



ÉCOLE NATIONALE SUPÉRIEURE DES MINES
DE SAINT-ÉTIENNE

MASTER 2 :
MATHEMATICAL IMAGING AND SPATIAL PATTERN ANALYSIS

Mémoire - Rapport de stage

**Analyse du comportement et amélioration de la
convergence des réseaux de neurones morphologiques**

Auteur : Antoine BOTTEMULLER

Tuteur EMSE : Yann GAVET

Encadrants : Guillaume TOCHON (LRE)
Jesús ANGULO (CMM)



Date : 18 septembre 2023

Préface

Le présent mémoire constitue une compilation et une synthèse des travaux de recherche entrepris pendant la période de mon stage de fin d'études s'étalant de mars à septembre 2023. Cette initiative s'inscrit dans le cadre de ma participation au programme de formation de niveau Master 2 intitulé "Mathematical Imaging and Spatial Pattern Analysis", dispensé par l'École Nationale Supérieure des Mines de Saint-Étienne (EMSE). Parallèlement, ce mémoire assume également le rôle de rapport final en vue de l'accomplissement des exigences du diplôme d'ingénieur généraliste décerné par l'École Nationale Supérieure des Mines d'Alès (EMA).

Le projet « Réseaux de neurones morphologiques » dont fait l'objet ce stage est porté conjointement par l'École Pour l'Informatique et les Techniques Avancées (EPITA), à travers Monsieur Guillaume TOCHON au sein du Laboratoire de Recherche de l'EPITA (LRE), et par l'École Nationale Supérieure des Mines de Paris (Mines Paris - PSL), à travers Monsieur Jesús ANGULO au sein du Centre de Morphologie Mathématique (CMM). Depuis plusieurs années, ce projet s'est développé continuellement entre différentes mains et a fait l'objet de plusieurs publications dans des revues à comité de lecture avant d'aboutir à ce stage, dont l'objectif est l'exploration et l'amélioration des réseaux déjà développés, et s'achevant sur ce mémoire.

Plusieurs personnes ont contribué à la réalisation et au bon déroulement de ce stage. Je tiens tout d'abord à remercier Valentin PENAUD-POLGE, doctorant au CMM, pour m'avoir initialement partagé l'offre associée à ce projet, ainsi que pour les précieux conseils donnés avant le début du stage. Je souhaite ensuite remercier les professeurs et les doctorants du LRE pour leur soutien et leur collaboration tout le long de mon séjour au sein du laboratoire, en particulier Caroline MAZINI-RODRIGUES, Adam DESCARPENTRIES, Charles VILLARD, Michaël ROYNARD et Baptiste ESTEBAN pour leur aide quotidienne apportée et les agréables moments partagés.

J'aimerais également remercier Romain HERMARY, de qui j'hérite ce projet, pour la présentation de ses travaux et pour son aide continue apportée dans la compréhension du code. Je tiens ensuite à remercier Yann GAVET et Léo THÉODON de l'EMSE pour leur aide et leurs précieux conseils apportés durant mon séjour au LRE. Enfin, je souhaite remercier de tout cœur Guillaume TOCHON et Jesús ANGULO pour m'avoir supervisé et soutenu tout le long de ce stage, et sans qui ce projet n'existerait.

Antoine BOTTEMULLER

Fait au Kremlin-Bicêtre (94270), le 18 septembre 2023

Table des matières

1	Introduction	1
2	État de l'art	3
2.1	Morphologie mathématique	3
2.1.1	Morphologie binaire	3
2.1.2	Morphologie en niveaux de gris	4
2.2	Structure des réseaux morphologiques	6
2.2.1	Construction d'une couche morphologique	6
2.2.2	Architecture d'un réseau morphologique	8
2.3	Les différentes couches morphologiques	10
2.3.1	Couche p -convolution	10
2.3.2	Couche \mathcal{LM} orph	11
2.3.3	Couche \mathcal{SM} orph	12
2.4	Efficacité des réseaux morphologiques	13
2.4.1	Méthodologie et description des expériences	13
2.4.2	Comparaison de la précision des réseaux	15
2.4.3	Problèmes de convergence des réseaux	19
3	Amélioration des réseaux	21
3.1	Réseau \mathcal{SM} orphNet-Tanh	21
3.1.1	Contexte et motivations	21
3.1.2	Formalisation	22
3.1.3	Comparaison avec \mathcal{SM} orphNet	23
3.2	Partage de poids	26
3.2.1	Contexte et motivations	26
3.2.2	Partage de poids dur	27
3.2.3	Partage de poids doux	28
3.2.4	Comparaison de l'efficacité	31
3.3	Contraintes sur les poids	34
3.3.1	Contexte et motivations	34
3.3.2	Contraintes géométriques sur les noyaux w	36
3.3.3	Contraintes sur les paramètres de contrôle α	42
3.3.4	Application et comparaison de l'efficacité	45
3.4	Modulation de l'initialisation	47
3.4.1	Motivations et principe	47
3.4.2	Comparaison de l'efficacité	48
4	Analyse des réseaux	50
4.1	Réseaux à deux couches pour une opération primaire	50
4.1.1	Motivations	50
4.1.2	Avec l'érosion	51

4.1.3	Avec la dilatation	52
4.1.4	Discussion	53
4.2	Impact du jeu de données d'entraînement	54
4.2.1	Description des banques MNIST et FashionMNIST	54
4.2.2	Différences de résultats entre les jeux	55
4.2.3	Résultats avec les jeux en négatif	56
4.3	Relation entre complexité et échec de convergence	57
4.3.1	Complexité de la forme de la fonction structurante cible	57
4.3.2	Complexité de la configuration du réseau	59
4.4	Autres rapides analyses	60
4.4.1	Importance des α dans la convergence des réseaux	60
4.4.2	Application au débruitage	61
5	Conclusion	63
Références		65

1 Introduction

Crée en 1964 par Georges Matheron et Jean Serra à l'École des Mines de Paris, la morphologie mathématique est un sous-domaine à la fois théorique et technique des mathématiques et de l'informatique, étroitement lié à l'algèbre, à la théorie des treillis et à la topologie, et qui est en particulier appliquée dans le domaine du traitement d'images [18]. Il met en œuvre des techniques de filtrage non linéaire à l'aide de formes géométriques, telles qu'un carré, une croix, un disque, ou toute autre forme encore, appelées « éléments structurants », qui sondent le contenu des images et en modifient les propriétés morphologiques [19].

Les opérations fondamentales de la morphologie mathématique sont l'érosion et la dilatation (fig. 1). En les combinant, on peut créer des opérations plus avancées, telles que l'ouverture (érosion puis dilatation) et la fermeture (dilatation puis érosion), ou encore le gradient morphologique (dilatation moins érosion). Ces opérations de morphologie mathématique sont typiquement appliquées dans des contextes où l'élément structurant est binaire, et où l'image elle-même peut être soit binaire, soit en niveaux de gris. Ces opérations peuvent également être généralisées à des fonctions structurantes, c'est-à-dire des éléments structurants en niveaux de gris [5, 20].



Figure 1 – Exemple d'une érosion \ominus (à gauche) et d'une dilatation \oplus (à droite) par une fonction structurante B circulaire sur une image en niveaux de gris de la banque MNIST.

La morphologie mathématique démontre une grande efficacité dans diverses applications de traitement d'images, telles que le débruitage et la détection d'objets [15, 8]. Cependant, ces performances sont souvent tributaires du niveau d'expertise de l'utilisateur, ainsi que de sa capacité à déterminer la séquence appropriée des opérations fondamentales et à définir adéquatement les éléments structurants associés à la tâche en question. Par conséquent, l'automatisation de la définition des types d'opérations morphologiques et de leurs éléments structurants pour des applications spécifiques représente une perspective particulièrement séduisante.

Récemment, plusieurs initiatives ont vu le jour en vue d'intégrer l'apprentissage automatique de séquences d'opérations morphologiques au sein des réseaux de neurones convolutionnels [13, 7, 2]. L'objectif est de remplacer les couches de convolution classiques par des couches morphologiques, capables d'apprendre et d'appliquer des érosions ou des dilatations. Les poids des filtres en jeu jouent alors un rôle analogue à celui des fonctions structurantes correspondantes.

Cependant, en raison du caractère localement non différentiable des opérations morphologiques, qui reposent sur des minimas (érosion) ou maximas (dilatation) ensemblistes, des approches détournées doivent être mises en œuvre pour assurer une convergence optimale du réseau durant la phase d'apprentissage. Deux stratégies ont émergé dans la littérature : l'emploi d'approximations continues et différentiables des opérateurs de minimum et de maximum [13, 21, 9, 7], ou le traitement de la couche morphologique de manière similaire à la gestion des opérations localement non différentiables dans les architectures classiques (comme le max-pooling) [14, 16].

Dans le cadre de l'emploi d'approximations différentiables du min et du max, la première couche morphologique fonctionnelle, la *pConv*, a été proposée en 2012 en utilisant la formule de la moyenne contre-harmonique [13]. Plus récemment encore, ont été proposées dans [9], et étendu dans [7], deux nouvelles couches morphologiques, la *LMorph* et la *SMorph*, qui reposent sur d'autres formules d'approximation. Ces couches sont capables d'apprendre de manière très performante les opérations d'érosion, de dilatation, d'ouverture et de fermeture ainsi que l'élément structurant associé. Cependant, certaines propriétés de ces couches restent mal comprises, voire incomprises, en particulier dans les cas où l'apprentissage est un échec.

L'objectif ici est donc d'explorer plus en profondeur le comportement de ces couches morphologiques différentiables, aussi bien sur des aspects théoriques (comportement asymptotique, morphologie de l'espace de la fonction de perte, convergence durant la phase d'apprentissage, impact de contraintes géométriques, etc.) que pratiques (analyse de la convergence des réseaux, des résultats d'apprentissage sous contrainte, de l'impact d'un partage de poids, etc.). Il s'agira en particulier d'apporter des solutions aux problèmes de convergence des réseaux associés.

Dans leur article, R. Hermary et al. [7] mettent en évidence de meilleures performances de convergence pour les couches *SMorph* par rapport aux autres couches morphologiques existant. Les études et résultats dont ce rapport fait la synthèse seront ainsi focalisées sur ces couches-ci. Les solutions apportées pour améliorer la convergence des réseaux associés à ces couches, dits *SMorphNet*, pourront aussi bien toucher la définition même des couches, que la structure des réseaux associés construits, ou que la définition de la fonction de perte lors de la phase d'entraînement.

Dans un premier temps seront passées en revue les différentes méthodes de la littérature : la définition des opérateurs classiques de la morphologie mathématique, la structure des différentes couches morphologiques proposées existant, l'efficacité de convergence des réseaux correspondants vers la structure neuronale cible, ainsi que les différents cas d'échec persistants. Dans un second temps seront présentées de manière théorique les différentes contributions apportées ainsi que les résultats pratiques de leur application, avant de comparer la convergence des réseaux améliorés à celle des réseaux existant de la littérature dans un dernier temps.

2 État de l'art

2.1 Morphologie mathématique

2.1.1 Morphologie binaire

La morphologie mathématique a, en 1964, d'abord été introduite sous sa forme binaire [19]. Dans ce cadre, les images peuvent être définies soit de manière ensembliste, où elles et les objets contenus en elles sont définis comme des parties d'un ensemble E structuré comme un groupe pour l'addition, soit, par équivalence, de manière fonctionnelle booléenne, où les images sont définies comme fonctions sur un sous-ensemble de E à valeurs dans $\{0, 1\}$, où 1 est associé aux objets et 0 au fond de l'ensemble de définition [18]. Dans un contexte discret, tel que les images, on a $E = \mathbb{Z}^2$.

Les opérations fondamentales de la morphologie mathématique sont l'érosion et la dilatation. Dans le cadre binaire, elles sont définies de manière ensembliste sur E .

Soit $B \subset E$ l'élément structurant, servant de structure aux opérations morphologiques. On définit B_x comme l'ensemble B translaté de $x \in E$: $B_x = \{b + x \mid b \in B\}$
On définit \check{B} comme l'ensemble symétrique de B : $\check{B} = \{-b \mid b \in B\}$

- L'érosion $\epsilon_B(X)$ d'une partie X de E par l'élément structurant B est définie par :

$$\epsilon_B(X) = X \ominus B = \{x \mid B_x \subset X\} \quad (1)$$

- La dilatation $\delta_B(X)$ de X par B est, quant à elle, définie par :

$$\delta_B(X) = X \oplus B = \{x \mid \check{B}_x \cap X \neq \emptyset\} \quad (2)$$

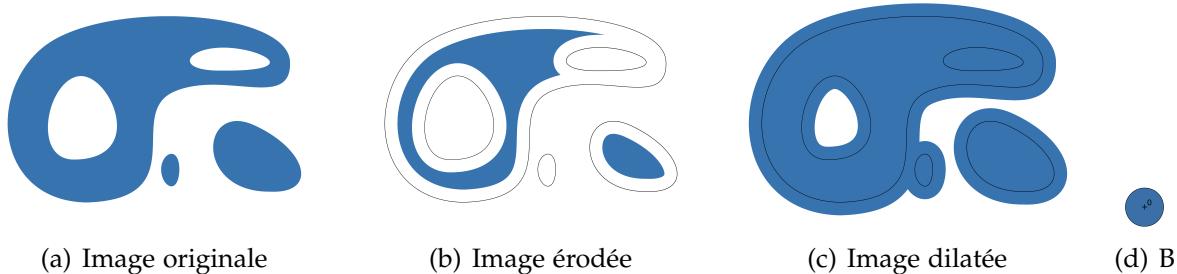


Figure 2 – Exemple de l'érosion \ominus 2(b) et de la dilatation \oplus 2(c) d'une image binaire 2(a) par un élément structurant binaire B en forme de disque centré en $(0, 0)$, fig. 2(d).

En combinant l'érosion (1) et la dilatation (2), illustrées fig. 2, on peut créer des opérations plus avancées, telles l'ouverture \circ : $\gamma_B(X) = \delta_B(\epsilon_B(X))$; la fermeture \bullet : $\phi_B(X) = \epsilon_B(\delta_B(X))$; ou encore le gradient morphologique : $\text{grad}_B(X) = \delta_B(X) - \epsilon_B(X)$.

2.1.2 Morphologie en niveaux de gris

La morphologie mathématique pour les images à niveaux a été développée après la morphologie binaire, dans le but de généraliser cette technique aux images fonctionnelles [5, 20]. Dans ce cas plus général, les images et éléments structurants sont considérés comme des applications sur une partie de E à valeurs dans une partie d'un autre ensemble tel que \mathbb{Z} ou \mathbb{R} pour les images à tons de gris, \mathbb{Z}^3 ou \mathbb{R}^3 pour les images en couleurs, ou tout autre ensemble construit comme groupe pour l'addition.

Les opérations fondamentales de la morphologie mathématique sont l'érosion et la dilatation. On étudiera par la suite uniquement le cas des images à tons de gris.

L'image fonctionnelle est notée $f : E \rightarrow G$. L'élément structurant est, comme l'image, une fonction notée $b : E \rightarrow G$, avec G groupe pour l'addition (\mathbb{Z} ou \mathbb{R}).

- L'érosion $\epsilon_b(f)$ de f par b est l'image $\epsilon_b(f) : E \rightarrow G$ définie pour tout $x \in E$ par :

$$\epsilon_b(f)(x) = (f \ominus b)(x) = \min_{y \in E} \{f(y) - b(y-x)\} \quad (3)$$

- La dilatation $\delta_b(f)$ de f par b est, quant à elle, l'image définie pour tout $x \in E$ par :

$$\delta_b(f)(x) = (f \oplus b)(x) = \max_{y \in E} \{f(y) + b(x-y)\} \quad (4)$$

Les processus d'érosion et de dilatation sont tous deux illustrés sur un exemple fig. 3.

On définit l'ouverture et la fermeture par les mêmes combinaisons entre l'érosion et la dilatation que dans la morphologie binaire. L'ouverture fonctionnelle \circ est ainsi définie par : $\gamma_b(f) = \delta_b(\epsilon_b(f))$; et la fermeture fonctionnelle \bullet par : $\phi_b(f) = \epsilon_b(\delta_b(f))$. Il en est de même pour le gradient morphologique : $\text{grad}_b(f) = \delta_b(f) - \epsilon_b(f)$.

Un lien direct entre la morphologie en niveaux de gris et la morphologie binaire peut se faire. Soit $B \subset E$ le support de l'élément structurant tel que défini dans la morphologie binaire. En définissant b comme fonction structurante plate telle que, pour tout $x \in E$:

$$b(x) = \begin{cases} 0 & \text{si } x \in B \\ -\infty & \text{sinon} \end{cases},$$

les formules 3 et 4 permettent à la morphologie en niveaux de gris d'avoir le même comportement que la morphologie binaire, pour des images binaires définies comme fonctions sur E à valeurs dans $\{0, 1\}$.

La formule de l'érosion devient alors : $\epsilon_b(f)(x) = \min_{y \in B_x} \{f(y)\}$
Et celle de la dilatation devient ainsi : $\delta_b(f)(x) = \max_{y \in \check{B}_x} \{f(y)\}$

$\min_{y \in B_x} \{f(y)\}$ correspond à l'inclusion de B_x dans les objets de l'image f au point x , et $\max_{y \in \check{B}_x} \{f(y)\}$ à la non vacuité de l'intersection entre \check{B}_x et les objets de f en x .

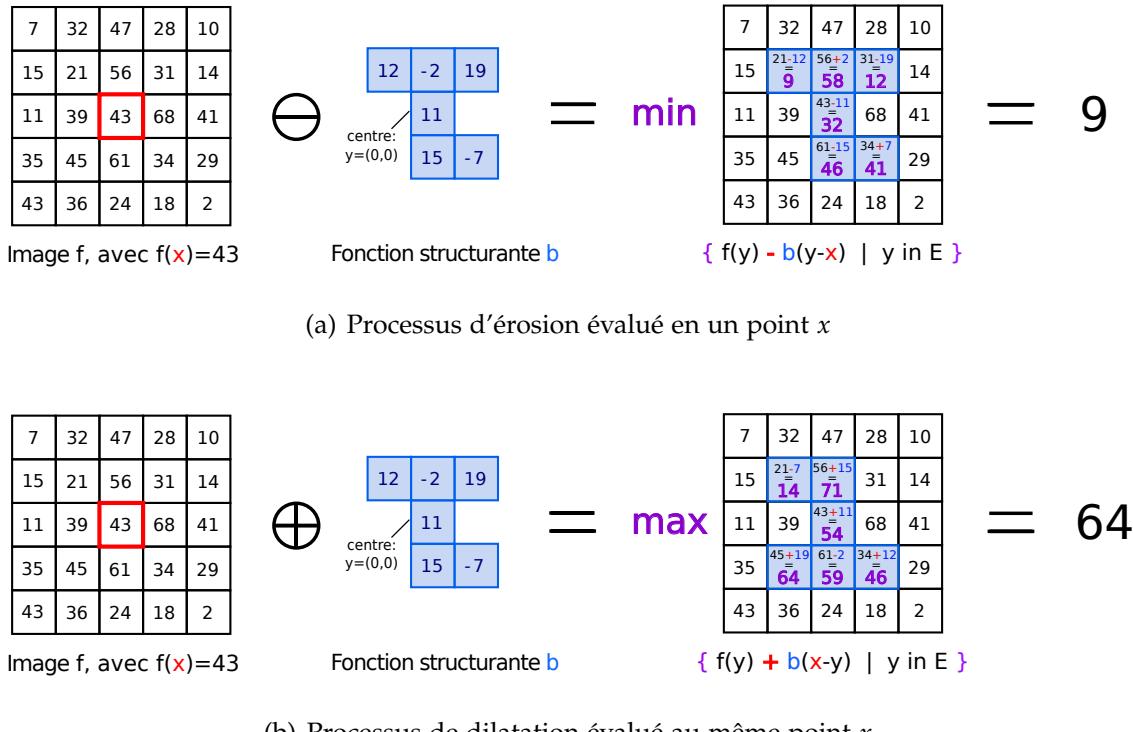


Figure 3 – Exemple du processus de l'érosion \ominus 3(a) et de celui de la dilatation \oplus 3(b) sur une image f en niveaux de gris, en un point $x \in \mathbb{Z}^2$, et par un élément structurant b .

Remarque pour la suite :

En morphologie mathématique, que ce soit en binaire ou en niveau de gris, la dilatation considère la symétrie $\bar{b} = -b$ de l'élément structurant b , contrairement à l'érosion qui considère simplement b [5]. En niveaux de gris, cela est illustré par l'inversion de x avec y , en argument de l'élément structurant b , entre l'érosion et la dilatation. Voir les formules précédentes, ainsi que l'illustration fig. 3. Il s'agit d'un élément important dans la suite de ce rapport, où les différentes formules d'approximation de ces opérations morphologiques considèrent soit uniquement b soit uniquement sa symétrique \bar{b} pour ces deux opérations à la fois, afin de conserver la continuité des formules lors du passage d'un effet d'érosion à un effet de dilatation, et inversement.

De plus, dans la morphologie en niveaux de gris, la seconde différence de construction entre l'ensemble affecté au min de l'érosion et celui affecté au max de la dilatation, est le signe devant l'élément structurant b . Voir formules associées 3 et 4. Cette seconde différence joue également un rôle important pour la suite, dans la définition des formules d'approximation de ces opérateurs morphologiques.

2.2 Structure des réseaux morphologiques

2.2.1 Construction d'une couche morphologique

La principale motivation dans le développement de couches morphologiques, initié à la fin des années 1980 [17, 4], est l'intégration d'opérateurs morphologiques automatiques, capables d'apprendre et d'appliquer des érosions ou des dilatations sur des images données, au sein même des réseaux de neurones convolutionnels [11], en remplacement des couches de convolution classiques dans ces réseaux-ci.

Ces nouvelles couches jouent ainsi, comme pour les couches classiques, le rôle de *filtres* appliqués aux images en entrée du réseau, dont le noyau à poids variables est de taille constante, mais qui sont ici régis par une fonction particulière d'approximation de l'érosion et de la dilatation (approximation lisse nécessaire due à la non-dérivabilité du min et du max des opérations d'érosion et de dilatation). Le noyau de ces couches morphologiques joue alors un rôle analogue à celui des fonctions structurantes correspondantes, dont il cherche à imiter le comportement [2]. Un réseau de neurones composé de telles couches est ainsi appelé « réseau de neurones morphologique ».

Soit $f : I \subseteq E \rightarrow \mathbb{R}$ une image, fonction sur une partie dénombrable I de E à valeurs dans \mathbb{R} , et soit $w : W \subseteq E \rightarrow \mathbb{R}$ le noyau d'une couche morphologique, également défini sur une partie dénombrable W de E . Comme précédemment, on écrit \check{W}_x pour désigner l'ensemble symétrique de W translaté de $x \in E$: $\check{W}_x = \{-w + x \mid w \in W\}$.

Attention : Ici, nous parlons de « noyau » ou de « filtre » d'une couche pour désigner la fonction w associée telle que définie ci-dessus et qui renvoie les poids $(w_y)_{y \in W} \in \mathbb{R}^{|W|}$ variables du filtre sur le support $W \subseteq E$. La « couche morphologique », quant à elle, est l'objet contenant à la fois ce noyau w ainsi qu'une fonction de transformation T , qui prend en argument w et qui, à partir d'une image f en entrée de couche, renvoie une nouvelle image (également application de I dans \mathbb{R}) en fonction de f et w .

D'une manière générale, la fonction de transformation T peut être décomposée en deux étapes de transformations, l'une s'inscrivant dans l'autre. La première, locale, considère f et w localement en produisant, pour un point $x \in I$ donné, des *familles de réels* à travers la mise en relation de la valeur $f(y)$ en $y \in \check{W}_x$ avec la valeur $w(x - y)$ au point $x - y$ symétrique de $x + y$ par rapport à x . La seconde étape, plus globale, met en relation, pour ce même $x \in I$, l'ensemble des éléments de ces *familles de réels* produites. Elle renvoie alors la valeur en ce point x associée à la nouvelle image. Voir fig. 4.

Par exemple, pour une couche de convolution classique, la relation locale R_l entre les $f(y)$ et $w(x - y)$ est le produit (\times), et la relation globale R_g entre les éléments de la famille de réels produite est l'addition sur $y \in \check{W}_x$. La nouvelle image f_c issue de cette transformation T est ainsi définie pour tout $x \in I$ par : $f_c(x) = \sum_{y \in \check{W}_x} f(y)w(x - y)$. Pour une dilatation classique, R_l est la somme (+), et R_g est l'opération de maximum.

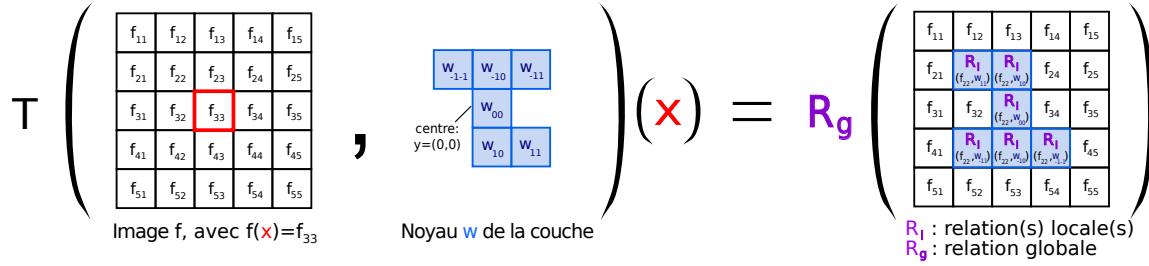


Figure 4 – Illustration de la décomposition de la fonction de transformation T en relations locales R_l et une globale R_g , sur une image f avec le noyau w et au point x .

La fonction T renvoie l'image $T(f, w)$ définie sur I dans \mathbb{R} . On écrira cette nouvelle image, sans expliciter les relations locales R_l entre les $f(y)$ et $w(x - y)$ ou globale R_g , simplement comme suit :

$$\begin{aligned} T(f, w) &: I \longrightarrow \mathbb{R} \\ x &\longmapsto T(f, w)(x) \end{aligned}$$

La fonction de transformation T , qui prend en argument le couple image-noyau (f, w) , transforme ainsi l'image f , en entrée de la couche, en une nouvelle image de même taille, en considérant, pour chacun des pixels x appartenant au support I de l'image d'entrée, le voisinage défini par le support $W \subseteq E$ du noyau w (illustration fig. 4).

La bonne définition d'une telle fonction de transformation T au sein d'une couche permettrait d'appliquer à l'image f l'effet désiré (convolution, max-pooling, morphologie, etc.), avec les poids du noyau w de la couche en question, grâce à des mises en relation locales entre $f(y)$ et $w(x - y)$ et à une relation globale sur toutes ces dernières.

Dans le cadre des réseaux de neurones, on va définir ces différentes relations de sorte que la fonction T soit *continue* et *dérivable* localement partout par rapport aux poids $(w_y)_{y \in W}$ variables du noyau w , afin de pouvoir calculer le gradient du réseau lors de la phase d'apprentissage et mettre à jour ces poids par rétropropagation du gradient.

Le but ici, dans le cadre des couches morphologiques, est alors de trouver une fonction dérivable T (et donc de définir les différentes relations locales et globale telles que décrites ci-dessus) adaptée, pour approximer au mieux l'effet des deux opérateurs morphologiques fondamentaux, l'érosion et la dilatation, et ce en une *unique* formule.

Ici, les noyaux des couches morphologiques seront toujours accompagnés d'un poids supplémentaire, noté p ou α , qui est indépendant des poids du noyau-même, et dont le rôle sera de faire une transition continue lisse entre le comportement d'érosion et celui de dilatation de la couche morphologique [7]. Pour mieux distinguer ce poids des autres, on notera la nouvelle image de T : $T(f, w, p)$. Les différentes définitions de la fonction de transformation T et l'intégration du poids p sont explicitées partie 2.3.

2.2.2 Architecture d'un réseau morphologique

Les couches morphologiques, telles que décrites précédemment, sont intégrées dans les réseaux de neurones morphologiques de la même manière que les couches de convolution dans les réseaux de neurones convolutionnels [6]. Dans le cadre de ce travail, nous construisons des réseaux comportant uniquement une ou plusieurs couches morphologiques successives, et pouvant comporter en plus (ou non) des couches de rééchelonnage de l'image, constantes ou avec des poids variables appris par le réseau lors de la phase d'entraînement, telles qu'une couche de convolution de taille 1 [14, 2].

Dans un tel réseau morphologique :

- l'entrée est une matrice 3D représentant un ensemble d'images 2D en niveaux de gris définies sur une partie I de \mathbb{Z}^2 , de taille $n \times l \times h$, avec n le nombre d'images, l la longueur verticale des images et h leur hauteur ;
- les images sont filtrées successivement par les différentes couches du réseau, à travers une succession ordonnée de couches morphologiques et de couches de rééchelonnage si incluses dans le réseau ;
- la sortie est une matrice 3D représentant l'ensemble des n images 2D filtrées, de même taille $n \times l \times h$ que la matrice d'entrée.

Le nombre de couches morphologiques à intégrer dans ces réseaux dépend de l'opération que l'on souhaite simuler ou du contexte dans lequel ils s'inscrivent.

Typiquement, pour des opérations d'érosion ou de dilatation, le réseau comportera une unique couche morphologique, accompagnée ou non de couches de rééchelonnage. Pour des opérations d'ouverture (érosion suivie de dilatation) ou de fermeture (dilatation suivie d'érosion), chaque couche morphologique ne pouvant simuler le comportement que de l'une des deux opérations morphologiques fondamentales, le réseau comportera deux couches morphologiques successives, dans le but que l'une simule le comportement d'une érosion, et que l'autre simule celui de l'érosion. Pour une opération de dessalage (i.e. supprimer les bruits poivre et sel sur les images), on voudra intégrer quatre couches morphologiques successives, pour tenter d'imiter le comportement d'une ouverture suivie d'une fermeture, ou inversement.

Il sera également possible de construire des réseaux comportant plusieurs couches morphologiques, dans le cadre d'une simple opération d'érosion ou de dilatation, afin de mieux comprendre leur fonctionnement et d'étudier le comportement et la convergence de ces réseaux durant la phase d'apprentissage. Un des points d'intérêt est l'étude de la forme que les noyaux adopteront pour décomposer en plusieurs parties une opération qu'une seule couche suffit à simuler. A l'inverse, il pourra être également pertinent d'étudier le comportement de réseaux composés d'une seule couche, pour des opérations plus complexes que l'érosion ou la dilatation.

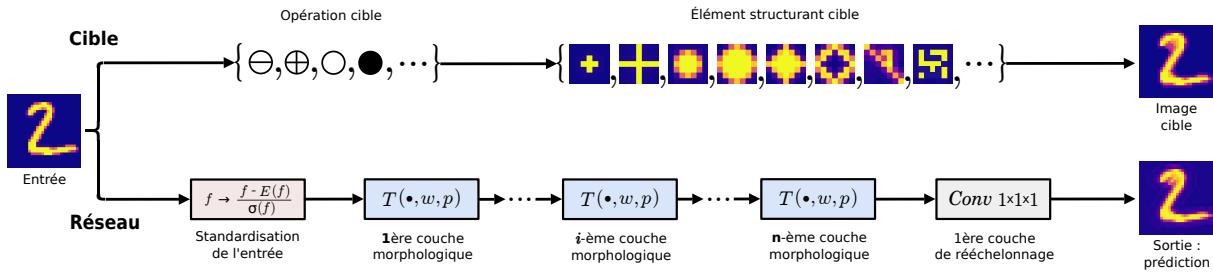


Figure 5 – Illustration de l'architecture type des réseaux de neurones morphologiques.
Cas où l'opération et l'élément structurant cibles sont définis.

De plus, l'intégration ou non de couches de rééchelonnage des valeurs des images dépend des besoins et du comportement de convergence du réseau. Un filtre de rééchelonnage peut être linéaire (normalisation linéaire dans un intervalle $[a, b]$, standardisation, etc.) ou de forme plus complexe (logarithmique, tangente hyperbolique, etc.), et peut être constant (fonction déterministe invariante) ou comportant des poids variables appris par le réseau lors de l'apprentissage. Dans ce dernier cas, les filtres linéaires sont typiquement de la forme : $f \rightarrow f \times s + b$, avec s l'échelle (« scale ») et b le biais (« bias »), tous deux formant les deux poids de cette couche de rééchelonnage. Il s'agit donc d'une couche de convolution classique, de taille $1 \times 1 \times 1$.

