Untrained Compact Neural Network Prior for High-Dimension Multispectral and Hyperspectral Data Fusion with Spectrally Varying Blurs

Dan Pineau, Françoois Orieux, Alain Abergel

Untrained Compact Neural Network Prior for High-Dimension Multispectral and Hyperspectral Data Fusion with Spectrally Varying Blurs

Dan Pineau [1,2,3], François Orieux [1,2], Alain Abergel [2,3]

***Abstract*—The information in a spectrally degraded dataset (multispectral data) and a spatially degraded one (hyperspectral data) can be processed jointly to reconstruct data with enhanced spatial and spectral resolutions. The problem of multispectral and hyperspectral (MS/HS) data fusion with spatial blurs is an ill-posed inverse problem, commonly solved by minimizing a mixed criterion containing data adequacy terms and a Tikhonov regularization term. However, such regularizations suffer from poor discriminative ability. Instead, the recently proposed Untrained Neural Network Priors (UNNP) have surpassed classical handcrafted priors in most imaging inverse problems without requiring prior learning with a labeled database. Few works have been proposed to solve the MS/HS fusion problem with UNNP, but none have explored the fusion with spectrally-varying spatial blurs. This paper aims to solve the MS/HS fusion problem with high dimension data degraded with spectrally-varying spatial blurs by relying on a simple and compact neural network architecture as prior, made possible by using a dimension reduction model. It demonstrates the superior performances of such a prior against handcrafted priors for the fusion of simulated realistic MS and HS data from the Mid-Infrared Instrument (MIRI) of the James Webb Space Telescope (JWST) in a high noise regime.**

***Index Terms*—inverse problem, deconvolution, data fusion, hyperspectral, untrained neural network, variant blur, JWST**

I. Introduction

Multispectral and hyperspectral (MS/HS) data fusion integrates the spatially rich information obtained from multispectral imaging with the spectrally detailed information captured by hyperspectral imaging. This technique enhances the quality and utility of remote sensing data, facilitating comprehensive analysis across various scientific disciplines, such as earth observation, remote sensing, and astrophysics [1].

More precisely, this study addresses the challenge of fusing high dimensional datasets characterized by spectrally varying spatial blurs spanning a broad spectral range. An example of such data is infrared astronomical data, for example obtained with the James Webb Space Telescope (JWST), highlighting the complexity and significance of the instrumental effects encountered [2]. To our knowledge, two notable studies are addressing this specific issue. The first [3] proposed to solve this inverse problem by minimizing a least squares criterion through an iterative algorithm. The second study [4] proposed an explicit and easily calculable solution to this fusion problem.

Nevertheless, the handcrafted priors used in these papers face limitations due to their low discriminative abilities [5]. Indeed, a significant portion of unnatural images also meet the criteria set by the forward models because of the ill-posedness of the inverse problem. Moreover, priors induced with Tikhonov regularizations [6] can introduce ringing effects in the reconstructed images, especially when the input data is noisy. In contrast, recent advancements in deep learning have led to significant progress in addressing inverse imaging problems [7] compared to traditional handcrafted priors. These methods include end-to-end learning-based methods, more specifically using self-attention such as [8] or invertible neural networks [9], based on directly learning the mapping from observed data to the desired output through a trainable neural network. Other works include the use of learning-based priors, where the regularization term is learned and involves a Variational Auto-Encoder (VAE) as in [10] or a denoising function [11]. Such priors can also be learned in a deep unrolling framework, as done in [12] and in [13]. However, in the context of MS/HS fusion with images from the JWST, the data available are scarce and therefore likely insufficient to effectively train the neural networks architecture used in the previously presented deep learning methods.

Nonetheless, another type of priors relies on untrained neural network, usually implemented as a diffusion model, for their observed inherent ability to better reconstruct natural images than noisy ones, as shown in [14], which proved to outperform handcrafted priors in many imaging inverse problems [15], such as denoising, inpainting, and super-resolution. These priors have thus emerged as a promising alternative to the limiting handcrafted priors and the data-demanding deep learning priors for such inverse problems, and they have already been used for solving MS/HS fusion problems [16] [17] [18]. However, the data formation processes considered in these papers do not consider any spatial blurs, and are therefore not fit for the inverse problem considered in this paper. Moreover, their proposed architectures aim to reproduce the whole, highly spatially and spectrally resolved cube reconstructed from the MS and HS datasets, which can be of high dimension for astronomical images. This could induce

Work funded by Agence Nationale de la Recherche (ANR) and Centre National d'Études Spatiales (CNES).

