# Efficient convolutional neural network for spectral-spatial hyperspectral denoising

Alessandro Maffei, Juan M. Haut, *Student Member, IEEE,* Mercedes E. Paoletti, *Student Member, IEEE,* Javier Plaza, *Senior Member, IEEE,* Lorenzo Bruzone, *Fellow, IEEE,* Antonio Plaza, *Fellow, IEEE,*

***Abstract*—Denoising is a common pre-processing step prior to analysis and interpretation tasks such as classification, unmixing and target detection, typically carried out for hyperspectral images (HSIs). In this paper we develop which performs spectral-spatial HSI denoising through a convolutional neural network (CNN). Our newly developed method, called single denoising CNN (HSI-SDeCNN), considers HSIs as 3D data cubes, performing the denoising process with only one single model. Experimental results on both synthetic and real data demonstrate that our newly developed HSI-SDeCNN outperforms other state-of-the-art HSI denoising methods.**

***Index Terms*—Hyperspectral images (HSIs), denoising, convolutional neural networks (CNNs), spatial-spectral information.**

## I. Introduction

HYPERSPECTRAL images (HSIs) are characterized by hundreds of spectral bands, acquired across the electromagnetic spectrum in several continuous narrow bands [1], [2]. This kind of high-dimensional data, collected by hyperspectral sensors (spectrometers) can be seen as 3D data cubes, which are characterized by their rich spectral information.

Recent Earth observation missions have acquired huge amounts of HSI data [3], fostering their use in a wide range of application domains [4], including classification [5], spectral unmixing [6] and target detection [7], among many others.

Since the acquisition process is noisy, leading to intra-class variability and inter-class similarity [8], a pre-processing denoising step prior to HSI data interpretation is typically carried out, in order to remove the significant amount of introduced noise.

In this paper, an improved CNN [9] architecture is developed to efficiently perform HSI denoising, called HSI single-denoising CNN (HSI-SDeCNN). The proposed model, rather than focusing only on spatial correlation as in 1D and 2D models, takes advantage of the high spectral correlation between adjacent bands present in HSIs. The newly developed HSI-SDeCNN performs denoising one band at time, taking as input data cubes rather than single bands. In this way it exploits the high correlation between adjacent bands, in order to recover the missing information corrupted during the acquisition process. Roughly speaking, the proposed network takes as input $K+1$ bands, (being $K$ the number of adjacent bands) and returns the central noise-free band.

The main advantages of the newly proposed HSI-SDeCNN with respect to existing HSI denoising methods, can be enumerated as follows:

1) It provides a fast solution to the HSI denoising problem, exploiting a down-sampling kernel that allows the network to perform very fast without losing in term of denoising performance.
2) It takes as input a noise-level map, i.e., an estimation of the noise-level present in the HSI to be denoised, which allows us to control the trade-off between denoising performance and detail preservation. This makes our network flexible and adaptive to multi-level noise, as well as to spatially-variant noise, without the need to train different models for different noise levels.
3) It outperforms other state-of-the-art HSI denoising methods, on both synthetic and real HSI images, demonstrating its full potential for practical HSI denoising applications.

## II. Methodology

In this section we present the HSI-SDeCNN model and how it is applied to HSIs. Hyperspectral data can be processed as 1D (spectral), 2D (spatial) or 3D (spectral-spatial) models [10]. Since both spatial and spectral information are useful in the noise removal process, we consider the HSI as a 3D data structure, i.e., a spectral-spatial model. In particular each HSI can be seen as a volume of size $h \times w \times B$, being $h$ and $w$ the spatial dimensions of the image and $B$ the spectral one.

The adopted strategy scans the spectral dimension in a raster way and performs denoising one band at a time. For this purpose, we use a network that receives as input a volume of size $h \times w \times K+1$, where $K$ is the number of adjacent bands with respect to the central one (i.e. the target band). The overall structure of the proposed network can be divided into four main operations: i) downsampling, ii) noise-level map concatenation, iii) non-linear mapping, and iv) upsampling.

The first operation (i) is performed by a downsampling kernel that reshapes the input HSI volume of size $h \times w \times K+1$ into downsampled sub-cubes. The scale-factor is set to 2. Roughly speaking, this operation downscales each band in 4 sub-cubes, each with size $w/2 \times h/2$, This process is applied

This work has been supported by the Spanish Ministry (FPU14/02012-FPU15/02090), and Junta de Extremadura (GR15005). Funding from MINECO project TIN2015-63646-C5-5-R is gratefully acknowledged.

