

# A HIGHLY INTERPRETABLE DEEP EQUILIBRIUM NETWORK FOR HYPERSPSPECTRAL IMAGE DECONVOLUTION

Alexandros Gkillas<sup>a</sup>, Dimitris Ampeliotis<sup>b</sup>, Kostas Berberidis<sup>a</sup>

<sup>a</sup>Dept. of Computer Engineering & Informatics, University of Patras, Greece

<sup>b</sup>Dept. of Digital Media and Communication, Ionian University, Greece

## ABSTRACT

In this paper, a novel technique for the hyperspectral image deconvolution problem is developed. First, considering the highly ill-posed nature of the examined problem, it is imperative to incorporate proper priors (regularizers) to capture the strong spectral and spatial dependencies of the hyperspectral images. Then, in light of this, a novel optimization problem is proposed by employing a convolutional neural network to act as a regularizer, which is learnt to reflect the properties of the signals of interest. To solve the proposed optimization problem, we use the half quadratic splitting methodology, thus designing an efficient iterative solver (iteration map). Based on the Deep Equilibrium (DEQ) modeling, which aims to express the proposed iterative solver as an equilibrium (fixed-point) computation, a highly interpretable deep learning-based network is derived, which can be trained end-to-end. Extensive numerical results using two publicly available datasets illustrate that the proposed method markedly outperforms other state-of-the-art approaches.

**Index Terms**— Hyperspectral images, deep equilibrium models, deconvolution, learnable regularizers, deep unrolling

## 1. INTRODUCTION

Over the years, many scientific and technological advancements have greatly improved hyperspectral imaging (HSI), which now constitutes a helpful tool in numerous domains and applications, such as remote sensing, medical science and autonomous driving [1, 2]. However, during the acquisition stage, hyperspectral images are often degraded in many ways, such as the addition of noise and various blurring effects, which heavily deteriorate the performance in HSI applications [3, 4]. Thus, the restoration of hyperspectral images constitutes a critical pre-processing stage and stresses the need for efficient denoising and deconvolution methods.

In literature, there is a plethora of works aiming to tackle the hyperspectral image deconvolution problem, following various perspectives and making different assumptions. In particular, several baseline methodologies employ filter-based

approaches, such as the 3D Wiener filter [5] and the Kalman filter [4], to tackle the deconvolution problem. Other approaches utilize the Fourier and wavelet domains to compute efficient solutions for the hyperspectral image restoration problem [6]. Considering the highly ill-posed nature of the examined problem, many approaches incorporate proper priors to enhance the restoration performance. More specifically, in [7] the deconvolution solution considers spatial and spectral priors under passivity constraints. In [8], an online algorithm is employed based on a sliding-block regularized Least Mean Squares (LMS) algorithm. Furthermore, in [9], a non negative regularized optimization problem was proposed, using minimum distance and maximum curvature criteria to estimate the regularization parameters.

Recently, several works have explored the use of regularization terms that are properly learnt from suitable training data to capture the properties of the signals of interest [10–12]. These regularization terms are used in place of the “hand-crafted” regularization terms traditionally employed in the cost functions of the studied restoration problems, and have the form of a suitable artificial neural network. This approach is known as the “plug-and-play” method. So far, such research efforts have mainly focused on 2D image processing inverse problems. The study in [13] considers the plug-and-play approach to tackle the HSI deconvolution problem. However, it should be highlighted that the plug-and-play methodology has a limitation, since the neural network (regularizer) is trained independently of the inverse problem at hand and the degraded data, thus requiring massive amounts of training data to obtain satisfactory results.