[1] Université Paris-Saclay, CNRS, CentraleSupélec, Laboratoire des Signaux et Systèmes, 3 rue Joliot Curie, 91190 Gif-sur-Yvette, France

[2] Institut d'Astrophysique Spatiale (IAS), CNRS, UMR 8617, Université Paris-Saclay, 91405 Orsay, France

[3] Innovative Common Laboratory For Space Spectroscopy (INCLASS), IAS, ACRI-ST, CNRS, Université Paris-Saclay, 91405 Orsay, France

a significantly high computational load and potential memory issues.

In comparison, this paper aims to solve a fusion problem with noised high dimension MS and HS data degraded with spectrally-varying spatial blurs with a simple and compact diffusion model architecture as prior and compares its performances with handcrafted priors. The reconstruction of the complete hyperspectral data is made possible by only reconstructing a compact version of this cube, introduced via the Linear Mixing Model, explained in Section II. The MS and HS data formation processes are also described in section II and are the same as the ones already given in a previous work [4]. Section III formulates the data fusion problem with a handcrafted prior and an UNNP, and presents the latter prior's architecture. Section IV shows the superior performances of using a relatively simple UNNP over handcrafted priors to fuse simulated realistic MIRI data when solving the problem under a specific, yet realistic, noise range.

II. Description of the data formation processes

A. Dimension reduction model

The reconstructed object is a three-dimensional hyperspectral datacube $\boldsymbol{x}$: two spatial dimensions, with I and J being the number of rows and columns respectively, and one spectral dimension, with L wavebands.

Instead of reconstructing the full datacube like in [16] [17] [18], we propose a linear mixing model [19] [3] [20] where the spectral content at any given location (i, j) of $\boldsymbol{x}$ is a linear combination of T known spectra $\boldsymbol{s}_t$

$$\boldsymbol{x}[i,j,l] = \sum_{t=1}^{T} \boldsymbol{a}_t[i,j]\ \boldsymbol{s}_t[l] \tag{1}$$

where $\boldsymbol{a}_t[i,j]$ is the abundance map of the spectrum $\boldsymbol{s}_t$ at the location (i, j).

This model offers several advantages for our problem. First, defining known spectra templates $\boldsymbol{s}_t$ for all wavelengths captured by the instruments enables efficient joint processing of their data due to the spectral correlations. Secondly, choosing $T \ll L$ significantly reduces the number of unknowns.

B. Multispectral data formation process

This section describes the processes leading to obtaining a set of C wideband images $\boldsymbol{y}_{\mathrm{m}}$ from $\boldsymbol{x}$. Such instruments are composed of mirrors to focus light, filters to select the waveband to observe, and a detector to spatially sample the light. However, due to the finite aperture of their mirrors, incoming light is diffracted, and the resulting images are spatially blurred. As the instruments' impulse responses are considered spatially invariant, the blurring effect can be modeled as a spatial convolution with the instruments' impulse responses. The resulting blurred version of the object $\boldsymbol{x}$ observed by the c^{th} imaging instrument is

$$\boldsymbol{x}^c_{\boldsymbol{h}_{\mathrm{m}}}[i,j,l] = (\boldsymbol{x} \underset{i,j}{*} \boldsymbol{h}^c_{\mathrm{m}})[i,j,l] \tag{2}$$

where $\boldsymbol{h}_{\mathrm{m}}$ of the impulse response of the c^{th} instrument, $\underset{i,j}{*}$ is the 2D spatial convolution operator on the axes of i and j.

The spectral filter profiles and detector's quantum efficiencies must also be taken into account for each imager model. These effects result in a spectrally weighted version of $\boldsymbol{x}^c_{\boldsymbol{h}_{\mathrm{m}}}$, denoted $\boldsymbol{x}^c_{\boldsymbol{w}_{\mathrm{m}}}$, and calculated as

$$\boldsymbol{x}^c_{\boldsymbol{w}_{\mathrm{m}}}[i,j,l] = \boldsymbol{x}^c_{\boldsymbol{h}_{\mathrm{m}}}[i,j,l]\ \boldsymbol{w}^c_{\mathrm{m}}[l] \tag{3}$$

where $\boldsymbol{w}^c_{\mathrm{m}}$ are the spectral weights for the c^{th} instrument.

