IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 25 March 2023, accepted 25 April 2023, date of publication 5 May 2023, date of current version 12 May 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3273486

RESEARCH ARTICLE

Deep Convolutional Network Aided by Non-Local Method for Hyperspectral Image Denoising

GABRIEL A. DE OLIVEIRA[1], LARISSA MEDEIROS DE ALMEIDA[2], EDUARDO R. DE LIMA[3], AND LUÍS GERALDO P. MELONI[1], (Member, IEEE)

[1]DECOM/FEEC, University of Campinas, Campinas, S ão Paulo 13083-852, Brazil
[2]Transmissora Aliança de Energia Elétrica S.A.—TAESA, Rio de Janeiro 20010-010, Brazil
[3]Department of Hardware Design, Instituto de Pesquisas Eldorado, Campinas, São Paulo 3083-898, Brazil

Corresponding author: Luís Geraldo P. Meloni (meloni@unicamp.br)

This work was supported by the Research and Development Project "Predictive Failure Analysis by Artificial Intelligence," funded by Transmissora Aliança de Energia Elétrica SA (TAESA) with resources from Agência Nacional de Energia Elétrica (ANEEL) Research, Development, and Innovation Program under Grant PD-07130-0062/2020.

ABSTRACT This paper introduces a new hyperspectral image denoising method called Non-local Convolutional Neural Network Denoiser (NL-CNND). The technique exploits data in four bands adjacent to the target one as additional information for the restoring process, and it uses a pre-denoising step based on BM4D. All the bands paired with their pre-denoised versions in a second step feed a Convolutional Neural Network. To network generalization, one of the inputs is the noise level of the input image, allowing a single model to work with different noise levels. This restoration technique overcomes quality when compared to current eight classical and neural methods. The results show higher peak signal to noise ratio, structural similarity index, and spectral angle mapper metrics than all the other restoration methods, surpassing those achieved using Block Matching and 4D Filtering alone. Besides, the results show a higher level of detail visually while at the same time reducing over-smoothing on the input images' features. The paper also includes an algorithm for complete image restoration, allowing for denoising full-sized hyperspectral images independent of their shape. The dataset creation used for network training is detailed, based on a small set of available hyperspectral images, encompassing data normalization, conversion, and storage.

INDEX TERMS Hyperspectral images, denoising, BM4D, convolutive neural network, NLM, band correlation.

I. INTRODUCTION

Nowadays, a great variety of sensors are used in different applications with the massive penetration of the Internet of Things. Among the most common examples are image sensors, providing data-rich in visually discernible details humans can interpret and analyze. Computers, however, can go further, extracting more information using the appropriate processes [1]. One of the many digital imaging applications involves capturing emitted or reflected electromagnetic waves from a target material, commonly present in remote sensing and land surveying [2], [3]. As it mimics human vision, conventional images are restricted to three different electromagnetic spectrum bands related to the red, green, and blue colors. Due to this type of capture, color images limit the analysis regarding the frequency response of captured scenes [1], [4]. To address this shortcoming, HSIs (hyperspectral images) offer another approach to image acquisition. Unlike its color counterparts, this capture type has no limited number of spectral bands. Instead, it divides a chosen range of the magnetic spectrum contiguously into several bands of the same width [5], [6]. Multi-band images reveal further details about their subjects that would not be present in color captures due to the exposure to many bands. Due to the physical properties of the captured materials, their absorbance and reflectance vary between the many bands captured by a hyperspectral imaging device. Consequently, this allows for classification and segmentation processes that extend the possibility of color images alone. For example, water tends to appear on lower frequencies of the spectrum, while land

The associate editor coordinating the review of this manuscript and approving it for publication was Ikramullah Lali.

TABLE 1. Denoising performance comparison for the tested methods for Gaussian noise of three different standard deviations. Their performance was evaluated using three different metrics, PSNR SSIM and SAM. The best results of the proposed method are highlighted in bold.

		BM4D	HyRes	fastHyDe	HyMiNor	rPCA	GRN	HSIDCNN	QRNN3D	MEMNET-Hy	NL-CNND
$\sigma = 30$	PSNR	35,689	27,180	33,870	28,739	16,225	22.96	28,935	30,093	24,648	**37,096**
	SSIM	0,961	0,714	0,962	0,794	0,181	0,523	0,849	0,848	0,583	**0,974**
	SAM	0,064	0,229	0,077	0,174	0,668	0,252	0,186	0,157	0,339	**0,046**
$\sigma = 50$	PSNR	30,480	25,787	32,126	27,320	16,155	22,116	28,413	28,309	23,316	**35,580**
	SSIM	0,911	0,651	0,945	0,741	0,176	0,664	0,840	0,783	0,513	**0,963**
	SAM	0,083	0,243	0,084	0,182	0,671	0,269	0,178	0,177	0,365	**0,059**
$\sigma = 70$	PSNR	26,570	24,450	30,397	25,893	15,790	19,732	27,691	26,534	22,001	**33,334**
	SSIM	0,859	0,590	0,924	0,685	0,165	0,594	0,824	0,707	0,448	**0,951**
	SAM	0,096	0,253	0,090	0,188	0,682	0,291	0,171	0,199	0,388	**0,087**

and crops appear throughout the spectrum in different levels across bands [7], [8].

Remote sensing has various applications, one of which is the non-invasive analysis of environments where direct measurements are unfeasible. This technique proves particularly useful for monitoring hazardous locations, such as power substations, for partial discharges and corona effects, which are indicative of potential equipment failure. Nevertheless, the effectiveness of automated monitoring is hindered by higher levels of noise in input data, which is a persistent issue in image capture, and more pronounced in hyperspectral images [9], [10].

