AgrIndia: AI-assisted Detection of chlorophyll content in leaves of food-crops using spectral signatures

Shourya Gupta
Medical Imaging and Theragnostic Lab
IIT Kharagpur
shourygupta358@gmail.com

Abstract

The AgrIndia project focuses on deployment of plant phenotyping and disease detection models using a handheld multispectral imager. In the proposed sub-project entitled "AgrIndia: AI-assisted Detection of chlorophyll content in leaves of food-crops using spectral signatures", the goal is to unmix three coloured images acquired by the imaging system and provide a chlorophyll mapped image of the region of interest. Digital imaging methods are used for mapping chlorophyll content of leaves and identifying any diseases present in crops using the multi-spectral images of leaves acquired by handheld device. Accurate measurement of chlorophyll will help in determining the nitrogen deficiency and recommend the correct usage of fertilisers.

1 Introduction

Spectral unmixing is the process by virtue of which a mixed image pixel is segregated into various constituent "endmembers," along with coefficients indicating the abundances or fractions of the endmembers present in the image pixel. Spectral unmixing generally consists of two major steps, the endmember identification and extraction and computing abundances of the endmembers pixel-wise.

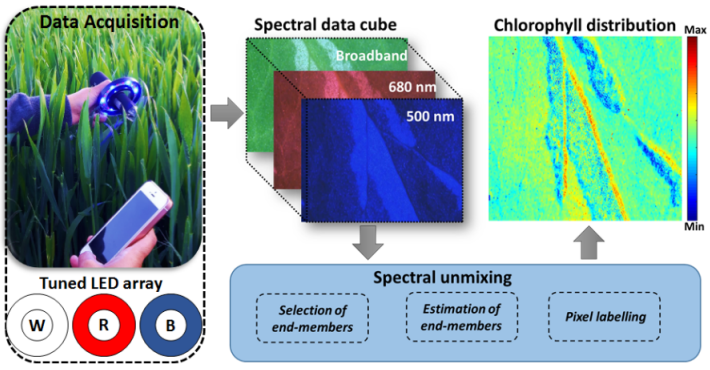


Figure 1: AgrIndia Project Pipeline

Multispectral unmixing implies application of multiple channels or spectral bands to extract and estimate endmembers from a given multispectral image. The project pipeline comprises the following steps:

1. Data Acquisition of three channels or spectral bands of the plant image at wavelengths of 680nm, 500nm and Broadband.

2. AI Algorithm to select endmembers and subsequent estimation and pixel labelling
3. Chlorophyll A and B abundance maps

Figure 1 gives a diagrammatic repersentation of the pipeline followed.

2 Literature Review

Various works have been done to solve the Spectral Unmixing (SU) Problem, which consists of linear as well as non-linear approaches. The Linear Mixing Model [1] and a lot of its variants have been employed to acquire the constituent spectra and their abundances. The LMM assumes that the observed HI pixel reflectance is obtained from a convex combination of pure material spectral signatures. The LMM is generally formulated by the following equation.

$$\mathbf{x} = \sum_{k=1}^{K} \mathbf{e}_k y_k + \eta = \mathbf{E}\mathbf{y} + \eta, \;\; s.t.\; y_k \geq 0, \; \sum_{k=1}^{K} y_k = 1 \tag{1}$$

This formula possesses an additional term η, which is the uncorrelated Gaussian noise, to account for the possible noise sources such as illumination variableness or sensor readout noise. The LMM is appropriate for a scene having segregated spatial regions of pure materials. The linear models are generally extended to geometrical volume-based algorithms to identify and quantify endmembers from hyperspectral data [2], [3], [4], [5]. The data distribution is treated as a simplex [6], and the vertices of the simplex set correspond to the endmembers. Vertex Component Analysis (VCA)[7] and similar methods [8] can extend the assumption by projecting and imposing an orthogonality condition onto endmembers in their estimations. However, these methods have several drawbacks [9].

