Going beyond p-convolutions to learn grayscale morphological operators

Alexandre Kirszenberg[1], Guillaume Tochon[1], Élodie Puybareau[1], and Jesus Angulo[2]

[1] EPITA Research and Development Laboratory (LRDE), Le Kremlin-Bicêtre, France
`surname.name@lrde.epita.fr`
[2] Centre for Mathematical Morphology, Mines ParisTech, PSL Research University, France
`jesus.angulo@mines-paristech.fr`

Abstract. Integrating mathematical morphology operations within deep neural networks has been subject to increasing attention lately. However, replacing standard convolution layers with erosions or dilations is particularly challenging because the min and max operations are not differentiable. Relying on the asymptotic behavior of the counter-harmonic mean, p-convolutional layers were proposed as a possible workaround to this issue since they can perform pseudo-dilation or pseudo-erosion operations (depending on the value of their inner parameter p), and very promising results were reported. In this work, we present two new morphological layers based on the same principle as the p-convolutional layer while circumventing its principal drawbacks, and demonstrate their potential interest in further implementations within deep convolutional neural network architectures.

Keywords: morphological layer, p-convolution, counter-harmonic mean, grayscale mathematical morphology.

1 Introduction

Mathematical morphology deals with the non-linear filtering of images [15]. The elementary operations of mathematical morphology amount to computing the minimum (for the erosion) or maximum (for the dilation) of all pixel values within a neighborhood of some given shape and size (the structuring element) of the pixel under study. Combining those elementary operations, one can define more advanced (but still non-linear) filters, such as openings and closings, which have many times proven to be successful at various image processing tasks such as filtering, segmentation or edge detection [18]. However, deriving the optimal combination of operations and the design (shape and size) of their respective structuring element is generally done in a tedious and time-consuming trial-and-error fashion. Thus, delegating the automatic identification of the right sequence of operations to use and their structuring element to some machine learning technique is an appealing strategy.

On the other hand, artificial neural networks are composed of units (or neurons) connected to each other and organized in layers. The output of each neuron is expressed as the linear combination of its inputs weighed by trainable weights, potentially mapped by a non-linear activation function [6]. Convolutional neural networks (CNNs) work in a similar fashion, replacing neurons with convolutional filters [8].

Because of the similarity between their respective operations, there has been an increasing interest in past years to integrate morphological operations within the framework of neural networks, and two major lines of research have emerged. The first one, tracing back to the end of the 80s, replaces the multiplication and addition of linear perceptron units with addition and maximum [4, 14, 20], resulting in so-called non-linear morphological perceptrons [19] (see [3, 21] for recent works in this domain). The second line, mainly motivated by the rise of deep CNNs, explores the integration of elementary morphological operations in such networks to automatically learn their optimal shape and weights, the major issue being that the min and max operations are not differentiable. A first workaround is to replace them by smooth differentiable approximations, making them suited to the conventional gradient descent learning approach via back-propagation [8]. In their seminal work, Masci *et al.* [10] relied on some properties of the counter-harmonic mean [2] (CHM) to provide p-convolutional ($PConv$) layers, the value of the trainable parameter p dictating which of the elementary morphological operation the layer ultimately mimics. The CHM was also used as an alternative to the standard max-pooling layer in classical CNN architectures [11]. LogSumExp functions (also known as multivariate softplus) were proposed as replacements of min and max operations to learn binary [17] and grayscale [16] structuring elements. An alternative approach was followed in [5, 12]: the non-linear morphological operations remained unchanged, and the backpropagation step was instead adapted to handle them in the same way the classical max-pooling layer is handled in standard CNNs. Finally, morphological operations were recreated and optimized as combinations of depthwise and pointwise convolution with depthwise pooling [13].

Looking at all recently proposed approaches (apart from [10], all other afore-mentionned works date back to no later than 2017) and the diversity of their evaluation (image classification on MNIST database [9], image denoising and restoration, edge detection and so on), it seems to us that the magical formula for integrating morphological operations within CNNs has yet to be derived. For this reason, we would like to draw the attention in this work back to the $PConv$ layer proposed in [10]. As a matter of fact, very promising results were reported but never investigated further. Relying on the CHM framework, we propose two possible extensions to the $PConv$ layer, for which we demonstrate potential interest in further implementations within deep neural network architectures.

In section 2, we review the work on p-convolutions [10], presenting their main advantages, properties and limitations. In section 3, we propose two new morphological layers, namely the $\mathcal{L}$Morph layer (also based on the CHM) and the $\mathcal{S}$Morph layer (based on the regularized softmax). Both proposed layers are

compatible with grayscale mathematical morphology and nonflat structuring elements. In Section 4, we showcase some results from our implementations, and proceed to compare these results to those of the p-convolution layer. Finally, Section 5 provides a conclusion and some perspectives from our contributions.

