Binary Morphological Neural Networks

Theodore Aouad, Hugues Talbot

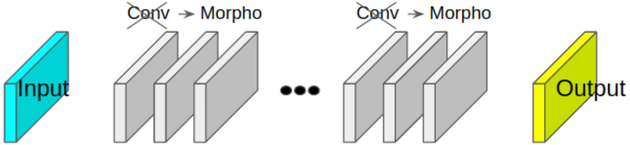
Center for Visual Computing - OPIS Inria - CentraleSupelec - University of Paris-Saclay

November 2021 - GDMM, Loria, Villers-lès-Nancy

Table of Contents

From convolution to morphological operators

- Convolution filters have been very successful in many vision problems.
- Convolutional Neural Networks (CNNs) learn the best filters for a given task.
- However, they are still cases where mathematical morphology makes more sense than convolution.
- Replacing "convolution" of CNNs by basic morphological operators (dilation and erosion) could be useful.
- How can we learn the structuring elements and the right sequence of operations for a given problem?

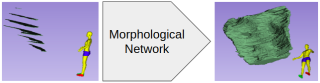


Related Work

- Learning morphological operators is not new. (Wilson [1993], Nakashizuka et al. [2010], Barrera et al. [1997])
- Recent hype on deep learning has motivated new techniques.
- Some researchers use the max-plus and min-plus definition of the dilation and erosion to perform grey-morphology on grey-scale images. (Mondal et al. [2019, 2020], Franchi et al. [2020])
- Others replace the max operation by a softmax. (Masci et al. [2013], Kirszenberg et al. [2021], Shen et al. [2019])
- Others try to learn a binary SE for grey-scale morphology (Nogueira et al. [2021])

Motivation

- Related research has primary worked on grey-scale images, with either grey-scale or binary structuring elements.
- Our aim is to perform shape analysis. For example, we want to be able to detect useful ROIs given input regions, or to infer the 3D shape using only a few slices.
- To do so, we want a system that can take work with fully binary inputs and binary structuring elements.



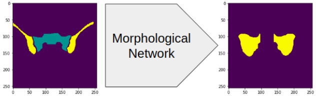


Table of Contents

Morphological operator from convolution

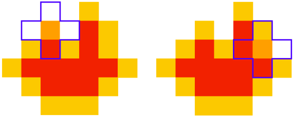
We rewrite the classical Minkowski addition and its dual operation with convolution.

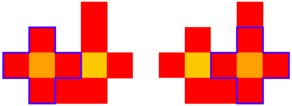
Proposition (Convolution for morphological operators)

Let $S \subset \mathbb{Z}^d$ be a binary structuring element and $X \subset \mathbb{Z}^d$ be a binary image.

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

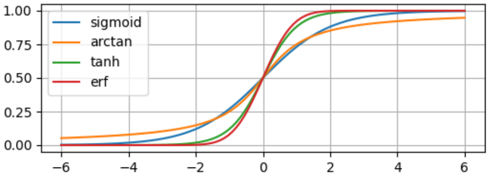
$$X \ominus S = \left(\mathbb{1}_X \circledast \mathbb{1}_S = card(S)\right) = \left\{ j \in \mathbb{Z}^d \middle| (\mathbb{1}_X \circledast \mathbb{1}_S)(j) = card(S) \right\} \tag{2}$$





Thresholding Weights

- Find $S \subset \mathbb{Z}^d$ using a weights matrix $W \in \{0,1\}^\Omega$, with Ω the support of the weights matrix (typically $\Omega = [-n, n]^d \cap \mathbb{Z}^d$, with $S \subset \Omega$).
- Relax W to be smooth and use a smooth thresholding $\xi : \mathbb{R} \mapsto]0,1[$ to ensure $\xi(W) \in [0,1]^\Omega$
  1. ξ must be increasing
  2. $\xi(0) = 0.5$
  3. $\lim(\xi(x))_{x \to -\infty} = 0$ and $\lim(\xi(x))_{x \to +\infty} = 1$



Binary Sructuring Element (BiSE) neuron

Definition (BiSE neuron)

Let $W \in \mathbb{R}^{\Omega}$ be a weight matrix, $b \in \mathbb{R}$ a bias, ξ a smooth thresholding function and $p \in \mathbb{R}_{+}^{*}$ a scaling number. We define a **BiSE neuron** as follow:

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

- The BiSE neuron is able to learn both an erosion and a dilation, as well as the associated structuring element.
- The weights W learn the structuring element.
- The bias b determines the operation, either dilation or erosion.
- The scaling number p determines how close to binary the output is.