The LMM and its variants work well as long as the endmembers do not interact and are spatially segregated [10]. Bilinear models for spectral unmixing assume that the observed spectra are structured as a bilinear combination of the pure endmember spectra present. Bilinear models are examples of nonlinear models, that is when the mixing process is nonlinear. Similarly, Multiple Endmember Bilinear Models also exist, which extend the mixing equation used in the Bilinear Models to more than two endmembers [11]. Bilinear models and their extensions possess several drawbacks as well. Since a bilinear term involves the multiplication of two endmember spectra that are smaller than one, hence the overall magnitude of the bilinear term will be lesser than the magnitude of the actual endmembers. Though the Linear Models are simplistic in nature, they do not take into account various physical phenomena such as Multiple Scattering Effects [12] and water-absorbing environments . Some approaches, such as nonlinear projection of data, such as [13], help to deal with such situations. However, the computational time required in these is very high.

Machine Learning and related fields, such as deep learning (DL), have been chosen to solve the problem of Spectral Unmixing [14]. Unsupervised methods such as Autoencoder-based [15] and belief networks [16] have been employed recently. However, their Fully Connected architecture induces a feature loss problem. Ramamurthy et al. [17] applied autoencoders and CNNs to perform image denoising and dimensionality reduction. The blind endmember extraction method based on neural networks [18] introduces a layer of components to the encoder layer of an encoder. [19] inculcates a variational AE that is capable of better inferring components and fractions from mixed materials.

CNNs have also been used for efficient feature extraction. Bera and Srivastava [20] applied CNN to analyse hyperspectral image classification by exploring different optimizers. Zhao and Du [21] explored CNNs to extract spatial information from HSI data by applying principal component analysis (PCA) for classification. The results failed to exploit the entire range of spectral information in the HSI. A morphological CNN-based architecture has been introduced by S. K. Roy et al. [22], which applies morphological operations for feature extraction to preserve essential structural characteristics of the image.

Apart from architecture, ANNs have also been in use. The ANN-based approach presented in [23] is based on non-linear PCA. It constructs an auto-associate neural network, where the output should be equal to the input but with a bottleneck layer. Zhong et al. [24] introduced specialized residual block networks to extract features from HI.

A novel architecture, an RBFNN[25], was proposed for non-linear spectral unmixing. Two sets of training data were obtained, one from the LMM model and the other from the Hapke Mixing equation. The network predicted the abundance values of the vectors, and further hyperparameter tuning led to further optimization. This showed better results than LMM on nonlinear data.

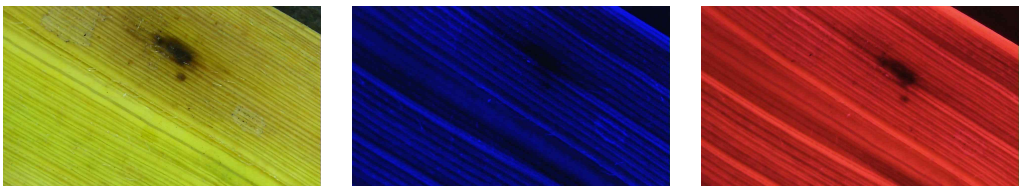
[26] and [27] showed the application of SVMs for fully constrained SVMs. However, this approach had drawbacks due to the class-specific requirements in SVMs. Various methods have combined SVMs with physical techniques, such as spectral band weighting techniques [28]. The motive behind these extended SVMs is to incorporate the information in each spectral band into a kernel function so as to give more importance to distinctive bands having higher essence.

Various novel deep learning networks have been proposed for the HSU problem. The authors in [28] introduce a general deep learning approach for hyperspectral unmixing by utilising the spectral properties of endmembers extracted from the hyperspectral imagery, called endmember-guided unmixing network (EGU-Net). The EGU-Net is a two-stream Siamese deep network, which features learning of an additional network from pure endmembers. This network shares network parameters with the autoencoder , hence progressing towards an accurate solution. The model is applicable to pixel-wise spectral unmixing as well as spatial information modelling.

A Deep Generative Endmember Model has been proposed in [29], which solves the problem of spectral variability of endmembers in HSI images. A deep generative EM model, estimated using a variational autoencoder (VAE) learns spectral variability directly from the observed data. The encoder and decoder which comprise the model are trained on pure spectral information which is extracted from the given HSI, hence allowing blind unsupervised unmixing.

3 Experimental Setup

The provided dataset consists of about 220 Multispectral Images of leaves of plants from the Poaceae family (rice, wheat, maize etc) scanned with the developed handheld device. A Multispectral plant leaf image or a Hypercube comprises three channels or spectral bands, at the wavelengths of 680 nm, 500 nm and the Broadband light. Hence, an image is of dimension $rows * columns * (numberofbands)$. Figure 2 displays the different bands for a multispectral leaf image.