In general, denoising comes before classification tasks in the image processing pipeline. As noise is intrinsic to image capture, many of the works in the literature suggest different ways to mitigate the various image corruption types. However, the high correlation between close bands can be exploited for hyperspectral images and is the core of many methods. Classical methods, which here are the ones unrelated to machine learning, are present ubiquitously in the literature and already been tried and validated, yielding good results. As examples of classical non-local methods, there are BM4D (Block Matching and 4D filtering) [11], and RPCA (Robust Principal Component Analysis) [12]. Unlike local methods, which are limited to only information present in adjacent pixels, these methods can apply more data by comparing the desired pixel to other similar areas of the image that have eventually been corrupted by noise. However, in recent years, the focus has shifted following the significant expansion of machine learning techniques, which have extended their usefulness to image processing. Newer methods such as those presented in [13] and [14] use the concept of deep convolutional neural networks to learn what consists of noise and what does not, resulting in an inference that restores images with greater detail and accuracy than classical methods.

This paper is organized as follows. After the current section, section II explains the motivation behind the work and the paper's major contributions. In section III, some background topics for better paper understanding are explored, briefly introducing the denoising methods used for result comparison. The section follows, discussing image characteristics, different types of noise, and the noise chosen model. Section IV presents the proposed denoising method and discusses how the network was constructed and trained, how bigger images were reassembled, and how images were pre-processed and rendered into a usable dataset. The results are discussed in section V, comparing the proposed model with nine other methods using objective metrics and a naked-eye analysis. The conclusion in Section VI summarizes the findings with notes on the meaning of results and their impact while presenting insight for future works.

II. MOTIVATION AND MAJOR CONTRIBUTIONS

Hyperspectral images are notoriously noisy and easily corrupted, given their specific sensor configurations and special image capture restrictions to accommodate multiple bands. Thermal noise is one of the most common types of image corruption and is the problem this article aims to address. A new method based in convolutional neural networks to reconstruct hyperspectral images from their noisy observations is presented. The method is an amalgam of different processes, updating them to work with hyperspectral images and changing their pre-processing step for a better-suited denoiser. BM4D acts as the first denoising step to make the network more robust, followed by the neural network itself. In FIGURE 1, the full process is laid out, from the inputs to the final restored image. It restores a central band of interest, using it and four additional adjacent bands to explore redundant information present in closer bands. Noise level information is also provided to the network to aid this process and make it more flexible, allowing using the neural network in different noise scenarios. This work differs from [15] and [14] as not being fixed to a noise level. It further differentiates itself from these works as the kernel depth used (5) is reduced compared to theirs (24 and 31, respectively), which lowers network complexity without sacrificing image quality. TABLE 1 shows the results of the proposed method, which fares better than other tested methods, classical and neural, including BM4D, which is part of its construction.

III. BACKGROUND TOPICS

This section presents several state-of-the-art noise reduction methods for comparison purposes with the proposed method.

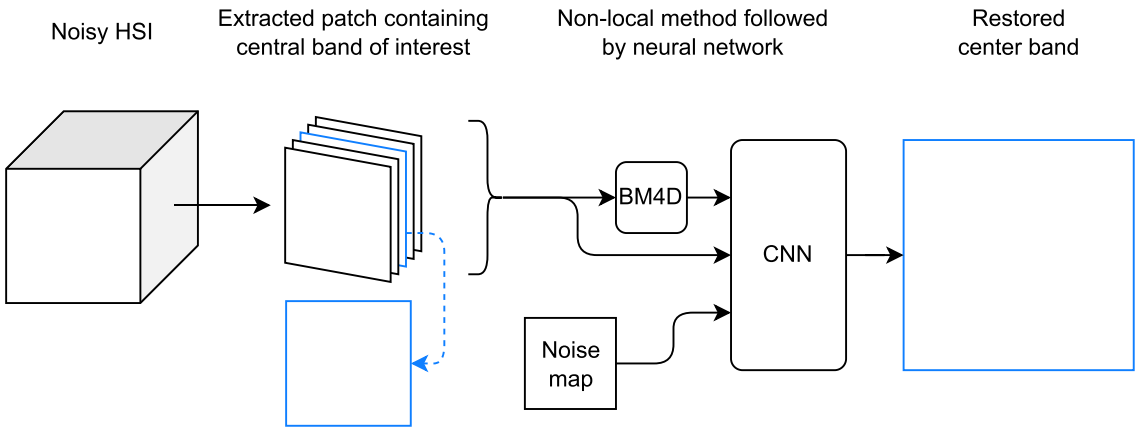

FIGURE 1. Summary of the proposed method: A chunk of a noisy HSI (Hyperspectral Image) is denoised using a total of five bands, a center band of interest and four adjacent bands. These bands are used as input for the convolutional neural network, having as its inputs the noisy chunk, the pre-denoised version of the chunk, and a noise map that contains the standard deviation of the noise corrupting the input image. The process's output is a restoration of the central band of interest.

They are followed by briefly explaining the various types of image corruption.

A. DENOISERS

This section reviews several denoising methods shown in TABLE 1. Classical and neural methods are present, showing their advances and how classical methods sometimes perform better than those based on machine learning. The proposed method was compared to the following classical methods: BM4D [11], HyRes [16], fastHyDe [17], HyMiNoR [18], rPCA [12], with default settings for Gaussian denoising. The following neural reconstruction methods were applied: GRN [19], HSIDCNN [14], QRNN3D [20], and MemNet (with a HyRes step) [20], [21].

1) BM4D

In contrast to local-based denoisers, non-local denoisers search for additional pixel information outside the neighborhood of a particular pixel. Block Matching and 3D Filtering (BM3D), a predecessor method, searches inside the target image for 2D patches that share the same characteristics. The assumption behind this is that the difference between those patches is mostly noise. These similar patches are stacked in a 3D configuration (hence the method's name). Denoising involves two steps: first, a thresholding process removes extreme values, followed by a second filtering process based on a Wiener filter [22].

This concept is extended for images with depth, such as color and hyperspectral images, searching these images for 3D chunks of high similarity, which are stacked in a 4D vector. Once again, a two-step process involving thresholding and Wiener filtering is applied. This process also assumes that the difference between the three-dimensional chunks is mostly noise.

2) HyRes

Based on low-rank modeling, Hyperspectral Restoration (HyRes) is a parameter-free method that estimates an unknown signal based on L1 penalization in the minimization problem. This type of denoising is widely used in HSI, exploring redundancy between the many different bands making a good low-rank approximation of the entire image.