Different from the above studies, in this work, we extend the potential of the learnable regularizers in the hyperspectral deconvolution problem. In more detail, by combining a neural network that acts as a spatial and spectral prior for hyperspectral images and a data fidelity term, we propose a novel cost function for hyperspectral deblurring. Accordingly, this problem is solved efficiently by employing the half quadratic splitting (HQS) method [14]. In the sequel, we consider that the iterative algorithm thus obtained can be incorporated into a deep unrolling methodology [15, 16], where a fixed number of iterations are unrolled to form a neural network with specific structure. Deep unrolling offers the possibility for

This work was supported in part by the University of Patras and the RPE, Cyprus, under the project INFRASTRUCTURES/1216/0017 (IRIDA).

end-to-end training of the resulting neural network, so as to better adapt to the problem at hand. However, due to computational constraints, the number of unrolled iterations must be kept small [17]. In view of this, in this work we focus on a more efficient architecture based on the Deep Equilibrium (DEQ) approach [18]. The deep equilibrium approach aims to express the entire deep learning architecture derived from the iterative algorithm as an equilibrium (fixed-point) computation, corresponding to a neural network with an equivalent infinite number of layers. A great benefit from this direction is the fact that the proposed DEQ network is highly interpretable, as its parameters have a direct correspondence to the parameters of the involved iterative algorithm, whose operation is well understood and justified [19]. Furthermore, the parameters of the neural network can be trained end-to-end, thus providing markedly better results as compared to the state-of-the-art, plug-and-play method in [13]. The authors of [20] have explored the applicability of this methodology to solve 2D image reconstruction problems. Different from that method, in our study the focus is on hyperspectral imaging aiming to derive a highly interpretable deep equilibrium network for the HSI deconvolution problem. *To the best of the authors' knowledge, this is the first study that employs the deep equilibrium modeling for tackling the hyperspectral deconvolution problem.*

## 2. PROBLEM FORMULATION

Consider a degraded hyperspectral image  $\mathbf{Y} \in \mathbb{R}^{M \times N \times d}$  and its corresponding ground truth (clean) image  $\mathbf{X} \in \mathbb{R}^{M \times N \times d}$ , where  $M, N$  denote the spatial dimensions of the images and  $d$  corresponds to the spectral dimension (i.e., number of spectral bands) of the images. Adopting the linear degradation model of [13, 21] and under the assumption that the blurring kernel remains constant across the spectral dimension (i.e.,  $\mathbf{H}_i = \mathbf{H}_c$  for  $i = 1 \dots d$ ), the  $i$ -th spectral band of the degraded image  $\mathbf{Y}_i \in \mathbb{R}^{M \times N}$  can be modelled as

$$\mathbf{Y}_i = \mathbf{H}_c \star \mathbf{X}_i + \mathbf{W}_i, \quad i = 1, \dots, d, \quad (1)$$

where  $\star$  denotes the convolution operator,  $\mathbf{H}_c$  is the blurring kernel that is common for all spectral bands,  $\mathbf{X}_i$  is the ground truth image at the  $i$ -th spectral band and  $\mathbf{W}_i$  is a zero-mean Gaussian noise term. Given the corrupted hyperspectral image  $\mathbf{Y} = \{\mathbf{Y}_i\}_{i=1}^d$ , the scope of this work is to recover the corresponding ground truth image  $\mathbf{X}$ . In this work we tackle this inverse problem by considering learnable regularizers, trained so as to capture the underlying inherent structure of the considered hyperspectral images, and leverage this idea by employing a deep equilibrium approach. In particular, we propose the optimization problem

$$\arg \min_{\mathbf{X}} \frac{1}{2} \sum_{i=1}^d \|\mathbf{Y}_i - \mathbf{H}_c \star \mathbf{X}_i\|_F^2 + \lambda \mathcal{R}(\mathbf{X}), \quad (2)$$

that consists of a data consistency term and a learnable regularizer  $\mathcal{R}(\cdot)$  aiming to capture the spatial and spectral dependencies of the whole estimated hyperspectral image  $\mathbf{X} = \{\mathbf{X}_i\}_{i=1}^d$ . Moreover,  $\lambda$  is the regularization parameter.

