HapkeCNN: Blind Nonlinear Unmixing for Intimate Mixtures Using Hapke Model and Convolutional Neural Network

Article *in* IEEE Transactions on Geoscience and Remote Sensing · January 2022

DOI: 10.1109/TGRS.2022.3202490

CITATIONS
34

READS
268

3 authors:

Behnood Rasti
Technische Universität Berlin
136 PUBLICATIONS **5,003** CITATIONS

SEE PROFILE

Bikram Koirala
University of Antwerp
41 PUBLICATIONS **579** CITATIONS

SEE PROFILE

Paul Scheunders
University of Antwerp
380 PUBLICATIONS **10,726** CITATIONS

SEE PROFILE

HapkeCNN: Blind Nonlinear Unmixing for Intimate Mixtures Using Hapke Model and Convolutional Neural Network

Behnood Rasti, *Senior Member, IEEE,* Bikram Koirala, *Member, IEEE,* Paul Scheunders, *Senior Member, IEEE,*

***Abstract*—© 2022 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.This paper proposes a blind nonlinear unmixing technique for intimate mixtures using the Hapke model and convolutional neural networks (HapkeCNN). We use the Hapke model and a fully convolutional encoder-decoder deep network for the nonlinear unmixing. Additionally, we propose a novel loss function that includes three terms; 1) a quadratic term based on the Hapke model, that captures the nonlinearity, 2) the reconstruction error of the reflectances, to ensure the fidelity of the reconstructed reflectance, and 3) a minimum volume total variation term that exploits the geometrical information to estimate the endmembers in the absence of pure pixels in the hyperspectral data. The proposed method is evaluated using two simulated and two real datasets. We compare the results of endmember and abundance estimation with a number of nonlinear, and projection-based linear unmixing techniques. The experimental results confirm that HapkeCNN considerably outperforms the state-of-the-art nonlinear approaches.The proposed method was implemented in Python (3.8) using PyTorch as the platform for the deep network and is available at: https://github.com/BehnoodRasti/HapkeCNN.**

***Index Terms*—Hyperspectral image, nonlinear unmixing, convolutional neural network, albedo, deep learning, Hapke model, endmember estimation, minimum simplex volume, blind unmixing**

I. Introduction

SPECTRAL sensors capture the spectral signature of materials in a range of wavelengths. The measured spectrum is generally a mixture of the pure spectra of the materials that are contained within a pixel. A mixing model models the observed spectral pixel as a function of the endmembers, corresponding to the pure spectra of the materials contained in that pixel, and each of the endmembers fractional abundances within that pixel's area. Unmixing is the task of estimating the fractional abundances, which often relies on the estimation or the extraction of the endmembers. The abundances are constraint to be nonnegative (abundance nonnegativity constraint (ANC)) and to sum to one (abundance sum-to-one constraint (ASC)). In blind unmixing both endmembers and abundances are estimated simultaneously. Alternatively, endmembers can be extracted before the abundance estimation using geometrical assumptions [1].

The mixing model is either linear or nonlinear. In linear unmixing, the endmembers are assumed to be linearly mixed. The linear model is valid when each light ray only interacts with one material before reaching the sensor. This is a common assumption in Earth observation applications, due to the macroscopic problems at hand. On the other hand, the linear approximation often fails in the case of intimate mixtures and when the light undergoes multiple reflections before reaching the sensor [1], [2], [3]. In that case, nonlinear models must be utilized [4], [5].

A popular group of nonlinear unmixing models are the bilinear models. In bilinear unmixing, it is assumed that the light interacts with at most two pure materials before reaching the sensor. Therefore, an additional mixing term is created by the Hadamard product between either of the endmembers. Researchers proposed a number of variations of this model, such as the Fan model [6], [7]. One disadvantage of the Fan model is that it does not perform well for linearly mixed datasets. To generalize the Fan model to be also applicable to the linear case, the polynomial post nonlinear mixing model (PPNM) ([8]), the generalized bilinear model (GBM) ([9]), and the linear-quadratic model (LQM) ([7]) have been proposed. These models contain hyperparameters to describe the trade-off between the linear and nonlinear terms. In [10], a nonlinear low-rank tensor unmixing algorithm was proposed to solve the GBM. Bilinear models have physical interpretation in some specific applications. However, they suffer from different drawbacks: they often include too many parameters and/or do not consider a shadow endmember, the interpretation of abundances is cumbersome, they often exclude self-interactions, they do not consider the reflections from objects outside of the instance field of view [7], and they are limited to secondary interactions. To explain the higher-order interactions of the incident light before reaching the sensor, several nonlinear mixing models have been developed (the multilinear mixing model (MLM), [11], the p-linear ($p > 2$) mixture model (pLMM) [12], [13], [14], etc.).

Kernel tricks have been employed for several nonlinear hyperspectral image processing and machine learning applications [15], [16], including unmixing [7]. Kernel functions represent data in a higher dimension space to linearize the

Behnood Rasti (corresponding author) is with Helmholtz-Zentrum Dresden-Rossendorf, Helmholtz Institute Freiberg for Resource Technology, Machine Learning Group, Chemnitzer Straße 40, 09599 Freiberg, Germany; b.rasti@hzdr.de

Bikram Koirala and Paul Scheunders are with Imec-Visionlab, Department of Physics, University of Antwerp (CDE) Universiteitsplein 1, B-2610 Antwerp, Belgium; Bikram.Koirala@uantwerpen.be; paul.scheunders@uantwerpen.be

problem. Therefore, linear algorithms become effective in the high-dimensional space. Kernelized fully constrained least-squares (kernel-FCLSU) was proposed in [17]. A kernel nonnegative matrix factorization (kernel-NMF) was proposed in [18] for nonlinear unmixing. The fully constraint linear spectral unmixing (FCLSU) problem can be solved using support vector machines (SVMs), if there are pure pixels present in the data that can serve as endmembers. The common support vectors are the endmembers which span the data simplex, and allow the estimation of the abundances of all the pixel enclosed within the simplex. Consequently, Kernel SVMs were proposed for nonlinear spectral unmixing [19].

