Binary Morphological Neural Networks

Theodore Aouad[1] and **Hugues Talbot**[1]

[1]Center for Visual Computing, CentraleSupelec, University of Paris-Saclay, Inria, France

Introduction

Conventional filters have been very successful at diverse tasks in computer vision, such as image filtering or feature extraction. Finding the right filters for a specific task is a still a challenge; convolutional neural networks (CNNs) can learn task-specific filters, given only couples of input and desired output. They have achieved outstanding results. Today, they are the go-to technology for almost any computer vision task, as long as enough data is available.

However, there still remain some tasks for which other methods are preferable. Mathematical morphology is one of them. For many applications, it is more suitable than convolution-based methods. However, finding the right sequence of operations, and the right structuring elements, can be difficult and time-consuming depending on the problem at hand. Our approach is to mimic the way CNNs are built on convolutional filters, and create a morphological network that can learn the best parameters.

Recent works have build a neural net that can learn to be either a dilation or an erosion, as well as its structuring element (SE). They either do so by using a max-plus algebra ([3]) or by replacing the non-differentiable max / min functions by smooth approximations ([2, 1, 4]). These methods deal with grey-scale inputs.

In this work, we seek to learn exact binary morphological operations with binary images as inputs. This could be very useful for shape analysis, such as shape reconstruction or classification.

Preliminary

We seek to create a single neuron that implement either an erosion or a dilation, with a learnable structuring element. We build on the optimized and distributed implementations of the CNN to create our neuron.

We rewrite the classical Minkowski addition and its dual operation with convolution. Let $S \subset \mathbb{Z}^2$ be a binary structuring element and $X \subset \mathbb{Z}^2$ be a binary image.

$$X \oplus S = \left\{ j \in \mathbb{Z}^2 \middle| (\mathbb{1}_X \circledast \mathbb{1}_S)(j) \geq 1 \right\} \quad (1)$$

$$X \ominus S = \left\{ j \in \mathbb{Z}^2 \middle| (\mathbb{1}_X \circledast \mathbb{1}_S)(j) = \text{card}(S) \right\} \quad (2)$$

We say that a function $\xi : \mathbb{R} \to]0,1[$ is a **smooth threshold function** if ξ is increasing, $\xi(0) = 0.5$, and $\lim_{x \to -\infty} \xi(x) = 0$, $\lim_{x \to +\infty} \xi(x) = 1$.

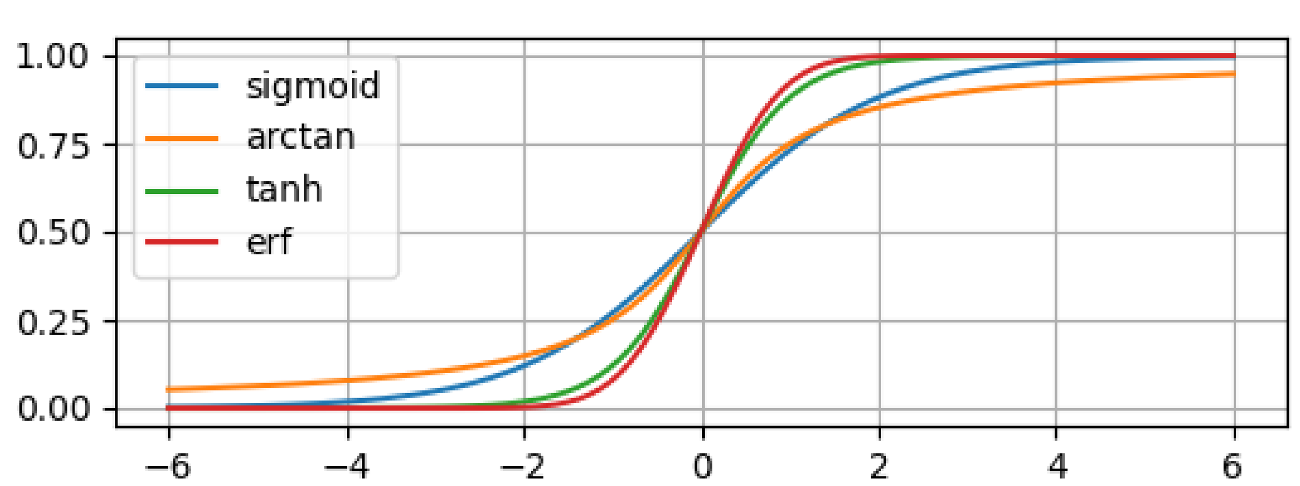


Fig. 1: Possible choices for ξ smooth-threshold function.

