SPECTRAL UNMIXING USING DEEP CONVOLUTIONAL ENCODER-DECODER

Behnood Rasti[1], Bikram Koirala[2], Paul Scheunders[2], and Pedram Ghamisi[1]

[1] Helmholtz-Zentrum Dresden-Rossendorf (HZDR),Chemnitzer Straße 40, 09599 Freiberg, Germany.
[2] Imec-Visionlab, University of Antwerp (CDE) Universiteitsplein 1, B-2610 Antwerp, Belgium.

ABSTRACT

© 2021 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. In this paper, we introduce 'Unmixing Deep Image Prior' (UnDIP), a deep learning-based technique for the linear hyperspectral unmixing problem. The proposed method contains two steps. First, the endmembers are extracted using a geometric endmember extraction method, i.e. a simplex volume maximization in a subspace of the dataset. Then, the abundances are estimated using a deep image prior. The proposed deep image prior uses a convolutional neural network to estimate the fractional abundances, relying on the extracted endmembers and the observed hyperspectral dataset. The results show considerable improvements compared to state-of-the-art methods.

***Index Terms*—** Hyperspectral Image; unmixing; denoising; linear mixing model; low-rank model; noise reduction; abundance estimation.

1. INTRODUCTION

Hyperspectral cameras generate rich contiguous spectra that allow to distinguish materials in a scene, relying on the fact that each material has a unique spectral signature. When the field of view contains more than one material, the observed spectrum is a mixture of the spectra of the constituent materials and spectral unmixing needs to be applied. From a modeling point of view, unmixing can be divided into two main groups; linear and nonlinear unmixing. In this paper, we aim at remote sensing applications and focus on the linear hyperspectral unmixing problem.

Deep learning-based networks are state-of-the-art in machine learning and computer vision applications. Inevitably, most of the remote sensing applications involving machine learning and image processing have been inspired by deep networks [1]. Recently, a variety of deep neural networks has been proposed for hyperspectral unmixing, mainly based on the variations of deep encoder-decoder networks. The inputs to the encoder networks are the spectra and the outputs are the abundances. The abundances are then decoded to the spectra again using linear layers with the endmembers as the weights. SNSA [2] and uDAS [3] are examples of such unmixing techniques. In these methods, the spatial information is ignored. More recently, spatial information has been exploited in deep learning unmixing algorithms, by adding convolutional layers to the encoder-decoder architecture [4]. First, the image was spatially divided into patches and then the convolution is applied on small patches of spectra. This was found useful for endmember estimation, since it supports the idea of endmember bundles that account for spectral variability. However, it degrades (and blurs) the estimated abundances [4], since small patches do not contain enough structure for the convolutions (filters) to perform better than a mere mean filter.

In this paper, we propose a deep convolutional neural network unmixing method: 'Unmixing Deep Image Prior' (UnDIP). UnDIP utilizes a geometrical approach for endmember extraction, and a new unmixing deep image prior with a deep convolutional neural network for the abundance estimation. UnDIP is designed to solve the abundance estimation as a regularized inverse problem, in which the regularizer is implicitly incorporated in the cost function. This controls overfitting of the fidelity term and makes the method robust to noise. Unlike the patchwise approaches from the literature, UnDIP is designed in a way that the convolutional network is applied globally on the entire spatial domain of an image, which leads to abundances as sharp as the ground truth, and enhances the robustness to noise.

1.1. Hyperspectral Modeling

We assume a linear model for HSI:

$$\mathbf{Y} = \mathbf{X} + \mathbf{N}, \tag{1}$$

where $\mathbf{Y} \in \mathbb{R}^{p \times n}$ is the observed HSI with n pixels and p bands, $\mathbf{X} \in \mathbb{R}^{p \times n}$ is the unknown image to be estimated, and $\mathbf{N} \in \mathbb{R}^{p \times n}$ is the model error, including noise. In spectral unmixing, we assume that:

$$\mathbf{Y} = \mathbf{E}\mathbf{A} + \mathbf{N}, \tag{2}$$

This research was funded by Alexander von Humboldt foundation and BELSPO (Belgian Science Policy Office) in the frame of the STEREO III programme – project GEOMIX (SR/06/357).

where $\mathbf{E} \in \mathbb{R}^{p\times r}$and $\mathbf{A} \in \mathbb{R}^{r\times n}$, $r \ll p$, contain r endmembers and their fractional abundances, respectively.

