

Self2Seg : Segmentation et débruitage conjoints autosupervisés d'une seule image

Nadja Gruber^{*,(†),(‡)} J. Schwab[§] N. Debroux[¶]
N. Papadakis^{||} M. Haltmeier[†]

30 avril 2024

Résumé

Nous développons Self2Seg, une méthode auto-supervisée pour la segmentation et le débruitage conjoints d'une seule image. À cette fin, nous combinons les avantages de la segmentation variationnelle avec l'apprentissage profond auto-supervisé. L'un des principaux avantages de notre méthode réside dans le fait que, contrairement aux méthodes basées sur les données, qui nécessitent d'énormes quantités d'échantillons étiquetés, Self2Seg segmente une image en régions significatives sans aucune base de données d'apprentissage. En outre, nous démontrons que le débruitage auto-supervisé lui-même est considérablement amélioré grâce à l'apprentissage spécifique à la région de Self2Seg. Par conséquent, nous introduisons une nouvelle fonction énergétique auto-supervisée dans laquelle le débruitage et la segmentation sont couplés de manière à ce que les deux tâches bénéficient l'une de l'autre. Nous proposons une stratégie d'optimisation unifiée et montrons numériquement que, pour les images microscopiques bruitées, l'approche conjointe que nous proposons est plus performante que sa contrepartie séquentielle et que les méthodes alternatives axées uniquement sur le débruitage ou la segmentation.

1 Introduction

Le débruitage et la segmentation des images sont des problèmes fondamentaux dans le traitement des images [30, 5, 16]. Dans de nombreuses applications biomédicales, telles que la microscopie à fluorescence ou la microscopie électronique à transmission, on s'intéresse tout particulièrement à la segmentation des images.

*Auteur correspondant nadja.gruber@uibk.ac.at

†Département de mathématiques, Université d'Innsbruck, Autriche

‡VASCAGE-Centre de recherche sur le vieillissement vasculaire et les accidents vasculaires cérébraux, Innsbruck, Autriche

§MRC Laboratory of Molecular Biology, Cambridge, Royaume-Uni

¶Institut Pascal, Université Clermont Auvergne, Clermont-Ferrand, France^{||} Institut de Mathématiques de Bordeaux, Bordeaux, France

dans la segmentation des objets. Les données d'entraînement pour cette tâche sont généralement rares et difficiles à obtenir en raison de la complexité intrinsèque et du bruit élevé de ces images, ainsi que du temps nécessaire aux experts pour les étiqueter. Il existe donc un besoin de méthodes non supervisées pour traiter les deux tâches d'imagerie de manière unifiée. Dans ce travail, nous proposons un tel cadre en combinant et en étendant les idées du modèle de segmentation Chan-Verse [9, 14] et de l'apprentissage auto-supervisé [19, 17].

Comme nous le démontrons dans cet article, le débruitage et la segmentation autosupervisés conjoints peuvent être très bénéfiques l'un pour l'autre. En identifiant des régions individuelles, la segmentation guide le processus de débruitage vers des structures d'image spécifiques. D'autre part, l'adaptation d'un débruiteur à une région spécifique guide avec succès le processus de segmentation.

Il existe au moins deux types d'approches principales pour s'attaquer aux deux tâches de manière distincte. La première classe de méthodes implique la minimisation d'une fonction d'énergie dans un cadre graphique ou variationnel. Le second type d'approches, récemment devenu populaire, prend en compte les techniques d'apprentissage profond, en particulier celles basées sur les réseaux neuronaux convolutifs [22].

Avant de passer en revue les travaux connexes qui abordent les tâches de débruitage et de segmentation de manière unifiée, nous présentons la motivation qui sous-tend la stratégie proposée.

1.1 Motivation et contributions

Pour surmonter les limites des algorithmes existants (voir 1.2), nous présentons dans cet article un nouvel algorithme de débruitage et de segmentation combinant les forces des modèles variationnels et de l'apprentissage profond.

Nous utilisons des exemples de jouets pour démontrer comment la segmentation guide efficacement le processus de débruitage. Tout d'abord, nous générons une image 256×256 présentant des motifs de rayures orientés différemment dans diverses régions, comme le montre la figure 1. Un bruit gaussien est ajouté manuellement pour corrompre l'image. Nous utilisons deux réseaux neuronaux linéaires, chacun servant d'"expert" dédié à l'avant-plan ou à l'arrière-plan. Ces réseaux sont constitués d'une seule couche convolutive avec un filtre 15×15 , formé à l'aide d'une version légèrement modifiée de la stratégie Noise2Fast [19], décrite en détail dans la section 4. Plus précisément, nous limitons l'entraînement du réseau aux régions respectives de l'image en masquant la fonction de perte et en restreignant l'entraînement à 30×30 boîtes, comme illustré dans la figure 1. Sans surprise, nous observons que les filtres appris s'adaptent à la structure du signal dans les régions désignées, ce qui conduit à des modèles d'erreur plus élevés dans les zones non entraînées. Cette observation constitue la base de l'exploitation du débruitage spécifique à une région pour la segmentation. D'autres détails expérimentaux pour

de cet exemple sont disponibles à la section 5.

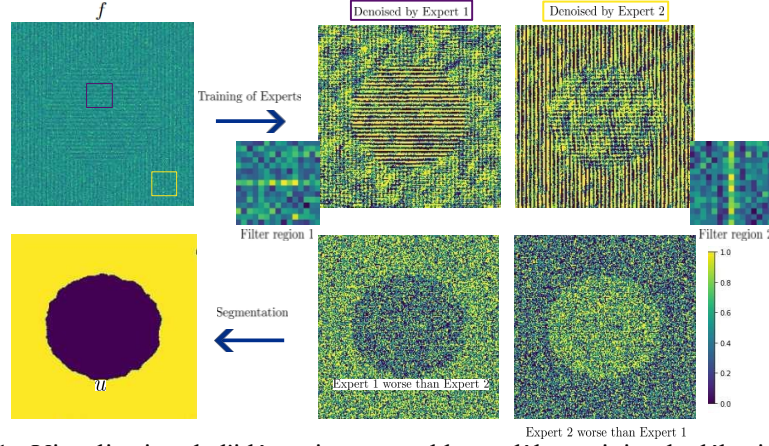


Figure 1 : Visualisation de l'idée qui sous-tend le modèle conjoint de débruitage et de segmentation proposé. Ici, nous avons entraîné deux réseaux composés d'un seul filtre en utilisant la stratégie Noise2Fast [18] et en limitant l'entraînement aux deux boîtes marquées dans l'image bruitée f . Sur les deux images binaires de droite de la rangée inférieure, nous observons que les deux experts en débruitage obtiennent de bien meilleurs résultats dans la région sur laquelle ils ont été entraînés. Les images de différence (image bruit moins débruitage par l'expert 1 (resp. 2) peuvent alors être utilisées dans le processus de segmentation, en exploitant le fait que les régions avec une petite erreur de débruitage pour le premier (resp. le second) expert peuvent être assignées comme premier plan (resp. arrière-plan).

L'effet positif de la segmentation sur le processus de débruitage est encore plus évident dans l'image naturelle de la figure 2. Nous avons utilisé l'architecture de réseau proposée par les auteurs dans [19], ce qui donne un réseau neuronal non linéaire. Le réseau de débruitage a d'abord été entraîné, puis appliqué à l'ensemble de l'image. La deuxième image montre le résultat obtenu avec deux réseaux neuronaux entraînés séparément. Cette stratégie donne un meilleur résultat visuel, confirmé par des valeurs PSNR de 19,69 et 19,12, respectivement. L'image bruitée a été générée en mettant à l'échelle l'image d'entrée RVB propre donnée sur $[0, 1]$ et en ajoutant un bruit gaussien distribué de manière aléatoire et mis à l'échelle de la valeur maximale du pixel, en veillant à ce que le bruit soit proportionnel à l'intensité de l'image. Ici, nous avons utilisé un masque généré manuellement pour région du zèbre et de l'arrière-plan et, pendant l'entraînement, nous avons calculé l'erreur quadratique moyenne (EQM) restreinte aux deux régions différentes, respectivement.

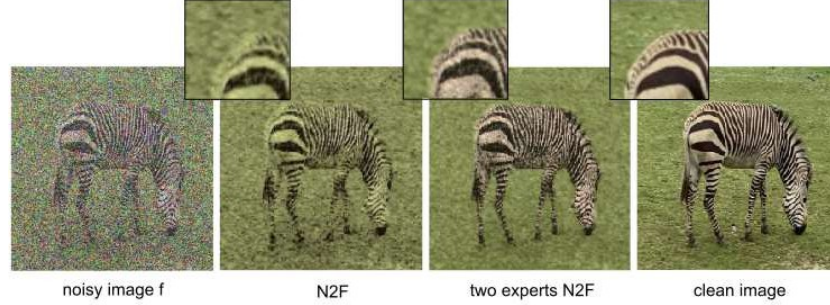


Figure 2 : Image d'entrée RVB bruitée (corrompue par un bruit gaussien, niveau de bruit = 0,75), image débruitée à l'aide de Noise2Fast sur l'ensemble de l'image, experts spécifiques à une région et image de référence. Nous observons clairement des bords plus nets et une meilleure récupération des informations sur les couleurs dans l'exemple "deux experts".

Plus précisément, nous minimisons la fonction de perte

$$S_{f,\lambda}(u, D_F, D_B) = \lambda \|u\|_{TV} + \int_{\Omega} (f(x) - D_F(f))^2 u(x) dx + \int_{\Omega} (f(x) - D_B(f)(x))^2 (1 - u(x)) dx \quad (1)$$

conjointement sur u, D_F et D_B , pris dans une classe spécifique $U \times D_F \times D_B$. Ici, u et $1 - u$ prennent des valeurs dans $[0, 1]$ et définissent les régions d'avant-plan et d'arrière-plan, et D_F et D_B sont des débruiteurs qui doivent s'adapter aux deux régions. Le couplage d'un modèle de segmentation variationnelle avec un CNN de débruitage auto-supervisé permet de s'affranchir de la nécessité de disposer de données étiquetées et d'un pré-entraînement. Nous intégrons les tâches de débruitage et de segmentation pour améliorer mutuellement les performances.

Plus précisément, nous développons des réseaux de débruitage dédiés aux régions de premier plan et d'arrière-plan, respectivement. De cette manière, la performance globale de débruitage est améliorée en exploitant les différences de performance entre les régions pour dériver le masque de segmentation.

Inversement, nous démontrons que les méthodes de débruitage auto-supervisées existantes, telles que Noise2Self [1] ou Noise2Fast [19], tirent profit de la division de l'image en régions en exploitant différentes structures et en effectuant un apprentissage spécifique à chaque région.

Contrairement à la plupart des méthodes d'apprentissage profond pour la segmentation, notre approche est auto-supervisée et permet d'obtenir des résultats comparables en utilisant une seule image. Contrairement aux approches existantes, nous exploitons la différence de performance des débruiteurs qui peut varier d'une région à l'autre.

dans l'image, et montrent que, grâce à la conception de notre approche, les tâches de débruitage et de segmentation ont une influence positive l'une sur l'autre.

