Robust Hyperspectral Inpainting via Low-Rank Regularized Untrained Convolutional Neural Network

Keivan Faghih Niresi and Chong-Yung Chi, *Life Fellow, IEEE*

***Abstract*— Over the past decade, many low-rank models, factorizations, or approximations have been applied to the restoration of hyperspectral images (HSIs) (e.g., denoising, inpainting, and super-resolution) from their incomplete and/or noisy measurements. Recently, deep learning (DL) has been shown to be a powerful method for solving inverse problems (including HSI restoration), but a large amount of training data is required. Since this is not possible for HSIs, unlike red green blue (RGB) images, in this work, a novel unsupervised framework for hyperspectral inpainting (HI) is proposed that can be implemented using an untrained convolutional neural network (CNN) for deep image prior (DIP), together with a recently reported differentiable regularization for the data rank and ℓ_2-norm squared loss function. Based on the proposed framework, we come up with a novel HI algorithm [denoted as deep low-rank hyperspectral inpainting (DLRHyIn)] and a robust DLRHyIn (denoted as R-DLRHyIn) which is robust against outliers, where the latter differs from the former only in the Huber loss function (HLF) (which has been justified robust to mixed noise) used instead. Then some simulation results and real-data experiments are provided to demonstrate the effectiveness of the proposed DLRHyIn and R-DLRHyIn. Finally, we draw some conclusions.**

***Index Terms*— Convolutional neural network (CNN), deep image prior (DIP), Huber loss function (HLF), hyperspectral inpainting (HI), inverse problems, low-rank regularization.**

I. Introduction

HYPERSPECTRAL imaging has been widely used in many applications such as remote sensing, environmental monitoring, planetary exploration, and so on, thanks to rich spectral information [1], [2]. Nevertheless, some deadlines (i.e., data missing on such lines) may occur in hyperspectral images (HSIs) due to various acquisition system errors during the scanning and/or transmission of the satellite/airborne sensor in addition to unfavorable environmental conditions such as shadow, cloud, and fog contamination [3], [4]. For this reason, inpainting is one of the most important preprocessing steps in hyperspectral imaging to fill missing pixels from a set of incomplete observations. Inpainting can be formulated as an inverse problem that consists of a data-fitting (data-fidelity) term and some regularization terms (due to the fact that the problem is generally ill-posed) based on our prior knowledge about the unknown clean image.

Due to the significant spectral–spatial correlation, low-rankness is one of the priors that works well for solving the inverse problem in HSIs [5], [6], [7]. The majority of previous efforts concentrated on finding a convex surrogate for the rank of the data matrix (nonconvex and nondifferentiable) followed by solving the formulated minimization problem. For instance, one of the most widely utilized solutions for HSI inverse problems has been the nuclear norm used as a convex surrogate for the nonconvex matrix rank [8].

In addition to low-rank regularizations, the plug-and-play (PnP) prior is an effective framework that directly applies an existing regularizer from a cutting-edge image denoiser to address an inverse image problem leading to potent image restoration results [9]. Deep learning (DL) has been regarded as a powerful framework for solving inverse problems with state-of-the-art results in many applications [10]. However, DL requires a large amount of training data which is a challenging issue in HSIs. Ulyanov et al. [11] recently proposed deep image prior (DIP), an unsupervised DL paradigm for solving imaging inverse problems, free from the pretraining on a large amount of data. The DIP uses a convolutional neural network (CNN) as an implicit regularization to recover images from the given noisy image data. Sidorov and Yngve Hardeberg [12] applied DIP to the restoration of HSIs. Similar to supervised approaches, DIP's optimization strategy for HSI restoration is established. Despite providing a novel paradigm for handling inverse problems without involving hand-crafted priors (regularization), DIP techniques do not perform as well as state-of-the-art frameworks.

