# BLIND UNMIXING USING A DOUBLE DEEP IMAGE PRIOR

*Chao Zhou, Miguel R.D. Rodrigues*

Dept. Electronic and Electrical Engineering, University College London
{chao.zhou.18, m.rodrigues}@ucl.ac.uk

## ABSTRACT

In this paper, we propose a novel network structure to solve the blind hyperspectral unmixing problem using a double Deep Image Prior (DIP). In particular, the blind unmixing problem involves two sub-problems: endmember estimation and abundance estimation. We, therefore, propose two sub-networks, endmember estimation DIP (EDIP) and abundance estimation DIP (ADIP), to generate the estimation of endmembers and estimation of corresponding abundances respectively. The overall network is then constructed by assembling these two sub-networks. The network is trained in an end-to-end manner by minimizing a novel composite loss function. The experiments on synthetic and real datasets show the effectiveness of the proposed method over state-of-art unmixing methods.

***Index Terms***— Hyperspectral unmixing, blind unmixing, neural networks, deep image prior (DIP)

## 1. INTRODUCTION

The hyperspectral image (HSI) is a multivariate image that captures spectral information over hundreds of spectral bands of a certain scene. The blind unmixing problem aims, for each pixel, to extract the spectra of the constituent materials and the corresponding fractional proportions, also known as endmembers and abundances, respectively. This type of problem often arises in areas where there is a need to understand materials of a scene/sample such as remote sensing [1, 2], art investigation [3], and others.

The blind unmixing problem involves performing two tasks [2]: endmember estimation and abundance estimation. Most endmember estimation methods are geometrical based approaches by assuming the data is embedded in a simplex, the vertices of which are the endmembers. Popular examples of such methods include vertex component analysis (VCA) [4] and simplex volume maximization (SiVM) [5]. On the other hand, most abundance estimation methods in the literature are based on linear mixture model (LMM) [6], which assumes the observed spectrum is a linear combination of the endmembers' spectra weighted by the corresponding abundance. When the endmembers are estimated by endmember estimation methods, the blind unmixing problem reduces to a least square problem, which can be solved by fully constrained least square (FCLS) [7] solver. When the endmembers are known in the form of a rich spectral library, the abundance estimation problem becomes a sparse regression (SR) problem [8], which has been solved by methods such as sparse unmixing by variable splitting and augmented Lagrangian (SunSAL) [8] and collaborative SUnSAL (CLSunSAL) [9]. However, these approaches can be computationally complex.

Many neural networks (NN) based approaches have also been proposed to tackle the unmixing challenge. Generally, these methods can be divided into supervised and unsupervised methods. In supervised methods, such as unfolding ADMM based abundance estimation network (U-ADMM-AENet) [10], the networks are fed with a set of pairs of HSI spectra and corresponding abundances. After learning, the network can map the spectra to corresponding abundances. However, such methods require access to the true abundance.

In unsupervised methods, such as MNN-BU [11] and uDAS [12], the networks are based on autoencoder structures, which take only the HSI spectra as input and enforce the output to reconstruct HSI spectra. After learning, the bottleneck of the autoencoder gives the abundance estimation and the weights of a linear decoder give the endmember estimation. These methods can however fail to deliver accurate endmembers and abundances [13].

Recently, UnDIP [14] has been proposed to overcome some of these challenges by using deep image prior (DIP) [15]. However, this approach only yields abundance estimation, relying on the availability of an estimation of the endmembers using other existing methods. Therefore, this approach is not suitable for blind unmixing purposes.

In this paper, we propose a novel blind unmixing network using double DIP which delivers improved endmember and abundance estimation. This paper is organized as follows: In Section 2, we introduce the blind unmixing problem. In Section 3, we present the proposed blind unmixing network using double DIP. Section 4 illustrates that our proposed network outperforms existing ones in unmixing tasks. Conclusions are drawn in Section 5.