The wideband images are then obtained from the spectral integration of $\boldsymbol{x}^c_{\boldsymbol{w}_{\mathrm{m}}}$. The full equation of the multispectral data formation process for the c^{th} instrument writes as

$$\boldsymbol{y}^c_{\mathrm{m}}[i,j] = \sum_{l} (\boldsymbol{x} \underset{i,j}{*} \boldsymbol{h}^c_{\mathrm{m}})[i,j,l]\ \boldsymbol{w}^c_{\mathrm{m}}[l] + \boldsymbol{n}^c_{\mathrm{m}}[i,j] \tag{4}$$

where $\boldsymbol{n}_{\mathrm{m}}$ is a cube of size $C \times P$ containing random white gaussian noise of standard deviation σ_{m}.

Integrating the linear mixing model from (1) into the latter equation and generalizing it for all instruments leads to rewriting it in a matrix form as

$$\boldsymbol{y}_{\mathrm{m}} = \boldsymbol{M}\boldsymbol{a} + \boldsymbol{n}_{\mathrm{m}} \tag{5}$$

where $\boldsymbol{M} \in \mathbb{R}^{CP \times KP}$.

C. Hyperspectral data formation process

The hyperspectral data in this article $\boldsymbol{y}_{\mathrm{h}}$ is obtained from K integral field spectrometers. As for imaging instruments, these instruments also use mirrors to focus light. The first transformation on the object $\boldsymbol{x}$ by the k^{th} spectrometer is thus calculated as

$$\boldsymbol{x}^k_{\boldsymbol{h}_{\mathrm{h}}}[i,j,l] = (\boldsymbol{x} \underset{i,j}{*} \boldsymbol{h}^k_{\mathrm{h}})[i,j,l] \tag{6}$$

where $\boldsymbol{h}^k_{\mathrm{h}}$ is the impulse response of the k^{th} spectrometer.

These instruments use a diffraction grating to separate input signal wavelengths, but do not transmit flux for all wavelengths equally. Taking this effect and the quantum efficiency of the detectors into account leads to adding a spectral weighting into the models with

$$\boldsymbol{x}^k_{\boldsymbol{w}_{\mathrm{h}}}[i,j,l] = \boldsymbol{x}^k_{\boldsymbol{h}_{\mathrm{h}}}[i,j,l]\ \boldsymbol{w}^k_{\mathrm{h}}[l] \tag{7}$$

where $\boldsymbol{w}^k_{\mathrm{h}}$ are the spectral weights of the k^{th} spectrometer.

Finally, the usually coarse spatial sampling of spectrometer detectors is modeled as a spatial integration of the flux over all regions of $d_i \times d_j$ pixels for every image in $\boldsymbol{x}^k_{\boldsymbol{w}_{\mathrm{h}}}$, d_i, $d_j \in \mathbb{N}^*$. The full equation of the spectrometer data formation process for the k^{th} instrument thus writes as

$$\boldsymbol{y}^k_{\mathrm{h}}[\bar{i},\bar{j},l] = \sum_{i=\bar{i}d_i}^{(\bar{i}+1)d_i} \sum_{j=\bar{j}d_j}^{(\bar{j}+1)d_j} (\boldsymbol{x} \underset{i,j}{*} \boldsymbol{h}^k_{\mathrm{h}})[i,j,l]\ \boldsymbol{w}^k_{\mathrm{h}}[l] + \boldsymbol{n}^k_{\mathrm{h}}[\bar{i},\bar{j},l] \tag{8}$$

where $\bar{i}$ and $\bar{j}$ are the pixels coordinates after spatial decimation of images, and $\boldsymbol{n}_{\mathrm{h}}$ is a cube of size $L \times P$ containing white gaussian noise of standard deviation σ_{h}. Generalizing

this equation for all spectrometers and integrating the mixing model from (1) leads to its matrix form

$$\boldsymbol{y}_{\mathrm{h}} = \boldsymbol{H}\boldsymbol{a} + \boldsymbol{n}_{\mathrm{h}} \tag{9}$$

where $\boldsymbol{H} \in \mathbb{R}^{LP' \times KP}$ with $P' = P/(d_i d_j)$.

