

# Fusion d'images non locale préservant la géométrie

Anaïs GASTINEAU<sup>1,2</sup>, Jean-François AUJOL<sup>1</sup>, Yannick BERTHOUMIEU<sup>2</sup>, Christian GERMAIN<sup>2</sup>

<sup>1</sup>Univ. Bordeaux, Bordeaux INP, CNRS, IMB, UMR 5251, F-33400 Talence, France

<sup>2</sup>Univ. Bordeaux, CNRS, IMS, UMR 5218, F-33400 Talence, France

anaïs.gastineau@math.u-bordeaux.fr, jean-francois.aujol@math.u-bordeaux.fr  
yannick.berthoumieu@ims-bordeaux.fr, christian.germain@ims-bordeaux.fr

**Résumé** – Le problème de pansharpening se situe dans la conception d'un procédé permettant de fusionner une image panchromatique haute résolution avec une image multispectrale basse résolution afin de générer une image multispectrale haute résolution. Dans le cadre de la recherche d'une solution par approche régularisée, nous étudions, dans cette communication, un terme d'attache aux données non local pour le problème de pansharpening fondé sur un terme exploitant la géométrie de l'image. Ce terme est basé sur l'hypothèse que l'on peut aligner le gradient de l'image multispectrale haute résolution que l'on cherche en chaque point en regardant la direction des vecteurs donnés par le gradient de l'image panchromatique au voisinage de ce point. Nous explicitons le gradient de ce terme et effectuons des tests sur des données simulées et des données satellites.

**Abstract** – The pansharpening problem allows to fuse a high resolution panchromatic image with a low resolution multispectral image in order to obtain a high resolution multispectral image. We propose, in this paper, a non local variational term for pansharpening based on a term exploiting the geometry of the image. This term is based on the assumption that the gradient of the high resolution multispectral image can be aligned at each point with the gradient of the panchromatic image by regarding the direction of the vectors given by the gradient in its neighborhood. We compute the gradient of the proposed term and we perform tests on simulated images and satellite images.

## 1 Introduction

Les satellites sont très importants pour l'observation de la Terre. Ils embarquent divers capteurs permettant l'acquisition d'images à des échelles variées selon l'objectif du satellite. En télédétection, la résolution spatiale est exprimée en fonction de la taille du terrain capturée par un pixel. Cela affecte alors la reproduction de détails présents dans la scène observée. En effet, quand la taille d'un pixel est réduite, plus de détails sont capturés et ainsi préservés dans l'image. La résolution spectrale est donnée à la fois par le nombre de canaux ou de capteurs et la bande passante du signal capturée par les capteurs produisant l'image. Plus la bande passante est étroite ou le nombre de canaux élevé, plus la résolution est élevée. Certains capteurs sont capables d'intégrer l'énergie à l'entrée du capteur sur un large spectre de longueur d'onde. Ils offrent peu d'information spectrale mais en contrepartie proposent une résolution spatiale plus élevée. Nous appelons ce type d'image, images panchromatiques, par opposition aux images dites multispectrales ou hyperspectrales, acquises par des capteurs sur des bandes beaucoup plus étroites du spectre, qui offrent une résolution spectrale élevée et une faible résolution spatiale. Comme la haute résolution spectrale provient des images multispectrales et la haute résolution spatiale des images panchromatiques, il est tentant d'essayer de générer des images multispec-

trales possédant la résolution spatiale disponible sur les images panchromatiques et la résolution spectrale disponible sur les images multispectrales. Le pansharpening est un moyen visant à répondre à cette ambition.

Le problème de pansharpening peut être formulé comme le fait de construire une image  $u$  ayant une haute résolution spatiale et spectrale à partir d'une image panchromatique  $P$  et d'une image multispectrale  $y$ . Cela revient à considérer le modèle suivant :

$$\begin{cases} y^k = SH^k u^k + B^k, & \forall k \leq N \\ P = \sum_{k \leq N} \alpha_k u^k + B \end{cases}, \quad (1)$$

où  $S$  est un opérateur de sous échantillonnage,  $H$  le noyau d'un opérateur de convolution,  $B$  un bruit gaussien de moyenne nulle et  $N$  le nombre de bandes. Exposé ainsi, le problème (1) suppose que l'image panchromatique peut être approchée par une combinaison linéaire des bandes de l'image multispectrale haute résolution que l'on cherche  $u$ . Ce modèle a été introduit par *Ballester et al.* [2] et est couramment utilisé [5, 10].