Hermay et al. [7] ont montré que l'ajout d'un filtre de standardisation des images en entrée du réseau, accompagnée d'une couche de rééchelonnage linéaire de l'image (convolution) en sortie de ce dernier, permet une meilleure convergence lors de la phase d'apprentissage. Ainsi, l'architecture type des réseaux de neurones morphologiques qu'on utilisera est illustrée fig 5. Elle définit d'abord un filtre de standardisation des images, puis n couches morphologiques successives, pour finir avec une couche de convolution de taille 1 pour rééchelonner l'image en sortie. L'illustration montre le cas où l'opération et l'élément structurant cibles sont définis. Mais ils peuvent ne pas l'être, comme dans le cas du dessalage, où l'on considère simplement des images cibles vers lesquelles les prédictions du réseau cherchent à converger.

A noter que le réseau construit n'est informé ni de l'opération cible, ni de l'élément structurant cible. Le réseau converge librement et se met à jour uniquement par rapport à l'image de sortie cible ; les poids (les w_i du noyau w et le paramètre de contrôle p/α) des couches morphologiques et ceux des couches de rééchelonnage se mettant à jour sur la base de l'erreur (MSE) entre l'image cible et l'image prédictive par le réseau. Ainsi, les couches morphologiques ne convergent pas nécessairement vers l'état espéré, et chacun n'adopte pas forcément un comportement morphologique (soit érosion ou soit dilatation) tel que souhaité, dû au lissage continue obligatoire de la transition entre un comportement d'érosion ou de dilatation de la couche. Là est tout l'enjeu d'une bonne définition de la fonction de transformation T des couches et de l'analyse du comportement et de la convergence des réseaux construits.

2.3 Les différentes couches morphologiques

2.3.1 Couche p -convolution

La toute première couche morphologique a été introduite par Masci et al. [13] suite aux travaux de Angulo et al. [1]. Il s'agit de la couche p -convolution. Dans une telle couche, l'approximation, qui définit T , des opérations d'érosion et de dilatation se base sur la moyenne contre-harmonique (CHM), également dite de Lehmer [3].

La CHM du vecteur $\mathbf{x} = (x_1, \dots, x_n) \in (\mathbb{R}_+)^n$ pondéré par $\mathbf{w} = (w_1, \dots, w_n) \in (\mathbb{R}_+)^n$ et d'ordre $p \in \mathbb{R}$ est :

$$CHM(\mathbf{x}, \mathbf{w}, p) = \frac{\sum_{i=1}^n w_i x_i^p}{\sum_{i=1}^n w_i x_i^{p-1}} \quad (5)$$

Étant donné que les vecteurs \mathbf{x} et \mathbf{w} ont des coordonnées positives, c'est le maximum (resp. minimum) des coordonnées de \mathbf{x} qui domine le numérateur et le dénominateur de l'équation 5 lorsque p tend vers $+\infty$ (resp. $-\infty$). On a ainsi aux limites :

$$\lim_{p \rightarrow -\infty} CHM(\mathbf{x}, \mathbf{w}, p) = \min_{i \in \llbracket 1, n \rrbracket} x_i \quad \lim_{p \rightarrow +\infty} CHM(\mathbf{x}, \mathbf{w}, p) = \max_{i \in \llbracket 1, n \rrbracket} x_i$$

La formule de la p -convolution se base sur celle de la CHM afin d'avoir le même comportement asymptotique que ce dernier en minimum ou maximum des coordonnées de \mathbf{x} . La p -convolution $pConv(f, w, p)$ d'une image $f : I \subseteq \mathbb{Z}^2 \rightarrow \mathbb{R}_+$ au pixel $x \in I$ avec un noyau de convolution positif $w : W \subseteq \mathbb{Z}^2 \rightarrow \mathbb{R}_+$ et avec $p \in \mathbb{R}$ est définie par :

$$pConv(f, w, p)(x) = (f *_p w)(x) = \frac{\sum_{y \in \check{W}_x} f(y)^p w(x-y)}{\sum_{y \in \check{W}_x} f(y)^{p-1} w(x-y)} \quad (6)$$

L'étude des propriétés asymptotiques de la $pConv$ a notamment été faite par Jesús Angulo et al. [1], qui montrent qu'aux limites de p , la $pConv$ se comporte ainsi :

$$(f *_p w)(x) \xrightarrow[p \rightarrow -\infty]{ } \min_{y \in \check{W}_x} \left\{ f(y) - \frac{1}{p} \log(w(x-y)) \right\} = \left(f \ominus \frac{1}{p} \log(w) \right)(x)$$

$$(f *_p w)(x) \xrightarrow[p \rightarrow +\infty]{ } \max_{y \in \check{W}_x} \left\{ f(y) + \frac{1}{p} \log(w(x-y)) \right\} = \left(f \oplus \frac{1}{p} \log(w) \right)(x)$$

On remarque ainsi que l'expression de la $pConv$ aux limites de p se rapproche bien de la formule (3) de l'érosion ($p \rightarrow -\infty$) et de celle (4) de la dilatation ($p \rightarrow +\infty$), à la fonction $\frac{1}{p} \log$ près sur w . Au sein d'une couche p -convolution de noyau w et $p \in \mathbb{R}$, la fonction de transformation T s'exprimera donc par : $T(\cdot, w, p) = pConv(\cdot, w, p)$. La $pConv$ étant dérivable partout sur \mathbb{R} par rapport aux poids w_i , le gradient peut être calculé et l'apprentissage du réseau peut se faire avec une telle couche p -convolution.

2.3.2 Couche \mathcal{LM} orph

La seconde couche morphologique a été introduite par Kirszenberg et al. [9]. Il s'agit de la couche \mathcal{LM} orph. Tout comme la p Conv, elle s'inspire de la formule de la CHM (5) afin d'adopter le même comportement asymptotique de min et de max que cette dernière. Le but de la \mathcal{LM} orph est de donner une meilleure approximation des opérations d'érosion et de dilatation que la p Conv, cette dernière présentant la fonction $\frac{1}{p} \log$ sur w dans ses expressions aux limites de p ($+\infty$ et $-\infty$).

L'opération \mathcal{LM} orph reprend ainsi la formule de la CHM, en associant au vecteur de pondération \mathbf{w} le vecteur unitaire 1 (on aura donc $\frac{1}{p} \log(w_i) = 0$ pour tout $w_i = 1$), et en associant au vecteur principal \mathbf{x} le vecteur $(f(y) + w(x - y))_{y \in \check{W}_x}$, pour tout x dans l'ensemble de définition de l'image considérée. Pour une image $f : I \subseteq \mathbb{Z}^2 \rightarrow \mathbb{R}_+$ et un noyau de couche $w : W \subseteq \mathbb{Z}^2 \rightarrow \mathbb{R}_+$, et avec $p \in \mathbb{R}$, \mathcal{LM} orph s'exprime donc ainsi au pixel $x \in I$:

$$\mathcal{LM}\text{orph}(f, w, p)(x) = \frac{\sum_{y \in \check{W}_x} (f(y) + w(x - y))^{p+1}}{\sum_{y \in \check{W}_x} (f(y) + w(x - y))^p} \quad (7)$$

On peut donc en déduire les formules de convergence asymptotique suivantes :

$$\lim_{p \rightarrow -\infty} \mathcal{LM}\text{orph}(f, w, p)(x) = \min_{y \in \check{W}_x} \{f(y) + w(x - y)\} = (f \ominus -\check{w})(x)$$

$$\lim_{p \rightarrow +\infty} \mathcal{LM}\text{orph}(f, w, p)(x) = \max_{y \in \check{W}_x} \{f(y) + w(x - y)\} = (f \oplus w)(x)$$

Tout comme pour la p Conv, on peut obtenir avec l'opération \mathcal{LM} orph soit une pseudo-dilatation ($p > 0$) soit une pseudo-érosion ($p < 0$), en fonction du signe de son paramètre de contrôle p . De plus, lorsque $p \rightarrow +\infty$, la \mathcal{LM} orph converge bien asymptotiquement vers la dilation de l'image f avec la fonction structurante w . Cependant, lorsque $p \rightarrow -\infty$, la \mathcal{LM} orph converge vers l'érosion de f avec la fonction $-\check{w}$, et non w .

Néanmoins, cela ne pose pas de problème dans un scénario d'apprentissage, car on peut facilement obtenir w à partir de sa négation $-w$ en vérifiant simplement si le signe de p est négatif. On pourrait par exemple ajouter devant w une transformation continue de p en son signe. De même, on peut facilement obtenir w à partir de sa symétrie \check{w} selon une transition douce du signe de p . Mais il faudra vérifier avant que de telles transitions n'engendrent pas de problèmes de convergence du réseau.

En s'appuyant sur la CHM comme pour la p Conv, la couche \mathcal{LM} orph proposée présente encore certaines limitations : l'image d'entrée f doit être positive et donc mise en échelle positive, et la fonction structurante (noyau) w également [9, 7].

On écrira ainsi la fonction de transformation $T : T(\cdot, w, p) = \mathcal{LM}\text{orph}(\cdot, w, p)$.

2.3.3 Couche SMorph

La dernière couche morphologique a été introduite par Hermary et al. [7]. Il s'agit de la couche *SMorph*. Elle s'inspire de la couche *LMorph*, en modifiant la formule dans l'objectif d'éviter les contraintes de positivité de l'image f en entrée et du noyau w de la couche. Pour cela, la *SMorph* prend la forme d'un α -softmax [10].

Soit $\alpha \in \mathbb{R}$ un réel. La fonction α -softmax de paramètre α est définie pour tout vecteur $\mathbf{x} = (x_i)_{i \in [\![1, n]\!]} \in \mathbb{R}^n$ par :

$$\mathcal{S}_\alpha(\mathbf{x}) = \frac{\sum_{i=1}^n x_i e^{\alpha x_i}}{\sum_{i=1}^n e^{\alpha x_i}} \quad (8)$$

Pour $\alpha = 0$, la formule devient la moyenne arithmétique des x_i . Quand le paramètre α devient très grand ($\alpha \rightarrow +\infty$), les $e^{\alpha x_i}$ deviennent très grands, et c'est la plus grande valeur des coordonnées de \mathbf{x} qui domine la somme des $e^{\alpha x_i}$, elle domine donc les sommes au dénominateur et au numérateur dans la formule du α -softmax. À l'inverse, quand α devient très petit ($\alpha \rightarrow -\infty$), les $e^{\alpha x_i}$ tendent très vite vers 0^+ , et c'est la plus petite valeur (non nulle) de \mathbf{x} qui domine la somme des $e^{\alpha x_i}$, et donc celles au dénominateur et au numérateur. Les propriétés asymptotiques sont donc :

$$\lim_{\alpha \rightarrow -\infty} \mathcal{S}_\alpha(\mathbf{x}) = \min_{i \in [\![1, n]\!]} x_i \quad \lim_{\alpha \rightarrow +\infty} \mathcal{S}_\alpha(\mathbf{x}) = \max_{i \in [\![1, n]\!]} x_i$$

L'opération *SMorph* reprend la formule du α -softmax en remplaçant le vecteur \mathbf{x} par le vecteur $(f(y) - w(x-y))_{y \in \tilde{W}_x}$. Ainsi, pour une image $f : I \subseteq \mathbb{Z}^2 \rightarrow \mathbb{R}$ et un noyau de couche $w : W \subseteq \mathbb{Z}^2 \rightarrow \mathbb{R}$, et avec $\alpha \in \mathbb{R}$, *SMorph* s'exprime donc ainsi au pixel $x \in I$:

$$\mathbf{SMorph}(f, w, \alpha)(x) = \frac{\sum_{y \in \tilde{W}_x} (f(y) + w(x-y)) e^{\alpha(f(y)+w(x-y))}}{\sum_{y \in \tilde{W}_x} e^{\alpha(f(y)+w(x-y))}} \quad (9)$$

On obtient les mêmes expressions de *SMorph* aux limites de α que pour *LMorph* aux limites de p ($+\infty$ et $-\infty$) :

$$\lim_{\alpha \rightarrow -\infty} \mathbf{SMorph}(f, w, \alpha)(x) = \min_{y \in \tilde{W}_x} \{f(y) + w(x-y)\} = (f \ominus -\check{w})(x)$$

$$\lim_{\alpha \rightarrow +\infty} \mathbf{SMorph}(f, w, \alpha)(x) = \max_{y \in \tilde{W}_x} \{f(y) + w(x-y)\} = (f \oplus w)(x)$$

L'avantage de la formule du *SMorph* est qu'elle est définie pour toute image f et tout noyau w à valeurs dans \mathbb{R} tout entier. Ainsi, il n'y a plus besoin de vérifier la positivité des valeurs de f et w et de les mettre en échelle positive comme pour *LMorph*.

On écrira ainsi la fonction de transformation $T : T(\cdot, w, \alpha) = \mathbf{SMorph}(\cdot, w, \alpha)$.

2.4 Efficacité des réseaux morphologiques

2.4.1 Méthodologie et description des expériences

A partir des trois différentes couches morphologiques, telles que définies dans la partie précédente, sont construits trois types de réseaux de neurones morphologiques correspondants : les réseaux *pConvNet* (composés de couches *pConv*), les réseaux *LMorphNet* (couches *LMorph*) et les réseaux *SMorphNet* (couches *SMorph*).

Sur la base de l'architecture générale des réseaux morphologiques, illustrée fig. 5, les structures considérées et expérimentées dans la littérature [9, 7] sont celles à 1 couche morphologique pour l'étude de l'érosion et de la dilatation, et celles à 2 couches morphologiques pour l'étude de l'ouverture et de la fermeture. Dans tous les cas, une standardisation en entrée du réseau et une couche de convolution $1 \times 1 \times 1$ en sortie sont imposées, afin d'améliorer la convergence des réseaux lors de leur entraînement (démontré dans [7]). Dans leur article, Hermary et al. exposent les résultats de leurs recherches sur la comparaison de l'efficacité de ces trois types de réseaux (récapitulés dans la sous-partie suivante) sur la base d'expériences telles que décrites ci-après.

Soyons d'abord précis sur les termes employés. On parle d'« expérience » ou de « scénario » pour désigner le paramétrage d'un réseau de neurones dans le cadre de son entraînement, pour lequel les hyperparamètres (taille des filtres, nombre de couches, éléments cibles, etc.) ont un état prédéfini particulier [14]. Si deux réseaux de neurones ont exactement le même paramétrage d'entraînement, alors il s'agit d'une seule et même expérience. À l'inverse, si deux réseaux ont au moins un hyperparamètre dont l'état diffère de l'un à l'autre, alors il s'agit de deux expériences différentes.

Les différents objets composant ces « paramètres de réseaux pré-entraînement » sont aussi bien les paramètres caractérisant l'architecture du réseau en question (type de réseau morphologique entre *pConvNet*, *LMorphNet* et *SMorphNet*; nombre de couches morphologiques dans le réseau; présence ou non de couches de rééchelonnage; type d'état initial des poids des couches; formule de la fonction de perte *loss*; ... | voir la section "réseau" en bas de la fig. 5) que les paramètres cibles de l'entraînement visée (opération cible entre érosion, dilatation, ouverture, fermeture ou dessallage; fonction(s) structurante(s) cible(s); ... | voir la section "cible" en haut de la fig. 5).

Pour toutes les expériences effectuées dans les travaux de Hermary et al., lorsqu'un entraînement est lancé, l'état des poids du noyau des différentes couches est aléatoire, défini par une loi normale d'espérance 0 et d'écart-type 1. Les poids de contrôle (p pour les *pConv* et les *LMorph*, et α pour les *SMorph*) sont, quant à eux, initialisés à 0. Enfin, dans la couche de convolution en sortie des réseaux, la valeur s de l'échelle (scale) est initialisée à 1, et la valeur b du biais (bias) à 0. De plus, la base de données d'images utilisée pour l'ensemble des entraînements de cette étude est MNIST, composée de 70000 images digitales en niveaux de gris entre 0 et 1 [12].

Dans cette étude sont également considérées 6 formes différentes de fonctions structurantes cibles sur un support de taille 7×7 : une croix plate de largeur 3, une seconde croix plate de largeur 7, un disque gradué de diamètre 5, un second disque gradué de diamètre 7, un losange gradué de diamètre 7, et enfin un cercle complexe gradué de diamètre 7 (voir fig. 5 partie cible, ou figures suivantes ligne supérieure).

Les différentes expériences réalisées sont divisées en 4 grands groupes, formés selon l'opération cible :

- 1er groupe : ensemble des expériences avec l'opération d'*érosion* (1 couche morphologique dans les réseaux)
- 2ème groupe : ensemble des expériences avec l'opération de *dilatation* (1 couche morphologique dans les réseaux)
- 3ème groupe : ensemble des expériences avec l'opération d'*ouverture* (2 couches morphologiques dans les réseaux)
- 4ème groupe : ensemble des expériences avec l'opération de *fermeture* (2 couches morphologiques dans les réseaux)

Pour chaque expérience, afin de pouvoir comparer la précision des trois types de réseaux de manière rigoureuse, des métriques d'intérêt sont définies et évaluées sur chacune des expériences faites. Dans l'étude de Hermary et al. [7], chaque expérience est réalisée au moins cinq fois, pour pouvoir établir de manière suffisamment pertinente une moyenne et un écart-type sur les métriques évaluées.

Les métriques choisies pour évaluer la précision des réseaux sont les suivantes :

- la perte globale du réseau (dite *loss*), qui est l'erreur moyenne sur la prédiction du réseau dans son état final, après convergence. Il s'agit ici de la moyenne des erreurs au carré (MSE) entre les images prédites, en sortie du réseau, et les images cibles, qui ont subi l'opération cible par la fonction structurante cible.
- l'erreur sur le noyau w des couches (dite *RMSE*), qui est la MSE entre le noyau w de la couche en question et la fonction structurante cible w_{cible} , accompagnée de l'aspect visuel de chacun. On espère en effet que w prenne la forme de w_{cible} qui a été utilisée pour créer par opération les images cibles (voir "cible" fig. 5).
- la valeur du paramètre de contrôle $|p|$ (ou α) des couches morphologiques du réseau après convergence. Plus p est éloigné de 0, plus l'effet de la couche se rapproche d'une réelle érosion ou réelle dilatation (et non une pseudo-opération si p est proche de 0), et donc mieux le réseau aura convergé.
- le nombre d'époques nécessaires à la convergence du réseau (moyenne et écart-type sur l'ensemble des runs faites pour chaque expérience), représentant la vitesse de convergence du réseau pour atteindre à l'état final.

Les résultats de comparaison de l'efficacité entre les trois différents réseaux, avec ces quatre métriques-ci, sont étudiés dans la sous-partie suivante.

2.4.2 Comparaison de la précision des réseaux

Il est pertinent de chercher à comprendre quelle couche morphologique, parmi les trois introduites dans la partie précédente, est la plus efficace pour à la fois apprendre automatiquement et effectuer de manière précise une des deux opérations morphologiques fondamentales (érosion et dilatation), dans un réseau de neurones. Pour ce faire, l'étude se fait à partir de l'analyse de la convergence et de l'état final des réseaux de neurones morphologiques construits suivant l'architecture fig. 5. La comparaison de l'efficacité des trois différents types de réseaux morphologiques (p ConvNet, \mathcal{L} MorphNet et \mathcal{S} MorphNet), se fait suivant les 4 groupes décrits précédemment selon le type d'opération cible (érosion, dilatation, ouverture, fermeture).

En reprenant les résultats de recherche de Hermay et al. [7], on obtient les résultats de convergence et de métriques ainsi que les figures présentés dans cette sous-partie.

Dans le premier cas, i.e. pour une opération d'érosion avec un réseau à 1 couche morphologique, on obtient les résultats de convergence et de caractéristiques suivants, sur 5 runs par expérience, fig 6 :

							
		<i>w</i>	<i>w</i>	<i>w</i>	<i>w</i>	<i>w</i>	<i>w</i>
<i>PConv</i>	<i>p</i>	-20.86 ± 0.45	-20.35 ± 0.20	-10.52 ± 0.004	-15.64 ± 0.06	-13.21 ± 0.02	-16.45 ± 0.07
	RMSE	$0.82 \pm 7 \times 10^{-4}$	$1.55 \pm 1 \times 10^{-3}$	$2.82 \pm 1 \times 10^{-4}$	$3.77 \pm 6 \times 10^{-4}$	$3.64 \pm 1 \times 10^{-4}$	$3.19 \pm 6 \times 10^{-3}$
	LOSS	$2.4 \times 10^{-5} \pm 1 \times 10^{-6}$	$6.2 \times 10^{-6} \pm 1 \times 10^{-7}$	$1.3 \times 10^{-4} \pm 1 \times 10^{-8}$	$2.6 \times 10^{-5} \pm 5 \times 10^{-8}$	$5.2 \times 10^{-5} \pm 2 \times 10^{-8}$	$1.2 \times 10^{-5} \pm 4 \times 10^{-8}$
	EPOCHS	209 ± 45	169 ± 22	49 ± 3	92 ± 7	66 ± 8	130 ± 18
\mathcal{L} Morph	<i>p</i>	-77.09 ± 0.15	-59.58 ± 0.20	-67.78 ± 0.30	-59.46 ± 0.18	-66.25 ± 0.17	-79.36 ± 0.93
	RMSE	$0.44 \pm 3 \times 10^{-4}$	$0.60 \pm 1 \times 10^{-2}$	$0.37 \pm 4 \times 10^{-2}$	$0.30 \pm 7 \times 10^{-3}$	$0.38 \pm 4 \times 10^{-2}$	$0.04 \pm 1 \times 10^{-3}$
	LOSS	$1.1 \times 10^{-6} \pm 5 \times 10^{-9}$	$3.6 \times 10^{-7} \pm 4 \times 10^{-9}$	$1.3 \times 10^{-6} \pm 6 \times 10^{-8}$	$3.7 \times 10^{-7} \pm 4 \times 10^{-9}$	$5.8 \times 10^{-7} \pm 3 \times 10^{-9}$	$1.1 \times 10^{-6} \pm 5 \times 10^{-8}$
	EPOCHS	158 ± 4	164 ± 20	146 ± 16	256 ± 3	231 ± 9	876 ± 139
\mathcal{S} Morph	<i>w</i>						
	α	-33.95 ± 0.03	-28.52 ± 0.013	-30.61 ± 0.016	-28.10 ± 0.006	-34.38 ± 0.003	-23.79 ± 0.011
	RMSE	$0.78 \pm 2 \times 10^{-2}$	$0.18 \pm 3 \times 10^{-2}$	$0.10 \pm 2 \times 10^{-3}$	$0.04 \pm 2 \times 10^{-2}$	$0.13 \pm 4 \times 10^{-2}$	$0.08 \pm 2 \times 10^{-2}$
	LOSS	$2.1 \times 10^{-5} \pm 5 \times 10^{-9}$	$4.0 \times 10^{-7} \pm 8 \times 10^{-9}$	$4.6 \times 10^{-7} \pm 8 \times 10^{-9}$	$4.2 \times 10^{-7} \pm 6 \times 10^{-9}$	$4.3 \times 10^{-7} \pm 3 \times 10^{-9}$	$9.5 \times 10^{-7} \pm 7 \times 10^{-9}$
	EPOCHS	40 ± 4	45 ± 4	38 ± 8	38 ± 7	42 ± 5	55 ± 9

Figure 6 – Poids des couches appris et valeur des moyennes et écarts-types des quatre métriques (p/α , RMSE, loss, et nombre d'époques) sur cinq runs, pour les trois types de réseaux et les six fonctions structurantes cibles, et pour l'opération cible d'érosion.

Dans le deuxième cas, i.e. pour une opération de dilatation avec 1 couche morphologique, on obtient les résultats suivants, sur 5 runs par expérience, fig 7 :

						
	<i>w</i>					
<i>pConv</i>	<i>p</i>	19.87 ± 0.23	22.58 ± 0.44	8.27 ± 0.002	9.43 ± 0.002	9.39 ± 0.001
	RMSE	$0.41 \pm 2 \times 10^{-3}$	$1.42 \pm 6 \times 10^{-4}$	$2.22 \pm 3 \times 10^{-4}$	$3.05 \pm 3 \times 10^{-4}$	$2.79 \pm 3 \times 10^{-4}$
	LOSS	$4.8 \times 10^{-5} \pm 1 \times 10^{-6}$	$9.0 \times 10^{-5} \pm 4 \times 10^{-6}$	$5.1 \times 10^{-4} \pm 5 \times 10^{-8}$	$6.3 \times 10^{-4} \pm 6 \times 10^{-9}$	$6.8 \times 10^{-4} \pm 5 \times 10^{-9}$
	EPOCHS	194 ± 17	222 ± 38	46 ± 8	41 ± 3	38 ± 4
						70 ± 7
	<i>w</i>					
<i>lMorph</i>	<i>p</i>	94.92 ± 0.18	95.89 ± 0.29	94.16 ± 0.17	94.25 ± 0.96	94.67 ± 0.31
	RMSE	$0.02 \pm 9 \times 10^{-5}$	$0.003 \pm 8 \times 10^{-6}$	$0.01 \pm 8 \times 10^{-5}$	$0.05 \pm 1 \times 10^{-4}$	$0.01 \pm 2 \times 10^{-4}$
	LOSS	$8.4 \times 10^{-6} \pm 4 \times 10^{-8}$	$1.1 \times 10^{-5} \pm 9 \times 10^{-8}$	$7.6 \times 10^{-6} \pm 4 \times 10^{-8}$	$1.2 \times 10^{-5} \pm 3 \times 10^{-7}$	$1.1 \times 10^{-5} \pm 1 \times 10^{-7}$
	EPOCHS	119 ± 12	159 ± 22	206 ± 11	193 ± 27	206 ± 9
						295 ± 35
	<i>w</i>					
<i>sMorph</i>	α	41.97 ± 0.016	52.01 ± 0.027	40.75 ± 0.014	50.06 ± 0.044	49.47 ± 0.04
	RMSE	$0.1 \pm 3 \times 10^{-3}$	$0.14 \pm 2 \times 10^{-3}$	$0.09 \pm 4 \times 10^{-4}$	$0.08 \pm 1 \times 10^{-3}$	$0.09 \pm 2 \times 10^{-3}$
	LOSS	$8.5 \times 10^{-7} \pm 3 \times 10^{-8}$	$1.5 \times 10^{-6} \pm 2 \times 10^{-8}$	$1.2 \times 10^{-6} \pm 1 \times 10^{-8}$	$1.8 \times 10^{-6} \pm 4 \times 10^{-8}$	$1.5 \times 10^{-6} \pm 9 \times 10^{-9}$
	EPOCHS	39 ± 8	39 ± 9	44 ± 9	45 ± 4	44 ± 4
						49 ± 7

Figure 7 – Poids des couches appris et valeur des moyennes et écarts-types des quatre métriques (p/α , RMSE, loss, et nombre d'époques) sur cinq runs, pour les trois types de réseaux et les six fonctions structurantes cibles, et pour l'opération cible de **dilatation**.

A partir des deux figures (fig. 6 pour l'ensemble des expériences avec l'érosion, et fig. 7 pour celui avec la dilatation), nous pouvons faire les constatations suivantes :

D'abord, en examinant la valeur des p/α de l'état final des réseaux sur la fig. 6 (resp. fig. 7), on constate que les trois types de couches morphologiques parviennent à trouver l'opération morphologique cible, i.e. une érosion (resp. une dilatation), car ces paramètres sont tous négatifs (resp. positifs) et leur amplitude est suffisamment élevée pour considérer l'effet des couches comme véritable opération morphologique.

Ensuite, à la fois pour l'érosion et la dilatation, on peut constater que la valeur des RMSE pour l'ensemble des fonctions structurantes cibles est bien plus élevée pour la couche *pConv* que pour les couches *lMorph* et *sMorph*, à l'exception de *cross3*. La forme du noyau des couches *pConv* est donc moins proche de la cible que les deux autres couches, ce qui peut également être constaté visuellement. La couche *pConv* semble souffrir d'un certain effet de « creux » dans la forme de son noyau.

La même chose peut être constatée avec la *loss* des réseaux dans leur état final : elle est bien plus élevée pour la couche *pConv*, de l'ordre de 10^1 à 10^2 fois plus, que pour les couches *LMorph* et *SMorph*, qui ont toutes deux le même ordre de grandeur. Ces deux dernières couches sont donc, à la fois pour l'érosion et la dilatation, bien meilleures que la *pConv*, car elles font de bien meilleures prédictions en moyenne d'après les valeurs de la *loss*. Dans le cadre de la dilatation, fig. 7, la *SMorph* semble cependant meilleure que la *LMorph*, d'un ordre de grandeur sur la *loss* en moyenne.

Le nombre d'époques nécessaire pour atteindre la convergence des réseaux est, pour érosion et dilatation, bien plus faible et régulier sur l'ensemble des expériences pour *SMorph* que pour *pConv* ou *LMorph*. La couche *SMorph* converge donc plus rapidement et avec plus de régularité que les deux autres couches. On remarque également que l'écart-type des métriques sont très faibles, montrant que les réseaux ont convergé sur les 5 cycles d'entraînement (runs) vers la même solution pour les trois couches.