A. Maffei and L. Bruzone are with the Remote Sensing Laboratory, Department of Information Engineering and Computer Science, University of Trento, 38123 Trento, Italy (e-mail: alessandro.maffei@studenti.unitn.it, lorenzo.bruzzone@unitn.it).

M. E. Paoletti, J. M. Haut, J. Plaza and A. Plaza are with the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, Escuela Politécnica, University of Extremadura, 10003 Cáceres, Spain (e-mail: mpaoletti@unex.es; juanmariohaut@unex.es; jplaza@unex.es; aplaza@unex.es).

to all the spectral channels, and the obtained sub-cubes are concatenated along the spectral dimension. This operation allows the network to be faster, without losing denoising performance.

The second layer of the network (ii) concatenates a noise-level map $M$ to the generated sub-cubes, obtaining a volume of size $h/2 \times w/2 \times 4(K+1)+1$. The noise-level map gives an estimation of the noise level $\sigma$ present in the image. It is inserted as a uniform map (all elements equal to $\sigma$) with the same spatial dimension than the sub-cubes, in order to avoid any mismatch in terms of dimensionality [11]. The noise-level map is given in input to a CNN as prior information, allowing the network to control the trade-off between denoising performance and detail preservation. This is because, opposed to common residual learning methods, adding a noise-level map makes the model parameters (i.e. weights and biases) invariant to the noise level of the input image. This is one of the main feature of the proposed method, that allows the network to handle different noise levels, as well as spatially-variant noise, with only one single model, only changing the input-noise level map $M$.

At this point, the obtained volume is fed to a standard CNN (iii) in order to learn a non-linear mapping function, able to recover the noise-free image from the input noisy data cube.

The final layer (iv) of our HSI-SDeCNN method is an upsampling kernel that takes as input the four downsampled, noise-free images (from the CNN) and outputs a clean single band. It performs an inverse operation with regards to the downsampling layer.

In summary, our HSI-SDeCNN method is inspired by a network typically used for grayscale and RGB images (named FFDNet), which has been adapted to perform denoising of hyperspectral images. The main reasons why the proposed method exhibits better performance that the standard network are twofold. First and foremost, our method takes as input a significantly larger number of bands, which allows the network to exploit the high spectral correlation between adjacent channels applying 3D convolutions instead of 2D (spatial) ones. Second, since our network considers overlapping bands, it can learn from a larger amount of data, resulting in much stronger denoising performance. In fact, the proposed HSI-SDeCNN exhibits more powerful performance (both in terms of denoising and computational time) when compared to other learning-based methods. This is mainly due to the downsampling layer, which allows the network to be faster without degrading performance, and also to the input noise-level map, which allows the CNN to achieve better denoising performance without the need to train different models for different noise levels.

## III. Experiments and Discussion

We have evaluated our newly proposed method using both synthetic and real HSIs. First, the effectiveness of the method has been verified through simulated experiments. Then, the method has been applied to real noisy images. In particular, the newly developed HSI-SDeCNN has been compared to the current mainstream approaches typically adopted in HSI denoising: hybrid spatial-spectral noise reduction (HSSNR) [12], low-rank tensor approximation (LRTA) [13], block matching and 4-D algorithm (BM4D) [14], low-rank matrix recovery (LRMR) [15], MH-prediction [16] and HSI denoising exploiting a spatial-spectral deep residual CNN (HSID-CNN) [17].

In order to train the proposed model, we have selected a part of the noise-free Washington DC Mall image (Fig. 1), obtained by the Hyperspectral Digital Imagery Collection Experiment (HYDICE) airborne sensor. The total Washington DC Mall image has been divided in two parts: one for training the proposed network and the other for testing purposes. In particular, experiments have been conducted on the following datasets:

- Washington DC Mall. A cropped part of the entire image has been employed for simulated experiments in which additive white gaussian noise (AWGN) is added to the original image.
- Indian Pines. This dataset, acquired by the Airborne Visible Infra-Red Imaging Spectrometer (AVIRIS), has been employed to test the effectiveness of the method in a real scenario.

### A. Simulated Data Experiments

In this section we present how the simulated experiments has been carried out and the results obtained on the Washington DC Mall dataset (test image). In order to perform the experiments, AWGN noise has been added to the noise-free HSI using the following strategy: we consider the same maximum level of noise for each band, where $\sigma_n = [5, 100]$. Here, $n \in [1, B]$ indicates a generic band. In order to evaluate the performances from a quantitative point of view three commonly employed metrics have been adopted: MPSNR (mean peak signal-to-noise-ratio), MSSIM (mean structural similarity index), and MSA (mean spectral angle).

.