Broandband 500nm 680nm

Figure 2: Figure showing the three spectral bands of the multispectral plant image

For blind spectral unmixing, each spectral Hypercube is generally considered as the dataset used by the network for training. This implies that the number of pixels in the input image becomes critical. Based on the exhaustive literature survey done for deep learning approaches to Spectral Unmixing, the images are generally resized to 200 rows by 200 columns, that is, 40,000 pixels. Here, each pixel is a mixed pixel from which endmember abundance fractions have to be estimated.

4 Autoencoder Based Network

An AutoEncoder based Network, as proposed in [30] has been followed to for unmixing Chlorophyll A and B signatures.

$X \in \mathbb{R}^{B \times P}$ denotes the input HSI having P pixels and B spectral bands. x_p is used for the observed mixed spectra of pixel number m, and $\hat{x}_p$ is used for the reconstructed mixed spectra of the pixel p, where p varies from 1 to P. The number of endmembers to be estimated is denoted by R. Each endmember is denoted by m_r, r varying from 1 to R. The endmember matrix is denoted as $M \in \mathbb{R}^{B \times R}$. The abundance array or vector for pixel p is represented by $sp = [s_{1,n}, \ldots, s_{R,p}]^\top$, and the matrix having all the abundance arrays as columns is represented by $S \in \mathbb{R}^{R \times P}$.

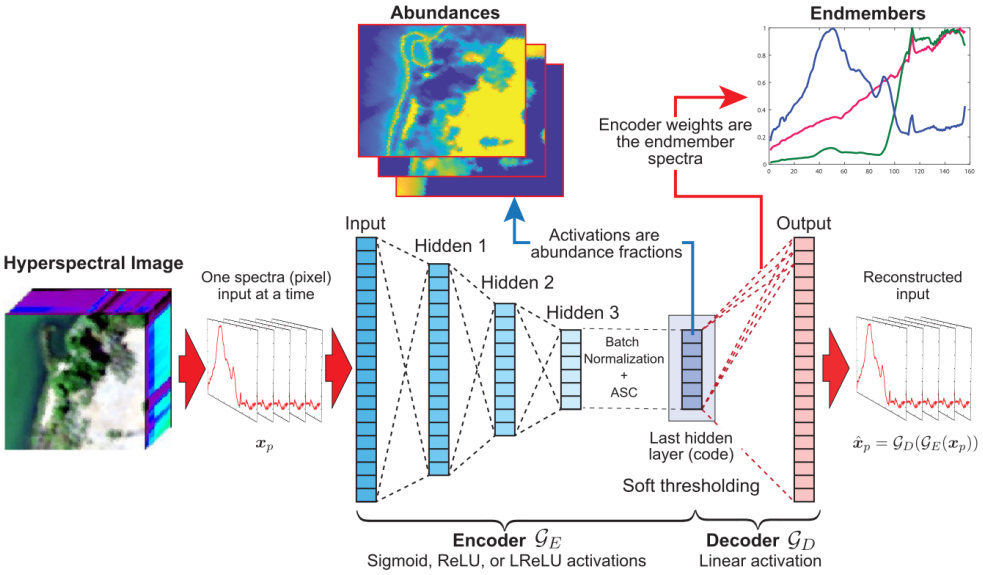


Figure 3: Autoencoder architecture described in [30]

4.1 Objective Functions

As proposed in the base paper, the MSE, SID and SAD measures as objective functions for the neural network. The MSE objective function is given by:

$$J_{\text{MSE}} = \frac{1}{P}\sum_{p=1}^{P} \|\boldsymbol{x}_p - \hat{\boldsymbol{x}}_p\|_2^2. \tag{2}$$

where x_p is the true spectrum and $\hat{x}_p$ is the estimated spectrum for pixel p. $||\cdot||_2$ denotes the Euclidean norm of a vector.

The SAD and measures the angle between spectra in their signal space in radians. It is given by:

$$J_{\text{SAD}} = \frac{1}{P}\sum_{p=1}^{P} \arccos\left(\frac{\langle \boldsymbol{x}_p, \hat{\boldsymbol{x}}_p\rangle}{\|\boldsymbol{x}_p\|_2 \|\hat{\boldsymbol{x}}_p\|_2}\right), \tag{3}$$

where x_p is the true spectrum and $\hat{x}_p$ is the estimated spectrum for pixel p.