Dans le troisième cas, pour une opération d'ouverture avec 2 couches fig. 8 :

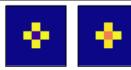
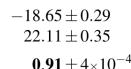
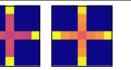
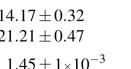
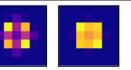
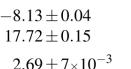
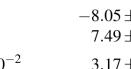
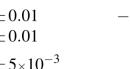
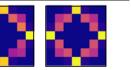
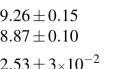
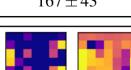
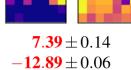
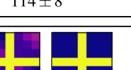
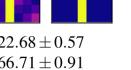
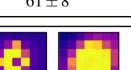
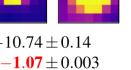
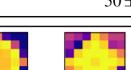
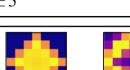
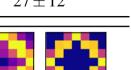
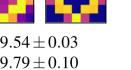
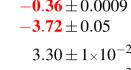
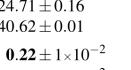
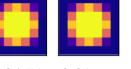
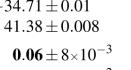
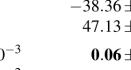
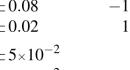
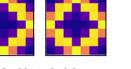
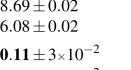
													
<i>PConv</i>	w_1/w_2												
	p_1	-18.65 ± 0.29	-14.17 ± 0.32	-8.13 ± 0.04	-8.59 ± 0.06	-8.05 ± 0.01	-9.26 ± 0.15						
	p_2	22.11 ± 0.35	21.21 ± 0.47	17.72 ± 0.15	7.03 ± 0.10	7.49 ± 0.01	8.87 ± 0.10						
	RMSE ₁	$0.91 \pm 4 \times 10^{-4}$	$1.45 \pm 1 \times 10^{-3}$	$2.69 \pm 7 \times 10^{-3}$	$3.13 \pm 3 \times 10^{-2}$	$3.17 \pm 5 \times 10^{-3}$	$2.53 \pm 3 \times 10^{-2}$						
	RMSE ₂	$0.32 \pm 1 \times 10^{-3}$	$0.87 \pm 2 \times 10^{-3}$	$2.49 \pm 2 \times 10^{-3}$	$2.51 \pm 1 \times 10^{-2}$	$2.43 \pm 4 \times 10^{-3}$	$2.49 \pm 1 \times 10^{-2}$						
	LOSS	$8.9 \times 10^{-5} \pm 3 \times 10^{-6}$	$7.5 \times 10^{-5} \pm 1 \times 10^{-6}$	$4.9 \times 10^{-4} \pm 2 \times 10^{-6}$	$3.9 \times 10^{-4} \pm 2 \times 10^{-5}$	$4.4 \times 10^{-4} \pm 4 \times 10^{-6}$	$2.2 \times 10^{-4} \pm 6 \times 10^{-6}$						
<i>LMorph</i>	EPOCHS	167 ± 43	114 ± 8	61 ± 8	47 ± 13	50 ± 5	27 ± 12						
	w_1/w_2												
	p_1	7.39 ± 0.14	-22.68 ± 0.57	-10.74 ± 0.14	-10.88 ± 0.11	-12.82 ± 0.08	-9.54 ± 0.03						
	p_2	-12.89 ± 0.06	66.71 ± 0.91	-1.07 ± 0.003	8.18 ± 0.08	12.55 ± 0.07	9.79 ± 0.10						
	RMSE ₁	$3.17 \pm 2 \times 10^{-3}$	$1.02 \pm 6 \times 10^{-2}$	$1.99 \pm 7 \times 10^{-4}$	$0.17 \pm 8 \times 10^{-3}$	$1.92 \pm 5 \times 10^{-2}$	$1.15 \pm 8 \times 10^{-2}$						
	RMSE ₂	$4.72 \pm 8 \times 10^{-3}$	$0.02 \pm 3 \times 10^{-4}$	$1.61 \pm 8 \times 10^{-3}$	$1.80 \pm 4 \times 10^{-2}$	$0.70 \pm 4 \times 10^{-3}$	$0.85 \pm 8 \times 10^{-3}$						
	LOSS	$8.7 \times 10^{-3} \pm 9 \times 10^{-6}$	$4.3 \times 10^{-5} \pm 2 \times 10^{-6}$	$2.0 \times 10^{-3} \pm 1 \times 10^{-5}$	$3.8 \times 10^{-4} \pm 1 \times 10^{-5}$	$3.2 \times 10^{-4} \pm 4 \times 10^{-6}$	$2.3 \times 10^{-4} \pm 2 \times 10^{-6}$						
<i>SMorph</i>	EPOCHS	51 ± 6	111 ± 8	165 ± 42	21 ± 5	46 ± 6	44 ± 4						
	w_1/w_2												
	α_1	-0.36 ± 0.0009	-24.71 ± 0.16	-34.71 ± 0.01	-29.15 ± 0.07	-38.36 ± 0.08	-18.69 ± 0.02						
	α_2	-3.72 ± 0.05	40.62 ± 0.01	41.38 ± 0.008	45.41 ± 0.02	47.13 ± 0.02	16.08 ± 0.02						
	RMSE ₁	$3.30 \pm 1 \times 10^{-2}$	$0.22 \pm 1 \times 10^{-2}$	$0.06 \pm 8 \times 10^{-3}$	$0.04 \pm 8 \times 10^{-3}$	$0.06 \pm 5 \times 10^{-2}$	$0.11 \pm 3 \times 10^{-2}$						
	RMSE ₂	$1.63 \pm 5 \times 10^{-3}$	$0.06 \pm 3 \times 10^{-2}$	$0.03 \pm 1 \times 10^{-2}$	$0.05 \pm 1 \times 10^{-2}$	$0.06 \pm 2 \times 10^{-2}$	$0.11 \pm 3 \times 10^{-2}$						
	LOSS	$4.8 \times 10^{-3} \pm 4 \times 10^{-6}$	$4.6 \times 10^{-7} \pm 9 \times 10^{-8}$	$5.5 \times 10^{-7} \pm 4 \times 10^{-8}$	$6.2 \times 10^{-7} \pm 8 \times 10^{-8}$	$5.5 \times 10^{-7} \pm 4 \times 10^{-8}$	$1.8 \times 10^{-6} \pm 4 \times 10^{-8}$						
	EPOCHS	22 ± 3	41 ± 5	35 ± 3	42 ± 4	42 ± 6	42 ± 4						

Figure 8 – Poids des couches appris et valeur des moyennes et écarts-types des quatre métriques (p/α , RMSE, loss, et nombre d'époques) sur cinq runs, pour les trois types de réseaux et les six fonctions structurantes cibles, et pour l'opération cible d'ouverture**.**

Dans le dernier cas, pour une opération de fermeture avec 2 couches fig. 9 :

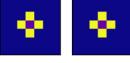
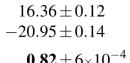
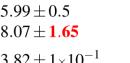
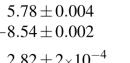
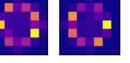
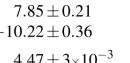
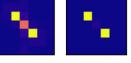
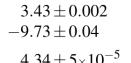
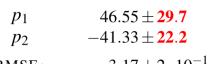
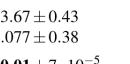
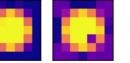
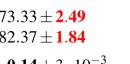
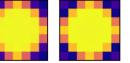
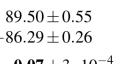
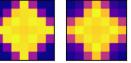
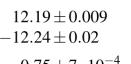
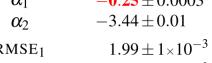
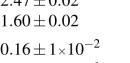
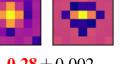
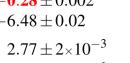
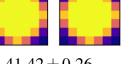
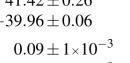
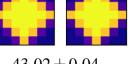
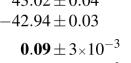
						
<i>PConv</i>	w_1/w_2	 	 	 	 	 
	p_1	16.36 ± 0.12	5.99 ± 0.5	5.78 ± 0.004	7.85 ± 0.21	3.43 ± 0.002
	p_2	-20.95 ± 0.14	-8.07 ± 1.65	-8.54 ± 0.002	-10.22 ± 0.36	-9.73 ± 0.04
	RMSE_1	$0.82 \pm 6 \times 10^{-4}$	$3.82 \pm 1 \times 10^{-1}$	$2.82 \pm 2 \times 10^{-4}$	$4.47 \pm 3 \times 10^{-3}$	$4.34 \pm 5 \times 10^{-5}$
	RMSE_2	$0.81 \pm 1 \times 10^{-3}$	$3.87 \pm 9 \times 10^{-2}$	$3.07 \pm 3 \times 10^{-4}$	$4.63 \pm 2 \times 10^{-1}$	$4.51 \pm 6 \times 10^{-5}$
	LOSS	$8.9 \times 10^{-5} \pm 1 \times 10^{-6}$	$6.5 \times 10^{-3} \pm 7 \times 10^{-4}$	$1.7 \times 10^{-3} \pm 6 \times 10^{-8}$	$1.0 \times 10^{-3} \pm 6 \times 10^{-5}$	$5.4 \times 10^{-3} \pm 5 \times 10^{-7}$
$\mathcal{L}\text{Morph}$	w_1/w_2	 	 	 	 	 
	p_1	46.55 ± 29.7	93.67 ± 0.43	73.33 ± 2.49	89.50 ± 0.55	12.19 ± 0.009
	p_2	-41.33 ± 22.2	-92.077 ± 0.38	-82.37 ± 1.84	-86.29 ± 0.26	-12.24 ± 0.02
	RMSE_1	$3.17 \pm 2 \times 10^{-1}$	$0.01 \pm 7 \times 10^{-5}$	$0.14 \pm 3 \times 10^{-3}$	$0.07 \pm 3 \times 10^{-4}$	$0.75 \pm 7 \times 10^{-4}$
	RMSE_2	$3.29 \pm 9 \times 10^{-1}$	$0.63 \pm 2 \times 10^{-3}$	$1.16 \pm 2 \times 10^{-2}$	$0.40 \pm 1 \times 10^{-3}$	$0.95 \pm 7 \times 10^{-4}$
	LOSS	$3.7 \times 10^{-3} \pm 6 \times 10^{-4}$	$1.3 \times 10^{-5} \pm 2 \times 10^{-7}$	$2.0 \times 10^{-4} \pm 1 \times 10^{-6}$	$1.7 \times 10^{-5} \pm 2 \times 10^{-7}$	$6.2 \times 10^{-4} \pm 1 \times 10^{-7}$
$\mathcal{S}\text{Morph}$	w_1/w_2	 	 	 	 	 
	α_1	-0.25 ± 0.0003	42.47 ± 0.02	-0.28 ± 0.002	41.42 ± 0.26	43.02 ± 0.04
	α_2	-3.44 ± 0.01	-41.60 ± 0.02	-6.48 ± 0.02	-39.96 ± 0.06	-42.94 ± 0.03
	RMSE_1	$1.99 \pm 1 \times 10^{-3}$	$0.16 \pm 1 \times 10^{-2}$	$2.77 \pm 2 \times 10^{-3}$	$0.09 \pm 1 \times 10^{-3}$	$0.09 \pm 3 \times 10^{-3}$
	RMSE_2	$3.77 \pm 3 \times 10^{-3}$	$0.09 \pm 8 \times 10^{-3}$	$3.81 \pm 9 \times 10^{-3}$	$0.01 \pm 3 \times 10^{-3}$	$0.10 \pm 5 \times 10^{-3}$
	LOSS	$1.9 \times 10^{-3} \pm 6 \times 10^{-7}$	$8.0 \times 10^{-7} \pm 3 \times 10^{-8}$	$4.5 \times 10^{-3} \pm 4 \times 10^{-6}$	$1.1 \times 10^{-6} \pm 3 \times 10^{-8}$	$8.7 \times 10^{-7} \pm 3 \times 10^{-8}$
	EPOCHS	18 ± 2	39 ± 6	15 ± 1	44 ± 4	39 ± 4

Figure 9 – Poids des couches appris et valeur des moyennes et écarts-types des quatre métriques (p/α , RMSE, loss, et nombre d'époques) sur cinq runs, pour les trois types de réseaux et les six fonctions structurantes cibles, et pour l'opération cible de fermeture.

A partir des deux figures 8 et 9 pour l'ouverture et la fermeture, on remarque la présence évidente d'échecs de convergence. Au travers des valeurs de RMSE et des noyaux des couches, on remarque que le réseau *pConvNet* réussit à plutôt bien deviner la forme de la fonction structurante cible sur l'ensemble des expériences pour l'ouverture et la fermeture, mais pas pour la fermeture où les RMSE et la loss sont tous très élevés. Dans les deux cas, *pConvNet* souffre toujours de l'effet « creux ».

Hors-mis la cross3, le réseau *L*MorphNet semble mieux converger que *pConvNet* sur l'ensemble des expériences, avec des RMSE et des loss bien plus faibles, à la fois pour l'ouverture et la fermeture. Cependant, la forme du noyau des couches n'est pas parfaite, et varie beaucoup entre la première et la seconde couche du réseau, ce qui n'est pas le cas pour le réseau *SMorphNet* qui, hors-mis avec le disk2 (fermeture), est meilleur encore et ne présente pas cette disymétrie de noyau entre les deux couches du réseau. *SMorphNet* surpassé presque toujours *pConvNet* et *L*MorphNet en termes de RMSE, de loss et de nombre d'époques, pour à la fois l'ouverture et la fermeture.

2.4.3 Problèmes de convergence des réseaux

L'analyse des résultats des différentes expériences précédentes montre une bien meilleure convergence et une meilleure précision des réseaux *SMorphNet* par rapport aux réseaux *LMorphNet* et *pConvNet*. Ces meilleures performances de la part de *SMorphNet* se traduit à la fois par une plus petite valeur de perte globale *loss* de l'état final des réseaux (et donc de meilleures prédictions), par une plus nette ressemblance des noyaux du réseau avec la fonction structurante cible (avec une RMSE bien plus faible), et par un nombre moyen d'époques bien plus bas pour atteindre la convergence (et donc une plus grande vitesse d'apprentissage). Et ce en moyenne, à la fois pour les expériences d'érosion, de dilatation, d'ouverture et de fermeture. Les réseaux *SMorphNet* sont donc plus performants et précis que les deux autres existant.

Cependant, l'analyse de ces résultats montre également que les différents réseaux morphologiques souffrent de problèmes de convergence sur différents aspects. Pour les expériences avec des réseaux à une seule couche morphologique (érosion fig. 6 ou dilatation fig. 7), aucun problème majeur hors-mis ceux évoqués précédemment n'est à déplorer. Par contre, pour les expériences avec des réseaux à deux couches morphologiques (ouverture fig. 8 ou fermeture fig. 9), les résultats obtenus soulèvent plusieurs problèmes en termes de comportement de convergence pour les couches *LMorph* et *SMorph*, en particulier pour les fonctions structurantes cibles *cross3* et *disk2* (fig. 9), où les couches n'arrivent pas à trouver la bonne forme de noyau, voire à trouver la bonne opération ciblée (par rapport au signe de p/α), et entraînent donc une *loss* sur le réseau plus élevée que les autres expériences et autres fonctions structurantes cibles.

Cela a conduit Hermay et al. [7] à effectuer une analyse plus approfondie de la convergence progressive des poids (noyau et paramètre de contrôle) de ces couches au fil des époques pendant la phase d'entraînement. Pour les réseaux à deux couches morphologiques, il en résulte ainsi les deux principales constatations suivantes :

- D'abord, la manière dont les poids des noyaux se mettent à jour tout au long du processus d'apprentissage présente plusieurs similitudes : la convergence vers la forme correcte est plus rapide pour la deuxième couche que pour la première (car ses poids sont mis à jour avant ceux de la première couche lors de la rétropropagation), de même que la convergence du paramètre p/α vers une plage de valeurs garantissant que l'opération appliquée par la couche est une bonne approximation d'une dilatation ou d'une érosion (et non que pseudo).
- Ensuite, le schéma de mise à jour des poids semble commencer sur les bords du noyau, pour ensuite se propager vers le centre. Cette dernière observation pourrait être la raison de l'échec de la convergence pour *cross3* et *disk2*, car l'étendue de ces fonctions structurantes est plus petite que leur support spatial, ne fournissant ainsi aucun point d'ancre sur les bords pour que les poids se propagent vers le centre du noyau. Mais cela ne reste qu'une hypothèse.

Toujours sur les cas d'échec avec les réseaux à deux couches morphologiques (fig. 8 et fig. 9), si l'on se penche sur les valeurs des paramètres de contrôle p/α , on remarque que, dans les cas d'échec, les couches des réseaux n'ont pas convergé vers la bonne opération (entre érosion et dilatation), avec des p/α de signe opposé à ce qu'ils devraient prendre, et de faible amplitude également, comme c'est le cas par exemple avec cross3 (ouverture et fermeture) pour les réseaux \mathcal{LM} orph et \mathcal{SM} orph, ou encore disk2 (fermeture) pour \mathcal{SM} orph. Le signe vers lequel tendent les p/α des deux couches d'un réseau ainsi que leur amplitude semblent en forte corrélation avec les échecs de convergence. La condition $|\alpha| > 10$ est en pratique suffisante pour reproduire une dilatation ou une érosion avec \mathcal{SM} orph, ce qui n'est pas le cas avec les échecs de convergence, où l'on a donc que des pseudo-opérations.

En conclusion, l'étude de Hermary et al. [7] montre que l'*étendue* d'une fonction structurante (i.e. l'espace occupé par les éléments de son support distingués du fond) par rapport à la taille de son support spatial semble être liée à la capacité des couches morphologiques à apprendre correctement cette fonction cible pour l'ouverture et la fermeture, mais cela n'explique pas à lui seul l'ensemble des cas d'échec présentés.

La convergence des réseaux vers des minima locaux à forte attraction est également une explication plausible, car l'identification d'une fonction structurante basée uniquement sur des images originales et des images transformées est une tâche sous-déterminée, plusieurs solutions pouvant exister. Par exemple, deux fonctions structurantes symétriquement décalées se compensent mutuellement et sont aussi valides en tant que solution que deux fonctions centrées (résultats que l'on cherche à obtenir).

Intégrer des informations *a priori* (symétrie, étendue spatiale) sur la fonction structurante recherchée dans la fonction de perte d'entraînement optimisée pourrait être une solution potentielle, mais ces informations *a priori* pourraient être difficiles à obtenir dans des cas d'utilisation pratiques, lorsque l'on ne connaît pas la forme cherchée.

Cette étude de Hermary et al. a permis de mieux comprendre l'impact de la rétro-propagation du gradient, lors de l'entraînement des réseaux, dans la tendance de la seconde couche à mieux converger vers la solution attendue, ses poids se mettant à jour en premier par rapport à la première couche. Elle montre également que la valeur, en particulier le signe, du paramètre de contrôle p/α des couches morphologiques semble fortement corrélée à la bonne ou mauvaise convergence des réseaux.

Pour la suite, la couche \mathcal{SM} orph ayant démontré de bien meilleurs résultats de convergence et de précision par rapport aux deux autres couches \mathcal{LM} orph et $p\text{Conv}$, nous travaillerons uniquement sur cette couche-ci. L'objectif sera alors d'étudier plus en profondeur le comportement des réseaux \mathcal{SM} orphNet, et de trouver des solutions pour améliorer la convergence de ces réseaux afin d'éviter les cas d'échecs tels que ceux présentés dans cette partie. Il s'agira d'un travail principalement exploratoire.

3 Amélioration des réseaux

3.1 Réseau $\mathcal{SM}\text{orphNet-Tanh}$

3.1.1 Contexte et motivations

La littérature montre que les réseaux $\mathcal{SM}\text{orphNet}$, composés de couches $\mathcal{SM}\text{orph}$ telles que décrites dans la partie état de l'art, sont les réseaux morphologiques existant les plus robustes dans leurs prédictions, à la fois en terme de perte globale *loss* (précision des images prédites par rapport aux images cibles), en terme de *RMSE* (similitude de la forme des noyaux w des couches morphologiques avec celle de la fonction structurante cible), en terme d'exactitude de l'opération prédictive ($\alpha << 0$ pour l'érosion, $\alpha >> 0$ pour la dilatation), et en terme de rapidité de convergence des réseaux lors de leur entraînement (avec le nombre d'époques nécessaire pour atteindre un minimum local dans la *loss*, et donc pour atteindre la convergence du réseau).

Cependant, bien que la couche $\mathcal{SM}\text{orph}$ pallie de nombreuses contraintes liées à la formulation des couches $p\text{Conv}$ et $\mathcal{LM}\text{orph}$ (telles que le fait que, pour ces deux couches, f et w doivent être à valeurs positives, et nécessitent donc un rééchelonnage en entrée), il reste une principale contrainte dans ses formulations asymptotiques.

En effet, lorsque $\alpha \rightarrow +\infty$, on a pour $x \in I$: $\lim_{\alpha \rightarrow +\infty} \mathcal{SM}\text{orph}(f, w, \alpha)(x) = (f \oplus w)(x)$, donc $\mathcal{SM}\text{orph}(f, w, \alpha)$ tend bien vers la véritable image dilatée de f par la fonction structurante w quand $\alpha \rightarrow +\infty$. Mais, à l'inverse, lorsque $\alpha \rightarrow -\infty$, on a pour $x \in I$: $\lim_{\alpha \rightarrow -\infty} \mathcal{SM}\text{orph}(f, w, \alpha)(x) = (f \ominus -\check{w})(x)$, donc $\mathcal{SM}\text{orph}(f, w, \alpha)$ tend vers l'image érodée de f par la fonction structurante $-\check{w}$ quand $\alpha \rightarrow -\infty$, avec $-\check{w}$ la *symétrique centrale* et la *négation* de w . Or, le but est d'avoir $\lim_{\alpha \rightarrow -\infty} \mathcal{SM}\text{orph}(f, w, \alpha)(x) = (f \ominus w)(x)$.

Une reformulation de $\mathcal{SM}\text{orph}$ doit donc être pensée, afin d'abord de pallier ce problème de signe des poids w . Cette nouvelle formule ne doit cependant pas générer d'irrégularités ou de trop hautes variations par rapport à la formule de $\mathcal{SM}\text{orph}$, pour éviter l'apparition de minima locaux en plus dans la fonction de perte *loss*, afin d'éviter le risque que les réseaux associés convergent moins bien (vers ces minima locaux-là). Il est donc requis que la nouvelle formule soit à minima davantage lisse (variations plus faibles) que $\mathcal{SM}\text{orph}$ en fonction des variables w et α .

Une nouvelle couche avec une nouvelle formulation satisfaisant les objectifs décrits a été ainsi pensée et initiée par Romain Hermay. Il s'agit de $\mathcal{SM}\text{orphTanh}$. La sous-partie suivante présente la formulation de cette nouvelle couche et ses propriétés.

Il serait possible, sur la même idée, de pallier le problème de symétrie \check{w} avec une transition continue et douce de w vers sa symétrie \check{w} quand α devient négatif au voisinage de 0. Mais il s'agit là d'un travail futur. Nous nous concentrerons d'abord sur la résolution de ce problème de signe devant w avec cette nouvelle formule $\mathcal{SM}\text{orphTanh}$.

3.1.2 Formalisation

L'idée dans la formule de cette nouvelle couche *SMorphTanh* dérivée de *SMorph* est simplement d'ajouter un facteur multiplicateur devant les valeurs du noyau w dans la formule de *SMorph* (9), facteur qui puisse être le reflet d'une transition lisse entre le signe + et - lors du passage de α d'un signe positive à un signe négatif, et inversement. Ce facteur est ici défini par la tangente hyperbolique de α : $\tanh(\alpha)$.

La formule de *SMorphNetTanh*, pour une image $f : I \subseteq \mathbb{Z}^2 \rightarrow \mathbb{R}$ et un noyau de couche $w : W \subseteq \mathbb{Z}^2 \rightarrow \mathbb{R}$, et avec $\alpha \in \mathbb{R}$, est ainsi définie pour tout $x \in I$ par :

$$\text{SMorphTanh}(f, w, \alpha)(x) = \frac{\sum_{y \in \check{W}_x} (f(y) + \tanh(\alpha)w(x-y)) e^{\alpha(f(y)+\tanh(\alpha)w(x-y))}}{\sum_{y \in \check{W}_x} e^{\alpha(f(y)+\tanh(\alpha)w(x-y))}} \quad (10)$$

Le facteur $\tanh(\alpha)$ permet bien, à la fois, d'ajuster le signe de w dans la formulation asymptotique de *SMorph*, et de lisser l'espace de la *loss* pour α au voisinage de 0, en réduisant l'impact des variations des poids w_i dans la formule de *SMorphTanh*, grâce à ce $\tanh(\alpha)$ devant w , qui est proche de 0 quand α est au voisinage de 0.

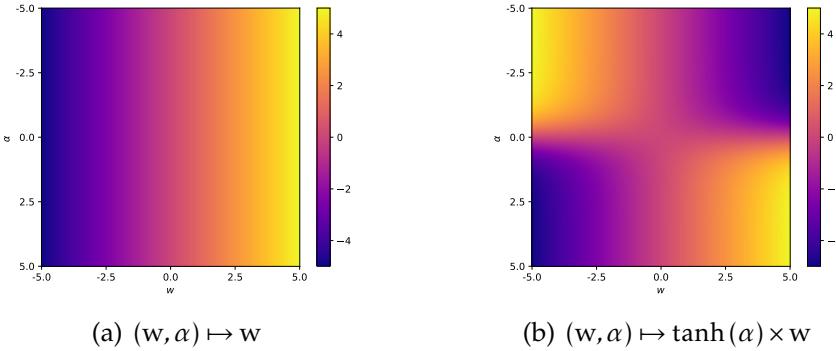


Figure 10 – Carte de la fonction $(w, \alpha) \mapsto w$ représentant *SMorphNet* (gauche 10(a)) et de la fonction $(w, \alpha) \mapsto \tanh(\alpha) \times w$ représentant *SMorphNetTanh* (droite 10(b)).

La figure 10 ci-dessus montre l'évolution des valeurs (en couleurs) de la fonction $(w, \alpha) \mapsto \tanh(\alpha) \times w$ en fonction de w (en abscisse) et α (en ordonnée). La carte 10(a) reflète le comportement de *SMorph* sur une image f nulle et avec un seul poids w du noyau w , et la carte 10(b) reflète celui de *SMorphTanh*. On remarque bien sur 10(a) le passage continu d'un signe positif à un signe négatif de α en ordonnée, ainsi que le lissage de la fonction au voisinage de $\alpha \approx 0$ selon le poids w en abscisse.

On obtient bien l'image $f \ominus \check{w}$ quand $\alpha \rightarrow -\infty$, et toujours $f \oplus w$ quand $\alpha \rightarrow +\infty$.

3.1.3 Comparaison avec SMorphNet

Pour savoir si les réseaux SMorphNetTanh sont des réseaux d'intérêt, au-delà de l'amélioration de la formulation asymptotique des couches SMorphTanh quand $\alpha \rightarrow -\infty$ par rapport à celle des couches SMorph telle que vue précédemment, il faut étudier les performances de ces nouveaux réseaux sur un ensemble d'expériences et comparer leurs résultats à ceux des réseaux SMorphNet simples.

Dans cet objectif, on reprend le même protocole d'expérimentation que dans l'état de l'art, où l'on étudie l'état des réseaux après leur entraînement avec certaines métriques. En particulier :

- on reprend les quatre mêmes métriques d'évaluation de la performance des réseaux que dans l'état de l'art (partie 2.4) ;
- on reste sur la banque d'images en niveaux de gris MNIST ;
- on fait cette fois-ci six runs par expérience, au lieu de cinq ;
- on considère cette fois huit fonctions structurantes, en prenant les six précédentes de l'état de l'art, et en y ajoutant une fonction structurante diagonale asymétrique *adiag* et une autre plate aléatoire *brand*.

On résume, comme dans l'état de l'art, les résultats de l'ensemble des expériences en quatre groupes : le premier pour les expériences avec l'érosion pour opération cible (réseaux à une couche morphologique), le deuxième avec la dilatation (1 couche), le troisième avec l'ouverture (2 couches), et le dernier avec la fermeture (2 couches).

Pour l'érosion sur des réseaux à une couche, on obtient les résultats suivants :

Opération : érosion									
Jeu : MNIST									
<i>w</i>									
<i>SMorph</i>	α	-37.10 ± 0.02	-31.21 ± 0.003	-33.67 ± 0.02	-30.87 ± 0.02	-37.34 ± 0.02	-25.92 ± 0.008	-22.93 ± 0.51	-31.22 ± 0.006
	RMSE	$0.11 \pm 1 \times 10^{-3}$	$0.03 \pm 5 \times 10^{-3}$	$0.02 \pm 3 \times 10^{-4}$	$0.007 \pm 3 \times 10^{-3}$	$0.02 \pm 7 \times 10^{-3}$	$0.01 \pm 2 \times 10^{-3}$	$0.06 \pm 4 \times 10^{-4}$	$0.07 \pm 6 \times 10^{-3}$
	LOSS	$2.4 \times 10^5 \pm 9 \times 10^{-7}$	$9.2 \times 10^7 \pm 5 \times 10^{-7}$	$1.8 \times 10^6 \pm 9 \times 10^{-7}$	$1.1 \times 10^{-6} \pm 1 \times 10^{-6}$	$8.8 \times 10^{-7} \pm 6 \times 10^{-7}$	$2.3 \times 10^6 \pm 9 \times 10^{-7}$	$1.9 \times 10^6 \pm 5 \times 10^{-7}$	$1.2 \times 10^4 \pm 7 \times 10^{-7}$
	EPOCHS	36 ± 7	37 ± 6	34 ± 5	32 ± 4	36 ± 6	53 ± 5	38 ± 9	41 ± 6
<i>w</i>									
<i>SMorphTanh</i>	α	-37.66 ± 0.03	-31.20 ± 0.006	-33.98 ± 0.03	-30.87 ± 0.03	-37.36 ± 0.008	-25.88 ± 0.04	-1.77 ± 0.004	-31.21 ± 0.009
	RMSE	$0.02 \pm 3 \times 10^{-4}$	$0.03 \pm 2 \times 10^{-3}$	$0.03 \pm 5 \times 10^{-4}$	$0.008 \pm 3 \times 10^{-3}$	$0.02 \pm 2 \times 10^{-3}$	$0.01 \pm 3 \times 10^{-3}$	$0.14 \pm 1 \times 10^{-3}$	$0.08 \pm 6 \times 10^{-3}$
	LOSS	$3.4 \times 10^6 \pm 2 \times 10^{-7}$	$9.6 \times 10^7 \pm 4 \times 10^{-7}$	$1.9 \times 10^6 \pm 5 \times 10^{-7}$	$1.3 \times 10^{-6} \pm 8 \times 10^{-7}$	$1.1 \times 10^{-6} \pm 9 \times 10^{-7}$	$1.8 \times 10^6 \pm 8 \times 10^{-7}$	$1.2 \times 10^4 \pm 1 \times 10^{-6}$	$1.4 \times 10^6 \pm 9 \times 10^{-7}$
	EPOCHS	31 ± 3	33 ± 3	37 ± 5	35 ± 5	42 ± 4	42 ± 6	32 ± 6	38 ± 4

Figure 11 – Comparaison des poids appris et des moyennes et écarts-types des métriques α , RMSE, loss et nombre d'époques, et ce sur six runs, entre SMorph et SMorphTanh à une couche, pour les huit fonctions structurantes cibles et l'opération d'érosion.

En rouge sont notées les moins bonnes performances moyennes du réseau par rapport à l'autre (grandes moyennes). En vert, les résultats instables (grands écarts-types).

Pour la dilatation sur des réseaux à une couche, on obtient les résultats suivants :

Opération : dilatation									
jeu : MNIST									
<i>SMorph</i>	α	44.48 \pm 1.01	54.72 \pm 1.42	43.96 \pm 0.92	52.87 \pm 1.40	52.43 \pm 0.84	36.77 \pm 0.40	53.53 \pm 2.96	52.56 \pm 1.19
	RMSE	0.01 \pm 5 \times 10 ⁻⁴	0.02 \pm 4 \times 10 ⁻⁴	0.01 \pm 4 \times 10 ⁻⁴	0.01 \pm 6 \times 10 ⁻⁴	0.01 \pm 3 \times 10 ⁻⁴	0.01 \pm 9 \times 10 ⁻³	0.03 \pm 2 \times 10 ⁻³	0.04 \pm 6 \times 10 ⁻³
	LOSS	1.8 \times 10 ⁻⁵ \pm 3 \times 10 ⁻⁶	2.9 \times 10 ⁻⁶ \pm 4 \times 10 ⁻⁶	2.7 \times 10 ⁻⁶ \pm 5 \times 10 ⁻⁶	3.5 \times 10 ⁻⁵ \pm 6 \times 10 ⁻⁶	3.5 \times 10 ⁻⁵ \pm 9 \times 10 ⁻⁶	3.9 \times 10 ⁻⁵ \pm 7 \times 10 ⁻⁶	1.8 \times 10 ⁻³ \pm 2 \times 10 ⁻⁵	1.9 \times 10 ⁻⁴ \pm 8 \times 10 ⁻⁶
	EPOCHS	30 \pm 8	33 \pm 7	28 \pm 4	33 \pm 7	29 \pm 3	29 \pm 6	29 \pm 3	31 \pm 4
<hr/>									
<i>SMorphTanh</i>	α	44.74 \pm 0.59	54.73 \pm 1.18	44.16 \pm 1.11	51.06 \pm 1.82	51.41 \pm 2.77	36.93 \pm 0.34	55.37 \pm 1.32	52.25 \pm 1.59
	RMSE	0.01 \pm 2 \times 10 ⁻⁴	0.02 \pm 4 \times 10 ⁻⁴	0.01 \pm 6 \times 10 ⁻⁴	0.01 \pm 7 \times 10 ⁻⁴	0.01 \pm 1 \times 10 ⁻³	0.01 \pm 6 \times 10 ⁻³	0.03 \pm 2 \times 10 ⁻³	0.04 \pm 3 \times 10 ⁻³
	LOSS	1.7 \times 10 ⁻⁵ \pm 4 \times 10 ⁻⁶	2.9 \times 10 ⁻⁶ \pm 1 \times 10 ⁻⁵	2.6 \times 10 ⁻⁵ \pm 2 \times 10 ⁻⁶	3.6 \times 10 ⁻⁵ \pm 4 \times 10 ⁻⁶	3.1 \times 10 ⁻⁵ \pm 7 \times 10 ⁻⁶	3.7 \times 10 ⁻⁵ \pm 8 \times 10 ⁻⁶	1.8 \times 10 ⁻³ \pm 2 \times 10 ⁻⁵	1.9 \times 10 ⁻⁴ \pm 9 \times 10 ⁻⁶
	EPOCHS	28 \pm 4	32 \pm 6	34 \pm 9	25 \pm 3	30 \pm 7	30 \pm 5	30 \pm 2	31 \pm 4

Figure 12 – Comparaison des poids appris et des moyennes et écarts-types des métriques α , RMSE, loss et nombre d'époques, et ce sur six runs, entre SMorph et SMorphTanh à une couche, pour les huit fonctions structurantes cibles et l'opération de **dilatation**.