## 2. THE BLIND UNMIXING PROBLEM

We consider the linear mixing model (LMM) given by [6]

$$\mathbf{Y} = \mathbf{E}\mathbf{A} + \mathbf{N} \tag{1}$$

where $\mathbf{Y} \in R^{p\times n}$ is a HSI data cube corresponding to $n$ pixels across $p$ spectral bands, $\mathbf{E} = [\mathbf{e}_1, ..., \mathbf{e}_r] \in R^{p\times r}$ is the endmember matrix containing $r$ endmembers across $p$ spectral bands, $\mathbf{A} = [\mathbf{a}_1, ..., \mathbf{a}_n] \in R^{r\times n}$ is the corresponding fractional abundance matrix, and $\mathbf{N} \in R^{p\times n}$ is the additive noise. Note that the abundance is subjected to non-negative constraint (ANC) and sum-to-one constraint (ASC), i.e., $\mathbf{A} \geq \mathbf{0}$ and $\mathbf{A}^T\mathbf{1}_r = \mathbf{1}_n$, where $\mathbf{1}_r$ is the all one vector with size $r \times 1$. Similarly, the endmember matrix is also subjected to non-negative constraint (ENC), $\mathbf{E} \geq \mathbf{0}$, in order to be physically meaningful.

The goal of blind unmixing is to estimate $\mathbf{E}$ and $\mathbf{A}$ given only $\mathbf{Y}$. A popular approach to solve this problem involves solving [16]:

$$\begin{aligned}\hat{\mathbf{E}}, \hat{\mathbf{A}} = \underset{\mathbf{E},\mathbf{A}}{\arg\min}\, \frac{1}{2}\|\mathbf{Y} - \mathbf{E}\mathbf{A}\|_F^2 + R(\mathbf{A}) \\ s.t., \mathbf{E} \geq \mathbf{0}, \mathbf{A} \geq \mathbf{0}, \mathbf{A}^T\mathbf{1}_r = \mathbf{1}_n\end{aligned} \tag{2}$$

where, $R$ is a regularizer depending on abundance matrix $\mathbf{A}$, such as total variation (TV) [16]. Generally, the choice of $R$ is heavily dependent on the prior knowledge about the task at hand.

## 3. BLIND UNMIXING USING DOUBLE DEEP IMAGE PRIOR

The blind unmixing problem (2) is usually solved by a two-stage cyclic descent method [16]: fix $\mathbf{A}$, update $\mathbf{E}$; and fix $\mathbf{E}$, update $\mathbf{A}$. We now propose an end-to-end network to solve the problem (2) using DIP techniques. DIP [15] was originally proposed to solve inverse problems such as denoising, given by:

$$\boldsymbol{x}^* = \underset{\boldsymbol{x}}{\arg\min}\, \|\boldsymbol{x} - \boldsymbol{x}_0\|_2^2 + R(\boldsymbol{x}) \tag{3}$$

where, $\boldsymbol{x}_0$ is a noisy image, and $R$ is a regularizer explicitly capturing the prior information about clean image $\boldsymbol{x}$. DIP however solves the optimization problem given by:

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta}}{\arg\min}\, \|f_{\boldsymbol{\theta}}(\boldsymbol{z}) - \boldsymbol{x}_0\|_2^2 \tag{4}$$

where, $f_{\boldsymbol{\theta}}(\boldsymbol{z})$ is a neural network parameterized by $\boldsymbol{\theta}$, with a random input $\boldsymbol{z}$. That is, DIP effectively replaces the regularizer in (3) with a neural network. After learning, the network parameterization would implicitly capture the prior and output the restored image given by $\boldsymbol{x}^* = f_{\boldsymbol{\theta}^*}(\boldsymbol{z})$. We next describe how to adapt this approach to solve our proposed blind unmixing problem.

### 3.1. Network Structure

Our proposed network consists of two sub-networks: Endmember estimation DIP (EDIP) and Abundance estimation DIP (ADIP). EDIP is responsible for endmember estimation whereas ADIP is responsible for abundance estimation.

**EDIP:** Let us first assume that, at the endmember estimation stage, we are given access to an estimate of the abundance $\tilde{\mathbf{A}}$, using some algorithm such as FCLS [7]. Then, the optimization problem (2) would reduce to:

$$\hat{\mathbf{E}} = \underset{\mathbf{E}}{\arg\min}\, \frac{1}{2}\|\mathbf{Y} - \mathbf{E}\tilde{\mathbf{A}}\|_F^2 \quad s.t., \mathbf{E} \geq \mathbf{0} \tag{5}$$

In this work, following the idea of DIP, we propose to use a network $f_{\boldsymbol{\theta}_E}$ with a random input $\boldsymbol{z}_E$ to estimate the endmembers. This leads to the optimization problem given by:

$$\hat{\boldsymbol{\theta}}_E = \underset{\boldsymbol{\theta}_E}{\arg\min}\, \frac{1}{2}\|\mathbf{Y} - f_{\boldsymbol{\theta}_E}(\boldsymbol{z}_E)\tilde{\mathbf{A}}\|_F^2 \tag{6}$$

Similar to [15], we use ResNet [17] structure for the proposed EDIP network $f_{\boldsymbol{\theta}_E}$. Note that, in order to meet the ENC, we use a sigmoid activation as the last layer of EDIP. After learning the parameter $\boldsymbol{\theta}_E^*$, the network would estimate the endmember, given by $\hat{\mathbf{E}} = f_{\boldsymbol{\theta}_E^*}(\boldsymbol{z}_E)$. The EDIP sub-network is illustrated in Fig.1.

