

Deep Hyperspectral Prior: Single-Image Denoising, Inpainting, Super-Resolution

Oleksii Sidorov and Jon Hardeberg
The Norwegian Colour and Visual Computing Laboratory, NTNU
Gjøvik, Norway

oleksiis@stud.ntnu.no, jon.hardeberg@ntnu.no



Abstract

Deep learning algorithms have demonstrated state-of-the-art performance in various tasks of image restoration. This was made possible through the ability of CNNs to learn from large exemplar sets. However, the latter becomes an issue for hyperspectral image processing where datasets commonly consist of just a few images. In this work, we propose a new approach to denoising, inpainting, and super-resolution of hyperspectral image data using intrinsic properties of a CNN without any training. The performance of the given algorithm is shown to be comparable to the performance of trained networks, while its application is not restricted by the availability of training data. This work is an extension of original “deep prior” algorithm to hyperspectral imaging domain and 3D-convolutional networks.

1. Introduction

Deep Convolutional Neural Networks (CNNs) are occupying more and more leading positions in benchmarks of various image processing tasks. Commonly, it is related to the excellent representative ability of hierarchical convolutional layers which allows CNNs to learn a large amount of visual data without any hand-crafted assumptions. Ulyanov *et al.* [30] were the first to show that not only a learning ability but also the inner structure of a CNN itself can be beneficial for processing of image data.

For example, the inverse task of image restoration, such as inpainting, noise removal, or super-resolution, can be formulated as an energy minimization problem as follows:

$$x^* = \min_x E(x, x_0) + R(x) \quad , \quad (1)$$

where $E(x, x_0)$ is a task related metric, x and x_0 are original and corrupted images, and $R(x)$ is a regularization term (image prior) which can be chosen manually or can be learned from data (as it happens in the vast majority of CNN-based methods). However, the theory of Ulyanov *et al.* states that image prior can be found in the space of the network’s parameters directly, through the optimization process, which allows removal of regularization term, and searching a solution as:

$$x^* = f_{\theta^*}(z), \quad \text{where } \theta^* = \arg \min_x E(f_{\theta}(z), x_0) \quad (2)$$

Here, f is a CNN with parameters θ , and z is a fixed input (noise). Thereby, an original image can be restored via optimization of the network’s weights using only a corrupted image.

This approach has a particularly high significance in the domain of hyperspectral imaging (HSI). Currently, HSI is a powerful tool which is widely used in remote sensing, agriculture, cultural heritage, food industry, pharmaceuticals, *etc.* The complexity of hyperspectral equipment and process of data acquisition make corruption of image data even more likely than it is for RGB imaging. Thus, it generates an increased demand for algorithms of hyperspectral

image restoration. But, at the same time, accurate learning-based methods can hardly be used due to the lack of data. The complexity of data acquisition does not allow gathering of large custom datasets for a particular task, and even openly available ones are very limited and rarely exceed one hundred images, sometimes consisting of just one image [3][15][31].

Our work aims to solve this problem and, altogether, our contributions can be formulated as follows:

- We propose an efficient algorithm for hyperspectral image restoration based on the theory of Ulyanov *et al.*
- We design a new 3D-convolutional implementation of the algorithm and prove the fundamental property of 3D convolutions to contain low-level image information which can be used as a prior.
- We demonstrate a potential application of the algorithm for HSI and evaluate its performance in comparison with other methods.
- Eventually, we make the source code publicly accessible and is ready to use out-of-the-box.

2. Related Works

In this section, we briefly introduce recent advances in hyperspectral denoising, inpainting, and super-resolution.

One way to perform HSI denoising is to apply 2D algorithms to each band separately. Such an approach can utilize bilateral [29] or NL-means filtering [7], total variation [27], block-matching 3D-filtering [9], or novel CNN-based techniques, *e.g.* DnCNN [39]. However, not taking spectral data into consideration may cause distortions and artifacts in the spectral domain. This, has given rise to a family of algorithms based on spatial-spectral features, such as spatio-spectral derivative-domain wavelet shrinkage [24], low-rank tensor approximation [26], low-rank matrix recovery [38], and most recently FastHyDe algorithm [40], which utilizes sparse representation of an image linked to its low-rank and self-similarity characteristics. A deep learning paradigm has been used in the 3D modification of DnCNN, and more advanced HSI-oriented network HSID-CNN [36].