3) FastHyDe

This method explores the relationship between data present on HSI and extremely compact representations of them. The sparsity of those representations is linked to their low-rank nature alongside their self-similarity characteristics. In other words, it exploits the same image characteristics of HyRes, and the difference lies in the method's speed, which uses a pre-learned subspace where its coefficients are self-similar. This allows for the denoising of these components by non-local patch-based denoisers.

4) HyMiNoR

Hyperspace Mixed Gaussian and Sparse Noise Reduction is a two-step process that first filters Gaussian noise using HyRes as the base and then offers a novel sparse noise removal (such as salt and pepper, missing pixels, and lines). This technique neglects sparse components, using an optimization problem based on L1-L1 norms. This optimization uses the L1 norm as the fidelity term instead of the L2 norm, as noise is by nature sparse. The penalization term is also an L1 norm in the spectral difference matrix to exploit the high correlation between the many bands of the target HSI. The optimization problem uses a split Bregman technique.

5) rPCA

Robust Principal Component Analysis extends the PCA statistical tool, which is hardened against sparse noise. In contrast, corrupted data renders the original PCA unusable, resulting in restoration that is, in Euclidean terms, far from the desired result. The Principal Component Pursuit recovers both the low rank and sparse components, minimizing a combination of nuclear and L1 norm, recovering a low-rank matrix L_0.

6) GRN

GRN is a Deep Spatial-Spectral Global Reasoning Network based on a U-Net architecture for hyperspectral image denoising. It differs from other works as it is a non-local neural network, where two global reasoning modules learn global spatial relationships and global interdependencies across channels of the feature map.

7) HSIDCNN

HSIDCNN employs HSI denoising as a spatial-spectral deep residual convolutional neural network where not only the spatial information is used as input for the network but also the spectral information. The 2D part of the method learns the features of a single band of interest, whereas the 3D part exploits the high correlation between adjacent bands.

8) QRNN3D

QRNN3D is a quasi-recurrent neural network aimed at hyperspectral reconstruction. It bases itself on 3D convolutions to extract and represent features in a lower dimensional space. This method also employs a network architecture similar to a U-Net, with the novelty that the skip connections alternate directions between the layers of the encoder and the decoder.

9) MemNet + HyRes

MemNet is a persistent memory network for image restoration, which adds a memory block to preserve multi-level representations across the large number of bands present in an HSI. The gate unit controls how much these representations influence the final image. In addition, this method includes a pre-processing step based on the previously mentioned HyRes.

B. IMAGE CORRUPTION AND NOISE MODEL

Image sensors, from simple to advanced, suffer from some level of noise that disfigures their captures [23]. Due to the way imaging sensors are built, the many field-effect transistors present in the construction of CMOS sensors can be subject to several effects that corrupt the true captured image. Noise comes in different flavors and the reconstruction problem pertaining to the original image may be different for each one of them [9]. Hyperspectral imaging presents an even harder challenge, as ranging from the sensor to the optical apparatus, the hardware has to be built differently. This higher complexity however allows for the capture of a larger number of bands, which ranges from a hundred to two hundred usually.

Thermal and quantization noise affect images in a way that does not depend on the image captured, because of this they are called signal-independent noise. For a hyperspectral image, this means there is no correlation between the noise affecting its bands. When this happens, the sensor which in essence is a photon counter miscounts the amount of light arriving individually at pixels, reconstructing a distorted version of the captured image [23]. This effect is exacerbated in captures where the subject is in low light, as the bulk of the photon count may not come from the captured scene, but from thermal-induced false positives. More expensive sensors, such as the ones attached to telescopes, are actively cooled in order to lower the amount of captured noise. Both types of noise (thermal and quantization) can be modeled as a Gaussian distribution [24], thus its effects influence a target image additively. Again, this noise has no correlation between the multiple bands of an HSI, however, its spectral signature remains the same across all of them.

There are some types of noise on the other hand that are influenced by the capture itself, such as Photonic or Shot Noise. It is the effect of the discrete nature of an electric charge, especially when an experiment is repeated a few times, such as low-light photon counting. It depends on the level of the signal, which is reflected by the variance of this noise. Furthermore, it is usually modeled by a Poisson distribution.

Another important type of noise is Sparse noise. It relates to corruption happening in only parts of the captured image, not affecting it entirely due to sensor malfunction. Salt and Pepper noise, for example, happens in some of the pixels maximizing or minimizing the value of certain pixels, appearing as black and white dots in an image. At times, entire lines are also missing from the final capture.

In case of reconstruction error, the way images are captured might also affect the final image. These are called Striping Noise and Fixed Pattern Noise. A sensor that scans lines from the image sequentially for all depths at the same time, known as push-broom, cause striping noises which can be traced back to calibration error or a varying sensitivity of the sensor given the captured dynamic range [25]. This results in lines of different contrast levels appearing along the image.

In this work, the proposed architecture aims to suppress signal-independent noises of Gaussian nature. Some other noise types may also be removed by this construction, as do other methods in the literature similarly. However, our network was not trained nor tested for them.

IV. THE PROPOSED DENOISING METHOD: NL-CNND

The process presented in this paper consists of a two-part denoising method for enhanced HSI restoration. In the first step, a classical method is used, the BM4D. The second step uses the first results along with the original noisy image. The output of the second step is the restored central band.

As the first step is decoupled from the second, other methods for further image restoration improvement may be applied. Currently, BM4D is considered the state-of-the-art in non-neural denoising.