**ADIP:** Let us now assume that, at the abundance estimation stage, we are given access to an estimate of the endmembers $\tilde{\mathbf{E}}$, using some algorithm such as SiVM [5]. Then, the optimization problem (2) would reduce to:

$$\begin{aligned}\hat{\mathbf{A}} = \underset{\mathbf{A}}{\arg\min}\, \frac{1}{2}\|\mathbf{Y} - \tilde{\mathbf{E}}\mathbf{A}\|_F^2 + R(\mathbf{A}) \\ s.t., \mathbf{A} \geq \mathbf{0}, \mathbf{A}^T\mathbf{1}_r = \mathbf{1}_n\end{aligned} \tag{7}$$

We also propose to use a network $f_{\boldsymbol{\theta}_A}$ with a random input $\boldsymbol{z}_A$ to generate the estimation of abundance. This leads to the optimization problem given by:

$$\hat{\boldsymbol{\theta}}_A = \underset{\boldsymbol{\theta}_A}{\arg\min}\, \frac{1}{2}\|\mathbf{Y} - \tilde{\mathbf{E}}f_{\boldsymbol{\theta}_A}(\boldsymbol{z}_A)\|_F^2 \tag{8}$$

We also adopt ResNet [17] structure for the proposed ADIP network $f_{\boldsymbol{\theta}_A}$. In order to meet ASC and ANC, we use softmax as the output layer of ADIP. After learning the parameter $\boldsymbol{\theta}_A^*$, the estimated abundance is given by $\hat{\mathbf{A}} = f_{\boldsymbol{\theta}_A^*}(\boldsymbol{z}_A)$. The ADIP sub-network is illustrated in Fig.1.

**Overall Structure:** After obtaining an estimation of endmember and abundance, $\hat{\mathbf{E}}$ and $\hat{\mathbf{A}}$, using EDIP and ADIP respectively, we can immediately generate a reconstruction of the HSI image, as follows:

$$\hat{\mathbf{Y}} = \hat{\mathbf{E}}\hat{\mathbf{A}} \tag{9}$$

Thus, the overall structure of the proposed network is obtained by assembling the proposed EDIP, ADIP, as shown in Fig.1. We name the proposed blind unmixing network using double DIP as BUDDIP.

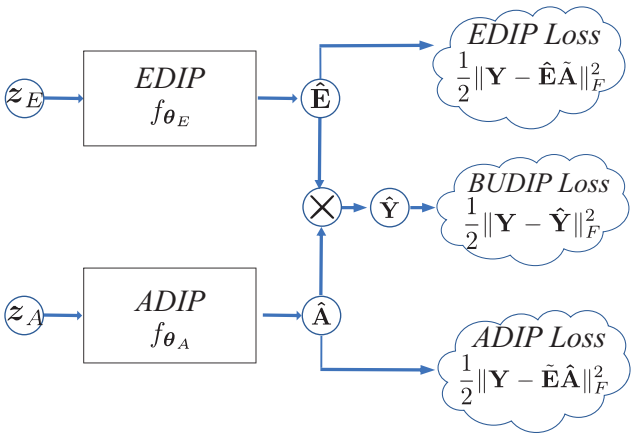



**Fig. 1**. Blind Unmixing using Double DIP (BUDDIP)

### 3.2. Network Optimization

We train our overall network using various loss functions. First, in line with optimization problem (6), we use the loss function given by:

$$L_{EDIP} = \frac{1}{2}\|\mathbf{Y} - f_{\boldsymbol{\theta}_E}(\boldsymbol{z}_E)\tilde{\mathbf{A}}\|_F^2 \tag{10}$$

Similarly, in line with optimization problem (8), we also use the loss function given by:

$$L_{ADIP} = \frac{1}{2}\|\mathbf{Y} - \tilde{\mathbf{E}}f_{\boldsymbol{\theta}_A}(\boldsymbol{z}_A)\|_F^2 \tag{11}$$

Note that, $\tilde{\mathbf{E}}$ in $L_{ADIP}$ and $\tilde{\mathbf{A}}$ in $L_{EDIP}$ are estimated by SiVM [5] and FCLS respectively.