The inpainting of grayscale and RGB images conventionally rely on patch-similarity and variational algorithms to propagate information from intact regions to holes [2][4][13] and may be used for HSI data in a band-wise manner. Novel inpainting approaches benefit from realistic reconstruction ability of GANs, which allows filling even large holes with remarkable accuracy [16][21][35]. However, they rely on large training datasets. There are also a number of HSI-specific inpainting methods [6][8][10]. Similar to our approach, Adesso *et al.* [1] address HSI

inpainting as an optimization task with (hand-crafted) collaborative total variation regularizer, while Yao *et al.* [34] designed a regularizer based on the Criminisi’s inpainting method. Recently, the FastHyIn algorithm [40] (an extension of FastHyDe) demonstrated state-of-the-art HSI inpainting accuracy, with the only remark that similar to [28] it utilizes information from intact bands, thus cannot be used in cases of all-bands corruption.

The majority of hyperspectral super-resolution (SR) algorithms perform a fusion of input hyperspectral image with a high-resolution multispectral image which is easier to obtain [11][25]. Single-image SR is a more sophisticated task. The attempts to solve it include spectral mixture analysis [19], low-rank tensor approximation [32], Local-Global Combined Network [18], MRF-based energy minimization [17], transfer learning [37], the recent method of 3D Full Convolutional Neural Network [23], and others.

3. Methodology

The main idea of the method is captured by Equations (1) and (2). The fully-convolutional encoder-decoder f_θ is designed to translate a fixed input z filled with noise to the original image x , conditioned on corrupted image x_0 . We use the paradigm of the “deep prior” method [30] which says that optimal weights θ of a network f_θ can be found from the intrinsic prior contained in a network structure instead of learning them from the data. Particularly, θ is approximated by the minimizer θ^* :

$$\theta^* = \arg \min_x E(f_\theta(z), x_0) \quad (3)$$

which can be obtained using an optimizer such as gradient descent from randomly initiated parameters. It is also possible to optimize over input z (not covered in this work).

The energy function $E(x, x_0)$ may be chosen accordingly to the application task. In the case of a basic reconstruction problem, it may be formulated as L_2 -distance:

$$E(x, x_0) = \|x - x_0\|^2 \quad (4)$$

It was shown [30] that optimization converges faster in cases of natural-looking images rather than random noise, *i.e.* the process demonstrates high impedance to noise and low impedance to signal. The latter can be used to interrupt the reconstruction before the noise will be recovered, which will lead to *blind image denoising*.

Also, $E(x, x_0)$ term can be modified to fill the missing regions in a *inpainting* problem with mask $m \in \{0, 1\}$:

$$E(x, x_0) = \|x - x_0 \circ m\|^2 \quad (5)$$

where \circ is a Hadamard product. Otherwise, downsampling operator $d(x, \alpha) : x^{\alpha N \times \alpha N \times C} \rightarrow x^{N \times N \times C}$ with factor α can be used in $E(x, x_0)$ to address *super-resolution* task as a

prediction of high-res image x which, when downsampled, is the same as low-res image x_0 :

$$E(x, x_0) = \|d(x) - x_0\|^2 \tag{6}$$

3.1. Implementation details

It was found, that different fully-convolutional encoder-decoder architectures (sometimes with skip-connections) are suitable for the implementation of the given method. For the exact description in details please see the source code¹. Although parameters differ for each sub-task, the general framework is common and is illustrated in Fig. 1.

We experiment with two versions of the networks – 2D and 3D. While 2D convolutions are still able to process multi-channel input, they cause shrinkage of spectral information already at the first convolutional layer and recover

¹<https://github.com/acecreamu/deep-hs-prior>

it back at the last one. In this case, filters of these layers would have an “elongated” shape with a depth equal to the depth of the hyperspectral image. A 3D convolution allows the use of smaller filters (e.g. $3 \times 3 \times 3$) along the whole network because its output is a 3D volume. This ability to preserve a 3D shape of the input is considered to be beneficial for the processing of hyperspectral data. It is worth mentioning, that unlike a conventional “hourglass” architecture, where data is consecutively downsampled/upsampled along two spatial dimensions, processing of 3D volumes allows doing the same with third dimension as well (see Fig. 1).