1.2. Endmember Extraction

In this paper, we use an algorithm, called simplex volume maximization (SiVM) [5], to extract the endmembers from the dataset. SiVM selects the endmembers by iteratively maximizing the simplex volume of the data:

$$\arg\max_{\mathbf{E}} \sqrt{\frac{(-1)^r \cdot \text{ cmd } (\mathbf{E})}{2^{r-1}(r-1)!}}, \tag{3}$$

where cmd is the Cayley–Menger determinant:

$$\text{cmd } (\mathbf{E}) = \det \begin{bmatrix} 0 & 1 & 1 & 1 & \dots & 1 \\ 1 & 0 & d^2_{1,2} & d^2_{1,3} & \dots & d^2_{1,r} \\ 1 & d^2_{2,1} & 0 & d^2_{2,3} & \dots & d^2_{2,r} \\ 1 & d^2_{3,1} & d^2_{3,2} & 0 & \dots & d^2_{3,r} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & d^2_{r,1} & d^2_{r,2} & d^2_{r,3} & \dots & 0 \end{bmatrix},$$

$d^2_{i,j}$ is the Euclidean distance between endmembers $\mathbf{e}_i$ and $\mathbf{e}_j$. Since (3) does not take into account nuisances such as noise, we first project the data on the subspace obtained by the spectral eigenvectors of a singular value decomposition.

1.3. Abundance Estimation

In this subsection, we propose a deep image prior [6] to solve the unmixing problem. The widely used classical method to estimate the abundances is to solve the optimization problem:

$$\hat{\mathbf{A}} = \arg\min_{\mathbf{A}} \frac{1}{2}||\mathbf{Y} - \mathbf{E}\mathbf{A}||^2_F \quad \text{s.t. } \mathbf{A} \geq 0, \mathbf{1}_r^T\mathbf{A} = \mathbf{1}_n^T, \tag{4}$$

i.e., the fully constrained least squares unmixing (FCLSU) [7], including both the abundance sum-to-one constraint (ASC) and abundance non-negativity constraint (ANC). Generally, regularized (or penalized) least squares techniques provide a better estimation of the abundances by taking into account prior knowledge of the data:

$$\hat{\mathbf{A}} = \arg\min_{\mathbf{A}} \frac{1}{2}||\mathbf{Y}-\mathbf{E}\mathbf{A}||^2_F+\lambda R(\mathbf{A})\text{s.t.}\mathbf{A} \geq 0, \mathbf{1}_r^T\mathbf{A} = \mathbf{1}_n^T, \tag{5}$$

where $R(\mathbf{A})$ is the regularizer or penalty term and λ is the regularization parameter. The choice of R is dependent on the available prior knowledge which can vary considerably in remote sensing images. To avoid having to make this choice, the regularizer can be implicitly substituted by a deep network and the problem is transformed into the optimization of the network's parameters:

$$\hat{\theta} = \arg\min_{\theta} \frac{1}{2}||\mathbf{Y} - \mathbf{E}f_\theta(\mathbf{Z})||^2_F \quad \text{s.t. } \hat{\mathbf{A}} = f_{\hat{\theta}}(\mathbf{Z}). \tag{6}$$

Problem (6) can be solved using a deep network. The only issue left to solve is to enforce the ASC and ANC constraints. This can be done by using a softmax function in the final layer of the network. As a result, the unmixing problem (4) can be solved. The convolutional encoder-decoder network (i.e., f_θ) used for UnDIP is shown in Fig. 1. The hyperparameters of the network are listed in Table 1.

Table 1: Hyperparameters used for UnDIP in the experiments .