III. Methodology

A. Formulation of the problem with handcrafted priors

The abundance maps $\boldsymbol{a}$, describing the original object $\boldsymbol{x}$, are defined as the minimizer of a criterion $\mathcal{J}$

$$\widehat{\boldsymbol{a}} = \underset{\boldsymbol{a}}{\operatorname{argmin}}\ \mathcal{J}(\boldsymbol{a}) \tag{10}$$

with

$$\mathcal{J}(\boldsymbol{a}) = \mu_{\mathrm{m}}\|\boldsymbol{y}_{\mathrm{m}} - \boldsymbol{M}\boldsymbol{a}\|_2^2 + \mu_{\mathrm{h}}\|\boldsymbol{y}_{\mathrm{h}} - \boldsymbol{H}\boldsymbol{a}\|_2^2 + R(\boldsymbol{a}) \tag{11}$$

where $\mu_{\mathrm{m}} = 1/2\sigma_{\mathrm{m}}^2$, $\mu_{\mathrm{h}} = 1/2\sigma_{\mathrm{h}}^2$, and $R(\boldsymbol{a})$ is a regularization term used to stabilize the problem.

Previous works [3] [4] solved this data fusion problem by imposing a spatial smoothness prior to the reconstructed maps $\boldsymbol{a}$. This prior is formulated by choosing R such that it quadratically (ℓ_2-norm) penalizes the differences between neighboring pixels, i.e.

$$R(\boldsymbol{a}) = \mu_{\mathrm{r}}\|\overline{\boldsymbol{D}}\boldsymbol{a}\|_2^2 \tag{12}$$

with $\mu_{\mathrm{r}} \in \mathbb{R}_+$ being the regularization parameter to control the influence of the spatial smoothness, and $\overline{\boldsymbol{D}}$ a finite difference matrix. This criterion was minimized with an iterative gradient-based method in [3], and with an explicit method in [4].

B. Formulation of the problem with an UNNP

Using an Untrained Neural Network Prior (UNNP) for regularizing our ill-posed inverse problem consists in optimizing the parameters $\boldsymbol{\theta}$ of a neural network $f_{\boldsymbol{\theta}}$ to reproduce the maps $\boldsymbol{a}$ from a fixed random input tensor $\boldsymbol{Z}_0$, such that the minimization problem becomes

$$\widehat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}}\ \mathcal{J}(\boldsymbol{\theta}) \tag{13}$$

with

$$\mathcal{J}(\boldsymbol{\theta}) = \mu_{\mathrm{m}}\|\boldsymbol{y}_{\mathrm{m}} - \boldsymbol{M}f_{\boldsymbol{\theta}}(\boldsymbol{Z}_0)\|_2^2 + \mu_{\mathrm{h}}\|\boldsymbol{y}_{\mathrm{h}} - \boldsymbol{H}f_{\boldsymbol{\theta}}(\boldsymbol{Z}_0)\|_2^2. \tag{14}$$

Once the optimal weights $\widehat{\boldsymbol{\theta}}$ have been estimated to specifically minimize (14), the reconstructed abundance maps $\widehat{\boldsymbol{a}}$ are then calculated with

$$\widehat{\boldsymbol{a}} = f_{\widehat{\boldsymbol{\theta}}}(\boldsymbol{Z}_0). \tag{15}$$

Unlike the learning-based method from the literature, these networks require no prior learning before being optimized to solve (14), thus removing potential learning biases and the need for a large and diverse database for training. Note that the use of UNNP allows the removal of the explicit regularization term $R(\boldsymbol{a})$, which was a necessity in (11). Indeed, the regularization here relies on untrained neural network structures' inherent ability to generate natural images, as demonstrated in [14].

C. Architecture of the UNNP

The neural network used in this work follows the Deep Decoder (DD) architecture proposed in [21], i.e. an underparametrized fully-connected neural network. This underparametrization compels the network to generalize images into a smaller parameter space, thus preventing the network from recovering the noise of the input data [21] [15].