The input of the network is uniform noise in range 0-0.1 of a shape equal to the shape of a processed hyperspectral image. Optionally, it is additionally perturbed at each iteration with Gaussian noise of specified σ . The activation function used is LeakyReLU [33]. Downsampling is per-

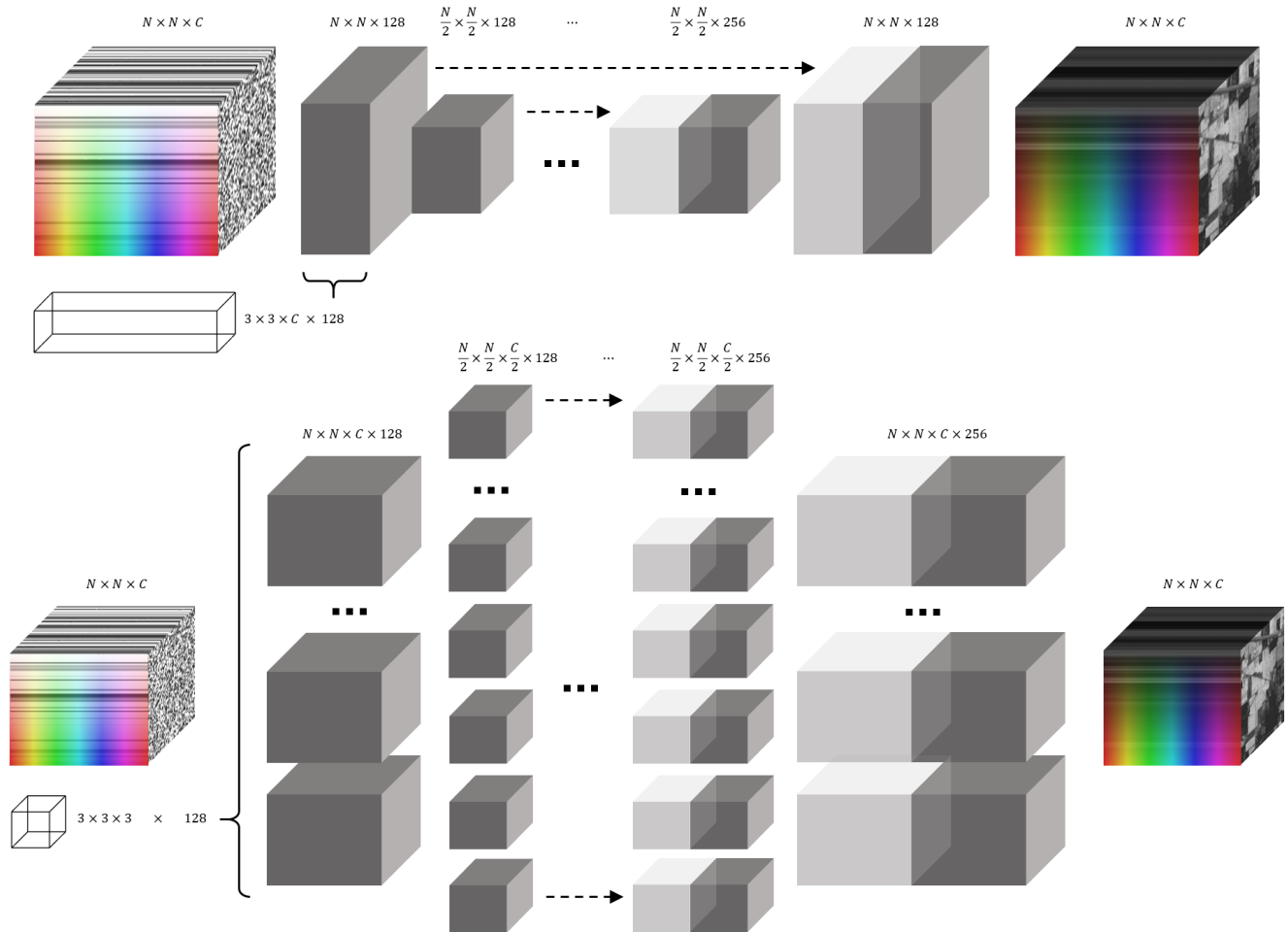


Figure 1: 2D (top) and 3D (bottom) convolutional architectures used in our experiments. Size of the filters of the first convolutional layer is illustrated under the input image. Pooling and activation layers are omitted for simplicity sake.