Radiative transfer models are the most advanced nonlinear mixing models ([7]). These models are often utilized to reconstruct the reflectance spectrum of intimately mixed materials. In remote sensing applications, the simplified version of the Hapke model ([20], [21]) is being used to predict the composition of intimate mixtures. This model estimates the areal fractions of the materials within the mixture by transforming the reflectance spectra to their single-scattering albedos (SSA) and applying linear unmixing. We should note that the existing Radiative transfer models assume that spectral reflectance of the pure materials are available for estimating fractional abundances.

Deep learning (DL) revolutionized image processing and computer vision applications. Deep autoencoder architectures have been broadly used for linear unmixing. On the other hand, fewer studies exist that exploit DL-architectures for nonlinear unmixing, particularly blind unmixing. In [22], a long short-term memory-based autoencoder was proposed for PPNM. In [23], a blind unmixing method was proposed, using a 3-dimensional convolutional autoencoder-based network for PPNM. A supervised autoencoder was used in [24] for Fan, bilinear, and PPNM, in which radial basis function (RBF) kernels and K-means clustering were used for the estimation of the number of endmembers and the endmember spectra, respectively. Most DL-based nonlinear unmixing techniques are autoencoder-based architectures based on PPNM and, therefore suffer from the drawbacks of the bilinear models mentioned above [25], [26], [27]. A deep autoencoder was proposed in [28] where the encoder utilizes an extra nonlinear layer to model the nonlinearity in the data. A generative adversarial network (GAN) was proposed in [29] for nonlinear unmixing. The proposed method exploits a cycle-consistent loss to ensure the reconstruction in addition to two GAN losses. Deep autoencoders with multitask learning were also suggested for bilinear unmixing [30].

Unlike the DL-based nonlinear unmixing models existing in the literature, in this paper, we investigate the advantage of DL-based architectures for the radiative transfer models. We propose a blind nonlinear unmixing method using the Hapke model and convolutional neural networks (HapkeCNN) [1]. We show that, unlike the conventional radiative transfer models, HapkeCNN can simultaneously estimate both the endmembers and abundances by relying on the linear mixture assumption of the albedos. The proposed loss function exploits a total variation-based minimum simplex volume term, which helps to estimate endmembers in the absence of pure spectral pixels, using geometrical information.

[1]The code is available at: https://github.com/BehnoodRasti/HapkeCNN.

A. Contributions and Novelties

The reflectance spectra of intimate mixtures of minerals do not follow the linear mixture model. On the other hand, their single-scattering albedos are linearly mixed. A simple strategy is to compute the SSA of the reflectance and apply linear unmixing techniques to estimate the abundances. However, this has several shortcomings that our proposed model will address.

The contribution of the proposed method is threefold.

1) Most of the nonlinear unmixing techniques, in particular the radiative transfer models proposed in the literature are supervised, i.e., they estimate the abundances assuming that the endmembers are known. Additionally, applying blind linear unmixing on single-scattering albedos is not the appropriate approach because minimization of the reconstruction error in albedo space might not be optimal in the reflectance space, as the error in the albedo space will magnify after transforming back to the reflectance space. To solve this, we propose a blind nonlinear unmixing technique that estimates both abundances and endmembers in the reflectance space.
2) We propose a fully convolutional deep encoder-decoder architecture that generates the endmembers and abundances for intimate mixtures. The convolutional operator helps to capture local homogeneity between abundances in intimate mixtures.
3) We propose a novel loss function that contains three terms: a quadratic term to capture the nonlinearity of the data using the Hapke model; a reconstruction error of reflectances; a geometrical penalty term applied to the albedos for estimating the endmembers in the absence of pure spectra.

In experiments, Hapke CNN was applied to two simulated and two real datasets. The results are compared for both endmember and abundance estimation with bilinear, nonlinear, and projection-based linear unmixing techniques. The experimental results confirm substantial improvements in terms of spectral angle distance (SAD) for the endmember estimation, and root mean square error (RMSE) for the abundance estimation.

The remaining of this paper is organized as follows. The proposed unmixing methodology is explained in detail in Section II. The experimental results are shown and discussed in Section III. Section IV concludes the paper.

II. Methodology

A. Reflectance Modeling of Intimate Mixtures

We assume intimate mixtures of densely packed grains or particles, such as minerals in rocks or sands. In the intimate mixture scenario, light interacts multiple times with the particles before reaching the sensor, and each particle might randomly scatter or absorb the photons. A transformation from

reflectance to single-scattering albedo makes it only dependent on the first material that the incoming light encounters, and thus makes it proportional to the materials areal abundance in the mixture. Consequently, SSA follows the linear mixture model. The wavelength-dependent SSA of a material is defined as the fraction of the photons scattered by a particle to the total number of photons affected by that particle:

$$w = \frac{S_c}{S_c + A_c} \tag{1}$$

where S_c and A_c are the wavelength-dependent material's scattering and absorption coefficient respectively. Therefore, $0 \leq w \leq 1$. When the particles completely scatter the photons, then $w = 1$ and for pure absorption, $w = 0$.