As presented in [21], the Deep Decoder architecture is composed of d layers, where each layer is composed of a pixel-wise linear combination of the input channels, computed as a 2D convolution with a 1×1 kernel, a bilinear upsampling unit to double the input's width and height, a ReLU activation, and a channelwise normalization (CN). These operations can be summed up as successive transformations of a tensor $\boldsymbol{Z}$ with

$$\boldsymbol{Z}_{i+1} = \mathrm{CN}(\mathrm{ReLU}(\boldsymbol{U}_i\boldsymbol{Z}_i\boldsymbol{\theta}_i)) \tag{16}$$

where $i \in [0, d-1]$, $\boldsymbol{Z}_i \in \mathbb{R}^{n_i \times k_i}$ the tensor at layer i composed of k_i channels of n_i variables each, $\boldsymbol{U}_i \in \mathbb{R}^{n_{i+1} \times n_i}$ the upsampling matrix, and $\boldsymbol{\theta} \in \mathbb{R}^{k_i \times k_{i+1}}$ the containing the weights for pixel-wise linear combinations of channels, also being the weights of the network. The dimensions of the random input tensor $\boldsymbol{Z}_0$ are set such that the spatial dimensions of $\boldsymbol{Z}_{d-1}$ match those of the network's output. Therefore, the layer $d-1$ does not require spatial upsampling, i.e $\boldsymbol{U}_{d-1} = \boldsymbol{I}$. Finally, the network's output is computed via a sigmoid function applied on a pixel-wise linear combination of the channels of $\boldsymbol{Z}_d$, such that its third dimension matches the one of the output (T in our case, see section II-A). This is written as

$$f_{\boldsymbol{\theta}}(\boldsymbol{Z}_0) = \mathrm{sigmoid}(\boldsymbol{Z}_d\boldsymbol{\theta}_d) \tag{17}$$

where $\boldsymbol{\theta}_d \in \mathbb{R}^{k_d \times T}$. Therefore, assuming k_i constant for $i \in [0, d]$ and equal to k, the number of parameters of this network is

$$N(d, k, T) = dk^2 + 2dk + Tk. \tag{18}$$

IV. Results

As mentioned in the introduction, this research primarily focuses on the fusion of astronomical imaging and spectrometry data degraded by spectrally-varying spatial blurs. The study case of this paper is the Mid-Infrared Instrument (MIRI) from the James Webb Space Telescope (JWST). We adapted the instrument models for MIRI's imager (MIRIM) and its Medium Resolution Spectrometer (MRS) to the multispectral and hyperspectral data formation processes presented in sections II-B and II-C respectively. The impulse response for both models, $\boldsymbol{h}$, was simulated using *webbpsf*, showcasing its spectral variability, as depicted in Fig. 1. MIRIM, equipped with 9 filters, incorporates the actual Photon Conversion Efficiency (PCE) curves for spectral response analysis. For the spectrometer model, we implemented decimation by a factor of 4 in both dimensions, resulting in spectral aliasing for signals under 20 µm. Although the MRS features 12 distinct PCE

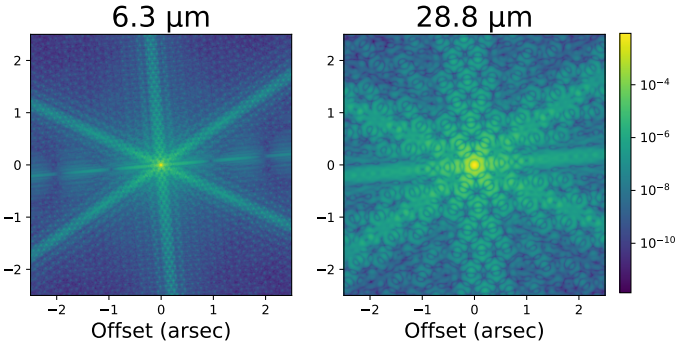


Fig. 1: MIRI PSF [22] (logarithmic scale).

curves across 4 channels and 3 sub-channels, our analysis simplifies this to a single channel by merging these curves.