Moreover, we also impose an additional loss function – the blind unmixing (BU) loss – given by:

$$L_{BU} = \frac{1}{2}\|\mathbf{Y} - \hat{\mathbf{Y}}\|_F^2 \tag{12}$$

This additional loss is necessary because otherwise EDIP and ADIP would yield endmember and abundance estimates close to $\tilde{\mathbf{E}}$ and $\tilde{\mathbf{A}}$, respectively.

The final loss function is a combination of the above losses, as follows:

$$L = \alpha_1 L_{EDIP} + \alpha_2 L_{ADIP} + \alpha_3 L_{BU} \tag{13}$$

where, $\alpha_1, \alpha_2, \alpha_3$ are the loss weights that control the relative importance of corresponding loss terms. $L_{EDIP}$ and $L_{ADIP}$ together can make sure the network outputs meaningful endmembers and abundances, that is, $\hat{\mathbf{E}}, \hat{\mathbf{A}}$ would be centered around $\tilde{\mathbf{E}}, \tilde{\mathbf{A}}$. $L_{BU}$, on the other hand, allow the network to search for a better estimation than $\tilde{\mathbf{E}}, \tilde{\mathbf{A}}$.

Different from the two-stage-cyclic descent method [16], the proposed network is trained in an end-to-end manner. Given only the HSI image $\mathbf{Y}$, the learnable parameters $\{\boldsymbol{\theta}_E, \boldsymbol{\theta}_A\}$ are learned by minimizing the composite loss $L$, using a variant of gradient descent optimizer, ADAM [18].

## 4. EXPERIMENTS

We now compare our proposed method BUDDIP, to other state-of-the-art methods such as UnDIP [14] and SiVM [5] + FCLS [7]. Note that, in UnDIP, SiVM is also used to generate the fixed endmember estimation. For a fair comparison, in this paper, we also use SiVM+FCLS to generate the $\tilde{\mathbf{E}}$ and $\tilde{\mathbf{A}}$ used in the proposed loss function. However, we will show later that the proposed network can yield better estimations.

We use averaged (over pixels) root mean square error (aRMSE) to measure the quality of abundance estimation and averaged (over endmembers) spectral angle distance (aSAD) to measure the quality of endmember estimation, given by:

$$aRMSE(\mathbf{A}, \hat{\mathbf{A}}) = \frac{1}{n}\sum_{i=1}^{n}\sqrt{\frac{1}{r}\|\mathbf{a}_i - \hat{\mathbf{a}}_i\|_2^2} \tag{14}$$

$$aSAD(\mathbf{E}, \hat{\mathbf{E}}) = \frac{1}{r}\sum_{i=1}^{r}\arccos\left(\mathbf{e}_i, \hat{\mathbf{e}}_i\right)\frac{180}{\pi} \tag{15}$$

We also use both synthetic and real data to evaluate the performance of the proposed approach.

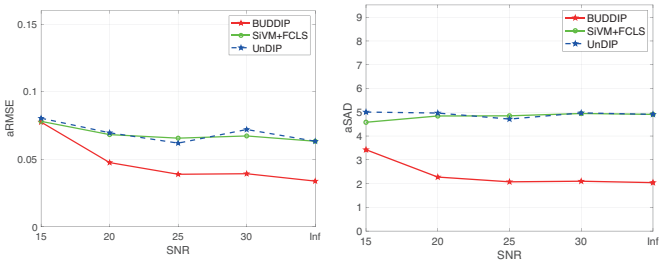
### 4.1. Evaluation on Synthetic Data

The synthetic data is generated according to the procedure in [10]. Six candidate endmember signatures are randomly chosen from the USGS spectral library [19]. A synthetic image of size $a \times a$ is divided into 100 disjoint patches, each of which is assigned with a spectrum mixed from two randomly selected endmembers from the six candidate signatures with fractions $[0.8, 0.2]$. The abundance map is then convolved with a Gaussian filter of size $11\times11$, followed by a pixel-wise re-scaling to meet ASC. Finally, the data is polluted with additive white gaussian noise (AWGN).