A. TWO-STEP DENOISING MODEL

The first step of the proposed denoising pipeline is image restoration using a non-local method. An essential feature of non-local methods is to use not only the information present at a pixel's vicinity, but also to compare features of a selected area with other areas of the image evaluated as having high

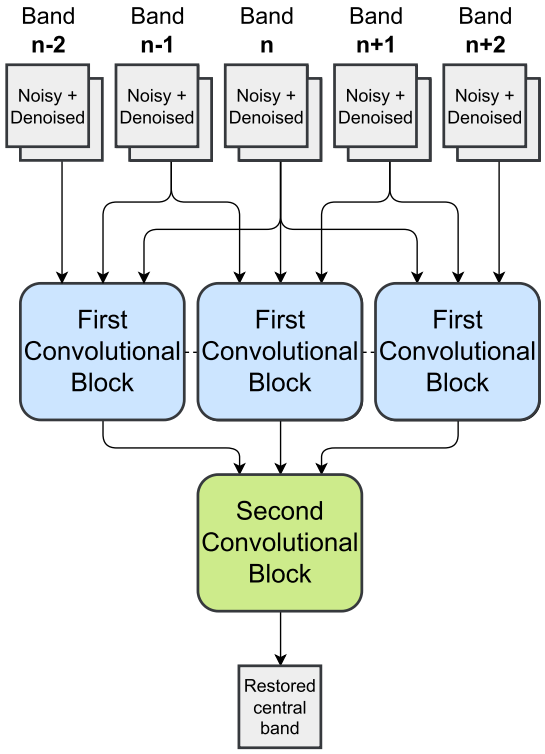

FIGURE 2. A top view of the neural network architecture showing the cascading denoising blocks that comprise it. For the inputs, each noisy band is fed alongside its pre-denoised one.

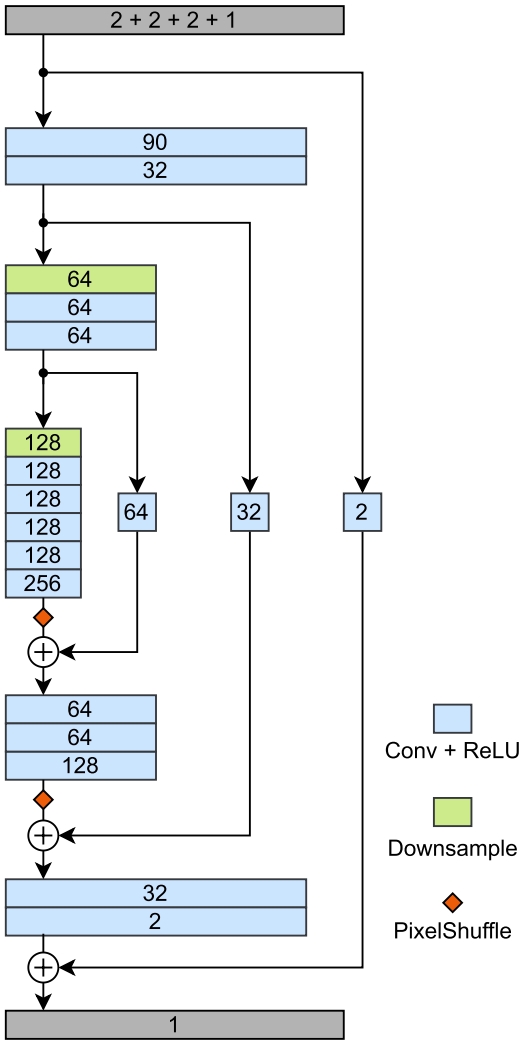

FIGURE 3. A look into the innards of the convolutional block. These take three bands, each pair containing their noisy and pre-denoised versions. The noise map is also used as input. The architecture also features convolutions in their residual connections.

similarity. Our previous method used BM3D as the first step denoiser. A drawback of BM3D is that it can only denoise grayscale images and it is unsuitable for higher-depth image denoising for not considering the correlation between different depths. The proposed method currently uses BM4D as a substitute for the previous one, which is more suitable for HSI containing many bands. The denoising process is similar for both techniques. Instead of stacking 2D patches in a 3D vector, as the input images are tridimensional, BM4D stacks similarly classified 3D chunks of the input image in a 4D vector. The noise removal consists of thresholding followed by Wiener filtering in the same manner as BM3D.

In the second step, the network presented is an extension of different works in the literature. First, DNCNN [26] suggests using an autoencoder (which is known as U-Net in reference to its shape) together with residual connections to create a denoising network. An important part of this work is the addition of the noise map, a 2D matrix with the same size of the input noisy image, where all its entries are the value of the standard deviation of the noise affecting the image. Noise mapping allows the model to learn different levels, making for a generalist network that does not need separate training for each noise level.

In [13], which extends the concepts of [26], it is proposed a network construction where the inputs consist not only of the noisy image but also a pre-denoised image using a non-local method. This allows for better reconstruction of the corrupted image without the drawbacks of BM3D alone, such as over-smoothing of features.

However, such previous works are not meant for images with higher depth, such as HSIs. A solution for this problem comes with the FastDVDNet implementation [27]. As it focuses on video denoising, a different approach is taken, using information present in different adjacent frames as the source of redundant information for noise reduction. The network in its core is a U-Net derived from [15], using that construction as sub-nets in a two-step convolutional denoiser. The first step comprises three convolutional autoencoder blocks that share weights, each taking three of the five input frames sequentially and outputting one denoised intermediate frame. The second step follows with another convolution block similar to the ones used in the first step, taking three output frames and outputting a single restored central frame.

In this work, we build a novel network topology encompassing some of the previously discussed features. The U-Net from [15] still shapes the core subnets, using convolutions

as its main operations. The inputs, on the other hand, have been changed to accommodate the first-step denoising data, as shown in FIGURE 2. The overall topology resembles that of [27]. However, instead of video frames, the network receives five bands of the HSI in their noisy and pre-denoised forms. The figure shows an overview of the network topology reduced to its building blocks. The kernel of the convolutional network has Image Shape (IS) of size $128 \times 128 \times 5$. The first three convolutional blocks take three bands each, which together encompass all five input bands. These blocks share neuron weights between themselves. Their output is a single band, which is passed to the second convolution block. This block outputs the restored central band. Reference [27] discussed and proved that this type of cascading of convolutional blocks employs better the information in neighboring bands, offering higher restoration results in this configuration rather than with a single denoising block taking five frames as input.

FIGURE 3 shows the various convolutional layers which shape the first and second convolutional blocks. Both blocks share the same topology, the only difference lies in the sharing of weights at the first three convolutional blocks which does not happen with the second. The subnet consists entirely of convolutional blocks arranged as an auto-encoder. Its defining features are the residual connections, aided by convolutions, and the presence of the pixel shuffle layers, which upsample the inputs converting depth to resolution.