1.2 Travaux connexes

1.2.1 Débruitage et segmentation conjoints

Afin d'améliorer la robustesse au bruit et de traiter les inhomogénéités d'intensité dans le modèle de Chan-Vese, plusieurs extensions basées sur les régions locales ont été proposées dans [20, 36, 27]. Néanmoins, ces méthodes restent sensibles aux caractéristiques créées manuellement et au contour initial. Des outils de pré-filtrage pour mieux préparer l'image à la segmentation sont envisagés dans [7, 21, 35]. Le travail de [24] propose un algorithme de débruitage d'image basé sur la segmentation pour le bruit dépendant du signal. Après un débruitage initial, la segmentation est appliquée à l'image pré-filtrée et, pour chaque segment, le niveau de bruit est estimé. Ensuite, pour chaque région, un débruiteur distinct est appliqué. Dans [6], un modèle abordant la segmentation d'images avec un niveau élevé de bruit ou de flou est présenté. À cette fin, ils proposent une approche variationnelle, associant une extension du modèle constant par morceaux de Mumford-Shah [26] à un modèle de restauration d'image, ce qui le rend plus robuste dans le traitement de l'image corrompue donnée. Une approche variationnelle pour la reconstruction et la segmentation conjointes est proposée dans [11], où un modèle composé d'une reconstruction régularisée par variation totale à partir de données sous-échantillonnées et d'une segmentation basée sur Chan-Vese est utilisé. Les auteurs montrent l'amélioration des performances de la reconstruction et de la segmentation conjointes par rapport à l'approche séquentielle. Dans un autre article de [29], la méthode des ensembles de niveaux de Mumford-Shah permet d'améliorer la qualité images reconstruites et la précision de la segmentation.

1.2.2 Débruitage d'image auto-supervisé

Récemment, les techniques d'apprentissage profond auto-supervisé ont fait l'objet d'une attention croissante dans le domaine de l'apprentissage automatique.

Dans [32], Ulyanov et al. exploitent le fait que la structure interne des CNN résonne de manière inhérente avec la distribution des images naturelles, et utilisent cette observation pour la restauration d'images sans avoir besoin de données d'entraînement supplémentaires. Pour chaque image à restaurer, cette méthode propose donc d'entraîner un CNN à reconstruire l'image considérée. L'idée est que l'arrêt précoce de l'entraînement permet de récupérer une image régularisée et débruitée. Une stratégie différente est proposée dans Noise2Noise [18], où des paires d'images bruitées sont mises en correspondance l'une avec l'autre. L'inconvénient de cette dernière méthode est qu'elle dépend toujours de la disponibilité de ces paires. Dans la pratique, même si l'on acquiert deux images bruyantes, il est difficile de les mettre en correspondance.

Il est souvent difficile d'obtenir des réalisations de la même image [5]. À cette , des méthodes d'apprentissage autosupervisées fonctionnant sur une seule image bruyante, telles que Noise2Void [17], Noise2Self [3] et, plus récemment, Noise2Fast [19], ont été proposées comme alternatives prometteuses. L'autosupervision est réalisée en excluant/masquant le centre (point aveugle) du champ réceptif du réseau. Dans ce type d'apprentissage, on suppose que le bruit est indépendant du pixel et que l'intensité réelle d'un pixel peut être prédite à partir du contexte local de l'image, à l'exception des taches aveugles [17]. L'un des ingrédients de notre méthode est la stratégie d'apprentissage adaptée de Noise2Fast [19], qui est une variante de Noise2Self/Noise2Void. Noise2Fast offre les avantages de l'efficacité de calcul tout en maintenant de bonnes performances. La méthode elle-même sera expliquée plus en détail dans les sections 2 et 4.

1.2.3 Méthodes d'apprentissage pour le débruitage et la segmentation

Dans le contexte des données de microscopie, des approches purement basées sur l'apprentissage profond traitant à la fois de la segmentation et du débruitage sont proposées dans [28, 5]. Dans [28], il est démontré sur divers ensembles de données de microscopie que l'utilisation de périeurs de débruitage autosupervisés améliore les résultats de la segmentation, en particulier lorsque seuls quelques masques de segmentation de la vérité au sol sont disponibles pour l'apprentissage. Dans [5], les auteurs proposent DENOISEG, qui se compose d'un réseau en U pour la segmentation d'images et du schéma de débruitage auto-supervisé Noise2Void [17], qui sont combinés et entraînés avec une perte commune. Il est démontré que l'optimisation globale est plus performante que la contrepartie séquentielle, où l'image est d'abord débruitée puis segmentée. Pour obtenir des performances de débruitage élevées, il est nécessaire de disposer d'une grande quantité de données bruitées. En outre, la fonction de perte est simplement la somme des pertes de segmentation et de débruitage. La fusion des deux tâches est réalisée en partageant le réseau ; la fonction de perte n'incorpore aucun couplage entre les deux tâches. Une approche légèrement différente est présentée dans [23]. Néanmoins, cette méthode nécessite des données de vérité terrain pour les deux tâches d'imagerie.

Le document est organisé comme suit : Dans la section 2, nous formulons les problèmes et passons en revue les composants algorithmiques. La section 3 détaille l'algorithme proposé, suivi de la mise en œuvre numérique dans la section 4. Nous appliquons la méthode aux données de microscopie dans la section 5 et démontrons sa polyvalence sur des images naturelles avec un minimum de conseils de la part de l'utilisateur. Enfin, nous concluons et discutons des pistes de recherche futures.

2 Description du problème

Nous allons maintenant fixer la notation, formaliser le problème et décrire les principaux ingrédients utilisés pour la méthode unifiée de débruitage et de segmentation proposée.

Dans ce qui suit, nous désignons par $\Omega \subset \mathbb{R}^2$ un ensemble limité avec une frontière de Lipschitz, et par un espace de fonctions $\rightarrow(d)$ avec d pour les images en niveaux de gris, et d dans le cas RGB. Nous considérons une image bruitée donnée $f \in F$, que nous voulons conjointement débruiter et diviser en C régions différentes.

Problème 1 (débruitage d'images). *L'objectif du débruitage d'images est de récupérer une image propre g à partir d'une observation bruitée f qui suit un modèle de dégradation d'image $f = g + n$, où n est le bruit de dégradation du signal que nous voulons éliminer.*

Bien que d'autres types de dégradation soient possibles, nous supposons ici un modèle additif et, plus précisément, nous considérons un bruit aléatoire dont la valeur attendue est égale à zéro.

Problème 2 (Segmentation d'images). *La segmentation d'images est le processus qui consiste à diviser automatiquement une image en régions significatives. Sur la base des caractéristiques spécifiques d'une image donnée $f \in F$, on cherche à diviser le domaine de l'image en deux (dans le cas d'une segmentation binaire) régions Σ et $\Omega \setminus \Sigma$. Dans le cas de la segmentation multiclass, l'objectif est de construire un modèle de segmentation multiclass.*

partition $\Omega = \bigcup_{i=1}^C \Sigma(i)$ du domaine de l'image en régions (classes) disjointes, où chacune des régions $\Sigma(1), \dots, \Sigma(C-1)$ représente une structure spécifique d'objets dans f , et $\Omega \setminus (\Sigma(1) \cup \dots \cup \Sigma(C-1))$ représente l'arrière-plan.

Dans ce travail, nous abordons ces deux problèmes simultanément en concevant une fonction énergétique de manière à ce que les deux tâches bénéficient l'une de l'autre. Nous examinons ensuite les deux principaux éléments qui constituent la base de notre approche.

2.1 Formulation convexe de Chan-Vese

Dans [10], Chan et al proposent d'assouplir le problème de segmentation binaire de Chan-Vese et de laisser la solution souhaitée u prendre des valeurs dans $[0, 1]$. L'énergie convexe résultante est

$$\min_{0 \leq u \leq 1} \int_{\Omega} |\nabla u(x)|^2 dx + \int_{\Omega} (c_1 - f(x))^2 u(x) dx + \int_{\Omega} (c_2 - f(x))^2 (1 - u(x)) dx. \quad (2)$$

Les auteurs ont montré que, pour toute constante fixe $c_1, c_2 \in \mathbb{R}$, on peut trouver un minimiseur global pour le problème non convexe en effectuant la minimisation dans (2), et en fixant $\Sigma = \{x : u(x) > \tau\}$ pour a.e. $\tau \in [0, 1]$.

Bien que le modèle soit convexe, il souffre encore de difficultés dans la segmentation des images où l'hypothèse de constance par morceaux n'est pas un a priori approprié pour les différentes régions de l'image, ou si l'image est corrompue par un bruit important. Ces questions constituent les principaux problèmes à résoudre dans le présent document en remplaçant c_1 et c_2 par des débruiteurs de réseaux neuronaux plus flexibles.

2.2 Débruitage auto-supervisé basé sur une seule image

Avant de présenter la perte auto-supervisée, nous introduisons une propriété très importante des fonctions de débruitage qui nous intéressent.

Formulées de très générale, pour une image bruitée donnée $f \in \mathbb{R}^n$, les méthodes de débruitage auto-supervisées sont basées sur une variante de la perte auto-supervisée.

$$L_f(D) = \int_{\Omega} (D(f)(x) - f(x))^{(2)} dx, D \in \mathcal{D} \quad (3)$$

minimisée à l'aide d'une stratégie d'optimisation qui exclut la fonction d'identité, $D = \text{Id}$, qui serait le minimiseur exact mais ne produit pas d'image débruitée.

Une stratégie pour surmonter ce problème est la méthode introduite dans [32] où un modèle génératif D pour minimiser (3) est combiné avec un arrêt précoce. Dans ce cadre, la structure convolutive et l'arrêt précoce empêchent D d'apprendre les caractéristiques fines de l'image (bruit) pour obtenir une image débruitée. Une autre stratégie est le filtrage linéaire avec une restriction sur le filtre [3, 17]. Par exemple, un filtre dont la position centrale est nulle et qui ne prend donc pas en compte l'information de ce pixel mais seulement les zones environnantes peut être utilisé pour débruiter une image en minimisant (3). Un autre type de méthode tente de minimiser une fonction légèrement différente. Motivés par l'exemple précédent, les auteurs de [19] introduisent N masques binaires aléatoires H_k qui suppriment des informations dans l'image. L'apprentissage est ensuite effectué à l'aide de la fonction de perte

$$L_f(D) = \frac{1}{N} \sum_{k=1}^N \int_{\Omega} (D(H_k - f)(x) - f(x))^2 - (1 - H_k) dx. \quad (4)$$

Cette stratégie d'apprentissage empêche le réseau d'apprendre l'opérateur d'identité. Dans ce travail, nous utilisons une variante de Noise2Fast [19], qui est numériquement

moins cher que Noise2Void ou Noise2Self, et où H_k $k = 1, \dots, N$, sont des pixels pairs et impairs de l'image, respectivement.