## 3. PROPOSED METHOD

### 3.1. HQS solver for the proposed optimization problem

To effectively tackle (2), the Half Quadratic Splitting (HQS) [14] methodology is employed, thus deriving an equivalent constrained version of (2), i.e.,

$$\arg \min_{\mathbf{X}} \frac{1}{2} \sum_{i=1}^d \|\mathbf{Y}_i - \mathbf{H}_c \star \mathbf{X}_i\|_F^2 + \lambda \mathcal{R}(\mathbf{Z}) \quad (3)$$

$$s.t. \quad \mathbf{Z} - \mathbf{X} = 0,$$

where  $\mathbf{Z} \in \mathbb{R}^{M \times N \times d}$  denotes an auxiliary variable. The corresponding augmented Lagrangian function then becomes

$$\mathcal{L} = \frac{1}{2} \sum_{i=1}^d \|\mathbf{Y}_i - \mathbf{H}_c \star \mathbf{X}_i\|_F^2 + \lambda \mathcal{R}(\mathbf{Z}) + \frac{b}{2} \|\mathbf{Z} - \mathbf{X}\|_F^2 \quad (4)$$

where  $b$  is a penalty parameter. Hence, from (4) a sequence of individual sub-problems emerges:

$$\mathbf{X}^{(k+1)} = \arg \min_{\mathbf{X}} \frac{1}{2} \sum_{i=1}^d \|\mathbf{Y}_i - \mathbf{H}_c \star \mathbf{X}_i\|_F^2 + \frac{b}{2} \|\mathbf{Z}^{(k)} - \mathbf{X}\|_F^2 \quad (5a)$$

$$\mathbf{Z}^{(k+1)} = \arg \min_{\mathbf{Z}} \lambda \mathcal{R}(\mathbf{Z}) + \frac{b}{2} \|\mathbf{Z} - \mathbf{X}^{(k+1)}\|_F^2. \quad (5b)$$

Focusing on sub-problem (5a), and considering that the blurring kernel is the same across the spectral dimension, a more compact solution can be derived by utilizing the convolution theorem, thus the problem in (5a) can be written in the Fourier domain as follows

$$\tilde{\mathbf{X}}^{(k+1)} = \arg \min_{\tilde{\mathbf{X}}} \frac{1}{2} \|\tilde{\mathbf{Y}} - \tilde{\mathbf{H}} \odot \tilde{\mathbf{X}}\|_F^2 + \frac{b}{2} \|\tilde{\mathbf{Z}}^{(k)} - \tilde{\mathbf{X}}\|_F^2 \quad (6)$$

where  $\tilde{\mathbf{Y}} = \{\mathcal{F}(\mathbf{Y}_i)\}_{i=1}^d$ ,  $\tilde{\mathbf{H}} = \{\mathcal{F}(\mathbf{H}_i)\}_{i=1}^d$ ,  $\tilde{\mathbf{X}} = \{\mathcal{F}(\mathbf{X}_i)\}_{i=1}^d$ ,  $\tilde{\mathbf{Z}} = \{\mathcal{F}(\mathbf{Z}_i)\}_{i=1}^d$  denote the concatenation of the discrete 2D Fourier transforms for each spectral band of the respective spatial domain signals (i.e.,  $\mathbf{Y}_i, \mathbf{H}_i, \mathbf{X}_i, \mathbf{Z}_i$ ),  $\mathcal{F}(\cdot)$  is the 2D Fourier transform and  $\odot$  is the element-wise multiplication operator. The solution of sub-problem (6) is given by

$$\tilde{\mathbf{X}}^{(k+1)} = (\tilde{\mathbf{H}} \odot \tilde{\mathbf{H}} + b \mathbb{1})^{-1} \odot (\tilde{\mathbf{H}} \odot \tilde{\mathbf{Y}} + b \tilde{\mathbf{Z}}^{(k)}). \quad (7)$$

where  $\mathbb{1}$  is a matrix with ones. Additionally, sub-problem (5b) can be written as follows

$$\mathbf{Z}^{(k+1)} = \arg \min_{\mathbf{Z}} \frac{1}{2(\sqrt{\lambda/b})^2} \|\mathbf{Z} - \mathbf{X}^{(k+1)}\|_F^2 + \mathcal{R}(\mathbf{Z}) \quad (8)$$