In all our simulated experiments, we have set the input noise-level map $M$ to the same level of the noise added to the image (we refer to this as ground-truth noise). Lower performances are obtained when the noise level map given in input to the network differs from the actual noise level present in the image. Roughly speaking, when we set the input noise level to be higher than the groundtruth noise. This means that we perform too much denoising, smoothing out some image details. On the other hand, if the input noise-level is lower than the ground-truth one, less denoising is performed, leaving some residual noise in the output image. Thus, a correct setting of the noise-level map (i.e., the input noise-level) is fundamental to obtain good performance. However, using a noise-level map that is slightly different from the ground-truth noise does not introduce significant performance degradation. This means that the network is flexible in the choice of this parameter. All the results obtained with HSI-SDeCNN are extracted with only one model, trained with different level of noise from 0 to 100. Table I shows the results of the simulated experiments obtained for noise levels 25 and 100, conducted with our method and also with other mainstream methods. In particular, our method outperforms all the other methods, but using only a single model.

TABLE I
QUANTITATIVE EVALUATION OF THE PROPOSED METHOD AGAINST THE MAINSTREAM METHODS FOR HSI DENOISING (SIMULATED EXPERIMENTS).

| Noise Level | Index | HSSNR | LRTA | BM4D | LRMR | HSID-CNN | Proposed |
|---|---|---|---|---|---|---|---|
| $\sigma_n = 25$ | MPSNR | $28.02 \pm 0.002$ | $30.67 \pm 0.003$ | $31.14 \pm 0.003$ | $33.03 \pm 0.002$ | $33.05 \pm 0.003$ | $\mathbf{33.44 \pm 0.008}$ |
| | MSSIM | $0.9461 \pm 0.0001$ | $0.9629 \pm 0.0002$ | $0.9685 \pm 0.0002$ | $0.9809 \pm 0.0001$ | $0.9813 \pm 0.0001$ | $\mathbf{0.9822 \pm 0.0000}$ |
| | MSA | $8.13 \pm 0.003$ | $5.80 \pm 0.006$ | $5.05 \pm 0.005$ | $4.61 \pm 0.003$ | $4.26 \pm 0.003$ | $\mathbf{3.91 \pm 0.008}$ |
| $\sigma_n = 100$ | MPSNR | $16.314 \pm 0.0065$ | $23.175 \pm 0.0048$ | $22.577 \pm 0.0054$ | $24.310 \pm 0.0047$ | $25.296 \pm 0.0043$ | $\mathbf{25.753 \pm 0.0121}$ |
| | MSSIM | $0.6049 \pm 0.0001$ | $0.8494 \pm 0.0003$ | $0.8119 \pm 0.0002$ | $0.8799 \pm 0.0002$ | $0.9014 \pm 0.0001$ | $\mathbf{0.9121 \pm 0.0002}$ |
| | MSA | $24.732 \pm 0.0065$ | $9.1219 \pm 0.0072$ | $9.7611 \pm 0.0068$ | $10.468 \pm 0.0074$ | $8.4061 \pm 0.0063$ | $\mathbf{7.3951 \pm 0.0182}$ |

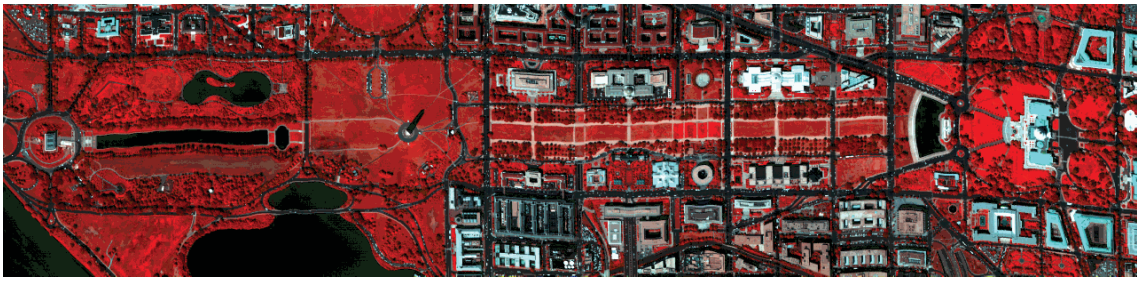

Fig. 1. Washington DC Mall, a color composite representation of the hyperspectral data employed for training and simulated experiments.

For visual comparison purposes, we have selected bands (57, 27, 17) to generate pseudo-color images. In Fig. 2 we display the visual results of one zoomed region (cropped from the Washington DC Mall test image) obtained with $\sigma_n = 100$.