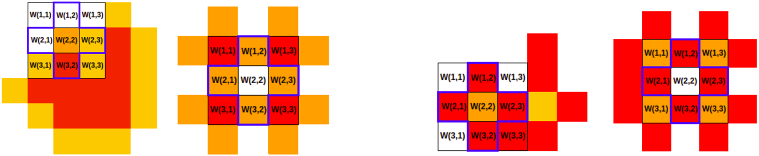
Equivalence BiSE / Dilation and Erosion - Binary input

Proposition

We assume the weights are thresholded: $W \in [0,1]^{\Omega}$. *Given a binary input,* $\epsilon_{W,b,+\infty}$ *is*

$$\text{a dilation by } S \text{ if and only if } \sum_{i \notin S} w_i \leq b < \min_{i \in S} w_i \quad (4)$$

$$\text{an erosion by } S \text{ if and only if } \max_{j \in S} \left(\sum_{i \neq j} w_i \right) \leq b < \sum_{i \in S} W_i \quad (5)$$



BiSE Output

Proposition (BiSE Output)

Let $u_1,\ u_2$ the bounds of the bias for dilation or erosion. (Ex for dilation, $u_1 := \sum_{i \notin S} w_i$, $u_2 := \min_{i \in S} w_i$). We assume that $u_1 \leq b < u_2$. Then we have:

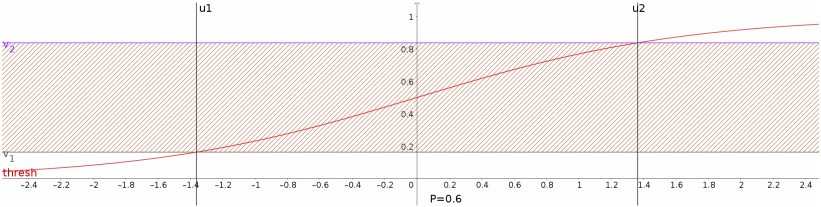
$$I \circledast W \notin \ ]u_1, u_2[\tag{6}$$

$$\epsilon_{W,b,p}(I) \notin \ ]\xi(p(u_1 - b)), \xi(p(u_2 - b))[\tag{7}$$

BiSE Output

Best case scenario, we have $b = \frac{v_1(p)+v_2(p)}{2}$. We show the possible output values depending on p.

$x \notin]u_1, u_2[\mapsto \xi(p * x) \notin]v_1(p), v_2(p)[$, with $x = I \circledast W(j) - \frac{u_1+u_2}{2}$:



Almost binary input

When a BiSE neuron is properly learned, the output will be either close to 1 or close to 0. Does it make sense to stack BiSE neurons if the outputs are not binary?

Definition (Almost Binary)

We say an image $I \in [0, 1]^{\mathbb{Z}^d}$ is **almost binary** if there exists $v_1 < v_2 \in [0, 1]$ such that $I(\mathbb{Z}^d) \notin]v_1, v_2[$.

We can extend the previously seen equivalence to almost binary inputs.

Equivalence BiSE / Dilation and Erosion - Almost Binary input

Proposition (Dilation Equivalence)

We assume the weights are thresholded: $W \in [0,1]^\Omega$. Given an almost binary input, $\epsilon_{W,b,+\infty}$ is a dilation by S if and only if

$$\sum_{i \notin S} w_i + v_1 \sum_{i \in S} w_i \leq b < v_2 \min_{i \in S} w_i \tag{8}$$

Proposition (Erosion Equivalence)

We assume the weights are thresholded: $W \in [0,1]^\Omega$. Given an almost binary input, $\epsilon_{W,b,+\infty}$ is an erosion by S if and only if

$$\max_{j \in S} \left(\sum_{i \neq j} (w_i) + v_1 w_j \right) \leq b < v_2 \sum_{i \in S} W_i \tag{9}$$

Binary Morphological Neural Network (BiMoNN)

Given an almost binary input, a properly trained BiSE neuron will always output an almost binary input. This allows the stacking of BiSE neurons sequentially.