2   P-convolutions: definitions, properties and pitfalls

In this section, we detail the notion of p-convolution as presented in [10].

2.1   Grayscale mathematical morphology

In mathematical morphology, an image is classically represented as a 2D function $f : E \to \mathbb{R}$ with $x \in E$ being the pixel coordinates in the 2D grid $E \subseteq \mathbb{Z}^2$ and $f(x) \in \mathbb{R}$ being the pixel value. In grayscale mathematical morphology, *i.e.* when both the image f and the structuring element b are real-valued (and not binary), erosion $f \ominus b$ and dilation $f \oplus b$ operations can be written as:

$$(f \ominus b)(x) = \inf_{y \in E} \{f(y) - b(x - y)\} \tag{1}$$

$$(f \oplus b)(x) = \sup_{y \in E} \{f(y) + b(x - y)\} \tag{2}$$

This formalism also encompasses the use of flat (binary) structuring elements, which are then written as

$$b(x) = \begin{cases} 0 & \text{if } x \in B \\ -\infty & \text{otherwise} \end{cases}, \tag{3}$$

where $B \subseteq E$ is the support of the structuring function b.

2.2   The counter-harmonic mean and the p-convolution

Let $p \in \mathbb{R}$. The counter-harmonic mean (CHM) of order p of a given non negative vector $\mathbf{x} = (x_1, \ldots, x_n) \in (\mathbb{R}^+)^n$ with non negative weights $\mathbf{w} = (w_1, \ldots, w_n) \in (\mathbb{R}^+)^n$ is defined as

$$CHM(\mathbf{x}, \mathbf{w}, p) = \frac{\sum_{i=1}^{n} w_i x_i^p}{\sum_{i=1}^{n} w_i x_i^{p-1}} . \tag{4}$$

The CHM is also known as the Lehmer mean [2]. Asymptotically, one has $\lim_{p \to +\infty} CHM(\mathbf{x}, \mathbf{w}, p) = \sup_i x_i$ and $\lim_{p \to -\infty} CHM(\mathbf{x}, \mathbf{w}, p) = \inf_i x_i$.
The p-convolution of an image f at pixel x for a given (positive) convolution kernel $w : W \subseteq E \to \mathbb{R}^+$ is defined as:

$$PConv(f, w, p)(x) = (f *_p w)(x) = \frac{(f^{p+1} * w)(x)}{(f^p * w)(x)} = \frac{\sum_{y \in W(x)} f^{p+1}(y) w(x - y)}{\sum_{y \in W(x)} f^p(y) w(x - y)} \tag{5}$$

where $f^p(x)$ denotes the pixel value $f(x)$ raised at the power of p, $W(x)$ is the spatial support of kernel w centered at x, and the scalar p controls the type of

operation to perform.

Based on the asymptotic properties of the CHM, the morphological behavior of the $PConv$ operation with respect to p has notably been studied in [1]. More precisely, when $p > 0$ (resp. $p < 0$), the operation is a pseudo-dilation (resp. pseudo-erosion), and when $p \to \infty$ (resp. $-\infty$), the largest (resp. smallest) pixel value in the local neighborhood $W(x)$ of pixel x dominates the weighted sum (5) and the $PConv(f, w, p)(x)$ acts as a non-flat grayscale dilation (resp. a non-flat grayscale erosion) with the structuring function $b(x) = \frac{1}{p}log(w(x))$:

$$\lim_{p \to +\infty} (f *_p w)(x) = \sup_{y \in W(x)} \left\{ f(y) + \frac{1}{p} \log\left(w(x-y)\right) \right\} \qquad (6)$$

$$\lim_{p \to -\infty} (f *_p w)(x) = \inf_{y \in W(x)} \left\{ f(y) - \frac{1}{p} \log\left(w(x-y)\right) \right\} \qquad (7)$$

In practice, equations (6) and (7) hold true for $|p| > 10$. The flat structuring function (3) can be recovered by using constant weight kernels, *i.e.*, $w(x) = 1$ if $x \in W$ and $w(x) = 0$ if $x \notin W$ and $|p| \gg 0$. As stated in [10], the $PConv$ operation is differentiable, thus compatible with gradient descent learning approaches via back-propagation.

2.3   Limits of the p-convolution layer

In order for the $PConv$ layer to be defined on all its possible input parameters, w and f must be strictly positive. Otherwise, the following issues can occur:

 – If $f(x)$ contains null values and p is negative, $f^p(x)$ is not defined;
 – If $f(x)$ contains negative values and p is a non-null, non-integer real number, $f^p(x)$ can contain complex numbers;
 – If $w(x)$ or $f^p(x)$ contain null values, $\frac{1}{(f^p * w)(x)}$ is not defined.