Lastly, the SID objective function is given by:

$$J_{\text{SID}} = \frac{1}{P}\sum_{p=1}^{P}\sum_{n=1}^{B} p_n \log\left(\frac{p_n}{q_n}\right) + \sum_{n=1}^{B} q_n \log\left(\frac{q_n}{p_n}\right),$$

where

$$p_n = \frac{\boldsymbol{x}_{i,n}}{\sum_{k=1}^{M} \boldsymbol{x}_{i,k}}, \quad q_n = \frac{\hat{\boldsymbol{x}}_{i,n}}{\sum_{k=1}^{M} \hat{\boldsymbol{x}}_{i,k}}, \tag{4}$$

are estimates of the probability mass functions of the target and estimated spectra, respectively

The SID and SAD objective functions differ from the MSE Loss function based on the aspect of scale invariance of the two , unlike MSE. The shortcoming of using SID and SAD objective functions potentially leads to incorrect estimation of endmember signatures, however, the ASC on the abundances reinforces the network towards correct endmember estimation.

5 Important Contributions

1. The exhaustive literature review showed that unmixing tasks generally require spectral bands ranging from 100-250, along with simulation of artificial data. However, for the given problem statement , only 3 spectral bands are available. To overcome this drawback, data simulation approaches were explored in order to create multispectral plant images similar to the given dataset.
   Specifically, GAN based approaches were explored to simulate data [31]. This approach is still a work in progress.

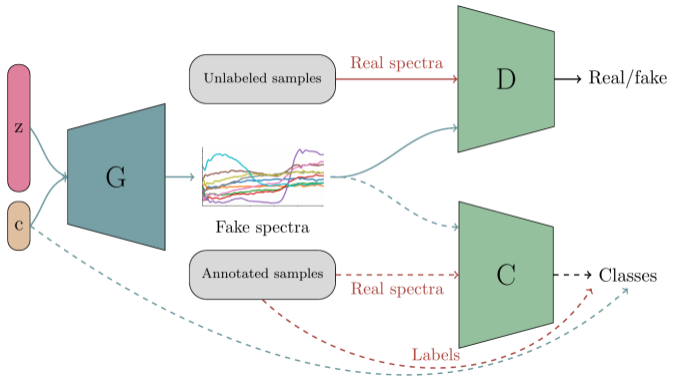


Figure 4: GAN architecture described in [31]

2. A novel objective function has been suggested, in order to improve the endmember estimation along with abundance mapping, called the "3-point measure". The 3-point measure is given by:

$$3\_point\_loss = \frac{1}{MN}\text{tr}\left[(Y-\hat{Y})^T(Y-\hat{Y})\right] \tag{5}$$

   In this equation, Y is the tensor containing true spectral absorbance values of the Chlorophyll A and B endmembers , $\hat{Y}$ represents the estimated values, and M and N are the number of spectral bands and number of endmembers respectively. For the give problem, M is 3 and N is 2.
   tr represents the trace operator, $(Y-\hat{Y})^T$ represents the transpose of the difference between Y and $\hat{Y}$, and $(Y-\hat{Y})$ represents the difference between the tensors Y and $\hat{Y}$. Finally, the result is divided by MN, which is the total number of elements in the tensors Y and $\hat{Y}$.
   The idea of the 3 point measure is simple: take spectral absorbance values of ChlA and ChlB at the given three wavelenghts and compare them with the estimated spectral absorbance values by the network. Figure 5 and 6 display the Chlorophyll A and B spectral signatures.
   The results were tested by adding the 3-point loss function to the reconstruction loss (SID/SAD). Based on the experimental results during training, adding weights to the two loss functions has been suggested:

$$\text{overall\_loss} = w_1 * \text{SID/SAD}\_loss + w_2 * \text{3\_point\_loss} \tag{6}$$

   However, using this loss function did not give optimal results with the current base network. Thus, it is suggested to include a sub-architecture into the base autoencoder in order to account for the 3-point measure to obtain accurate endmember signatures.