We generated multispectral $\boldsymbol{y}_{\mathrm{m}}$ and hyperspectral $\boldsymbol{y}_{\mathrm{h}}$ data by applying these models to a set of $T = 5$ spectral abundance maps, simulating the Linear Mixing Model (1). These maps, derived from the Orion Bar's observational data, contain 128×256 pixels each, covering the 5 to 28 µm MIRI band with $L = 300$ spectral points. The datasets $\boldsymbol{y}_{\mathrm{m}}$ and $\boldsymbol{y}_{\mathrm{h}}$ were subsequently corrupted with additive white Gaussian noise, with standard deviations σ_{m} and σ_{h} respectively. These standard deviations were first adjusted such that the signal-to-noise ratio $\mathrm{SNR}_{\mathrm{dB}}$ of both MS and HS inputs, $\boldsymbol{y}_{\mathrm{m}}$ and $\boldsymbol{y}_{\mathrm{h}}$, is 25 dB. The reconstruction with two different handcrafted smoothness priors are compared in this study : one with an ℓ_1-norm, and another with an ℓ_2-norm [3] [4]. These reconstructions involves the setting of the hyperparameter μ_{r}, see (12), but it was here chosen to minimize the true Normalized Root Mean Squared Error (NRMSE) [23] with respect to the original data in order to maximize their performances. The reconstruction results for these priors are shown Fig. 2c and Fig. 2d. Such results are compared with another prior: a Deep Decoder architecture presented in section III-C. This neural network is initialized with $d = 6$ layers with $k = 128$ parameters each. The dimensions of the random input $\boldsymbol{Z}_0$ are $128 \times 4 \times 8$, and those of the network output are $5 \times 128 \times 256$, i.e., the dimensions of the abundance maps. The network is thus correctly underparametrized (see (18)) with $N(6, 128, 5) = 100\,480$ parameters, whereas the output has $163\,840$ variables. The imager and spectrometer models are then used to generate data from the network output to compute the loss (14). The network is optimized using the ADAM algorithm [24] with $60\,000$ iterations to ensure convergence. Only the weights $\boldsymbol{\theta}$ minimizing $\mathcal{J}(\boldsymbol{\theta})$ (and not the true NRMSE that relies on the ground truth) across all iterations were saved to reconstruct data, which are shown Fig.2e. Tab. I compares the performances of those three priors for four different quality reconstruction metrics, with superior results for the UNNP for each of them.

Fig. 3 shows the superior PSNR of the reconstructed data with the UNNP for all input data with $\mathrm{SNR}_{\mathrm{dB}} \leq 40$ dB, demonstrating a stronger resilience of this prior for realistic noise cases.

V. Conclusion

A multispectral and hyperspectral data fusion problem with spectrally-varying spatial blurs was solved using a simple and compact Untrained Neural Network Prior, based on a Deep Decoder architecture [21]. This problem includes denoising, deconvolution, and unmixing aspects, which had only been solved with handcrafted priors in the literature. This UNNP aimed to be a compromise between the performance-limiting handcrafted priors and the data-demanding deep learning methods, as the latter cannot be easily used when dealing with astronomical data.

The introduction of a Linear Mixing Model (section II-A) induced three benefits : significantly reducing the dimension of the object to reconstruct by estimating a small set of abundance maps $\boldsymbol{a}$ instead of a hyperspectral cube $\boldsymbol{x}$, thus allowing the use of a compact neural network as prior. Secondly, this model makes the network's architecture independent of the spectral dimension of $\boldsymbol{x}$, which opens the possibility for the fusion of high-dimension data. Finally, the spectral correlations incorporated in $\boldsymbol{x}$ via this model are thought to help with the deconvolution aspect of this problem. Specifically, the high spatial resolution of short-wavelength images can be leveraged to enhance the deconvolution of long-wavelength images.

A comparison for the fusion of simulated realistic MIRI/JWST data between handcrafted priors and UNNP highlighted higher performances for the latter for cases where the input MS and HS data had a signal-to-noise ratio equal or inferior to 40 dB, thus concluding a higher noise resilience when using the UNNP for this study case. These promising results could pave the way for using an alternative prior without requiring labeled databases. Interesting avenues of research could include implementing both priors in a single criterion or developing an underparametrized convolutional network (instead of a fully-connected one) for an even more compact network.

References

[1] N. Yokoya *et al.*, "Hyperspectral and multispectral data fusion: A comparative review of the recent literature," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 2, pp. 29–56, 2017.

[2] STScI, "JWST documentation website," https://jwst-sdocs.stsci.edu, accessed: 2023-03-06.