As such, before feeding an image to the p-convolution operation, it first must be rescaled between $[1, 2]$:

$$f_r(x) = 1.0 + \frac{f(x) - \min_{x \in E} f(x)}{\max_{x \in E} f(x) - \min_{x \in E} f(x)} \qquad (8)$$

Moreover, if several $PConv$ layers are concatenated one behind the other (to achieve (pseudo-) opening and closing operations for instance), a rescaling must be performed before each layer. Particular care must also be taken with the output of the last $PConv$ layer, since it must also be rescaled to ensure that the range of the output matches that of the target. This is done by adding a trainable scale/bias 1x1x1 convolution layer at the end of the network.

Last but not least, a notable drawback of the $PConv$ layer when it comes to learning a specific (binary or non-flat) structuring element is that it tends to be hollow and flattened out in the center (see further presented results in Section 4).

3   Proposed $\mathcal{L}$Morph and $\mathcal{S}$Morph layers

As exposed in the previous section 2.3, the $PConv$ layer has a few edge cases and drawbacks. For this reason, we now propose two new morphological layers, based upon the same fundamental principle as the $PConv$ operation, with the intent of making them compatible with general grayscale mathematical morphology.

3.1   Introducing the $\mathcal{L}$Morph operation

Our main objective is still to circumvent the non-differentiability of min and max functions by replacing them with smooth and differentiable approximations. In the previous section 2, we presented the CHM, whose asymptotic behavior is exploited by the $PConv$ layer [10]. Relying on this behavior once more, we now propose to define the following $\mathcal{L}$Morph (for $\mathcal{L}$ehmer-mean based Morphological) operation:

$$\mathcal{L}\text{Morph}(f,w,p)(x) = \frac{\sum_{y \in W(x)} (f(y) + w(x-y))^{p+1}}{\sum_{y \in W(x)} (f(y) + w(x-y))^p} \tag{9}$$

where $w : W \to \mathbb{R}^+$ is the structuring function and $p \in \mathbb{R}$. Defined as such, we can identify $\mathcal{L}\text{Morph}(f,w,p)$ with the CHM defined by the equation (4): all weights w_i (resp. entries x_i) of equation (4) correspond to 1 (resp. $f(y) + w(x-y)$) in the equation (9), from which we can deduce the following asymptotic behavior:

$$\lim_{p \to +\infty} \mathcal{L}\text{Morph}(f,w,p)(x) = \sup_{y \in W(x)} \{f(y) + w(x-y)\} = (f \oplus w)(x) \tag{10}$$

$$\lim_{p \to -\infty} \mathcal{L}\text{Morph}(f,w,p)(x) = \inf_{y \in W(x)} \{f(y) + w(x-y)\} = (f \ominus -w)(x) \tag{11}$$

By changing the sign of p, one can achieve either pseudo-dilation (if $p > 0$) or pseudo-erosion (if $p < 0$). Figure 1 displays examples of applying the $\mathcal{L}$Morph function with a given non-flat structuring element for different values of p. In practice $|p| > 20$ is sufficient to reproduce non-flat grayscale dilation or non-flat grayscale erosion. Note however that the applied structuring function is $-w$ in the case of an erosion.
Relying on the CHM like the $PConv$ layer brings over some shared limitations: the input image f must be positive and rescaled following equation (8), and the structuring function w must be positive or null.

3.2   Introducing the $\mathcal{S}$Morph operation

Deriving a morphological layer based on the asymptotic behavior of the CHM has a major drawback in that the input must be rescaled within the range $[1, 2]$. In order to circumvent this issue, we now propose to leverage the α-softmax function [7], which is defined as:

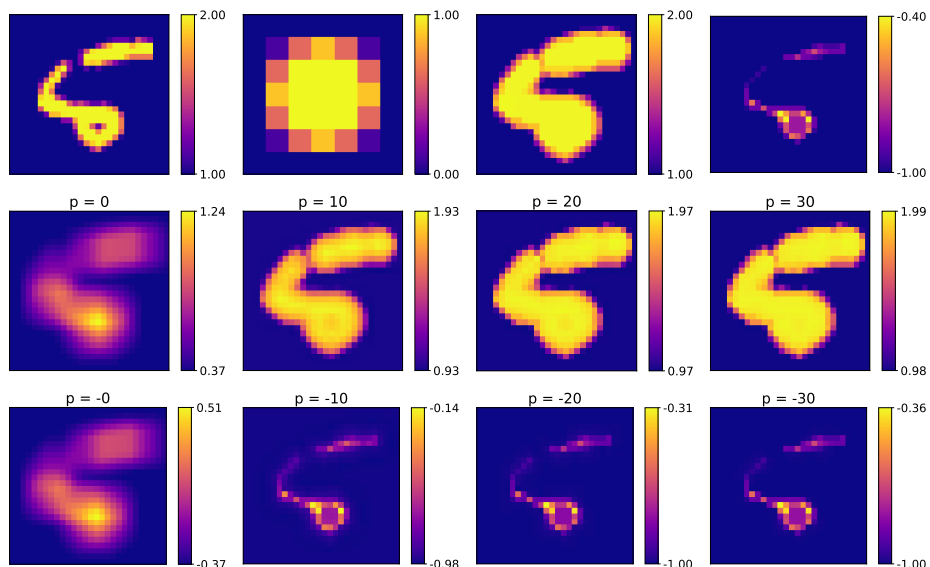$$\mathcal{S}_\alpha(\mathbf{x}) = \frac{\sum_{i=1}^n x_i e^{\alpha x_i}}{\sum_{i=1}^n e^{\alpha x_i}} \tag{12}$$