Liu et al. [13], DIP is further improved by employing total variation (TV) regularization as an explicit regularization in the optimization process. Motivated by [13], the low-rank property is utilized as an explicit regularization incorporated in the objective function to restrict the solutions provided by CNNs to those that are of low rank in this work. However, unlike TV, using a rank function or even its convex envelope (nuclear norm) in a CNN brings forth analytical challenges because neither is a differentiable function of the

Manuscript received 29 August 2022; revised 6 December 2022 and 21 January 2023; accepted 28 January 2023. Date of publication 31 January 2023; date of current version 10 February 2023. This work was supported by the Ministry of Science and Technology, Taiwan, under Grant MOST 110-2221-E-007-048 and Grant MOST 111-2221-E-007-047-MY2. *(Corresponding author: Keivan Faghih Niresi.)*

Keivan Faghih Niresi is with the Institute of Communications Engineering, National Tsing Hua University, Hsinchu 30013, Taiwan (e-mail: keivan@m109.nthu.edu.tw; keyvan.faghih@gmail.com).

Chong-Yung Chi is with the Institute of Communications, Department of Electrical Engineering, National Tsing Hua University, Hsinchu 30013, Taiwan (e-mail: cychi@ee.nthu.edu.tw).

Digital Object Identifier 10.1109/LGRS.2023.3241161

image matrix/tensor. Instead, we consider a differentiable data-matrix rank approximation proposed by Ye et al. [14] in the deep CNN optimization process, together with a robust statistics-based data-fitting function, so-called Huber loss function (HLF) [15], to establish a framework for hyperspectral inpainting (HI), that turns out to be effective in the presence of both missing lines (i.e., deadlines) and outliers in the noisy HSI.

The main contributions of this letter are summarized as follows.

1) Free from pretraining, an unsupervised HSI inpainting algorithm [denoted deep low-rank hyperspectral inpainting (DLRHyIn)] is proposed using ℓ_2-norm square as the data fitting function, DIP, and a recently reported smooth low-rank regularization; a robust DLRHyIn (denoted R-DLRHyIn) is proposed except for the HLF used for the data-fitting function, which is robust against outliers.
2) Both simulation results and real-data experiments are provided to demonstrate the effectiveness of the proposed DLRHyIn and R-DLRHyIn, with superior performance compared to some state-of-of-the-art algorithms.

The rest of the letter is organized as follows. Section II reviews related background for ease of the ensuing presentation. In Section III, the two proposed algorithms are presented. In Section IV, simulated and real-data experiments are conducted to evaluate the performance of the proposed algorithm together with a comparison with some state-of-the-art methods. Finally, some conclusions are drawn in Section V.

II. Related Backgrounds

Throughout the letter, third-order tensors are represented as bold calligraphic letters, for example, $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{H\times W\times B}$ (real tensor with dimension $H \times W \times B$) with its (i, j, k)th entry denoted as a_{ijk}. Mode-3 unfolding operator for 3-D tensor is defined as $\text{unfold}_3(\cdot) : \mathbb{R}^{H\times W\times B} \to \mathbb{R}^{B\times(WH)}$, which converts a tensor $\boldsymbol{\mathcal{A}}$ into a matrix $\mathbf{A} = \{a_{ij}\}_{B\times(WH)}$. $\mathbf{A}^T$ denotes the transpose of $\mathbf{A}$. $\mathbf{diag}(a_1, \ldots, a_n)$ denotes a diagonal matrix with n diagonal entries $a_1, \ldots, a_n$. The Frobenius norm and ℓ_1-norm are defined as $\|\mathbf{A}\|_F = (\sum_{ij} a_{ij}^2)^{1/2}$ and $\|\mathbf{A}\|_1 = \sum_{ij} |a_{ij}|$, respectively.

A. Inverse Problems

Consider the following degradation model:

$$\mathbf{Y} = \mathcal{T}(\mathbf{X}) + \mathbf{S} \tag{1}$$

where $\mathbf{Y} \in \mathbb{R}^{n_1\times n_2}$, $\mathbf{X} \in \mathbb{R}^{m_1\times m_2}$ (while $n_1 n_2 \le m_1 m_2$), and $\mathbf{S} \in \mathbb{R}^{n_1\times n_2}$ are mode-3 unfolded of corrupted HSI ($\boldsymbol{\mathcal{Y}}$), clean HSI ($\boldsymbol{\mathcal{X}}$), and sparse noise ($\boldsymbol{\mathcal{S}}$), respectively. Moreover, $\mathcal{T}(\cdot)$ is any known linear degradation operator such that it is an identity transformation for the denoising task. In this work, $\mathcal{T}(\mathbf{X}) = \mathbf{M} \odot \mathbf{X}$ is considered, where $\mathbf{M}$ is an all-one matrix formed by all the nonmissing entries of $\mathbf{X}$ and $\odot$ is Hadamard (element-wise) product. While $\mathbf{S} = \mathbf{0}$ is considered in the conventional inpainting task, for robust hyperspectral tensor completion problem which will be studied in Section III, $\mathbf{S}$ is a sparse noise matrix yielded by outliers. Even $\mathbf{S} = \mathbf{0}$, since the inverse problems in the real scenario are ill-posed, it is needed to use the regularization term based on our prior knowledge about clean HSI besides the data-fitting term to restrict the solution. Hence, the inpainting task can be rewritten as the regularized least-squares problem

$$\mathbf{X}^* = \underset{\mathbf{X}}{\text{argmin}} \left\{ \|\mathbf{Y} - \mathcal{T}(\mathbf{X})\|_F^2 + \lambda W(\mathbf{X}) \right\} \tag{2}$$

where inside the braces, the first term (the second term, the regularization weighted by parameter $\lambda > 0$) is for the consistency between the restored image and the observed image (our prior information).

B. Deep Image Prior

Recently, Ulyanov et al. [11] proposed the DIP framework which is capable of solving various inverse problems in imaging without the need for pretraining. In other words, CNN which is the core element of DIP can implicitly capture some inherent image attributes. So, the regularization can be waived from (2). Hence, the optimization problem for DIP can be formulated as

$$\Theta^* = \underset{\Theta}{\text{argmin}} \|\mathbf{Y} - \mathcal{T}(f_\Theta(\mathbf{Z}))\|_F^2 \text{ and } \mathbf{X}^* = f_{\Theta^*}(\mathbf{Z}) \tag{3}$$

where $\mathbf{Z}$ (a random initialization with the same dimension of $\mathbf{X}$) is the input of CNN $f_\Theta(\cdot)$ parameterized by the column vector Θ. Since all terms in (3) are differentiable, a gradient-based algorithm can efficiently solve the optimization problem (3) to update the network parameter vector Θ at each iteration so that the output of the network is approaching the target.

C. Low-Rank Prior

Recently, it has been claimed that HSIs are inherently low-rank [5]. A nonsmooth rank function can be expressed as $\text{Rank}(\mathbf{T}) = \sum_i u(\sigma_i(\mathbf{T}))$, where $\sigma_i(\mathbf{T})$ is the ith singular value of matrix $\mathbf{T} \in \mathbb{R}^{B\times(WH)}$ and $u(t)$ is the nonsmooth unit step function. The nuclear norm of $\mathbf{T}$, known as the convex envelope of $\text{Rank}(\mathbf{T})$, is a nondifferentiable function, thus preventing its application to CNNs. Ye et al. [14] proposed a differentiable approximation to $\text{Rank}(\mathbf{T})$ as follows:

$$R_\eta(\mathbf{T}) = \sum_{i=1}^{N} 1 - \exp\left(-\frac{\sigma_i(\mathbf{T})}{\eta}\right) \tag{4}$$

where $N = \min(B, \text{WH})$ and $\eta > 0$ is a parameter that determines the smoothness and approximation error. Moreover, $\lim_{\eta\to 0} R_\eta(\mathbf{T}) = \text{Rank}(\mathbf{T})$. Its gradient is given by

$$\frac{\partial R_\eta(\mathbf{T})}{\partial \mathbf{T}} = \mathbf{U}\mathbf{diag}\left(\frac{1}{\eta}\exp\left(-\frac{\sigma_1(\mathbf{T})}{\eta}\right), \ldots, \frac{1}{\eta}\exp\left(-\frac{\sigma_N(\mathbf{T})}{\eta}\right)\right)\mathbf{V}^T$$

where all the quantities are taken from the singular value decomposition (SVD) $\mathbf{T} = \mathbf{U}\mathbf{diag}(\sigma_1(\mathbf{T}), \ldots, \sigma_N(\mathbf{T}))\mathbf{V}^T$.