3. A tradeoff was observed between the accuracy of abundance maps and endmember signatures obtained from the network. In order to resolve this tradeoff, a Reinforcement Learning based approach was explored. This approach involves introducing an RL layer at the end of the network, which basically penalizes the network decoder weights by comparing with the true Chlrophyll spectra. Hence, the network receives feedback in the form of rewards or penalties for its actions and learns to adjust its behavior accordingly to maximize the $3\_point\_measure$ over time. The code for building this layer is in the works.

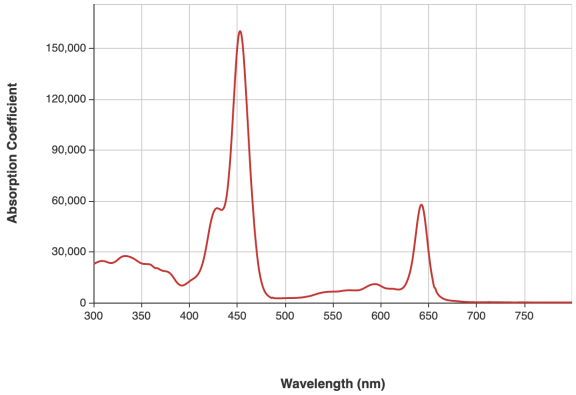


Figure 5: Chlorophyll A signatures

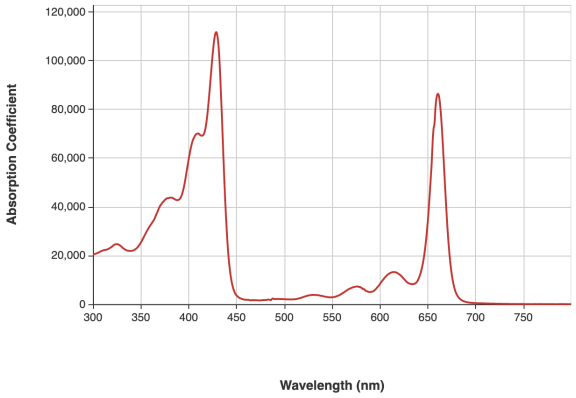


Figure 6: Chlorophyll B signatures

6 Results and Discussion

The performance of the network while training is evaluated using the MSE, SAD and SID objective functions. The function given as the argument for the network is assigned as the reconstruction loss.

The primary hyperparameters tuned for training include the number of epochs, the type of encoder(deep or shallow), learning rate, gaussian dropout, and the activation function used in the network layers.

The following table presents the best permutation of hyperparameters , obtained after procedural hyperparameter tuning.

Hyperparameters	**Value**
Epochs	250
Encoder Type	Deep
Learning Rate	1e-3
Gaussian Dropout	0.2
Activation Function	Leaky-ReLU

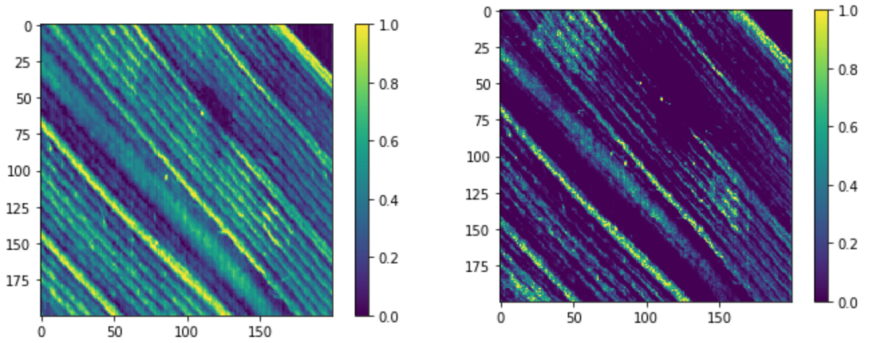
Table 1: Best hyperparameters

6.1 Comparison with Existing Blind Unmixing Models

The final abundance maps obtained for images from the network were compared with the existing blind unmixing models to measure the accuracy of the maps relative to the standardised methods. Two blind unmixing models were compared, such as UCLS (Unconstrained least-squares) [32] and FCLS(Fully-constrained least-squares) [32]. Table summarises the MSE differences between the estimated vs existing abundance maps for an image from the dataset.