Fig. 1: Top row: input image, non-flat structuring element, target dilation, target erosion. Middle row: $\mathcal{L}$Morph pseudo-dilation for increasing value of p. Bottom row: $\mathcal{L}$Morph pseudo-erosion for increasing value of $|p|$.

for some $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$. In fact, $\mathcal{S}_\alpha$ has the desired properties that $\lim_{\alpha \to +\infty} \mathcal{S}_\alpha(\mathbf{x}) = \max_i x_i$ and $\lim_{\alpha \to -\infty} \mathcal{S}_\alpha(\mathbf{x}) = \min_i x_i$. This function is less restrictive than the CHM since it does not require the elements of $\mathbf{x}$ to be strictly positive. A major benefit is that it is no longer necessary to rescale its input.

Exploiting this property, we define in the following the $\mathcal{S}$Morph (standing for Smooth Morphological) operation:

$$\mathcal{S}\text{Morph}(f, w, \alpha)(x) = \frac{\sum_{y \in W(x)} (f(y) + w(x-y)) e^{\alpha(f(y)+w(x-y))}}{\sum_{y \in W(x)} e^{\alpha(f(y)+w(x-y))}}, \qquad (13)$$

where $w : W \to \mathbb{R}$ plays the role of the structuring function. We can see from the properties of $\mathcal{S}_\alpha$ that the following holds true:

$$\lim_{\alpha \to +\infty} \mathcal{S}\text{Morph}(f, w, \alpha)(x) = (f \oplus w)(x) \qquad (14)$$

$$\lim_{\alpha \to -\infty} \mathcal{S}\text{Morph}(f, w, \alpha)(x) = (f \ominus -w)(x) \qquad (15)$$

As such, just like the $PConv$ and $\mathcal{L}$Morph layers, the proposed $\mathcal{S}$Morph operation can alternate between a pseudo-dilation ($\alpha > 0$) and a pseudo-erosion ($\alpha < 0$). Furthermore, when $\alpha \gg 0$ (resp. $\alpha \ll 0$), this function approximates the grayscale dilation (resp. the grayscale erosion).

Figure 2 showcases examples of applying the $\mathcal{S}$Morph function with a given non-flat structuring element for different values of α. We can see that, as $|\alpha|$ increases, the operation better and better approximates the target operation.
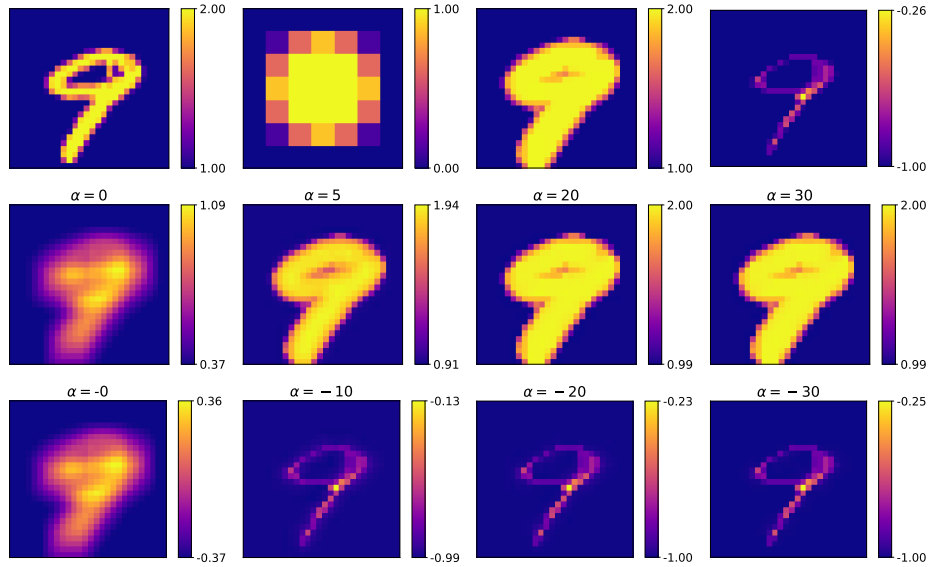
Fig. 2: Top row: input image, non-flat structuring element, target dilation, target erosion. Middle row: $\mathcal{S}$Morph pseudo-dilation for increasing value of α. Bottom row: $\mathcal{S}$Morph pseudo-erosion for increasing value of $|\alpha|$.
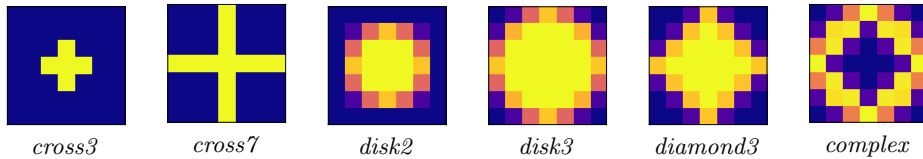

Fig. 3: 7×7 target grayscale structuring elements.