III. Proposed Methods

This section presents two novel methods for HI. One is designed by integrating low-rank prior with DIP to capture the spatial–spectral correlation of HSIs, referred to as DLRHyIn. The other is made by using robust statistics so that it is effective in the presence of outliers.

A. Deep Low-Rank Hyperspectral Inpainting

The proposed DLRHyIn algorithm aims to solve the following problem:

$$\Theta^* = \text{argmin}_\Theta\{\|\mathbf{Y} - \mathcal{T}(f_\Theta(\mathbf{Z}))\|_F^2 + \lambda R_\eta(f_\Theta(\mathbf{Z}))\} \quad \text{and} \quad \mathbf{X}^* = f_{\Theta^*}(\mathbf{Z}). \tag{5}$$

Since (5) is a differentiable optimization problem, any gradient-based algorithm can be used to update the network parameter vector Θ as in handling (3).

B. Robust Deep Low-Rank Hyperspectral Inpainting

We would like to mention that (3) and (5) are derived under model (2), that is, in the absence of sparse noise. However, sparse noise (e.g., outliers) is present in practical scenarios. For this reason, we further propose an R-DLRHyIn algorithm, at first by solving

$$\min_{\mathbf{X},\mathbf{S}} \|\mathbf{Y} - \mathcal{T}(\mathbf{X}) - \mathbf{S}\|_F^2 + \delta\|\mathbf{S}\|_1 + \lambda R_\eta(\mathbf{X}). \tag{6}$$

It is noteworthy that (6) is the commonly known model-based method with a convex envelope $\|\mathbf{S}\|_1$ of $\|\mathbf{S}\|_0$ used as the regularization term for handling the sparse noise ($\mathbf{S}$). To handle (6) more efficiently, based on Proposition 1 of [16], it can be shown that (6) is equivalent to the below problem if the CNN is not applied [i.e., $f_{\Theta^*}(\mathbf{Z})$ replaced by $\mathbf{X}$ in (7)]. Therefore, we come up with the proposed R-DLRHyIn as follows:

$$\Theta^* = \text{argmin}_\Theta \left\{H_\delta(\mathbf{Y} - \mathcal{T}(f_\Theta(\mathbf{Z}))) + \lambda R_\eta(f_\Theta(\mathbf{Z}))\right\} \quad \text{and} \quad \mathbf{X}^* = f_{\Theta^*}(\mathbf{Z}). \tag{7}$$

where $H_\delta(\mathbf{E} = \{e_{ij}\}_{m\times n}) = \sum_{i=1}^m \sum_{j=1}^n p_\delta(e_{ij})$, in which $\mathbf{E} \triangleq \mathbf{Y} - f_\Theta(\mathbf{Z})$, and $p_\delta(\cdot)$ is HLF [15] with parameter $\delta > 0$ defined as

$$p_\delta(\omega) = \begin{cases} \frac{1}{2}\omega^2, & |\omega| \le \delta \\ \delta|\omega| - \frac{1}{2}\delta^2, & |\omega| > \delta. \end{cases}$$

It is noticeable that R-DLRHyIn with large δ [cf. (7)] reduces to DLRHyIn [cf. (5)]. The robustness of the former to sparse noise is thanks to the fact that when $\mathcal{T}(\mathbf{X}) = \mathbf{X}$ (i.e., no missing pixels) and $\lambda = 0$, (7) is exactly the recently reported robust denoiser HLF-DIP [16].

In this work, the DIP architecture is based on a convolutional encoder–decoder with some skip connections to tackle the vanishing gradient. A random noise realization [e.g., with a uniform distribution $\mathbf{U}(0, 0.1)$] is generated as the CNN input, which is processed by the encoder to yield embedded features. The embedded features are then used by the decoder to reconstruct the HSI. For a fair comparison, we would like to emphasize that the DIP architecture employed is identical to [12] and [16] for highlighting the low-rank regularization and HLF impact on the performance of the proposed algorithms. Finally, the optimization problem is solved by an adaptive moment estimation (Adam) [17] optimizer with a learning rate equal to 0.001.