Our main assumption is that the popular Hapke model [21], [31] is valid for a closely packed particulate medium, i.e., the particles are much larger than the light wavelength. Additionally, we assume that particles are spherical and scatter isotropically. Another assumption is that the phase angle (i.e., the angle between the incoming light and the outgoing radiation) is large enough. With these assumptions, the bidirectional reflectance of the particulate medium can be modeled by its optical and physical parameters as a function of SSA w:

$$y(w, \mu, \mu_0) = \frac{w\mu_0 H(w, \mu_0) H(w, \mu)}{4\pi(\mu + \mu_0)} \tag{2}$$

where μ_0 and μ are the cosines of the angles of the incoming and outgoing radiation, respectively, and H denotes the Chandrasekhar's isotropic scattering function. The latter can be approximated by: [21], [31]

$$H(w, \mu) = \frac{1 + 2\mu}{1 + 2\mu\sqrt{1 - w}} \tag{3}$$

When using Eq. (2), the relative bidirectional reflectance with respect to that of a pure scattering panel (i.e., $w = 1$) is given by:

$$y(w, \mu, \mu_0) = \frac{wH(w, \mu_0) H(w, \mu)}{H(1, \mu_0) H(1, \mu)} \tag{4}$$

which can be rewritten (using Eq. (3)) as:

$$y(w, \mu, \mu_0) = R(w) = \frac{w}{(1 + 2\mu\sqrt{1 - w})(1 + 2\mu_0\sqrt{1 - w})} \tag{5}$$

This equation is invertible and the SSA w can be written as a function of the reflectance:

$$w = R^{-1}(y) = 1 - \left(\frac{[(\mu + \mu_0)^2 y^2 + (1 + 4\mu\mu_0 y)(1 - y)]^{0.5} - (\mu + \mu_0)y}{1 + 4\mu\mu_0 y}\right)^2 \tag{6}$$

Note that, for the sake of the simplicity, the expressions above are all wavelength dependent scalar-valued. Therefore, function R can be element-wise applied on the entire image.

B. Hyperspectral Modeling Using Albedo

In intimate mixture scenarios, the linear mixture model fails, and therefore it is essential to utilize nonlinear models. In general, in a nonlinear model, the reflectance depends on the endmembers and abundances as:

$$\mathbf{Y} = F(\mathbf{A}, \mathbf{E}) \tag{7}$$

where F is a nonlinear function, $\mathbf{Y} \in \mathbb{R}^{p \times n}$ is the observed spectral image, containing n pixels and p spectral bands, $\mathbf{E} \in \mathbb{R}^{p \times r}$, and $\mathbf{A} \in \mathbb{R}^{r \times n}$, $r \ll p$ contain the r endmembers (reflectance) and fractional abundances, respectively.

Generally speaking, such nonlinear modeling can be obtained using radiative transfer theory [32]. In this work, we assume that Eq. (7) is described by the simplified Hapke model. This model reconstructs the spectral reflectance of an intimate mixture by nonlinearly combining endmembers and fractional abundances. The major advantage of this model is that it provides the transformation function (see Eq. (6)) that linearizes the nonlinear dataset. This linear space is often described as the single-scattering albedo space. In this space, the spectra of the pure materials ($\mathbf{W} = R^{-1}(\mathbf{E}) \in \mathbb{R}^{p \times r}$, $r \ll p$) are mixed linearly as follows:

$$\mathbf{X} = \mathbf{WA}, \quad \text{s.t.} \quad \mathbf{A} \geq 0, \mathbf{1}_r^T \mathbf{A} = \mathbf{1}_n^T, 0 \leq \mathbf{W} \leq 1 \tag{8}$$

where $\mathbf{X} = R^{-1}(\mathbf{Y}) \in \mathbb{R}^{p \times n}$ contains the albedos of the observed reflectances. $\mathbf{1}_n$ indicates an n-component column vector of ones. A drawback of (8) is that the noise distribution is not well-defined. In blind unmixing scenarios, the task is to estimate both $\mathbf{E}$ and $\mathbf{A}$ simultaneously. A possible strategy contains three steps: 1) applying Eq. (6) on the observed reflectances; 2) applying a linear blind unmixing technique to obtain $\mathbf{A}$ and the endmember albedos; 3) using Eq. (5) to convert back the endmembers albedo to the reflectance. A disadvantage of such a strategy is that the noise and model errors are affected by the nonlinear transformation and, as we will observe in experiments, downgrades the performance of the algorithms. Alternatively, we propose the following model:

$$\begin{aligned} \mathbf{Y} =& R(R^{-1}(\mathbf{E})\mathbf{A}) + \mathbf{N}, \\ & \text{s.t.} \quad \mathbf{A} \geq 0, \mathbf{1}_r^T \mathbf{A} = \mathbf{1}_n^T, 0 \leq \mathbf{E} \leq 1 \end{aligned} \tag{9}$$

where $\mathbf{N} \in \mathbb{R}^{p \times n}$ is noise. The main advantage of Eq. (9) is that it utilizes the linear mixing attribute of the albedos in the spectral domain and therefore, the noise will not be affected by the nonlinear transformation. To simultaneously estimate $\mathbf{E}$ and $\mathbf{A}$, the optimization is performed with respect to $\mathbf{E}$ and $\mathbf{A}$. However, such optimization is nontrivial and computationally expensive due to the nonlinearity of R. We propose a CNN to solve this problem.

C. HapkeCNN

Fig. 2 shows the proposed CNN used for the simultaneous estimation of $\mathbf{E}$ and $\mathbf{A}$. The deep network utilizes five well-established CNN modules. The color code defines different layers. Dark blue is Softmax and green is the Upsampler module. Orange is a Convolutional layer (Conv) and light blue is a batch normalization (BN) layer. The blocks shown in purple contain a Conv layer, a BN layer, and an activation function (shown in red). The BN layers speed up the learning process and provide more robustness for selecting the hyper-parameters. We use Leaky ReLU (rectified linear unit [33]) as the nonlinear activation function for all CNN modules. We use softmax to enforce the ASC and ANC:

$$softmax(\mathbf{A}) = \frac{e^{\mathbf{A}_{ij}}}{\sum_{i=1}^{r} e^{\mathbf{A}_{ij}}} \quad \forall i, j \tag{10}$$