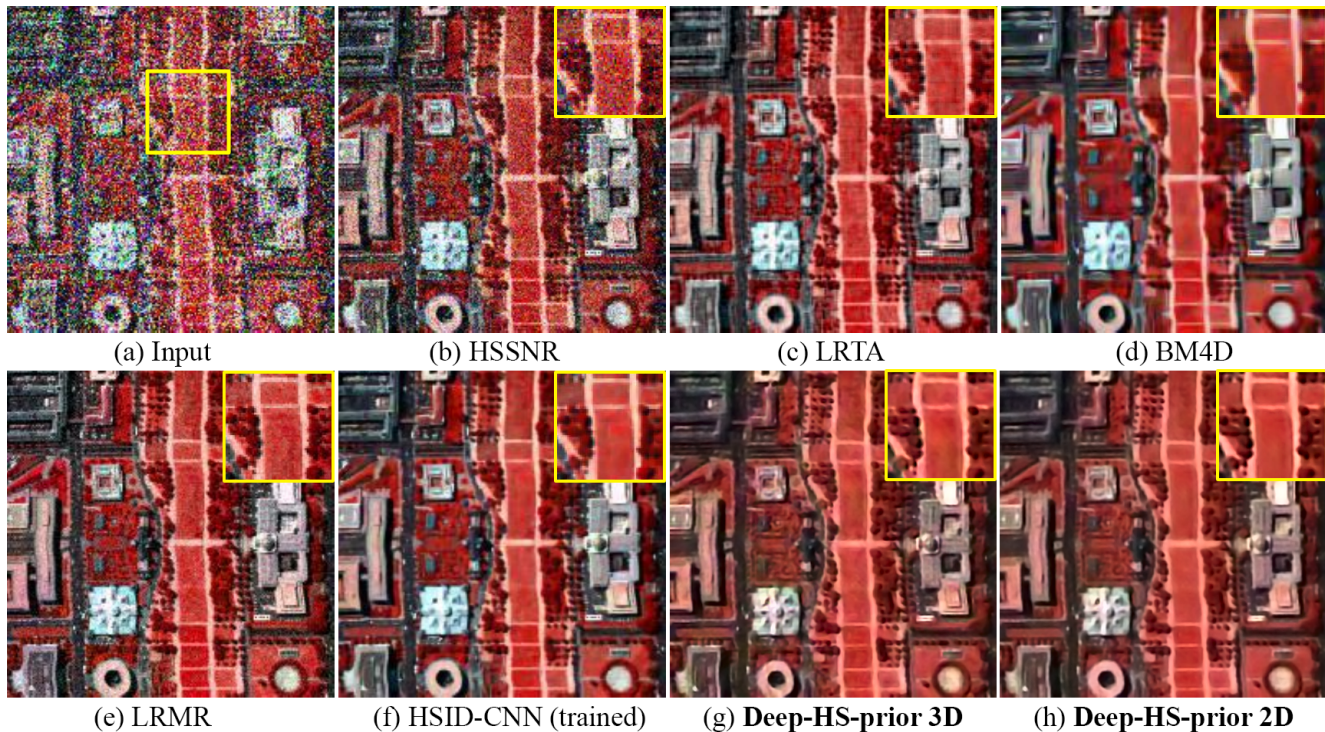


Figure 2: HSI denoising results. HYDICE DC Mall image; false-color with bands (57, 27, 17).

formed using a stride of convolutions, while upsampling is either “nearest” or bilinear (trilinear for a 3D case). Other methods can also be used but these prevailed in our experiments. ADAM algorithm was used for optimization.

4. Experimental setup

Denoising. We evaluate the ability of an algorithm to remove noise using HYDICE DC Mall data [31] with synthetically added Gaussian noise of $\sigma = 100$. The image consists of 191 channels and was cropped to 200×200 pixels size. Results (Fig. 2, Table 1) are compared to HSSNR [24], LRTA [26], BM4D [22], LRMR [38], and HSID-CNN [36] methods.

Inpainting. The Indian Pines dataset [3] ($145 \times 145 \times 200$) from AVIRIS sensor was used to test the proposed inpainting method. The mask of corrupted strips was applied to all bands. Results (Fig. 4, Table 1) are compared to Mumford-Shah [14] and fourth-order total variation ($TV-H^{-1}$) [5] 2D methods, as well as the state-of-the-art HSI inpainting method FastHyIn [40].

Super-resolution. The experiment was conducted using ROSIS-03 image of Pavia Center [15] (102 spectral bands). A patch of 150×150 pixels was cropped from the orig-

inal image and downsampled by a factor of 2 by spatial dimensions. The evaluation includes “nearest” and bilinear upsampling, learning-based method SRCNN [12] applied band-wise (msiSRCNN) or by groups of 3 bands (3B-SRCNN [20]), and 3D-FCNN [23]. Results are presented in Fig. 3 and Table 1.