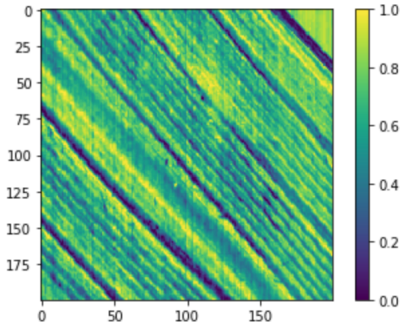
Method	Chlorophyll A Map		Chlorophyll B Map	
	MSE	MAE	MSE	MAE
UCLS	0.07774	0.24681	0.17456	0.37367
FCLS	0.14792	0.34926	0.12937	0.30205

Table 2: Comparison using MSE and MAE functions

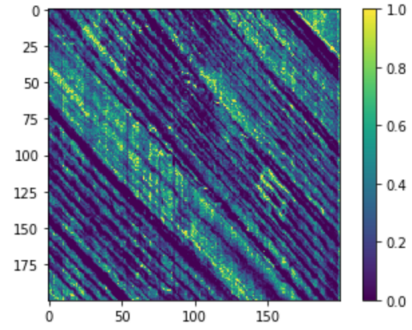


(a) UCLS Map For ChlA

(b) Predicted Map For ChlA



(c) UCLS Map For ChlB



(d) Predicted Map For ChlB

Figure 7: Abundance maps using standardized UCLS method vs estimated abundance maps for an image

7 Future Work

The future work to be carried out as part of the project involves improving the 3_point_loss. The loss function is critical, since it helps us reinforce the network to estimate accurate endmember spectra, which helps in better generalisation of the model.

Currently, to obtain optimal results, a new network has to be trained from scratch, since the generalisation capabilities of the model have a scope for improvement. This issue is proposed to be improved

by using multiple images for training a generalised model, which can estimate the chlorophyll endmembers and their abundance maps for any given multispectral plant leaf image.

8 Conclusion

To conclude, the project focussed on devising a procedural AI-based pipeline to unmix Chlorophyll A and B endmembers from plant leaf images, which were taken using a handheld device. The workflow established was aimed at improving the unmixing accuracy beyond the standardised mathematical methods such as VCA, UCLS and FCLS. AI based methods are efficient at generalisation , which helps save computational resources and ensures the scalability of the pipeline to further deploy it on hardware devices.

References

[1] Bruce Hapke. Bidirectional reflectance spectroscopy: 1. theory. *Journal of Geophysical Research: Solid Earth*, 86(B4):3039–3054, 1981.

[2] José MP Nascimento and José MB Dias. Vertex component analysis: A fast algorithm to unmix hyperspectral data. *IEEE transactions on Geoscience and Remote Sensing*, 43(4):898–910, 2005.

[3] Joseph W Boardman. Geometric mixture analysis of imaging spectrometry data. In *Proceedings of IGARSS'94-1994 IEEE International Geoscience and Remote Sensing Symposium*, volume 4, pages 2369–2371. IEEE, 1994.

[4] Joseph C Harsanyi and C-I Chang. Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach. *IEEE Transactions on geoscience and remote sensing*, 32(4):779–785, 1994.

[5] Michael E Winter. N-findr: An algorithm for fast autonomous spectral end-member determination in hyperspectral data. In *Imaging Spectrometry V*, volume 3753, pages 266–275. SPIE, 1999.

[6] Joseph W Boardman. Geometric mixture analysis of imaging spectrometry data. In *Proceedings of IGARSS'94-1994 IEEE International Geoscience and Remote Sensing Symposium*, volume 4, pages 2369–2371. IEEE, 1994.

[7] José MP Nascimento and José MB Dias. Vertex component analysis: A fast algorithm to unmix hyperspectral data. *IEEE transactions on Geoscience and Remote Sensing*, 43(4):898–910, 2005.

[8] Joseph C Harsanyi and C-I Chang. Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach. *IEEE Transactions on geoscience and remote sensing*, 32(4):779–785, 1994.

[9] Marian-Daniel Iordache, José M Bioucas-Dias, and Antonio Plaza. Sparse unmixing of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 49(6):2014–2039, 2011.

[10] Nirmal Keshava and John F Mustard. Spectral unmixing. *IEEE signal processing magazine*, 19(1):44–57, 2002.

[11] Rob Heylen, Mario Parente, and Paul Gader. A review of nonlinear hyperspectral unmixing methods. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6):1844–1868, 2014.