3 Proposition de débruitage et de segmentation conjoints

Nous présentons maintenant notre modèle conjoint, inspiré des observations décrites à la section 1.1. Pour contrôler la segmentation binaire, nous proposons d'entraîner deux réseaux de neurones de débruitage, chacun se concentrant sur la performance dans 'une des régions à segmenter (cf. figures 1 et 2). Nous désignons ces "experts" par D_F pour l'avant-plan et D_B pour l'arrière-plan. Ces experts sont des réseaux de neurones avec les paramètres F et B , qui sont entraînés avec une stratégie de débruitage modifiée. Mentionnons que le modèle est présenté dans le cas de deux régions, mais l'extension à plusieurs classes est simple, en suivant par exemple le cadre de [14, 1, 25]

3.1 Énergie articulaire fonctionnelle

Dans ce qui suit, nous désignons par $BV(\Omega)$ l'espace de toutes les fonctions intégrables $u \rightarrow \mathbb{R}$ à variation totale bornée $\|u\|_{TV}$, et nous considérons l'ensemble admissible

$$A := \{u \in BV(\Omega) \mid \|u\|_{TV} \leq 1\}.$$

De plus, laissons $i_A : BV(\Omega) \rightarrow [0, \infty]$ la fonction indicatrice associée, qui est à l'intérieur de A et ∞ ailleurs. Les paramètres des deux experts en débruitage, D_F et D_B sont notés $W = (F, B) \in \mathbb{R}^{L_1 \times L_2}$, et sont respectivement dédiés à l'avant-plan et à l'arrière-plan. Notez que le nombre de paramètres pour D_F et D_B ne doit pas nécessairement être le même. Ces deux experts sont des réseaux de neurones entraînés selon la stratégie proposée dans [19]. Nous considérons le modèle conjoint

$$\begin{aligned} S_{f,\lambda}(u, D_F, D_B) = & i_A(u) + \lambda \|u\|_{TV} + \int_{\Omega} (f(x) - D_F(f)(x))^2 u(x) dx \\ & + \int_{\Omega} (f(x) - D_B(f)(x))^2 (1 - u(x)) dx. \end{aligned} \quad (5)$$

Notons que pour des paramètres de réseau W fixes, l'énergie proposée est convexe en u . De plus, nous pouvons seuiller le résultat et toujours avoir un optimum global (voir le Théorème 4).

La figure 3 illustre l'idée qui sous-tend le modèle de segmentation proposé. Pour les images en niveaux de gris, l'algorithme peut être initialisé par le seuillage des valeurs de l'image. Dans les cas plus complexes, l'utilisateur peut être invité à fournir des boîtes représentatives pour les régions d'arrière-plan et de premier plan. Ensuite, alternativement, les experts en débruitage sont formés sur des sous-ensembles des deux régions segmentées différentes et les segmentations sont mises à jour. Dans la pratique, le terme de fidélité des données dans (5) est mis à jour en fonction des performances de débruitage des deux experts D_F et D_B . Pour des paramètres de réseau W fixes, l'énergie (5) est minimisée. En répétant cette procédure jusqu'à ce qu'un critère de convergence soit atteint, nous obtenons le masque de segmentation u , ainsi que l'image débruitée $g \approx u \odot_{D(F)}(f) + (1 - u) \odot_{D(B)}(f)$.

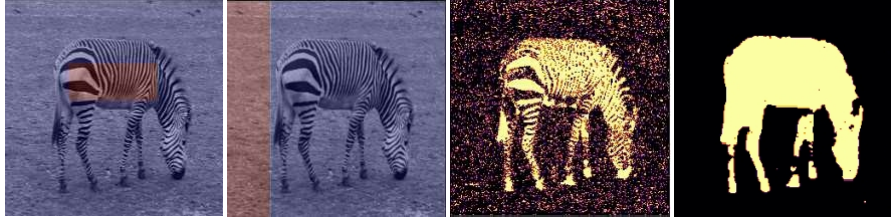


Figure 3 : La première image montre l'image d'entrée en niveaux de gris f , et les boîtes définies par l'utilisateur représentant les régions grossières de premier plan et d'arrière-plan. La troisième image met en évidence les pixels pour lesquels le débruiteur expert d'avant-plan est plus performant que celui d'arrière-plan, tandis que la dernière image représente le résultat de la segmentation obtenu en minimisant l'énergie proposée (5.)

Exemple 1. Nous donnons ici des exemples de réseaux neuronaux qui agissent comme des dénois- seurs et nous nous référons aux approches existantes.

- *Arrière-plan constant* : Dans le cas où l'arrière-plan est supposé constant, on peut simplement supposer que $D_B(f) = B \odot 1$, ce qui correspond à l'estimation d'une valeur scalaire B représentant la valeur moyenne de l'image donnée à l'intérieur de la région correspondante, comme dans le modèle original de Chan et Vese.
- *Filtre linéaire* : Dans ce cas, le réseau est linéaire par rapport aux paramètres du réseau B , plus précisément, $D_B(f) = \omega_B * f$, ce qui conduit à une fonction d'énergie bi-convexe (5). Dans l'exemple de la figure 1, nous avons appliqué un tel réseau linéaire composé d'un seul filtre de taille de noyau 15×15 .

- *Filtrage du terme de fidélité des données* : Lorsque l'une des régions est supposée constante et que des niveaux de bruit élevés sont présents, le filtrage de la moyenne améliore les résultats. Les termes de fidélité des données de l'énergie (5) peuvent alors être remplacés par les termes suivants

$$\int_{\Omega} K(\sigma * (f - (F)f))^{(2)} \text{ et } \int_{\Omega} K(\sigma * (f - (B)f))^{(2)} (1 - \cdot),$$
re-similaire de manière spéculative, où $K(\sigma)$ est un filtre moyen avec une taille de noyau σ . Une méthode a été réalisée dans [20], où une version plus robuste du modèle de Chan-Vese [9] a été proposée.
- *CNN générique* : Tout réseau neuronal de débruitage typique (Deep image prior [32], Noise2Void [17]) peut être utilisé dans notre cadre de travail. Notez que dans ce cas, la bi-convexité de l'énergie (5) n'est plus assurée.

Dans le paragraphe suivant, nous examinons plus en détail la procédure d'optimisation conjointe alternée que nous proposons pour minimiser l'énergie (5.)

3.2 Optimisation conjointe

Nous proposons d'optimiser itérativement le problème (5) à l'aide d'une procédure alternée [12]. Dans le cas où l'étape de débruitage ne minimise pas exactement l'énergie (5), nous alternons en fait entre la minimisation de deux fonctionnelles légèrement différentes. Pour des raisons de lisibilité, cela n'est pas indiqué dans la notation. Nous commençons par initialiser le masque de segmentation u . Cette initialisation est réalisée soit par seuillage pour les images en niveaux de gris, soit, comme le montre la figure 3, en choisissant manuellement des boîtes représentant les différentes régions à segmenter dans l'image. Ensuite, sur la base de l'estimation initiale, les experts en débruitage D_F et D_B sont formés sur les masques initiaux donnés. À cette fin, nous utilisons l'optimiseur ADAM [15] jusqu'à convergence. Dans une étape suivante, pour des paramètres de réseau fixes $W = (F, B)$, nous mettons à jour le masque de segmentation u . Pour W fixe, la fonction d'énergie (5) est convexe et toutes les hypothèses nécessaires à l'application de l'algorithme dual primal [8] sont remplies. Une description plus détaillée des schémas discrets considérés est fournie à la section 4 (voir l'algorithme 2.) Ces étapes alternatives sont répétées tant que la diminution de l'énergie (5) est supérieure à 15 % (c'est-à-dire $p = 0,15$), ce qui, d'après nos observations empiriques, constitue un bon compromis entre la vitesse de calcul et la qualité des résultats.

Le schéma global d'optimisation conjointe est présenté dans l'algorithme 1. Pour les deux experts en débruitage, nous utilisons la même architecture de réseau, à savoir un réseau neuronal convolutionnel composé de quatre instances séquentielles de couches convolutionnelles, chacune d'entre elles étant constituée de 64 canaux. La dernière couche est à nouveau une convolution avec un noyau de taille 1×1 , et la sortie finale est obtenue par l'application d'une activation sigmoïde. Dans chaque étape de mise à jour du débruitage, nous

utilisent la même stratégie que celle proposée dans [19], à savoir qu'après chaque période d'apprentissage, l'EQM entre l'image d'entrée bruitée et la sortie "débruitée" du réseau est calculée. Si cette métrique ne change pas beaucoup pendant 100 époques, la formation est interrompue et la moyenne des cent dernières images de validation constitue la sortie débruitée finale. La figure 4 donne un aperçu de la procédure alternée.

ENTRÉE : image d'entrée bruitée f
INITIALISATION : $u^0 \rightarrow 1_{\{f>0\}}$, $W^0 = W_0$, choisir $\lambda > 0$ et fixer $p \rightarrow 0,15$
tandis que $S_{f,\lambda}^k(u^k, W^k) / S_{f,\lambda}^{k-1}(u^{k-1}, W^{k-1}) \geq p - S_{f,\lambda}^{k-1}(u^{k-1}, W^{k-1}) / S_{f,\lambda}^{k-2}(u^{k-2}, W^{k-2})$ do
 $W^{k+1} \rightarrow \argmin_W S_{f,\lambda}^k(u^{k+1}, W)$ {avec quelques itérations ADAM pour F et Chan et Vese actualisent pour l'arrière-plan si $D_B(f) = D_B \mathbf{1}$ }
 $u^{k+1} \rightarrow \argmin_u S_{f,\lambda}^k(u, W^k)$ {avec l'algorithme 2}
terminer pendant que
Algorithme 1 : Schéma d'optimisation alternée.

Dans le paragraphe suivant, nous examinerons la propriété de convergence de l'algorithme 1.

3.3 Résultats théoriques

Dans cette section, nous examinons certains résultats théoriques de la fonction d'énergie proposée et de l'algorithme alternatif présenté. Notez que ces résultats sont valables si le débruiteur est formé en minimisant (3).

Remarque 1 (Monotonie de la minimisation alternée). *La fonction énergétique proposée (5) est continue et limitée par le bas. De plus, pour chaque $k > 0$, les relations suivantes sont valables*

$$\begin{aligned} S_{f,\lambda}(u^{(k)}, W^{(k+1)}) &\leq S_{f,\lambda}(u^{(k)}, W^{(k)}) \\ S_{f,\lambda}(u^{(k+1)}, W^{(k+1)}) &\leq S_{f,\lambda}(u^{(k)}, W^{(k)}) \end{aligned}$$

Par conséquent, la séquence générée $\{S_{f,\lambda}(u^{(k)}, W^{(k)})\}_{k \in \mathbb{N}}$ converge de manière monotone.