5. Results and discussion

As can be seen, the proposed method outperforms all single-image algorithms and demonstrate performance comparable to *trained* CNNs, while *not being trained* on any dataset before. Surprisingly, the 2D version of the Deep HS prior outperformed the 3D-convolutional one in all experiments. Note that the 2D implementation has nothing to do with band-wise processing; instead, it captures spectral information in filters of the first convolutional layer and uses combinations of them at the subsequent layers. Besides, the 3D implementation requires significantly more memory and computational time for processing the same amount of data because it works with tensors of higher dimensionality (roughly speaking a 3D version will have C times more parameters, where C is a number of bands). Eventually, we find 2D implementation was sufficient for processing of hyperspectral data. And yet, it is worth mentioning that 3D version demonstrated comparable performance that is im-

Denoising							
	HSSNR	LRTA	BM4D	LRMR	HSID-CNN (trained)	Deep HS prior 3D	Deep HS prior 2D
MPSNR	16.31	23.17	22.57	24.31	25.29	23.24	25.05
MSSIM	0.605	0.849	0.812	0.879	0.901	0.852	0.889
SAM	24.73	9.122	9.761	10.46	8.406	9.910	8.606

Inpainting							
		Do Nothing	Mumford- Shah	$TV-H^{-1}$	FastHyIn	Deep HS prior 3D	Deep HS prior 2D
MPSNR		17.75	24.74	27.68	28.08	35.34	37.54
MSSIM		0.722	0.890	0.911	0.920	0.966	0.979
SAM		Inf	5.429	3.855	3.032	1.133	0.856

Super-resolution							
	Nearest	Bicubic	msiSRCNN (trained)	3B-SRCNN (trained)	3D-FCNN (trained)	Deep HS prior 3D	Deep HS prior 2D
MPSNR	29.98	31.10	32.48	32.69	33.92	32.31	33.67
MSSIM	0.921	0.937	0.957	0.960	0.969	0.945	0.967
SAM	4.786	4.592	4.617	4.661	4.140	4.692	4.211

Table 1: Quantitative evaluation of the results.