Based on Bayesian estimation theory, relation (8) can be interpreted as a Gaussian denoiser with noise level  $(\sqrt{\lambda/b})$  [11]. In light of this, we employ a neural network  $f_{\theta}(\cdot)$  to act as denoiser whose weights ( $\theta$ ), can be learned from training data, and thus it can be designed with properties adapted to the signals of interested. Thus, equation (8) can be written as

$$\mathbf{Z}^{(k+1)} = f_{\theta}(\mathbf{X}^{(k+1)}) \quad (9)$$

Thus, the iteration map consists of the following equations:

$$\mathbf{Z}^{(k)} = f_{\theta}(\mathbf{X}^{(k)}) \quad (10a)$$

$$\tilde{\mathbf{Z}}^{(k)} = \{\mathcal{F}(\mathbf{Z}_i^{(k)})\}_{i=1}^d \quad (10b)$$

$$\tilde{\mathbf{X}}^{(k+1)} = (\tilde{\mathbf{H}} \odot \tilde{\mathbf{H}} + b\mathbb{1})^{-1} \odot (\tilde{\mathbf{H}} \odot \tilde{\mathbf{Y}} + b\tilde{\mathbf{Z}}^{(k)}) \quad (10c)$$

$$\mathbf{X}^{(k+1)} = \{\mathcal{F}^{-1}(\tilde{\mathbf{X}}_i^{(k+1)})\}_{i=1}^d \quad (10d)$$

where  $\mathcal{F}^{-1}(\cdot)$  denotes the 2D inverse Fourier transform. Note that the neural-network (denoiser)  $f_{\theta}$  can be pre-trained offline via the loss function  $\sum_{p=1}^P \|f_{\theta}(\mathbf{X}^p + \mathbf{W}; \theta) - \mathbf{X}^p\|_F^2$ , employing pairs of generated noisy hyperspectral images (using Gaussian noise, denoted as  $\mathbf{W}$ ) and their corresponding ground truth versions denoted as  $\{\mathbf{X}^p + \mathbf{W}, \mathbf{X}^p\}_{p=1}^P$ .

### 3.2. Proposed Deep equilibrium model

**Preliminaries:** Deep equilibrium (DEQ) models aim to design (equivalent) infinite-depth neural networks by expressing the entire deep architecture as an equilibrium (fixed point) computation [18]. Consider a generic  $K$ -layer deep feedforward model expressed by the following recursion

$$x^{(k+1)} = g_{\theta}^{(k)}(x^{(k)}; y), \quad k = 0, 1 \dots K-1, \quad (11)$$

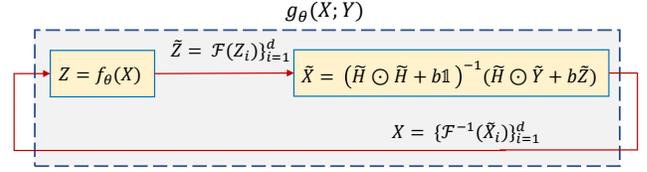
where  $k$  is the layer index,  $x^{(k)}$  denotes the output of the  $k$ -th layer,  $y$  is an input common to all the layers and  $g_{\theta}^{(k)}(\cdot)$  stands for some nonlinear transformation that corresponds to the operation of the  $k$ -th layer of the neural network.