[12] Ben Somers, Kenneth Cools, Stephanie Delalieux, Jan Stuckens, Dimitry Van der Zande, Willem W Verstraeten, and Pol Coppin. Nonlinear hyperspectral mixture analysis for tree cover estimates in orchards. *Remote Sensing of Environment*, 113(6):1183–1193, 2009.

[13] Sam T Roweis and Lawrence K Saul. Nonlinear dimensionality reduction by locally linear embedding. *science*, 290(5500):2323–2326, 2000.

[14] Shutao Li, Weiwei Song, Leyuan Fang, Yushi Chen, Pedram Ghamisi, and Jon Atli Benediktsson. Deep learning for hyperspectral image classification: An overview. *IEEE Transactions on Geoscience and Remote Sensing*, 57(9):6690–6709, 2019.

[15] Liangpei Zhang, Lefei Zhang, and Bo Du. Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geoscience and remote sensing magazine*, 4(2):22–40, 2016.

[16] Yushi Chen, Xing Zhao, and Xiuping Jia. Spectral–spatial classification of hyperspectral data based on deep belief network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(6):2381–2392, 2015.

[17] Madhumitha Ramamurthy, Y Harold Robinson, Shanmuganathan Vimal, and Annamalai Suresh. Auto encoder based dimensionality reduction and classification using convolutional neural networks for hyperspectral images. *Microprocessors and Microsystems*, 79:103280, 2020.

[18] Frosti Palsson, Jakob Sigurdsson, Johannes R Sveinsson, and Magnus O Ulfarsson. Neural network hyperspectral unmixing with spectral information divergence objective. In *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 755–758. IEEE, 2017.

[19] Yuanchao Su, Jun Li, Antonio Plaza, Andrea Marinoni, Paolo Gamba, and Somdatta Chakravortty. Daen: Deep autoencoder networks for hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 57(7):4309–4321, 2019.

[20] Somenath Bera and Vimal K Shrivastava. Analysis of various optimizers on deep convolutional neural network model in the application of hyperspectral remote sensing image classification. *International Journal of Remote Sensing*, 41(7):2664–2683, 2020.

[21] Wenzhi Zhao and Shihong Du. Spectral–spatial feature extraction for hyperspectral image classification: A dimension reduction and deep learning approach. *IEEE Transactions on Geoscience and Remote Sensing*, 54(8):4544–4554, 2016.

[22] Swalpa Kumar Roy, Ranjan Mondal, Mercedes E. Paoletti, Juan M. Haut, and Antonio Plaza. Morphological convolutional neural networks for hyperspectral image classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14:8689–8702, 2021.

[23] Giorgio A Licciardi and Fabio Del Frate. Pixel unmixing in hyperspectral data by means of neural networks. *IEEE transactions on Geoscience and remote sensing*, 49(11):4163–4172, 2011.

[24] Zilong Zhong, Jonathan Li, Zhiming Luo, and Michael Chapman. Spectral–spatial residual network for hyperspectral image classification: A 3-d deep learning framework. *IEEE Transactions on Geoscience and Remote Sensing*, 56(2):847–858, 2017.

[25] Kerri J Guilfoyle, Mark L Althouse, and Chein-I Chang. A quantitative and comparative analysis of linear and nonlinear spectral mixture models using radial basis function neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 39(10):2314–2318, 2001.

[26] Martin Brown, Steve R Gunn, and Hugh G Lewis. Support vector machines for optimal classification and spectral unmixing. *Ecological Modelling*, 120(2-3):167–179, 1999.

[27] Martin Brown, Hugh G Lewis, and Steve R Gunn. Linear spectral mixture models and support vector machines for remote sensing. *IEEE Transactions on geoscience and remote sensing*, 38(5):2346–2360, 2000.

[28] Liguo Wang and Xiuping Jia. Integration of soft and hard classifications using extended support vector machines. *IEEE Geoscience and Remote Sensing Letters*, 6(3):543–547, 2009.

[29] 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*, 33(11):6518–6531, 2021.

[30] Burkni Palsson, Jakob Sigurdsson, Johannes R Sveinsson, and Magnus O Ulfarsson. Hyperspectral unmixing using a neural network autoencoder. *IEEE Access*, 6:25646–25656, 2018.

[31] N. Audebert, B. Le Saux, and S. Lefèvre. Generative adversarial networks for realistic synthesis of hyperspectral samples. In *2018 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, July 2018.

[32] C Therien. Welcome to the pysptools documentation, 2021.