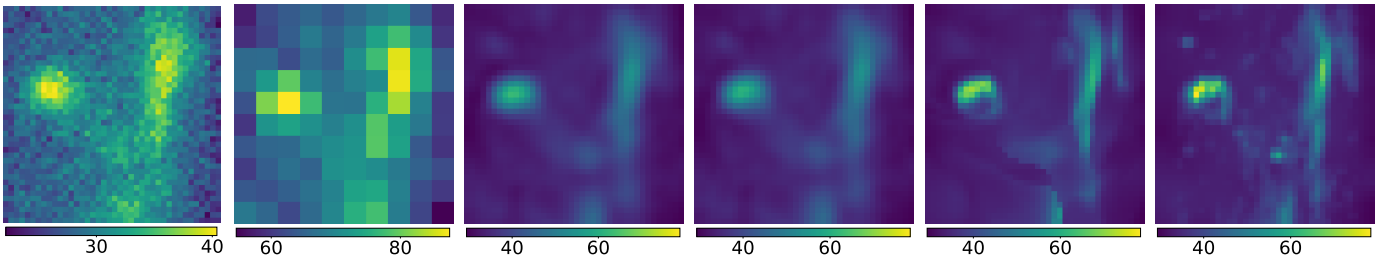
[3] C. Guilloteau *et al.*, "Hyperspectral and multispectral image fusion under spectrally varying spatial blurs–application to high dimensional infrared astronomical imaging," *IEEE Transactions on Computational Imaging*, vol. 6, pp. 1362–1374, 2020.

[4] D. Pineau, F. Orieux, and A. Abergel, "Exact solution for multispectral and hyperspectral fusion via hessian inversion," in *Workshop on Hyperspectral Images and Signal Processing: Evolution in Remote Sensing*, 2023.

[5] C. Hegde, "Algorithmic aspects of inverse problems using generative models," in *2018 56th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2018, pp. 166–172.

[6] G. H. Golub, P. C. Hansen, and D. P. O'Leary, "Tikhonov regularization and total least squares," *SIAM journal on matrix analysis and applications*, vol. 21, no. 1, pp. 185–194, 1999.

[7] G. Ongie, A. Jalal, C. A. Metzler, R. G. Baraniuk, A. G. Dimakis, and R. Willett, "Deep learning techniques for inverse problems in imaging," *IEEE Journal on Selected Areas in Information Theory*, vol. 1, no. 1, pp. 39–56, 2020.



(a) $\boldsymbol{y}_{\mathrm{m}}$ (b) $\boldsymbol{y}_{\mathrm{h}}$ (c) ℓ_1 smoothness (d) ℓ_2 smoothness (e) UNNP (f) Original

Fig. 2: From left to right : multispectral data $\boldsymbol{y}_{\mathrm{m}}$ and hyperspectral data $\boldsymbol{y}_{\mathrm{h}}$ with $\mathrm{SNR}_{\mathrm{dB}} = 25$ dB, reconstruction with the handcrafted ℓ_2 smoothness prior, the ℓ_1 one, reconstruction with the untrained neural network prior, and the original image. All images taken at 18 µm. The last four images share the same dynamic range.

Methods	PSNR [dB]	NRMSE $\times 10^{-3}$	SSIM $\times 10^{-2}$	SAM $\times 10^{-3}$
ℓ_1 smoothness prior	46.7	29.2	99.0	5.9
ℓ_2 smoothness prior [3] [4]	46.7	29.3	99.0	5.8
Untrained neural network prior	**48.4**	**23.8**	**99.2**	**4.4**

TABLE I: Comparison of reconstructions with handcrafted smoothness priors against an UNNP when the input MS and HS data have $\mathrm{SNR}_{\mathrm{dB}} = 25$ dB for four different metrics : Peak Signal-to-Noise Ratio (PSNR), Normalized Root Mean Square Error (NRMSE), Structural Similarity Index Measure (SSIM), and the Spectral Angle Mapper (SAM). Best results in bold.