Théorème 3 (Convergence de l'algorithme 1). *Supposons que l'ensemble de niveaux $S^0 = \{(u, W) : S_{f,\lambda}(u, W) \leq S_{f,\lambda}(u^0, W^0)\}$ de $S_{f,\lambda}$ définie en (5) est compacte et que $S_{f,\lambda}$ est continue sur S^0 . Alors, la séquence $\{(u^k, W^k)\}$ générée par l'algorithme 1 est définie et bornée. De plus, chaque point de cluster de $\{(u^k, W^k)\}$ est un point stationnaire de $S_{f,\lambda}$.*

et c_2 , nous examinons les sorties fixes du réseau $D_F(f)$ et $D_B(f)$. En réécrivant le terme $\|u\|_{TV}$ avec la formule de la zone de coopération, nous montrons que $S_{f,\lambda}(u, W) = \int_0^1 S(\Sigma(\tau), W) - C \, d\tau$, avec $C = \int_\Omega (f - D_F(f))^2 \, dx$ indépendant de u . Ainsi, nous concluons que si u est un minimiseur de l'énergie (5) pour W fixé, alors pour tout $\tau \in [0, 1]$ l'ensemble $\Sigma(\tau)$ doit être un minimiseur de (6). \square

4 Mise en œuvre numérique

Dans ce qui suit, nous décrivons la mise en œuvre numérique de la méthode proposée.

4.1 Étape de segmentation

Nous pouvons réécrire notre sous-problème de segmentation sous la forme suivante

$$\min_{u \in X} \mathcal{J}(K(u)) + G(u), \quad (7)$$

où $K(u) := \nabla u$, $\mathcal{J}(v) := \|v\|_{1,2}$ et $G(u) := \int_\Omega (u - (f - D_F(f)))^2 \, dx + \int_\Omega (f - D_B(f)) \, dx$. Il est établi que $K : X \rightarrow Y$ est un mappage linéaire entre les espaces de Hilbert X, Y et que $\mathcal{J} : Y \rightarrow [0, \infty]$ et $G : X \rightarrow [0, \infty]$ sont convexes et que des fonctions semi-continues inférieures, 'est-à-dire que toutes les hypothèses nécessaires à l'application du cadre de l'algorithme dual primal proposé dans [8] sont remplies.

4.1.1 Discrétion

Dans ce qui suit, nous fixons la notation que nous utiliserons tout au long de cette section. Nous travaillons avec des images discrètes dans $H := \mathbb{R}^{N_1 \times N_2}$, représentant un espace de Hilbert de dimension finie équipé d'un produit intérieur $\langle u, v \rangle = \sum_i u[i]v[i]$ pour $u, v \in H$ avec $i = (i_1, i_2) \in \{1, \dots, N_1\} \times \{1, \dots, N_2\}$. Le gradient discret $\nabla = (\nabla_1, \nabla_2) : H \rightarrow H \times H$ est définie par des différences directes avec des conditions aux limites de Neumann,

$$((\nabla u)[i]) := \begin{cases} (u[i_1+1, i_2] - u[i_1, i_2])/h & \text{si } i_1 < N_1 \\ 0 & \text{si } i_1 = N_1 \end{cases}$$

$$((\nabla u)[i]) := \begin{cases} (u[i_1, i_2+1] - u[i_1, i_2])/h & \text{si } i_2 < N_2 \\ 0 & \text{si } i_2 = N_2. \end{cases}$$

Son adjoint est donné par $\nabla^* (v_1, v_2) = \nabla^* v_1 + \nabla^* v_2 = \begin{pmatrix} \cdot \\ \cdot \end{pmatrix} - \text{div}(v_1, v_2)$ où $\text{div} : H \times H \rightarrow H$ est l'opérateur de divergence discrète et pour $(v_1, v_2) \in H \times H$ nous avons

$$\begin{aligned} (\nabla_{1v(1)}^*)[i] &= \begin{cases} -(v_1[i_1, i_2] - v_1[i_1 - 1, i_2])/h & \text{si } 1 < i_1 < N_1 \\ -v_1[1, i_2] & \text{si } i_1 = 1 \\ v_1[N_1 - 1, i_2] & \text{si } i_1 = N_1 \end{cases} \\ (\nabla_{2v(2)}^*)[i] &= \begin{cases} -(v_2[i_1, i_2] - v_2[i_1, i_2 - 1])/h & \text{si } 1 < i_2 < N_2 \\ -v_2[i_1, 1] & \text{si } i_2 = 1 \\ v_2[i_1, N_2 - 1] & \text{si } i_2 = N_2. \end{cases} \end{aligned}$$

La demi-norme TV discrète et isotrope d'une image $u \in H$ est définie comme suit

$$\|\nabla u\|_{1,2} := \sum_i \sqrt{(\nabla_1 u[i])^2 + (\nabla_2 u[i])^2}.$$

Les versions discrètes de l'ensemble admissible et de la fonction indicatrice correspondante sont $A = \{u \in H \mid 0 \leq u \leq 1\}$, et i_A . La discrétisation du terme de fidélité des données de l'énergie (5) s'écrit $\sum_i F(u[i], W)$, où

$$\begin{aligned} F(u, W) &:= d(u, F) + d(1 - u, B) \\ d(u, F) &:= u - (D_F(f) - f)^2 \\ d(1 - u, B) &:= (1 - u) - (D_B(f) - f)^2. \end{aligned} \tag{8}$$

En utilisant ces notations, la version discrète de l'énergie (5) se lit comme suit

$$S_{f,\lambda}^{cv}(u, W) = i_A(u) + \lambda \|\nabla u\|_{1,2} + \sum_i F(u[i], W). \tag{9}$$

Le problème d'optimisation (9) est en général un problème non convexe et difficile à résoudre. Nous utiliserons la minimisation alternée, avec l'algorithme de Chambolle-Pock [8] pour l'étape de mise à jour du masque de segmentation u , tandis que pour la mise à jour des paramètres du réseau $W = (F, B)$, nous appliquons l'optimisation ADAM [15].

4.1.2 Algorithme de segmentation

Nous donnons ici des détails sur la minimisation de la fonction (5) par rapport à u pour W fixé. Cela correspond à la résolution du problème (7) avec

$$X = H, Y = H^2, \gamma = \lambda \|\cdot\|_{1,2}, K = \nabla, G = i_A + \sum_i F(u[i], W).$$

Comme l'opérateur K est linéaire et que les fonctions f et G sont convexes et semi-continues inférieures, toutes les conditions requises pour l'application de l'algorithme dual primal proposé dans [8] sont remplies.

Pour mettre en œuvre cet algorithme, il est nécessaire de calculer le conjugué de Fenchel f^* de f , ainsi que les mappings proximaux de f^* et G . Nous commençons par la dérivation du conjugué de Fenchel de f . Pour $\|\cdot\|_{1,(2)}$, il correspond à la fonction indicatrice de la boule unitaire de la norme duale, ce qui se traduit par

$i_{2,\infty} = \|\cdot\|_{2,\infty}^*$. On a donc $f^*(v) = i_{2,\infty}(v/\lambda)$, la fonction indicatrice de $\{v \mid \|v\|_{2,\infty} \leq 1\} \subset H^2$. Dans un deuxième, nous calculons les opérateurs proximaux de f^* et G . Rappelons que l'opérateur proximal de la fonction indicatrice i_C d'un ensemble C est donné par la projection orthogonale sur C . La projection $P_{2,\infty} : H \rightarrow H^2$ sur la boule unitaire dans la norme $(2,\infty)$ est donc obtenue par

$$(P_{2,\infty}(v))[i, k] = \frac{v[i, k]}{\max\{1, (\frac{v[i, k]^2}{1} + \frac{v[k, k]^2}{2})^{1/2}\}}.$$

Ainsi, l'opérateur proximal de f^* se traduit par

$$\text{prox}_{F^*}(v) = P_{2,\infty,\lambda}(v) := P_{2,\infty}(v/\lambda).$$

De plus, en introduisant $\tilde{f} = (f - D_F(f))^2 - (f - D_B(f))^2$, on peut enfin montrer que

$$\text{prox}_{\tau G}(u_0[i]) = P_A u_0[i] - \tau \tilde{f}[i].$$

L'algorithme dual primal 2 est résumé ci-dessous.

```

ENTRÉE : image d'entrée bruitée  $f \in H$ 
INITIALISATION :  $v^0 \in H, u^0, \bar{u}^0 \in H$  while  $\|u^{n+1} - u^n\| > \varphi$  do
     $v^{n+1} \rightarrow P_{2,\infty,\lambda}(v^n + \sigma u \nabla^{-n})$ 
     $u^{n+1} \rightarrow P_A(u^n - \tau \nabla^{\top} v^{(n)+1} - \tau f)$ 
     $\bar{u}^{n+1} \rightarrow u^{n+1} + \eta(u^{n+1} - u^n)$ 
end while return  $u^{n+1}$ 

```

Algorithme 2 : Algorithme de segmentation basé sur la minimisation de l'indice de masse corporelle.

fonctionnelle de l'énergie (5) par rapport à u pour un W fixe.

4.2 Accélération avec un masque préalable

Dans nos expériences, nous avons observé que l'un des débruiteurs (celui qui est formé sur la région la plus complexe) tend à s'améliorer sur l'autre région.

Une fois que cette progression a commencé, il est assez difficile d'empêcher le masque de segmentation de s'étendre et de converger vers un minimum indésirable qui est le résultat constant de la segmentation. Inspirés par les travaux de [2, 34], nous proposons maintenant de surmonter le problème de la recherche d'un critère d'arrêt approprié, en ajoutant un terme de fidélité, garantissant que le masque de segmentation mis à jour u^k ne s'écarte pas trop de sa supposition initiale. Supposons que nous disposons d'un masque de référence u^0 , puis, dans le continu, nous considérons les problèmes :

$$S_{f,\lambda}^k(u, D_F, D_B) := i_A(u) + \lambda \|u\|_{TV} + \int_{\Omega} (f - D_F(f))^2 u \, dx + \int_{\Omega} (f - D_B(f))^2 (1 - u) \, dx + \frac{\mu}{2} \|u - u^k\|_R^2. \quad (10)$$

Nous pouvons donc optimiser itérativement le problème (10) avec la procédure alternative présentée dans l'algorithme 3. Notez que dans ce cas, comme l'énergie globale est modifiée à chaque itération, nous n'avons plus de garantie de convergence pour la procédure alternée.

ENTRÉE : image d'entrée bruitée f

INITIALISATION : $u^0 \rightarrow f$ et $W^0 = W_0$, u^0 , choisir $\lambda > 0$

Pour $k = 1, \dots, N$ faire

$W^{k+1} \rightarrow \arg\min_W S_{f,\lambda}^k(u^{k+1}, W)$ {avec quelques itérations ADAM pour F et Chan et Vese actualisent pour l'arrière-plan si $D_B(f) = D_B - 1$ }

$u^{k+1} \rightarrow \arg\min_u S_{f,\lambda}^k(u, W^k)$ {avec l'algorithme 4}

$u_R^{k+1} = u^{k+1}$ (mise à jour du masque de référence)

fin pour

Algorithme 3 : Schéma d'optimisation alternée avec accélération.