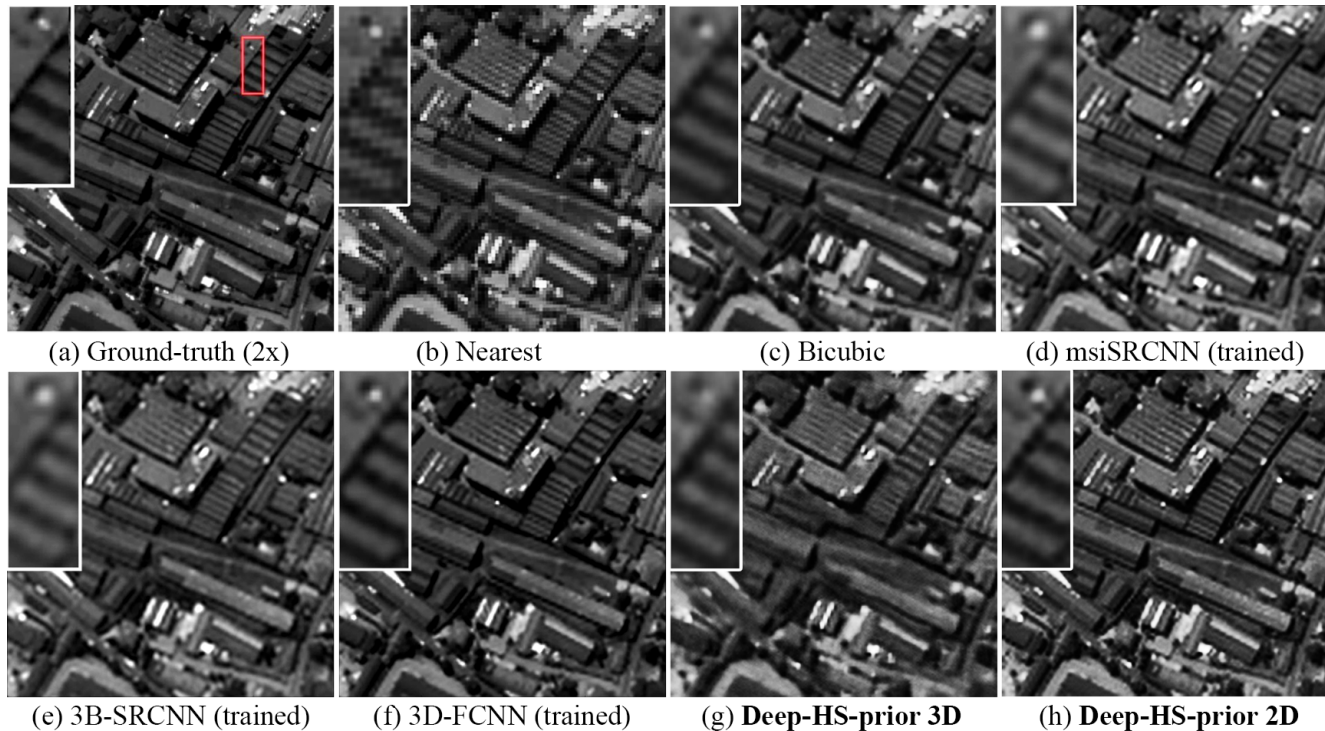


Figure 3: HSI super-resolution results. Pavia Center image; rescale factor 2; band 25 visualization.

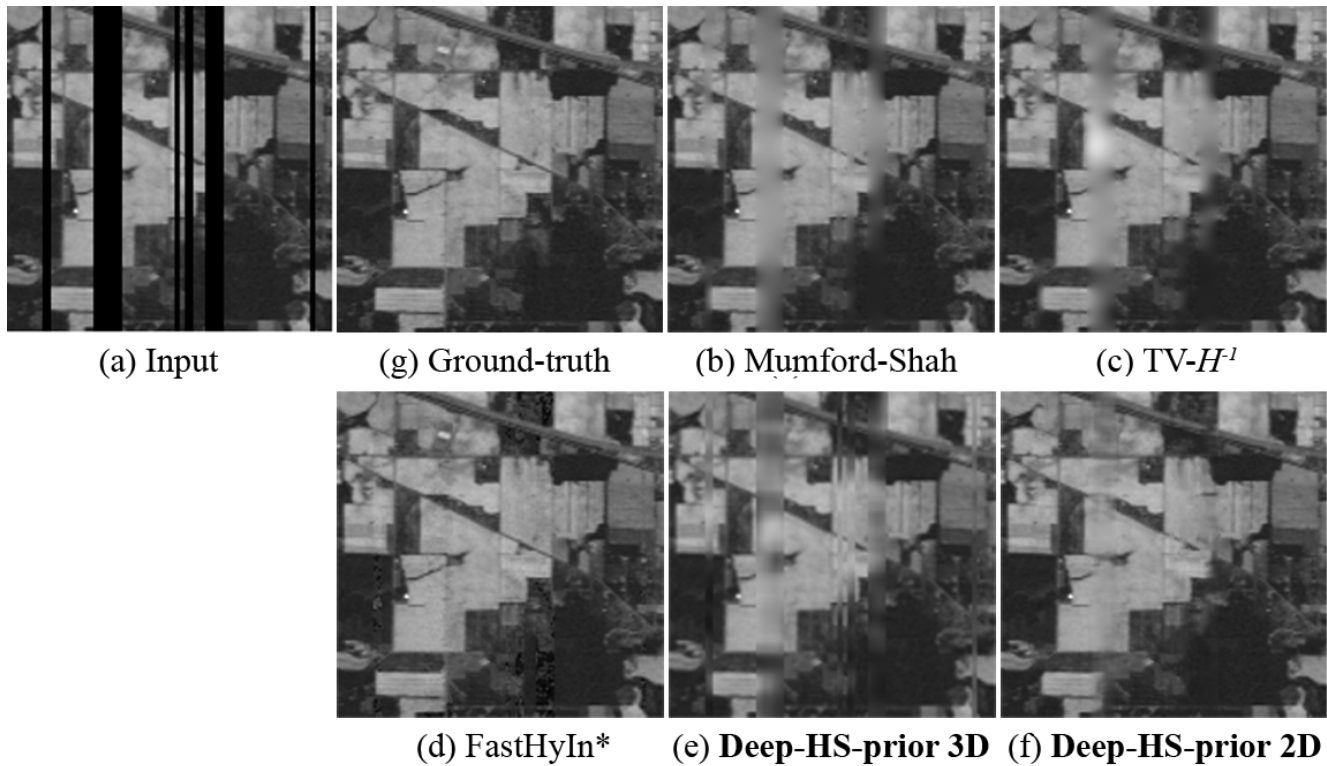


Figure 4: HSI inpainting results. AVIRIS Indian Pines dataset; band 150. The mask of corrupted stripes (a) was applied to all bands, except case (d) where only bands 25:175 (75%) were affected.

portant from a theoretical point of view because it proves that CNN architectures based on 3D convolutions also contain the image prior within the intrinsic parameters.