De nombreuses approches permettent d'aborder le problème de pansharpening. Elles peuvent être divisées en plusieurs catégories. Les méthodes par substitution de composantes regroupent les méthodes telles que l'ACP [7], la transformation de Gram-Schmidt [1] ou l'IHS (Intensity Hue Saturation) [4] par exemple qui utilisent une

transformation linéaire sur les données pour séparer les détails spatiaux des détails spectraux afin de substituer les détails spatiaux par l'image panchromatique. Ces méthodes sont souvent utilisées pour leur faible temps de calcul et pour leur bonne capacité à retenir les détails spatiaux. Les méthodes d'analyse multirésolution, qui regroupent les méthodes décomposant l'image en séquences de signaux ou pyramides avec un contenu d'information décroissant. Nous pouvons y associer des méthodes basées sur les ondelettes [11, 6], la pyramide de Laplace [3], etc. Les méthodes variationnelles regroupent les approches par optimisation déterministe, c'est à dire celles qui considèrent un modèle régularisé en donnant a priori sur la solution recherchée [2], [5]. Enfin, les méthodes d'apprentissage permettent de modéliser efficacement la relation entre les variables par la composition de plusieurs niveaux (convolution, etc.) [9], [12].

D'après la littérature récente, le terme de régularisation non local NLV de *Duran et al.* [5] donne des résultats de l'état de l'art. Cependant, le terme P+XS proposé par *Ballester et al.* [2] permet de préserver la géométrie de l'image panchromatique en la transférant à la solution que l'on cherche. C'est pour cela que nous proposons une version non locale du terme d'attache aux données proposé par *Ballester et al.*

La suite de cet article est organisée de la façon suivante : le paragraphe 2 décrit le modèle que nous proposons et le paragraphe 3 présente et compare les résultats que nous avons obtenus avec différentes méthodes.

## 2 Terme proposé

Le modèle P+XS est une approche introduite par *Ballester et al.* [2] qui minimise les deux termes d'attaches aux données (1) en ajoutant un terme forçant l'alignement des lignes de niveau de chaque canal multispectral de l'image que l'on cherche avec celles de l'image panchromatique. Cela revient à transférer la géométrie de l'image panchromatique à l'image multispectrale recherchée de la manière suivante :

$$\sum_{k \leq N} \int_{\Omega} |\theta^{\perp} \cdot \nabla u^k|^2 dx, \quad (2)$$

où  $\theta = \frac{\nabla P}{|\nabla P|}$ ,  $\Omega$  correspond au domaine de la solution  $u$ . Ici et dans la suite de cet article, le gradient est défini de la façon suivante :

$$\nabla u = (\partial_x u, \partial_y u).$$

Ce modèle est sensible au bruit contrairement au modèle NLV proposé par *Duran et al.* [5] qui considère un terme de régularisation non local inspiré de l'algorithme de débruitage NL-means :

$$\sum_{k \leq N} \int_{\Omega} \int_{\mathcal{N}_x} (u^k(x) - u^k(z))^2 \omega(x, z) dz dx, \quad (3)$$

où  $\mathcal{N}_x = \{z \in \Omega \text{ t.q. } |x - z| \leq r\}$  représente le voisinage de  $x$ . Pour ce modèle, les poids  $\omega$  au point  $x$  sont calculés de la façon suivante :

$$\omega(x, z) = \frac{1}{C(x)} \exp\left(-\frac{d(P(x), P(z))}{h^2}\right), \quad (4)$$

où  $C(x)$  est la constante de normalisation et  $d$  la distance entre les patchs  $P(x)$  et  $P(z)$  de l'image panchromatique,  $z \in \mathcal{N}_x$  et  $h$  est un paramètre de filtrage. Les poids décroissent quand la dissimilarité augmente : la moyenne est calculée entre des patchs de région très similaires et préserve ainsi l'image tout en réduisant les petites fluctuations, en particulier celles causées par le bruit.

D'après les différents modèles présentés précédemment, nous proposons ici un terme non local (PXSNL) combinant les idées présentes dans les deux termes précédents (2) et (3). C'est à dire que nous proposons un terme préservant la géométrie tel que celui proposé par *Ballester et al* tout en gardant le caractère non local de *Duran et al* :

$$\sum_{k \leq N} \int_{\Omega} \int_{\mathcal{N}_x} |\theta^{\perp}(z) \cdot \nabla u^k(x)|^2 \omega(x, z) dz dx. \quad (5)$$

Ce terme permet d'aligner le gradient de  $u$  au point  $x$  en prenant en compte les vecteurs au voisinage de ce point. Ainsi, comme le gradient est assez sensible au bruit, utiliser les vecteurs au voisinage permet d'atténuer le bruit.