The final convolutional layer contains the same number of filters as the number of endmembers. The endmembers are the weights of the final convolutional layer. The skip connection can easily learn the identity function and therefore avoids vanishing gradients in a deep network. Table I depicts the hyperparameters used for the network shown in Fig.2. The convolutional layers use 3×3 kernels, except for the one in the skip connection and the last convolutional layer, which uses a 1×1 kernel. We should note that the last convolutional layer does not use bias. We use a stride of two for the first convolutional layer, followed by an upsample layer. We set the negative slope of Leaky ReLU to 0.1 and we use bilinear interpolation for the upsampling. To preserve the spatial size of the image after the convolutions, we use reflection padding. We generate random noise with the same size as the dataset and use it for the input to train the network. Inspired by deep image prior [34], [35], the network is trained in an unsupervised manner mapping the observed dataset to the desired reconstructed data. This type of training was found to be effective for image reconstruction [34], [35] and spectral unmixing [36], [37]. We set the number of iterations to 8000 and use exponentially weighted averaging over the outputs to make the algorithm robust to the variance of the loss function at the stopping point. Adam optimizer is used with a learning rate of 0.001 to minimize the loss function. The loss contains three terms:

$$\mathcal{L}(\mathbf{Y},\hat{\mathbf{Y}},\mathbf{A},\mathbf{E}) = \frac{1}{2}||\mathbf{Y} - R(R^{-1}(\mathbf{E})\mathbf{A})||_F^2 + \frac{\alpha}{2}||\mathbf{Y}-\hat{\mathbf{Y}}||_F^2 + \lambda TV(R^{-1}(\mathbf{E})) \quad (11)$$

where $\hat{\mathbf{Y}}$ is the estimated (or reconstructed) data and $\|.\|_F$ denotes the Frobenius norm. The first term uses the albedo model of Eq. (9) to capture the nonlinearity of the data. The second term is the fidelity term, that ensures minimal reconstruction error of the reflectance. The third term is a geometrical total variation penalty which enforces the data simplex to have minimum volume [38]. The tuning parameters α and λ control the trade-off between the terms. With $\mathbf{W} = R^{-1}(\mathbf{E})$, the total variation geometrical penalty is given by [38], [2]:

$$\begin{aligned} TV(\mathbf{W}) &= \sum_{i,j=1}^{r} \frac{1}{2}||\mathbf{w}_i - \mathbf{w}_j||_2^2 = \\ &\frac{1}{2}\sum_{i,j=1}^{r} ||\mathbf{w}_i - \bar{\mathbf{w}} - \mathbf{w}_j + \bar{\mathbf{w}}||_2^2 = \\ &\frac{1}{2}\sum_{i=1}^{r}||\mathbf{w}_i - \bar{\mathbf{w}}||_2^2 + \frac{1}{2}\sum_{j=1}^{r}||\mathbf{w}_j - \bar{\mathbf{w}}||_2^2 - \\ &\left(\sum_{i=1}^{r}(\mathbf{w}_i - \bar{\mathbf{w}})\right)^T \sum_{j=1}^{r}(\mathbf{w}_j - \bar{\mathbf{w}}) \end{aligned} \quad (12)$$

where $\bar{\mathbf{w}} = \frac{1}{r}\mathbf{W}\mathbf{1}_r$ and the third term is zero. Thus:

$$\begin{aligned} TV(\mathbf{W}) &= \sum_{i=1}^{r}||\mathbf{w}_i - \bar{\mathbf{w}}||_2^2 = \\ &\sum_{i=1}^{r}||\mathbf{w}_i - \frac{1}{r}\mathbf{W}\mathbf{1}_r||_2^2 = ||\mathbf{W}(\mathbf{I}_r - \frac{1}{r}\mathbf{1}_r\mathbf{1}_r^T)||_F^2. \end{aligned} \quad (13)$$

In [39], another popular geometrical penalty to enforce a minimum volume simplex was proposed, which enforces the endmembers towards the center of the data simplex. However, we selected the total variation penalty since it does not rely on the data mean, and therefore we can directly apply it to the albedos, i.e., $R^{-1}(\mathbf{E})$. The performance of the different geometrical penalties were discussed in [2]. The geometrical performance of the TV penalty is illustrated in Fig. 1. Minimizing the total variation of the endmembers pulls the vertices towards each other and minimizes the initial data simplex. The main purpose of the geometrical penalties is to estimate the endmembers in the absence of pure pixels.

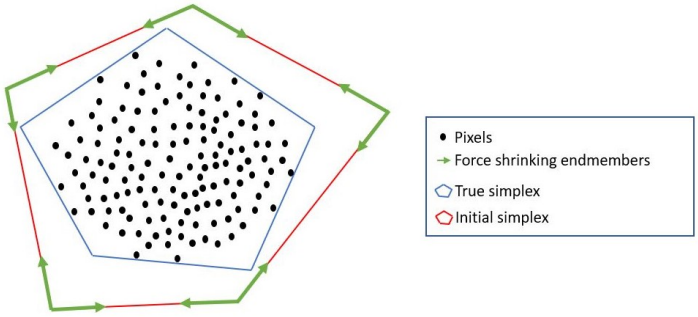


Fig. 1. A geometrical illustration of how the total variation penalty pulls the endmembers towards each other and enforces the minimum volume constraint.

TABLE I
HYPERPARAMETERS USED FOR HAPKECNN.

Hyperparameters				
	Input Ch.	Ouput Ch.	Filter Size	Stride
Conv1	p	256	3x3	2
Conv2	256	256	3x3	1
Conv3	260	256	3x3	1
Conv4	256	256	3x3	1
Conv5	256	r	3x3	1
Conv6	r	p	1x1	1
ConvSkip	p	4	1x1	1
	Negative Slope			
Leaky ReLU	0.1			
	Scale Factor		Mode	
Upsample	2		Bilinear	
	Type	Learning Rate	Iterations	
Optimizer	Adam	0.001	8000	

III. EXPERIMENTAL RESULTS

We performed experiments on two simulated, one realistic synthetic, and one real dataset. The description of the datasets is given below.