Definition (Binary Morphological Neural Networks (BiMoNN))

We call a **Binary Morphological Neural Networks (BiMoNN)** a composition of multiple BiSE neurons. Let $K \in \mathbb{N}^*$, let $W = (W_1, ..., W_K) \in \mathbb{R}^{\Omega_1} \times ... \times \mathbb{R}^{\Omega_K}$ the set of weights for each BiSe neuron, $b = (b_1, ..., b_K) \in \mathbb{R}^K$ the set of biases and $p = (p_1, ..., p_K) \in \mathbb{R}^K$ the set of scaling numbers. We denote the BiMoNN as:

$$\phi_{W,b,p} = \epsilon_{W_K, b_K, p_K} \circ ... \circ \epsilon_{W_1, b_1, p_1} \quad (10)$$

Deep Learning Optimization

- Given a set of N couples of inputs-targets $\{(X_i, Y_i) \mid i \in \{1, ..., N\}\}$, given a loss function $\mathcal{L} : \mathbb{R}^2 \mapsto R$, we minimize:

$$\min_{W,b,p} \frac{1}{N} \sum_{i=1}^{N} \mathcal{L}(\phi_{W,b,p}(X_i), Y_i) \qquad (11)$$

- We optimize the loss using derivatives of stochastic gradient descent (batch-SGD, ADAM, ...).
- The BiMoNN is totally differentiable, and its convolutional structure make it optimized to compute the gradient for each parameter using back propagation.

Table of Contents

Training Data Generation

We train on generated binary data. The data is generated as follow:

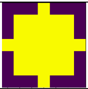
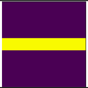
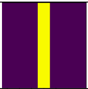
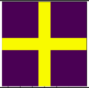
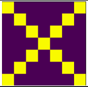
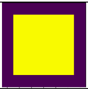
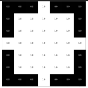
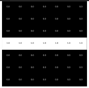
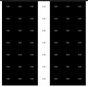
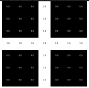
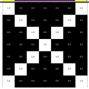
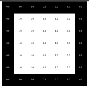
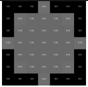
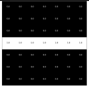
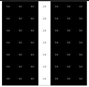
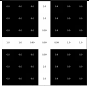
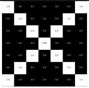
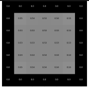
1. We generate $N = 30$ shapes
2. A shape is either a disk or a rotated rectangle box
3. We add random Bernoulli noise
4. We apply complementation with probability 0.5
5. We set the borders at 0 (depending on the kernel size Ω)

Training Regimen

- We tested on dilation, erosion, opening, closing.
- We tested 6 structuring elements of size 7x7.
- We generate images X_i and true targets Y_i with the true operation.
- For dilation and erosion, we train on $N = 200k$ images. For opening and closing, we train on $N = 1M$ images.
- The loss function $\mathcal{L}$ we use is the binary cross-entropy:

$$\mathcal{L}(\hat{y}, y^*) = y^* \log(\hat{y}) + (1 - y^*) \log(1 - \hat{y}) \tag{12}$$

Results - Final Weights on Dilation / Erosion

Operation	Disk	Hstick	Vstick	Scross	Dcross	Square
Target						
Dilation $\oplus$	 OK	 OK	 OK	 OK	 OK	 OK
Erosion $\ominus$	 OK	 OK	 OK	 OK	 OK	 OK

Results - Final Weights on Opening / Closing

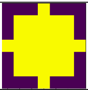
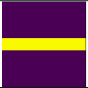
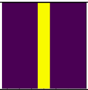
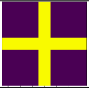
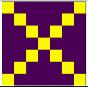
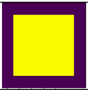
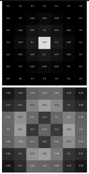
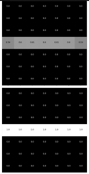
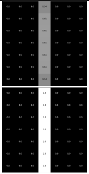
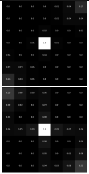
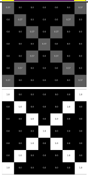
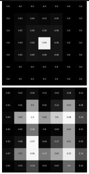
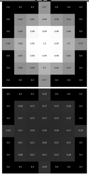
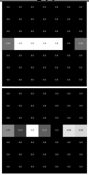
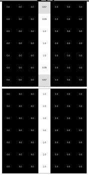
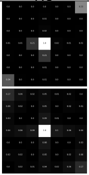
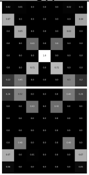
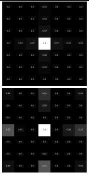
Operation	Disk	Hstick	Vstick	Scross	Dcross	Square
Target						
Opening ∘	 KO	 OK	 OK	 KO	 OK	 KO
Closing •	 OK	 KO	 OK	 KO	 KO	 KO