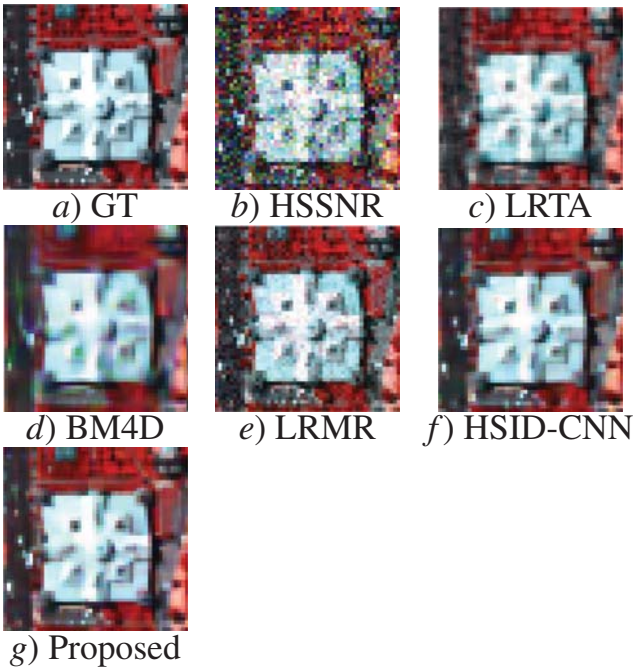


Fig. 2. Zoomed denoising results for the Washington DC Mall image (simulated experiments, with $\sigma_n = 100$). Bands (57, 27, 17) are selected to generate false-color images.

We can see that HSID-CNN and our proposed method obtain much better visual results than those obtained with the other methods. In particular, the denoised images provided by HSSNR and LRMR present residual noise, while the images produced by BM4D and LRTA contain visual artifacts. Instead, HSID-CNN and the proposed HSI-SDeCNN generate denoised images that are very similar to the ground-truth one. However with an accurate analysis and especially in presence of details our method exhibit much powerfull performances.

Thus, both a visual and quantitative standpoint, we can conclude that our method outperforms all the other mainstream methods in our simulated experiments.

### *B. Real Data Experiments*

In this section we discuss the experiments performed on the Indian Pines real HSI. In order to verify the effectiveness of the proposed method classification experiments are conducted. This because a ground-truth, noise-free image is not available for real data. As a result, the performance of the method is measured by analyzing the classification accuracy before and after the denoising process, on 16 ground-truth classes. The metrics adopted are the overall accuracy (OA) and the kappa coefficient. A support vector machines (SVM) with linear kernel has been employed as a simple classifier in our experiments.

For the training of the classifier, we randomly select $10\%$ of the available labeled samples from each class, and use the remaining labeled samples for testing purposes.

Since the noise level is unknown in real HSIs, the proposed denoising algorithm has been applied by empirically setting the input noise level-map to the one that shows the best performance among the following noise levels: $\sigma = 5, 25, 50, 75, 100$. In particular, we have found that the best results are achieved setting a noise-level map with $\sigma = 50$.

The obtained results are shown in Table II, where is possible to see the high improvement achieved with the proposed method, going from an OA of $75.96\%$, obtained without any denoising process, to an OA of $95.58\%$.

### *C. Running Time*

In order to evaluate the computational efficiency of the proposed denoising algorithm, we compare the running times

TABLE II
CLASSIFICATION RESULTS OBTAINED AFTER DENOISING THE INDIAN PINES IMAGE USING DIFFERENT METHODS

| | **Original** | **Proposed** |
|---|---|---|
| **OA** | 75.96 | **95.58** |
| **Kappa (x100)** | 72.20 | **94.97** |

of the proposed HSI-SDeCNN with regards to HSID-CNN. This because the HSID-CNN has achieved the best results in terms of computational time among the other methods (see results in [17]). The running time has been calculated on the Washington DC Mall dataset (i.e., simulated experiments), and the obtained results are provided in Table III. Specifically, we can observe that our method is more than two times faster than the HSID-CNN.

TABLE III
AVERAGE RUNTIME (IN SECONDS) MEASURED FOR THE HSID-CNN AND THE PROPOSED HSI-SDECNN METHOD.

| **Dataset** | **Size** | **HSID-CNN** | **Proposed** |
|---|---|---|---|
| Wash. DC Mall | $200 \times 200 \times 191$ | $7.23 \pm 0.016$ | $\mathbf{3.08 \pm 0.024}$ |

## IV. CONCLUSION AND FUTURE LINES

In this paper, we have presented a new learning-based method for HSI denoising, called single denoising convolutional neural network (HSI-SDeCNN). This method takes as input a data cube instead of a single band forcing the network to consider both spatial and spectral correlation present in HSIs. The two main features of the proposed network are: a downsampling layer that allows the network to be faster without losing in terms of denoising performance, and a noise-level map that is used to give as input to the network an estimation of the noise. Our newly developed method outperforms other mainstream methods commonly adopted in HSI denoising, with only one single model. However, as with any new approach, there are still some future research avenues that can be further explored. Specifically, our proposed network makes the denoising at only one level for all the bands (i.e. same input noise level for all the bands). However, in HSIs normally the noise differs from one band to another. For this reason, a further improvement of the method can focus on adapting the input noise-level to each specific band.

