

Algorithmes proximaux rapides déroulés pour l’analyse d’images fractales homogènes par morceaux.

Hoang T.V LE¹, Barbara PASCAL², Nelly PUSTELNIK^{1,3}, Marion FOARE^{4,5}, Patrice ABRY¹

¹CNRS, ENS de Lyon, Laboratoire de Physique, Lyon, France.

²Univ. Lille, CNRS, Centrale Lille, UMR 9189 CRISTAL, F-59000 Lille, France.

³ISPGROUP/ICTEAM, UCLouvain, Belgium.

⁴Univ Lyon, Ens de Lyon, CNRS, INRIA, LIP, Lyon, France.

⁵ CPE Lyon, Villeurbanne, France.

hoang.le@ens-lyon.fr, barbara.pascal@univ-lille.fr, nelly.pustelnik@ens-lyon.fr,
marion.foare@ens-lyon.fr, patrice.abry@ens-lyon.fr

Résumé – Nous proposons deux architectures déroulées de réseaux profonds construites à partir des algorithmes proximaux FISTA et Chambolle-Pock afin d’estimer la régularité locale dans les images fractales homogènes par morceaux. Ces deux réseaux, reposant sur un compromis entre approche variationnelle non-supervisée standard par minimisation d’une fonction de coût convexe non-lisse et approche supervisée deep-learning « boîte noire », ont des performances proches entre elles et conduisent à de bien meilleures performances d’estimation de régularité locale et également de détection de contours entre des changements de régularité locale.

Abstract – We propose two unrolled deep networks architectures built from the FISTA and Chambolle-Pock proximal algorithms to estimate local regularity in piecewise homogeneous fractal images. These two networks, built from a trade-off between a standard unsupervised variational approach by minimizing a non-smooth convex cost function and a deep-learning “black box” supervised approach, have similar performances and lead to much better performance in local regularity estimation and also in edge detection between local regularity changes.

1 Introduction

Contexte. L’autosimilarité est un outil dont l’intérêt a été montré dans de nombreux domaines de l’analyse des signaux et des images et plus particulièrement dans les applications biomédicales, comme la classification de signaux cardiaques fœtaux [1] ou la détection précoce de cancer à partir d’imagerie médicale [2, 3]. Une estimation précise de l’exposant d’échelle est en effet utile dans l’obtention de diagnostics fiables. Pour être précise, celle-ci doit cependant être réalisée sur un domaine suffisamment étendu pour réduire la variance d’estimation [4], une limitation importante pour une estimation plus locale, requise par exemple en segmentation. Cette estimation locale de l’exposant d’échelle, envisagé comme une mesure de régularité locale, constitue un enjeu appliqué majeur, revisité ici à l’aide d’algorithmes d’apprentissage profonds déroulés.

Etat-de-l’art. Pour améliorer cette estimation locale, une première approche consiste à combiner la procédure de régression de quantités multiéchelles locales contre les échelles, par minimisation des moindres carrés, avec introduction d’*a priori* sur le comportement spatial de ces descripteurs (e.g., constance ou linéarité par morceaux,...) conduisant à la minimisation d’un critère fortement convexe non-lisse, résolu efficacement à l’aide d’algorithmes proximaux [5, 6]. Ceux-ci peuvent être combinés à une procédure de type SUGAR pour l’estimation des hyperparamètres de régularisation [7, 8]. D’autres approches reposent sur des stratégies

de régularisation bayésienne [9]. Il a également été envisagé de créer une base de données d’apprentissage proche du modèle *a priori*, puis d’apprendre un réseau de neurones profond que l’on qualifiera de « boîte-noire », la construction n’étant guidée par aucune connaissance du modèle [10]. Pour la segmentation d’images texturées, supposées caractérisées par leurs seules propriétés statistiques à petites échelles, une comparaison entre approches variationnelles, reposant sur la régularité locale [11], et deep-learning, construite à partir du réseau FCNN, a été menée dans [10]. Elle met en évidence que le réseau FCNN supervisé se compare favorablement à une approche variationnelle non-supervisée, mais avec une robustesse moindre et une détection d’interfaces de moins bonne qualité.

Dans la littérature dédiée aux réseaux de neurones profonds, on rencontre de plus en plus le concept d’algorithmes d’apprentissage profonds déroulés [12], permettant d’intégrer des informations de modèle dans la construction du réseau, que nous proposons d’explorer dans cette étude.