The blocks' inputs are three bands alongside a noise map, feeding the network with information regarding the standard deviation of the noise. Both first-step and second-step convolution blocks receive the noise map.

B. NETWORK TRAINING

A dataset was derived from several hyperspectral images for the network training, which correspond to roughly two hundred thousand IS samples. As discussed in [28], a larger number of samples impacts network performance significantly more than changing hyperparameters, influencing the choice of a more extensive dataset. All the entries are shuffled every epoch so as not to introduce bias due to their ordering, given that the same images are present for four different noise levels. The number of total epochs for the training was not set in advance, and the training stopped monitoring the loss metric with a patience parameter of 10 epochs.

There are a large number of loss functions available to evaluate how well the output of a network matches the expected output. We choose the L1 norm, which yields better results. This corroborates with the results presented in [29], where it is shown that the L2 norm is not the best for images as supposed. In their work, the L2 norm is shown to be worse than the L1 norm alone, followed by a proposed mixture of L1 and L2 norms which shows improvement over the L2 and L1 norms used separately. Here, however, the L1 norm was chosen for the sake of simplicity.

The optimizer chosen for transversing the loss hyperplane was NADAM set with default parameters: the learning rate equals to 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\epsilon = 1e - 07$.

C. FULL IMAGE RECONSTRUCTION

Given the chosen kernel size for the network, an image with dimensions different from this size needs to be sliced and processed individually. To maintain the original image size and compatibility with the kernel size, images with dimensions that are not multiples of it are padded to a dimension multiple of it. This padding is then filled with the blurred image's stretched borders, which are then added noise to simulate a real image. This false-image extension prevents the network from dealing with an abrupt change devoid of real image characteristics, which has been shown to induce artifacts in the real denoised area.

When reconstructing parts of a bigger image separately, the end result shows seams where these parts are re-joined. The applied solution consists of extracting and denoising a third part positioned where the seam would appear in the full image, which is then averaged with the two denoised adjacent parts that create the seam.

D. DATASET CREATION

Hyperspectral captures are not standardized, rarely appearing in the same shapes and depths. In general, the images come from multiple sources and are captured by different sensors [30], [31], further exacerbating their differences. A single representation must be chosen so the neural network's inputs remain the same across all samples. To make these corrections, the image is reformatted, normalized, and resized in chunks to maintain the same network input requirements.

For the construction of the training dataset, a mixture of AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) scenes with some images from the ICVL (Imperial Computer Vision Learning Lab) dataset were used. Each one of these images generated a number of hyperspectral chunks with dimensions compatible with the network's input. By the end of the process, the dataset amounted to 125400 images added of Gaussian noise of different levels. For the test dataset, ten images not previously selected were chosen from both AVIRIS scenes and the ICVL dataset. The test dataset was composed of 160 noisy images with standard deviation in the same range as previously presented.

1) IMAGE FORMAT

Hyperspectral images are distributed in many different formats such as Matlab files [31], [32], tiff files (Images exported directly from HSI cameras such as from Cubert [33]), and many others [34]. In order to operate on images with the same characteristics, the raw files are converted to the same format. For each image loaded, its contents are saved as a NumPy array with a floating point precision of 64 bits. The input images' pixels are also mapped into a range between 0 and 1 and saved as NumPy file.

2) NORMALIZATION

After the initial conversion, another problem still needs to be solved. Most of the time, captured images do not span

the whole numerical range. This effect happens even when comparing bands of the same image, which sometimes have different dynamic ranges. This is a result of the limited representation available due to the sensor's dynamic range, affecting certain bands more than others. Numerically, this means pixel values for pure brightness or pure darkness are other than 0 and 1, respectively, for these bands. To mitigate this problem, a per-band normalization is performed, adjusting their maxima and minima to the expected range accordingly to $P_{\text{adjusted}} = P_{\text{value}}/P_{\text{maximum}}$, where P_{value} is a pixel value before normalization, and P_{maximum} is the maximum pixel value found in the entire image [1]. Each normalized band has its normalization factor saved for posterior restoration if needed.

3) RESIZING

With the hyperspectral images converted and normalized, the last step in the dataset creation pipeline is to generate the samples used as input for the network training. The network inputs require an image size of $128 \times 128 \times 5$ pixels. Chunks are acquired by choosing random positions inside the image, and this process repeats for all available bands. This procedure is useful for creating a more extensive dataset from a limited one.

4) NOISE ADDITION

As discussed previously in subsection III-B, additive Gaussian noise is used to allow for a more generalist neural network that encompasses different noise levels. All the dataset images are added by Gaussian noise with standard deviation equals to $\sigma = 10, 30, 50, 70$. In visual terms, the image corruption varies from almost imperceptible noise to intensely degraded, barely visible scenes.

V. DENOISERS PERFORMANCE ANALYSIS

This section first presents a brief overview of the image metrics for denoising results comparison. Three different methods are chosen for better analyzing the image characteristics, thus aiming to evaluate the similarity of the restored images through various aspects. The images used here were not present in the test dataset, coming from a different set of hyperspectral images to ensure that the trained network learned to generalize well. In the sequel, simulation results are presented in terms of objective numerical comparison, and subjective visual comparison.

A. IMAGE COMPARISON METRICS

At denoising quality evaluation, it is important to choose more than one metric as some of them may have problems with specific scenarios, such as small brightness differences or small visual anomalies [35]. All the metrics presented here are full-reference, which means each noisy image to be evaluated has to be compared against a clean, noiseless version of it.

1) PSNR

Peak Signal to Noise Ratio (PSNR) is a full-reference metric based on the L2 norm, averaging the squared differences between the pixels of a reference image and a target image. The maximum signal and noise power ratio is evaluated for all pairs of pixels of both images.

The PSNR is widely used to analyze image degradation in video and image compression processes. Although it is mathematically simple to calculate and interpret, it does not evaluate perceived visual quality. With the increase of the likeness between two images, the PSNR value increases towards infinity. As pointed out in [35], specific scenarios, such as certain brightness shifts, produce no difference in PSNR while visually degrading the image. Thus, other metrics are needed as well.