## REFERENCES


[1] R. P. Iyer, A. Raveendran, S. K. T. Bhuvana, and R. Kavitha, "Hyperspectral image analysis techniques on remote sensing," in *2017 Third International Conference on Sensing, Signal Processing and Security (ICSSS)*, May 2017, pp. 392–396.

[2] T. Ado, J. Hruka, L. Pdua, J. Bessa, E. Peres, R. Morais, and J. Sousa, "Hyperspectral imaging: A review on uav-based sensors, data processing and applications for agriculture and forestry," *Remote Sensing*, vol. 2017, p. 1110, 10 2017.

[3] J. Transon, R. dAndrimont, A. Maugnard, and P. Defourny, "Survey of hyperspectral earth observation applications from space in the sentinel-2 context," *Remote Sensing*, vol. 10, p. 157, 01 2018.

[4] J. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *Geoscience and Remote Sensing Magazine, IEEE*, vol. 1, pp. 6–36, 06 2013.

[5] M. E. Paoletti, J. M. Haut, R. Fernandez-Beltran, J. Plaza, A. Plaza, J. Li, and F. Pla, "Capsule networks for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 4, pp. 2145–2160, April 2019.

[6] R. Fernandez-Beltran, A. Plaza, J. Plaza, and F. Pla, "Hyperspectral unmixing based on dual-depth sparse probabilistic latent semantic analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 11, pp. 6344–6360, Nov 2018.

[7] B. Yang, M. Yang, A. Plaza, L. Gao, and B. Zhang, "Dual-mode fpga implementation of target and anomaly detection algorithms for real-time hyperspectral imaging," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 6, pp. 2950–2961, June 2015.

[8] P. Ghamisi, N. Yokoya, J. Li, W. Liao, S. Liu, J. Plaza, B. Rasti, and A. Plaza, "Advances in hyperspectral image and signal processing: A comprehensive overview of the state of the art," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 4, pp. 37–78, Dec 2017.

[9] M. Paoletti, J. Haut, J. Plaza, and A. Plaza, "A new deep convolutional neural network for fast hyperspectral image classification," *ISPRS journal of photogrammetry and remote sensing*, vol. 145, pp. 120–147, 2018.

[10] B. Rasti, P. Scheunders, P. Ghamisi, G. Licciardi, and J. Chanussot, "Noise reduction in hyperspectral imagery: Overview and application," *Remote Sensing*, vol. 10, p. 482, 03 2018.

[11] K. Zhang, W. Zuo, and L. Zhang, "Ffdnet: Toward a fast and flexible solution for cnn-based image denoising," *IEEE Transactions on Image Processing*, vol. 27, no. 9, pp. 4608–4622, Sep. 2018.

[12] H. O. and, "Noise reduction of hyperspectral imagery using hybrid spatial-spectral derivative-domain wavelet shrinkage," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 2, pp. 397–408, Feb 2006.

[13] C. Li, Y. Ma, J. Huang, X. Mei, and J. Ma, "Hyperspectral image denoising using the robust low-rank tensor recovery," *Journal of the Optical Society of America A*, vol. 32, 09 2015.

[14] C. Jiang, H. Zhang, L. Zhang, H. Shen, and Q. Yuan, "Hyperspectral image denoising with a combined spatial and spectral weighted hyperspectral total variation model," *Canadian Journal of Remote Sensing*, vol. 42, no. 1, pp. 53–72, 2016.

[15] 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*, vol. 52, no. 8, pp. 4729–4743, Aug 2014.

[16] C. Chen, W. Li, E. W. Tramel, M. Cui, S. Prasad, and J. E. Fowler, "Spectralspatial preprocessing using multihypothesis prediction for noise-robust hyperspectral image classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 4, pp. 1047–1059, April 2014.

[17] Q. Yuan, Q. Zhang, J. Li, H. Shen, and L. Zhang, "Hyperspectral image denoising employing a spatial-spectral deep residual convolutional neural network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 2, pp. 1205–1218, Feb 2019.