Contributions. Dans cette étude, nous proposons de nous concentrer sur la tâche d’analyse de textures à travers l’estimation de la régularité locale, lorsque celle-ci est constante par morceaux à travers une image. Nous proposons d’adapter deux réseaux profonds déroulés que nous avons récemment développés dans le contexte du débruitage d’images [13] à celui de l’estimation de la régularité locale. Les deux réseaux proposés

sont : DFH (pour Deep Fista Hloc) et DSH (pour Deep Strong convexity Chambolle-Pock Hloc). Ces réseaux reposent sur une version déroulée des algorithmes FISTA et Chambolle-Pock avec stratégie d'accélération. Les performances obtenues avec ces réseaux sont comparées à une approche réseaux de neurones profonds de type « boîte-noire » concurrentielle à ces réseaux dans le contexte du débruitage d'images. Nous comparons également les performances avec les résultats obtenus via une approche variationnelle plus standard non-supervisée [14]. **Plan.** La section 2 rappelle le concept de régularité locale et les procédures standards d'estimation. La section 3 décrit les réseaux DFH et DSH. La section 4 fournit des performances de débruitage et de segmentation systématiques permettant de comparer les réseaux proposés aux méthodes de l'état-de-l'art.

2 Estimation locale de l'autosimilarité

Régularité locale. La *régularité locale* est une quantité mathématique qui peut être estimée par régression de quantités multi-échelles (e.g., le logarithme de la valeur absolue des coefficients d'ondelettes ou des coefficients dominants [4, 6, 8]). La régularité locale autour de la position n est mesurée par l'exposant dit de Hölder h_n tel que si h_n est proche de 0 (resp. 1), l'image est localement très irrégulière (resp. un champ localement lisse). Formellement, si on note $\bar{x} = (\bar{x}_n)_{1 \leq n \leq N} \in \mathbb{R}^N$ l'image que l'on cherche à analyser, et si on note $\mathcal{L}_{j,n}$ une quantité multi-échelle associée à l'image \bar{x} définie pour les échelles $j \in \{1, \dots, J\}$ et la position n , la régularité locale est reliée à cette quantité par la relation suivante :

$$\mathcal{L}_{j,n} \simeq \eta_n 2^{jh_n}, \quad (1)$$

quand $2^j \rightarrow 0$ où η_n est proportionnelle à la variance locale de \bar{x} à la position n . L'estimation locale de h_n peut être obtenue par régression linéaire à travers les échelles :

$$\hat{h}_n^{(RL)} = \sum_j w_{j,n} \log_2 \mathcal{L}_{j,n} \quad (2)$$

où $(w_{j,n})_{j,n}$ modélisent les poids de régression à l'échelle j et position n . On parle d'estimation non biaisée lorsque $\sum_j w_{j,n} \equiv 0$ et $\sum_j j w_{j,n} \equiv 1$ [4].

Estimation avec a priori spatial constant par morceaux
Dans les études dédiées à l'estimation de la régularité locale, il est usuel de considérer un champ homogène. L'estimation se fait alors en moyennant les valeurs des estimées $\hat{h}_n^{(RL)}$ obtenues en chaque position n de l'image. Afin de procéder à une estimation locale tout en réduisant la variance d'estimation, [5, 6] ont proposé d'imposer un a priori spatial de type constant par morceaux, conduisant au problème de minimisation suivant :

$$\hat{h}^{(TV)} = \arg \min_h \frac{1}{2} \|h - \hat{h}^{(RL)}\|_2^2 + \lambda \|Dh\|_{\bullet}. \quad (3)$$

où D modélise l'opérateur de différence finie et où $\|\cdot\|_{\bullet}$ modélise une norme imposant un a priori de parcimonie, usuellement $\|\cdot\|_{\bullet} = \|\cdot\|_{1,2}$ lors d'une pénalisation de type variation totale. Il s'agit d'un problème d'optimisation fortement convexe non-lisse, permettant d'estimer

$\hat{h}^{(TV)}$ efficacement par des approches proximales [15, 16]. Notre étude se concentre sur deux schémas algorithmiques rapides (convergence en $O(1/k^2)$ sur la fonctionnelle) présentés dans [17] pour l'analyse de régularité locale : Fast Iterative Soft Thresholding Algorithm (FISTA) et Chambolle Pock avec forte convexité (ScCP). Une succession de K iterations de ces schémas algorithmiques est respectivement notée $f^{\text{FISTA}}(z; D, D^{\top}, \lambda, \|\cdot\|_{\bullet}, \tau, \alpha, K)$ et $f^{\text{ScCP}}(z; D, D^{\top}, \lambda, \|\cdot\|_{\bullet}, \tau, \sigma, \alpha, K)$ permettant de mettre en évidence la dépendance des itérations en fonction des paramètres du problème de minimisation (3) et des pas de descente des stratégies algorithmiques (τ, α pour FISTA et τ, σ, α pour ScCP).