TABLE I
EXPERIMENTAL RESULTS IN THE PRESENCE OF DEADLINES BUT WITHOUT ADDITIVE NOISE, WHERE BOLDFACE (UNDERLINED) NUMBERS DENOTE THE BEST (SECOND BEST) PERFORMANCE

Method	MPSNR	MSSIM	MFSIM	MSAM	Running Time (sec.)
Corrupted	18.53	0.708	0.788	0.451	-
FastHyIn	33.95	0.972	0.983	0.057	**10**
TNN-DCT	35.15	0.985	0.987	<u>0.040</u>	<u>351</u>
DHP	39.99	0.990	<u>0.992</u>	0.045	543
DIP-TV	<u>40.17</u>	<u>0.991</u>	<u>0.992</u>	0.043	546
DLRHyIn†	**41.22**	**0.993**	**0.995**	**0.036**	545

† DLRHyIn and R-DLRHyIn are equivalent for this experiment.

IV. EXPERIMENTAL RESULTS

A. Experimental Settings

1) Simulated Dataset: The proposed methods are evaluated based on the Washington DC Mall (WDC Mall) dataset,[1] which was captured over WDC Mall by the Hyperspectral Digital Imagery Collection Experiment (HYDICE) sensor which has a 1-m spatial resolution and a 10-nm band spacing covering the spectral range of 400–2500 nm. Spectral bands between 0.9 and 1.4 μm are omitted from the dataset since the atmosphere was opaque. The experiments are conducted on a patch of size $200 \times 200 \times 191$ (i.e., $H = W = 200$ and $B = 191$) from the WDC Mall image. Following the same procedure in [16], the generated data $\mathbf{Y}$, and the input ($\mathbf{Z}$) and output $f_\Theta(\mathbf{Z})$ of the CNN are all within the interval [0, 1].

2) Performance Evaluation: To evaluate the effectiveness of the proposed methods in two different scenarios (to be presented in Sections IV-B and IV-C), four competing methods have been chosen. FastHyIn [18] is a subspace-based approach in which low-rank and self-similarity characteristics of HSIs are considered besides the well-known BM3D [19] denoiser plugged in to fill the missing pixels. Tensor nuclear norm (TNN)-discrete cosine transform (DCT) [20] is a tensor robust principal component analysis (PCA) approach based on TNN under linear transformation such as DCT. Deep hyperspectral prior (DHP) [12] is an unsupervised approach that applies DIP to HI with neither pretraining nor any explicit regularization, hence corresponding to the proposed DLRHyIn with $\lambda = 0$. DIP-TV [13] exploits DHP and TV as an explicit regularization in boosting the DHP performance. Parameters of competing algorithms are tuned based on suggested or default values according to the research papers and/or available source codes provided by the authors, while for the proposed R-DLRHyIn, we set $\eta = 0.04$ [14], $\lambda = 0.1$, and $\delta = 0.001$ ($\delta = 1$) for the case that sparse noise is present (not present) [16]. Let us emphasize that R-DLRHyIn is actually equivalent to the proposed DLRHyIn defined by (5) for large δ (e.g., $\delta = 1$ due to $|e_{ij}| \le 1$). The mean peak signal-to-noise ratio (MPSNR), mean structural similarity index measure (MSSIM) [21], mean feature similarity index measure (MFSIM) [22], and mean spectral angle mapper (MSAM) [23] are four quantitative performance metrics are used as performance indexes, for which the larger the values of the first three indexes, the