Pour l'ouverture sur réseaux à deux couches, on obtient les résultats suivants :

Opération : ouverture									
jeu : MNIST									
<i>SMorph</i>	α_1	-0.47 \pm 0.002	-29.73 \pm 0.05	-0.93 \pm 0.003	-34.45 \pm 0.03	-45.35 \pm 0.02	-21.65 \pm 0.04	-3.00 \pm 0.12	-30.03 \pm 0.05
	α_2	-4.93 \pm 0.008	47.32 \pm 0.02	-0.54 \pm 0.004	52.66 \pm 0.02	54.50 \pm 0.03	17.02 \pm 0.02	-0.29 \pm 0.01	49.48 \pm 0.02
	RMSE ₁	0.48 \pm 1 \times 10 ⁻³	0.03 \pm 4 \times 10 ⁻³	0.43 \pm 4 \times 10 ⁻⁴	0.01 \pm 5 \times 10 ⁻³	0.02 \pm 3 \times 10 ⁻³	0.12 \pm 4 \times 10 ⁻³	0.43 \pm 2 \times 10 ⁻²	0.09 \pm 5 \times 10 ⁻³
	RMSE ₂	0.23 \pm 9 \times 10 ⁻⁴	0.01 \pm 4 \times 10 ⁻³	0.35 \pm 2 \times 10 ⁻³	0.01 \pm 2 \times 10 ⁻³	0.01 \pm 2 \times 10 ⁻³	0.05 \pm 9 \times 10 ⁻³	0.40 \pm 3 \times 10 ⁻³	0.02 \pm 7 \times 10 ⁻³
	LOSS	4.7 \times 10 ⁻³ \pm 4 \times 10 ⁻⁵	2.2 \times 10 ⁻⁶ \pm 1 \times 10 ⁻⁶	3.4 \times 10 ⁻³ \pm 6 \times 10 ⁻⁵	1.9 \times 10 ⁻⁶ \pm 1 \times 10 ⁻⁶	1.7 \times 10 ⁻⁶ \pm 5 \times 10 ⁻⁷	8.0 \times 10 ⁻⁶ \pm 2 \times 10 ⁻⁶	1.5 \times 10 ⁻³ \pm 3 \times 10 ⁻⁵	2.0 \times 10 ⁻⁶ \pm 7 \times 10 ⁻⁷
	EPOCHS	25 \pm 5	32 \pm 4	27 \pm 4	38 \pm 4	36 \pm 2	36 \pm 5	32 \pm 8	33 \pm 2
<hr/>									
<i>SMorphTanh</i>	α_1	-0.53 \pm 0.001	-29.67 \pm 0.02	-1.12 \pm 0.001	-33.94 \pm 0.46	-45.34 \pm 0.04	-21.71 \pm 0.04	-6.64 \pm 2.32	-30.00 \pm 0.06
	α_2	-5.18 \pm 0.008	47.32 \pm 0.06	-0.63 \pm 0.003	52.63 \pm 0.07	54.49 \pm 0.02	17.00 \pm 0.01	-0.63 \pm 0.06	49.47 \pm 0.03
	RMSE ₁	0.46 \pm 9 \times 10 ⁻⁴	0.04 \pm 9 \times 10 ⁻³	0.44 \pm 2 \times 10 ⁻³	0.02 \pm 7 \times 10 ⁻³	0.11 \pm 3 \times 10 ⁻³	0.42 \pm 3 \times 10 ⁻²	0.09 \pm 4 \times 10 ⁻³	
	RMSE ₂	0.23 \pm 1 \times 10 ⁻³	0.01 \pm 5 \times 10 ⁻³	0.36 \pm 1 \times 10 ⁻³	0.02 \pm 8 \times 10 ⁻³	0.01 \pm 1 \times 10 ⁻³	0.05 \pm 6 \times 10 ⁻³	0.41 \pm 3 \times 10 ⁻³	0.02 \pm 8 \times 10 ⁻³
	LOSS	4.7 \times 10 ⁻³ \pm 9 \times 10 ⁻⁶	1.6 \times 10 ⁻⁶ \pm 5 \times 10 ⁻⁷	3.5 \times 10 ⁻³ \pm 7 \times 10 ⁻⁵	1.9 \times 10 ⁻⁶ \pm 8 \times 10 ⁻⁷	1.8 \times 10 ⁻⁶ \pm 4 \times 10 ⁻⁷	6.9 \times 10 ⁻⁶ \pm 1 \times 10 ⁻⁶	1.6 \times 10 ⁻³ \pm 9 \times 10 ⁻⁵	2.1 \times 10 ⁻⁶ \pm 5 \times 10 ⁻⁷
	EPOCHS	22 \pm 7	32 \pm 5	24 \pm 7	39 \pm 8	37 \pm 6	37 \pm 6	42 \pm 14	34 \pm 4

Figure 13 – Comparaison des poids appris et des moyennes et écarts-types des métriques α , RMSE, loss et nombre d'époques, et ce sur six runs, entre SMorph et SMorphTanh à deux couches, pour les huit fonctions structurantes cibles et l'opération d'**ouverture**.

Pour la fermeture sur réseaux à deux couches, on obtient les résultats suivants :

Opération : fermeture									
Jeu : MNIST									
<i>SMorph</i>		w_1/w_2	-0.31 ± 0.001	49.73 ± 0.30	-0.33 ± 0.001	49.26 ± 0.21	50.74 ± 0.34	28.71 ± 0.06	24.21 ± 8.94
α_1			-4.54 ± 0.008	-49.21 ± 0.19	-8.44 ± 0.028	-47.38 ± 0.12	-50.22 ± 0.17	-40.74 ± 0.09	37.49 ± 0.17
α_2									-25.36 ± 0.11
RMSE ₁			$0.28 \pm 2 \times 10^{-4}$	$0.02 \pm 2 \times 10^{-3}$	$0.38 \pm 7 \times 10^{-4}$	$0.021 \pm 3 \times 10^{-4}$	$0.02 \pm 1 \times 10^{-4}$	$0.02 \pm 6 \times 10^{-3}$	$0.14 \pm 2 \times 10^{-2}$
RMSE ₂			$0.55 \pm 9 \times 10^{-4}$	$0.01 \pm 7 \times 10^{-4}$	$0.54 \pm 1 \times 10^{-3}$	$0.004 \pm 2 \times 10^{-4}$	$0.02 \pm 2 \times 10^{-4}$	$0.02 \pm 4 \times 10^{-3}$	$0.12 \pm 3 \times 10^{-3}$
LOSS			$1.8 \times 10^{-3} \pm 6 \times 10^{-6}$	$1.4 \times 10^{-3} \pm 3 \times 10^{-6}$	$4.5 \times 10^{-3} \pm 8 \times 10^{-6}$	$1.8 \times 10^{-5} \pm 6 \times 10^{-6}$	$1.1 \times 10^{-5} \pm 3 \times 10^{-6}$	$1.9 \times 10^{-5} \pm 1 \times 10^{-6}$	$3.2 \times 10^{-4} \pm 1 \times 10^{-5}$
EPOCHS			17 ± 2	30 ± 9	18 ± 4	32 ± 5	34 ± 7	32 ± 8	33 ± 8
<i>SMorphTanh</i>									
w_1/w_2			-0.43 ± 0.001	49.82 ± 0.44	1.48 ± 0.004	49.18 ± 0.22	50.50 ± 0.44	28.70 ± 0.05	4.35 ± 6.16
α_1			-5.11 ± 0.012	-49.20 ± 0.36	-1.78 ± 0.003	-47.29 ± 0.17	-50.03 ± 0.34	-40.72 ± 0.07	-3.87 ± 5.02
α_2									37.26 ± 0.16
RMSE ₁			$0.32 \pm 1 \times 10^{-3}$	$0.02 \pm 8 \times 10^{-4}$	$0.57 \pm 7 \times 10^{-4}$	$0.018 \pm 2 \times 10^{-4}$	$0.01 \pm 4 \times 10^{-4}$	$0.02 \pm 3 \times 10^{-3}$	$0.12 \pm 4 \times 10^{-3}$
RMSE ₂			$0.57 \pm 9 \times 10^{-4}$	$0.01 \pm 3 \times 10^{-4}$	$0.65 \pm 5 \times 10^{-4}$	$0.004 \pm 4 \times 10^{-4}$	$0.01 \pm 3 \times 10^{-4}$	$0.03 \pm 9 \times 10^{-3}$	$0.13 \pm 3 \times 10^{-3}$
LOSS			$2.0 \times 10^{-3} \pm 4 \times 10^{-5}$	$1.5 \times 10^{-5} \pm 8 \times 10^{-6}$	$1.8 \times 10^{-3} \pm 6 \times 10^{-6}$	$1.4 \times 10^{-5} \pm 4 \times 10^{-6}$	$1.4 \times 10^{-5} \pm 4 \times 10^{-6}$	$2.5 \times 10^{-5} \pm 9 \times 10^{-6}$	$1.8 \times 10^{-3} \pm 5 \times 10^{-4}$
EPOCHS			17 ± 1	32 ± 8	23 ± 4	27 ± 2	28 ± 4	31 ± 4	27 ± 6

Figure 14 – Comparaison des poids appris et des moyennes et écarts-types des métriques α , RMSE, loss et nombre d'époques, et ce sur six runs, entre SMorph et SMorphTanh à deux couches, pour les huit fonctions structurantes cibles et l'opération de **fermeture**.

Pour les réseaux à une couche : il n'existe pas de grandes différences entre les résultats portés par SMorphNet et ceux portés par SMorphNetTanh. L'ensemble des résultats semble plutôt bon pour ces deux réseaux uni-couche, aussi bien au niveau de la *loss*, de la *RMSE*, du signe et de l'amplitude de α , et de la vitesse de convergence. Il existe toutefois deux expériences dont les résultats entre ces deux réseaux divergent légèrement : la première avec *cross3* où SMorphNetTanh semble meilleur en terme de *RMSE* et de *loss*; la seconde avec *adiag* où cette fois SMorphNet semble meilleur.

Pour les réseaux à deux couches : on retrouve les mêmes problèmes de convergence avec *cross3* et *disk2* que dans l'état de l'art (partie 2.4). On remarque que *adiag* pose ici également des problèmes pour les deux réseaux. Cependant, là où SMorphNet et SMorphNetTanh se comportent de la même manière sur ces trois structures cibles avec l'ouverture, ils se comportent plutôt différemment avec la fermeture. Pour les autres structures, les deux réseaux donnent des résultats à la fois bons et similaires.

Ces expériences ont été réalisées avec une autre banque d'images, FashionMNIST, et les mêmes conclusions de similarité d'efficacité des deux réseaux ont pu être établies.

En conclusion, les deux modèles ont une précision et une vitesse de convergence similaires. Quelques rares différences dans les résultats subsistent sur certaines expériences, autant en faveur de SMorphNetTanh qu'en sa défaveur. SMorphTanh est donc équivalent à SMorph en terme de performances. Il résout néanmoins le problème de signe devant w dans la formulation asymptotique de SMorph lorsque $\alpha \rightarrow -\infty$. On gardera alors le SMorphNetTanh comme réseau d'étude pour la suite.

3.2 Partage de poids

3.2.1 Contexte et motivations

La partie précédente montre que les réseaux morphologiques donnent de bons résultats lorsqu'ils ne possèdent qu'une seule couche morphologique et qu'ils n'ont qu'une simple opération d'érosion ou de dilatation à faire. Cependant, lorsque l'on considère des réseaux à deux couches pour des opérations cibles plus complexes comme l'ouverture et la fermeture, qui sont la succession de deux opérations morphologiques fondamentales, l'ensemble des configurations possibles se multiplie, et la convergence des réseaux se comporte plutôt mal dans plusieurs cas, car ils convergent vers des minima locaux sur la fonction de perte *loss* lors de leur entraînement.

C'est pourquoi l'objectif principal devient la recherche de solutions à ces problèmes de convergence pour les réseaux morphologiques à deux couches. Dans les expériences précédentes, on remarque que les cas d'échec sont corrélés à une différence de forme entre les noyaux des deux couches. Or, en morphologie mathématique, dans le cadre d'une ouverture ou d'une fermeture (2.1), c'est le même élément structurant qui est utilisé pour à la fois l'érosion et la dilatation. De ce fait, dans le cadre d'une opération cible d'ouverture ou de fermeture dans les réseaux morphologiques, on peut avoir l'idée de créer une certaine influence entre les poids de la première couche du réseau et ceux de la seconde, qui puisse faire tendre les deux noyaux à avoir symétriquement les mêmes poids selon le signe de leur paramètres de contrôle (α_1/α_2) respectif.

Deux méthodes d'influence des poids (dites méthodes de « partage de poids ») des deux couches d'un réseau ont été implémentées et testées. La première est une méthode de partage de poids dite « dure ». Il s'agit là de simplement transférer les poids du noyau de la première couche à la seconde couche du réseau lors du passage de la matrice d'images. Dans cette méthode de transfert direct des poids, on peut même considérer les deux couches comme une seule et unique couche, dite « duale », qui sera appliquée deux fois successivement (mais de manière différente, i.e. symétriquement selon les conditions) à l'image d'entrée de la couche, et qui possède un seul et unique noyau w , mais bien deux paramètres de contrôle indépendants, α_1 et α_2 .

La seconde méthode de partage des poids entre la première et la seconde couche morphologique du réseau est une méthode plus indirecte, dite « douce ». Il s'agit de la formulation d'une métrique de contrainte entre les poids des deux couches, qui est simplement ajoutée dans la *loss*. La correspondance symétrique des deux noyaux n'est ainsi pas directement imposée, mais elle est contrainte dans la fonction de perte : le réseau va alors, lors de son entraînement, tendre vers la similarité symétrique entre les deux noyaux (hors des α) en même temps qu'il tend vers une configuration favorable aux bonnes prédictions avec la *MSE* entre images cibles et prédictions, et ce de manière plus ou moins prioritaire sur cette dernière en fonction d'un paramètre λ .

3.2.2 Partage de poids dur

La méthode de partage de poids dite « dure » fait un partage direct des poids du noyau de la première couche avec ceux du noyau de la seconde couche du réseau. Les deux couches ayant ainsi exactement le même noyau, on peut les considérer comme une seule et même couche, dite « *duale* », mais avec bien deux paramètres de contrôle indépendants, notés α_1 et α_2 . Ces dernières permettent aux poids du même noyau w d'être appliqués deux fois successivement mais de manière différente à l'image en entrée de couche, la première fois sous le contrôle de α_1 , la seconde fois sous celui de α_2 .

Cependant, d'après les formules aux limites de $\mathcal{SM}\text{orphTanh}$ (p. 22), il est nécessaire de considérer la symétrie de w quelque part lorsque α_1 et α_2 sont de signe opposé, et donc pour que le réseau puisse effectivement converger vers une ouverture ou une fermeture de f par w . On va d'abord considérer la formule de $\mathcal{SM}\text{orphTanh}$ sous symétrie conditionnelle, notée $\mathcal{SMT}_{\text{cond}}$, qui est définie pour toute image $f : I \subseteq \mathbb{Z}^2 \rightarrow \mathbb{R}$, tout noyau $w : W \subseteq \mathbb{Z}^2 \rightarrow \mathbb{R}$ et tout réel $\alpha \in \mathbb{R}$ de la manière suivante (11) :

$$\mathcal{SMT}_{\text{cond}}(f, w, \alpha) = \begin{cases} \mathcal{SM}\text{orphTanh}(f, w, \alpha) & \text{si } \alpha \geq 0 \\ \mathcal{SM}\text{orphTanh}(f, \check{w}, \alpha) & \text{sinon} \end{cases} \quad (11)$$

Malgré la conditionnalité sur α dans la définition de la fonction $\mathcal{SMT}_{\text{cond}}$, la transition du noyau w à sa symétrique, quand α passe d'un signe négatif à un signe positif ou inversement dans la formule de $\mathcal{SM}\text{orphTanh}$, reste lisse, rendant $\mathcal{SMT}_{\text{cond}}$ continue (mais pas forcément dérivable) en $\alpha = 0$, grâce au $\tanh(\alpha)$ devant w .

La fonction de transformation de la couche $\mathcal{SM}\text{orphTanhDual}$ est ainsi définie, pour toute image $f : I \subseteq \mathbb{Z}^2 \rightarrow \mathbb{R}$ et tout noyau de couche $w : W \subseteq \mathbb{Z}^2 \rightarrow \mathbb{R}$, et avec $\alpha_1, \alpha_2 \in \mathbb{R}$, par :

$$\mathcal{SM}\text{orphTanhDual}(f, w, \alpha_1, \alpha_2) = \mathcal{SMT}_{\text{cond}}(\mathcal{SMT}_{\text{cond}}(f, w, \alpha_1), w, \alpha_2) \quad (12)$$

Ainsi, si α_1 et α_2 ont un signe opposé, on obtient l'ouverture (si $\alpha_1 \ll 0$ et $\alpha_2 \gg 0$) ou la fermeture (si $\alpha_1 \gg 0$ et $\alpha_2 \ll 0$) de l'image f par la fonction structurante w .

On a donc les limites asymptotiques de $\mathcal{SM}\text{orphTanhDual}$ suivantes :

$$\lim_{\substack{\alpha_1 \rightarrow -\infty \\ \alpha_2 \rightarrow +\infty}} \mathcal{SM}\text{orphTanhDual}(f, w, \alpha_1, \alpha_2) = f \circ w$$

$$\lim_{\substack{\alpha_1 \rightarrow +\infty \\ \alpha_2 \rightarrow -\infty}} \mathcal{SM}\text{orphTanhDual}(f, w, \alpha_1, \alpha_2) = f \bullet w$$

On obtient bien avec $\mathcal{SM}\text{orphTanhDual}$, aux limites, l'ouverture et la fermeture.

3.2.3 Partage de poids doux

La méthode de partage de poids dite « douce », quant à elle, ne modifie pas la structure-même du réseau. Elle se base sur l'expression d'une métrique de contrainte entre les poids des noyaux des deux couches morphologiques du réseau, qui est alors ajoutée dans la fonction de perte globale *loss*, et qui va ainsi faire tendre ces deux noyaux symétriquement vers la même forme durant l'entraînement du réseau, sans donc passer par un transfert direct des poids d'une couche à l'autre. La contrainte fait ainsi prendre la même forme aux deux noyaux, mais de manière plus douce.

Rappelons d'abord la formule de l'erreur quadratique moyenne (MSE) entre deux fonctions f_1 et f_2 définies sur le même ensemble dénombrable I à valeurs dans \mathbb{R} , qui est utilisée dans notre cas pour calculer l'éloignement entre deux images :

$$\text{MSE}(f_1, f_2) = \mathbb{E}((f_2 - f_1)^2) = \frac{1}{|I|} \sum_{x \in I} (f_2(x) - f_1(x))^2 \quad (13)$$

La fonction de perte *loss* du réseau, qui permet d'entraîner et d'évaluer les performances de prédiction de ce dernier, se formule classiquement comme la MSE entre les images cibles F_{cible} et les images prédites par le réseau F_{pred} : $\text{loss} = \text{MSE}(F_{\text{cible}}, F_{\text{pred}})$. Dans la *loss*, est alors ici ajoutée à $\text{MSE}(F_{\text{cible}}, F_{\text{pred}})$ une contrainte de similarité C_{sim} entre le noyau de la première couche, w_1 , et celui de la seconde, w_2 . Cette contrainte est pondérée par l'hyperparamètre λ qui permet de moduler son influence dans la convergence du réseau lors de l'entraînement, par rapport à celle de $\text{MSE}(F_{\text{cible}}, F_{\text{pred}})$.

On étudie trois formulations de cette métrique de contrainte entre w_1 et w_2 . Elles sont explicitées à la page suivante. Pour mieux les comprendre, définissons d'abord deux fonctions de rééchelonnage linéaire primaires, la normalisation et la standardisation.

On définit la « normalisation » comme la fonction $\Omega_{(a,b)}$ qui à une image $f : I \rightarrow \mathbb{R}$ associe son rééchelonnage linéaire (rescale) dans un intervalle $[a, b] \subset \mathbb{R}$:

$$\Omega_{(a,b)}(f) = \frac{f - \inf_{x \in I} \{f(x)\}}{\sup_{x \in I} \{f(x)\} - \inf_{x \in I} \{f(x)\}} (b - a) + a \quad (14)$$

On définit la « standardisation » comme la fonction \mathfrak{S} qui à une image $f : I \rightarrow \mathbb{R}$ associe son rééchelonnage linéaire de moyenne $\mathbb{E}(\mathfrak{S}(f)) = 0$ et d'écart-type $\sigma(\mathfrak{S}(f)) = 1$ sur l'ensemble de ses valeurs d'arrivée $f(I)$:

$$\mathfrak{S}(f) = \frac{f - \mathbb{E}(f)}{\sigma(f)} = \frac{f - \frac{1}{|I|} \sum_{x \in I} f(x)}{\sqrt{\frac{1}{|I|} \sum_{x \in I} (f(x) - \frac{1}{|I|} \sum_{x \in I} f(x))^2}} \quad (15)$$

a. Première formulation

La première formulation de la métrique de contrainte C_{sim} est définie à partir de la MSE entre les deux noyaux w_1 et w_2 , symétrisés l'un par rapport à l'autre et normalisés (14) entre 0 et 1. Ainsi, C_{sim} est définie pour tous noyaux w_1 et w_2 par :

$$C_{\text{sim}}(w_1, w_2) = \text{MSE}\left(\Omega_{(0,1)}(w_1), \Omega_{(0,1)}(\check{w}_2)\right) \quad (16)$$

b. Deuxième formulation

La deuxième formulation de la métrique de contrainte C_{sim} est définie à partir de la MSE entre les deux noyaux w_1 et w_2 , symétrisés l'un par rapport à l'autre et cette fois standardisés (15). Ainsi, C_{sim} est définie pour tous noyaux w_1 et w_2 par :

$$C_{\text{sim}}(w_1, w_2) = \text{MSE}(\mathfrak{S}(w_1), \mathfrak{S}(\check{w}_2)) \quad (17)$$

c. Troisième formulation

La troisième formulation de la métrique de contrainte C_{sim} est définie à partir de la MSE entre les deux noyaux w_1 et w_2 , symétrisés l'un par rapport à l'autre et standardisés, mais multipliés par l'écart-type des poids d'un des deux noyaux, afin d'éviter de changer et d'influencer, avec un écart relatif, la taille de la gamme des poids des noyaux. C_{sim} est alors définie pour tous noyaux w_1 et w_2 par :

$$C_{\text{sim}}(w_1, w_2) = \text{MSE}\left(w_1 - \mathbb{E}(w_1), (\check{w}_2 - \mathbb{E}(\check{w}_2)) \frac{\sigma(w_1)}{\sigma(\check{w}_2)}\right) \quad (18)$$

Remarque :

Si on prend ici la symétrique de \check{w}_2 de w_2 , c'est parce que, comme précédemment, le problème de symétrie dans l'expression asymptotique de la formule de SMorphTanh (p. 22) nécessite de prendre la symétrie d'un des deux noyaux w . De plus, si les trois formulations se basent sur un rééchelonnage des noyaux, c'est parce qu'ils peuvent tous deux s'exprimer dans deux intervalles de \mathbb{R} très différents liés à la standardisation en entrée du réseau et au filtre de convolution de taille 1 en sortie. On imagine alors qu'il doit exister une relation linéaire entre la disposition dans \mathbb{R} des poids du premier noyau w_1 et la disposition dans \mathbb{R} de ceux du second noyau w_2 .

Avec C_{sim} définie par l'une des trois formulations présentées (16, 17, 18), la nouvelle formule de la *loss* dans le cadre de ce partage de poids doux s'écrit ainsi :

$$\text{loss} = \text{MSE}\left(F_{\text{cible}}, F_{\text{pred}}\right) + \lambda \times C_{\text{sim}}(w_1, w_2) \quad (19)$$

Comparaison rapide de résultats de convergence entre ces trois formulations

Pour savoir quelle formulation de la contrainte de similarité C_{sim} entre w_1 et w_2 est la plus adaptée à nos besoins, on exécute plusieurs expériences d'entraînement de réseau SMorphNetTanh à deux couches, avec séparément les trois différentes formulations de C_{sim} telles que décrites précédemment, ajoutées dans la *loss*. On observe alors rapidement la forme des deux noyaux w_1 et w_2 après convergence du réseau. La figure suivante présente les résultats, associés par paires (la première image pour w_1 , la seconde qui suit pour w_2), sur l'ensemble des expériences avec les huit structures cibles, et deux banques d'images (MNIST et FashionMNIST).

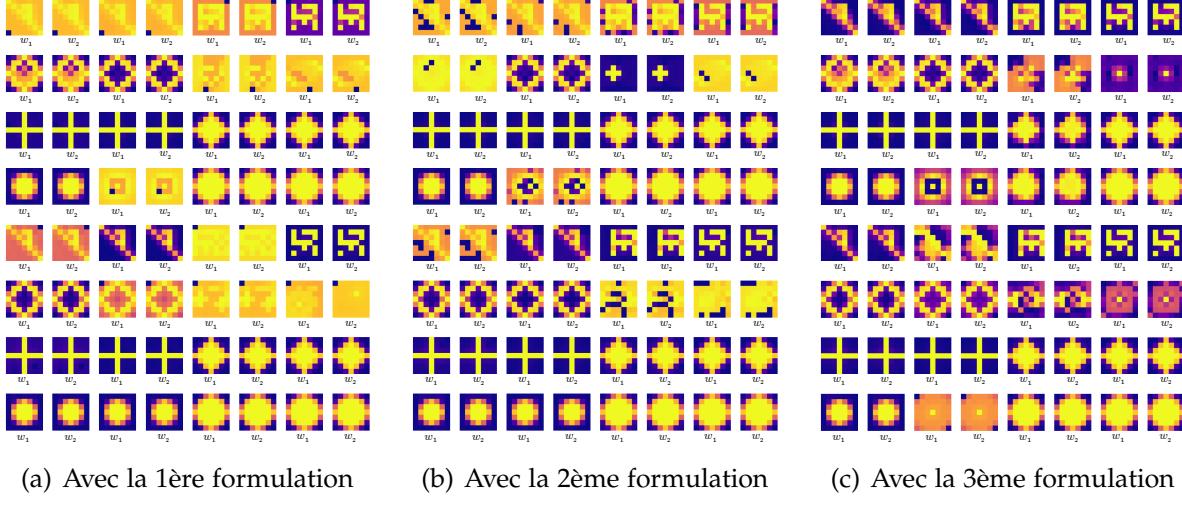


Figure 15 – Forme des noyaux w_1 et w_2 par paires du réseau SMorphNetTanh à deux couches après convergence, sur un ensemble de 32 expériences, avec dans la *loss* la première formulation de C_{sim} (15(a)), la deuxième (15(b)) et la troisième (15(c)).

Sur la figure, chaque emplacement d'un résultat d'une des trois cartes à l'autre représente la même expérience. On remarque tout de suite le problème avec la première formulation 15(a) basée sur la normalisation (14) : elle a tendance à binariser les filtres sur certaines expériences (première ligne, les deux premières paires par exemple), en éloignant un unique pixel très bas dans les valeurs (en bleu foncé, jusqu'à -80). Mais cela n'est que visuel, il n'y a pas d'impact sur l'opération car, en dessous d'un certain seuil, toute valeur a le même effet que $-\infty$. De même pour la seconde formulation 15(b) basée, elle, sur la standardisation (15). Cependant, elle a tendance à éloigner plusieurs pixels très bas dans les valeurs, plutôt qu'un unique pixel. Quant à la troisième, il n'y a pas ce problème de binarisation, car pas de mise en échelle des gammes. Elle donne les résultats qu'on attendait, similaires à ceux avec la *loss* sans contrainte, mais avec les deux noyaux de chaque paire qui tendent bien vers la même forme.

On remarque néanmoins que, sur certaines expériences, là où la troisième formulation donne les mêmes mauvais résultats (échecs) de convergence que les expériences sans contrainte dans la *loss* (uniquement la MSE entre les images cibles et les prédictions), les deux autres formulations permettent au réseau de trouver la vraie fonction structurante cible pour ces quelques expériences cibles-là (par exemple, avec le *disk2* en fermeture pour MNIST, dernière ligne et deuxième paire à partir de la gauche).

Ces bons résultats des deux premières formulations (16 et 17), qui surpassent la troisième (18) sur ces quelques échecs de convergence, semblent liés au rééchelonnage de la gamme de valeurs des poids w_1 et w_2 , qui fait qu'un éloignement d'une partie des valeurs entraîne une binarisation de ces dernières, qui réduit alors drastiquement la valeur de la contrainte $C_{\text{sim}}(w_1, w_2)$. Ils pourront être considérés plus tard dans le cadre de l'amélioration de la convergence des réseaux sur ces expériences-là.

On préférera la troisième formulation dans cette optique de partage de poids doux, car seule cette dernière a le comportement que l'on visait initialement, c'est-à-dire faire uniquement tendre les deux noyaux vers la même forme sans créer d'éloignement de valeurs et de binarisation, pour ensuite voir les résultats d'une telle contrainte et estimer si elle permet à elle seule une amélioration de la convergence des réseaux.

3.2.4 Comparaison de l'efficacité

Pour comparer l'efficacité des partages de poids dur et doux, on reprend le même protocole et les mêmes expériences que précédemment avec *SMorphNetTanh* : on considère les quatre mêmes métriques d'évaluation de la performance des réseaux (*loss*, *RMSE*, valeurs de α_1 et α_2 , et nombre d'époques pour converger) ; on reprend la banque d'images en niveaux de gris MNIST (on fait également, à part, les expériences sur FashionMNIST) ; on fait toujours six runs par expérience ; et on considère enfin les huit fonctions structurantes cibles telles qu'illustrées sur les figures suivantes.

Dans le cas du partage de poids doux, l'hyperparamètre λ est fixé à 0.01. Plusieurs expériences ont été réalisées avec différentes valeurs de λ (inutile de les analyser dans ce rapport), et c'est cette valeur-ci qui semble donner les meilleurs résultats. Comme expliqué précédemment, on considère ici la formule de la *loss* 19, avec C_{sim} la troisième formulation (18) de la contrainte de similarité sur les deux noyaux w_1 et w_2 . Concernant le partage de poids dur, la *loss* considérée est simplement la classique MSE entre les images cibles et les prédictions. Comme le réseau ne comporte qu'une seule couche, *SMorphTanhDual*, telle que décrite précédemment, avec un unique noyau, seul ce noyau en fin de convergence du réseau et sa *RMSE* associée seront affichés.

Les deux figures suivantes 16 et 17 font la synthèse des résultats de l'ensemble des expériences réalisées, et fait la comparaison entre ceux obtenus sans partage de poids, ceux avec un partage dur (dual), et ceux avec un partage doux (contrainte).