3 Algorithmes déroulés : DFH et DSH

À partir d'un ensemble d'apprentissage $\mathcal{S} = \{(\bar{h}_s, \hat{h}_s^{(RL)})\}_{s=1, \dots, I}$ où \bar{h}_s désigne la régularité locale « vraie » et $\hat{h}_s^{(RL)}$ l'estimation de la régularité locale par régression linéaire (2) à partir de la transformée multi-échelle de l'image texturée x_s . Notre objectif est d'apprendre les paramètres Θ d'un réseau f_{Θ} , afin de minimiser l'erreur empirique :

$$\text{minimiser}_{\Theta} E(\Theta) := \frac{1}{I} \sum_{s=1}^I \|\bar{h}_s - f_{\Theta}(\hat{h}_s^{(RL)})\|_2^2. \quad (4)$$

Les réseaux de neurones profonds standards formés de K couches se formulent, pour tout $h \in \mathbb{R}^N$,

$$f_{\Theta}(h) = \eta^{[K]} \left(W^{[K]} \dots \eta^{[1]} \left(W^{[1]} h + b^{[1]} \right) \dots + b^{[K]} \right), \quad (5)$$

où, pour chaque $k \in \{1, \dots, K\}$, $W^{[k]}$ désigne une transformation linéaire telle que des convolutions ou décimations pour réduire le nombre de paramètres, $b^{[k]}$ un vecteur de biais, et une fonction d'activation non linéaire $\eta^{[k]}$.

Nous proposons deux réseaux profonds déroulés reposant sur la structure algorithmique de FISTA et ScCP pour la minimisation de (3) reposant sur le lien entre la fonction d'activation HardTanh et l'opérateur proximal de la fonction conjuguée de la norme ℓ_1 [13, 18].

Architecture 1. Le réseau d'apprentissage DFH pour l'analyse de régularité locale est défini par

$$f_{\Theta}(h) := f^{\text{FISTA}}(h; D_1^{[k]}, D_2^{[k]}, 1, \|\cdot\|_1, 1, \alpha_k, K)$$

où $D_1^{[k]} \in \mathbb{R}^{|\mathbb{F}| \times N}$ et $D_2^{[k]} \in \mathbb{R}^{N \times |\mathbb{F}|}$ sont associés à une concaténation de $|\mathbb{F}|$ filtres et où $\Theta = \{D_1^{[k]}, D_2^{[k]}, \alpha_k\}_k$. Plus précisément, l'architecture est, pour tout $k \in \{2, \dots, K-1\}$:

$$\begin{cases} W^{[1]} = \begin{bmatrix} D_1^{[1]} \\ (\text{Id}_{|\mathbb{F}|} - D_1^{[1]} D_2^{[1]}) D_1^{[1]} \end{bmatrix}, \\ b^{[1]} = \begin{bmatrix} 0 \\ D_1^{[1]} \hat{h}^{(RL)} \end{bmatrix}, \eta^{[1]} = \begin{cases} \text{Id}_{|\mathbb{F}|} \\ \text{HardTanh}_{\lambda} \end{cases}, \\ W^{[k]} = \begin{bmatrix} 0 & \text{Id}_{|\mathbb{F}|} \\ -\alpha_{k-1} (\text{Id}_{|\mathbb{F}|} - D_1^{[k]} D_2^{[k]}) & (1 + \alpha_{k-1}) (\text{Id}_{|\mathbb{F}|} - D_1^{[k]} D_2^{[k]}) \end{bmatrix}, \\ b^{[k]} = \begin{bmatrix} 0 \\ D_1^{[k]} \hat{h}^{(RL)} \end{bmatrix}, \eta^{[k]} = \begin{cases} \text{Id}_{|\mathbb{F}|} \\ \text{HardTanh}_{\lambda} \end{cases}, \\ W^{[K]} = \begin{bmatrix} 0 & -D_2^{[K]} \end{bmatrix}, b^{[K]} = \hat{h}^{(RL)}, \eta^{[K]} = \text{Id}_N. \end{cases}$$