[1] http://lesun.weebly.com/hyperspectral-data-set.html

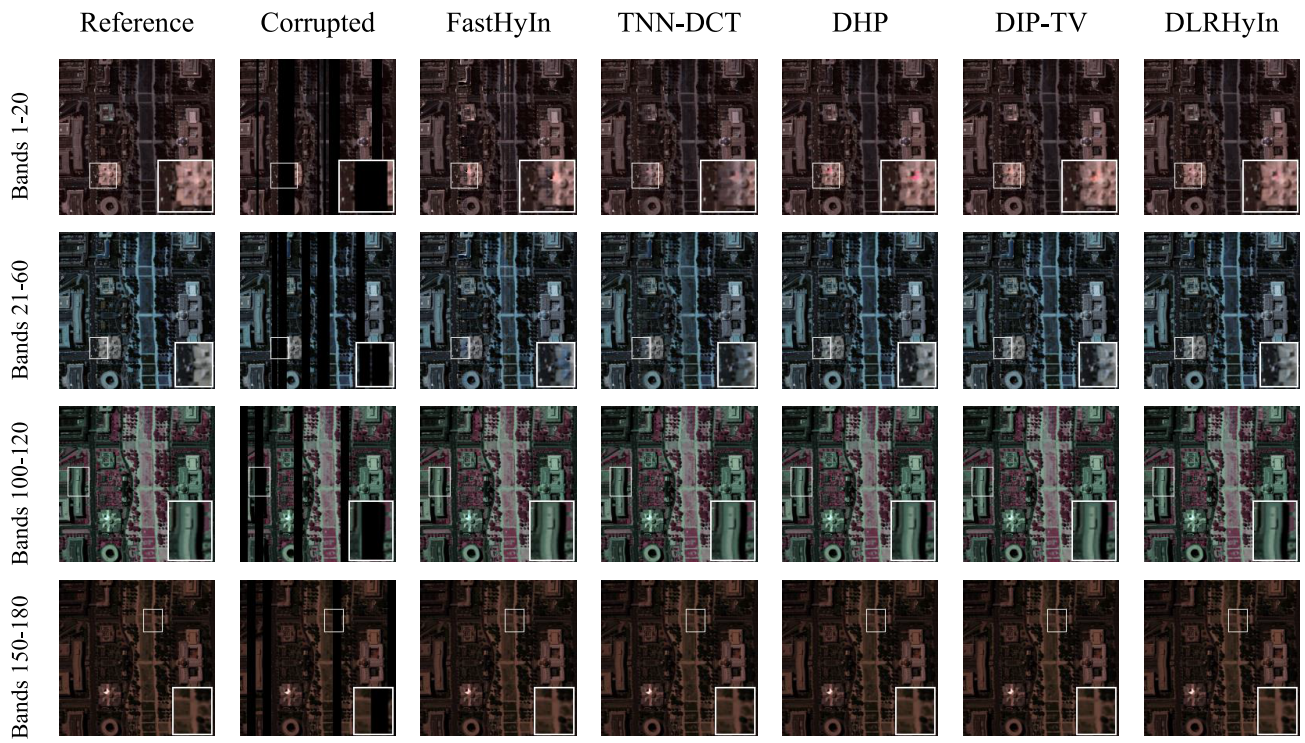

Fig. 1. Visual quality assessment of various HI methods using the WDC Mall dataset for the case in the presence of deadlines but without additive noise, where the recovered HSIs obtained under four different patterns of deadlines (shown in the second column) over nonoverlapping band sets 1–20 (first row), 21–60 (second row), 100–120 (third row), and 150–180 (fourth row), respectively, are displayed by the pseudocolor (3-channel) images for each band set.

TABLE II

EXPERIMENTAL RESULTS IN THE PRESENCE OF BOTH DEADLINES AND SPARSE NOISE, WHERE BOLDFACE (UNDERLINED) NUMBERS DENOTE THE BEST (SECOND BEST) PERFORMANCE

Method	MPSNR	MSSIM	MFSIM	MSAM	Running Time (sec.)
Corrupted	14.92	0.361	0.691	0.820	-
FastHyIn	27.80	0.884	0.947	0.137	**11**
TNN-DCT	27.35	0.905	0.947	0.168	<u>363</u>
DHP	30.02	0.915	0.952	0.137	521
DIP-TV	30.21	0.917	0.956	0.135	530
DLRHyIn	<u>30.37</u>	<u>0.919</u>	<u>0.961</u>	<u>0.133</u>	527
R-DLRHyIn	**41.09**	**0.993**	**0.995**	**0.037**	541