Nous considérons deux termes d'attaches aux données pour le modèle (1) ainsi, cela revient à minimiser :

$$\begin{aligned} \underset{u}{\operatorname{argmin}} \quad & \sum_{k \leq N} \int_{\Omega} |SH^k u^k - y^k|^2 + \lambda \int_{\Omega} \left( \sum_{k \leq N} \alpha_k u^k - P \right)^2 \\ & + \mu \sum_{k \leq N} \int_{\Omega} \int_{\mathcal{N}_x} |\theta^{\perp}(z) \cdot \nabla u^k(x)|^2 \omega(x, z) dz dx. \end{aligned} \quad (6)$$

Pour cela, nous utilisons l'algorithme de descente du gradient à pas fixe. En effet, si nous considérons

$$F(u) = \int_{\Omega} \int_{\mathcal{N}_x} |\theta^{\perp}(z) \cdot \nabla u(x)|^2 \omega(x, z) dz dx, \quad (7)$$

le gradient est alors :

$$\nabla F(u) = -2 \operatorname{div} \left( \int_{\mathcal{N}_x} (\theta^{\perp}(z) \cdot \nabla u(x)) \theta^{\perp}(z) \omega(x, z) dz \right), \quad (8)$$

avec  $\operatorname{div}(v) = \partial_x v_1 + \partial_y v_2$ , où  $v = (v_1, v_2)$ .

## 3 Résultats

Dans ce paragraphe, nous avons comparé les méthodes mentionnées en simulant une image panchromatique et une image multispectrale à partir des images de la Figure 1, en suivant le modèle (1).



FIGURE 1 – Images de références utilisées pour les tests.

Image	Modèle	CC	SAM	RMSE	PSNR
	Valeurs idéales	1	0	0	max
Cactus	P+XS	<u>0.9917</u>	<u>0.0656</u>	<u>6.07</u>	<u>32.53</u>
	NLV	0.992	0.054	5.96	32.66
	PXSNL	0.9923	0.0515	5.85	32.83
	PXSNL + NLV	<b>0.9925</b>	<b>0.047</b>	<b>5.77</b>	<b>32.95</b>
Burano	P+XS	<u>0.997</u>	<u>0.086</u>	<u>5.99</u>	<u>32.7</u>
	NLV	0.9977	0.0585	5.33	33.73
	PXSNL	0.9979	0.0616	5.07	34.12
	PXSNL + NLV	<b>0.9980</b>	<b>0.0515</b>	<b>4.87</b>	<b>34.46</b>
Maisons	P+XS	<u>0.996</u>	<u>0.1257</u>	<u>5.37</u>	<u>33.59</u>
	NLV	0.9973	0.0769	4.41	35.26
	PXSNL	0.9974	0.083	4.32	35.42
	PXSNL + NLV	<b>0.9977</b>	<b>0.0645</b>	<b>4.05</b>	<b>35.99</b>
Ballons	P+XS	<u>0.995</u>	<u>0.0299</u>	<u>5.66</u>	<u>33.07</u>
	NLV	0.9977	0.0191	3.85	36.84
	PXSNL	0.998	0.0201	3.54	37.12
	PXSNL + NLV	<b>0.9985</b>	<b>0.0157</b>	<b>3.05</b>	<b>38.41</b>
Feuilles	P+XS	<u>0.9966</u>	<u>0.0750</u>	<u>6.06</u>	<u>32.81</u>
	NLV	0.9978	0.0460	4.83	34.85
	PXSNL	<b>0.9979</b>	0.0479	<b>4.71</b>	35.01
	PXSNL + NLV	0.9975	<b>0.0391</b>	5.19	<b>35.20</b>