A. Hyperspectral Data Description

1) Simulated Dataset 1: A hyperspectral dataset of 105×105 pixels (see Fig. 3(a)) is simulated by combining six endmembers (Fig. 3(b)) nonlinearly. The Hapke model is used as a nonlinear mixing model. Each pixel contains reflection values for 200 different spectral bands covering the wavelength region [1000-2500] nm. In Fig. 4(a), we show a principal component analysis (PCA) reduced data manifold.

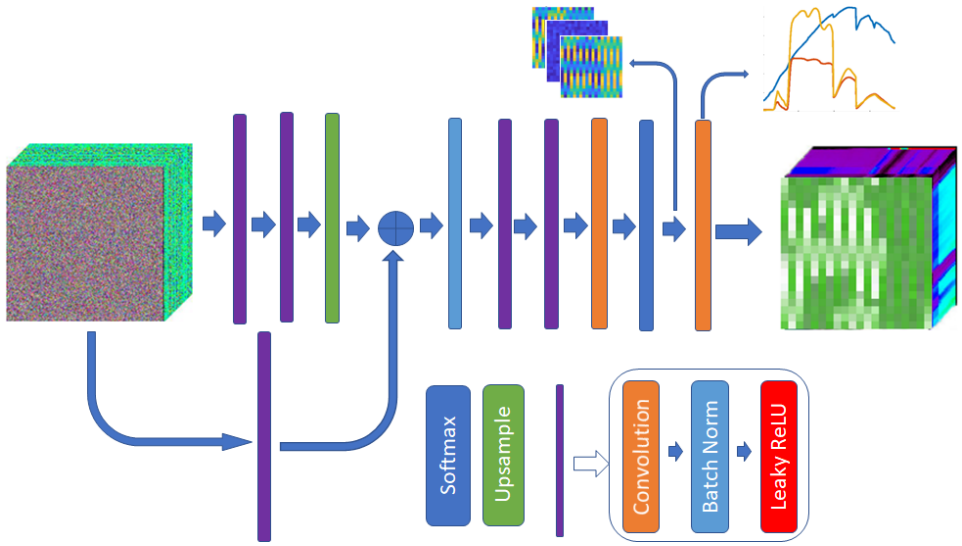


Fig. 2. The architecture of HapkeCNN. HapkeCNN uses a CNN with a skip connection and six convolutional layers. The softmax layer generates the abundances and the endmembers are given by the weights of the final convolutional layer.

We can observe that pure pixels (red circles) are available in this dataset. This makes it relatively easy to extract the endmembers by applying geometrical endmember extraction techniques.

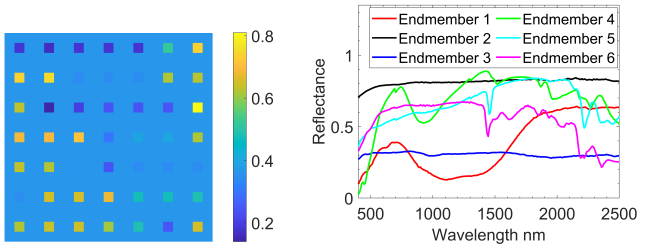


Fig. 3. Simulated Dataset 1: a) Band number 70 (1050 nm) b) Endmembers.

2) Simulated Dataset 2: By combining the same six endmembers (Fig. 3(b)) nonlinearly, another hyperspectral dataset of 105×105 pixels is simulated using the Hapke model. In comparison to the first simulated dataset, this dataset does not contain pure pixels, but at least two nonlinearly mixed data points are available on each facet of the nonlinear data manifold (see Fig. 4(b)) to geometrically reconstruct virtual endmembers.

3) Ray Tracing Vegetation Dataset: This dataset of an orchard scene ([40]) was generated by utilizing a fully calibrated virtual citrus orchard ray tracer [41]. This hyperspectral dataset (see Fig. 5 (b)) contains 20×20 pixels. The spatial resolution of this image is 2 m. There are three materials (see Fig. 5 (a) for their spectral reflectance) in this scene: soil, weed patches, and citrus trees. Each pixel contains reflection values for 216 different spectral bands covering the wavelength region [350-2500] nm. Reflection values for 31 water absorption bands (1330:1440 nm, 1770:1950 nm) were removed for further processing, and 185 bands remained. For each hyperspectral pixel, exact fractional abundance values are available. Although the interaction of light with vegetated areas is usually described

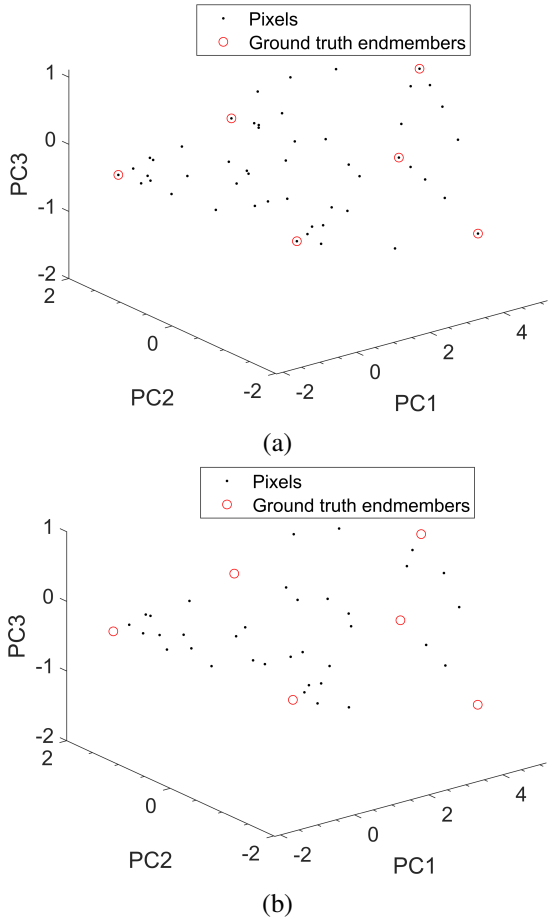


Fig. 4. PCA reduced data manifolds: a) Simulated dataset 1: there are pure pixels for each endmember. b) Simulated dataset 2: there are no pure pixels, but at least two nonlinearly mixed data points on each facet of the nonlinear simplex are available.