Hyperparameters				
	Input Ch.	Ouput Ch.	Filter Size	Stride
Conv1	r	256	3x3	2
Conv2	256	256	3x3	1
Conv3	260	256	3x3	1
Conv4	256	256	1x1	1
Conv5	256	r	1x1	1
ConvSkip	r	4	1x1	1
	Negative Slope			
Leaky ReLU	0.1			
	Scale Factor		Mode	
Upsample	2		Bilinear	
	Type	Learning Rate		Iterations
Optimizer	Adam	0.001		3000

2. EXPERIMENTAL SETUP

2.1. The Data

The Samson hyperspectral dataset contains 95×95 pixels. The spectral signatures contain 156 bands in the wavelength range [401-889] nm. There are three main materials: soil, tree, and water. The endmembers were extracted using SiVM and the ground truth fractional abundances were generated using FCLSU.

2.2. Experimental Setup

Seven unmixing techniques from different categories were used as competing methods in the experiments: FCLSU [7], applying the ground truth endmembers; NMF-QMV, a blind unmixing method [8] that estimates the endmembers along with the abundances; Collab, a sparse unmixing method [9] that uses an overcomplete set of endmembers, obtained by Vertex Component Analysis to estimate the abundances and uDAS [3] and SNSA [2], 2 deep unmixing methods. All the parameters for the competing methods were selected according to the reported default values.

HSIs generally contain different levels and types of noise [10]. It has been shown that hyperspectral unmixing techniques are often remarkably robust to noise and can be used as denoisers [11]. To compare the robustness of the techniques w.r.t. the image SNR, we added white zero-mean Gaussian noise to the data to generate the observed data $\mathbf{Y}$. Images are generated with SNR= 20, 30, 40, and 50 dB. All the experiments are repeated five times with random noise realizations. Mean results and standard deviations are shown.

For the dataset, ground truth abundance maps are available. In the experiments, the results are compared based on

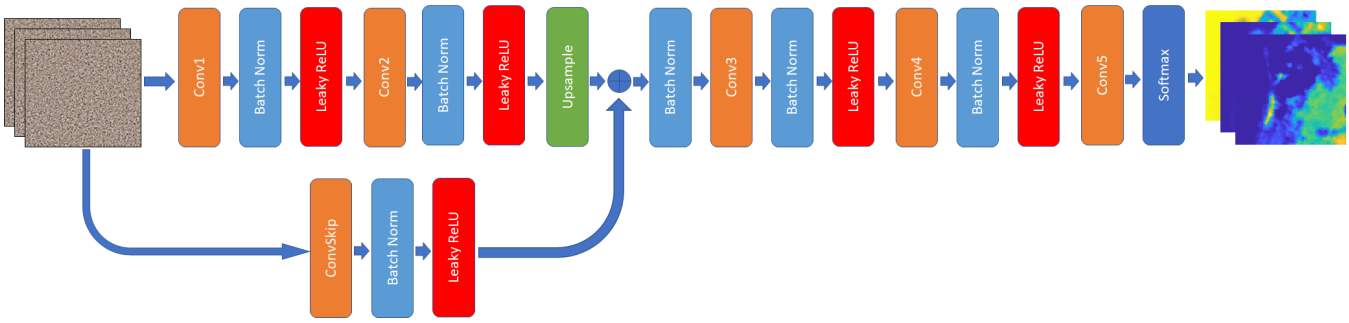

Fig. 1: The proposed convolutional network architecture with one skip connection. This network is used as f_θ for UnDIP in the experiments. Different layers in the network are shown with specific colors.