DEQ models assume that the nonlinear transformation is exactly the same for all the layers of the considered neural network. We denote this common transformation as  $g_{\theta}(\cdot)$ . Under this “weight tying” strategy, in [18] it was recognised that every output of such an infinite depth should be a fixed point of the transformation (iteration map)  $g_{\theta}(\cdot)$ , thus obeying the equation

$$x^* = g_{\theta}(x^*; y). \quad (12)$$

**Proposed model:** Starting from the optimization algorithm in (10), some possible directions to proceed further could be the following. One approach would be the so-called “plug and play” method, that is, to pretrain a neural network and use it in ((10a)) assuming that its parameters remain constant during the iterative process. Another approach would be to consider a fixed number of iterations of (10), unroll these iterations to generate an equivalent deep learning architecture, and train the parameters of this model end-to-end, an approach known as deep unrolling. Thus, for  $K$  iterations of the HQS solver, we can derive a deep learning network with  $K$  layers. However, due to computational constraints the number of unrolled iterations must be kept small [17]. Given the drawbacks of the plug-and-play approach [20] and the complexity requirements of the deep unrolling method [17], in this work we follow an alternative approach, known as the deep equilibrium method, corresponding to an infinite-depth and more efficient architecture. In particular, we denote the operation of the four equations in (10) as  $g_{\theta}(\mathbf{X}, \mathbf{Y})$ , and rewrite it as

$$\mathbf{X}^{(k+1)} = g_{\theta}(\mathbf{X}^{(k)}; \mathbf{Y}). \quad (13)$$



**Fig. 1.** An illustration of the proposed end-to-end Deep equilibrium model, derived from the iteration map in eq. 10.

Furthermore, similarly to algorithm unrolling, we consider that equation (13) is used to construct an equivalent deep learning model, but, different from algorithm unrolling, we consider that this model has an infinite number of layers. Interestingly, if we consider that the iteration in (13) is applied an infinite number of times, then the resulting output should be a fixed-point of the iteration map  $g_{\theta}(\mathbf{X}, \mathbf{Y})$  and thus obey the equation

$$\mathbf{X}^* = g_{\theta}(\mathbf{X}^*; \mathbf{Y}). \quad (14)$$

This fact can be used to significantly accelerate the computation of the outputs of the considered infinite-length model, termed deep equilibrium model, as also discussed in the sequel. Also, the parameters of the proposed model (i.e., the neural network  $f_{\theta}(\cdot)$  and the penalty parameter  $b$ ) can be trained end-to-end, using the loss function

$$\sum_{p=1}^P \|g_{\theta}(\mathbf{X}^{*,p}; \mathbf{Y}^p) - \mathbf{X}^{*,p}\|_F^2, \quad (15)$$

where  $\{\mathbf{X}^p, \mathbf{Y}^p\}$  represent  $P$  suitable training pairs of clean hyperspectral images and the corresponding noisy and blurred images. Also,  $g_{\theta}(\mathbf{X}^{*,p}; \mathbf{Y}^p) = \mathbf{X}^{*,p}$  denotes the output of the supposed infinite length neural network when the degraded hyperspectral image  $\mathbf{Y}^p$  is applied at the input. Figure (1) illustrates the proposed architecture.

*Efficient computation of the forward pass:* During the training and the testing phases of the proposed model, several fixed points must be calculated given the iteration map in (10) and some degraded hyperspectral images. A simple, yet computationally demanding, approach is to apply the equations given in (13) until convergence, i.e., until  $\mathbf{X}^{(k+1)}$  and  $\mathbf{X}^{(k)}$  are sufficiently close to each other. However, the computation of fixed points for iterative mappings is a well studied problem, and various acceleration methods have been proposed. In this work, the Anderson acceleration methodology [22] is used to accelerate the process of estimating fixed points.

*Efficient computation of the backward pass:* During the backpropagation stage, in order to avoid backpropagating through a large number of fixed point iterations, the methodology proposed in [18], [23] is employed to express the backpropagation parameter update step as a fixed point computation problem, which can be integrated in standard automatic differentiation tools. Due to space limitations the details of this part are omitted.