3.2 Partage de poids

3 AMÉLIORATION DES RÉSEAUX

Opération : ouverture jeu : MNIST									
w_1/w_2									
Sans partage	α_1	-0.53 ± 0.001	-29.67 ± 0.02	-1.12 ± 0.001	-33.94 ± 0.46	-45.34 ± 0.04	-21.71 ± 0.04	-6.64 ± 2.32	-30.00 ± 0.06
	α_2	-5.18 ± 0.008	47.32 ± 0.06	-0.63 ± 0.003	52.63 ± 0.07	54.49 ± 0.02	17.00 ± 0.01	-0.63 ± 0.06	49.47 ± 0.03
	RMSE1	0.46 ± 9 × 10 ⁻⁴	0.04 ± 9 × 10 ⁻³	0.44 ± 2 × 10 ⁻³	0.02 ± 7 × 10 ⁻²	0.02 ± 7 × 10 ⁻³	0.11 ± 3 × 10 ⁻³	0.42 ± 3 × 10 ⁻²	0.09 ± 4 × 10 ⁻³
	RMSE2	0.23 ± 1 × 10 ⁻³	0.01 ± 5 × 10 ⁻³	0.36 ± 1 × 10 ⁻³	0.02 ± 8 × 10 ⁻³	0.01 ± 1 × 10 ⁻³	0.05 ± 6 × 10 ⁻³	0.41 ± 3 × 10 ⁻³	0.02 ± 8 × 10 ⁻³
	LOSS	4.7 × 10 ⁻⁶ ± 9 × 10 ⁻⁶	1.6 × 10 ⁻⁶ ± 5 × 10 ⁻⁷	3.5 × 10 ⁻³ ± 7 × 10 ⁻⁵	1.9 × 10 ⁻⁶ ± 8 × 10 ⁻⁷	1.8 × 10 ⁻⁶ ± 4 × 10 ⁻⁷	6.9 × 10 ⁻⁶ ± 1 × 10 ⁻⁶	1.6 × 10 ⁻³ ± 9 × 10 ⁻⁵	2.1 × 10 ⁻⁶ ± 5 × 10 ⁻⁷
	EPOCHS	22 ± 7	32 ± 5	24 ± 7	39 ± 8	37 ± 6	37 ± 6	42 ± 14	34 ± 4
Partage dur	α_1	-0.51 ± 0.003	-30.28 ± 0.02	-1.04 ± 0.002	-32.29 ± 1.91	-45.09 ± 0.01	-29.70 ± 0.06	-4.31 ± 0.32	-23.85 ± 1.47
	α_2	-1.17 ± 0.003	47.80 ± 0.05	-0.49 ± 0.002	52.06 ± 2.32	55.93 ± 0.01	19.05 ± 0.03	-0.38 ± 0.03	40.69 ± 2.29
	RMSE	0.37 ± 1 × 10 ⁻³	0.03 ± 2 × 10 ⁻³	0.47 ± 2 × 10 ⁻³	0.03 ± 4 × 10 ⁻³	0.02 ± 5 × 10 ⁻³	0.10 ± 2 × 10 ⁻³	0.48 ± 1 × 10 ⁻²	0.04 ± 4 × 10 ⁻³
	LOSS	5.5 × 10 ⁻³ ± 5 × 10 ⁻⁵	1.6 × 10 ⁻⁶ ± 4 × 10 ⁻⁷	3.6 × 10 ⁻³ ± 1 × 10 ⁻⁴	1.9 × 10 ⁻⁶ ± 9 × 10 ⁻⁸	1.8 × 10 ⁻⁶ ± 5 × 10 ⁻⁷	5.7 × 10 ⁻⁶ ± 9 × 10 ⁻⁷	2.1 × 10 ⁻³ ± 1 × 10 ⁻⁵	2.1 × 10 ⁻⁶ ± 4 × 10 ⁻⁷
	EPOCHS	21 ± 2	36 ± 7	29 ± 9	35 ± 5	40 ± 5	37 ± 2	31 ± 8	35 ± 6
Partage doux	α_1	-0.52 ± 0.002	-29.98 ± 0.04	-1.21 ± 0.01	-34.89 ± 0.02	-45.51 ± 0.01	-23.24 ± 0.03	-3.70 ± 0.008	-29.35 ± 0.10
	α_2	-5.25 ± 0.011	47.58 ± 0.01	-0.61 ± 0.01	53.79 ± 0.03	55.07 ± 0.01	17.55 ± 0.05	-0.36 ± 0.003	49.59 ± 0.05
	RMSE1	0.43 ± 3 × 10 ⁻⁴	0.01 ± 2 × 10 ⁻³	0.53 ± 1 × 10 ⁻²	0.004 ± 6 × 10 ⁻⁴	0.005 ± 2 × 10 ⁻³	0.11 ± 7 × 10 ⁻⁴	0.51 ± 5 × 10 ⁻⁴	0.05 ± 2 × 10 ⁻³
	RMSE2	0.42 ± 2 × 10 ⁻³	0.01 ± 2 × 10 ⁻³	0.53 ± 1 × 10 ⁻²	0.004 ± 6 × 10 ⁻⁴	0.005 ± 2 × 10 ⁻³	0.11 ± 7 × 10 ⁻⁴	0.51 ± 4 × 10 ⁻⁴	0.05 ± 2 × 10 ⁻³
	LOSS	4.9 × 10 ⁻³ ± 3 × 10 ⁻⁵	1.5 × 10 ⁻⁷ ± 6 × 10 ⁻⁷	3.6 × 10 ⁻³ ± 3 × 10 ⁻⁵	1.4 × 10 ⁻⁶ ± 4 × 10 ⁻⁷	1.5 × 10 ⁻⁶ ± 4 × 10 ⁻⁷	5.1 × 10 ⁻⁶ ± 5 × 10 ⁻⁷	2.0 × 10 ⁻³ ± 2 × 10 ⁻⁵	1.6 × 10 ⁻⁶ ± 5 × 10 ⁻⁷
	EPOCHS	18 ± 2	37 ± 6	17 ± 4	34 ± 5	44 ± 3	39 ± 6	37 ± 6	36 ± 9

Figure 16 – Comparaison des poids appris et des moyennes et écarts-types des métriques α , RMSE, loss et nombre d'époques, et ce sur six runs, entre partage de poids dur, doux et sans partage, pour les huit fonctions structurantes cibles et l'opération d'ouverture**.**

Opération : fermeture jeu : MNIST									
w_1/w_2									
Sans partage	α_1	-0.43 ± 0.001	49.82 ± 0.44	1.48 ± 0.004	49.18 ± 0.22	50.50 ± 0.44	28.70 ± 0.05	4.35 ± 6.16	37.26 ± 0.16
	α_2	-5.11 ± 0.012	-49.20 ± 0.36	-1.78 ± 0.003	-47.29 ± 0.17	-50.03 ± 0.34	-40.72 ± 0.07	-3.87 ± 5.02	-25.29 ± 0.07
	RMSE1	0.32 ± 1 × 10 ⁻³	0.02 ± 8 × 10 ⁻⁴	0.57 ± 7 × 10 ⁻⁴	0.018 ± 2 × 10 ⁻⁴	0.01 ± 4 × 10 ⁻⁴	0.02 ± 3 × 10 ⁻³	0.44 ± 1 × 10 ⁻¹	0.12 ± 4 × 10 ⁻³
	RMSE2	0.57 ± 9 × 10 ⁻⁴	0.01 ± 3 × 10 ⁻⁴	0.65 ± 2 × 10 ⁻⁴	0.004 ± 6 × 10 ⁻⁴	0.01 ± 3 × 10 ⁻⁴	0.03 ± 8 × 10 ⁻³	0.43 ± 1 × 10 ⁻¹	0.13 ± 3 × 10 ⁻³
	LOSS	2.0 × 10 ⁻³ ± 4 × 10 ⁻⁵	1.5 × 10 ⁻⁵ ± 8 × 10 ⁻⁶	1.8 × 10 ⁻³ ± 6 × 10 ⁻⁶	1.4 × 10 ⁻⁵ ± 4 × 10 ⁻⁶	1.4 × 10 ⁻⁵ ± 4 × 10 ⁻⁶	2.5 × 10 ⁻⁵ ± 9 × 10 ⁻⁶	1.8 × 10 ⁻³ ± 5 × 10 ⁻⁴	3.1 × 10 ⁻⁴ ± 4 × 10 ⁻⁶
	EPOCHS	17 ± 1	32 ± 8	23 ± 4	27 ± 2	28 ± 4	31 ± 4	27 ± 6	28 ± 3
Partage dur	α_1	-0.38 ± 0.001	50.62 ± 0.06	1.60 ± 0.003	56.55 ± 0.24	52.89 ± 0.05	34.80 ± 0.09	49.84 ± 0.88	44.89 ± 0.20
	α_2	-0.73 ± 0.003	-49.63 ± 0.05	-1.97 ± 0.006	-50.97 ± 0.16	-51.91 ± 0.08	-45.21 ± 0.14	-19.41 ± 0.10	-34.36 ± 0.19
	RMSE	0.26 ± 9 × 10 ⁻⁴	0.02 ± 8 × 10 ⁻³	0.61 ± 2 × 10 ⁻⁴	0.02 ± 1 × 10 ⁻⁴	0.02 ± 2 × 10 ⁻⁵	0.03 ± 8 × 10 ⁻³	0.16 ± 2 × 10 ⁻³	0.11 ± 3 × 10 ⁻³
	LOSS	3.2 × 10 ⁻³ ± 6 × 10 ⁻⁵	1.2 × 10 ⁻⁵ ± 1 × 10 ⁻⁶	1.5 × 10 ⁻³ ± 8 × 10 ⁻⁶	1.2 × 10 ⁻⁵ ± 1 × 10 ⁻⁶	1.3 × 10 ⁻⁵ ± 1 × 10 ⁻⁶	2.1 × 10 ⁻⁵ ± 3 × 10 ⁻⁶	7.3 × 10 ⁻⁴ ± 7 × 10 ⁻⁶	3.2 × 10 ⁻⁴ ± 4 × 10 ⁻⁶
	EPOCHS	42 ± 9	35 ± 5	22 ± 3	31 ± 5	36 ± 7	26 ± 2	33 ± 5	30 ± 5
Partage doux	α_1	-0.42 ± 0.001	50.25 ± 0.06	1.51 ± 0.05	54.82 ± 0.23	52.21 ± 0.22	29.24 ± 0.06	39.68 ± 7.81	38.01 ± 0.06
	α_2	-5.73 ± 0.014	-49.46 ± 0.05	-1.91 ± 0.09	-49.41 ± 0.15	-50.90 ± 0.22	-40.95 ± 0.08	-17.53 ± 0.08	-28.65 ± 0.09
	RMSE1	0.22 ± 7 × 10 ⁻⁴	0.02 ± 8 × 10 ⁻⁴	0.56 ± 2 × 10 ⁻⁴	0.005 ± 2 × 10 ⁻⁴	0.007 ± 2 × 10 ⁻³	0.03 ± 2 × 10 ⁻³	0.07 ± 8 × 10 ⁻³	0.13 ± 8 × 10 ⁻⁴
	RMSE2	0.23 ± 1 × 10 ⁻³	0.02 ± 7 × 10 ⁻⁴	0.56 ± 2 × 10 ⁻²	0.005 ± 2 × 10 ⁻⁴	0.007 ± 4 × 10 ⁻⁴	0.03 ± 2 × 10 ⁻³	0.08 ± 8 × 10 ⁻³	0.13 ± 7 × 10 ⁻⁴
	LOSS	2.1 × 10 ⁻³ ± 2 × 10 ⁻⁵	1.1 × 10 ⁻⁴ ± 2 × 10 ⁻⁶	1.8 × 10 ⁻⁴ ± 9 × 10 ⁻⁵	1.4 × 10 ⁻⁵ ± 3 × 10 ⁻⁶	1.4 × 10 ⁻⁵ ± 6 × 10 ⁻⁶	2.2 × 10 ⁻⁵ ± 3 × 10 ⁻⁶	5.2 × 10 ⁻⁴ ± 1 × 10 ⁻⁵	3.2 × 10 ⁻⁴ ± 3 × 10 ⁻⁶
	EPOCHS	30 ± 8	34 ± 7	19 ± 6	33 ± 11	29 ± 4	30 ± 5	34 ± 5	32 ± 4

Figure 17 – Comparaison des poids appris et des moyennes et écarts-types des métriques α , RMSE, loss et nombre d'époques, et ce sur six runs, entre partage de poids dur, doux et sans partage, pour les huit fonctions structurantes cibles et l'opération de **fermeture.**

Le réseau avec un partage de poids dur, comportant donc une couche duale, semble converger aussi bien voire un peu mieux que le réseau sans partage de poids. Dans les expériences où les réseaux sans partage de poids convergent initialement bien (*cross7*, *disk3*, *diamond3*, *complex* et *brand*), ceux avec un partage de poids dur convergent au moins aussi bien, voire mieux dans certains cas (dans l'expérience avec la structure *brand* et l'opération cible d'ouverture (fig. 16), il y a par exemple moins d'artéfacts sur le noyau de la couche duale que sur les noyaux du réseau sans partage).

Dans les expériences où les réseaux sans partage de poids convergent, à l'inverse, initialement mal (forme des noyaux et valeurs des métriques trop mauvaises, comme *cross3*, *disk2* et *adiag*), ceux avec un partage de poids dur convergent de la même manière, à l'exception de l'expérience avec la structure *adiag* pour la fermeture (fig. 17), où le noyau de la couche converge ici vers la bonne forme cible, et où les métriques de performance du réseau présentent une perte *loss* et une *RMSE* bien plus faibles.

Le réseau avec un partage doux, quant à lui, présente exactement les mêmes conclusions que le réseau avec un partage dur, bien que les formes des noyaux des réseaux et les métriques de performance soient légèrement différentes. Il présente des résultats similaires au réseau sans partage de poids. Comme pour le partage dur, seule l'expérience avec la structure *adiag* pour l'opération de fermeture montre une bien meilleure convergence du réseau, avec une forme des noyaux proche de celle de la cible et de plutôt bons résultats au niveau des métriques de performance.

En conclusion, ces deux méthodes de partage de poids des noyaux ne permettent pas forcément au réseau d'être plus efficace sur chaque expérience, mais résolvent tout de même le problème de convergence de *adiag* pour la fermeture. De plus, elles donnent toutes deux des résultats similaires, bien que la forme des noyaux puisse être légèrement différente de l'un à l'autre sur certaines expériences. Les conclusions sont les mêmes avec la banque FashionMNIST, bien que les résultats soient assez différents.

Remarque :

Ici, on fait un partage de poids uniquement sur les noyaux. On aurait également pu partager inversement les α des deux couches, mais on considère qu'il s'agit là d'une bien trop forte hypothèse dans de nombreux contextes. En effet, pour des images cibles dont on ne connaît pas l'ensemble des opérations duquel ces images résultent à partir d'images initiales, imposer au réseau bi-couches considéré de partir soit vers une opération de (pseudo-)ouverture soit vers une opération de (pseudo-)fermeture (car, dans un tel partage de poids, on imposerait une valeur opposée aux deux alphas du réseau) est une hypothèse déjà trop forte pour un tel système cible d'opérations inconnues, là où, au contraire, on peut se dire qu'il est plutôt plausible que les deux noyaux du réseau tendent à converger vers la même forme. Un tel partage inverse des α est cependant à considérer si l'on sait à l'avance que l'on veut soit un comportement d'ouverture soit de fermeture pour de tels réseaux à deux couches (partie suivante).

3.3 Contraintes sur les poids

3.3.1 Contexte et motivations

Malgré l'ajout d'un partage de poids dans le réseau *SMorphNetTanh*, certaines expériences restent un échec dans le cadre d'une opération cible d'ouverture ou de fermeture : le réseau aura convergé vers un minimum local. Dans ces échecs, c'est à la fois l'aspect visuel du noyau des couches qui est considéré comme trop éloigné de la forme de la fonction structurante cible, et à la fois les métriques d'évaluation des performances du réseau qui sont considérées comme mauvaises par rapport aux succès. Les poids des noyaux w des couches du réseau, avec les paramètres de contrôle α associés, étant responsables des bonnes ou mauvaises performances du réseau.

En analysant, sur ces échecs, les poids des couches du réseau, que ce soit le noyau w ou le paramètre α , on observe le manque d'une certaine harmonie : d'un côté, les poids dans les noyaux w sont très étalés sur le support, et leurs valeurs tournent autour de la moyenne et manquent d'une certaine binarisation, surtout aux bords du support, rendant difficile la distinction d'un « objet » par rapport au « fond » du support ; et d'un autre côté, on remarque que la valeur des paramètres de contrôle α sont très proches de 0, ne générant que des pseudo-opérations morphologiques, et sont bien souvent du mauvais signe (ils ne prennent pas la « bonne » opération cible). Une idée est alors de créer une « influence » sur ces poids, qui puisse les guider vers la structure souhaitée, sans donner des informations trop précises sur la structure cible.

Cette « influence »-là se base sur le même concept que celui utilisé pour le partage de poids doux : on ajoute, dans la fonction de perte *loss*, une métrique de contrainte sur les poids de couches visés, pondérée par un hyperparamètre λ (voir formule 19). Cette contrainte, fonction définie sur les poids w/α à valeurs dans \mathbb{R} , doit illustrer l'erreur sur l' $«$ aspect $»$ que l'on souhaite ajouter dans les poids du réseau. Il peut s'agir d'un aspect géométrique, d'un aspect morphologique, ou encore d'une distance. En ajoutant cette contrainte dans la *loss*, le réseau va alors se forcer à minimiser cette contrainte lors de son entraînement, faisant tendre les poids vers l'aspect souhaité.

Attention : Ces contraintes-là sont appliquées uniquement dans des cas particuliers où l'on fait une hypothèse forte sur l'aspect des noyaux w et des paramètres α des couches du réseau. Elles sont utilisées seulement dans des contextes où on a des hypothèses sur la forme des poids des couches qui peuvent être justifiées, où l'on sait que la contrainte doit être respectée et est pertinente dans le cas en question. Si on ne sait rien à priori sur la forme des poids que l'on cherche, elles ne sont pas pertinentes.

On distingue ici deux types de contraintes sur les réseaux morphologiques : les contraintes géométriques sur les noyaux w , et les contraintes sur les paramètres de contrôle α . Ces deux types de contraintes sont étudiés dans les pages suivantes.

Mais avant de plonger dans ces métriques et de voir leur impact sur la forme des noyaux et plus généralement sur la convergence des réseaux, rappelons qu'il est important que les fonctions associées soient continues partout, et, si possible, dérivables partout, afin de pouvoir calculer le gradient de la *loss* par rapport aux poids du réseau, nécessaire à la rétropropagation du gradient lors de l'entraînement.

L'application $| \cdot |^\zeta : \mathbb{R} \rightarrow \mathbb{R}$, définie pour $\zeta \in \mathbb{R}$, et qui à un réel x associe $|x|^\zeta$, est beaucoup utilisée dans les différentes métriques de contraintes citées. Étudions donc d'abord rapidement ses propriétés de continuité et de dérivabilité sur \mathbb{R} en fonction de $\zeta \in \mathbb{R}$. On cherche en particulier à connaître les conditions nécessaires et suffisantes sur le paramètre ζ pour que l'application $x \mapsto |x|^\zeta$ soit dérivable partout sur \mathbb{R} .

Sur \mathbb{R}^* :

On sait que l'application $x \mapsto x^\zeta$ est continue et dérivable sur \mathbb{R}_+^* pour tout $\zeta \in \mathbb{R}$. Sur \mathbb{R}_+^* , l'application $x \mapsto |x|^\zeta$ devient $x \mapsto x^\zeta$, qui est donc continue et dérivable $\forall \zeta \in \mathbb{R}$. Sur \mathbb{R}_-^* , $x \mapsto |x|^\zeta$ devient $x \mapsto (-x)^\zeta$, qui est également continue et dérivable $\forall \zeta \in \mathbb{R}$, puisque $\forall x \in \mathbb{R}_-^*, -x \in \mathbb{R}_+^*$. Ainsi, $x \mapsto |x|^\zeta$ est continue et dérivable sur \mathbb{R}^* , $\forall \zeta \in \mathbb{R}$.

En 0 :

On sait que, sur \mathbb{R}_+ , l'application $x \mapsto x^\zeta$ n'est pas définie en 0 pour $\zeta \in \mathbb{R}_-$, car x^ζ tend vers $+\infty$ quand x tend vers 0^+ si $\zeta < 0$. Si $\zeta = 0$, l'application $x \mapsto x^\zeta$ définie sur \mathbb{R}_+^* peut être prolongée par continuité en 0, avec $x^\zeta = 1$ en $x = 0$, car $\forall x \in \mathbb{R}_+, x^0 = 1$. Enfin, si $\zeta \in \mathbb{R}_+^*$, l'application $x \mapsto x^\zeta$ est bien définie et continue sur \mathbb{R}_+ , et donc en 0. Ainsi, l'application $x \mapsto |x|^\zeta$ est définie et continue en 0 si et seulement si $\zeta \geq 0$.

Étudions à présent la dérivabilité de $x \mapsto |x|^\zeta$ en 0, en partant de la définition-même de la dérivabilité. Soit $\zeta > 0$ (excluons le cas $\zeta = 0$ pour l'instant), et soit $h > 0$.

$$\text{À gauche : } \frac{|0-h|^\zeta - |0|^\zeta}{0-h-0} = \frac{|-h|^\zeta - |0|^\zeta}{-h} = \frac{|-h|^\zeta}{-h} = \frac{h^\zeta}{-h} = -h^{\zeta-1}$$

$$\text{À droite : } \frac{|0+h|^\zeta - |0|^\zeta}{0+h-0} = \frac{|h|^\zeta - |0|^\zeta}{h} = \frac{|h|^\zeta}{h} = \frac{h^\zeta}{h} = h^{\zeta-1}$$

L'application $x \mapsto |x|^\zeta$ est donc dérivable en 0 ssi : $\lim_{h \rightarrow 0^+} (-h^{\zeta-1}) = \lim_{h \rightarrow 0^+} (h^{\zeta-1})$. Donc, si $\zeta \neq 0$, elle est dérivable en 0 ssi $\zeta > 1$. Si $\zeta = 0$, elle est également dérivable en 0 (en considérant le prolongement par continuité en 0).

⇒ L'application $x \mapsto |x|^\zeta$ est donc dérivable partout sur \mathbb{R} ssi $\zeta > 1$ ou $\zeta = 0$.

La figure 18 ci-après illustre la courbe de la fonction $x \mapsto |x|^\zeta$ pour quelques valeurs de ζ . Elle permet de comprendre facilement le comportement et l'évolution de cette application en fonction de la valeur de ζ . On constate l'aplatissement de la courbe en 0 quand ζ augmente, et on devine bien la dérivabilité en 0 pour $\zeta > 1$.

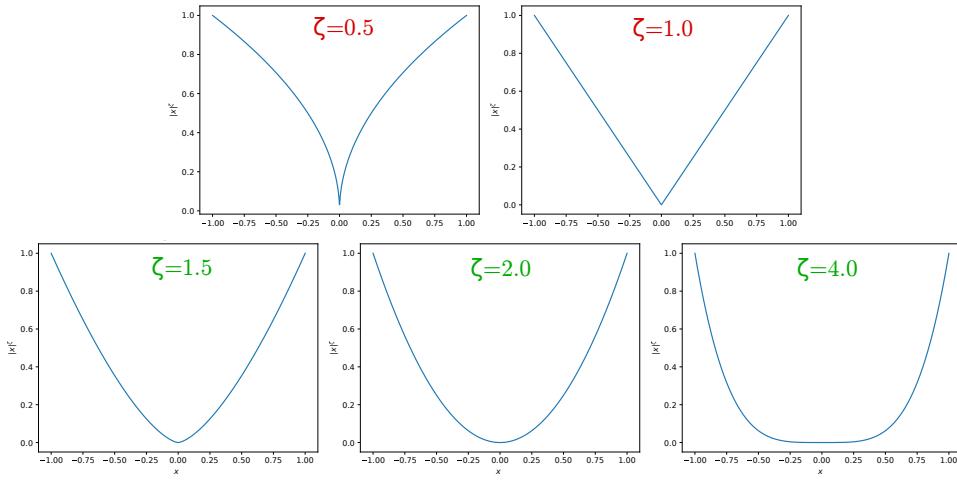


Figure 18 – Courbes de la fonction $x \mapsto |x|^\zeta$ dans l'intervalle $[-1, 1]$, pour différentes valeurs de ζ . En-haut : non dérivable en 0 ($1 \geq \zeta > 0$) ; en-bas : dérivable en 0 ($\zeta > 1$).

La propriété de cette application $x \mapsto |x|^\zeta$ d'être dérivable partout, en particulier en $x = 0$, si et seulement si $\zeta > 1$ (ou $\zeta = 0$, mais ce cas ne sera jamais utilisé) justifient, dans les métriques suivantes, l'utilisation par défaut de $\zeta = 2$ dans les fonctions mises en puissance. Cette propriété est également valable pour la norme euclidienne dans un espace à n dimensions, d'où le fait que ζ soit aussi égal à 2 par défaut dans les formules 21 et 22 de la première et de la deuxième contrainte géométrique (ci-après).

3.3.2 Contraintes géométriques sur les noyaux w

Le premier type de contraintes sur les poids des réseaux morphologiques est celui porté sur la géométrie des noyaux w des couches de ces réseaux. Sont développées et étudiées dans cette sous-partie les différentes métriques de contraintes géométriques utilisées pendant le stage. On distingue ici cinq contraintes géométriques : l'erreur sur le centre de masse C_{center} ; l'erreur sur la dispersion C_{disp} ; l'erreur sur la symétrie C_{sym} ; l'erreur sur la binarité C_{binary} ; et l'erreur sur la norme C_{norm} .

a. Erreur sur le centre de masse

La première métrique de contrainte géométrique sur les noyaux w des réseaux morphologiques est l'erreur sur le centre de masse, notée C_{center} . Soit $w : W \subseteq \mathbb{Z}^2 \rightarrow \mathbb{R}$ un noyau de couche. On prolonge \mathbb{Z}^2 dans \mathbb{R}^2 , et on considère \mathbb{R}^2 comme un espace vectoriel muni de la norme euclidienne $\|\cdot\|$. Le centre de masse $p_w \in \mathbb{R}^2$ de w est défini par :

$$p_w = \frac{1}{\sum_{x \in W} w(x)} \sum_{x \in W} w(x)x \quad (20)$$

Soit $p_c \in \mathbb{R}^2$ le point « central » visé, et soit $\zeta > 0$ (par défaut, $\zeta = 2$). L'erreur C_{center} sur le centre de masse p_w de w par rapport à p_c est définie par :

$$C_{\text{center}}(w, p_c)_\zeta = \frac{\|p_w - p_c\|^\zeta}{\sup_{p \in W} \|p - p_c\|^\zeta} \quad (21)$$

Dans la pratique, on voudra faire tendre le centre de masse p_w du noyau w vers le centre de son support W (de taille finie) dans \mathbb{R}^2 . Dans ce cas : $p_c = \frac{1}{|W|} \sum_{x \in W} x$.

Prenons par exemple l'expérience avec *disk3* pour la fermeture sur FashionMNIST. La figure ci-après montre l'évolution de la convergence des deux noyaux du réseau, avec des checkpoints à différents stades durant l'entraînement, sans et avec cette contrainte C_{center} dans la *loss* (avec $\lambda = 0.01$). Dans les deux cas, le réseau possède le partage de poids doux tel que décrit précédemment. On obtient les mêmes résultats sur 6 runs.

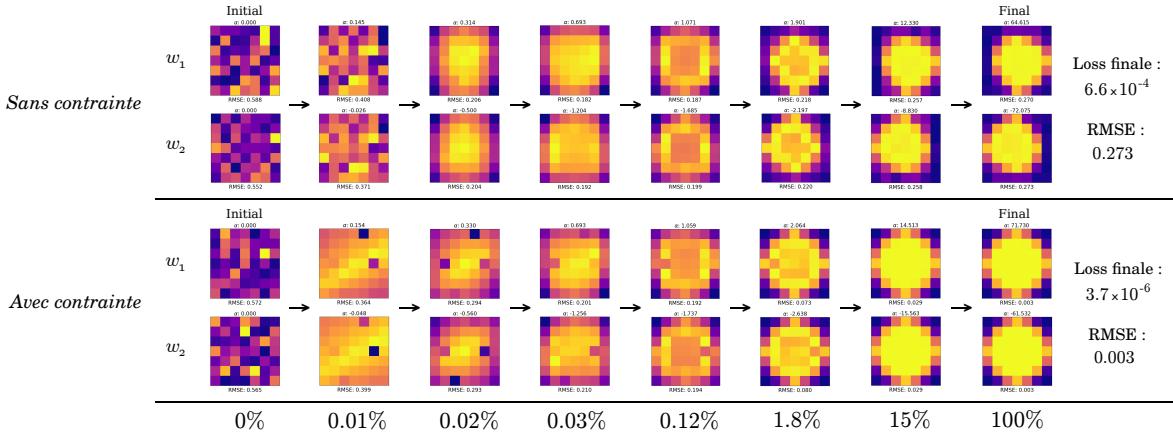


Figure 19 – Évolution de la forme du noyau des deux couches du réseau et de sa RMSE en fonction de la progression de l'entraînement (en % avant l'état final), pour le *disk3* et la fermeture sur la banque FashionMNIST, avec et sans la contrainte C_{center} .

On voit sur la figure 19 que, sur cette expérience par exemple, avec la structure *disk3* pour la fermeture et la banque FashionMNIST, cette métrique est très efficace pour avoir à la fois une RMSE faible (la forme du noyau w ressemble bien à la structure cible) et à la fois une *loss* faible (le réseau a bien convergé). Cependant, cette contrainte ne peut typiquement pas être appliquée dans le cas de fonctions structurantes cibles pour lesquelles le centre de masse est éloigné du centre du support, ou est inconnu. Comme expliqué dans les motivations, on ne pourra ainsi appliquer cette contrainte que dans des contextes particuliers où les informations nécessaires sont connues.

b. Erreur sur la dispersion

La deuxième métrique de contrainte géométrique sur les noyaux w est l'erreur sur la dispersion, notée C_{disp} . Alors que l'erreur sur le centre de masse peut être associée à une moyenne (biais), l'erreur sur la dispersion peut être, quant à elle, associée à une variance. Soit $p_c \in \mathbb{R}^2$ le point de « concentration » visé (en pratique, $p_c = p_w$), et soit $\zeta > 0$ (par défaut, $\zeta = 2$). L'erreur C_{disp} sur la dispersion de w par rapport à p_c est définie par :

$$C_{\text{disp}}(w, p_c)_\zeta = \frac{1}{\sup_{p \in W} \|p - p_c\|^\zeta} \frac{1}{\sum_{p \in W} w(p)} \sum_{p \in W} \|p - p_c\|^\zeta w(p) \quad (22)$$

Prenons par exemple l'expérience avec *disk2* pour la fermeture sur MNIST. La figure ci-après montre l'évolution de la convergence des deux noyaux du réseau, avec des checkpoints à différents stades durant l'entraînement, sans et avec cette contrainte C_{disp} dans la *loss* (avec $\lambda = 0.01$). Dans les deux cas, le réseau possède encore le partage de poids doux décrit précédemment. On obtient les mêmes résultats sur 6 runs.

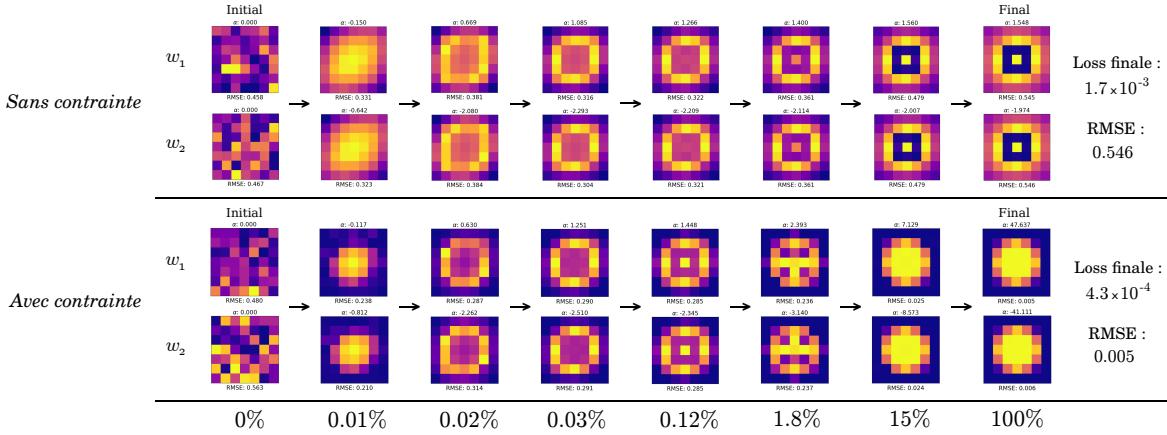


Figure 20 – Évolution de la forme du noyau des deux couches du réseau et de sa RMSE en fonction de la progression de l'entraînement (en % avant l'état final), pour le *disk2* et la fermeture sur la banque MNIST, avec et sans la contrainte C_{disp} .

On voit sur la figure 20 que, sur cette expérience par exemple, avec la structure *disk2* pour la fermeture et la banque MNIST, cette métrique est très efficace pour avoir à la fois une RMSE faible (la forme du noyau w ressemble bien à la structure cible) et à la fois une *loss* faible (le réseau a bien convergé). Cependant encore, cette contrainte ne peut pas être appliquée dans le cas de fonctions structurantes cibles pour lesquelles les poids sont dispersés sur le support, ou dont la configuration est inconnue.

c. Erreur sur la symétrie

La troisième métrique de contrainte géométrique sur les noyaux w des réseaux morphologiques est l'erreur sur la symétrie, notée C_{sym} . Soit $p_c \in \mathbb{R}^2$ le centre de symétrie visé (en pratique, $p_c = p_w$), et soit $\zeta > 0$ (par défaut, $\zeta = 2$). L'erreur C_{sym} sur la symétrie centrale de w par rapport au point p_c est définie par :

$$C_{\text{sym}}(w, p_c)_\zeta = \frac{\sum_{p \in W} \begin{cases} |w(p) - w(2p_c - p)|^\zeta & \text{si } 2p_c - p \in W \\ (w(p) - \inf_{q \in W} \{w(q)\})^\zeta & \text{sinon} \end{cases}}{|W| \left(\sup_{p \in W} \{w(p)\} - \inf_{p \in W} \{w(p)\} \right)^\zeta} \quad (23)$$

Prenons par exemple l'expérience avec *diamond3* pour la fermeture sur la banque FashionMNIST. La figure ci-après montre l'évolution de la convergence des deux noyaux du réseau, avec des checkpoints à différents stades durant l'entraînement, sans et avec cette contrainte C_{sym} dans la *loss* (avec $\lambda = 0.01$). Dans les deux cas, le réseau possède toujours un partage de poids doux. On obtient les mêmes résultats sur 6 runs.

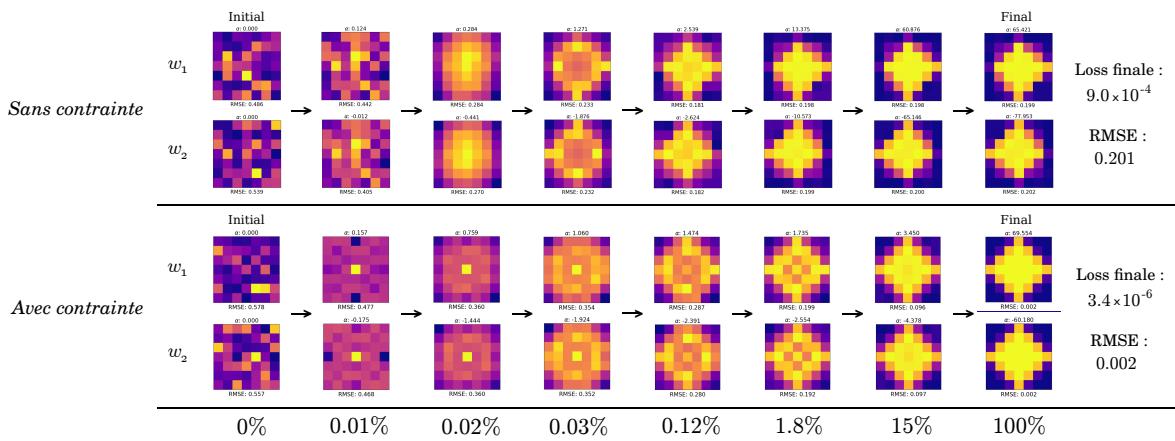


Figure 21 – Évolution de la forme du noyau des deux couches du réseau et de sa *RMSE* en fonction de la progression de l'entraînement (en % avant l'état final), pour le *diamond3* et la fermeture sur la banque FashionMNIST, avec et sans la contrainte C_{sym} .

