (Chua & Roska, 2002) for the CNN-UM 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 or saturated output filter. In this way, virtually any spatial filter from DIP theory (Jain, 1989) can be implemented on such a feed-forward driven CNN, which ensures its output stability.

For instance, the EDGE template defined by

$$A = 0,\; B_{EDGE} = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix},\; I = -1 \qquad (5)$$

is designed to work correctly for binary inputs, giving black (+1) output pixels in the input locations where a black edge pixel exists (i.e. if a black pixel has 1 white neighbour), and white (–1) pixels elsewhere.

However, when a grey-scale input image is fed to this CNN, the output may not be a binary image. To solve this potential problem, the following modification is performed over the EDGE CNN:

$$A = 2,\; B = B_{EDGE},\; I = -0.5 \qquad (6)$$

The definition of a centre feedback absolute value greater than 1 in (**6**) ensures a binary output and thus output network stability. The B template used in these CNN is of the *Laplacian* type, having the important property that all surrounding input synaptic weights are inhibitory (i.e. negative) and identical, but the centre synaptic weight is excitatory, and the average of all input synaptic weights is zero.