## 4. EXPERIMENTAL PART

**Dataset and experimental set up:** To highlight the efficacy of the proposed deep equilibrium model for tackling the HSI deconvolution problem, we employed two publicly available hyperspectral datasets, that is CAVE [24] dataset consisting of 32,  $512 \times 512 \times 31$  images and Harvard [25] dataset comprised of 50,  $1024 \times 1024 \times 31$  images, with  $d = 31$  spectral bands of 10 nm, covering the visible spectrum 400-700 nm. Following the exact same experimental set up with study in [13], we used the following blurring kernels and Gaussian noise with standard deviation  $\sigma$  to create the degraded images: (a)  $15 \times 15$  Gaussian kernel with bandwidth  $\sigma_k = 1.6$ , and  $\sigma = 0.01$ , (b)  $15 \times 15$  Gaussian kernel with bandwidth  $\sigma_k = 2.4$ , and  $\sigma = 0.01$ , (c)  $15 \times 15$  Gaussian kernel with bandwidth  $\sigma_k = 1.6$ , and  $\sigma = 0.03$ , (d) Circle kernel with diameter 7, and  $\sigma = 0.01$ , (e) Square kernel with side length of 5, and  $\sigma = 0.01$ . Similarly to [13], the first 20 images from the CAVE dataset were employed for training and the rest were used as test set. Regarding the Harvard dataset, the first 30 were used as training set and the rest as test set.

**Neural Network Architecture:** Concerning, the neural network  $f_\theta(\cdot)$ , we employed a network consisting of 5 layers, where each layer has 128 filters with size  $3 \times 3$ . Additionally, the ReLU was used as activation function. The CNN network first was pre-trained using pairs of noisy images corrupted by Gaussian noise and their corresponding ground truth images. During this training process the ADAM optimizer was used with learning rate equal  $1e - 04$  and batch size equal to 16, whereas the number of epochs was set to 300.

**Proposed Model-Parameter Setting:** Focusing on the proposed model, we employed the Anderson acceleration procedure [22] for the forward and backward pass fixed-point iterations. In particular, the number of fixed-point iteration was set to 20. Finally, during the end-to-end training stage of the proposed model, the ADAM optimizer was employed with learning rate equal  $1e - 04$ , batch size equal to 6 and number of epochs to 250.

**Compared methods:** To validate the quality of the estimated hyperspectral images, we used the Peak Signal to Noise Ratio (PSNR), the Structural Similarity Index (SSIM) [26] metrics and the Spectral Angle Mapper (SAM) [27], comparing the proposed deep equilibrium model with the approaches of [13], [28] and [7]. Method [13] is a plug-and-play approach learning offline an unknown regularizer, whereas the other methods employ well-designed spectral and spatial priors. The methods [7,28] were selected because they constitute well-established/benchmark methods. Note that the state-of-the-art method [13] was also compared with the previously mentioned approaches.

**Hyperspectral Image Deconvolution Results :** Tables 1 and 2 summarize the average quantitative results of in comparison with several well-established hyperspectral deblurring algorithms. It is evident that the proposed deep equi-

**Table 1. Results on the CAVE dataset**

Metric	Method 1 [27]	Method 2 [7]	Method 3 [13]	Proposed	
(a)	PSNR	37.42	36.53	41.17	<b>41.78</b>
	SSIM	0.9257	0.9474	0.9602	<b>0.9793</b>
	SAM	11.19	8.24	6.27	<b>5.94</b>
(b)	PSNR	35.34	34.69	35.96	<b>38.51</b>
	SSIM	0.9058	0.9255	0.9199	<b>0.9647</b>
	SAM	10.98	8.49	7.77	<b>6.68</b>
(c)	PSNR	30.04	35.65	37.42	<b>39.65</b>
	SSIM	0.6140	0.9161	0.9395	<b>0.9678</b>
	SAM	25.83	12.23	6.46	<b>6.13</b>
(d)	PSNR	36.10	35.60	39.88	<b>41.77</b>
	SSIM	0.9116	0.9350	0.9477	<b>0.9801</b>
	SAM	11.52	8.41	6.80	<b>6.01</b>
(e)	PSNR	37.05	36.53	42.71	<b>45.21</b>
	SSIM	0.9211	0.9463	0.9669	<b>0.9884</b>
	SAM	11.77	8.33	6.29	<b>5.55</b>