Table of Contents

Conclusion

Conclusion

- We introduce the BiSE neuron, that can learn the erosion, dilation and structuring element
- We introduce the BiMoNN, which in theory can learn any composition of dilations and erosions
- We managed to learn perfectly the erosion and dilation
- We managed to learn some structuring elements for opening and closing

Future Work

Future work

- Make the opening and closing work for more structuring elements
- Learn more complicated filters with multiple openings and closings
- Extend the BiMoNN to more complicated operations (for example intersection / union of dilations / erosions)
- Extend the network to classification
- Extend to shape analysis on real data
- Explore different BiSE possibilities

References I

Junior Barrera, Edward R Dougherty, and Nina Sumiko Tomita. Automatic programming of binary morphological machines by design of statistically optimal operators in the context of computational learning theory. *Journal of Electronic Imaging*, 6 (1):54–67, 1997.

Gianni Franchi, Amin Fehri, and Angela Yao. Deep morphological networks. *Pattern Recognition*, 102:107246, 2020.

Alexandre Kirszenberg, Guillaume Tochon, Élodie Puybareau, and Jesus Angulo. Going beyond p-convolutions to learn grayscale morphological operators. In *International Conference on Discrete Geometry and Mathematical Morphology*, pages 470–482. Springer, 2021.

References II

Jonathan Masci, Jesús Angulo, and Jürgen Schmidhuber. A learning framework for morphological operators using counter–harmonic mean. In *International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing*, pages 329–340. Springer, 2013.

Ranjan Mondal, Deepayan Chakraborty, and Bhabatosh Chanda. Learning 2d morphological network for old document image binarization. In *2019 International Conference on Document Analysis and Recognition (ICDAR)*, pages 65–70. IEEE, 2019.

Ranjan Mondal, Moni Shankar Dey, and Bhabatosh Chanda. Image restoration by learning morphological opening-closing network. *Mathematical Morphology-Theory and Applications*, 4 (1):87–107, 2020.

References III

Makoto Nakashizuka, Shinji Takenaka, and Youji Iiguni. Learning of structuring elements for morphological image model with a sparsity prior. In *2010 IEEE International Conference on Image Processing*, pages 85–88. IEEE, 2010.

Keiller Nogueira, Jocelyn Chanussot, Mauro Dalla Mura, and Jefersson A Dos Santos. An introduction to deep morphological networks. *IEEE Access*, 9:114308–114324, 2021.

Yucong Shen, Xin Zhong, and Frank Y Shih. Deep morphological neural networks. *arXiv preprint arXiv:1909.01532*, 2019.

Stephen S Wilson. Training structuring elements in morphological networks. In *Mathematical Morphology in Image Processing*, pages 1–41. CRC Press, 1993.

Table of Contents

Table of Contents

Complementation

If we can force the BiSE neuron to be a dilation, we can learn only the structuring element as well as a complementation.

Definition (Smooth complementation)

Let $\alpha \in [0, 1]$. We define the **smooth complementation** No_α as:

$$No_\alpha : x \in [0, 1] \mapsto \alpha \cdot x + (1 - \alpha) \cdot (1 - x) \tag{13}$$

If we pass a binary image $X \in \{0, 1\}^{\mathbb{Z}^d}$ through the complementation before giving it to a BiSE, then the input is $No_\alpha(X) \in \{\alpha, 1 - \alpha\}^{\mathbb{Z}^d}$.
We need to approximate the bias b to force a dilation.