4   Conducted experiments

4.1   Experimental protocol

In the following, we evaluate the ability of the proposed $\mathcal{L}$Morph and $\mathcal{S}$Morph layers to properly learn a target structuring element and compare with results obtained by the $PConv$ layer. To do so, we apply in turn dilation $\oplus$, erosion $\ominus$, closing $\bullet$ and opening $\circ$ to all 60000 digits of the MNIST dataset [9], with the target structuring elements displayed by Figure 3. For dilation and erosion (resp. closing and opening), each network is composed of a single (resp. two) morphological layer(s) followed by a scale/bias $Conv$ 1x1x1 to rescale the output into the range of the target images. Note that for both $PConv$ and $\mathcal{L}$Morph networks, the image also has to be rescaled in the range $[1, 2]$ before passing through the morphological layer. We train all networks with a batch size of 32, optimizing for the mean squared error (MSE) loss with the Adam optimizer (with starting learning rate $\eta = 0.01$). The learning rate of the optimizer is scheduled
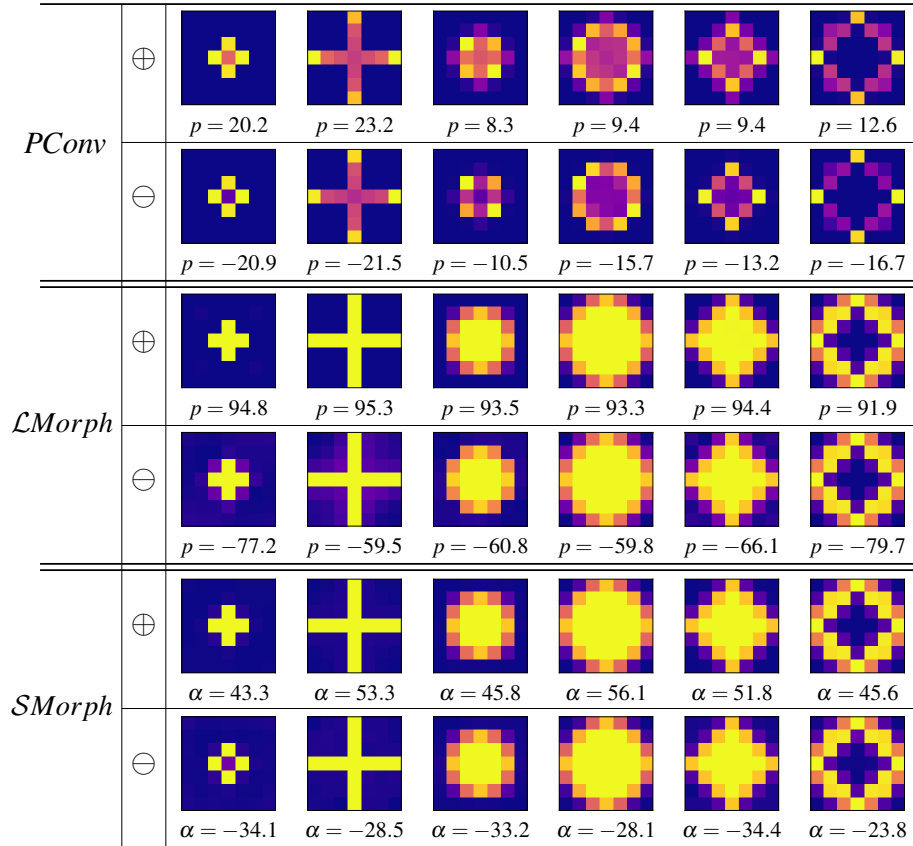
Fig. 4: Learned structuring element (with corresponding p or α at convergence) for $PConv$, $\mathcal{L}$Morph and $\mathcal{S}$Morph layers on dilation $\oplus$ and erosion $\ominus$ tasks.

to decrease by a factor of 10 when the loss plateaus for 5 consecutive epochs. Convergence is reached when the loss plateaus for 10 consecutive epochs. For the $PConv$ layer, the filter is initialized with 1s and $p = 0$. For $\mathcal{L}$Morph, the filter is initialized with a folded normal distribution with standard deviation $\sigma = 0.01$, and $p = 0$. For the $\mathcal{S}$Morph layer, the filter is initialized with a centered normal distribution with standard deviation $\sigma = 0.01$ and $\alpha = 0$. In all instances, the training is done simultaneously on the weights and the parameter p or α.