By default, the random input $\boldsymbol{z}_A$ and abundance $\hat{\mathbf{A}}$ have the same size $R^{r\times n}$. Similarly, $\boldsymbol{z}_E$ and $\hat{\mathbf{E}}$ have the same size $R^{p\times r}$. The default value of $a$ is 100, that is the training image consists of $100\times100$ pixels, which is then contaminated with AWGN leading to $SNR = 30$ dB. The network is trained using ADAM optimizer [18] with learning rate set to $1e-4$, and the number of epochs set to 4500. We set $\alpha_1 = 0.1, \alpha_2 = 0.01, \alpha_3 = 1.0$, respectively, via grid search techniques, as the proposed method is unsupervised.

**Performance *vs*. SNR** We first evaluate the impact of SNR, which is shown in Fig.2. It can be seen that when the SNR is 15 dB, the proposed methods can achieve better aSAD than the other algorithms. With the increase of SNR, all algorithms can benefit from the decrease of noise. Overall, the proposed methods achieve better unmixing performance than the competitors.

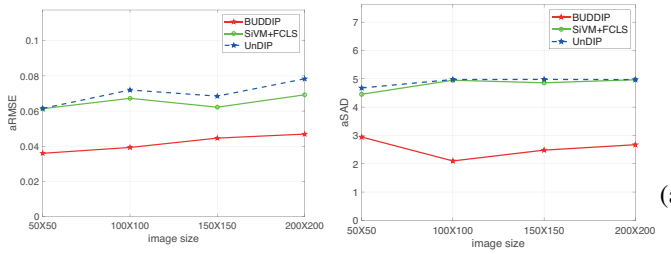
**Performance *vs*. Image Size** We now evaluate the impact of HSI image size on unmixing performance. The results are shown in Fig.3. It can be seen that UnDIP basically can not surpass SiVM+FCLS. On the other hand, as the image




(a) aRMSE vs. SNR (b) aSAD vs. SNR

**Fig. 2**. Performance vs. SNR.

size becomes larger, the performance of SiVM+FCLS deteriorates. As a result, UnDIP and BUDDIP, which rely on SiVM and/or FCLS to provide endmember and/or abundance estimation as guidance, exhibit worse performance. But BUDDIP still achieves better aRMSE and aSAD than the competitors.




(a) aRMSE vs. image size (b) aSAD vs. image size

**Fig. 3**. Performance vs. image size.

### 4.2. Evaluation on Real Data

We also report performance comparison on real HSI dataset, Jasper Ridge, which has four endmembers: Road, Soil, Water and Tree. Due to complexity consideration, a sub-image of size $100\times100$ pixels are used in this experiment. After removing bands due to water vapour effects, 198 bands are used in this paper.

In this experiment, we set the number of epochs to be equal to 24000, and $\alpha_1 = 1.0, \alpha_2 = 1.0, \alpha_3 = 1.0$, respectively, via grid search techniques, as the proposed method is unsupervised. The performance metrics are summarised in Table 1. As we showed before, UnDIP can not achieve better performance than SiVM+FCLS. On the other hand, the proposed BUDDIP can deliver both better endmember estimation and abundance estimation than the competitors. The visual comparisons are illustrated in Fig. 4 and Fig. 5, which also shows the superiority of the proposed method.

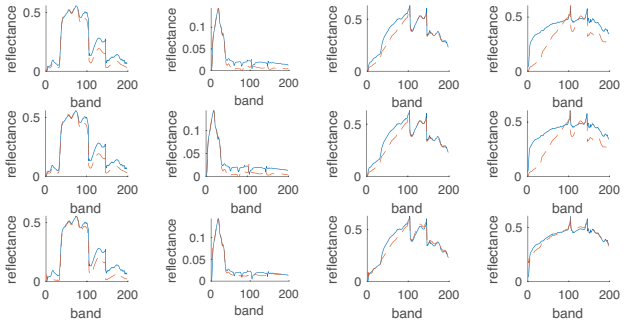
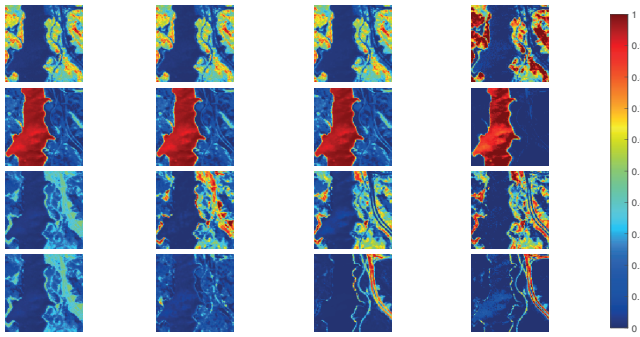



**Fig. 4**. Endmember estimation. From top to bottom: SiVM, UnDIP and BUDDIP. From left to right: Tree, Water, Soil and Road. Solid line is the true value, while dot line is the estimated value.