**Table 2. Results on the Harvard dataset**

Metric	Method 1 [27]	Method 2 [7]	Method 3 [13]	Proposed	
(a)	PSNR	38.18	38.65	41.11	<b>42.57</b>
	SSIM	0.9133	0.9353	0.9486	<b>0.9626</b>
	SAM	8.41	5.06	4.84	<b>4.10</b>
(b)	PSNR	36.88	37.22	38.85	<b>40.10</b>
	SSIM	0.8903	0.9110	0.9202	<b>0.9335</b>
	SAM	8.07	5.10	4.77	<b>4.54</b>
(c)	PSNR	30.51	37.07	38.95	<b>40.87</b>
	SSIM	0.5947	0.9046	0.9165	<b>0.9439</b>
	SAM	23.24	8.22	4.58	<b>4.02</b>
(d)	PSNR	37.35	38.08	40.50	<b>42.66</b>
	SSIM	0.8992	0.9232	0.9384	<b>0.9612</b>
	SAM	8.58	5.08	4.82	<b>4.33</b>
(e)	PSNR	37.92	38.77	41.80	<b>43.13</b>
	SSIM	0.9090	0.9347	0.9541	<b>0.9669</b>
	SAM	8.86	5.08	4.87	<b>4.21</b>

librium method constantly outperforms the other approaches in various blurring scenarios. Additionally, for high levels of noise, the proposed end-to-end model is able to maintain high metric values in all cases. Compared to the plug-and-play method in [13], the proposed methodology provides notably better reconstruction results, since the CNN denoiser along with the penalty parameters derived from the proposed optimization problem are optimized end-to-end based on the quality of the estimated results and adapted to the problem at hand. It should be highlighted that the superiority of the proposed method is achieved with no additional complexity as compared to the other approaches.

## 5. CONCLUSIONS

In this work, the problem of hyperspectral image deconvolution was studied. Considering the ill-posed nature of the examined problem, a regularized optimization problem was proposed, utilizing a learnable CNN to capture the underlying spectral and spatial dependencies of the hyperspectral data. A Deep equilibrium model algorithm was developed to solve efficiently the considered problem, thus forming a highly interpretable and infinite-depth deep learning based architecture. Extensive numerical results demonstrated the efficacy and applicability of the proposed method.