6. Conclusions

Starting from the paradigm that the image prior can be found within a CNN itself and not be learned from training data or designed manually, we developed an effective single-hyperspectral-image restoration algorithm. Qualitative and quantitative evaluation of the results demonstrated superior effectiveness of the proposed algorithm compared to other single-image algorithms.

References

- [1] P. Addesso, M. Dalla Mura, L. Condat, R. Restaino, G. Vivone, D. Picone, and J. Chanussot. Hyperspectral image inpainting based on collaborative total variation. In *IEEE International Conference on Image Processing (ICIP)*, pages 4282–4286, Sep. 2017. 2
- [2] C. Barnes, E. Shechtman, A. Finkelstein, and D. B. Goldman. PatchMatch: A randomized correspondence algorithm for structural image editing. *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 28(3), Aug. 2009. 2
- [3] M. F. Baumgardner, L. L. Biehl, and D. A. Landgrebe. 220 band aviris hyperspectral image data set: June 12, 1992 indian pine test site 3, Sep 2015. 2, 4
- [4] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester. Image inpainting. In *Proceedings of the conference on Computer graphics and interactive techniques*, pages 417–424. ACM Press/Addison-Wesley Publishing Co., 2000. 2
- [5] A. Bertozzi and C.-B. Schönlieb. Unconditionally stable schemes for higher order inpainting. *Communications in Mathematical Sciences*, 9(2):413–457, 2011. 4
- [6] F. Bousefsaf, M. Tamaazousti, S. H. Said, and R. Michel. Image completion using multispectral imaging. *IET Image Processing*, 12(7):1164–1174, 2018. 2
- [7] A. Buades, B. Coll, and J.-M. Morel. A non-local algorithm for image denoising. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR’05)*, volume 2, pages 60–65. IEEE, 2005. 2
- [8] Q. Cheng, H. Shen, L. Zhang, and P. Li. Inpainting for remotely sensed images with a multichannel nonlocal total variation model. *Geoscience and Remote Sensing, IEEE Transactions on*, 52:175–187, 01 2014. 2
- [9] K. Dabov, A. Foi, and K. Egiazarian. Video denoising by sparse 3d transform-domain collaborative filtering. In *European Signal Processing Conference*, pages 145–149. IEEE, 2007. 2