Architecture 2. Le réseau d'apprentissage DSH pour l'analyse de régularité locale est défini par

$$f_{\Theta}(\mathbf{h}) := f^{\text{FISTA}}(\mathbf{h}; \mathbf{D}_1^{[k]}, \mathbf{D}_2^{[k]}, 1, \|\cdot\|_1, 1, \sigma_k, \alpha_k, K)$$

où $\mathbf{D}_1^{[k]} \in \mathbb{R}^{|\mathbb{F}| \times N}$ et $\mathbf{D}_2^{[k]} \in \mathbb{R}^{N \times |\mathbb{F}|}$ sont associés à une concaténation de $|\mathbb{F}|$ filtres et où $\Theta = \{\mathbf{D}_1^{[k]}, \mathbf{D}_2^{[k]}, \sigma_k, \alpha_k\}_{1 \leq k \leq K}$. Plus précisément, l'architecture est, pour chaque $k \in \{2, \dots, K-1\}$:

$$\begin{cases} \mathbf{W}^{[1]} = \begin{bmatrix} \text{Id}_N \\ 2\mathbf{D}_1^{[1]} \end{bmatrix}, \mathbf{b}^{[1]} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \eta^{[1]} = \begin{cases} \text{Id}_N \\ \text{HardTanh}_{\lambda} \end{cases}, \\ \mathbf{W}^{[k]} = \begin{bmatrix} \frac{1}{1+\sigma_{k-1}} & -\frac{\sigma_{k-1}}{1+\sigma_{k-1}} \mathbf{D}_2^{[k-1]} \\ \frac{1+\alpha_k}{1+\sigma_{k-1}} \mathbf{D}_1^{[k]} - \alpha_k \mathbf{D}_1^{[k]} & \text{Id}_{|\mathbb{F}|} - \frac{(1+\alpha_k)\sigma_{k-1}}{1+\sigma_{k-1}} \mathbf{D}_1^{[k]} \mathbf{D}_2^{[k-1]} \end{bmatrix}, \\ \mathbf{b}^{[k]} = \begin{bmatrix} \frac{\sigma_{k-1}}{1+\sigma_{k-1}} \widehat{\mathbf{h}}^{(\text{RL})} \\ \frac{(1+\alpha_k)\sigma_{k-1}}{1+\sigma_{k-1}} \mathbf{D}_1^{[k]} \widehat{\mathbf{h}}^{(\text{RL})} \end{bmatrix}, \eta^{[k]} = \begin{cases} \text{Id}_N \\ \text{HardTanh}_{\lambda} \end{cases}, \\ \mathbf{W}^{[K]} = \begin{bmatrix} \text{Id}_N & 0 \end{bmatrix}, \mathbf{b}^{[K]} = 0, \eta^{[K]} = \text{Id}_N. \end{cases}$$

Dans un contexte d'apprentissage profond, les paramètres $(\mathbf{D}_1^{[k]}, \mathbf{D}_2^{[k]}, \sigma_k, \alpha_k)_{k \in \{1, \dots, K\}}$ sont appris. Mais notons qu'en choisissant $\mathbf{D}_1^{[k]} = \lambda \mathbf{D}$, $\mathbf{D}_2^{[k]} = \tau \lambda \mathbf{D}^{\top}$, $\|\cdot\|_{\bullet} = \|\cdot\|_1$, et des choix spécifiques sur τ_k , σ_k et α_k , ces deux architectures permettent de construire une suite convergente vers (3).

4 Résultats expérimentaux

Données. Pour évaluer les performances de DFH et DSH, nous créons une base de données d'images texturées \bar{x}_s générées à partir d'un processus fractal homogène par morceaux décrit dans [11] permettant d'assembler Q textures fractales. Chaque texture est un champ gaussien stationnaire dont la structure de covariance est pleinement définie par sa variance Σ et son paramètre fractal H . Une illustration de texture générée \bar{x}_s et la carte des valeurs de régularité locale « vraies » associées \bar{h}_s sont représentées Figure 2(a)-(b). En suivant la procédure décrite dans [10], trois bases de données sont générées : un jeu de données d'apprentissage composé de 2000 images avec $Q = 2$ régions, $H_1 = 0.5$, $\Sigma_1^2 = 0.6$, $H_2 = 0.8$, $\Sigma_2^2 = 1.1$ (config I), un jeu de données de tests composé de 100 images avec les mêmes paramètres, un jeu de données de tests de 100 images avec $H_1 = 0.5$, $\Sigma_1^2 = 0.33$, $H_2 = 0.65$, $\Sigma_2^2 = 1$ (config II).