In order to assess the performance of the morphological networks for all scenarios (one scenario being one morphological operation $\oplus$, $\ominus$, $\bullet$ and $\circ$ and one target structuring element among those presented by Figure 3), we computed the root mean square error (RMSE) between the filter learned at convergence and the target filter. The loss at convergence as well as the value of the parameter p or α also serve as quantitative criteria.

Table 1: MSE loss at convergence and RMSE between the learned structuring element displayed by Figure 4 and the target for *PConv*, $\mathcal{L}$Morph and $\mathcal{S}$Morph layers on dilation $\oplus$ and erosion $\ominus$ tasks. Best (lowest) results are in bold.

			cross3	cross7	disk2	disk3	diamond3	complex	
PConv	$\oplus$	LOSS	0.5	0.8	5.1	6.4	6.8	3.3	$\times 10^{-4}$
		RMSE	0.41	1.42	2.22	3.09	2.80	2.38	
	$\ominus$	LOSS	2.4	0.62	13	2.6	5.2	1.2	$\times 10^{-5}$
		RMSE	0.82	1.55	2.82	3.77	3.63	2.76	
$\mathcal{L}$Morph	$\oplus$	LOSS	0.84	1.2	0.78	1.2	1.2	2.1	$\times 10^{-5}$
		RMSE	**0.02**	**0.00**	**0.02**	**0.05**	**0.01**	**0.48**	
	$\ominus$	LOSS	1.1	**0.37**	1.3	**0.37**	0.58	**1.1**	$\times 10^{-6}$
		RMSE	**0.44**	0.63	0.37	0.29	0.38	**0.48**	
$\mathcal{S}$Morph	$\oplus$	LOSS	**1.8**	**4.0**	**2.5**	**4.9**	**3.9**	**1.7**	$\mathbf{\times 10^{-7}}$
		RMSE	0.10	0.13	0.08	0.06	0.09	0.51	
	$\ominus$	LOSS	210	4.1	**0.8**	4.1	**4.3**	9.5	$\mathbf{\times 10^{-7}}$
		RMSE	0.78	**0.15**	**0.09**	**0.05**	**0.09**	0.51	

4.2   Obtained results

Figure 4 gathers the structuring elements learned by the *PConv*, $\mathcal{L}$Morph and $\mathcal{S}$Morph layers for dilation and erosion, and the value of their respective parameter. Looking at the sign of the parameter, all three networks succeed at finding the correct morphological operation. The magnitude of the parameter at convergence also confirms that the operation applied by all networks can be considered as dilation or erosion (and not simply pseudo-dilation or pseudo-erosion). However, looking at the shape of the learned structuring element, it is clear that the *PConv* layer suffers from the hollow effect mentionned in section 2.3, while both $\mathcal{L}$Morph and $\mathcal{S}$Morph layers accurately retrieve the target structuring element. This is confirmed by the RMSE values between those structuring elements and their respective targets, as presented by Table 1. More particularly, $\mathcal{L}$Morph always achieves the lowest RMSE value for dilation tasks, while $\mathcal{S}$Morph succeeds better on erosion. In any case, the loss at convergence of the $\mathcal{S}$Morph network is almost consistently lower than that of the $\mathcal{L}$Morph network by one or two orders of magnitude, and by two to three with respect to the *PConv* network.

Figure 5 displays the structuring elements learned by the *PConv*, $\mathcal{L}$Morph and $\mathcal{S}$Morph layers for all 6 target structuring elements for closing and opening operations. This time, since each network comprises of two morphological layers (and a scale/bias *Conv* 1x1x1 layer), it is worth mentioning that the two filters evolve independently from each other. Nevertheless, the two morphological layers are expected to learn filters having exactly the same shape, with parameter p or α of opposite signs, once training converges.