2) SSIM

To better assess the similarity between two images, the Structural Similarity Index Measure (SSIM) is based on the hypothesis that the human visual system is highly adapted to extract structural similarity, which is the interdependence of parts of an image that are closer to each other. After normalizing the pixels for contrast and luminance, the method compares local information into pixel intensities. It evaluates over a range from zero to one [36].

3) SAM

As a metric aimed at hyperspectral images, the Spectral Angle Mapper (SAM) compares spectral curves between a reference image and the target one. For an image with n_b bands, the spectral response of a single pixel is a vector on an n-dimensional space. The metric compares the angle of the mean spectral response vectors for all pixels between an image and a reference. Equation (1) shows the α calculation, the angle between the spectral curves of a pair of objects, where t and r are the reference and analyzed spectral responses respectively [37].

$$\alpha = cos^{-1}\left[\frac{\sum_{i=1}^{n_b} t_i r_i}{(\sum_{i=1}^{n_b} t_i^2)^{\frac{1}{2}}(\sum_{i=1}^{n_b} r_i^2)^{\frac{1}{2}}}\right] \tag{1}$$

As it measures the angle between two spectral curves, when the likeness between two images increase, the angle approaches zero.

B. RESULTS

The proposed method was compared with the following classical methods: BM4D, HyRes, fastHyDe, HyMiNoR, rPCA, with default settings for Gaussian denoising. The following were chosen as the neural reconstruction methods: GRN, HSIDCNN, QRNN3D, and MemNet (with a HyRes step). The default models were used, aimed at Gaussian denoising for the selected networks. Many of the used denoising methods presented here were used through their HyDe implementation [38], whereas others such as BM4D were installed via the official pip package [39], for GRN the official

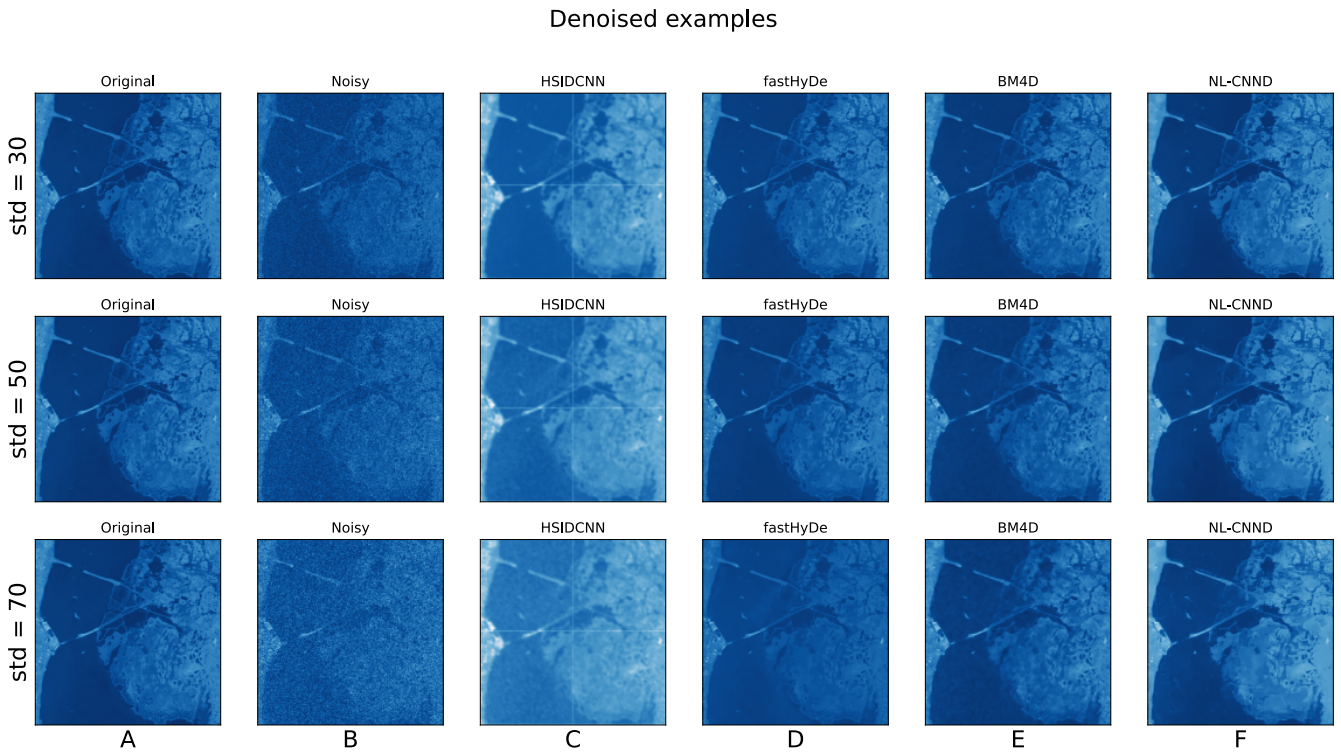

FIGURE 4. Denoised results for the four most performant methods tested, done for three different noise levels. A single band was chosen, and the grayscale images were colored artificially. The right column presents the results for the proposed method, NL-CNND.

implementation available on GitHub was used [40]. To assess the performance of these methods, three images were selected from the ICVL dataset [31], and one from the NASA AVIRIS archives, from the Kennedy Space Center [32]. After being processed, the total number of images amounted to 160 IS samples. The HSIs used for the comparison were not present in the pool used for network training.

Concerning an objective comparison between methods, TABLE 1 (already shown) precises results where the proposed method surpasses the other methods in all of the three chosen metrics. However, it is interesting to point out how expressive are the FastHyDe results, since it is a classical method that does not rely on the novelty features present in neural networks.