Paramètres d'apprentissage– DFH et DSH sont entraînés sous Pytorch avec ADAM [21] sur 500 itérations, un *batch* de taille 10 et un taux d'apprentissage de $1e^{-4}$. Deux configurations de réseaux sont considérées, variant le nombre de filtres $|\mathbb{F}|$ et le nombre de couches, avec 5×10^3 coefficients (resp. 3×10^4) correspondant à $K = 13$ et $|\mathbb{F}| = 21$ (resp. $K = 45$ et $|\mathbb{F}| = 37$).

Comparaisons– Nous comparons les performances de DFH et DSH avec une méthode standard non supervisée variationnelle minimisant (3) lorsque $\|\cdot\|_{\bullet} = \|\cdot\|_{1,2}$ et λ est choisi automatiquement par une stratégie SUGAR décrite dans [11]. Une comparaison avec la méthode de référence DnCNN [19], qui s'avère concurrentielle aux réseaux proposés Deep-FISTA-GD et Deep-ScCP-GD ayant la même structure que DFH et DSH mais considérée dans un contexte de débruitage d'images [13],

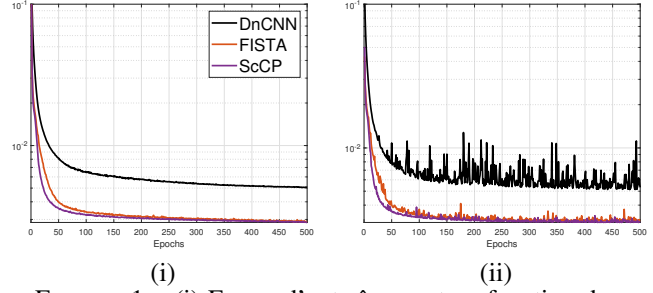


FIGURE 1 – (i) Erreur d'entraînement en fonction du nombre d'itérations sur la config I (ii) Erreur d'estimation en fonction du nombre d'itérations sur la base de test de la Config.I

est également présentée. Les nombres de couches et de paramètres du réseau DnCNN sont choisis de façon à obtenir un nombre total de paramètres équivalent à celui considéré avec les stratégies proposées DFH et DSH (i.e. $K = 9$ ou $K = 10$). **Performance d'estimation et de segmentation** – La Fig. 1 illustre un meilleur comportement de nos architectures tant en terme d'erreur d'apprentissage qu'en termes d'erreur de test comparé à une approche standard DnCNN. De plus les réseaux proposés sont plus stables (cf. oscillations sur la courbe de l'erreur calculée sur les données de test). Le Tableau 1 (ligne 1) permet quant à lui de confirmer le gain de DFH et DSH par rapport à une approche non supervisée TV dans un contexte où l'apprentissage et le test se font sur une base de données avec les mêmes caractéristiques (Config I) et également par rapport à DnCNN. Dans le cas où le test se fait sur une configuration différente que celle de l'apprentissage, les performances en termes d'erreur d'estimation restent meilleures avec les approches supervisées (cf. Tableau 1- ligne 2). Le réseau proposé est une architecture pour effectuer une tâche d'estimation, cependant nous regardons les performances obtenues en terme de segmentation après avoir effectué un k-means sur les estimées. Les performances restent systématiquement meilleures que DnCNN cependant lorsque le réseau est évalué sur la Config 2, la procédure non-supervisée permet d'obtenir de meilleures performances. Une illustration des performances d'estimation et segmentation sont présentées sur les Figs. 2.

5 Conclusion

Nous avons construit deux architectures de réseaux profonds déroulés pour l'estimation de la régularité locale dans des images texturées en dépliant/déroulant deux schémas algorithmiques proximaux accélérés (FISTA et SCCP) pour la minimisation d'une fonction fortement convexe non lisse. Les expériences numériques complètent une étude préliminaire menée dans [13] mettant en évidence qu'un réseau de neurones déroulé se compare favorablement à une approche variationnelle plus classique non supervisée et à un réseau de neurones de type « boîte noire » avec de très bons résultats en terme d'estimation et une détection d'interfaces de qualité.