the abundance mean absolute error (MAE), the reconstruction error (RE), the spectral RMSE, and the spectral angle distance (SAD). All results, except for SAD, are reported as percentages. The abundance MAE is given by the mean of the absolute errors (in percent) between the estimated abundances and the ground truth abundances:

$$\text{Abundance MAE} = \frac{1}{rn}\sum_{k=1}^{r}\sum_{i=1}^{n}\left|\hat{\mathbf{A}}_{ki} - \mathbf{A}_{ki}\right| \times 100, \quad (7)$$

the reconstruction error is the RMSE (in percent) between the obtained reconstructed image $\hat{\mathbf{X}}$ and the observed (noisy) image $\mathbf{Y}$:

$$\text{RE} = \sqrt{\frac{1}{pn}\sum_{j=1}^{p}\sum_{i=1}^{n}\left(\hat{\mathbf{X}}_{ji} - \mathbf{Y}_{ji}\right)^2} \times 100, \quad (8)$$

The Spectral RMSE is the RMSE (in percent) between the obtained reconstructed image $\hat{\mathbf{X}}$ and the original noise-free image $\mathbf{X}$:

$$\text{Spectral RMSE} = \sqrt{\frac{1}{pn}\sum_{j=1}^{p}\sum_{i=1}^{n}\left(\hat{\mathbf{X}}_{ji} - \mathbf{X}_{ji}\right)^2} \times 100, \quad (9)$$

SAD (in degree) is the spectral angle distance between an estimated and the ground truth endmember as:

$$\text{SAD}(\mathbf{e}_{(i)}, \hat{\mathbf{e}}_{(i)}) = \arccos\left(\frac{\langle \mathbf{e}_{(i)}, \hat{\mathbf{e}}_{(i)} \rangle}{\|\mathbf{e}_{(i)}\| \, \|\hat{\mathbf{e}}_{(i)}\|}\right)\frac{180}{\pi},$$

3. EXPERIMENTS

Fig. 2 shows the results of the unmixing experiments applied on the Samson dataset, and Fig. 3 shows the estimated abundance maps. It can be observed that FCLSU, UnDIP, and NMF-QMV obtain the best abundance estimation performances (Fig. 2 (a)) and produce similar abundance maps close to the ground truth (Fig. 3). Both UnDIP and NMF-QMV obtain a lower RE and Spectral RMSE than FCLSU. The Abundance MAE of uDAS increases with increasing noise power although the RE and Spectral RMSE remain low. One can conclude that uDAS performs better as a denoiser than as an unmixer. This is due to the denoising constraint applied in the encoder of the uDAS network. The abundance estimation performance of SNSA is moderate, and the poorest of all methods for 20 dB, which shows that it is not robust w.r.t. noise. The abundance estimation performance of collaborative unmixing is poor for all SNRs, which makes it very sensitive to noise (notice the large variance for 20 dB), as can also be observed from the abundance maps in Fig. 3. Fig. 2 (d) shows that SiVM (applied in the proposed method UnDIP) and uDAS perform better for the estimation of endmembers than the other methods. A very low SAD is obtained by collaborative unmixing for 20 dB, despite the poor abundance estimation. This can be attributed to the averaging effect of the overcomplete set of endmembers that considerably helps to decrease the SAD.

4. CONCLUSIONS

In this paper, we proposed a deep prior unmixing technique called UnDIP. UnDIP first extracts the endmembers using a geometrical simplex volume maximization technique. Relying on the extracted endmembers, UnDIP estimates the fractional abundances using a deep convolutional network. The network is inspired by the theory behind the deep image prior that implicitly induces a regularizer on the cost function via the network parameters. Experiments were carried out on the Samson dataset. Comparative assessments were performed using blind, sparse and deep unmixing methods. Experimental results based on quality metrics and visual assessment confirm that UnDIP is competitive to the other unmixing techniques.