Image	Modèle	CC	SAM	RMSE	PSNR
	Valeurs idéales	1	0	0	max
Phare	P+XS	<u>0.9936</u>	<u>0.0281</u>	<u>5.66</u>	<u>33.09</u>
	NLV	0.9953	0.0257	4.87	34.39
	PXSNL	0.9956	0.0239	4.69	34.73
	PXSNL + NLV	<b>0.996</b>	<b>0.0215</b>	<b>4.49</b>	<b>35.11</b>
Satellite4	P+XS	<u>0.9896</u>	0.0520	<u>8.09</u>	<u>29.99</u>
	NLV	0.98898	<u>0.0534</u>	8.08	30.00
	PXSNL	0.9892	0.0517	7.98	30.10
	PXSNL + NLV	<b>0.9894</b>	<b>0.0516</b>	<b>7.94</b>	<b>30.16</b>
Satellite3	P+XS	<u>0.986</u>	<u>0.0910</u>	<u>8.44</u>	<u>29.61</u>
	NLV	0.986	0.0905	8.40	29.65
	PXSNL	0.9866	0.0881	8.26	29.80
	PXSNL + NLV	<b>0.9867</b>	<b>0.0871</b>	<b>8.21</b>	<b>29.85</b>
Satellite2	P+XS	<u>0.9717</u>	<u>0.1096</u>	<u>4.89</u>	<u>34.34</u>
	NLV	0.9869	0.0692	3.34	37.63
	PXSNL	0.9854	0.0829	3.48	37.28
	PXSNL + NLV	<b>0.9890</b>	<b>0.0583</b>	<b>3.07</b>	<b>38.35</b>
Satellite1	P+XS	<u>0.9733</u>	<u>0.0980</u>	<u>5.59</u>	<u>33.18</u>
	NLV	0.9876	0.0603	4.06	35.94
	PXSNL	0.9861	0.0698	4.14	35.76
	PXSNL + NLV	<b>0.9898</b>	<b>0.0511</b>	<b>3.79</b>	<b>36.53</b>

TABLE 1 – Résultats quantitatifs obtenus avec les images de la Figure 1, les moins bons résultats sont soulignés et les meilleurs sont en gras. Nous pouvons voir que le modèle P+XS donne les moins bons résultats sur chaque image testée et que le terme PXSNL couplé avec le terme NLV donne les meilleurs résultats.

Nous avons utilisé les mesures de qualité globale PSNR et RMSE, ainsi que les mesures SAM et CC qui calculent respectivement la qualité spectrale et spatiale des images fusionnées [8]. Les paramètres  $\lambda$  et  $\mu$  de l'énergie à minimiser (6) ont été choisis afin de maximiser le PSNR.

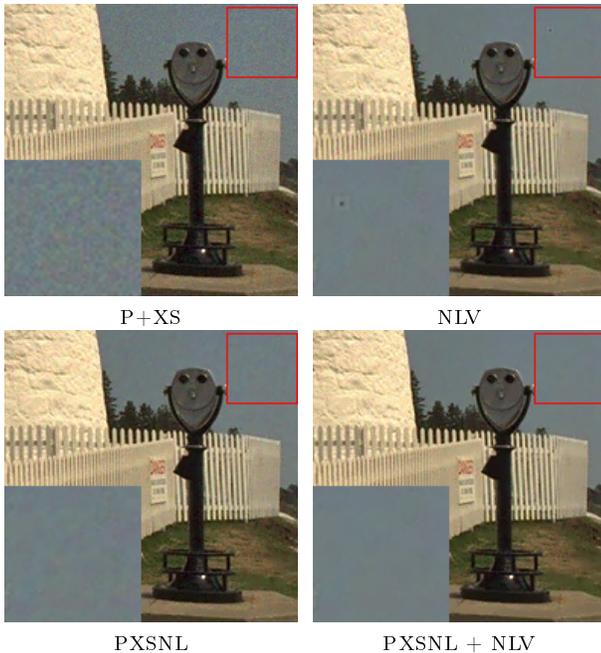


FIGURE 2 – Zoons sur les résultats obtenus avec l'image "Phare". Les termes PXSNL et NLV débruitent la solution contrairement au modèle P+XS.

Nous pouvons voir qu'avec les données simulées, le terme

PXSNL proposé comparé aux modèles P+XS et NLV, donne de meilleurs résultats quantitatifs (Table 1) pour la majorité des images testées. Pour les images où les résultats quantitatifs sont inférieurs au modèle NLV, considérer les deux termes (NLV et PXSNL) permet d'obtenir un gain de performance. De plus, nous pouvons observer que contrairement au modèle P+XS le modèle que nous proposons débruite la solution (Figure 2).