Method

We want a "neuron" element that can learn the Minkowski addition (dilation) and its dual (erosion). Given a smooth threshold function $\xi : \mathbb{R} \mapsto]0,1[$, some weights $W \in \mathbb{R}^(K \times K)$, a bias $b \in \mathbb{R}$ and a scaling parameter $p \in \mathbb{R} \cup \{+\infty\}$, we define the Binary Structuring Element (BiSE) neuron as follow:

BiSE

$$\epsilon_{W,b,p} : x \in [0,1]^{W \times L} \mapsto \xi(p(x \circledast \xi(W) - b)) \in [0,1]^{W \times L} \quad (3)$$

We approximate a binary SE with thresholded weights $\xi(W)$. Then, the bias b determines whether the operation is an erosion or a dilation. We can stack two BiSE neurons to learn closings and openings: $\epsilon_{W_1,b_1,p_1}(\epsilon_{W_2,b_2,p_2}(x))$. As with classical deep learning, we learn the parameters W, b, p by minimizing a loss function $\mathcal{L} : ([0,1]^{W \times L})^2 \to \mathbb{R}$ by back propagation.

Let $W, b \in [0,1]^{N \times N} \times \mathbb{R}$ be a set of thresholded weights and a bias. Let $S \subset [\![1, N]\!]^2$. We say an image $X : \mathbb{Z} \to [0,1]$ is "almost binary" if $\exists v_1 < v_2 \in [0,1]$, $X(\mathbb{Z}) \notin]v_1, v_2[$.

Proposition 1. *Given an almost binary input, $\epsilon_{W,b,+\infty}$ is a dilation by S if and only if*

$$\sum_{i \notin S} w_i + v_1 \sum_{i \in S} w_i \leq b < v_2 \min_{i \in S} w_i$$

Proposition 2. *Given an almost binary input, $\epsilon_{W,b,+\infty}$ is an erosion by S if and only if*

$$\max_{j \in S} \left(\sum_{i \neq j} (w_i) + v_1 w_j \right) \leq b < v_2 \sum_{i \in S} W_i$$

To train the BiSE, we generate binary images consisting of unions of random disks and rotated rectangles with binomial noise, and we apply complementation with probability 0.5.

Fig. 2: Input generated data

References

[1] *Going Beyond p-convolutions to Learn Grayscale Morphological Operators | SpringerLink.* 2021. URL: https://link.springer.com/chapter/10.1007/978-3-030-76657-3_34 (visited on 10/28/2021).