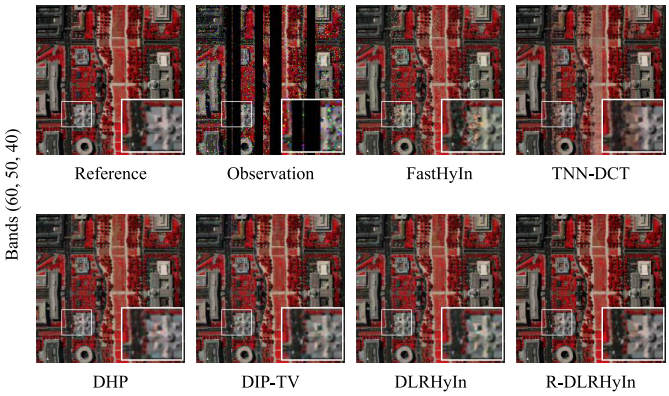

Fig. 2. Visual quality assessment of various HI methods in the presence of outliers using the WDC Mall dataset.

better the performance, while this is the reverse for MSAM. We would like to mention that in our experiments, some bands are clean and not corrupted by any deadlines. For this reason, for evaluating the performance of each method, we drop these bands in assessing the performance based on the aforementioned metrics.

B. Hyperspectral Inpainting

For the HI experiment, 112 spectral bands (1–20, 21–60, 100–120, 150–180) of WDC Mall are corrupted by deadlines in different patterns that can be observed in Fig. 1. It is worth mentioning that 60 consecutive bands are corrupted which is a serious case (deadlines in bands 1–20 are in the same pattern but they are different from bands 21–60). Quantitative results of various methods for HI are illustrated in Table II, showing that the proposed methods outperform the other four existing methods[2] based on almost all performance metrics. Moreover, one can observe that low-rank prior $R_\eta(\cdot)$ plays a positive role in HSI recovery since the performance of DLRHyIn (R-DLRHyIn) is higher than DHP by nearly 1.3 dB margin in MPSNR. A visual comparison of the methods in different bands is depicted in Fig. 1. It can be clearly seen from the first two rows of this figure that FastHyIn is able to fill some missing pixels, but some artifacts have remained in restored HSIs; DLRHyIn performs best in spite of slightly better performance than DIP-TV, DHP, and TNN-DCT. However, the bottom two rows in Fig. 1 show a similar performance to all the algorithms under test, perhaps due to fewer deadlines and smaller amounts of consecutive corrupted bands than in the top two rows.

[2]Hyperion Bhilwara Hyperspectral Data Cube. Accessed: Aug. 2, 2020. [Online]. Available: https://earthexplorer.usgs.gov/

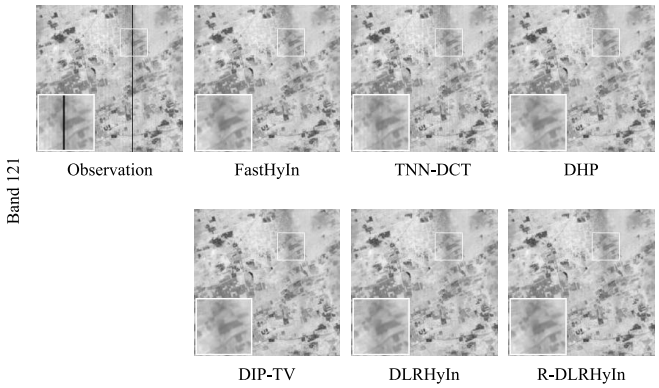

Fig. 3. Visual quality assessment of various methods for band 121 of NASA's Hyperion real hyperspectral satellite images over Bhilwara, India.

C. Robust Hyperspectral Inpainting

For the robust HI experiment, sparse noise (5% salt and pepper noise) is further added to the corrupted image in Section IV-B. Table II shows that R-DLRHyIn significantly outperforms all the other algorithms under test, thereby demonstrating its robustness to outliers. Fig. 2 illustrates the visual quality assessment among all the methods under test. It can be observed that R-DLRHyIn performs best in filling missing pixels and removing sparse noise. Furthermore, compared to other methods, details of the restored HSI in R-DLRHyIn are better maintained, thereby justifying that low-rank regularization, DIP, and HLF are effective for inpainting in the presence of sparse noise.

D. Real Data Experiment

This section assesses the proposed method's performance on a real HSI obtained by the Hyperion hyperspectral sensor over Bhilwara, India. The real Hyperion HSI on band 121 is noisy and contaminated by a deadline [24] is processed by all the algorithms under test, and the recovered HSIs are shown in Fig. 3. From this figure, it can be observed that FastHyIn, DHP, DIP-TV, DLRHyIn, and R-DLRHyIn yield visually quite similar results in this noisy low-rank data set, while TNN-DCT can fill the missing line but some vertical stripes have remained in the restored image.