As can be seen on Figure 5, the *PConv* network always succeeds at learning the right morphological operation: the first (resp. second) layer converges to $p > 0$
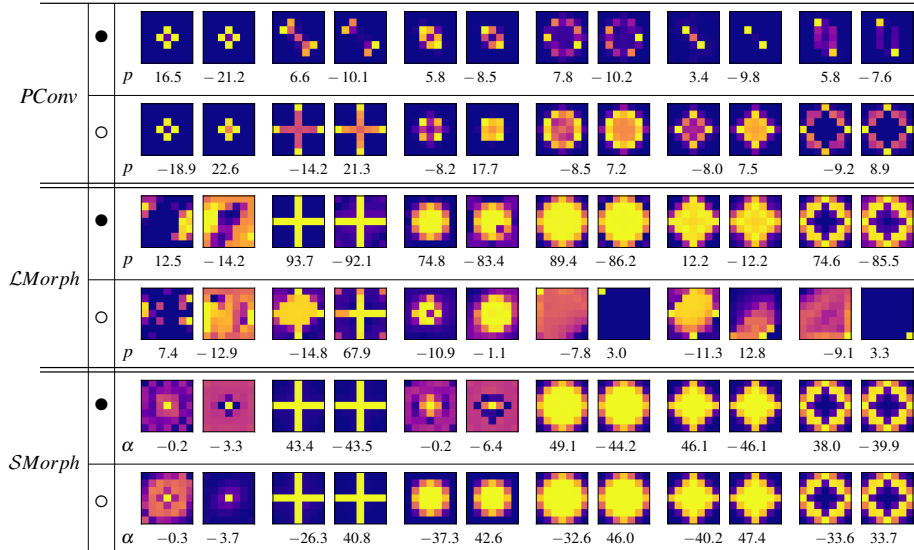
Fig. 5: Learned structuring elements (with corresponding p or α value for each layer) for $PConv$, $\mathcal{L}$Morph and $\mathcal{S}$Morph layers on closing ● and opening ○ tasks.

(resp. $p < 0$) for the closing, and the opposite behavior for the opening. However, quite often $|p| < 10$, indicating that the layer is applying pseudo-dilation or pseudo-erosion only. In addition, the learned structuring element suffers again from the hollow effect for the opening, and does not find the correct shape for the closing. The $\mathcal{L}$Morph network succeeds in learning the correct operation and shape for the closing operation with large target structuring elements (all but *cross3* and *disk2*). For the opening operation however, it consistently fails at retrieving the shape of the target structuring element. This counter-performance is up to now unexplained. The $\mathcal{S}$Morph network also struggles with small structuring elements for both the opening and closing, but perfectly recovers large ones. The edge case of small target structuring elements could come from the scale/bias $Conv$ 1x1x1, which over-compensates for the gain or loss of average pixel intensities. Thus, when back-propagating the error during the learning phase, all filters behind the scale/bias $Conv$ 1x1x1 layer might not learn the right operation with a parameter p or α not converging toward the correct sign domain.

Table 2 presents the MSE loss at convergence and RMSE value between the learned filters and the target structuring elements for closing and opening scenarios. Except for the aforementioned edge case, the $\mathcal{S}$Morph layer consistently achieves the lowest loss value and RMSE for opening, while the best results for closing are obtained either by $\mathcal{L}$Morph or $\mathcal{S}$Morph layers. Overall, apart from small structuring elements for closing or opening operations, the proposed $\mathcal{S}$Morph layer outperforms its $PConv$ and $\mathcal{L}$Morph counterparts. Last but not least, it should also be noted that the $\mathcal{S}$Morph layer is also numerically more stable. As a

Table 2: MSE loss at convergence and RMSE between the learned structuring elements displayed by Figure 5 and the target for $PConv$, $\mathcal{L}$Morph and $\mathcal{S}$Morph layers on closing ● and opening ○ tasks. Best (lowest) results are in bold.

			cross3	cross7	disk2	disk3	diamond3	complex	
PConv	●	LOSS	0.09	5.1	1.7	1.0	5.4	4.9	$\times 10^{-3}$
		RMSE	**0.83** | **0.81**	3.94 | 3.97	2.82 | 3.07	4.44 | 4.64	4.34 | 4.51	3.68 | 3.65	
	○	LOSS	**0.86**	0.76	4.8	3.9	4.4	2.3	$\times 10^{-4}$
		RMSE	**0.91** | **0.32**	1.45 | 0.87	2.70 | 2.49	3.10 | 2.52	3.15 | 2.43	2.10 | 2.08	
$\mathcal{L}$Morph	●	LOSS	**0.5**	0.13	**2.0**	0.18	6.2	0.14	$\times 10^{-4}$
		RMSE	3.58 | 5.12	**0.01** | 0.63	**0.14** | **1.16**	**0.07** | 0.40	0.75 | 0.95	**0.47** | 0.99	
	○	LOSS	8.7	36	2.0	2.7	4.6	2.5	$\times 10^{-3}$
		RMSE	3.17 | 4.74	2.90 | 1.20	1.99 | 1.61	2.81 | 5.56	1.66 | 3.54	2.72 | 3.98	
$\mathcal{S}$Morph	●	LOSS	1700	**0.31**	4400	**0.42**	**0.37**	**0.17**	$\times 10^{-6}$
		RMSE	2.10 | 3.59	0.14 | **0.08**	2.78 | 3.68	0.08 | **0.01**	**0.08** | **0.10**	0.52 | **0.52**	
	○	LOSS	4700	**0.14**	**0.22**	**0.25**	**0.24**	**0.09**	$\times 10^{-6}$
		RMSE	3.40 | 1.63	**0.18** | **0.02**	**0.05** | **0.01**	**0.02** | **0.02**	**0.02** | **0.01**	**0.51** | **0.52**	

matter of fact, raising to the power of p in the $PConv$ and $\mathcal{L}$Morph layers faster induces floating point accuracy issues.