On voit sur la figure 21 que, sur cette expérience par exemple, avec la structure cible symétrique *diamond3* pour la fermeture et la banque FashionMNIST, cette métrique est efficace pour avoir à la fois une *RMSE* faible (le noyau w est proche de la structure cible) et à la fois une *loss* faible (le réseau a bien convergé). Encore une fois, cette contrainte ne peut cependant pas être appliquée dans le cas de fonctions structurantes cibles non symétriques (centralement), ou si le centre de symétrie est inconnu.

d. Erreur sur la binarité

La quatrième métrique de contrainte géométrique sur les noyaux w des réseaux morphologiques est l'erreur sur la binarité, notée C_{binary} . Elle permet de faire converger les poids de w vers deux bornes pré-définies, une inférieure $b_{\text{inf}} \in \mathbb{R}$ et une supérieure $b_{\text{sup}} \in \mathbb{R}$, avec $b_{\text{sup}} > b_{\text{inf}}$. En pratique, $b_{\text{inf}} = \inf_{p \in W} \{w(p)\}$ et $b_{\text{sup}} = \sup_{p \in W} \{w(p)\}$. Pour $\zeta > 0$ (par défaut, $\zeta = 2$), l'erreur C_{binary} sur la binarité des poids de w aux bornes b_{inf} et b_{sup} est définie par :

$$C_{\text{binary}}(w, (b_{\text{inf}}, b_{\text{sup}}))_{\zeta} = \frac{1}{|W|} \sum_{p \in W} \left| 1 - \left| \frac{2w(p) - b_{\text{sup}} - b_{\text{inf}}}{b_{\text{sup}} - b_{\text{inf}}} \right|^{\zeta} \right| \quad (24)$$

Prenons par exemple l'expérience avec *brand* pour la fermeture sur la banque MNIST. La figure ci-après montre l'évolution de la convergence des deux noyaux du réseau, avec des checkpoints à différents stades durant l'entraînement, sans et avec cette contrainte C_{binary} dans la *loss* (avec $\lambda = 0.0001$). Dans les deux cas, le réseau possède toujours un partage de poids doux. On obtient les mêmes résultats sur 6 runs.

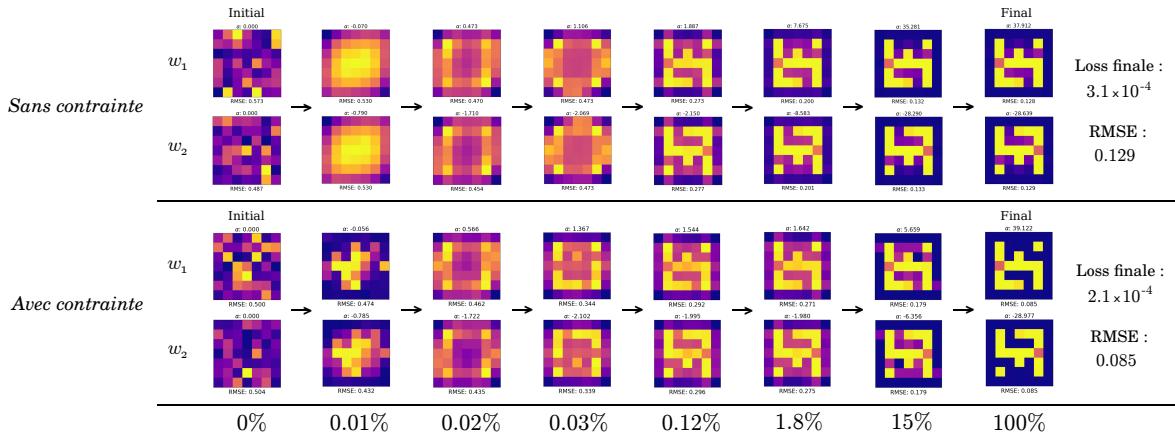


Figure 22 – Évolution de la forme du noyau des deux couches du réseau et de sa RMSE en fonction de la progression de l'entraînement (en % avant l'état final), pour le *brand* et la fermeture sur la banque MNIST, avec et sans la contrainte C_{binary} .

On voit sur la figure 22 que, sur cette expérience par exemple, avec la structure cible *brand* pour la fermeture et la banque MNIST, cette métrique permet effectivement de mieux binariser les poids du noyau, et permet également une légère amélioration de la convergence du réseau, aussi bien sur la *loss* que sur la RMSE. Note : les images affichées sont les symétriques des noyaux w . Précisons également que cette contrainte n'a pas de pertinence dans le cas de fonctions structurantes cibles non binaires.

e. Erreur sur la norme

La dernière métrique de contrainte géométrique sur les noyaux w est l'erreur sur leur norme, notée C_{norm} . Elle permet à la norme du noyau de converger vers une valeur cible prédéfinie $l_c \in \mathbb{R}_+$. Par défaut, $l_c = 0$. L'erreur C_{norm} sur la norme n , avec $n > 0$ (par défaut, $n = 1$), du noyau w est définie, avec $\zeta > 0$ (par défaut, $\zeta = 2$), par :

$$C_{\text{norm}}(w, l_c)_{n,\zeta} = \frac{\left| \left(\sum_{p \in W} |w(p)|^n \right)^{1/n} - l_c \right|^{\zeta}}{|W|^{\zeta/n} \left(\sup_{p \in W} \{w(p)\} - \inf_{p \in W} \{w(p)\} \right)^{\zeta}} \quad (25)$$

Par défaut, $n = 1$, car la norme 1 est le meilleur approximateur convexe de la norme 0 (on s'intéresse à la norme 0 pour calculer la taille de l'« objet » du support).

Prenons l'expérience avec *complex* pour la fermeture sur la banque FashionMNIST. La figure ci-après montre l'évolution de la convergence des deux noyaux du réseau durant l'entraînement, sans et avec la contrainte C_{norm} dans la *loss* (avec $\lambda = 0.001$). Le réseau a toujours un partage de poids doux. On obtient ces résultats sur 6 runs.

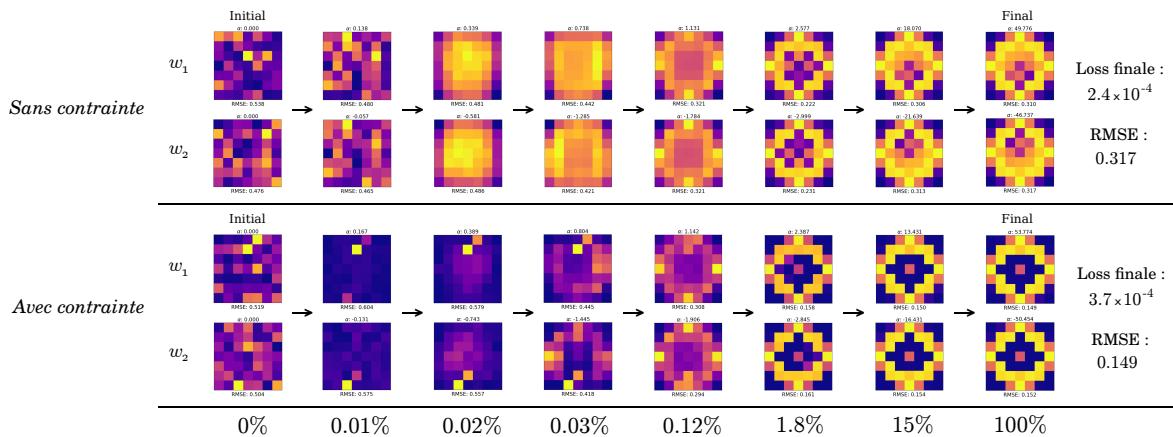


Figure 23 – Évolution de la forme du noyau des deux couches du réseau et de sa RMSE en fonction de la progression de l'entraînement (en % avant l'état final), pour le *complex* et la fermeture sur la banque FashionMNIST, avec et sans la contrainte C_{norm} .

Cette contrainte permet de réduire le nombre de pixels nuisibles et d'artéfacts sur w si l'hyperparamètre λ est bien ajusté dans la *loss*. Les résultats de la figure 23 avec cette contrainte (où $l_c = 0$ et $\lambda = 0.001$) sont meilleurs en terme de RMSE, mais un peu moins bons en terme de *loss* par rapport à ceux sans C_{norm} . Cette contrainte étant utilisée en particulier pour minimiser le nombre de pixels apparents dans w , elle peut donc être utilisée dans tout contexte si on veut avoir un objet minimal dans w sur W .

Remarque pour l'ensemble des contraintes sur w :

Dans les cas pratiques, on utilisera, pour ces cinq métriques de contraintes étudiées, le noyau w normalisé, écrit $\Omega(w)$ (formule 14). Cela permet en particulier de distinguer l'« objet » (dont les valeurs sur $\Omega(w)$ sont près de 1) du « fond » (dont les valeurs sur $\Omega(w)$ sont près de 0) sur le support W . De plus, si on est avec les réseaux SMorphNet (classique, sans tangente hyperbolique Tanh) ou LMorphNet, et si le paramètre α/p associé au noyau w est négatif, alors on prend la négation $-w$ au lieu de w .

3.3.3 Contraintes sur les paramètres de contrôle α

Le second type de contraintes sur les poids des réseaux est celui sur les paramètres de contrôle α des couches morphologiques (ou p pour LMorph et p Conv). On définit ici deux contraintes sur ces paramètres : l'erreur sur l'éloignement individuelle des α de 0, notée C_{away} , et l'erreur sur l'éloignement opposé des α , notée C_{awayOPP} . La première métrique évalue les paramètres α de manière indépendante, là où la seconde évalue deux paramètres de contrôle de deux couches, α_1 et α_2 , en tant que couple. Elles permettent toutes deux de s'éloigner des cas de pseudo-opération.

a. Erreur sur l'éloignement

La première métrique de contrainte sur les paramètres de contrôle α (ou p selon le type de réseau) est l'erreur sur leur éloignement individuelle de 0, notée C_{away} . Elle permet de contraindre la fonction de perte *loss* en la forçant de manière douce à faire éloigner de 0 le poids α d'une couche morphologique, loin dans les positifs ou loin dans les négatifs selon la direction de convergence du réseau lors de l'entraînement, jusqu'à atteindre une valeur minimale objectif de $|\alpha|$, notée α_{\min} .

Avec une valeur de $\alpha_{\min} > 0$ fixée (par défaut, $\alpha_{\min} = 80$), l'erreur C_{away} sur l'éloignement du paramètre α d'une couche morphologique est définie, avec $\zeta > 0$ (par défaut, $\zeta = 2$), par :

$$C_{\text{away}}(\alpha, \alpha_{\min})_{\zeta} = \begin{cases} \left| \frac{|\alpha|^{\zeta} - \alpha_{\min}^{\zeta}}{\alpha_{\min}^{\zeta}} \right|^{\zeta} & \text{si } |\alpha| < \alpha_{\min} \\ 0 & \text{sinon} \end{cases} \quad (26)$$

Dans cette contrainte à nouveau, le paramètre ζ joue le rôle de régulateur de la fonction $x \mapsto |x|^{\zeta}$ en permettant d'adoucir la courbure de cette dernière, en particulier au voisinage de 0, et de la rendre dérivable. La métrique C_{away} est ainsi continue et dérivable partout pour $\zeta > 1$. En particulier donc en $\alpha = 0$, ainsi qu'en $|\alpha| = \alpha_{\min}$, là où la dérivée de la fonction est nulle pour $\zeta > 1$, et donc où la condition dans la définition de C_{away} selon la valeur de $|\alpha|$ n'empêche pas la dérivabilité de cette fonction.

Prenons par exemple l'expérience avec *complex* pour l'ouverture sur la banque d'images MNIST. La figure ci-après montre l'évolution de la convergence des deux noyaux w_1 et w_2 du réseau SMorphNetTanh durant l'entraînement, d'abord sans puis avec la contrainte C_{away} dans la fonction de perte *loss* (avec $\lambda = 0.01$). Comme précédemment, le réseau a toujours un partage de poids doux entre les noyaux de ses deux couches morphologiques. On obtient les résultats suivants sur 6 runs.

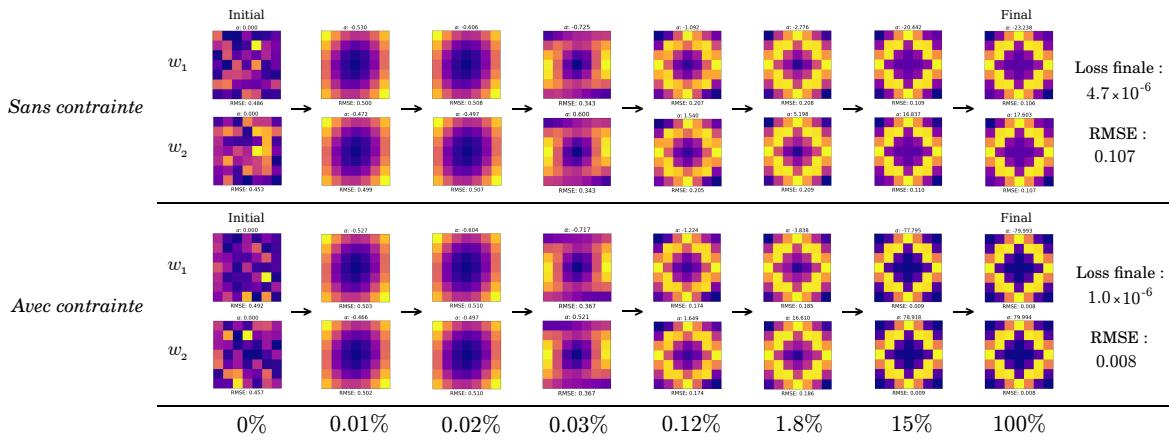


Figure 24 – Évolution de la forme du noyau des deux couches du réseau et de sa RMSE en fonction de la progression de l'entraînement (en % avant l'état final), pour le *complex* et l'ouverture sur la banque MNIST, avec et sans la contrainte C_{away} .

À partir de ces résultats, on remarque que cette contrainte permet bien sur cet exemple d'éloigner les deux α de 0, en les amenant dans ce cas près de 80 en valeur absolue. On remarque également qu'elle permet ici d'améliorer automatiquement la forme des deux noyaux w du réseau, en les rapprochant davantage de la structure cible (les artefacts violets au centre et aux bords des noyaux sont réduits), avec une valeur de RMSE bien plus faible, et cela simplement grâce à cet éloignement des α . La valeur de la *loss* est également plus faible, donc les prédictions sont meilleures grâce à cette contrainte.

Cette contrainte semble efficace lorsque les noyaux w convergent initialement, sans contrainte, vers une forme proche de celle de la fonction structurante cible, et que les paramètres de contrôle α associés sont encore relativement proches de 0 mais bien du signe espéré et opposé pour les opérations cibles d'ouverture et de fermeture. Cependant, après différentes expériences réalisées, cette métrique ne semble fonctionner que moyennement pour les expériences qui sont originellement des échecs sans cette contrainte, et pour lesquelles les paramètres α des deux couches du réseau ont le même signe. C'est le cas par exemple de l'expérience avec *adiag* pour l'ouverture et la banque MNIST, pour laquelle C_{away} ne permet pas l'éloignement des α .

b. Erreur sur l'éloignement opposé

La seconde métrique de contrainte sur les paramètres de contrôle α (ou p) est l'erreur sur leur éloignement opposé, notée C_{awayOPP} . Elle considère deux couches morphologiques en tant que couple, et permet d'éloigner les deux paramètres de contrôle associés, α_1 et α_2 , l'un de l'autre dans une direction opposée dans \mathbb{R} . Elle est définie pour tout couple $(\alpha_1, \alpha_2) \in \mathbb{R}^2$, avec un paramètre de régularisation contrôlée $\alpha_{\text{ctr}} > 0$, par :

$$C_{\text{awayOPP}}((\alpha_1, \alpha_2), \alpha_{\text{ctr}}) = \frac{\alpha_{\text{ctr}}}{\alpha_{\text{ctr}} - \alpha_1 \alpha_2} \quad (27)$$

Cette métrique est continue et dérivable partout tant que $\alpha_1 \alpha_2 < \alpha_{\text{ctr}}$. Par expérience, on sait que si α_1 et α_2 sont du même signe alors qu'ils ne le devraient pas, alors chacun ne dépassera pas 10 en valeur absolue. Ainsi, pour que l'inégalité présentée soit toujours vraie en pratique, on donnera par défaut $\alpha_{\text{ctr}} = 100$.

Prenons l'expérience avec *adiag* pour l'ouverture sur la banque MNIST. La figure ci-après montre l'évolution de la convergence des deux noyaux du réseau durant l'entraînement, sans et avec la contrainte C_{awayOPP} dans la *loss* (avec $\lambda = 0.01$). Le réseau a toujours un partage de poids doux. On obtient les résultats suivants sur 6 runs.

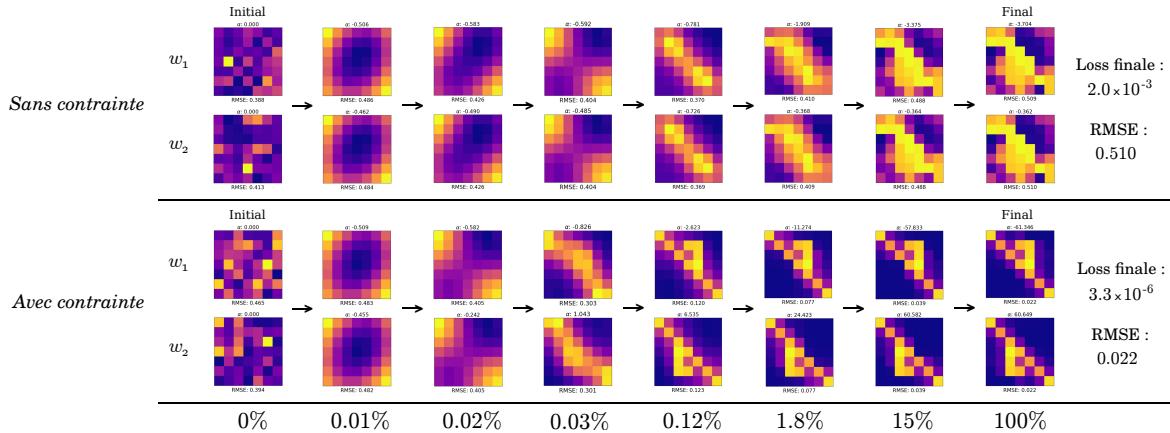


Figure 25 – Évolution de la forme du noyau des deux couches du réseau et de sa RMSE en fonction de la progression de l'entraînement (en % avant l'état final), pour le *adiag* et l'ouverture sur la banque MNIST, avec et sans la contrainte C_{awayOPP} .

Comme le montrent les résultats de cet exemple, cette métrique est très performante dans les échecs de convergence où les α sont de même signe alors qu'ils ne le devraient pas. Cependant, il faut garder à l'esprit le fait que cette métrique ne peut être utilisée que si on sait à priori que les deux α doivent être de signe opposé.

3.3.4 Application et comparaison de l'efficacité

À partir de ces différentes contraintes définies précédemment, celles sur les noyaux w et celles sur les paramètres de contrôle α des couches morphologiques, on crée finalement une *loss* personnalisable, dans laquelle on ajoute les contraintes que l'on souhaite, chacune pondérée par un hyperparamètre propre à elle-même. C'est à l'utilisateur de choisir les contraintes à intégrer selon les contextes, selon la géométrie visée des noyaux du réseau, et selon les valeurs des α recherchées. Cependant, plus il y a de contraintes différentes ajoutées dans la fonction de perte *loss*, plus le réseau aura tendance à « délaisser » certaines contraintes pour d'autres, lors de son entraînement, et moins le réseau convergera bien et sera performant à la fin.

Dans notre cas, les huit fonctions structurantes cibles étudiées ayant des propriétés géométriques bien différentes, l'application d'une même contrainte géométrique parmi celles développées n'est pas adaptée. On peut cependant, dans le cadre des échecs de la partie précédente sur les réseaux à deux couches, faire l'hypothèse forte (que l'on n'utilisera pas dans des contextes où l'on ignore l'hypothèse) que l'on veut que les réseaux se comportent soit comme une ouverture, soit comme une fermeture. On prend ainsi la seconde contrainte sur les paramètres de contrôle α , l'erreur sur l'éloignement opposé (formule 27), qu'on applique sur l'ensemble des expériences de la partie précédente, avec toujours un partage de poids doux en plus.

En reprennant la même méthodologie que dans les parties précédentes avec des réseaux *SMorphNetTanh* à deux couches, on obtient, pour l'opération cible d'ouverture et la banque d'images d'entraînement MNIST, les résultats suivants (fig. 26) :

Opération : ouverture									
Jeu : MNIST		<i>cross3</i>	<i>cross7</i>	<i>disk2</i>	<i>disk3</i>	<i>diamond3</i>	<i>complex</i>	<i>adiag</i>	<i>brand</i>
<i>w</i> ₁ / <i>w</i> ₂									
<i>Sans C_{awayOPP}</i>		-0.59 ± 0.002 -5.25 ± 0.011	-29.98 ± 0.04 47.58 ± 0.01	-1.21 ± 0.01 -0.61 ± 0.01	-34.89 ± 0.02 53.79 ± 0.03	-45.51 ± 0.01 55.07 ± 0.01	-23.24 ± 0.03 17.55 ± 0.05	-3.70 ± 0.008 -0.36 ± 0.003	-29.35 ± 0.10 49.59 ± 0.05
RMSE ₁	0.43 ± 3 × 10 ⁻⁴	0.01 ± 2 × 10 ⁻³	0.53 ± 1 × 10 ⁻²	0.004 ± 6 × 10 ⁻⁴	0.005 ± 2 × 10 ⁻³	0.11 ± 7 × 10 ⁻⁴	0.51 ± 5 × 10 ⁻⁴	0.05 ± 2 × 10 ⁻³	
RMSE ₂	0.42 ± 2 × 10 ⁻³	0.01 ± 2 × 10 ⁻³	0.53 ± 1 × 10 ⁻²	0.004 ± 6 × 10 ⁻⁴	0.005 ± 2 × 10 ⁻³	0.11 ± 7 × 10 ⁻⁴	0.51 ± 4 × 10 ⁻⁴	0.05 ± 2 × 10 ⁻³	
LOSS	4.9 × 10 ⁻⁵ ± 3 × 10 ⁻⁵	1.5 × 10 ⁻⁶ ± 6 × 10 ⁻⁷	3.6 × 10 ⁻³ ± 3 × 10 ⁻⁵	1.4 × 10 ⁻⁶ ± 4 × 10 ⁻⁷	1.5 × 10 ⁻⁶ ± 4 × 10 ⁻⁷	5.1 × 10 ⁻⁶ ± 5 × 10 ⁻⁷	2.0 × 10 ⁻³ ± 2 × 10 ⁻⁵	1.6 × 10 ⁻⁴ ± 5 × 10 ⁻⁷	
EPOCHS	18 ± 2	37 ± 6	17 ± 4	34 ± 5	44 ± 3	39 ± 6	37 ± 6	36 ± 9	
<i>Avec C_{awayOPP}</i>									
<i>w</i> ₁ / <i>w</i> ₂		-0.40 ± 0.001 -1.29 ± 0.004	-59.42 ± 0.06 63.19 ± 0.07	-58.12 ± 0.10 63.23 ± 0.02	-58.47 ± 0.07 65.01 ± 0.04	-60.90 ± 0.05 65.04 ± 0.01	-39.62 ± 0.66 67.98 ± 0.68	-61.29 ± 0.07 60.72 ± 0.06	-58.97 ± 0.12 63.88 ± 0.02
α ₁	0.39 ± 1 × 10 ⁻³	0.01 ± 8 × 10 ⁻⁴	0.03 ± 5 × 10 ⁻⁴	0.004 ± 7 × 10 ⁻⁴	0.004 ± 1 × 10 ⁻³	0.01 ± 2 × 10 ⁻³	0.03 ± 8 × 10 ⁻³	0.05 ± 5 × 10 ⁻³	
α ₂	0.39 ± 1 × 10 ⁻³	0.01 ± 8 × 10 ⁻⁴	0.03 ± 5 × 10 ⁻⁴	0.004 ± 8 × 10 ⁻⁴	0.004 ± 1 × 10 ⁻³	0.01 ± 3 × 10 ⁻³	0.03 ± 8 × 10 ⁻³	0.05 ± 5 × 10 ⁻³	
RMSE ₁	1.6 × 10 ⁻² ± 2 × 10 ⁻⁵	2.1 × 10 ⁻⁶ ± 5 × 10 ⁻⁷	6.2 × 10 ⁻⁶ ± 7 × 10 ⁻⁷	1.2 × 10 ⁻⁶ ± 2 × 10 ⁻⁷	1.9 × 10 ⁻⁶ ± 2 × 10 ⁻⁷	2.3 × 10 ⁻⁶ ± 1 × 10 ⁻⁶	4.0 × 10 ⁻⁶ ± 7 × 10 ⁻⁷	1.8 × 10 ⁻⁶ ± 6 × 10 ⁻⁷	
RMSE ₂	29 ± 7	38 ± 4	33 ± 5	41 ± 6	36 ± 6	40 ± 6	45 ± 13	39 ± 10	
EPOCHS									

Figure 26 – Comparaison des poids appris et des moyennes et écarts-types des métriques α , RMSE, loss et nombre d'époques, et ce sur six runs, entre sans et avec la contrainte C_{awayOPP} , pour les huit fonctions structurantes cibles et l'opération d'ouverture**.**

On obtient également, pour l'opération cible de fermeture et la banque MNIST avec des réseaux SMorphNetTanh à deux couches, les résultats suivants (fig. 27) :

Opération : fermeture									
Jeu : MNIST									
<i>Sans C_{awayOPP}</i>		w_1/w_2	-0.42 ± 0.001	50.25 ± 0.06	1.51 ± 0.05	54.82 ± 0.23	52.21 ± 0.22	29.24 ± 0.06	39.68 ± 7.81
α_1		-5.73 ± 0.014	-49.46 ± 0.05	-1.91 ± 0.09	-49.41 ± 0.15	-50.90 ± 0.22	-40.95 ± 0.08	-17.53 ± 0.08	-28.65 ± 0.09
α_2		$0.22 \pm 7 \times 10^{-4}$	$0.02 \pm 8 \times 10^{-4}$	$0.56 \pm 2 \times 10^{-2}$	$0.005 \pm 2 \times 10^{-4}$	$0.007 \pm 2 \times 10^{-4}$	$0.03 \pm 2 \times 10^{-3}$	$0.07 \pm 8 \times 10^{-3}$	$0.13 \pm 8 \times 10^{-4}$
RMSE ₁		$0.23 \pm 1 \times 10^{-3}$	$0.02 \pm 7 \times 10^{-4}$	$0.56 \pm 2 \times 10^{-2}$	$0.005 \pm 2 \times 10^{-4}$	$0.007 \pm 4 \times 10^{-4}$	$0.03 \pm 2 \times 10^{-3}$	$0.08 \pm 8 \times 10^{-3}$	$0.13 \pm 7 \times 10^{-4}$
RMSE ₂									
LOSS		$2.1 \times 10^{-8} \pm 2 \times 10^{-5}$	$1.1 \times 10^{-7} \pm 2 \times 10^{-6}$	$1.8 \times 10^{-3} \pm 9 \times 10^{-5}$	$1.4 \times 10^{-5} \pm 3 \times 10^{-6}$	$1.4 \times 10^{-5} \pm 6 \times 10^{-6}$	$2.2 \times 10^{-5} \pm 3 \times 10^{-6}$	$5.2 \times 10^{-4} \pm 1 \times 10^{-5}$	$3.2 \times 10^{-4} \pm 3 \times 10^{-6}$
EPOCHS		30 ± 8	34 ± 7	19 ± 6	33 ± 11	29 ± 4	30 ± 5	34 ± 5	32 ± 4
<i>Avec C_{awayOPP}</i>									
w_1/w_2		-0.34 ± 0.001	62.56 ± 0.39	62.14 ± 0.03	64.42 ± 0.14	63.18 ± 0.34	47.21 ± 0.52	81.37 ± 3.29	55.17 ± 0.38
α_1		-2.37 ± 0.005	-62.42 ± 0.29	-60.55 ± 0.01	-62.16 ± 0.06	-63.08 ± 0.25	-62.33 ± 0.66	-20.73 ± 0.49	-41.24 ± 0.15
α_2		$0.28 \pm 3 \times 10^{-4}$	$0.01 \pm 8 \times 10^{-4}$	$0.009 \pm 1 \times 10^{-4}$	$0.005 \pm 1 \times 10^{-4}$	$0.005 \pm 2 \times 10^{-4}$	$0.03 \pm 3 \times 10^{-3}$	$0.06 \pm 2 \times 10^{-3}$	$0.12 \pm 2 \times 10^{-3}$
RMSE ₁		$0.28 \pm 6 \times 10^{-4}$	$0.01 \pm 7 \times 10^{-4}$	$0.009 \pm 6 \times 10^{-5}$	$0.005 \pm 2 \times 10^{-4}$	$0.005 \pm 3 \times 10^{-4}$	$0.03 \pm 3 \times 10^{-3}$	$0.07 \pm 2 \times 10^{-3}$	$0.12 \pm 2 \times 10^{-3}$
RMSE ₂									
LOSS		$1.3 \times 10^{-2} \pm 2 \times 10^{-5}$	$1.2 \times 10^{-5} \pm 2 \times 10^{-6}$	$1.4 \times 10^{-5} \pm 3 \times 10^{-6}$	$1.3 \times 10^{-5} \pm 2 \times 10^{-6}$	$1.5 \times 10^{-5} \pm 2 \times 10^{-6}$	$1.9 \times 10^{-5} \pm 3 \times 10^{-6}$	$5.2 \times 10^{-4} \pm 5 \times 10^{-6}$	$3.2 \times 10^{-4} \pm 6 \times 10^{-6}$
EPOCHS		25 ± 5	32 ± 8	37 ± 4	33 ± 4	31 ± 5	33 ± 7	34 ± 7	30 ± 5

Figure 27 – Comparaison des poids appris et des moyennes et écarts-types des métriques α , RMSE, loss et nombre d'époques, et ce sur six runs, entre sans et avec la contrainte C_{awayOPP} , pour les huit fonctions structurantes cibles et l'opération de fermeture.

Les résultats de convergence de SMorphNetTanh sur les figures 26 et 27, montrant l'état final des noyaux des réseaux et les valeurs des métriques de performances associées pour les différentes expériences réalisées, permettent de mettre en évidence l'efficacité de la métrique de contrainte C_{awayOPP} , qui influe sur les paramètres de contrôle α , et qui est ici associée à un partage de poids doux, dans le cadre d'une opération cible d'ouverture (26) et de fermeture (27) sur des réseaux à deux couches.

Les résultats en couleur verte sont ceux pour lesquels les réseaux avec cette contrainte C_{awayOPP} ont de bien meilleurs résultats. Cela concerne en particulier les expériences avec *disk2*, *complex* et *adiag* pour l'ouverture, et l'expérience avec *disk2* pour la fermeture. On obtient les mêmes résultats avec les expériences faites sur FashionMNIST. Pour ces expériences-là, la contrainte C_{awayOPP} couplée à un partage de poids doux a permis au réseau de bien converger vers l'état cible, avec une forme des noyaux proche de celle de la fonction structurante cible, et avec des valeurs de α du bon signe et éloignées de 0 (on a donc plus une opération exacte qu'une pseudo-opération).

Cependant, on remarque qu'il existe toujours une fonction structurante cible pour laquelle cette contrainte n'a pas d'impact, et ce à la fois pour l'ouverture et la fermeture : *cross3*. En examinant les valeurs de α , on remarque qu'elles restent très proches de 0 et sont du même signe, malgré la présence de la contrainte C_{awayOPP} . Elle n'est donc pas encore suffisante pour améliorer la convergence de l'ensemble des échecs.

3.4 Modulation de l'initialisation

3.4.1 Motivations et principe

Les résultats précédents montrent qu'il reste encore quelques cas d'échecs de convergence pour les réseaux SMorphNetTanh à deux couches morphologiques, avec les opérations cibles d'ouverture et de fermeture, et ce malgré l'ajout d'un partage doux des poids entre les noyaux w des deux couches et l'ajout dans la *loss* d'une contrainte C_{awayOPP} entre les deux paramètres de contrôle α (ou p) associés.