by multilayer mixing models, particularly for the mixture of trees and grass or trees and soil, we observed for the Ray tracing dataset that the performance of the (supervised) Hapke model is much better than competing (supervised) bilinear or multilinear mixing models. This could be explained by the fact that the mixture of weed and soil seems to behave similarly as an intimate mixture.

The main reason for selecting this dataset is that the Hapke model performs well on this dataset (abundance error of less than 10%). This might be because the mixture of weed and soil behaves as an intimate mixture.

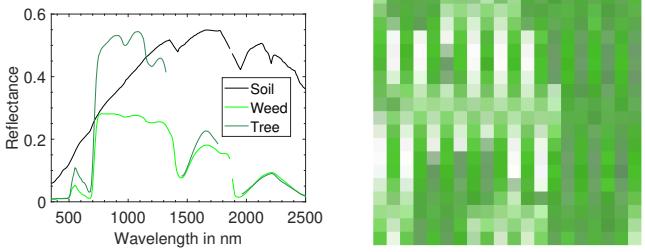


Fig. 5. Ray tracing Dataset: a) Ground truth endmember spectra of soil, weed, and citrus tree. b) True-color image (red: 650 nm, green: 550 nm, and blue: 450 nm).

4) RELAB dataset: This hyperspectral image of 60×60 pixels (see Fig. 6 (a)) is generated by utilizing the real spectral reflectances of mineral mixtures. In total, there are nine unique spectral reflectances in this image. These nine spectra were acquired in the NASA Reflectance Experiment Laboratory (RELAB) at Brown University ([42]) by mixing three minerals (Anorthite (An), Bronzite (Br) and Olivine (Ol)). The spectral reflectances of these three minerals are shown in Fig. 6 (b). Each spectral reflectance contains reflection values for 461 different spectral bands covering the wavelength region [300-2600] nm. The size of the particles (of the order of 100 μm) is much larger than the wavelength of the light. All nine mixtures are binary mixtures, i.e., mixtures of An-Br, of Br-Ol, and of An-Ol. For each mineral pair, the spectra of three mixtures (0.75/0.25, 0.5/0.5, and 0.25/0.75) are available. In the experimental part, we will demonstrate that these nine unique spectra are sufficient to geometrically reconstruct the virtual endmembers. These nine spectra were selected because the Hapke model can accurately estimate the fractional abundances of these mixtures. We should note that, although the mixed spectral pixels are real, the abundance maps are purely synthetic. We utilized ground truth fractional abundances of nine mixtures to create homogeneous abundance patches, i.e., there is no variation in the fractional abundances within a patch. However, there are 13 homogeneous patches (see Fig. 6(a)), and the ground truth fractional abundances of more than one patch can be the same. These synthetically generated abundance maps were the basis for creating the hyperspectral image. All pixels with the same fractional abundances have the same spectral reflectance. We utilized the spectral reflectance of one mixture as the background of this hyperspectral image.

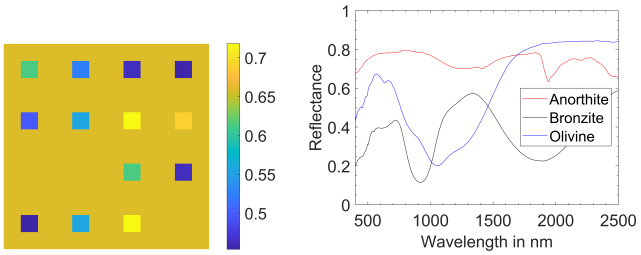


Fig. 6. RELAB Dataset 1: a) Band number 250 (1545 nm) b) Endmembers.

B. Experimental Setup

In experiments, we evaluate HapkeCNN and compare it with several unmixing techniques, both in unsupervised (endmembers and abundances are estimated simultaneously) and supervised (endmembers are known a priori) scenario's. The unmixing methods in our experiments can be divided into four categories: 1- Blind Linear Unmixing (NMF-QMV), 2-(endmember) supervised linear unmixing (FCLSU and UnDIP), 3- (endmember) supervised nonlinear unmixing (PPNM, MLM, and LR-NTF), and 4- Blind nonlinear unmixing (MAC-U and HapkeCNN). The linear mixing models (groups 1 and 2) cannot be used in the reflectance space. The errors will be high. Therefore, endmembers are extracted in the single scattering albedo space, and transformed back into the reflectance space. The supervised nonlinear models (group 3) do not estimate endmembers. Therefore, we considered two scenarios: 1- a supervised scenario in which we provided the ground truth endmembers 2- we used endmember extraction techniques in the albedo space for fair comparisons with the other methods. The last group (4) estimates the spectral reflectance of the endmembers. Here, we briefly discuss the unmixing algorithms used in the experiments.