Pour résoudre le problème de segmentation, nous reformulons l'optimisation problème (10) pour W fixe comme $\min_u \mathcal{G}(u) + G^k(u)$, avec

$$\mathcal{G}(u) = i_A(u) + \frac{\mu}{2} \|u - u^k\|_R^2 + \int_{\Omega} (f(x) - D_F(f)(x))^2 u(x) \, dx + \int_{\Omega} (f(x) - D_B(f)(x))^2 (1 - u(x)) \, dx.$$

Rappelant que $\tilde{f} = (f(x) - D_F(f)(x))^{(2)} - (f(x) - D_B(f)(x))^{(2)}$, on peut montrer que :

$$\text{prox}_{\tau G^k}(u^0[i]) = P_A \frac{u^0[i] + \tau \mu u_R^k[i] \tilde{f}[i]}{u^0[i] + \tau \mu u_R^k[i] + \tau \mu \tilde{f}[i]}.$$

En observant que G^k est μ -fortement convexe en u , nous considérons l'algorithme dual primal accéléré de [8] pour résoudre le problème (10.)

ENTRÉE : image d'entrée bruitée $f \in H$
INITIALISATION : $v^0 \in H, u^0, \bar{u}^0 \in H, \lambda > 0, \sigma > 0, \tau > 0, W$ fixe
while $\|u^{n+1} - u^n\| > \rho$ do
 $v^{n+1} \rightarrow P_{2,\infty,\lambda}(v^n + \sigma \nabla u^n)$
 $u^{n+1} \rightarrow P_A(u - \tau \nabla^\top v^{n+1} + \tau \mu u_R^k - \tau f) / (1 + \tau \mu)$
 $\eta = \frac{1}{1+2\mu\tau}, \tau = \tau\eta, \sigma = \sigma\eta$
 $\bar{u}^{n+1} \rightarrow u^{n+1} + \eta(u^{n+1} - u^n)$
end while return u^{n+1}

Algorithme 4 : Algorithme de segmentation basé sur la minimisation de l'indice de masse corporelle.

fonctionnelle de l'énergie (10) par rapport à u pour un W fixe.

Comme nous avons discuté de la mise en œuvre numérique de l'étape de segmentation, nous présentons maintenant la configuration discrète et la mise en œuvre de l'étape de débruitage.

4.3 Étape de débruitage à l'aide de la stratégie Noise2Fast

Nous détaillons ici le débruitage d'une image 2D discrétisée $f \in \mathbb{R}^{m \times n}$ composée d'un signal propre $g \in \mathbb{R}^{n \times n}$ et d'un bruit $n \in \mathbb{R}^{m \times n}$, c'est-à-dire

$$f = g + n.$$

Par souci d'exhaustivité, nous introduisons u^k qui, pour $k = 0$, correspond à l'initialisation de la région d'arrière-plan. Ces masques peuvent être obtenus par seuillage de l'image ou être donnés sous forme de boîtes fournies par l'utilisateur.

Pour les étapes de mise à jour suivantes, c'est-à-dire $k = 1, \dots, B$, il est vrai que $u^k = 1 - u^k$. En utilisant

Avec ces notations, pour $u^k = (u^k, u^k)$ fixe à la k -ième étape de débruitage de notre procédure alternée, la fonction d'énergie (5) se réduit à

$$\min_W \sum_i F(u^k[i], W) = \min_{F,B} \sum_i (D_F(f)[i] - f[i]^{(2)} - u^k[i] + (D_B(f)[i] - f[i]^{(2)} - u^k[i]), \quad B \quad (11)$$

où $D_F(f)$ et $D_B(f)$ sont à nouveau les experts (profonds) respectivement dédiés au débruitage du premier plan et de l'arrière-plan.

Nous construisons nos débruiteurs sur la base de la méthode Noise2Fast introduite par Lequyer et al. dans [19]. Les auteurs proposent un débruiteur aveugle rapide pour une seule image, en utilisant une stratégie spéciale de sous-échantillonnage. Plus précisément, leur méthode

consiste à diviser une image donnée en parties plus petites en utilisant une stratégie de sous-échantillonnage en damier. À partir d'une seule image, quatre images sont ainsi générées, en supprimant la moitié de tous les pixels et en déplaçant les pixels restants pour combler les vides. Ensuite, un réseau est formé pour apprendre les correspondances entre les paires d'images sous-échantillonnées résultantes. En raison de la redondance interne sous forme de patches récurrents présents dans les images et du degré élevé d'auto-similarité, le réseau neuronal sera également capable de débruiter l'image entière au lieu des images sous-échantillonnées [4, 37, 13]. Pour une description plus détaillée de la stratégie d'apprentissage Noise2Fast, telle que l'architecture du réseau, nous renvoyons le lecteur à [19].

Dans notre approche, nous utilisons une fonction de perte différente de celle décrite dans les travaux de Lequyer et al [19]. Au lieu de considérer l'ensemble du domaine de l'image pour l'apprentissage, nous limitons le processus d'optimisation pour l'expert en premier plan $D_F(f)$ (resp. en arrière-plan $D_B(f)$) aux masques de segmentation actuels u^k (resp. $1 - u^k$) obtenus par l'algorithme 2.

Dans un premier temps, comme dans [19], les images d'apprentissage sous-échantillonnées sont générées de la manière suivante

$$\begin{aligned} f_{\text{even}}(i, j) &= f(i, 2j + (i \bmod 2)) \in \mathbb{R}^{m \times \frac{n}{2}} \\ f_{\text{odd}}(i, j) &= f(i, 2j + (i \bmod 2) + 1) \in \mathbb{R}^{m \times \frac{n}{2}} \\ f_{\text{even}}(i, j) &= f(2i + (i \bmod 2), j) \in \mathbb{R}^{\frac{m}{2} \times n} \\ f_{\text{odd}}(i, j) &= f(2i + (i \bmod 2) + 1, j) \in \mathbb{R}^{\frac{m}{2} \times n}, \end{aligned}$$

et nous répétons cette procédure de sous-échantillonnage pour les masques de segmentation u^k et u^k , pour $k = 0, \dots, N$ également. Nous notons

$$\mathbf{J}^k = \{(f_{\text{even}}, f_{\text{odd}}), u_{\text{imp}}^k, u_{B, \text{odd}}^k\}, (f_{\text{impair}}, f_{\text{pair}}, u_{\text{mém}}^k, u_{B, \text{voire}}^k), \\ (f_{\text{mém}}', f_{\text{impair}}', u_{\text{impair}}^{\text{air}}, u_{B, \text{odd}}^{\text{air}}), (f_{\text{impair}}', f_{\text{mém}}', u_{\text{mém}}^{\text{e}}, u_{B, \text{even}}^{\text{e}})\}$$

l'ensemble d'apprentissage pour $k = 0, \dots, N$, N étant le nombre d'itérations de la minimisation alternée.

Nous formons ensuite les deux réseaux de débruitage, $D_F(f)$ et $D_B(f)$, limités aux régions données, u^k et u^k , c'est-à-dire que pour $(\tilde{f}, \tilde{g}, \tilde{u}, \tilde{u}_B) \in \mathbf{J}^k$, nous minimisons

$$\mathcal{L}^k(W) = \sum_i D_F(\tilde{f})[i] - \tilde{g}[i])^2 - \tilde{u}[i] + D_B(\tilde{f})[i] - \tilde{g}[\tilde{u}_B])^2 - \tilde{u}_B[i]. \quad (12)$$

Ainsi, les débruiteurs auto-supervisés apprennent à reconstruire les pixels pairs (resp. impairs) et \tilde{f} de l'image grâce aux pixels impairs (resp. pairs) \tilde{g} . Comme mentionné

ci-dessus, causée par la redondance de l'autosimilarité, en minimisant (12), L^k , on obtient permet également de résoudre le problème (11.)

Dans le paragraphe suivant, nous démontrons les applications possibles de trois variantes différentes de la méthode conjointe de débruitage et de segmentation proposée. Le code est fourni sur GitHub (<https://github.com/Nadja1611/Self2Seg.git>).

5 Expériences et résultats

Comme première application, nous testons notre méthode sur le jeu de données de noyaux de cellules microscopiques du jeu de données DSB2018¹ issu du défi Kaggle 2018 Data Science Bowl. Les données consistent en une collection diversifiée de noyaux cellulaires imagés par différents microscopes à fluorescence. Les patches sont de taille 128×128 , et sont accompagnés de vérités de base de segmentation générées manuellement. Plus précisément, nous utilisons des données sans bruit et ajoutons manuellement du bruit gaussien avec trois niveaux de bruit différents, à savoir 10, 30 et 50. Dans nos expériences, nous avons considéré le même sous-ensemble d'images que celui utilisé dans [5], où les auteurs ont démontré que la segmentation des données bruitées peut être améliorée en traitant le débruitage et la segmentation de manière coopérative (mais pas totalement conjointe).

Dans les expériences suivantes, nous utilisons la métrique de Dice pour évaluer la performance de la segmentation et la métrique de similarité structurelle (SSIM) pour mesurer la performance du débruitage dans les expériences, en choisissant le rapport signal/bruit maximal (PSNR).

Nous arrêtons notre algorithme alternatif 1 dès que la diminution de l'énergie (5) est inférieure à 15 % du taux de diminution précédent. Nous avons plusieurs stratégies différentes, et celle-ci s'est avérée être la plus prometteuse. Nous avons en effet observé qu'un critère basé sur la variation de la décroissance de l'énergie est robuste à différentes échelles pour le paramètre de régularisation λ , et qu'il s'adapte également à différents types d'images.

Nous comparons les performances de segmentation de notre approche conjointe de débruitage et de segmentation avec le modèle convexe de Chan-Vese de [10] appliqué soit directement sur les images bruitées, soit sur les données préalablement débruitées dans le cadre d'une approche séquentielle. Pour l'approche conjointe proposée et l'approche séquentielle, nous utilisons la même image débruitée comme point de départ pour des comparaisons équitables. En outre, nous testons notre méthode par rapport à la méthode en deux étapes [7] et au cadre de débruitage et de segmentation partiellement conjoint basé sur l'apprentissage profond de [5].

¹ <https://www.kaggle.com/c/data-science-bowl-2018>

Segmentation avec l'hypothèse d'un fond constant

Nous commençons par évaluer notre méthode sur un sous-ensemble des données DSB2018 sur les noyaux cellulaires qui ont été corrompues manuellement (niveaux de bruit 10, 30 et 50). Pour cela, nous entraînons un débruiteur d'avant-plan, $D_F(f)$, et nous supposons que l'arrière-plan est constant, c'est-à-dire $D_B = B \odot 1$. Pour ce type particulier d'images, cette hypothèse est utile, tandis que pour les images présentant des motifs plus structurels, ce choix peut ne pas être raisonnable et deux experts en débruitage peuvent être nécessaires.