Références

- [1] J. Spilka, and J. Frecon, and R. Leonarduzzi, and N. Pustelnik, and P. Abry, and M. Doret. *Sparse Support Vector Machine for Intrapartum*

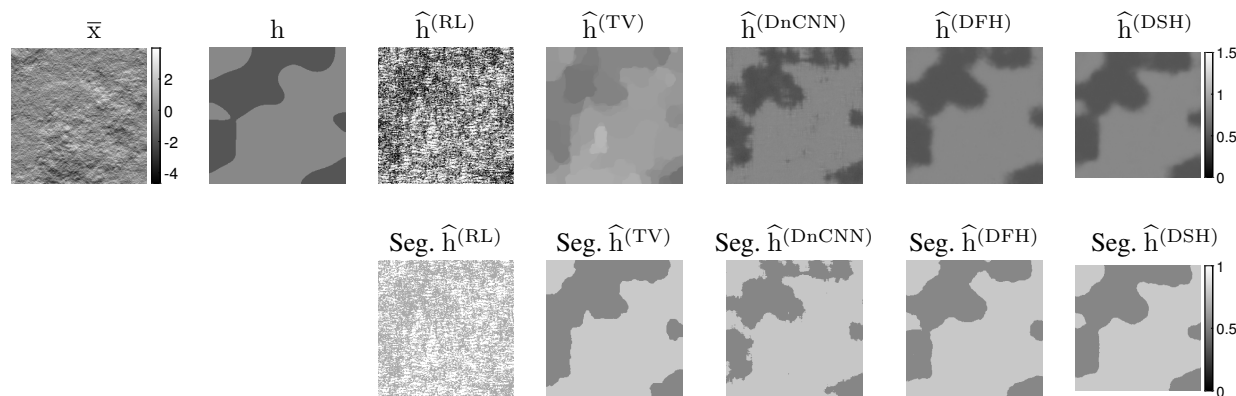


FIGURE 2 – Résultats d’estimation (ligne 1) et de segmentation (ligne 2) sur une image de texture de la Config I. Configuration où les méthodes supervisées (DnCNN, DFH et DSH avec 3×10^4 poids) sont entraînés sur la Confi I et évalué sur la Confi I. Segmentation obtenue par k-means des estimées \hat{h} .

	TV-SUGAR	DnCNN 5.10^3	DnCNN 3.10^4	DFH 5.10^3	DFH 3.10^4	DSH 5.10^3	DSH 3.10^4
Erreur d’estimation							
Test config I	0.339+/-0.048	0.113+/-0.011	0.104+/-0.010	0.073+/-0.008	0.069+/-0.007	0.072+/-0.007	0.069+/-0.007
Test config II	0.306+/-0.029	0.145+/-0.014	0.130+/-0.012	0.115+/-0.011	0.120+/-0.012	0.116+/-0.012	0.119+/-0.012
Score de classification							
Test config I	83.7+/-2.68	81.0+/-1.73	86.2+/-1.29	94.8+/-0.45	95.2+/-0.41	94.8+/-0.42	95.3+/-0.41
Test config II	73.2+/-2.51	68.4+/-1.36	67.9+/-1.45	70.7+/-2.65	70.9+/-2.67	68.3+/-2.74	70.6+/-2.71

TABLE 1 – Comparaison en termes d’erreur normalisée ou de précision de segmentation entre différentes méthode.