5   Conclusion

We present two new morphological layers, namely $\mathcal{L}$Morph and $\mathcal{S}$Morph. Similarly to the $Pconv$ layer of Masci *et al.* [10], the former relies on the asymptotic properties of the CHM to achieve grayscale erosion and dilation. The latter instead relies on the α-softmax function to reach the same goal, thus sidestepping some of the limitations shared by the $PConv$ and $\mathcal{L}$Morph layers (namely being restricted to strictly positive inputs, rescaled in the range $[1, 2]$, and positive structuring functions). In order to evaluate the performances of the proposed morphological layers, we applied in turn dilation, erosion, closing and opening on the whole MNIST dataset images, with 6 target structuring elements of various sizes and shapes, and the morphological layers were trained to retrieve those structuring elements. Qualitative and quantitative comparisons demonstrated that the $\mathcal{S}$Morph layer overall outperforms both the $PConv$ and $\mathcal{L}$Morph layers. Future work includes investigating the edge cases uncovered for both proposed layers, as well as integrating them into more complex network architectures and evaluating them on concrete image processing applications.

References

1. Angulo, J.: Pseudo-morphological image diffusion using the counter-harmonic paradigm. In: International Conference on Advanced Concepts for Intelligent Vision Systems. pp. 426–437. Springer (2010)

2. Bullen, P.S.: Handbook of means and their inequalities, vol. 560. Springer Science & Business Media (2013)
3. Charisopoulos, V., Maragos, P.: Morphological perceptrons: geometry and training algorithms. In: International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing. pp. 3–15. Springer (2017)
4. Davidson, J.L., Ritter, G.X.: Theory of morphological neural networks. In: Digital Optical Computing II. vol. 1215, pp. 378–388. International Society for Optics and Photonics (1990)
5. Franchi, G., Fehri, A., Yao, A.: Deep morphological networks. Pattern Recognition **102**, 107246 (2020)
6. Hassoun, M.H., et al.: Fundamentals of artificial neural networks. MIT press (1995)
7. Lange, M., Zühlke, D., Holz, O., Villmann, T., Mittweida, S.G.: Applications of Lp-Norms and their Smooth Approximations for Gradient Based Learning Vector Quantization. In: ESANN. pp. 271–276 (2014)
8. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. nature **521**(7553), 436–444 (2015)
9. LeCun, Y., Cortes, C., Burges, C.J.: The MNIST database of handwritten digits. URL http://yann. lecun. com/exdb/mnist **10**(34), 14 (1998)
10. Masci, J., Angulo, J., Schmidhuber, J.: A learning framework for morphological operators using counter–harmonic mean. In: International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing. pp. 329–340. Springer (2013)
11. Mellouli, D., Hamdani, T.M., Ayed, M.B., Alimi, A.M.: Morph-cnn: a morphological convolutional neural network for image classification. In: International Conference on Neural Information Processing. pp. 110–117. Springer (2017)
12. Mondal, R., Dey, M.S., Chanda, B.: Image restoration by learning morphological opening-closing network. Mathematical Morphology-Theory and Applications **4**(1), 87–107 (2020)
13. Nogueira, K., Chanussot, J., Dalla Mura, M., Schwartz, W.R., Santos, J.A.d.: An introduction to deep morphological networks. arXiv preprint arXiv:1906.01751 (2019)
14. Ritter, G.X., Sussner, P.: An introduction to morphological neural networks. In: Proceedings of 13th International Conference on Pattern Recognition. vol. 4, pp. 709–717. IEEE (1996)
15. Serra, J.: Image analysis and mathematical morphology. Academic Press, Inc. (1983)
16. Shen, Y., Zhong, X., Shih, F.Y.: Deep morphological neural networks. arXiv preprint arXiv:1909.01532 (2019)
17. Shih, F.Y., Shen, Y., Zhong, X.: Development of deep learning framework for mathematical morphology. International Journal of Pattern Recognition and Artificial Intelligence **33**(06), 1954024 (2019)
18. Soille, P.: Morphological image analysis: principles and applications. Springer Science & Business Media (2013)
19. Sussner, P.: Morphological perceptron learning. In: Proceedings of the 1998 IEEE International Symposium on Intelligent Control (ISIC). pp. 477–482. IEEE (1998)
20. Wilson, S.S.: Morphological networks. In: Visual Communications and Image Processing IV. vol. 1199, pp. 483–495. International Society for Optics and Photonics (1989)
21. Zhang, Y., Blusseau, S., Velasco-Forero, S., Bloch, I., Angulo, J.: Max-plus operators applied to filter selection and model pruning in neural networks. In: International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing. pp. 310–322. Springer (2019)