## 6. REFERENCES

- [1] N. Kussul, M. Lavreniuk, S. Skakun, and A. Shelestov, "Deep learning classification of land cover and crop types using remote sensing data," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 5, pp. 778–782, 2017.
- [2] R. Calvini, A. Ulrici, and J. M. Amigo, "Growing applications of hyperspectral and multispectral imaging," *Data Handling in Science and Technology*, vol. 32, pp. 605–629, 01 2019.
- [3] Y. Song, D. Brie, E.-H. Djermoune, and S. Henrot, "Regularization parameter estimation for non-negative hyperspectral image deconvolution," *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5316–5330, 2016.
- [4] S. Bongard, F. Soulez, E. Thiebaut, and E. Pecontal, "3D deconvolution of hyper-spectral astronomical data," *Monthly Notices of the Royal Astronomical Society*, vol. 418, no. 1, 11 2011.
- [5] J.-M. Gaucel, M. Guillaume, and S. Bourennane, "Adaptive-3d-wiener for hyperspectral image restoration: Influence on detection strategy," in *2006 14th European Signal Processing Conference*, 2006, pp. 1–5.
- [6] R. Neelamani, H. Choi, and R. Baraniuk, "Forward: Fourier-wavelet regularized deconvolution for ill-conditioned systems," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 418–433, 2004.
- [7] S. Henrot, C. Soussen, and D. Brie, "Fast positive deconvolution of hyperspectral images," *IEEE Transactions on Image Processing*, vol. 22, no. 2, pp. 828–833, 2013.
- [8] Y. Song, E.-h. Djermoune, J. Chen, C. Richard, and D. Brie, "Online Deconvolution for Industrial Hyperspectral Imaging Systems," *SIAM Journal on Imaging Sciences*, vol. 12, no. 1, pp. 54–86, 2019.
- [9] Y. Song, D. Brie, E.-H. Djermoune, and S. Henrot, "Regularization parameter estimation for non-negative hyperspectral image deconvolution," *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5316–5330, 2016.
- [10] D. Gilton, G. Ongie, and R. Willett, "Learned patch-based regularization for inverse problems in imaging," in *2019 IEEE 8th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, 2019, pp. 211–215.
- [11] K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep cnn denoiser prior for image restoration," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.
- [12] G. Ongie, A. Jalal, C. A. Metzler, R. G. Baraniuk, A. G. Dimakis, and R. Willett, "Deep learning techniques for inverse problems in imaging," *IEEE Journal on Selected Areas in Information Theory*, vol. 1, no. 1, pp. 39–56, 2020.
- [13] X. Wang, J. Chen, C. Richard, and D. Brie, "Learning spectral-spatial prior via 3ddncnn for hyperspectral image deconvolution," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 2403–2407.
- [14] D. Geman and C. Yang, "Nonlinear image recovery with half-quadratic regularization," *IEEE Transactions on Image Processing*, vol. 4, no. 7, pp. 932–946, 1995.
- [15] Y. Yang, J. Sun, H. Li, and Z. Xu, "Admm-csnet: A deep learning approach for image compressive sensing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 3, pp. 521–538, 2020.
- [16] I. Y. Chun, Z. Huang, H. Lim, and J. Fessler, "Momentum-net: Fast and convergent iterative neural network for inverse problems," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–1, 2020.
- [17] V. Monga, Y. Li, and Y. C. Eldar, "Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing," *IEEE Signal Processing Magazine*, vol. 38, no. 2, pp. 18–44, 2021.
- [18] S. Bai, J. Z. Kolter, and V. Koltun, "Deep Equilibrium Models," in *Advances in Neural Information Processing Systems*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, Eds., vol. 32, 2019.
- [19] A. Gkillas, D. Ampeliotis, and K. Berberidis, "Connections between deep equilibrium and sparse representation models with application to hyperspectral image denoising," *IEEE Transactions on Image Processing*, vol. 32, pp. 1513–1528, 2023.
- [20] D. Gilton, G. Ongie, and R. Willett, "Deep equilibrium architectures for inverse problems in imaging," *IEEE Transactions on Computational Imaging*, vol. 7, pp. 1123–1133, 2021.
- [21] S. Henrot, C. Soussen, and D. Brie, "Fast positive deconvolution of hyperspectral images," *IEEE Transactions on Image Processing*, vol. 22, no. 2, pp. 828–833, 2013.
- [22] H. F. Walker and P. Ni, "Anderson Acceleration for Fixed-Point Iterations," <http://dx.doi.org/10.1137/10078356X>, vol. 49, no. 4, pp. 1715–1735, aug 2011.
- [23] "Deep Implicit Layers - Neural ODEs, Deep Equilibrium Models, and Beyond." [Online]. Available: <http://implicit-layers-tutorial.org/>
- [24] F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar, "Generalized assorted pixel camera: Postcapture control of resolution, dynamic range, and spectrum," *IEEE Transactions on Image Processing*, vol. 19, no. 9, pp. 2241–2253, 2010.
- [25] A. Chakrabarti and T. Zickler, "Statistics of Real-World Hyperspectral Images," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2011, pp. 193–200.
- [26] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [27] F. A. Kruse, A. B. Lefkoff, J. W. Boardman, K. B. Heidebrecht, A. T. Shapiro, P. J. Barloon, and A. F. H. Goetz, "The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data," *Remote Sensing of Environment*, vol. 44, no. 2, pp. 145–163, 1993.
- [28] D. Krishnan and R. Fergus, "Fast Image Deconvolution using Hyper-Laplacian Priors," in *Advances in Neural Information Processing Systems*, Y. Bengio, D. Schuurmans, J. Lafferty, C. Williams, and A. Culotta, Eds., vol. 22. Curran Associates, Inc., 2009.