For the visual comparison, FIGURE 4 shows an example of a clean, and image with added noise using $\sigma = 30, 50, 70$, at columns A and B respectively. These columns are followed by the four most performing denoising methods. Regarding the reconstruction results, the proposed NL-CNND method (column F) shows the best restoration. Even though some small artifacts are present, it shows the best contrast in complex areas such as the terrain. The second-best method, BM4D (column E), shows a good restoration that, on the other hand, has a lot of blurry artifacts that can be seen in areas of the image where the colors should have been more consistent, such as the water body. FastHyDe (column D) also shows a good restoration, but as the noise level increases, the contrast of the images continuously decreases. HSIDCNN (column C) has a more blurred reconstruction, with lower contrast, and also suffers from a reconstruction problem where the denoised sub-images are reassembled back, a common problem with neural restoration fixed to a specific kernel size.

VI. CONCLUSION

This paper has addressed the issue of noisy hyperspectral images corrupted by Gaussian noise. These images suffer significantly from this type of corruption due to the number of bands acquired simultaneously, and quantization noise. Here, we have presented a brief overview of different classical and neural-based noise removal techniques. Our work is based on a mixture of those two methods, BM4D and Neural Networks, showing that the proposed recovery method yields better results when compared to the other techniques, supported by three different metrics PSNR, SSIM, and SAM. Subjective metrics also show that the method provides high-performance restoration results. In NL-CNND, the processing time of BM4D is an order of magnitude higher than the deep neural network inference time. As future works, optimizing the process for BM4D is appealing for real-time applications.

REFERENCES

[1] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. London, U.K.: Pearson, 2018.
[2] J. R. Schott, *What is Remote Sensing (As far as We're Concerned)*. London, U.K.: Oxford Univ. Press, 2007.
[3] R. A. Schowengerdt, *The Nature of Remote Sensing*. Amsterdam, The Netherlands: Elsevier, 2011, ch. 1.
[4] R. Hirsch, *Exploring Colour Photography: A Complete Guide*. London, U.K.: Laurence King, 2005.

[5] N. Hagen and M. W. Kudenov, "Review of snapshot spectral imaging technologies," *Opt. Eng.*, vol. 52, no. 9, Sep. 2013, Art. no. 090901, doi: 10.1117/1.OE.52.9.090901.

[6] H. Grahn and P. Geladi, *Techniques and Applications of Hyperspectral Image Analysis*. Hoboken, NJ, USA: Wiley, 2007.

[7] T. Adão, J. Hruška, L. Pádua, 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 Sens.*, vol. 9, no. 11, p. 1110, Oct. 2017. [Online]. Available: https://www.mdpi.com/2072-4292/9/11/1110

[8] M. J. Khan, H. S. Khan, A. Yousaf, K. Khurshid, and A. Abbas, "Modern trends in hyperspectral image analysis: A review," *IEEE Access*, vol. 6, pp. 14118–14129, 2018.

[9] P. Ghamisi, "Advances in hyperspectral image and signal processing: A comprehensive overview of the state of the art," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 37–78, Dec. 2017.

[10] B. Rasti, P. Scheunders, P. Ghamisi, G. Licciardi, and J. Chanussot, "Noise reduction in hyperspectral imagery: Overview and application," *Remote Sens.*, vol. 10, no. 3, p. 482, Mar. 2018. [Online]. Available: https://www.mdpi.com/2072-4292/10/3/482

[11] M. Maggioni, V. Katkovnik, K. Egiazarian, and A. Foi, "Nonlocal transform-domain filter for volumetric data denoising and reconstruction," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 119–133, Apr. 2013.

[12] E. J. Candes, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" 2009, *arXiv:0912.3599*.

[13] Y. Guo, A. Davy, G. Facciolo, J.-M. Morel, and Q. Jin, "Fast, nonlocal and neural: A lightweight high quality solution to image denoising," *IEEE Signal Process. Lett.*, vol. 28, pp. 1515–1519, 2021.

[14] Q. Yuan, Q. Zhang, J. Li, H. Shen, and L. Zhang, "Hyperspectral image denoising employing a spatial–spectral deep residual convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 1205–1218, Feb. 2018.

[15] K. Zhang, W. Zuo, and L. Zhang, "FFDNet: Toward a fast and flexible solution for CNN-based image denoising," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4608–4622, Sep. 2018.

[16] B. Rasti, M. O. Ulfarsson, and P. Ghamisi, "Automatic hyperspectral image restoration using sparse and low-rank modeling," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 12, pp. 2335–2339, Dec. 2017.

[17] L. Zhuang and J. M. Bioucas-Dias, "Fast hyperspectral image denoising and inpainting based on low-rank and sparse representations," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 730–742, Mar. 2018.

[18] B. Rasti, P. Ghamisi, and J. A. Benediktsson, "Hyperspectral mixed Gaussian and sparse noise reduction," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 3, pp. 474–478, Mar. 2020.

[19] D. Hong, L. Gao, N. Yokoya, J. Yao, J. Chanussot, Q. Du, and B. Zhang, "More diverse means better: Multimodal deep learning meets remote-sensing imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 4340–4354, May 2021, doi: 10.1109/TGRS.2020.3016820.

[20] K. Wei, Y. Fu, and H. Huang, "3-D quasi-recurrent neural network for hyperspectral image denoising," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 1, pp. 363–375, Jan. 2021.

[21] Y. Tai, J. Yang, X. Liu, and C. Xu, "MemNet: A persistent memory network for image restoration," 2017, *arXiv:1708.02209*.

[22] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[23] B. Keelan, *Handbook of Image Quality: Characterization and Prediction*. Boca Raton, FL, USA: CRC Press, 2002.

[24] C. Boncelet, "Image noise models," in *The Essential Guide to Image Processing*, A. Bovik, Ed. New York, NY, USA: Academic, 2009, pp. 143–167. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B978012374457900007X

[25] L. Gómez-Chova, L. Alonso, L. Guanter, G. Camps-Valls, J. Calpe, and J. Moreno, "Correction of systematic spatial noise in push-broom hyperspectral sensors: Application to CHRIS/PROBA images," *Appl. Opt.*, vol. 47, no. 28, p. F46, Oct. 2008. [Online]. Available: https://opg.optica.org/ao/abstract.cfm?URI=ao-47-28-F46

[26] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian Denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.