- Geometrical unmixing: Both endmember extraction and abundance estimation are performed in the single scattering albedo space (see Eq. (6) for the transformation function). VCA [43] is applied for endmember extraction and FCLSU [44] for abundance estimation. After this, the endmember albedos are converted back to the endmember reflectances (see equation 5 for the transformation function). In this work, we will refer to this procedure as FCLSU.
- Geometrical and blind linear unmixing: NMF-QMV [2]. Blind unmixing is performed in the single scattering albedo space (i.e., the linear space), after which the endmember albedos are converted to the endmember reflectances. In this work, we will refer to this procedure as NMF-QMV.
- Bilinear unmixing: PPNM [6]. Endmembers are obtained in the single scattering albedo space, using VCA, and transformed back into the reflectance space. Hereafter, abundances are estimated by bilinear unmixing.
- Multilinear unmixing: MLM [11]. Endmembers are obtained in the single scattering albedo space, using VCA, and transformed back into the reflectance space. Hereafter, abundances are estimated by multilinear unmixing.

- Bilinear unmixing: Nonnegative Tensor Factorization (LR-NTF) [10]. Endmembers are obtained in the single scattering albedo space, using VCA, and transformed back into the reflectance space. Hereafter, abundances are estimated by Nonnegative Tensor Factorization.
- Deep learning-based unmixing method: UnDIP [36]. Both endmember extraction and abundance estimation are performed in the single scattering albedo space. Endmembers are extracted using SiVM [45]. In this work, we will refer to this procedure as UnDIP. Hereafter, endmembers are transformed back into the reflectance space.
- Blind nonlinear unmixing: Deep Autoencoder based on PPNM (MAC-U) [25]. Both endmember extraction and abundance estimation are performed in the reflectance space. For initialing endmembers, the VCA algorithm was applied in the albedo space, and the spectrum was transformed back into the reflectance space.

The following supervised unmixing approaches have been compared to: FCLSU, PPNM, MLM, and LR-NTF. In all these techniques, we provide the ground truth endmembers. For FCLSU the unmixing is performed in the albedo space.

Note that for NMF-QMV, we choose the "TV" as the optional penalty term to have a fair comparison with HapkeCNN. For the Hapke model, we set parameters $\mu_0 = 1$ and $\mu = 1$. For HapkeCNN we select all the hyperparameters as discussed in subsection II-C, and we set $\alpha = 0.0001$ and $\lambda = 0.1$. We would like to stress that, in all the experimental results provided in this paper, we did not fine-tune the hyperparameter of HapkeCNN for a fair comparison. We choose all the parameters for the competing methods according to the reported default values.

Quantitative results are provided by the root mean squared error (RMSE) in percentage between the estimated and ground-truth abundance fractions:

$$\text{RMSE}(\hat{\mathbf{A}}, \mathbf{A}) = 100 \times \sqrt{\frac{1}{rn} \sum_{i=1}^{r} \sum_{j=1}^{n} \left(\hat{\mathbf{A}}_{ij} - \mathbf{A}_{ij} \right)^2}, \quad (14)$$

and by the spectral angle distance (SAD) in degree between the estimated and ground truth endmembers:

$$\text{SAD}(\mathbf{E}, \hat{\mathbf{E}}) = \frac{1}{r} \sum_{i=1}^{r} \arccos \left(\frac{\left\langle \mathbf{e}_{(i)}, \hat{\mathbf{e}}_{(i)} \right\rangle}{\left\| \mathbf{e}_{(i)} \right\|_2 \left\| \hat{\mathbf{e}}_{(i)} \right\|_2} \right) \frac{180}{\pi}, \quad (15)$$

where $\langle . \rangle$ denotes the inner product and $\mathbf{e}_{(i)}$ indicates the ith column of $\mathbf{E}$. We should note that we use ten runs for every experiment and report the means and standard deviations for the above metrics.

C. Unsupervised Unmixing Experiments

1) Simulated Dataset 1: Tables II and III report the abundance RMSE in percentage and SAD in degree for all unmixing methods applied to simulated dataset 1. It can be observed that HapkeCNN outperforms the other techniques except for 50dB, at which FCLSU performs the best. Both UnDIP and NMF-QMV perform similarly and outperform the other compared techniques for $20 \leq \text{SNR} \leq 40$. Recall that the method FCLSU, NMF-QMV, and UnDIP are applied in albedo space. PPNM, MLM, MAC-U and LR-NTF do not seem to capture the nonlinearity of the simulated dataset. From the tables, one can observe that the performance of the linear models i.e., NMF-QMV, UnDIP, and FCLSU deteriorates fast with increasing noise levels, which confirms our earlier statement regarding models 8 and 9 that the noise is largely affected by the nonlinear transformation of the endmembers.

Figs. 7 and 8 compare visually the obtained abundance maps and endmembers from the different unmixing techniques for the 40dB SNR case. From the abundance maps, UnDIP, NMF-QMV, and HapkeCNN perform equally well and outperform the others. However, the visual comparison of the endmembers shows the advantage of HapkeCNN compared to UnDIP and NMF-QMV, again caused by the endmember estimation in the albedo space in the latter situations, which is further discussed in subsection III-G.

2) Simulated Dataset 2: Simulated Dataset 2 does not contain pure pixels, and therefore it is harder to estimate the endmembers. It is expected that endmember extraction techniques, such as SiVM and VCA will fail. Tables IV and V show the abundance RMSE and SAD for all unsupervised unmixing methods applied to simulated dataset 2. The unmixing results confirm the superior performance of HapkeCNN for all cases. The other unmixing techniques cannot cope with the absence of pure pixels, resulting in poor results, even for high SNR. NMF-QMV is designed to cope with no-pure pixel scenarios, but the nonlinear transformation considerably weakens its performance. Figs. 9 and 10 compare the abundance maps and the estimated endmembers obtained by the different unmixing techniques on simulated dataset 2 with SNR=40 dB.