Nous avons également testé ces mêmes méthodes sur des images du satellite Pléiades. En utilisant notre modèle sur ces images, nous pouvons voir (Figure 3) qu'il est plus difficile d'observer de grandes différences entre les modèles comparés. Cependant, le modèle que nous proposons conserve mieux la structure de l'image panchromatique.

Nous avons également analysé les résultats donnés par deux méthodes de deep learning. Ces deux méthodes considèrent chacune soit un Convolutional Neural Network [9] (Pansharpening Neural Network) ou un Deep Residual Neural Network [12] (Deep Residual Pansharpening Neural Network) adaptés pour le pansharpening. Ces réseaux donnent une bonne résolution spatiale mais une résolution spectrale plus faible que les méthodes variationnelles considérées. En effet, on peut voir que les résultats obtenus avec les images Pléiades présentent des distorsions au niveau du spectre (Figure 3). Cela peut être expliqué par les données utilisées lors de l'apprentissage de ces réseaux.

## 4 Conclusion

Dans cet article, nous avons proposé un terme d'attache aux données non local permettant de transférer la géométrie de l'image panchromatique à l'image multispectrale haute résolution recherchée. De plus, le terme que nous proposons permet de débruite la solution en considérant

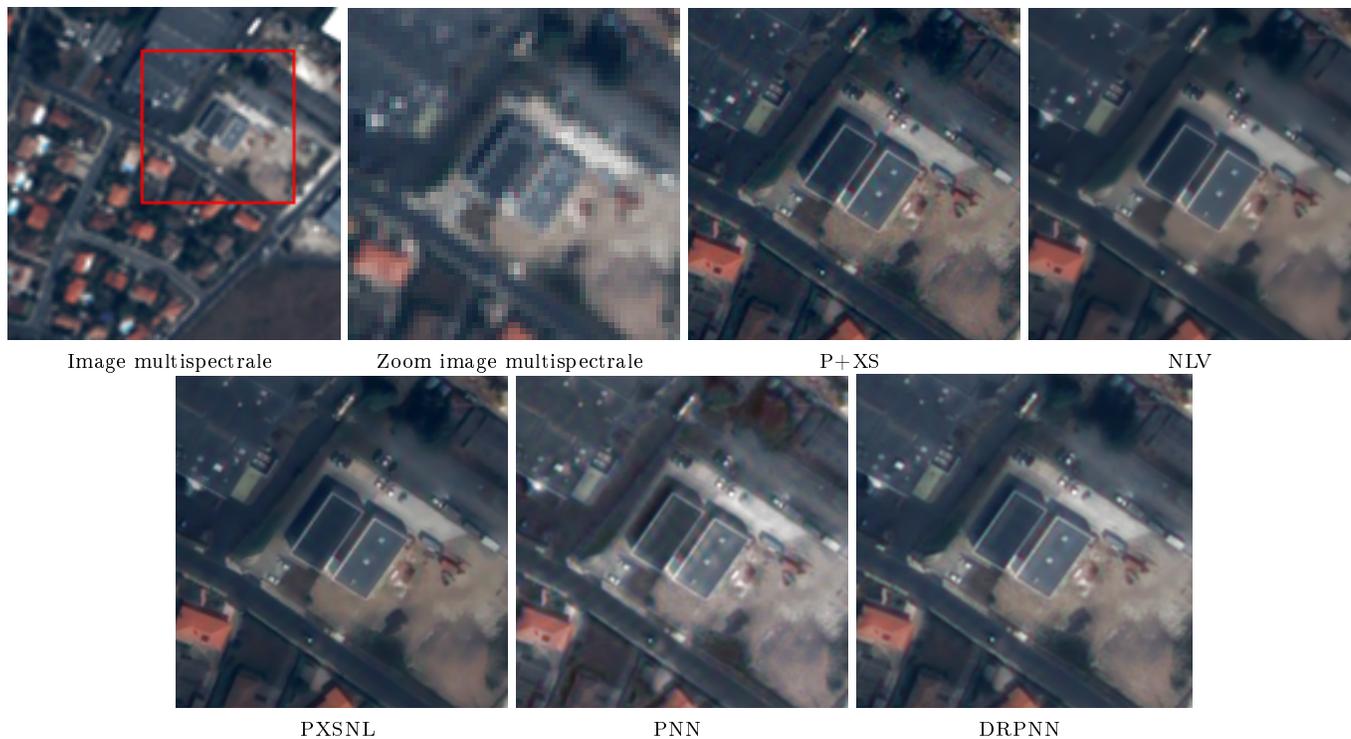


FIGURE 3 – Zooms sur des résultats obtenus avec une image Pléiades. Nous pouvons voir que le modèle que nous proposons (PXSNL) préserve mieux la géométrie que le modèle NLV, les contours sont plus nets. Les résultats obtenus avec les méthodes basées réseaux donnent une bonne résolution spatiale mais apporte un changement au niveau du spectre.

les gradients donnés par le voisinage en chaque point. Les résultats obtenus sur des données simulées et satellites montrent une amélioration visuelle et quantitative par rapport aux méthodes comparées.