5. REFERENCES

[1] B. Rasti, D. Hong, R. Hang, P. Ghamisi, X. Kang, J. Chanussot, and J. A. Benediktsson, "Feature extrac-

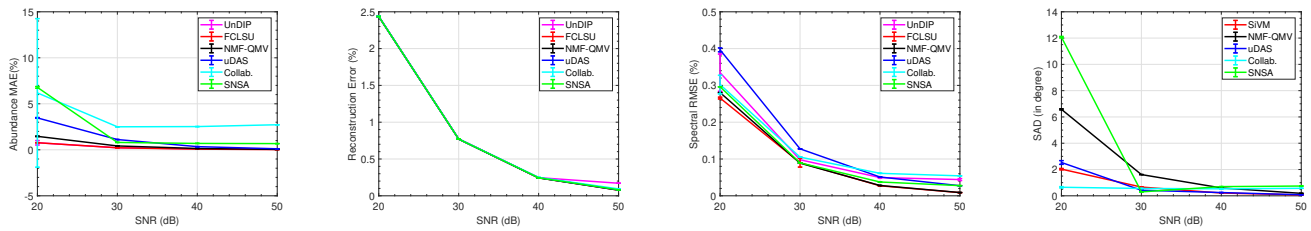

Fig. 2: Samson dataset - The results of unmixing in terms of (a) Abundance MAE, (b) Reconstruction Error, (c) Spectral RMSE, and (d) SAD (in degree) w.r.t. different noise level of the observed image (in SNR).

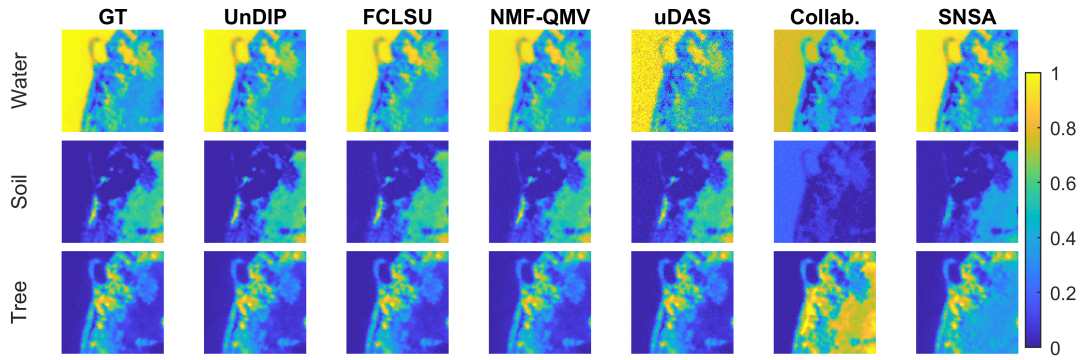

Fig. 3: Samson dataset - Abundance maps obtained by applying different unmixing techniques (20 dB).

tion for hyperspectral imagery: The evolution from shallow to deep (overview and toolbox)," *IEEE Geoscience and Remote Sensing Magazine*, pp. 0–0, 2020.

[2] Y. Su, A. Marinoni, J. Li, J. Plaza, and P. Gamba, "Stacked nonnegative sparse autoencoders for robust hyperspectral unmixing," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 9, pp. 1427–1431, 2018.

[3] 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, 2019.

[4] B. Palsson, M. O. Ulfarsson, and J. R. Sveinsson, "Convolutional autoencoder for spectral-spatial hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–15, 2020.

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

[6] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," *International Journal of Computer Vision*, vol. 128, no. 7, pp. 1867–1888, Mar 2020.

[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] L. Zhuang, C. Lin, M. A. T. Figueiredo, and J. M. Bioucas-Dias, "Regularization parameter selection in minimum volume hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 12, pp. 9858–9877, 2019.

[9] L. Drumetz, T. R. Meyer, J. Chanussot, A. L. Bertozzi, and C. Jutten, "Hyperspectral image unmixing with endmember bundles and group sparsity inducing mixed norms," *IEEE Transactions on Image Processing*, vol. 28, no. 7, pp. 3435–3450, 2019.

[10] B. Rasti, P. Scheunders, P. Ghamisi, G. Licciardi, and J. Chanussot, "Noise reduction in hyperspectral imagery: Overview and application," *Remote Sensing*, vol. 10, no. 3, pp. 482, 2018.

[11] B. Rasti, B. Koirala, P. Scheunders, and P. Ghamisi, "How hyperspectral image unmixing and denoising can boost each other," *Remote Sensing*, vol. 12, no. 11, pp. 1728, May 2020.