Pour appliquer notre approche conjointe, nous commençons par débruiter l'image donnée en utilisant la stratégie Noise2Fast de la manière décrite à la section 4, et nous utilisons l'image débruitée rafraîchie (avec le seuil ϱ fixé à 0,5) comme initialisation. Pour le niveau de bruit 10, nous avons appliqué l'algorithme de segmentation 1 avec l'hypothèse d'un fond constant, tandis qu'un expert Noise2Fast a été envisagé pour des niveaux de bruit plus élevés. Nous rappelons que le processus global de résolution de la segmentation et du débruitage conjoints est présenté dans l'algorithme 1. Selon le type d'image, le processus alternatif nécessite entre deux et six itérations pour satisfaire aux critères de convergence.

Pour chaque méthode, nous avons mené les expériences avec dix valeurs différentes du paramètre de régularisation λ réparties uniformément dans l'intervalle $[0, 1]$, puis nous avons sélectionné pour chaque image le résultat présentant valeur de Dice la plus élevée.

À titre de comparaison supplémentaire, nous avons appliqué le modèle convexe de Chan-Vese de [10] directement sur les images bruitées. Les résultats obtenus sont présentés dans les figures 5 à 7, tandis que les mesures d'évaluation de la segmentation sont résumées dans le tableau 1. Nous observons que pour les trois niveaux de bruit, la méthode séquentielle et la méthode de Chan-Vese de [10] sont confrontées à des inhomogénéités d'intensité des cellules. Ces exemples mettent en évidence la force de l'approche unifiée proposée, qui est capable de segmenter les cellules avec des intensités proches de la valeur moyenne de l'arrière-plan. Notez que l'approche proposée ne donne pas de bons résultats dans le dernier exemple en raison de la présence d'inhomogénéités d'intensité, attribuées à un champ variant dans l'espace, le champ de biais, dans le coin supérieur gauche de l'image. Veuillez noter que dans ce cas, l'évaluation des performances de débruitage n'est pas appropriée, car nous supposons un arrière-plan constant et nous n'appliquons pas de débruitage à l'arrière-plan.

En outre, nous l'avons comparé à l'approche en deux étapes décrite par Cai et al. dans [7]. Nous avons utilisé le code MATLAB fourni et affiné le paramètre λ pour trois niveaux de bruit distincts : n10 avec $\lambda = 4,5$, n30 avec $\lambda = 3,5$ et n50 avec $\lambda = 3,2$. Après analyse des résultats, il est devenu évident que cette méthode rencontre des difficultés avec les inhomogénéités d'intensité et tend à produire des résultats excessivement lissés.

Le tableau 1 résume les résultats obtenus par la méthode supervisée de débruitage et de segmentation DENOISEG [5]. Ici, nous avons exécuté le code fourni de manière aléatoire en utilisant 10 images d'entraînement. Plus précisément, nous avons utilisé l'ensemble de données DSB2018 avec un niveau de bruit nul et avons ajouté du bruit gaussien de la même manière que précédemment à l'ensemble des 4320 images, dont 3750 ont été utilisées pour l'entraînement, 670 pour la validation et les mêmes 50 que nous avons utilisées pour nos expériences pour le test. Il convient de mentionner que pour la validation, les 570 images de validation annotées sont utilisées dans [5], ce qui donne un nombre total de 580 images annotées au cours du processus d'apprentissage. Comme le montre le tableau 1, cette méthode est la plus performante. Pour obtenir une comparaison plus équitable en termes de données d'entraînement, nous avons décidé d'adapter leur méthode en utilisant 10 images au total (DENOISEG 10 dans le tableau 1), 7 pour l'entraînement et 3 pour la validation. Dans ce cadre, toutes les données disponibles sont encore utilisées pour l'entraînement du débruiteur, alors que pour le réseau de segmentation, les masques de vérité terrain pour toutes les images sauf les dix images d'entraînement sont mis à zéro. Avec ce niveau de supervision plus faible, notre approche est plus performante que la méthode de [5].

Pour des niveaux de bruit plus élevés, il est nécessaire de filtrer le terme de fidélité de l'arrière-plan. Cela permet d'éviter de considérer des valeurs plus élevées du paramètre de régularisation λ , qui peuvent conduire à une sur-segmentation de l'arrière-plan et à une diminution globale de la performance de la segmentation. Pour les niveaux de bruit 30 et 50, comme mentionné dans la section 3.1, nous minimisons donc

$$S_{f,\lambda}(u, W) = i_A(u) + \lambda \|u\|_{TV} + \int_{\Omega} [K_{\sigma} * (f - D_F(f))]^2 u(x) dx + \int_{\Omega} [K_{\sigma} * (f - D_B(f))]^2 (1 - u(x)) dx$$

K_{σ} étant un filtre moyen avec $\sigma = 3$.

Le paragraphe suivant présente les résultats expérimentaux obtenus en appliquant notre idée de former des experts en débruitage pour les deux régions.

Bruit	Base de référence	Séquentiel	Proposé	TwoStage [7]	DenoiSeg [5]	DenoiSeg 10
n10	0.820	0.799	<u>0.851</u>	0.829	0.864	0.843
n30	0.773	0.777	<u>0.825</u>	0.800	0.848	0.820
n50	0.582	0.735	<u>0.786</u>	0.758	0.818	0.750

Tableau 1 : Valeurs de dé obtenues sur 50 images du jeu de données DSB2018 pour les méthodes comparées, et trois niveaux de bruit différents. Ici, la ligne de base est la méthode convexe de Chan-Vese [10] directement appliquée aux données bruitées, tandis que pour la méthode séquentielle, nous dénaturons d'abord l'image à l'aide de Noise2Fast [19].

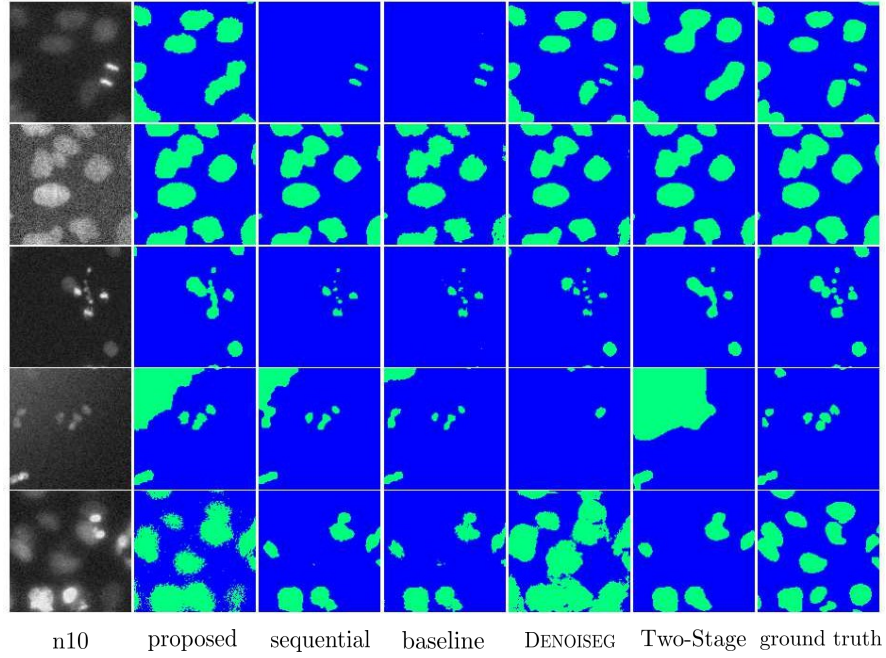


Figure 5 : Comparaison visuelle des résultats de la segmentation des données avec un niveau de bruit de 10. De gauche à droite, cette figure montre : l'entrée bruitée, les résultats obtenus avec l'approche conjointe proposée, l'approche séquentielle, la ligne de base de Chan-Vese et les masques de segmentation de la vérité terrain. Pour toutes les méthodes comparées, le λ maximisant le score de Dice a été choisi.

Segmentation à l'aide de deux débruiteurs

Dans l'exemple de la figure 1 de la section 1.1, nous avons formé deux experts en débruitage (dans ce cas, nous avons utilisé un réseau linéaire composé d'un filtre de taille 15×15) initialisés par les boîtes jaune et violette de taille 30×30 . Nous avons itéré entre le débruitage et la segmentation trois fois, jusqu'à ce que la diminution de l'énergie soit inférieure à 10 %. Pour la segmentation, nous avons fixé le paramètre de régularisation λ à 0,02. Après la première étape de segmentation, les fonctions de perte des débruiteurs ont été limitées à u et $1 - u$ respectivement.

La figure 8 est un exemple typique montrant la force de l'algorithme proposé par rapport aux approches basées sur l'intensité. Dans cette expérience, nous avons prétraité l'image donnée de taille 256×256 de manière à ce que les deux régions aient la même valeur moyenne, et nous avons ajouté un bruit gaussien comme décrit précédemment,

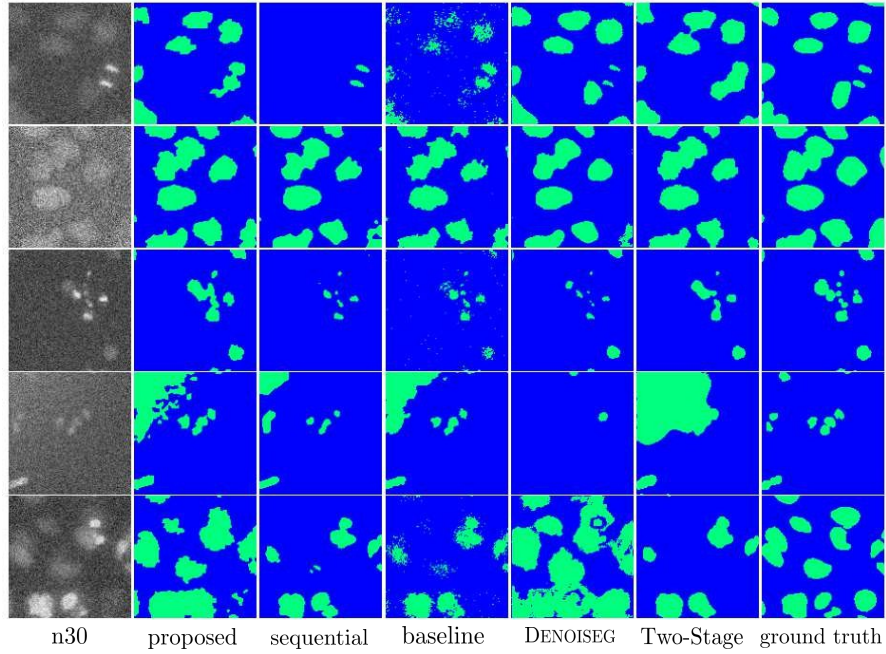


Figure 6 : Comparaison visuelle des résultats de la segmentation des données avec un niveau de bruit de 30.

avec un niveau de bruit de 10. Par conséquent, l'algorithme classique de Chan-Vese échoue totalement sur cet exemple. Ce modèle peut néanmoins donner de bons résultats avec une transformation manuelle adaptée de l'image à segmenter. Comme l'illustrent les deux dernières images de la figure 8, lorsqu'il est alimenté par la carte du gradient normalisé de l'image au lieu des intensités de l'image originale, le modèle de Chan-Vese est capable de segmenter les deux parties de l'image.