## Remerciements

Les auteurs souhaitent remercier l'équipe de la recette thématique utilisateurs du programme ORFEO d'accompagnement Pléiades pour les images Pléiades (©CNES (2012,2013), distribution Airbus DS/Spot Image) utilisées dans le cadre de cette étude.

Cette étude a été menée avec le soutien financier de l'État français, géré par l'Agence Nationale de la Recherche dans le cadre du projet "Investissements d'avenir". Programme IdEx Bordeaux-Sysnum (ANR-10-IDEX-03-02).

## Références

- [1] B. Aiazzi, S. Baronti, M. Selva, and L. Alparone. Enhanced Gram-Schmidt spectral sharpening based on multivariate regression of MS and PAN data. *IEEE Inter. Conf. on Geosci. and Remote Sens. Symposium*, 2006.
- [2] C. Ballester, V. Caselles, L. Igual, J. Verdera, and B. Rougé. A Variational Model for P+XS Image Fusion. *IJCV*, vol. 69, no. 1, pp. 43-59, 2006.
- [3] P. Burt and E. Adelson. The Laplacian pyramid as a compact image code. *IEEE Transactions on communications*, vol. 31, No. 4, pp.532-540, 1983.
- [4] W. Carper, T. Lillesand, and R. Kiefer. The use of Intensity-Hue-Saturation transformations for merging SPOT panchromatic and multispectral image data. *Photogramm Eng Remote Sensing*, vol. 56, No. 4, pp. 459-467, 1990.
- [5] J. Duran, A. Buades, B. Coll, and C. Sbert. A non local variational model for pansharpening image fusion. *SIAM*, vol. 7, no. 2, pp. 761-796, 2015.
- [6] M. Gonzalez-Audicana, X. Otazu, O. Fors, and A. Seco. Comparison between Mallat's and the 'A trous' discrete wavelet transform algorithms for the fusion of multispectral and panchromatic images. *Int J Remote Sens*, vol. 26, no. 3, pp. 595-614, 2005.
- [7] M. Gonzalez-Audicana, J. Saleta, R. Garcia Catalan, and R. Garcia. Fusion of Multispectral and Panchromatic Images Using Improved IHS and PCA Mergers based on Wavelet Decomposition. *IEEE TGRS*, vol. 42, No. 6, pp.1291-1299, 2004.
- [8] L. Loncan, L. Almeida, X. Briottet, J. Bioucas-Dias, J. Chanussot, N. Dobigeon, S. Fabre, W. Liao, G. Licciardi, M. Simoes, J.-Y. Tourneret, M. Veganzones, G. Vivone, Q. Wei, and N. Yokoya. Hyperspectral Pansharpening : A review. *IEEE Geosc. and Remote Sens.*, vol. 3, no. 3, pp. 27-46, 2015.
- [9] G. Masi, D. Cozzolino, L. Verdolina, and G. Scarpa. Pansharpening by Convolutional Neural Network. *Remote Sensing*, vol. 8, no. 7, pp. 594-616, 2016.
- [10] F. Palsson, J. Sveinsson, M. Ulfarsson, and J. Benediktsson. A New Pansharpening Method Using an Explicit Image Formation Model Regularized via Total Variation. *IEEE International Geoscience and Remote Sensing Symposium*, 2012.
- [11] G. Vivone, L. Alparone, J. Chanussot, M. Dalla Mura, A. Garzelli, G. Licciardi, R. Restaino, and L. Wald. A Critical Comparison Among Pansharpening Algorithms. *IEEE TGRS*, vol. 53, no. 5, pp. 2565-2586, 2015.
- [12] Y. Wei, H. Yuan, H. Shen, and L. Zhang. Boosting the Accuracy of Multispectral Image Pansharpening by Learning a Deep Residual Network. *IEEE GRSL*, pp. 99, 2017.