- [10] K. Degraux, V. Cambareri, L. Jacques, B. Geelen, C. Blanch, and G. Lafruit. Generalized inpainting method for hyperspectral image acquisition. In *IEEE International Conference on Image Processing (ICIP)*, pages 315–319, 2015. 2
- [11] R. Dian, L. Fang, and S. Li. Hyperspectral image super-resolution via non-local sparse tensor factorization. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3862–3871, July 2017. 2
- [12] C. Dong, C. C. Loy, K. He, and X. Tang. Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(2):295–307, 2015. 4
- [13] A. A. Efros and W. T. Freeman. Image quilting for texture synthesis and transfer. In *Proceedings of the conference on Computer graphics and interactive techniques*, pages 341–346. ACM, 2001. 2
- [14] S. Esedoglu and J. Shen. Digital inpainting based on the muford–shah–euler image model. *European Journal of Applied Mathematics*, 13(4):353–370, 2002. 4
- [15] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton. Advances in spectral-spatial classification of hyperspectral images. *Proceedings of the IEEE*, 101(3):652–675, 2012. 2, 4
- [16] S. Iizuka, E. Simo-Serra, and H. Ishikawa. Globally and Locally Consistent Image Completion. *ACM Transactions on Graphics (Proc. of SIGGRAPH 2017)*, 36(4):107:1–107:14, 2017. 2
- [17] H. Irmak, G. B. Akar, and S. E. Yuksel. A map-based approach for hyperspectral imagery super-resolution. *IEEE Transactions on Image Processing*, 27(6):2942–2951, June 2018. 2
- [18] S. Lei, Z. Shi, and Z. Zou. Super-resolution for remote sensing images via localglobal combined network. *IEEE Geoscience and Remote Sensing Letters*, 14(8):1243–1247, Aug 2017. 2
- [19] J. Li, Q. Yuan, H. Shen, X. Meng, and L. Zhang. Hyperspectral image super-resolution by spectral mixture analysis and spatialspectral group sparsity. *IEEE Geoscience and Remote Sensing Letters*, 13(9):1250–1254, Sep. 2016. 2
- [20] L. Liebel and M. Körner. Single-image super resolution for multispectral remote sensing data using convolutional neural networks. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 41:883–890, 2016. 4
- [21] G. Liu, F. A. Reda, K. J. Shih, T.-C. Wang, A. Tao, and B. Catanzaro. Image inpainting for irregular holes using partial convolutions. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 85–100, 2018. 2
- [22] M. Maggioni, V. Katkovnik, K. Egiazarian, and A. Foi. Non-local transform-domain filter for volumetric data denoising and reconstruction. *IEEE Transactions on Image Processing*, 22(1):119–133, Jan 2013. 4
- [23] S. Mei, X. Yuan, J. Ji, Y. Zhang, S. Wan, and Q. Du. Hyperspectral image spatial super-resolution via 3d full convolutional neural network. *Remote Sensing*, 9(11):1139, 2017. 2, 4
- [24] H. Othman and Shen-En Qian. Noise reduction of hyperspectral imagery using hybrid spatial-spectral derivative-domain wavelet shrinkage. *IEEE Trans. on Geoscience and Remote Sensing*, 44(2):397–408, Feb 2006. 2, 4
- [25] Y. Qu, H. Qi, and C. Kwan. Unsupervised sparse dirichlet-net for hyperspectral image super-resolution. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2511–2520, 2018. 2
- [26] N. Renard, S. Bourennane, and J. Blanc-Talon. Denoising and dimensionality reduction using multilinear tools for hyperspectral images. *IEEE Geoscience and Remote Sensing Letters*, 5(2):138–142, April 2008. 2, 4
- [27] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: nonlinear phenomena*, 60(1-4):259–268, 1992. 2
- [28] H. Shen, X. Li, L. Zhang, D. Tao, and C. Zeng. Compressed sensing-based inpainting of aqua moderate resolution imaging spectroradiometer band 6 using adaptive spectrum-weighted sparse bayesian dictionary learning. *IEEE Transactions on Geoscience and Remote Sensing*, 52(2):894–906, Feb 2014. 2
- [29] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *IEEE International Conference on Computer Vision, ICCV '98*, pages 839–, Washington, DC, USA, 1998. IEEE Computer Society. 2
- [30] D. Ulyanov, A. Vedaldi, and V. Lempitsky. Deep image prior. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 9446–9454, 2018. 1, 2
- [31] Q. Wang, L. Zhang, Q. Tong, and F. Zhang. Hyperspectral imagery denoising based on oblique subspace projection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6):2468–2480, 2014. 2, 4
- [32] Y. Wang, X. Chen, Z. Han, S. He, et al. Hyperspectral image super-resolution via nonlocal low-rank tensor approximation and total variation regularization. *Remote Sensing*, 9(12):1286, 2017. 2
- [33] B. Xu, N. Wang, T. Chen, and M. Li. Empirical evaluation of rectified activations in convolutional network. *arXiv preprint arXiv:1505.00853*, 2015. 3
- [34] D. Yao, L. Zhuang, L. Gao, B. Zhang, and B.-D. J. M. Hyperspectral image inpainting based on low-rank representation: a case study on Tiangong-1 data. In *IEEE International Geoscience and Remote Sensing Symposium*, 2017. 2
- [35] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang. Generative image inpainting with contextual attention. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 5505–5514, June 2018. 2
- [36] Q. Yuan, Q. Zhang, J. Li, H. Shen, and L. Zhang. Hyperspectral image denoising employing a spatialspectral deep residual convolutional neural network. *IEEE Trans. on Geoscience and Remote Sensing*, 57(2):1205–1218, Feb 2019. 2, 4
- [37] Y. Yuan, X. Zheng, and X. Lu. Hyperspectral image super-resolution by transfer learning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(5):1963–1974, May 2017. 2

- [38] H. Zhang, W. He, L. Zhang, H. Shen, and Q. Yuan. Hyperspectral image restoration using low-rank matrix recovery. *IEEE Transactions on Geoscience and Remote Sensing*, 52(8):4729–4743, Aug 2014. [2](#), [4](#)
- [39] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *Trans. Img. Proc.*, 26(7):3142–3155, July 2017. [2](#)
- [40] L. Zhuang and J. M. Bioucas-Dias. Fast hyperspectral image denoising and inpainting based on low-rank and sparse representations. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(3):730–742, March 2018. [2](#), [4](#)