V. Conclusion

We have presented an HI algorithm (DLRHyIn), using a CNN-based DIP and a differentiable approximation to the rank of HSIs as the explicit regularization, and a robust HI algorithm (R-DLRHyIn) that further uses the HLF (derived from robust statistics) as the data-fitting function. The former (latter) is effective in the absence (presence) of outliers or sparse noise. Experimental results demonstrate their effectiveness and superior overall performance over some state-of-the-art algorithms.

References

[1] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. M. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, Jun. 2013.

[2] S. Nakhostin, H. Clenet, T. Corpetti, and N. Courty, "Joint anomaly detection and spectral unmixing for planetary hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 6879–6894, Dec. 2016.

[3] P. Duan, S. Hu, X. Kang, and S. Li, "Shadow removal of hyperspectral remote sensing images with multiexposure fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–11, 2022.

[4] X. Kang, Z. Fei, P. Duan, and S. Li, "Fog model-based hyperspectral image defogging," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–12, 2022.

[5] H. Zhang, W. He, L. Zhang, H. Shen, and Q. Yuan, "Hyperspectral image restoration using low-rank matrix recovery," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 4729–4743, Aug. 2013.

[6] H. Fan, Y. Chen, Y. Guo, H. Zhang, and G. Kuang, "Hyperspectral image restoration using low-rank tensor recovery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 10, pp. 4589–4604, Oct. 2017.

[7] Q. Wang, X. He, and X. Li, "Locality and structure regularized low rank representation for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 911–923, Feb. 2018.

[8] M. Golbabaee and P. Vandergheynst, "Hyperspectral image compressed sensing via low-rank and joint-sparse matrix recovery," in *Proc. IEEE ICASSP*, Mar. 2012, pp. 2741–2744.

[9] L. Zhuang, M. K. Ng, and X. Fu, "Hyperspectral image mixed noise removal using subspace representation and deep CNN image prior," *Remote Sens.*, vol. 13, no. 20, p. 4098, Oct. 2021.

[10] R. Hang, Q. Liu, and Z. Li, "Spectral super-resolution network guided by intrinsic properties of hyperspectral imagery," *IEEE Trans. Image Process.*, vol. 30, pp. 7256–7265, 2021.

[11] V. Lempitsky, A. Vedaldi, and D. Ulyanov, "Deep image prior," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 9446–9454.

[12] O. Sidorov and J. Y. Hardeberg, "Deep hyperspectral prior: Single-image denoising, inpainting, super-resolution," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshop (ICCVW)*, Oct. 2019, pp. 3844–3851.

[13] J. Liu, Y. Sun, X. Xu, and U. S. Kamilov, "Image restoration using total variation regularized deep image prior," in *Proc. IEEE ICASSP*, May 2019, pp. 7715–7719.

[14] H. Ye, H. Li, B. Yang, F. Cao, and Y. Tang, "A novel rank approximation method for mixture noise removal of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4457–4469, Jul. 2019.

[15] P. J. Huber, "Robust estimation of a location parameter," *Ann. Math. Statist.*, vol. 35, no. 1, pp. 73–101, 1964.

[16] K. F. Niresi and C.-Y. Chi, "Unsupervised hyperspectral denoising based on deep image prior and least favorable distribution," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 5967–5983, 2022.

[17] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2015, pp. 1–15.

[18] 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.

[19] 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.

[20] C. Lu, X. Peng, and Y. Wei, "Low-rank tensor completion with a new tensor nuclear norm induced by invertible linear transforms," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 5996–6004.

[21] 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.

[22] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.

[23] F. A. Kruse et al., "The spectral image processing system (SIPS)-interactive visualization and analysis of imaging spectrometer data," *Remote Sens. Environ.*, vol. 44, pp. 145–163, May 1993.

[24] C.-H. Lin, Y.-C. Lin, and P.-W. Tang, "ADMM-ADAM: A new inverse imaging framework blending the advantages of convex optimization and deep learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–16, 2022.