[27] M. Tassano, J. Delon, and T. Veit, "FastDVDNet: Towards real-time deep video denoising without flow estimation," 2019, *arXiv:1907.01361*.

[28] M. Banko and E. Brill, "Scaling to very very large corpora for natural language disambiguation," in *Proc. 39th Annu. Meeting Assoc. Comput. Linguistics*, Toulouse, France, Jul. 2001, pp. 26–33. [Online]. Available: ttps://aclanthology.org/P01-1005

[29] H. Zhao, O. Gallo, I. Frosio, and J. Kautz, "Loss functions for image restoration with neural networks," *IEEE Trans. Comput. Imag.*, vol. 3, no. 1, pp. 47–57, Mar. 2017.

[30] *Aviris Airborne Visible/Infrared Imaging Spectrometer Instrument*. Accessed: May 9, 2023. [Online]. Available: https://aviris.jpl.nasa.gov/aviris/instrument.html

[31] B. Arad and O. Ben-Shahar, "Sparse recovery of hyperspectral signal from natural RGB images," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2016, pp. 19–34.

[32] *Aviris Airborne Visible/Infrared Imaging Spectrometer*. Accessed: May 9, 2023. [Online]. Available: https://aviris.jpl.nasa.gov/

[33] *Cubert Video Spectroscopy Ultris x20*. Accessed: May 9, 2023. [Online]. Available: https://www.cubert-hyperspectral.com/products/ultris-x20

[34] O. Ben-Ahmed, T. Urruty, N. Richard, and C. Fernandez-Maloigne, "Toward content-based hyperspectral remote sensing image retrieval (CB-HRSIR): A preliminary study based on spectral sensitivity functions," *Remote Sens.*, vol. 11, no. 5, p. 600, Mar. 2019. [Online]. Available: https://www.mdpi.com/2072-4292/11/5/600

[35] M. Hasan and M. R. El-Sakka, "Improved BM3D image denoising using SSIM-optimized Wiener filter," *EURASIP J. Image Video Process.*, vol. 2018, no. 1, pp. 1–12, Dec. 2018.

[36] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[37] F. Kruse, A. Lefkoff, J. Boardman, K. Heidebrecht, A. Shapiro, P. Barloon, and A. Goetz, "The spectral image processing system (SIPS)—Interactive visualization and analysis of imaging spectrometer data," *Remote Sens. Environ.*, vol. 44, no. 2, pp. 145–163, 1993. [Online]. Available: https://www.sciencedirect.com/science/article/pii/003442579390013N

[38] D. Coquelin, B. Rasti, M. Gotz, P. Ghamisi, R. Gloaguen, and A. Streit, "Hyde: The first open-source, Python-based, GPU-accelerated hyperspectral denoising package," in *Proc. 12th Workshop Hyperspectral Imag. Signal Process., Evol. Remote Sens. (WHISPERS)*, Sep. 2022, pp. 1–5.

[39] Y. Mäkinen. *BM4D on Pypi*. Accessed: May 9, 2023. [Online]. Available: https://pypi.org/project/bm4d/

[40] X. Cao, C. Xu, and D. Meng. *GRN/GRN-HSI-Denoising at Main B7 Xiangyongcao/GRN*. Accessed: May 9, 2023. [Online]. Available: https://github.com/xiangyongcao/GRN/tree/main/GRN-HSI-Denoising

GABRIEL A. DE OLIVEIRA was born in Juiz de Fora, Minas Gerais, in 1993. He received the degree in electrical engineering with an emphasis on electronics from Universidade Federal de Juiz de Fora. He is currently pursuing the master's degree with the Electrical and Computer Engineering Program, Universidade Estadual de Campinas, with a focus on image processing using neural networks and hardware. His research interests include image processing, hardware design, embedded systems, and neural networks.

LARISSA MEDEIROS DE ALMEIDA received the M.Sc. degree in electrical engineering, and mechatronics engineering degree from UFAM, Brazil. She was a researcher with RD institutes in the areas of artificial intelligence, digital image processing, and data science, the Coordinator of Undergraduate Research with universities, and a professor, teaching subjects, such as process automation, theory of mechanical structures, control engineering, and servomechanisms. She is currently the Research and Development Project Manager of TAESA, Brazil. She is also a technologist in analysis and systems development. She received the Graduate Certificate in Big Data and Analytics.

EDUARDO R. DE LIMA received the degree in EE from the University of São Paulo State—UNESP and the Ph.D. degree from the Technical University of Valéncia—UPV, Spain. He is currently the Research and Development Manager of the Exploratory Hardware Design Department, Eldorado Research Institute, and a Visiting Professor with UNICAMP, Campinas, Brazil. He has published more than 20 years of experience in telecommunications systems. He also coordinates several research and development projects related do microelectronics, embedded systems, smart grids, and the IoT. His current interests include the implementation and theoretic aspects of physical layers of wireless and wired communications systems. He is also a MCTI/CNPq Fellow of Technological Productivity.

LUÍS GERALDO P. MELONI (Member, IEEE) received the B.E. degree in electrical engineering/electronics and the M.Sc. degree in electrical engineering/automation from the University of Campinas (UNICAMP), Brazil, in 1980 and 1982, respectively, and the Ph.D. degree in automation/signal processing from the University of Nancy I, France, in 1985. He has vast academic and industrial experience in telecommunications. Previously, he worked in several telecommunication companies. He was a member of the Council of the Brazilian System for Digital Television Forum, from 2012 to 2016, for coordinating works of standardization of technologies for interactive channel with the Brazilian Association of Technical Standards (ABNT). He was the Coordinator of the Faculty Outreach, FEEC, for four years. He was also a Lecturer with the University of Brasilia (1990–1993). He is currently a Professor with the School of Electrical and Computer Engineering (FEEC), UNICAMP. He is also the current Director of Outreach University and UNICAMP. He has many publications in international journals and symposia and has also been an instructor in continuing education programs for telecommunication professionals. His current research interests include new technologies for wireless communications, hypercomplex algebra, artificial intelligence, middleware architectures for home energy management systems, the Internet of Things, and software-defined radio technologies.