En , notre approche est capable d'apprendre automatiquement une transformation pertinente des données d'image et fournit une excellente segmentation sans aucune astuce préalable. La raison en est encore une fois que les poids appris par les deux experts en débruitage dépendent fortement du véritable signal sous-jacent qui, contrairement à l'intensité moyenne, est différent dans les deux régions. Ici, les deux experts en débruitage ont été initialisés par des boîtes de taille 50×50 centrées sur les régions. Nous avons utilisé un paramètre de régularisation λ de 0,06 et fixé le taux d'apprentissage à 0,001. En utilisant le même critère d'arrêt que dans l'exemple des cellules, ces résultats ont été obtenus après 3 itérations de la procédure alternant les étapes de débruitage et de segmentation.

Dans la figure 9, nous affichons l'image propre considérée dans l'expérience

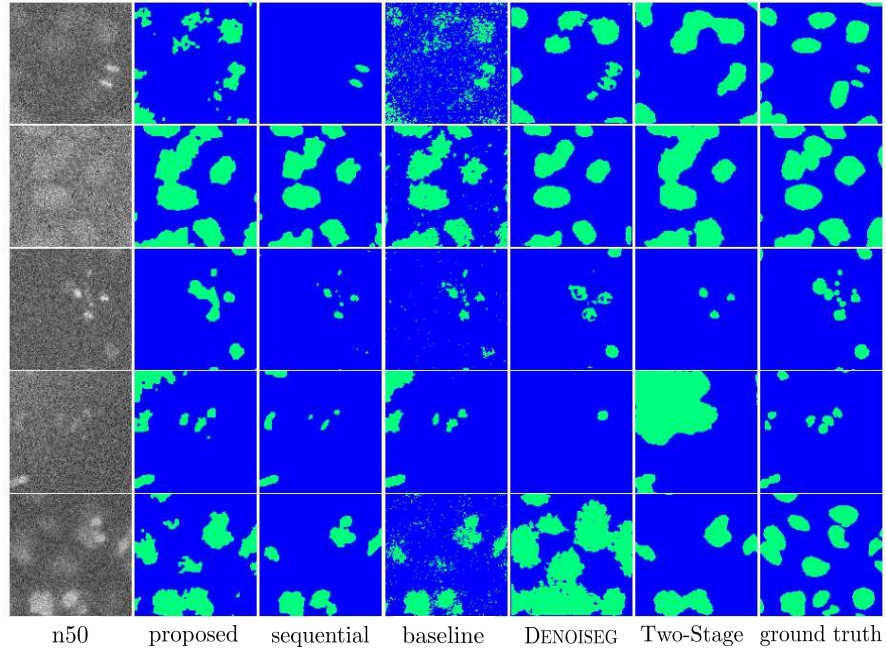


Figure 7 : Comparaison visuelle des résultats de la segmentation des données avec un niveau de bruit de 50.

de la figure 8, ainsi que différentes images débruitées avec leurs mesures quantitatives correspondantes. Plus précisément, la deuxième image de la figure est obtenue en appliquant la stratégie Noise2Fast à l'ensemble de l'image, tandis que la troisième image est le résultat de la procédure d'optimisation conjointe proposée, où l'image est composée à l'aide du masque de segmentation u et des images débruitées par les deux experts en débruitage. En particulier dans les régions de gauche, nous pouvons observer une meilleure performance de débruitage de la méthode proposée, ce qui est plus évident en examinant les valeurs PSNR (20,36 contre 19,815) et SSIM (0,753 contre 0,696).

Segmentation avec un masque de référence à l'aide de l'algorithme 3

La figure 10 montre un autre exemple de segmentation d'image pour trois niveaux de bruit différents à l'aide de l'algorithme 4. La principale difficulté de cette image réside dans les intensités partagées par l'objet à segmenter et l'arrière-plan. Par conséquent, nous avons choisi une boîte représentative pour initialiser l'écureuil, qui comprend à la fois des zones sombres et lumineuses, afin d'en-

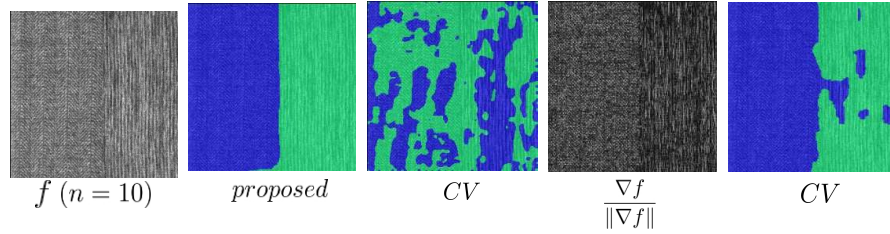


Figure 8 : Segmentation d'une image bruyante de Brodatz composée de deux textures différentes. Les trois premières images montrent l'entrée bruitée f , le minimiseur d'énergie (5) et le résultat obtenu en appliquant directement l'algorithme de contour actif [9]. La quatrième image montre le gradient normalisé de f , et la dernière est le résultat obtenu en appliquant l'algorithme classique de Chan-Vese sur la carte de gradient normalisé.

L'expert en débruitage d'avant-plan peut ainsi mieux généraliser sur la région d'avant-plan composée de zones sombres et lumineuses. Naturellement, comme l'écureuil et l'arrière-plan ne diffèrent pas beaucoup en termes de propriétés structurales, le débruiteur d'avant-plan, $D_F(f)$, est également performant sur l'arrière-plan, ce qui entraîne une augmentation du masque de segmentation u . *Afin de contrôler ce comportement, nous avons appliqué notre deuxième stratégie qui comprend un masque de référence récursif tel que décrit dans le chapitre précédent.* Afin de contrôler ce comportement, nous avons appliqué notre deuxième stratégie qui inclut un masque de référence récursif tel que décrit dans l'algorithme 3, empêchant ainsi le masque de segmentation obtenu à l'itération $k+1$ de trop s'écarter du masque précédent à l'itération k . Plus précisément, les paramètres que nous avons utilisés pour le niveau de bruit 10 étaient $\mu=0,0001$, $\lambda=0,005$, pour le niveau de bruit 30 nous avons fixé $\mu=0,005$, $\lambda=0,005$, tandis que pour un niveau de bruit 50, nous avons considéré $\mu=0,00015$ et $\lambda=0,005$.

Dans ce qui suit, nous discutons de certaines extensions possibles et des limites actuelles de l'approche conjointe de débruitage et de segmentation proposée.

6 Extensions et limitations

Dans cette section, nous montrons d'abord comment notre cadre unifié peut être étendu au cas de la segmentation multiclasse (multicanal). Nous examinons ensuite les limites actuelles de l'approche conjointe de débruitage et de segmentation proposée.

6.1 Modèle multi-classes à valeurs vectorielles

Afin de segmenter une image vectorielle corrompue par le bruit et représentée par $\mathbf{f} = (f_1, \dots, f_L)$ en C régions différentes, nous pouvons considérer C réseaux

neuronaux dédiés agissant en tant qu'experts en débruitage pour chaque région.
Dans ce , les

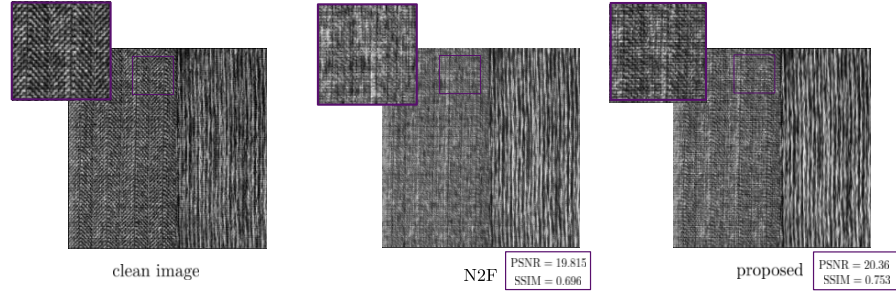


Figure 9 : Comparaison des performances de débruitage avec différentes stratégies Noise2Fast. Sur l'image du milieu, Noise2Fast est appliqué à l'ensemble de l'image. Sur l'image de droite, nous présentons l'image débruitée finale obtenue à partir des deux débruiteurs distincts appris avec le cadre de travail proposé.

L'objectif est d'estimer C masques de segmentation $\{u_i\}_{i=1}^C$ satisfaisant au critère simplifié, c'est-à-dire $\sum_{i=1}^C \int_{\Omega} u_i = 1$, ainsi que l'ensemble des paramètres du réseau $\mathcal{W}^{\text{MC}} = (\mathcal{W}_1^{\text{MC}}, \dots, \mathcal{W}_C^{\text{MC}})$. Avec ces notations, l'énergie (5) peut être étendue à la segmentation d'images vectorielles corrompues par le bruit \mathbf{f} comme suit

$$S_{f,\lambda}(u, \mathcal{W}) := \int_{\Omega} \|\mathbf{f} - \mathbf{D}_{\mathcal{W}^{\text{MC}}}(f_j)_i\|_{\text{TV}^+}^2 u_i. \quad (13)$$

Comme précédemment, il peut ne pas être nécessaire d'entraîner C réseaux de débruitage différents, car certaines régions peuvent être supposées constantes et, dans ce cas, l'"expert" pour la région i peut être remplacé par la valeur moyenne de l'image à l'intérieur de la région i .

6.2 Limites

Une limitation est rencontrée lors de l'apprentissage de deux débruiteurs, l'un d'eux pouvant englober des parties des deux régions pendant la segmentation, ce qui conduit à une convergence indésirable des masques. Bien que nous proposons une stratégie pour remédier, l'ajout d'une contrainte pour donner la priorité aux performances de débruitage dans les régions initiales pourrait être bénéfique.