(a) SiVM+FCLS (b) UnDIP (c) BUDDIP (d) Reference

**Fig. 5**. Abundance estimation maps. From top to bottom: Tree, Water, Soil and Road.

**Table 1**. Unmixing performance by Different Algorithms.

| | SiVM+FCLS | UnDIP | BUDDIP |
|---|---|---|---|
| aRMSE | 0.1480 | 0.1748 | **0.1023** |
| aSAD | 11.3492 | 11.3493 | **6.8489** |
| SAD of Tree | **8.5545** | 8.5545 | 9.2215 |
| SAD of Water | 14.4876 | 14.4877 | **9.2517** |
| SAD of Soil | 6.5558 | 6.5558 | **5.4084** |
| SAD of Road | 15.7991 | 15.7991 | **3.5141** |

## 5. CONCLUSION

In this paper, we have proposed a novel blind unmixing network using double DIP techniques (BUDDIP). In particular, we build two DIP sub-networks to estimate the endmember and abundance respectively, which are coined as EDIP and ADIP. The final network is constructed by assembling EDIP and ADIP based upon the linear mixture model. The experiments both on a synthetic and real dataset demonstrate the superiority of the proposed method over the state-of-the-art including UnDIP [14] and SiVM [5] + FCLS [7].

## 6. REFERENCES


[1] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 44–57, Aug. 2002.

[2] G. Camps-Valls, D. Tuia, L. Gómez-Chova, and S. Jiménez, *Remote Sensing Image Processing*, Morgan and Claypool, CO, USA, 2011.

[3] Neda Rohani, Emeline Pouyet, Marc Walton, Oliver Cossairt, and Aggelos K. Katsaggelos, "Pigment unmixing of hyperspectral images of paintings using deep neural networks," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 3217–3221.

[4] J. M. P. Nascimento and J. M. B. Dias, "Vertex component analysis: a fast algorithm to unmix hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 4, pp. 898–910, Apr. 2005.

[5] Rob Heylen, Dževdet Burazerovic, and Paul Scheunders, "Fully constrained least squares spectral unmixing by simplex projection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4112–4122, 2011.

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

[7] D. C. Heinz and Chein-I-Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 3, pp. 529–545, 2001.

[8] J. M. Bioucas-Dias and M. A. T. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *2010 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, 2010, pp. 1–4.

[9] Marian-Daniel Iordache, José M. Bioucas-Dias, and Antonio Plaza, "Collaborative sparse regression for hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 1, pp. 341–354, 2014.

[10] Chao Zhou and Miguel R.D. Rodrigues, "An admm based network for hyperspectral unmixing tasks," in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 1870–1874.

[11] Y. Qian, F. Xiong, Q. Qian, and J. Zhou, "Spectral mixture model inspired network architectures for hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 10, pp. 7418–7434, Oct. 2020.

[12] Y. Qu and H. Qi, "udas: An untied denoising autoencoder with sparsity for spectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 3, pp. 1698–1712, Mar. 2019.

[13] Danfeng Hong, Lianru Gao, Jing Yao, Naoto Yokoya, Jocelyn Chanussot, Uta Heiden, and Bing Zhang, "Endmember-guided unmixing network (egu-net): A general deep learning framework for self-supervised hyperspectral unmixing," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–14, 2021.

[14] Behnood Rasti, Bikram Koirala, Paul Scheunders, and Pedram Ghamisi, "Undip: Hyperspectral unmixing using deep image prior," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–15, 2021.

[15] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky, "Deep image prior," *International Journal of Computer Vision*, vol. 128, no. 7, pp. 1867–1888, Mar 2020.

[16] Jakob Sigurdsson, Magnus Orn Ulfarsson, and Johannes R. Sveinsson, "Blind hyperspectral unmixing using total variation and $\ell_q$ sparse regularization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 11, pp. 6371–6384, 2016.

[17] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," 2015.

[18] Diederik P. Kingma and Jimmy Ba, "Adam: A method for stochastic optimization," 2014.

[19] "Usgs library," https://www.usgs.gov/labs/spec-lab.