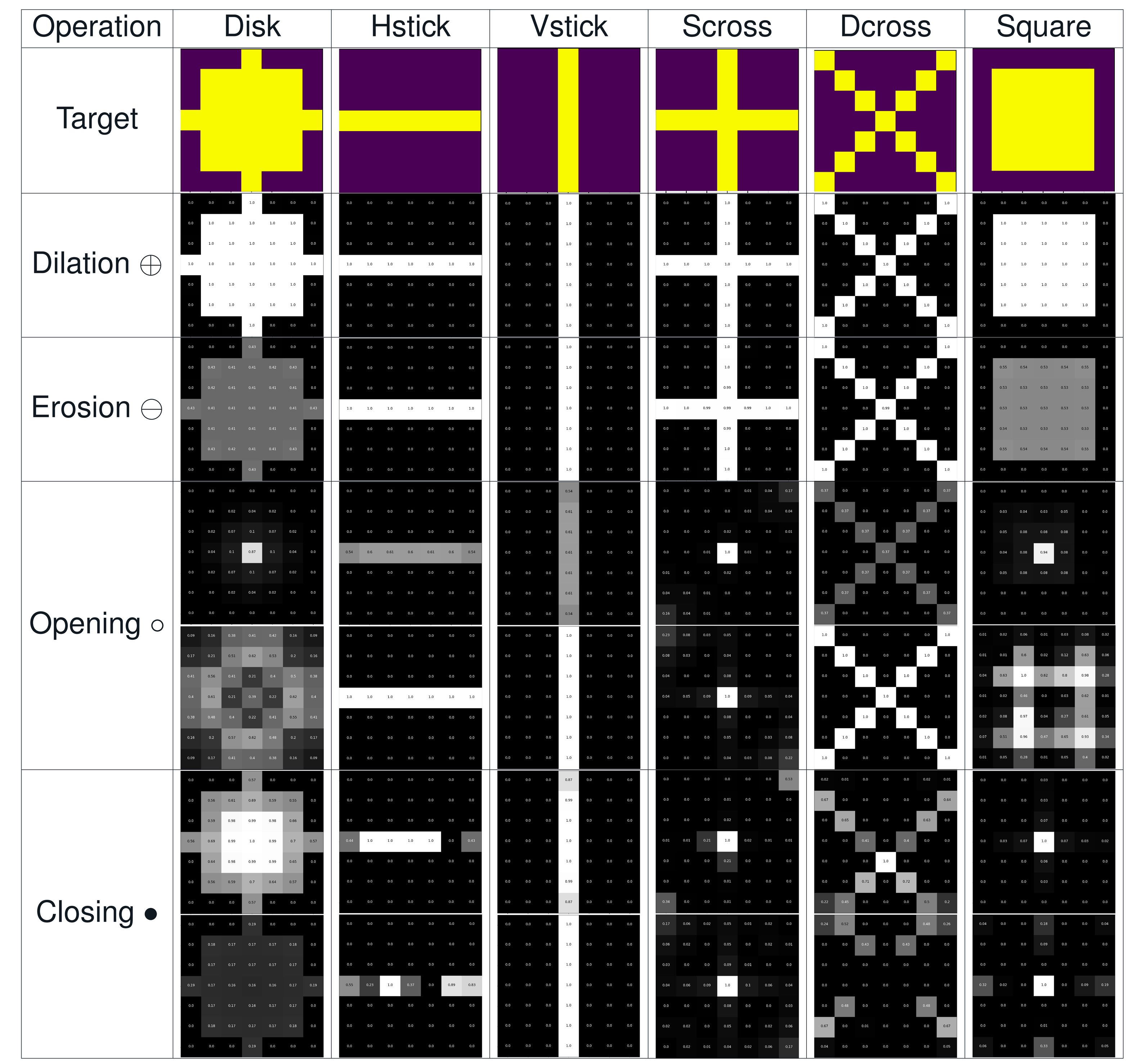
[2] Jonathan Masci, Jesus Angulo, and Jürgen Schmidhuber. "A Learning Framework for Morphological Operators Using Counter-Harmonic Mean". en. In: vol. 7883. Springer, May 2013, p. 329. DOI: 10.1007/978-3-642-38294-9_28. URL: https://hal-mines-paristech.archives-ouvertes.fr/hal-00834523 (visited on 10/28/2021).

[3] Ranjan Mondal, M. Dey, and B. Chanda. "Image Restoration by Learning Morphological Opening-Closing Network". In: *Math. Morphol. Theory Appl.* (2020). DOI: 10.1515/mathm-2020-0103.

[4] Yucong Shen, Xin Zhong, and Frank Y. Shih. "Deep Morphological Neural Networks". In: *arXiv:1909.01532 [cs, eess]* (Sept. 2019). arXiv: 1909.01532. URL: http://arxiv.org/abs/1909.01532 (visited on 10/28/2021).

Results

We report the results on erosion, dilation, opening and closing.



We succeed in learning the erosion and dilation on all presented SE perfectly.

The opening is more challenging for the disk, straight cross and square. If it fails to learn the right SE, its weights do not make sense.

The closing behaves differently. The disk does not fail anymore. It almost succeeds for the stick and the diagonal cross: the operation is the right one while the SE is not.

We believe the difference in learning performance between openings and closings is due to the training regimen, which is not currently self-dual.

Future Research

- Make the opening and closing work in practice on all SE
- Try stacking more opening / closing to filter
- Try new operations with union / intersection (top-hat, ...)
- Extend the network for shape analysis.