En outre, le débruiteur Noise2Fast peut ne pas convenir à certaines images, comme l'exemple du zèbre, en raison de son fonctionnement en filtre local. Il peut en résulter des résultats similaires au modèle de Chan-Vese à constante par morceaux. Nous avons modifié la stratégie Noise2Fast et suggérons d'explorer d'autres débruiteurs, tels que le deep image prior [32]. En outre, l'utilisation de

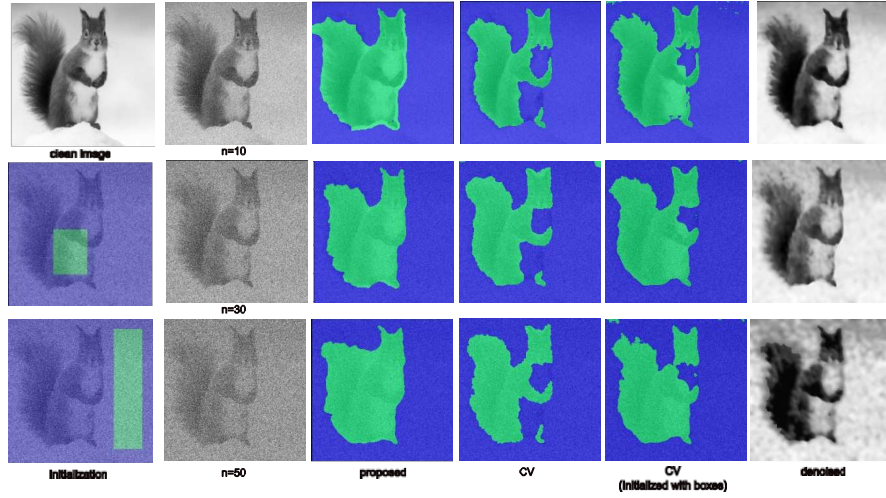


Figure 10 : Résultats de la segmentation obtenus sur les images bruitées montrant un écureuil corrompu par trois niveaux de bruit différents. La première colonne montre l'image d'entrée propre et l'initialisation des régions d'avant-plan et d'arrière-plan, tandis que la deuxième colonne présente les versions bruitées de l'image donnée. Les autres colonnes présentent les résultats de segmentation obtenus en utilisant la stratégie proposée avec l'algorithme de segmentation 4, les masques de segmentation utilisant l'algorithme de Chan-Vese fourni par skimage [33] avec une initialisation en damier et une initialisation en boîte, respectivement. La dernière colonne montre les images débruitées qui sont obtenues en composant le masque de segmentation obtenu et les sorties du débruiteur expert.

l'introduction de nouveaux termes d'adaptation des données axés sur les similitudes structurelles au sein des classes pourrait améliorer les performances de la segmentation.

7 Conclusion

Dans ce travail, nous avons proposé une nouvelle fonction énergétique pour le débruitage et la segmentation d'images. Notre cadre combine les avantages de modèles variationnels bien établis avec des stratégies modernes d'apprentissage profond auto-supervisé. L'un des principaux atouts de la méthode réside dans le fait qu'elle peut traiter des images uniques sans qu'il soit nécessaire de recourir à des masques de segmentation de référence ou à des paires d'entraînement propres et bruyantes. En outre, la fonction énergétique est conçue manière à ce que les deux tâches bénéficient l'une de l'autre, ce qui a également été confirmé par des expériences.

Références

- [1] Egil Bae et Ekaterina Merkurjev. Convex variational methods on graphs for multiclass segmentation of high-dimensional data and point clouds. *Journal of Mathematical Imaging and Vision*, 58:468-493, 2017.
- [2] Antonio Baeza, Vicent Caselles, Pau Gargallo et Nicolas Papadakis. A narrow band method for the convex formulation of discrete multilabel problems. *Multiscale Modeling & Simulation*, 8(5):2048-2078, 2010.
- [3] Joshua Batson et Loic Royer. Noise2Self : Blind denoising by self supervision. In *International Conference on Machine Learning*, pages 524-533. PMLR, 2019.
- [4] Antoni Buades, Bartomeu Coll et Jean-Michel Morel. Self-similarity- based image denoising (débruitage d'images basé sur l'autosimilarité). *Communications of the ACM*, 54(5):109-117, 2011.
- [5] Tim-Oliver Buchholz, Mangal Prakash, Deborah Schmidt, Alexander Krull et Florian Jug. DenoiSeg : joint denoising and segmentation. In *Computer Vision- ECCV 2020 Workshops : Glasgow, UK, August 23-28, 2020, Proceedings, Part I*, pages 324-337. Springer, 2021.
- [6] Xiaohao Cai. Modèle variationnel de segmentation d'image couplé à des réalisations de restauration d'image . *Pattern Recognition*, 48(6):2029-2042, 2015.
- [7] Xiaohao Cai, Raymond Chan et Tiejong Zeng. A two-stage image segmentation method using a convex variant of the Mumford-Shah model and thresholding. *SIAM Journal on Imaging Sciences*, 6(1):368- 390, 2013.
- [8] Antonin Chambolle et Thomas Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of mathematical imaging and vision*, 40:120-145, 2011.
- [9] Tony Chan et Luminita Vese. Un modèle de contour actif sans bords. In *Scale-Space Theories in Computer Vision : Second International Conference, Scale-Space'99 Corfu, Greece, September 26-27, 1999 Proceedings 2*, pages 141-151. Springer, 1999.
- [10] Tony F Chan, Selim Esedoglu et Mila Nikolova. Algorithms for finding global minimizers of image segmentation and denoising models. *SIAM journal on applied mathematics*, 66(5):1632-1648, 2006.

- [11] Veronica Corona, Martin Benning, Matthias J Ehrhardt, Lynn F Glad- den, Richard Mair, Andi Reci, Andrew J Sederman, Stefanie Reichelt et Carola- Bibiane S nlieb. Amélioration de la reconstruction et de la segmentation conjointes avec l'itération de Bregman non convexe. *Inverse Problems*, 35(5):055001, 2019.
- [12] Imre Csisz´ ar. Information geometry and alternating minimization pro- cedures. *Statistics and Decisions*, 1:205-237, 1984.
- [13] Daniel Glasner, Shai Bagon et Michal Irani. Super-résolution à partir d'une seule image. In *2009 IEEE 12th international conference on computer vision*, pages 349-356. IEEE, 2009.
- [14] Nadja Gruber, Johannes Schwab, S´ ebastien Court, Elke R Gizewski et Markus Haltmeier. Lifting-based variational multiclass segmenta- tion algorithm : Design, convergence analysis, and implementation with applications in medical imaging. *Journal of Computational and Applied Mathematics*, 439:115583, 2024.
- [15] Diederik P Kingma et Jimmy Ba. Adam : A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [16] Sreedhar Kollem, Katta Rama Linga Reddy et Duggirala Srinivasa Rao. Une revue des méthodes de débruitage et de segmentation d'images basées sur des images médicales. *International Journal of Machine Learning and Computing*, 9(3):288-295, 2019.
- [17] Alexander Krull, Tim-Oliver Buchholz et Florian Jug. Noise2Void- apprentissage du débruitage à partir d'images bruitées uniques. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2129-2137, 2019.
- [18] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala et Timo Aila. Noise2Noise : Learning image restoration without clean data. *arXiv preprint arXiv:1803.04189*, 2018.
- [19] Jason Lequyer, Reuben Philip, Amit Sharma, Wen-Hsin Hsu et Lau- rence Pelletier. A fast blind zero-shot denoiser. *Nature Machine Intel- ligence*, 4(11):953-963, 2022.
- [20] Chunming Li, Chiu-Yen Kao, John C Gore et Zhaohua Ding. Implicit active contours driven by local binary fitting energy. In *2007 IEEE con- ference on computer vision and pattern Recognition*, pages 1-7. IEEE, 2007.

- [21] Xu Li, Xiaoping Yang et Tieyong Zeng. A three-stage variational image segmentation framework incorporating intensity inhomogeneity information. *SIAM Journal on Imaging Sciences*, 13(3):1692-1715, 2020.
- [22] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen AWM Van Der Laak, Bram Van Ginneken et Clara I Sánchez. Une étude sur l'apprentissage profond dans l'analyse d'images médicales. *Medical image analysis*, 42:60-88, 2017.
- [23] Ding Liu, Bihan Wen, Xianming Liu, Zhangyang Wang et Thomas S Huang. Quand le débruitage d'images rencontre des tâches de vision de haut niveau : A deep learning approach. *arXiv preprint arXiv:1706.04284*, 2017.
- [24] Xinhao Liu, Masayuki Tanaka et Masatoshi Okutomi. Signal dependent noise removal from a single image. In *2014 IEEE International Conference on Image Processing (ICIP)*, pages 2679-2683. IEEE, 2014.
- [25] Niklas Mevenkamp et Benjamin Berkels. Variational multi-phase segmentation using high-dimensional local features (ségrégation variationnelle multiphase à l'aide de caractéristiques locales de haute dimension). Dans *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1-9. IEEE, 2016.
- [26] David Bryant Mumford et Jayant Shah. Optimal approximations by piecewise smooth functions and associated variational problems (Approximations optimales par des fonctions lisses par morceaux et problèmes variationnels associés). *Communications on pure and applied mathematics*, 1989.
- [27] Sijie Niu, Qiang Chen, Luis De Sisternes, Zexuan Ji, Zeming Zhou et Daniel L Rubin. Modèle de contour actif robuste basé sur la région de bruit via un facteur de similarité local pour la segmentation d'images. *Pattern Recognition*, 61:104-119, 2017.
- [28] Mangal Prakash, Tim-Oliver Buchholz, Manan Lalit, Pavel Tomancak, Florian Jug et Alexander Krull. Leveraging self-supervised denoising for image segmentation. In *2020 IEEE 17th international symposium on biomedical imaging (ISBI)*, pages 428-432. IEEE, 2020.
- [29] Ronny Ramlau et Wolfgang Ring. A Mumford-Shah level-set approach for the inversion and segmentation of X-ray tomography data. *Journal of Computational Physics*, 221(2):539-557, 2007.
- [30] Otmar Scherzer, Markus Grasmair, Harald Grossauer, Markus Haltmeier et Frank Lenzen. *Variational methods in imaging*, volume 167. Springer, 2009.

- [31] Paul Tseng. Convergence of a block coordinate descent method for nondifferentiable minimization. *Journal of optimization theory and applications*, 109(3):475, 2001.
- [32] Dmitry Ulyanov, Andrea Vedaldi et Victor Lempitsky. Deep image prior. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 9446-9454, 2018.
- [33] Stéfan van der Walt, Johannes L. Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D. Warner, Neil Yager, Emmanuelle Gouillart, Tony Yu, et les contributeurs de scikit-image. scikit-image : image processing in Python. *PeerJ*, 2:e453, 6 2014.
- [34] Romain Yildizoglu, Jean-François Aujol, et Nicolas Papadakis. Contours actifs sans jeux de niveaux. In *2012 19th IEEE International Conference on Image Processing*, pages 2549-2552. IEEE, 2012.
- [35] Tianming Zhan, Jun Zhang, Liang Xiao, Yunjie Chen et Zhihui Wei. An improved variational level set method for MR image segmentation and bias field correction. *Magnetic Resonance Imaging*, 31(3):439-447, 2013.
- [36] Kaihua Zhang, Lei Zhang, Huihui Song et Wengang Zhou. Active contours with selective local or global segmentation : a new formulation and level set method. *Image and Vision computing*, 28(4):668-676, 2010.
- [37] Maria Zontak et Michal Irani. Statistiques internes d'une image naturelle unique. In *CVPR 2011*, pages 977-984. IEEE, 2011.