Dilation Approximation

Proposition

Let $\alpha \in [0,1]$. *Let* $a = \min(\alpha, 1-\alpha)$ *and* $A = \max(\alpha, 1-\alpha)$. $S \subset \mathbb{Z}^d$ *be a structuring element. We consider all the possible images of values in* $\{\alpha, 1-\alpha\}$. *We denote* $\mathcal{V} = \cup_{I \in \{\alpha, 1-alpha\}^{\mathbb{Z}^d}} I \circledast \mathbb{1}_S(\mathbb{Z}^d)$. *Then* $card(\mathcal{V}) = card(S) + 1$. *Let* $v_0 < ... < v_{card(S)} \in \mathcal{V}$. *Then:*

$$\forall i \in \{0, ..., card(S)\}, v_i = a * (card(S) - i) + A * i \tag{14}$$

Definition (Bias dilation function)

The best bias for dilation is $b = \frac{v_0 + v_1}{2}$. Given weights $W \in \mathbb{R}^\Omega$ and a smooth thresholding ξ, we approximate card(S) by $\sum_{i \in \Omega} \xi(W(i))$. We define the **bias dilation function**:

$$b : \alpha \in \mathbb{R} \mapsto \min(\xi(\alpha), 1 - \xi(\alpha))\Big(\sum_{i \in \Omega} W(i) - 1\Big) + 0.5 \tag{15}$$

BiSEC

Definition (BiSEC neuron)

Let $W \in \mathbb{R}^{\Omega}$ be a weight matrix, ξ a smooth thresholding function and $p \in \mathbb{R}_{+}^{*}$ a scaling number. Let $\alpha \in \mathbb{R}$. Let $b(\alpha)$ be the bias dilation function. We define a **BiSEC neuron** as follow, with No_{α} a smooth complementation function:

$$\hat{\epsilon}_{W,\alpha,p} : x \in [0,1]^{\mathbb{Z}^d} \mapsto No_{\xi(\alpha * \infty)} \circ \epsilon_{W,b(\alpha),p} \circ No_{\xi(\alpha)} \tag{16}$$

Table of Contents

Tropical BiSE

We use the fact that the dilation can be written with max instead, inspired from grey-scale morphology.

Proposition (Tropical Dilation)

Let $\Omega \subset \mathbb{Z}^d$ a support kernel and $S \in \Omega$ a structuring element. If $W = -\infty \cdot \mathbb{1}_{\Omega \setminus S}$, then

$$\forall X \subset \mathbb{Z}^d \,,\, \forall j \in \mathbb{Z}^d \,,\, \mathbb{1}_{X \oplus S}(j) = \max_{i \in \Omega}(\mathbb{1}_X(j-i)+W(i)) =: \delta_W(X)_j \quad (17)$$

Definition (Tropical BiSE)

Let $\Omega \subset \mathbb{Z}^d$ be a support kernel and $W \in \mathbb{R}^\Omega$ a set of weights. Let $\alpha \in \mathbb{R}$ and No_α a smooth complementation function. We define the **tropical BiSE** as:

$$\bar{\epsilon}_{W,\alpha} = No_{\xi(\alpha * \infty)} \circ \delta_W \circ No_{\xi(\alpha)} \quad (18)$$

Table of Contents

BiSE Convolution Definition

It is the classic convolution. The difference with BiSE is that we remove the smooth thresholding of weights.

Definition (BiSE Convolution)

Let $\Omega \subset \mathbb{Z}^d$ and $W \in \mathbb{R}^{\Omega}$. Let $b \in \mathbb{R}$ and $p \in \mathbb{R}_+^*$. Let ξ be a smooth thresholding function. We define the **BiSE Convolution** as:

$$CONV_{W,b,p} : I \in [0,1]^{Z^d} = \xi(p(I \circledast W - b)) \tag{19}$$

BiSE Convolution Properties

Proposition (Dilation Equivalence)

We take the same notation as the previous definition. Let $S \subset \Omega$. Then $CONV_{W,b,+\infty}$ is a dilation by S if and only if

$$\max_{K \in \mathcal{P}(\bar{S})} \left(\sum_{i \in K} W_i \right) \leq b < \min_{K \in \mathcal{P}(S), K \neq \emptyset} \left(\sum_{i \in K} W_i \right) \tag{20}$$

Proposition (Erosion Equivalence)

We take the same notation as the previous definition. Let $S \subset \Omega$. Then $CONV_{W,b,+\infty}$ is an erosion by S if and only if

$$\max_{K \in \mathcal{P}(\bar{S})} \left(\sum_{i \in K} W_i \right) + \max_{j \in S} \left(\sum_{i \in S, i \neq j} W_i \right) \leq b < \sum_{i \in S} W_i + \min_{K \in \mathcal{P}(\bar{S})} \left(\sum_{i \in K} W_i \right) \tag{21}$$