La modulation de l'initialisation des noyaux w du réseau devient alors une solution potentielle à ces derniers échecs. Alors que les poids des noyaux sont originellement initialisés selon une loi normale centrée réduite afin de brasser les différents états de convergence possibles à partir d'initialisations aléatoires différentes sur plusieurs expériences, on peut finalement se dire que cette forme initiale des noyaux peut être fixée et définie de manière déterminée au préalable, selon des caractéristiques morphologiques générales à priori que l'on connaît ou que l'on recherche sur ces noyaux.

En l'occurrence, on peut facilement s'imaginer que ce que l'on recherche ou favorise dans nos différents cas d'étude, c'est une forme de noyaux centrée sur son support et peu dispersée. On choisira alors une initialisation sous la forme d'une gaussienne 2D, allant de 0 aux coins de l'image des noyaux à 1 en son centre (voir figure 28 suivante).

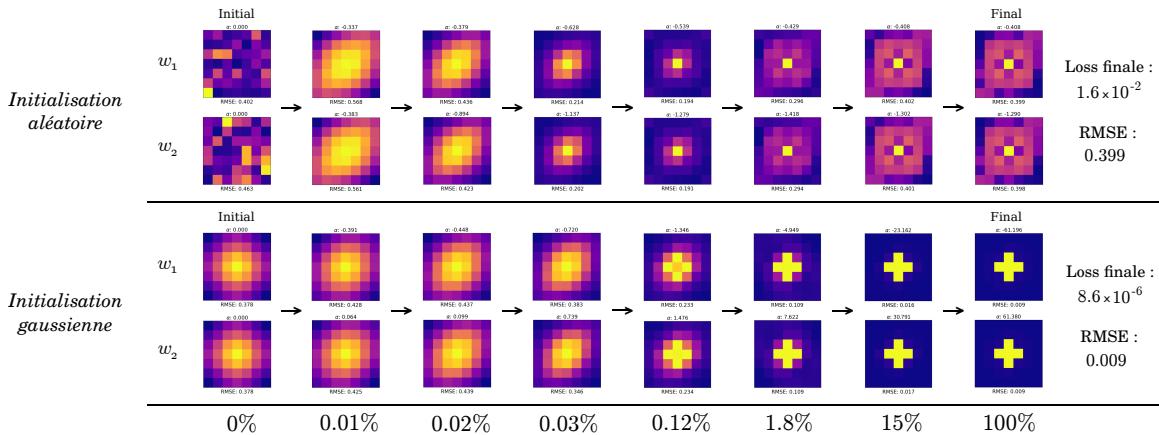


Figure 28 – Évolution de la forme du noyau des deux couches du réseau et de sa *RMSE* en fonction de la progression de l'entraînement (en % avant l'état final), pour le *cross3* et l'ouverture sur la banque MNIST, avec initialisation aléatoire et init. gaussienne.

L'exemple fig. 28, obtenu sur six runs, montre bien ici l'efficacité de cette initialisation gaussienne des noyaux w par rapport à l'aléatoire normale centrée réduite sur *cross3* pour l'ouverture sur MNIST, avec des valeurs de *RMSE* et de *loss* bien plus faibles.

3.4.2 Comparaison de l'efficacité

La première expérience avec la fonction structurante cible *cross3* pour l'opération cible d'ouverture et avec la banque d'images d'entraînement MNIST, expérience pour laquelle le réseau ne converge originellement pas avec une initialisation aléatoire des poids des noyaux w issue d'une loi normale centrée réduite, s'avère ici être un succès grâce à l'initialisation gaussienne 2D de ces noyaux (voir figure précédente 28). Cependant, pour estimer avec davantage de certitude l'efficacité globale de cette modulation de l'initialisation des noyaux w des couches du réseau, il faut observer son impact sur l'ensemble des expériences de notre étude, donc également sur les expériences qui sont originellement un succès avec une initialisation aléatoire des w .

Considérons alors ici le même paramétrage du réseau SMorphNetTanh à deux couches morphologiques que dans la partie précédente, c'est-à-dire munie, dans la fonction de perte *loss*, d'un partage de poids doux entre les noyaux w de ses deux couches (formule 18), et de la contrainte C_{awayOPP} sur l'éloignement opposé entre ses deux paramètres de contrôle α respectifs (formule 27). On modifie donc simplement ici, sur l'ensemble des expériences de notre étude, l'initialisation des poids des noyaux w par une gaussienne 2D telle que décrite précédemment (illustrée fig. 28 en début d'entraînement, sur la seconde ligne, à 0%), et les résultats de convergence issus de cette initialisation déterministe sont comparés à ceux issus de l'initialisation aléatoire.

En reprenant donc la même méthodologie que précédemment avec des réseaux SMorphNetTanh à deux couches, on obtient, pour l'opération cible d'ouverture et la banque d'images d'entraînement MNIST, les résultats suivants (fig. 29) :

Opération : ouverture Jeu : MNIST									
	w_1/w_2								
<i>Initialisation aléatoire</i>	α_1	-0.40 \pm 0.001	-59.42 \pm 0.06	-58.12 \pm 0.10	-58.47 \pm 0.07	-60.90 \pm 0.05	-39.62 \pm 0.66	-61.29 \pm 0.07	-58.97 \pm 0.12
	α_2	-1.29 \pm 0.004	63.19 \pm 0.07	63.23 \pm 0.02	65.01 \pm 0.04	65.04 \pm 0.01	67.98 \pm 0.68	60.72 \pm 0.06	63.88 \pm 0.02
	RMSE1	0.39 \pm 1 \times 10 ⁻³	0.01 \pm 8 \times 10 ⁻⁴	0.03 \pm 5 \times 10 ⁻⁴	0.004 \pm 7 \times 10 ⁻⁴	0.004 \pm 1 \times 10 ⁻³	0.01 \pm 2 \times 10 ⁻³	0.03 \pm 8 \times 10 ⁻³	0.05 \pm 5 \times 10 ⁻³
	RMSE2	0.39 \pm 1 \times 10 ⁻³	0.01 \pm 8 \times 10 ⁻⁴	0.03 \pm 5 \times 10 ⁻⁴	0.004 \pm 8 \times 10 ⁻⁴	0.004 \pm 1 \times 10 ⁻³	0.01 \pm 3 \times 10 ⁻³	0.03 \pm 8 \times 10 ⁻³	0.05 \pm 5 \times 10 ⁻³
	LOSS	1.6 \times 10 ⁻² \pm 2 \times 10 ⁻⁵	2.1 \times 10 ⁻⁶ \pm 5 \times 10 ⁻⁷	6.2 \times 10 ⁻⁶ \pm 7 \times 10 ⁻⁷	1.2 \times 10 ⁻⁶ \pm 2 \times 10 ⁻⁷	1.9 \times 10 ⁻⁶ \pm 2 \times 10 ⁻⁷	2.3 \times 10 ⁻⁶ \pm 1 \times 10 ⁻⁶	4.0 \times 10 ⁻⁴ \pm 7 \times 10 ⁻⁷	1.8 \times 10 ⁻⁶ \pm 6 \times 10 ⁻⁷
	EPOCHS	29 \pm 7	38 \pm 4	33 \pm 5	41 \pm 6	36 \pm 6	40 \pm 6	45 \pm 13	39 \pm 10
<i>Initialisation gaussienne</i>	w_1/w_2								
	α_1	-61.07 \pm 0.25	-59.33 \pm 0.10	-48.68 \pm 21.34	-58.46 \pm 0.03	-60.94 \pm 0.03	-39.04 \pm 0.44	-61.30 \pm 0.02	-58.97 \pm 0.12
	α_2	61.35 \pm 0.05	63.21 \pm 0.01	52.62 \pm 23.75	65.02 \pm 0.01	65.05 \pm 0.01	68.55 \pm 0.46	60.69 \pm 0.03	63.90 \pm 0.01
	RMSE1	0.009 \pm 2 \times 10 ⁻⁴	0.01 \pm 2 \times 10 ⁻³	0.11 \pm 2 \times 10 ⁻¹	0.004 \pm 1 \times 10 ⁻³	0.003 \pm 4 \times 10 ⁻⁴	0.01 \pm 2 \times 10 ⁻³	0.02 \pm 6 \times 10 ⁻³	0.05 \pm 4 \times 10 ⁻³
	RMSE2	0.009 \pm 2 \times 10 ⁻⁴	0.01 \pm 2 \times 10 ⁻³	0.11 \pm 2 \times 10 ⁻¹	0.004 \pm 1 \times 10 ⁻³	0.003 \pm 4 \times 10 ⁻⁴	0.01 \pm 2 \times 10 ⁻³	0.02 \pm 6 \times 10 ⁻³	0.05 \pm 4 \times 10 ⁻³
	LOSS	8.3 \times 10 ⁻⁶ \pm 4 \times 10 ⁻⁷	2.1 \times 10 ⁻⁶ \pm 6 \times 10 ⁻⁷	2.3 \times 10 ⁻³ \pm 2 \times 10 ⁻³	1.8 \times 10 ⁻⁶ \pm 3 \times 10 ⁻⁷	1.5 \times 10 ⁻⁶ \pm 5 \times 10 ⁻⁷	1.6 \times 10 ⁻⁶ \pm 3 \times 10 ⁻⁷	3.8 \times 10 ⁻⁶ \pm 9 \times 10 ⁻⁷	2.1 \times 10 ⁻⁶ \pm 6 \times 10 ⁻⁷
	EPOCHS	34 \pm 4	34 \pm 4	33 \pm 7	37 \pm 6	42 \pm 9	43 \pm 5	47 \pm 13	33 \pm 3

Figure 29 – Comparaison des poids appris et des moyennes et écarts-types des métriques α , RMSE, loss et nombre d'époques, et ce sur six runs, entre initialisation aléatoire et gaussienne, pour les huit fonctions structurantes cibles et l'opération d'ouverture.

On obtient également, pour l'opération cible de fermeture et la banque MNIST avec des réseaux SMorphNetTanh à deux couches, les résultats suivants (fig. 30) :

Opération : fermeture		Jeu : MNIST							
	w_1/w_2								
Initialisation aléatoire	α_1	-0.34 ± 0.001	62.56 ± 0.39	62.14 ± 0.03	64.42 ± 0.14	63.18 ± 0.34	47.21 ± 0.52	81.37 ± 3.29	55.17 ± 0.38
	α_2	-2.37 ± 0.005	-62.42 ± 0.29	-60.55 ± 0.01	-62.16 ± 0.06	-63.08 ± 0.25	-62.33 ± 0.66	-20.73 ± 0.49	-41.24 ± 0.15
	RMSE ₁	0.28 ± 3 × 10 ⁻⁴	0.01 ± 8 × 10 ⁻⁴	0.009 ± 1 × 10 ⁻⁴	0.005 ± 1 × 10 ⁻⁴	0.005 ± 2 × 10 ⁻⁴	0.03 ± 3 × 10 ⁻³	0.06 ± 2 × 10 ⁻³	0.12 ± 2 × 10 ⁻³
	RMSE ₂	0.28 ± 6 × 10 ⁻⁴	0.01 ± 7 × 10 ⁻⁴	0.009 ± 6 × 10 ⁻⁵	0.005 ± 2 × 10 ⁻⁴	0.005 ± 3 × 10 ⁻⁴	0.03 ± 3 × 10 ⁻³	0.07 ± 2 × 10 ⁻³	0.12 ± 2 × 10 ⁻³
	LOSS	1.3 × 10 ⁻² ± 2 × 10 ⁻⁵	1.2 × 10 ⁻⁵ ± 2 × 10 ⁻⁶	1.4 × 10 ⁻⁵ ± 3 × 10 ⁻⁶	1.3 × 10 ⁻⁵ ± 2 × 10 ⁻⁶	1.5 × 10 ⁻⁵ ± 2 × 10 ⁻⁶	1.9 × 10 ⁻⁵ ± 3 × 10 ⁻⁶	5.2 × 10 ⁻⁷ ± 5 × 10 ⁻⁶	3.2 × 10 ⁻⁴ ± 6 × 10 ⁻⁶
Initialisation gaussienne	EPOCHS	25 ± 5	32 ± 8	37 ± 4	33 ± 4	31 ± 5	33 ± 7	34 ± 7	30 ± 5
	w_1/w_2								
	α_1	-0.34 ± 0.001	62.78 ± 0.14	-0.33 ± 0.001	64.07 ± 0.73	63.28 ± 0.35	46.72 ± 0.62	68.23 ± 14.45	55.29 ± 0.22
	α_2	-2.36 ± 0.005	-62.60 ± 0.07	-3.24 ± 0.004	-61.99 ± 0.34	-63.16 ± 0.20	-62.83 ± 0.85	-38.56 ± 18.10	-41.20 ± 0.19
	RMSE ₁	0.28 ± 9 × 10 ⁻⁴	0.01 ± 8 × 10 ⁻⁴	0.36 ± 2 × 10 ⁻⁴	0.004 ± 1 × 10 ⁻⁴	0.005 ± 2 × 10 ⁻⁴	0.03 ± 3 × 10 ⁻³	0.16 ± 9 × 10 ⁻³	0.12 ± 2 × 10 ⁻³
	RMSE ₂	0.28 ± 1 × 10 ⁻³	0.01 ± 7 × 10 ⁻⁴	0.38 ± 8 × 10 ⁻⁴	0.004 ± 8 × 10 ⁻⁵	0.005 ± 1 × 10 ⁻⁴	0.03 ± 3 × 10 ⁻³	0.15 ± 8 × 10 ⁻²	0.12 ± 2 × 10 ⁻³
	LOSS	1.3 × 10 ⁻⁴ ± 1 × 10 ⁻⁵	9.7 × 10 ⁻⁶ ± 3 × 10 ⁻⁶	1.6 × 10 ⁻² ± 3 × 10 ⁻⁵	1.8 × 10 ⁻⁵ ± 1 × 10 ⁻⁵	1.3 × 10 ⁻⁵ ± 3 × 10 ⁻⁶	2.4 × 10 ⁻⁹ ± 5 × 10 ⁻⁶	1.2 × 10 ⁻⁴ ± 7 × 10 ⁻⁴	3.2 × 10 ⁻⁴ ± 6 × 10 ⁻⁶
	EPOCHS	20 ± 1	36 ± 8	27 ± 6	34 ± 8	36 ± 5	41 ± 10	30 ± 5	34 ± 7

Figure 30 – Comparaison des poids appris et des moyennes et écarts-types des métriques α , RMSE, loss et nombre d'époques, et ce sur six runs, entre initialisation aléatoire et gaussienne, pour les huit fonctions structurantes cibles et l'opération de fermeture.

Les données en rouge illustrent les cas où l'initialisation gaussienne donne de moins bons résultats que l'initialisation aléatoire, tandis que ceux en vert illustrent, à inverses, les cas où c'est l'initialisation gaussienne qui donne de meilleurs résultats.

On remarque en effet que l'expérience avec *cross3* pour l'ouverture avec MNIST est un net succès pour l'initialisation gaussienne contrairement à l'initialisation aléatoire, aussi bien en terme visuel avec une grande ressemblance des noyaux w par rapport à la structure cible *cross3*, qu'en termes de valeurs des métriques de performances, avec une RMSE bien plus faible pour les deux noyaux w , une forte valeur absolue des deux paramètres de contrôle α avec le bon signe, et une perte globale *loss* mineure.

Cependant, il s'agit là de la seule expérience dont les résultats sont meilleurs avec l'initialisation gaussienne. Il y a, à l'inverse, plusieurs expériences qui ont cette fois moins bien convergé avec une telle initialisation, telles que celles avec *disk2* à la fois pour l'ouverture et la fermeture, ou encore celle avec *adiag* pour la fermeture (qui a des écarts-types élevés, illustrant l'échec de certaines runs), sur MNIST. Certains succès avec une initialisation aléatoire sont ainsi devenus des échecs avec la gaussienne.

En conclusion, on ne peut pas dire que l'initialisation aléatoire donne de meilleurs résultats de convergence sur ces réseaux, d'autant plus que l'on obtient les mêmes résultats, voir pires, avec la banque FashionMNIST. On s'interroge en particulier sur la raison de l'échec des expériences avec la structure cible *disk2*, qui a pourtant une forme visuelle proche de la gaussienne 2D utilisée pour l'initialisation des noyaux w .

4 Analyse des réseaux

4.1 Réseaux à deux couches pour une opération primaire

4.1.1 Motivations

Les réseaux morphologiques sont construits sous l'architecture telle que décrite et présentée dans la partie état de l'art, figure 5, avec n couches morphologiques du type considéré (p Conv, \mathcal{LM} orph, \mathcal{SM} orph/Tanh), disposées successivement entre une standardisation au début du réseau et une couche de convolution de taille $1 \times 1 \times 1$ en fin du réseau. En l'occurrence, pour les réseaux \mathcal{SM} orphNet (resp. \mathcal{SM} orphNetTanh) que l'on considère ici, le réseau possèdera n couches \mathcal{SM} orph (resp. \mathcal{SM} orphTanh).

Comme vu dans les parties précédentes, chaque couche morphologique, par sa définition, ne pourra avoir l'effet que d'une (pseudo-)opération morphologique *primaire*, à savoir soit celle d'une *érosion* soit celle d'une *dilatation*, définie en fonction de la valeur du paramètre de contrôle α (ou p) de la couche. En particulier, si α (ou p) tend vers $-\infty$, alors l'effet de la couche tendra vers celle d'une érosion exacte, et si α (ou p) tend vers $+\infty$, alors son effet tendra vers celle d'une dilatation exacte.

De ce fait, pour que le réseau en lui-même puisse avoir l'effet, sur les images en entrée, d'une opération primaire, une seule et unique couche morphologique suffit au sein de ce réseau. Pour qu'il ait l'effet d'une ouverture ou d'une fermeture, qui sont tous deux la composition de deux opérations primaires successives, on munira naturellement le réseau de deux couches morphologiques, en espérant que chacune tende soit vers l'érosion exacte soit vers la dilatation exacte.

Vient alors une question : que se passe-t-il si l'on munit le réseau de plusieurs couches morphologiques successives, deux par exemple, pour une opération cible *primaire* ?

On souhaiterait ici comprendre comment le réseau se comporterait dans une telle configuration. On aimerait savoir s'il arriverait à trouver une décomposition exacte de l'élément (ou fonction) structurant cible en deux éléments successifs, et s'il arriverait à converger correctement (avec des paramètres de contrôle α/p de grande amplitude, et une valeur de *loss* faible sur les prédictions faites). Par exemple, une décomposition exacte d'une érosion ou d'une dilatation par un élément structurant *carré* de taille 3×3 est la même opération successive par un segment vertical de longueur 3 (et de largeur 1) puis par un segment horizontal de même taille, ou bien inversement.

On décomposera cette analyse en deux parties : celle pour l'opération cible d'*érosion* sur un réseau à deux couches, et celle pour l'opération cible de *dilatation* sur un réseau à deux couches. On considérera ici les réseaux \mathcal{SM} orphNet et \mathcal{SM} orphNetTanh.

4.1.2 Avec l'érosion

Commençons d'abord cette analyse avec l'opération cible d'érosion. On munit un réseau SMorphNetTanh classique de deux couches morphologiques, sans partage de poids et sans contrainte (on considère la fonction de perte *loss* classique, c'est-à-dire seulement la MSE entre les images cibles et les prédictions du réseau), et on l'entraîne sur la banque MNIST avec les huit fonctions structurantes cibles de nos études précédentes, et avec toujours l'opération cible d'**érosion**. On ajoute également à ces huit fonctions structurantes la nouvelle structure binaire carrée de taille 3x3, qui nous intéresse dans cette étude afin de voir si le réseau peut par exemple la décomposer en deux segments, un vertical et un horizontal, comme expliqué précédemment.

En lançant l'entraînement des différentes expériences, on remarque que l'ensemble des résultats de convergence se divisent en deux parties : celle dans laquelle les noyaux du réseau ont convergé vers une forme distinguable et harmonieuse avec une valeur de perte *loss* faible (le réseau a « bien » convergé, c'est un « succès »), et celle dans laquelle les noyaux ont convergé vers une forme peu distinguable et peu harmonieuse, avec en général des valeurs de paramètre de contrôle α proches de 0, et avec une valeur de *loss* élevée (le réseau a « mal » convergé, c'est un « échec »). En prenant l'exemple du carré de taille 3x3, on obtient une « mauvaise » convergence, dont l'état final des deux couches est la suivante, illustrée figure 31 :

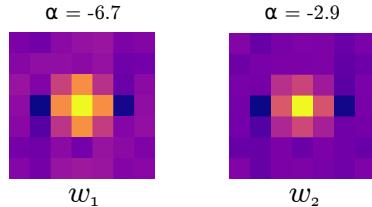


Figure 31 – État final du réseau SMorphNetTanh à deux couches morphologiques, pour la fonction structurante cible carrée 3x3 et l'opération cible d'**érosion** sur la banque d'images d'entraînement MNIST, sans contrainte. Résultat : échec.

On remarque que, pour cet exemple, le réseau a mal convergé, car la forme des noyaux est peu reconnaissable, les paramètres de contrôle α sont de très faible amplitude, et on obtient une valeur de perte sur les prédictions élevée, de l'ordre de 10^{-3} .

On peut cependant, pour cette structure carrée 3x3, tenter de la forcer à mieux converger, en lui imposant une contrainte dans la fonction de perte *loss*. Dans notre cas, on prendra la contrainte sur la norme L_1 , avec $l_c = 0$ (formule 25). La convergence devient alors bien meilleure : on obtient l'état final des deux couches suivant, illustré fig. 32.

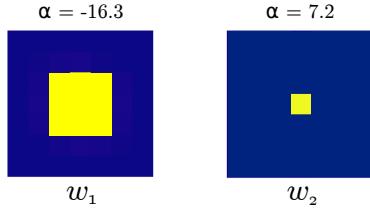


Figure 32 – État final du réseau SMorphNetTanh à deux couches morphologiques, pour la fonction structurante cible carrée 3x3 et l’opération cible d’**érosion** sur la banque MNIST, avec une contrainte des noyaux sur la norme L_1 . Résultat : succès.

On remarque que, pour cet exemple encore, le réseau a, cette fois-ci, bien convergé, grâce à l’ajout de cette contrainte des noyaux w sur la norme l_1 : la forme des noyaux est très reconnaissable (le premier noyau a la forme de la structure cible i.e. le carré 3x3, et le second a la forme de l’identité i.e. un unique pixel distinguable au centre du support), les paramètres de contrôle α ont une plus grande amplitude, et on obtient une valeur de perte sur les prédictions du réseau bien plus faible, de l’ordre de 10^{-6} .

Parmi les deux groupes « succès » et « échecs » des résultats de convergence que l’on a distingués sur l’ensemble des expériences, en examinant uniquement les *succès*, on peut remarquer que l’amplitude des paramètres α n’est jamais très élevée (de l’ordre de 10), et aussi, mais surtout, que l’on obtient *toujours* la configuration des couches suivantes : la première couche a un noyau w_1 de la forme de la fonction structurante cible, avec un α associé négatif, et la seconde couche a un noyau w_2 de la forme de l’identité (un seul pixel actif), avec un α associé positif. On n’obtient jamais de composition originale ou plus complexe sur ces succès, même avec l’aide de contraintes pour favoriser certaines configurations, et on a toujours le même ordre : d’abord la structure cible, puis l’identité. C’est le cas par exemple du carré 3x3, aidé de la contrainte L_1 , telle qu’illustrée figure 32 (d’abord carré 3x3, puis identité).

4.1.3 Avec la dilatation

On reprend ici, pour l’opération cible de dilatation, le même protocole d’expérimentation que pour l’érosion : on considère à nouveau un réseau SMorphNetTanh qu’on munit de deux couches morphologiques, et on fixe ici l’opération cible à celle de la dilatation. On lance les différents entraînements du réseau avec l’ensemble des huit fonctions structurantes cibles considérées, avec en plus le carré binaire de taille 3x3, et ce toujours sur la banque MNIST. Tout comme pour l’érosion, on peut distinguer ici pour la dilatation deux groupes de résultats : les résultats à « succès » de convergence, et ceux à « échec » de convergence, tels que définis précédemment.

Pour le carré 3x3 par exemple, à nouveau contraint par la norme L_1 (avec $l_c = 0$) pour aider le réseau à converger correctement, on obtient l'état des deux couches morphologiques du réseau suivant, illustré figure 33 :

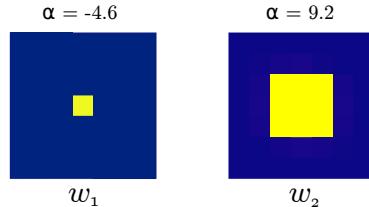


Figure 33 – État final du réseau SMorphNetTanh à deux couches morphologiques, pour la fonction structurante cible carrée 3x3 et l’opération cible de **dilatation** sur la banque MNIST, avec une contrainte des noyaux sur la norme L_1 . Résultat : succès.

On peut à nouveau constater ici, dans le groupe des succès, l’uniformisation des solutions trouvées dans la décomposition de la fonction structurante cible par le réseau à deux couches. Sauf que, pour le cas de la dilatation, l’ordre entre structure cible et identité est inversée par rapport à celle de l’érosion : on a d’abord l’identité sur le noyau w_1 de la première couche, avec un α négatif, puis on a la forme de la structure cible sur le noyau w_2 de la seconde couche, avec un α positif. Pour l’exemple du carré 3x3, on a ainsi d’abord l’identité sur w_1 , puis le carré 3x3 sur w_2 (fig. 33). On remarque à nouveau que, d’une manière générale, l’amplitude des α reste faible.

4.1.4 Discussion

Dans le cas des expériences à succès, le réseau converge donc toujours vers la structure cible pour w_1 et l’identité pour w_2 avec l’érosion, et, à l’inverse, vers l’identité pour w_1 et la structure cible pour w_2 avec la *dilatation*. Ce constat, surprenant à priori, reste à ce jour inexpliqué. Il s’agit d’une configuration très solide, puisqu’il est difficile d’obtenir l’ordre inverse respectif des rôles entre les deux couches, malgré l’ajout de contraintes forçant cette inversion des rôles. Bien qu’elle soit théoriquement tout aussi valide que l’ordre initial, la seule manière qui a fonctionné dans notre étude pour inverser les rôles des deux couches, c’est avec une initialisation de ces dernières dans la configuration visée (forme des deux noyaux w avec l’ordre inversé correspondant, et paramètres α du bon signe et légèrement éloignés de 0, de l’ordre de 0.1).

L’ordre systématique entre structure cible et identité, selon l’érosion ou la dilatation, doit être lié aux propriétés de transformation de ces deux opérations sur les images d’entrée, et à l’ordre d’exécution des deux couches du réseau. La rétropropagation du gradient doit également jouer un rôle. Mais la cause de ce résultat reste inconnu.

4.2 Impact du jeu de données d'entraînement

4.2.1 Description des banques MNIST et FashionMNIST

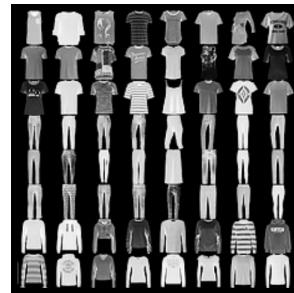
Dans l'ensemble de ce rapport, jusqu'à maintenant, nous considérons tout particulièrement la banque d'images MNIST en tant que jeu de données d'entraînement des réseaux étudiés. Il s'agit d'un jeu composé de 70000 images digitales en niveaux de gris, échelonnées entre 0 et 1 et de taille 28x28 pixels, représentant des digits (chiffres) écrits à la main [12]. Les objets sur les images (ensemble des pixels de valeurs strictement positives, qui se détache visuellement du fond de l'image, dont les valeurs sont proches de 0) de ce jeu de données sont particulièrement fins, et peuvent être considérés topologiquement comme des filaments entrelacés. Cet aspect-là rend ce jeu très spécifique dans le cadre d'opérations morphologiques, et peut donner sur les réseaux morphologiques étudiés des résultats bien particuliers.

L'ensemble des expériences citées durant ce rapport a également été réalisé, en parallèle de MNIST, avec un autre jeu de données : FashionMNIST. Il s'agit d'une banque de 70000 images digitales en niveaux de gris, échelonnées entre 0 et 1 et de taille 28x28 pixels également, mais qui représente cette fois-ci des vêtements de tout type (pantalons, chemises, robes, etc.). Contrairement à MNIST, ce jeu de données représente donc des objets bien plus larges et imposants face au support des images, et les résultats d'entraînement des réseaux morphologiques avec un tel jeu peuvent être bien différents de ceux avec MNIST.

La figure suivante (34) permet de visualiser une partie des images des deux banques MNIST (34(a)) et FashionMNIST (34(b)).

3	8	6	9	6	4	5	3
1	5	0	5	9	7	4	1
1	3	6	8	0	7	7	6
8	4	4	1	2	9	8	1
7	2	7	3	1	4	0	5
4	0	6	1	9	2	1	3
2	8	6	9	7	0	9	1
8	6	8	7	8	8	6	9

(a) MNIST



(b) FashionMNIST

Figure 34 – Echantillon des jeux de données MNIST (34(a)) et FashionMNIST (34(b)).

Le but est alors de comparer les résultats de convergence d'un même réseau pour les mêmes expériences entre ces deux jeux de données MNIST et FashionMNIST, et ainsi mieux comprendre leur impact sur la convergence des réseaux morphologiques.

4.2.2 Différences de résultats entre les jeux

On souhaite comparer les résultats de convergence de réseaux entraînés sur MNIST avec ceux entraînés sur FashionMNIST pour les mêmes expériences, dans le but d'en tirer des caractéristiques et une différence globale de comportement des réseaux liés à ces deux jeux de données. Prenons par exemple le réseau *SMorphNetTanh* composé de deux couches morphologiques, et que l'on munit simplement d'un partage de poids doux (contrainte entre w_1 et w_2 dans la *loss*, eq. 19). On fait alors une analyse rapide en constatant uniquement la forme des noyaux du réseau dans son état final, moyennée sur 6 runs par expérience. On obtient les résultats représentés sur la figure 35, avec MNIST à gauche (35(a)) et FashionMNIST à droite (35(b)).

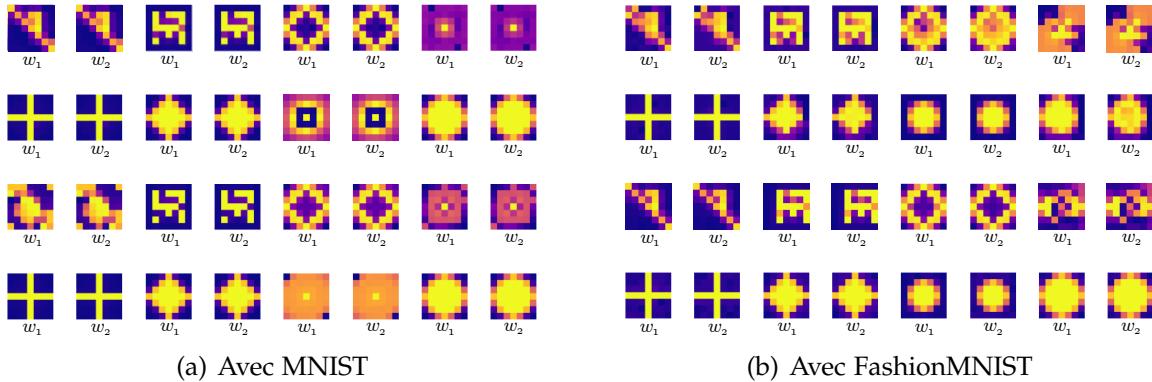


Figure 35 – Noyaux w associés par paires du réseau SMorphNetTanh après convergence, entraîné sur les jeux de données MNIST (35(a)) et FashionMNIST (35(b)). Les cases à gauche et à droite de même emplacement correspondent à la même expérience.

Les différences de résultats des noyaux w pour ces expériences-là entre les deux jeux MNIST et FashionMNIST sont très notables. Certains résultats semblent meilleurs avec la banque FashionMNIST, comme ceux avec les structures cibles *adiag* pour l'ouverture (troisième ligne [milieu bas] première paire [tout à gauche]) ou *disk2* (deuxième ligne troisième paire, et dernière ligne troisième paire), tandis que d'autres semblent meilleurs avec FashionMNIST, comme ceux avec les structures cibles *brand* (première ligne deuxième paire, et troisième ligne deuxième paire) ou *complex* pour la fermeture (première ligne troisième paire). Les résultats sont très diversifiés, et ne plaident pas significativement en la faveur d'un des deux jeux de données.

Les mêmes observations et la même conclusion peuvent être faites avec les autres réseaux morphologiques (*pConvNet*, *LMorphNet* et *SMorphNet*) et dans les autres configurations de réseaux (avec/sans partage de poids, avec/sans contrainte géométrique sur w ou sur α/p , initialisation aléatoire / initialisation gaussienne, etc.).