- Fetal Heart Rate Classification* IEEE Journal of Biomedical and Health Informatics, vol. 21, no. 3, pp. 664–671, 2017.
- [2] Z. Marin, and K. A. Batchelder, and B. C. Toner, and L. Guimond, and E. Gerasimova-Chechkina, and Harrow, A. R. and A. Arneodo, and A. Khalil. *Mammographic evidence of microenvironment changes in tumorous breasts*. Medical Physics, vol. 44, no. 4, pp. 1324–1336, 2017.
- [3] L. Vidyaratne and M. Alam and Z. Shboul and K. M. Iftekharuddin. *Deep learning and texture-based semantic label fusion for brain tumor segmentation*. Medical Imaging 2018 : Computer-Aided Diagnosis, International Society for Optics and Photonics, 105750D, 2018.
- [4] H. Wendt and P. Abry and S. Jaffard. *Bootstrap for Empirical Multifractional Analysis*. IEEE Signal Process. Mag., vol. 24, no. 4, pp. 38–48, 2007.
- [5] C. Naornita, A. Isar, and J. D. B. Nelson. *Regularised, semi-local hurst estimation via generalised lasso and dual-tree complex wavelets*. Proc. IEEE Int. Conf. Image Process, Paris, France, Oct. 2014, pp. 2689–2693, 2014.
- [6] N. Pustelnik, H. Wendt, P. Abry, N. Dobigeon, *Combining local regularity estimation and total variation optimization for scale-free texture segmentation*. IEEE Trans. Comput. Imaging, vol. 2, no. 4, pp. 468–479, 2016.
- [7] C.-A. Deledalle, S. Vaïter, J. Fadili, and G. Peyré. *Stein unbiased gradient estimator of the risk (SUGAR) for multiple parameter selection*. SIAM J. Imaging Sci., 7(4) :2448–2487, 2014
- [8] B. Pascal, S. Vaïter, N. Pustelnik, and P. Abry. *Automated data-driven selection of the hyperparameters for Total-Variation based texture segmentation*. Journal of Mathematical Imaging and Vision, pp 1-30, 2021
- [9] H. Wendt and S. Combexelle and Y. Altmann and J.-Y. Tournet and S. McLaughlin and P. Abry *Multifractional analysis of multivariate images using gamma Markov random field priors* SIAM J. on Imaging Sciences (SIIMS), 11(2), pp.1294-1316, 2018.
- [10] B. Pascal, V. Mauduit, P. Abry, and N. Pustelnik, *Scale-free Texture Segmentation : Expert Feature-based versus Deep Learning strategies*. Proc. EUSIPCO, Amsterdam, The Netherlands, Jan. 18-22, 2021.
- [11] B. Pascal, and N. Pustelnik, and P. Abry. *Strongly Convex Optimization for Joint Fractal Feature Estimation and Texture Segmentation*. Appl. Comp. Harm. Analysis, vol. 54, pp. 303–322, 2021.
- [12] K. Gregor and Y. LeCun. *Learning fast approximations of sparse coding*. Proc. International Conference on Machine Learning, Haifa, Israel, pp. 399–406, Jun. 21-24 2010.
- [13] H.T.V. Le, N. Pustelnik, and M. Foare, *The faster proximal algorithm, the better unfolded deep learning architecture ? The study case of image denoising*. submitted EUSIPCO, 2022.
- [14] B. Pascal, and N. Pustelnik, and P. Abry, J.-C. Géminard, V. Vidal. *Parameter-free and fast nonlinear piecewise filtering : application to experimental physics*. Annals of Telecommunications, 7vol. 5, no. 11, 655–671, 2020
- [15] H. H. Bauschke and P. L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. New York : Springer, 2017.
- [16] P. L. Combettes and J.-C. Pesquet *Proximal splitting methods in signal processing*. In Fixed-point algorithms for inverse problems in science and engineering, Springer, New York, NY, pp. 185-212, 2011.
- [17] B. Pascal, and N. Pustelnik, and P. Abry, and M. Serres, and V. Vidal. *Joint Estimation of Local variance and Local regularity for texture segmentation. Application to multiphase flow characterization*. Proc. IEEE Int. Conf. Image Process., Athens, Greece, pp. 2092–2096, 2018.
- [18] P. L. Combettes and J.-C. Pesquet. *Lipschitz certificates for layered network structures driven by averaged activation operators*, SIAM J. Math. Data Sci., vol. 2, no. 2, pp. 529–557, 2020.
- [19] Zhang K., Zuo W., Chen Y., Meng, D. and Zhang L, *Beyond a Gaussian denoiser : Residual learning of deep CNN for image denoising*. IEEE. j-ieee-tip, vol. 26, pp. 3142–3155, 2017.
- [20] Chambolle, A. and Pock, T., *A first-order primal-dual algorithm for convex problems with applications to imaging*. Journal of mathematical imaging and vision, 40(1), pp.120-145, 2011.
- [21] Kingma D. P. and Ba J. *Adam : A method for stochastic optimization*, arXiv preprint arXiv :1412.6980, 2014