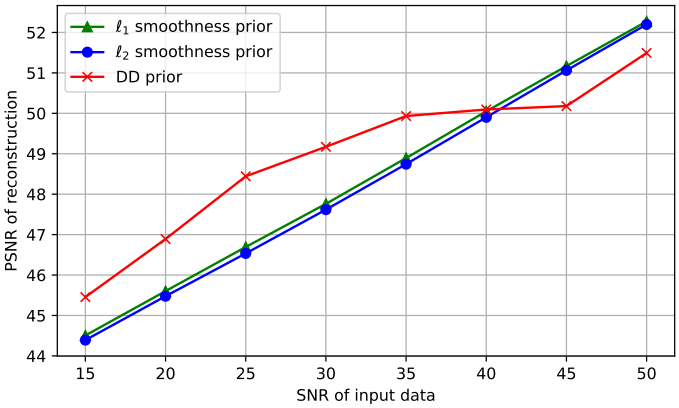


Fig. 3: Comparison of the PSNR of the reconstructed data with handcrafted smoothness priors and a Deep Decoder (DD) untrained neural network prior for multiple $\mathrm{SNR}_{\mathrm{dB}}$ for the input MS and HS data.

[8] J.-F. Hu, T.-Z. Huang, L.-J. Deng, H.-X. Dou, D. Hong, and G. Vivone, "Fusformer: A transformer-based fusion network for hyperspectral image super-resolution," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.

[9] L. Ardizzone, J. Kruse, S. Wirkert, D. Rahner, E. W. Pellegrini, R. S. Klessen, L. Maier-Hein, C. Rother, and U. Köthe, "Analyzing inverse problems with invertible neural networks," *arXiv preprint arXiv:1808.04730*, 2018.

[10] T. Oberlin and M. Verm, "Regularization via deep generative models: an analysis point of view," in *2021 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2021, pp. 404–408.

[11] Y. Romano, M. Elad, and P. Milanfar, "The little engine that could: Regularization by denoising (red)," *SIAM Journal on Imaging Sciences*, vol. 10, no. 4, pp. 1804–1844, 2017.

[12] Y. Huang, E. Chouzenoux, and J.-C. Pesquet, "Unrolled variational bayesian algorithm for image blind deconvolution," *IEEE Transactions on Image Processing*, vol. 32, pp. 430–445, 2022.

[13] D. Gilton, G. Ongie, and R. Willett, "Deep equilibrium architectures for inverse problems in imaging," *IEEE Transactions on Computational Imaging*, vol. 7, pp. 1123–1133, 2021.

[14] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 9446–9454.

[15] A. Qayyum, I. Ilahi, F. Shamshad, F. Boussaid, M. Bennamoun, and J. Qadir, "Untrained neural network priors for inverse imaging problems: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 5, pp. 6511–6536, 2023.

[16] T. Uezato, D. Hong, N. Yokoya, and W. He, "Guided deep decoder: Unsupervised image pair fusion," in *European Conference on Computer Vision*. Springer, 2020, pp. 87–102.

[17] J. Gao, J. Li, and M. Jiang, "Hyperspectral and multispectral image fusion by deep neural network in a self-supervised manner," *Remote Sensing*, vol. 13, no. 16, p. 3226, 2021.

[18] S. Liu, S. Miao, J. Su, B. Li, W. Hu, and Y.-D. Zhang, "Umag-net: A new unsupervised multiattention-guided network for hyperspectral and multispectral image fusion," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 7373–7385, 2021.

[19] M. E. Hadj-Youcef, "Spatio spectral reconstruction from low resolution multispectral data: application to the mid-infrared instrument of the james webb space telescope," Ph.D. dissertation, Université Paris-Saclay (ComUE), 2018.

[20] R. Abi-Rizk *et al.*, "Super-resolution hyperspectral reconstruction with majorization-minimization algorithm and low-rank approximation," *IEEE Transactions on Computational Imaging*, vol. 8, pp. 260–272, 2022.

[21] R. Heckel and P. Hand, "Deep decoder: Concise image representations from untrained non-convolutional networks," *arXiv preprint arXiv:1810.03982*, 2018.

[22] M. D. Perrin *et al.*, "Simulating point spread functions for the james webb space telescope with webbpsf," in *Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave*, vol. 8442. SPIE, 2012, pp. 1193–1203.

[23] J. R. Fienup, "Invariant error metrics for image reconstruction," *Appl. Opt.*, vol. 36, no. 32, pp. 8352–8357, Nov 1997. [Online]. Available: https://opg.optica.org/ao/abstract.cfm?URI=ao-s36-s32-s8352

[24] S. J. Reddi, S. Kale, and S. Kumar, "On the convergence of adam and beyond," *arXiv preprint arXiv:1904.09237*, 2019.