Plusieurs hypothèses peuvent expliquer ces différences de résultats entre les deux jeux de données. En particulier, l'aspect « fin » des objets sur les images MNIST doit impacter l'entraînement des réseaux de manière différente que ne le font les aspects « épais » et « étendu » des objets sur les images FashionMNIST. Les objets fins sur les images peuvent par exemple s'effacer intégralement avec une petite opération d'*érosion*, et donc la prédiction sur l'érosion serait moins facile à faire pour les réseaux morphologiques ciblant cette opération ou bien ciblant l'ouverture (qui commence par une érosion, et donc peut directement effacer les objets avant même la dilatation).

C'est en effet ce que l'on observe sur la figure 35 : on constate avec MNIST (35(a)) que ce sont les résultats sur la fermeture (moitié supérieure) qui semblent plus précis que ceux sur l'ouverture (moitié inférieure), notamment en observant les structures *adiag* et *complex*. Avec FashionMNIST (35(b)), on constate l'inverse : ce sont les résultats sur l'ouverture (dessous) qui semblent plus précis que ceux sur la fermeture (dessus). Cela peut s'expliquer par la trop petite place laissée au fond de l'image sur son support, à cause de l'épaisseur des objets, et donc par le fait qu'une dilatation (et donc une fermeture) supprimerait une trop grande partie du fond, ce qui rendrait les prédictions sur la dilatation et la fermeture avec FashionMNIST plus délicates.

4.2.3 Résultats avec les jeux en négatif

On considère ici les images des banques MNIST et FashionMNIST en *négatif*, c'est-à-dire qu'on leur applique : $f \mapsto 1 - f$. Ainsi, les propriétés de chacune de ces deux banques d'images sont inversées : les images de MNIST en négatif comportent des objets (parties du support dont la valeur est proche de 1) très épais qui s'étalent le long des quatre bords du support et occupent une grande partie de l'espace sur ce dernier, englobant un fond (parties du support de valeur proche de 0) filamentaire fin et centré sur le support, qui disparaît facilement avec une dilatation ; tandis que les images de FashionMNIST en négatif comportent plutôt des objets qui s'étalent également le long des bords, mais qui, à l'inverse de MNIST, sont fins et disparaissent facilement avec une érosion, et qui entourent un fond épais centré sur le support.

En relançant les mêmes entraînements que précédemment sur *SMorphNetTanh* à deux couches morphologiques muni d'un partage de poids doux, mais avec cette fois-ci les deux jeux de données MNIST et FashionMNIST en négatif, on obtient les conclusions inverses par rapport à celles liées aux images en positif : sur MNIST, les résultats de convergence du réseau sont ici globalement meilleurs avec l'opération cible d'ouverture qu'avec la fermeture, et sur FashionMNIST, à l'inverse, les résultats sont globalement meilleurs pour la fermeture que pour l'ouverture (ce dernier cas présentant les mêmes défauts et irrégularités dans la forme des noyaux w que pour les résultats de fermeture avec les images en positif). Ce qui est cohérent avec les caractéristiques inversées des images négatives par rapport aux images positives.

4.3 Relation entre complexité et échec de convergence

4.3.1 Complexité de la forme de la fonction structurante cible

Dans leur article, Hermary et al. [7], pour tenter de donner une explication aux différents échecs de convergence de leurs expériences, émettaient l'hypothèse que le fait qu'un élément (ou fonction) structurant « touche » le bord de son support (i.e. le fait qu'il existe un pixel au bord du support de valeur strictement supérieure à 0 pouvant être associée à l'objet de l'image) aide le réseau, lors de son entraînement, à converger de manière plus précise et certaine vers l'élément structurant cible.

En reproduisant les différentes expériences de leur article, avec notamment leurs six fonctions structurantes cibles telles que décrites dans les grandes parties précédentes (*cross3*, *cross7*, *disk2*, *disk3*, *diamond3* et *complex*), on peut effectivement constater que les cas d'échec interviennent uniquement sur les structures dont aucune partie ne touche le bord de leur support (*cross3* et *disk2*), là où les structures touchant le bord de leur support donnent systématiquement lieu à un succès de convergence (*cross7*, *disk3*, *diamond3* et *complex*), du moins pour ces fonctions structurantes-là (fig. 8 et 9).

Une explication donnée est le fait que, lors de la rétropropagation du gradient pendant l'apprentissage du réseau, ce sont les poids des bords du support des noyaux w qui semblent se mettre à jour en premier, ou du moins qui convergent plus rapidement. Par exemple, sur la figure 25 montrant l'évolution des deux noyaux du réseau pour la fonction structurante cible *adiag*, avec ou sans contrainte C_{awayOPP} , on voit lors des premières époques que ce sont effectivement les poids au bord du support qui semblent converger en premier vers leur état final : les coins supérieur gauche et inférieur droit, par lesquels passe la diagonale, prennent en premier les plus grandes valeurs, et les coins supérieur droit et inférieur gauche prennent de plus petites valeurs, là où le centre du support reste creux et surtout très homogène. La forme de la structure semble se dessiner en partant des bords jusqu'au centre du support.

Cependant, avec les deux autres fonctions structurantes créées en plus (*adiag* et *brand*), on remarque que cette propriété n'est ici pas respectée : malgré le fait que *adiag* touche largement les bords d'une part et d'autre de son support, les réseaux *SMorphNet* et *SMorphNetTanh*, sans contrainte, ont tous deux du mal à converger vers cette forme diagonale, pour laquelle les métriques de performances des réseaux sont moyennes voire mauvaises (de l'ordre de 10^{-3} pour la *loss*). À l'inverse, pour la fonction structurante *brand*, les deux réseaux arrivent très bien à trouver la forme de la structure cible, malgré le fait que cette dernière ne touche pas le bord (fig. 13 et 14).

Des expériences complémentaires ont été réalisées en ajoutant, dans *brand*, un unique pixel au bord du support. Et les résultats de convergence des réseaux avec cette structure cible-là, qui touche donc le bord, sont même légèrement moins bons que ceux avec la structure *brand* originelle, comme le montre la figure 36 ci-après.

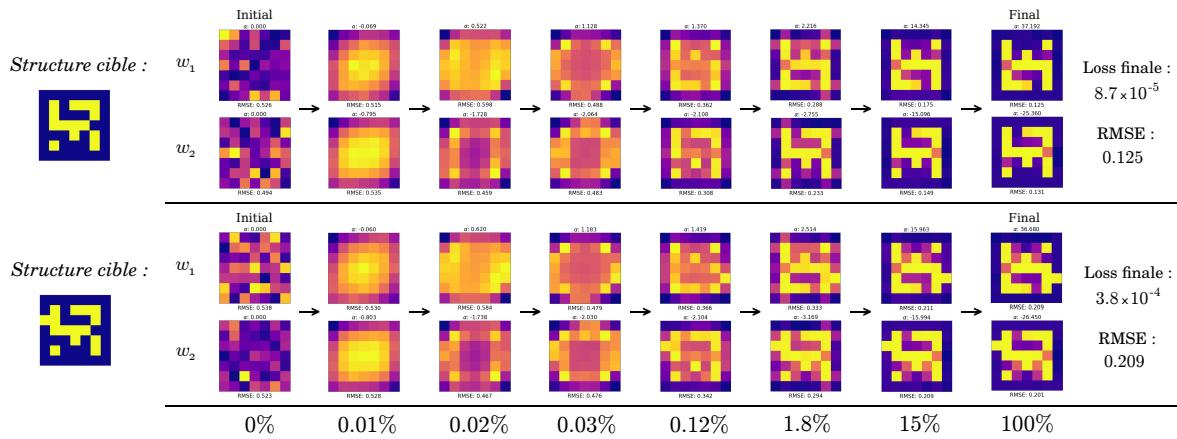


Figure 36 – Évolution de la forme des deux noyaux du réseau et de leur RMSE en fonction de la progression de l’entraînement (en % avant l’état final), pour la fermeture sur MNIST, avec le *brand* (haut) et le *brand* avec un pixel en plus au bord (bas).

L’hypothèse que la structure cible touche (resp. ne touche pas) le bord de son support puisse déterminer le succès (resp. l’échec) de convergence du réseau ne peut donc pas, à elle seule, expliquer l’ensemble des échecs de convergence. En revanche, en exploitant les résultats de convergence de diverses expériences, avec davantage de formes différentes de fonctions structurantes cibles, il semble se dessiner un lien - une correlation positive - entre une certaine « complexité » de la forme de la structure cible et la réussite du réseau dans sa convergence. En effet, plus la forme de la structure cible semble « complexe », plus le réseau semble converger de manière précise et juste, avec des noyaux w de forme davantage proche de celle de la structure cible.

Cette notion de « complexité » des fonctions structurantes cibles reste très subjective et mal définie. Elle prend en compte à la fois la *taille* de la structure par rapport à son support (si elle est plus petite, alors les configurations à deux couches où elle est symétriquement translatée sont strictement équivalentes), la multiplicité des *nuances de gris* dont l’élément est composé, et les *caractéristiques topologiques* de cette structure (rotundité, connexité, convexité, extrémités...). Ce premier aspect (élément plus petit que le support) peut expliquer les échecs des cas où l’élément «ne touche pas le bord».

Cette notion de « complexité » d’une structure morphologique serait inversement corrélée au nombre de décompositions exactes possibles de cet élément dans le cadre d’une érosion ou d’une dilatation. Par exemple, le carré 3x3 peut être considéré comme « peu complexe », car il existe une multitude de décompositions de cette structure, par exemple celle en deux segments successifs de longueur 3 (et de largeur 1), l’un horizontal et l’autre vertical, dans le cadre d’une érosion ou d’une dilatation. C’est une explication fort probable des nombreux échecs de convergence des réseaux à deux couches avec comme structure cible le carré 3x3 sur un support de taille 7x7 (en plus de la taille du support qui multiplie les configurations possibles).

Une explication possible à cette corrélation positive entre complexité de la forme de la structure cible et succès de convergence des réseaux, est le fait que plus la structure est complexe, plus elle est « unique » : il existerait moins de formes différentes possibles « proches » morphologiquement parlant de cette structure, et moins de décompositions possibles de cet élément en plusieurs sous-éléments distincts (dans le cadre de réseaux multi-couches). Cette diminution du nombre de formes théoriques possibles « proches » de la structure visée équivaut à une diminution du nombre de minima locaux possibles dans la fonction de perte *loss* qui sont « proches » du minimum global associé à la structure cible. Ce qui ferait ainsi que les noyaux w du réseau arrivent mieux à converger vers cette structure cible, et avec plus de précision et de rapidité, car il y aurait moins de minima locaux proches de ce minimum global visé. À l'inverse, une structure cible trop « élémentaire » impliquerait une multiplicité des solutions locales possibles (i.e. minima locaux dans la *loss*), et donc amènerait davantage le réseau à converger vers l'un de ces minima locaux (un mauvais résultat).

Ces explications restent néanmoins au stade d'hypothèses, qui doivent être vérifiées et validées. La notion de « complexité » d'une forme reste également très subjective, et une définition précise de ce terme doit être formulée. Son rôle dans les succès ou les échecs de convergence des réseaux pourrait alors être évalué et quantifié.

4.3.2 Complexité de la configuration du réseau

En plus de la corrélation positive établie entre « complexité » de la fonction structurante cible et succès de convergence du réseau, il semble y avoir également corrélation, mais cette fois-ci négative, entre ces succès et une certaine « complexité » dans la *configuration* du réseau : son architecture, le nombre de couches morphologiques, la taille des noyaux des couches, ou plus généralement l'ensemble des hyperparamètres du réseau. Cette notion de « complexité » de la configuration d'un réseau peut se traduire en particulier par la taille et la densité de ce dernier : un réseau morphologique devient davantage « complexe » si la *taille du support* de ses noyaux w augmente ou si le *nombre de couches* morphologiques successives qu'il contient augmente, donc de manière générale si le nombre de poids apprenables par le réseau augmente.

Cette corrélation peut s'expliquer simplement : plus le réseau est grand et dense, plus il existe de configurations possibles du réseau, et donc plus il y aura de minima locaux dans la fonction de perte *loss* vers lesquels le réseau risque de converger. Les configurations possibles du réseau pour des caractéristiques ciblées (comme la forme des noyaux w) étant moins nombreuses avec des petits réseaux peu complexes qu'avec de grands réseaux denses. Mais ces explications restent encore une fois des hypothèses, qui doivent être vérifiées. Il doit encore exister une multitude de caractéristiques sur les réseaux ou sur les éléments cibles pouvant avoir un rôle concret dans les succès ou les échecs de convergence des réseaux, car ce problème reste sous-déterminé.

4.4 Autres rapides analyses

4.4.1 Importance des α dans la convergence des réseaux

Lors des différentes expériences effectuées, notamment sur les cas d'échec de convergence, une observation à priori étonnante a pu être faite sur l'impact de la valeur des paramètres de contrôle α (ou p), notamment celui de leur signe, sur la direction dans laquelle converge le réseau lors de son entraînement. Les couches morphologiques, en particulier *SMorph* et *SMorphTanh*, se comportent en effet comme des moyenneurs quand leur paramètre de contrôle α est proche de 0, peu importe son signe. On pouvait alors espérer que la différence de comportement d'un réseau, dont le paramètre α d'une couche est proche de 0, serait minime voire négligeable entre un α positif et un α négatif. Or, plusieurs expériences permettent de montrer que le signe des paramètres α , même très proches de 0, a un impact majeur, voire plus important que la forme des noyaux w , sur la direction de convergence des réseaux.

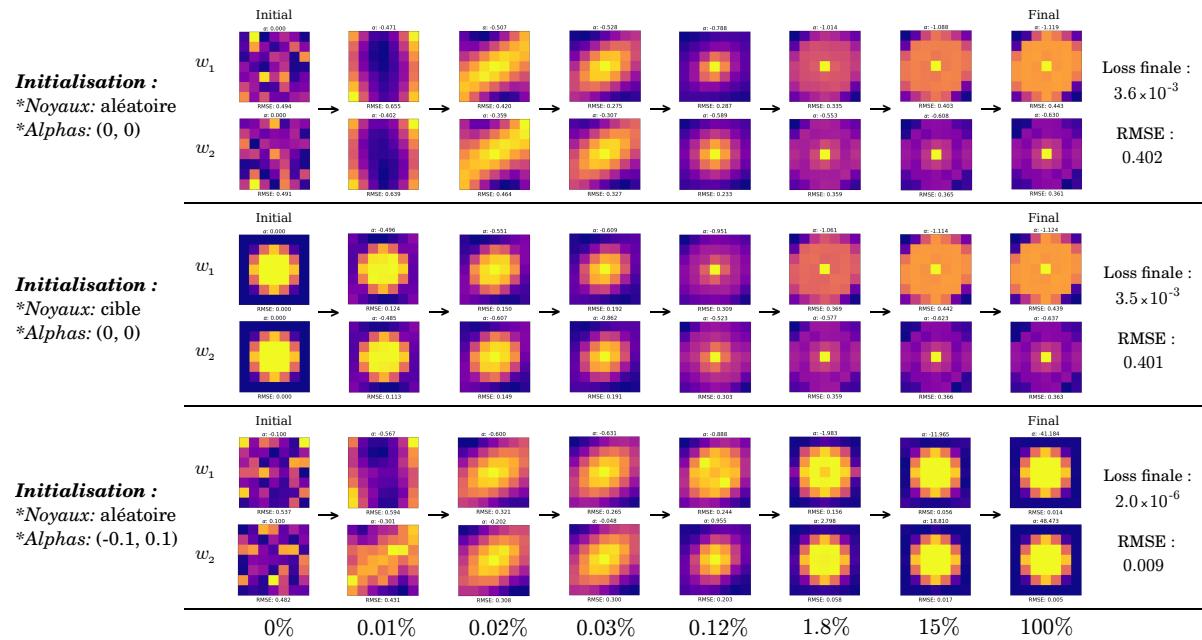


Figure 37 – Évolution de la forme des deux noyaux w du réseau *SMorphNetTanh* et de leur RMSE en fonction de la progression de l'entraînement (en % avant l'état final), pour l'ouverture sur MNIST avec le *disk2*, sur trois types d'initialisation.

La figure 37 ci-dessus illustre le cas du réseau *SMorphNetTanh* à deux couches morphologiques - muni d'aucune contrainte sur ces dernières ni d'aucun partage de poids, dont l'opération cible est l'ouverture et la structure cible est le *disk2*, et entraîné sur MNIST -, qui converge initialement vers un échec (tiers supérieur de la figure).

Ce même réseau, en initialisant ses deux noyaux par la structure cible exacte recherchée, converge exactement vers le même état d'échec (tiers central). Cependant, ce même réseau encore, avec cette fois uniquement une initialisation des paramètres de contrôle α de très faible amplitude ($|\alpha| = 0.1$) mais du signe recherché (négatif pour α_1 et positif pour α_2 , comme on est dans le cadre d'une opération cible d'ouverture), converge effectivement bien vers l'état cible de succès (tiers inférieur).

Cela montre que le déplacement des paramètres de contrôle α , même de très faible amplitude, vers leur signe cible respectif, a un impact bien plus important sur la réussite de la convergence que ne l'a le déplacement des poids des noyaux vers la forme exacte de la structure cible. Cette même conclusion peut être faite sur d'autres expériences encore qui sont initialement des échecs, et dont le signe des α dans leur état final n'est pas le bon recherché. Une bonne valeur (à l'initialisation ou au cours de l'entraînement) des paramètres de contrôle α , et en particulier le bon signe de ces paramètres, a donc une importance majeure, probablement supérieure à celle de la forme des noyaux w , dans le succès de la convergence des réseaux.

4.4.2 Application au débruitage

Dans l'ensemble des travaux entrepris dans le cadre de ce stage, nous nous sommes intéressés en particulier à l'aspect théorique des couches morphologiques et des réseaux associées, et à l'amélioration et l'analyse comportementales de ces réseaux de neurones, sans vraiment nous pencher sur leur côté applicatif. Cette dernière partie est donc dédiée à une analyse rapide des résultats de l'une des applications pratiques possibles de tels réseaux morphologiques. Ces applications sont nombreuses, en particulier en traitement d'images. L'une d'elles est classiquement le débruitage.

Dans le cadre de l'application pratique classique de débruitage poivre-et-sel d'images par un réseau de neurones morphologique, on considère un réseau *SMorphNetTanh*, que l'on munit de quatre couches morphologiques *SMorphTanh*. Cela afin que le réseau puisse simuler successivement les opérations d'ouverture et de fermeture, ou inversement : il faut donc deux couches simulant une érosion, et deux autres simulant une dilatation, toutes quatre disposées en rimes embrassées. On entraîne alors le réseau construit avec uniquement un ensemble d'images en entrée du réseau (input), sur lesquelles on aura ajouté au préalable le bruit poivre et sel (avec une certaine proportion surfacique d'occupation du bruit), ainsi qu'avec l'ensemble des images cibles correspondantes visées par le réseau, qui sont les images d'origine avant l'ajout manuel du bruit. Il n'y a donc ici pas d'élément structurant cible, ni vraiment d'opération morphologique cible, bien qu'on espère que le réseau simule la succession d'une ouverture et d'une fermeture. On prendra ici le jeu d'entraînement FashionMNIST.

Un exemple des résultats de débruitage du réseau entraîné est présenté figure 38.

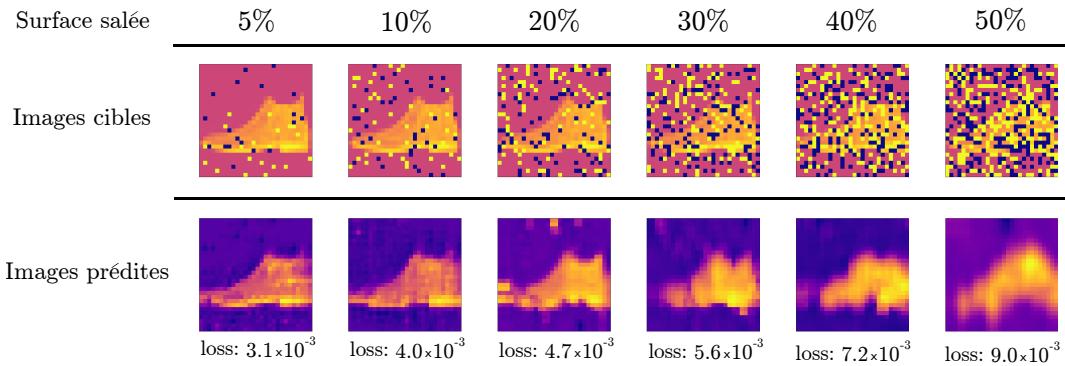


Figure 38 – Résultats de débruitage d’une image représentant une chaussure par le réseau SMorphNetTanh à quatre couches entraîné sur FashionMNIST, avec plusieurs niveaux de bruit poivre et sel (en proportion de surface occupée).

Malgré des valeurs de *loss* plutôt élevées de l’ordre de 10^{-3} à 10^{-2} , les résultats visuels de débruitage du réseau semblent plutôt bons à l’oeil nu : le réseau retrouve toujours globalement la forme de la chaussure, même à très haute proportion de bruit, et les pixels poivre et sel disparaissent très bien de l’image bruitée, même à 50%. Une opération morphologique à la main ne pourrait faire que légèrement mieux.

Cependant, bien que les résultats de débruitage semblent « plutôt bons », l’état du réseau l’est bien moins, car la forme des noyaux w est très aléatoire et manque de régularité, et surtout, les paramètres de contrôle α sont toujours très proches de 0.

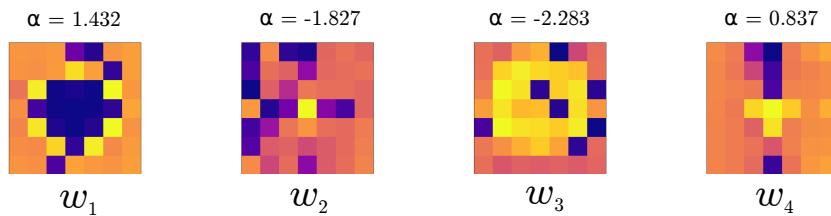


Figure 39 – Noyaux w et paramètres de contrôle α des quatre couches du réseau SMorphNetTanh après entraînement sur FashionMNIST bruitée à 10%.

La figure 39 ci-dessus présente l’état des quatre couches du réseau après un entraînement sur les images de la banque FashionMNIST bruitées à 10%. D’après l’ordre des signes des α , on peut en déduire que le réseau a le comportement d’une pseudo-fermeture suivie d’une pseudo-ouverture, ce que l’on visait initialement. Mais les α restent très proches de 0. On remarque sur les autres proportions de bruitage que plus cette proportion est élevée, plus les α sont proches de 0, et donc plus les couches ont le comportement d’un moyenneur, ce qui peut être observé sur la figure 38.

5 Conclusion

Les travaux menés dans le cadre de ce projet « Réseaux de neurones morphologiques » sont multiples et variés. Ils ont abouti à différents résultats pour l'amélioration de la convergence de ces réseaux, ainsi qu'à des analyses de leurs propriétés sur diverses aspects. Dans la continuité des constats de l'état de l'art, notamment ceux de Kirszenberg et al. [9] et de Hermary et al. [7], c'est le réseau *SMorphNet*, plus efficace que les autres de la littérature, qui a été en particulier étudié pour ce projet.

Ce réseau *SMorphNet* converge, lors de son entraînement, avec une haute précision et une grande rapidité, notamment dans le cadre d'opérations cibles élémentaires (érosion ou dilatation) avec une unique couche morphologique. Mais il fait cependant face à de nombreux échecs de convergence, en particulier pour les opérations plus complexes d'ouverture et de fermeture avec deux couches morphologiques au sein du réseau. Plusieurs pistes ont alors été explorées dans le but d'améliorer la convergence de ce réseau morphologique, et de pallier notamment ces échecs-là de convergence.

Quatre pistes principales se distinguent dans ces travaux :

- La définition d'une nouvelle couche morphologique *SMorphTanh*, inspirée directement de la couche *SMorph*. Ses performances au sein d'un réseau semblent équivalentes à celles de *SMorph*, mais la formulation de sa fonction de transformation permet néanmoins de résoudre le problème de signe devant w dans l'expression asymptotique de *SMorph* lorsque $\alpha \rightarrow -\infty$;
- La mise en place de deux méthodes différentes de partage de poids entre les noyaux w de deux couches morphologiques, ne servant pas à une amélioration nette de la convergence des réseaux, mais qui permet à ces deux noyaux-là de converger vers la même forme de structure dans le cadre d'une opération cible d'ouverture ou de fermeture pour des réseaux bi-couches ;
- Le développement de différentes métriques de contraintes sur les noyaux w et sur les paramètres de contrôle α des couches morphologiques, permettant une meilleure convergence de ces dernières lors de l'entraînement du réseau, sur la base d'informations géométriques connues, à priori, à propos de caractéristiques visées de ces noyaux et paramètres de contrôle ;
- La modulation de l'initialisation des noyaux w , passant d'une initialisation aléatoire de loi normale centrée réduite à une initialisation gaussienne 2D, permettant de mener la convergence du réseau à un succès pour certaines fonctions structurantes cibles seulement, telles la *cross3*, mais n'apportant pas d'amélioration nette sur l'ensemble des structures cibles.

Aucune solution plus générale et universelle n'a été trouvée ou développée, mais davantage des solutions locales s'inscrivant dans des tâches cibles spécifiques.

Les analyses de ces réseaux morphologiques, quant à elles, ont pu déboucher sur plusieurs conclusions et davantage d'hypothèses et de questions. L'étude de tels réseaux est en effet complexe, et la plupart des observations ne sont pas encore expliquées. À la suite de ces travaux et de l'exploration des réseaux morphologiques, plus de questions se sont ouvertes que ne se sont résolues, et la majorité reste en suspens.

Ces analyses ont néanmoins permis de mieux comprendre certains aspects du comportement des réseaux morphologiques. En particulier, elles ont permis d'établir une corrélation entre la complexité des structures et les échecs de convergence des réseaux, de mieux comprendre l'impact du jeu de données d'entraînement sur l'état final des poids des couches, et de constater la direction de convergence forcée d'un réseau bi-couches pour une opération cible primaire, qui est strictement symétrique entre dilatation et érosion. Elles ont également permis de mieux comprendre l'impact des paramètres de contrôle α , et en particulier leur signe, dans la direction de convergence des couches morphologiques sur les cas d'échec, probablement plus importante que la forme des noyaux w . Une application rapide à du débruitage automatique poivre & sel a enfin été étudiée, et a permis de montrer une certaine efficacité visuelle des réseaux sur les résultats de prédiction, malgré un état des poids peu satisfaisant.

Bien d'autres choses encore, d'autres méthodes et d'autres procédés, ont été implémentés et testés sur ces réseaux morphologiques, mais la plupart n'a pas donné les résultats espérés. Par exemple, une modulation de la définition de la fonction de perte *loss* évoluant au cours du temps et des époques d'entraînement a été pensée et testée : il s'agissait de rendre *dynamiques* (i.e. qui évoluent au cours des époques) les pondérations λ des métriques de contrainte dans la fonction *loss*. Cette évolution peut se faire de manière continue linéaire, jusqu'à ce que la valeur de toutes les pondérations λ atteigne 0 ; de manière discrète par étapes, avec une valeur des pondérations λ qui décrémente à chaque époque d'une valeur fixe ; ou encore de manière cyclique, par exemple sinusoïdale cardinale, afin de rappeler au réseau les contraintes attendues toutes les n périodes. Ce projet garde encore beaucoup de questions ouvertes et de méthodes à tester et évaluer. Une suite à ces travaux serait souhaitable.

Dans un travail futur, une solution possible pour pallier le problème restant de symétrie des noyaux w dans la formulation asymptotique de la couche SMorphTanh, lorsque $\alpha \rightarrow -\infty$, serait de remplacer, dans la formule de SMorphTanh en question, les « $w(x - y)$ » par :

$$w(x - y) \frac{1 + \tanh(\alpha)}{2} + w(y - x) \frac{1 - \tanh(\alpha)}{2}$$

À condition que le support du noyau w soit symétrique (sinon, si $y - x$ n'est pas dans le support, on remplace $w(y - x)$, qui n'est pas définie, par $-\infty$). Elle permettrait également de lisser davantage l'espace de perte. Il serait donc suggérable de l'explorer.

Références

- [1] Jesús ANGULO : Pseudo-morphological image diffusion using the counter-harmonic paradigm. *Advanced Concepts for Intelligent Vision Systems Conference*, pages 426–437, 2010. pages 10
- [2] Isabelle BLOCH, Samy BLUSSEAU, Ramón Pino PÉREZ, Élodie PUYBAREAU et Guillaume TOCHON : On some associations between mathematical morphology and artificial intelligence. *International Joint Conference on Discrete Geometry and Mathematical Morphology*, pages 457–469, 2021. pages 1, 6, 8
- [3] P. S. BULLEN : Handbook of means and their inequalities. *Springer, Berlin*, vol. 560, 2013. pages 10
- [4] Jennifer L. DAVIDSON et Gerhard X. RITTER : Theory of morphological neural networks. *Photonics West - Lasers and Applications in Science and Engineering*, pages 378–388, 1990. pages 6
- [5] Robert M. HARALICK, Stanley R. STERNBERG et Xinhua ZHUANG : Image analysis using mathematical morphology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1987. pages 1, 4, 5
- [6] Mohamad H. HASSOUN, Nathan INTRATOR, Susan MCKAY et Wolfgang CHRISTIAN : Fundamentals of artificial neural networks. *Computers in Physics*, vol. 10:137, 1996. pages 8
- [7] Romain HERMARY, Guillaume TOCHON, Élodie PUYBAREAU, Alexandre KIRSZENBERG et Jesús ANGULO : Learning grayscale mathematical morphology with smooth morphological layers. *Journal of Mathematical Imaging and Vision*, vol. 64:736–753, 2022. pages 1, 2, 7, 9, 11, 12, 13, 14, 15, 19, 20, 57, 63
- [8] HORGAN : Mathematical morphology for analysing soil structure from images. *European Journal of Soil Science*, vol. 49:161–173, 1998. pages 1
- [9] Alexandre KIRSZENBERG, G. TOCHON, É. PUYBAREAU et J. ANGULO : Going beyond p-convolutions to learn grayscale morphological operators. *International Joint Conference on Discrete Geometry and Mathematical Morphology*, pages 470–482, 2021. pages 2, 11, 13, 63
- [10] Mandy LANGE, Dietlind ZÜHLKE, Olaf HOLZ et Thomas VILLMANN : Applications of l_p -norms and their smooth approximations for gradient based learning vector quantization. *The European Symposium on Artificial Neural Networks*, pages 271–276, 2014. pages 12
- [11] Yann LEcun, Yoshua Bengio et Geoffrey E. Hinton : Deep learning. *Nature*, vol. 521:436–444, 2015. pages 6
- [12] Yann LEcun et Corinna Cortes : The mnist database of handwritten digits. <http://yann.lecun.com/exdb/mnist>, 1998. pages 13, 54

- [13] Jonathan MASCI, Jesús ANGULO et Jürgen SCHMIDHUBER : A learning framework for morphological operators using counter-harmonic mean. *arXiv : Computer Vision and Pattern Recognition*, 2012. pages 1, 2, 10
- [14] Keiller NOGUEIRA, Jocelyn CHANUSSOT, Mauro Dalla MURA et Jefersson A. dos SANTOS : An introduction to deep morphological networks. *arXiv :1906.01751*, 2019. pages 2, 8, 13
- [15] Ii. R.A. PETERS : A new algorithm for image noise reduction using mathematical morphology. *IEEE Transactions on Image Processing*, 1995. pages 1
- [16] Swalpa Kumar Roy, Ranjan MONDAL, Mercedes E. PAOLETTI, Juan M. HAUT et Antonio PLAZA : Morphological convolutional neural networks for hyperspectral image classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14:8689–8702, 2021. pages 2
- [17] Stephen S. et WILSON : Morphological networks. *Visual Communications and Image Processing IV*, vol. 1199:483–495, 1989. pages 6
- [18] Jean SERRA : Image analysis and mathematical morphology. *Academic Press, London*, 1983. pages 1, 3
- [19] Jean SERRA : Introduction to mathematical morphology. *Graphical Models graphical Models and Image Processing computer Vision, Graphics, and Image Processing*, 1986. pages 1, 3
- [20] Jean SERRA et Luc VINCENT : An overview of morphological filtering. *Circuits Systems and Signal Processing*, vol. 11:47–108, 1992. pages 1, 4
- [21] Frank Y. SHIH, Yucong SHEN et Xin ZHONG : Development of deep learning framework for mathematical morphology. *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 33:1954024, 2019. pages 2