3) Ray Tracing Dataset: Tables VI and VII show the RMSE and SAD results on the Ray Tracing Dataset. Figs. 11 and 12 show the visual results for the 40 dB case. As can be seen, HapkeCNN considerably outperforms the competing techniques in terms of RMSE for all noise levels. However, in terms of SAD , a very high error is obtained, compared to NMF-QMV, MAC-U, VCA, and SiVM. This often can be observed in real data scenarios and is caused by spectral variability and scaling effects of the spectra. Scaling effects are not visible by SAD, since it scales the spectra by their norms. However, this scaling effect can considerably affect the performance of the abundance estimation, which is the main goal of unmixing. The extracted endmembers by SiVM are scaled versions of the ground truth endmembers, and therefore the obtained SAD values are lower than with HapkeCNN, while the abundance estimation by UnDIP is poor compared to HapkeCNN. The visual comparison reveals that HapkeCNN considerably outperforms the other methods for Weed's endmember and abundance estimation.

4) RELAB Dataset: The unmixing results for the RELAB dataset are given in Tables VIII and IX in terms of abundance RMSE and SAD, respectively. HapkeCNN significantly outperforms the other techniques in terms of RMSE and SAD, and NMF-QMV provides the second-best results. The other methods show poor performances. Additionally, HapkeCNN shows robustness in terms of noise power, while NMF-QMV shows poor performance in the very noisy scenario, i.e., 20

TABLE II
RMSE (SIMULATED 1). THE BEST PERFORMANCES ARE SHOWN IN BOLD.

	FCLSU	PPNM	MLM	LR-NTF	MAC-U	UnDIP	NMF-QMV	HapkeCNN
20dB	22.19±1.24	22.67±4.9	21.06±4.63	20.07±6.03	19.67±2.89	10.88±0.45	15.14±0.56	**4.83±0.09**
30dB	16.93±2.09	17.13±1.21	17.29±0.6	15.92±1.22	15.71±2.47	6.06±0.43	8.74±0.24	**1.75±0.02**
40dB	11.84±1.3	16.13±1.32	15.34±0.99	15.03±0.6	15.1±2.46	2±1.87	1.26±0.11	**0.8±0.03**
50dB	**0.29±0.08**	14.31±0.03	13.95±0.04	13.29±0.04	17.74±0.15	0.55±0.05	0.34±0.04	0.63±0.02

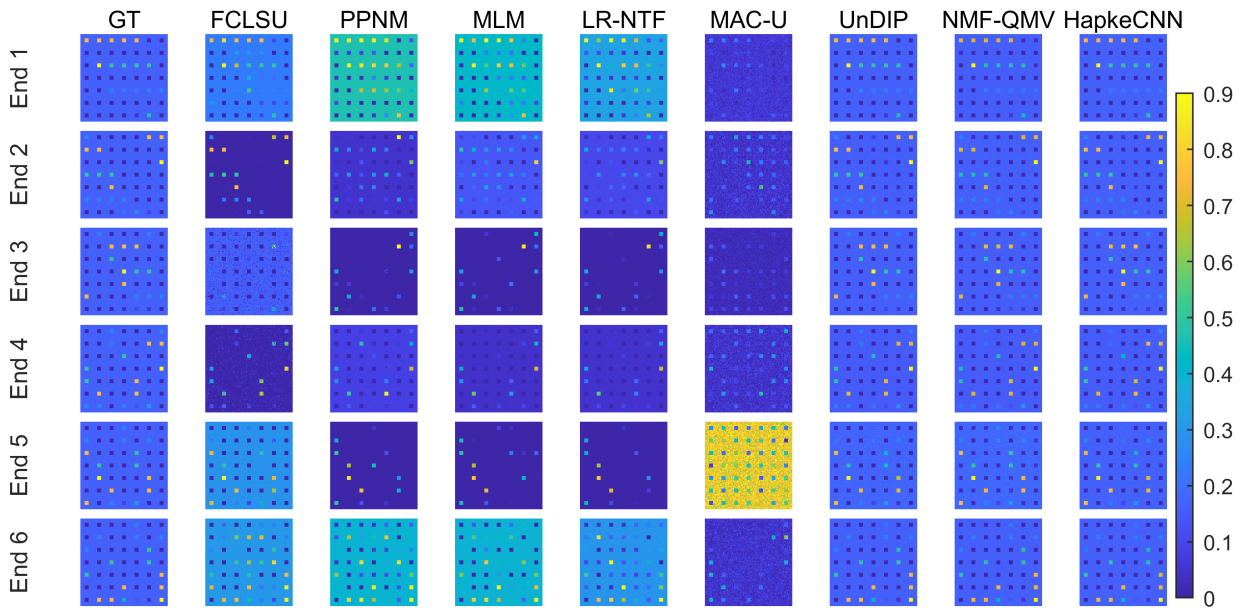


Fig. 7. Simulated dataset 1 (40dB) - Visual comparison of the abundance maps obtained by different unmixing techniques.

TABLE III
SAD (SIMULATED 1). THE BEST PERFORMANCES ARE SHOWN IN BOLD.

	SiVM	VCA	MAC-U	NMF-QMV	HapkeCNN
20dB	10.42±0.45	12.4±1.45	12.19±1.51	12.65±1.08	**1.79±0.05**
30dB	3.37±0.18	5.56±0.76	5.53±0.82	5.08±0.45	**0.48±0.02**
40dB	1.91±0.47	2.77±0.37	3±0.42	0.84±0.23	**0.39±0.02**
50dB	0.12±0.03	0.1±0.02	0.52±0.01	**0.1±0.02**	0.39±0.01

dB. Figs. 13 and 14 visually compare the performance of the unmixing technique on the RELAB dataset for 40 dB SNR. The visual comparison reveals that HapkeCNN considerably outperforms the other methods for both endmember and abundance estimation.

D. Supervised Unmixing Experiments

Table X compares the results of HapkeCNN with (endmember-) supervised techniques. As we already discussed, most nonlinear unmixing techniques assume that the endmembers are known. This considerably simplifies the problem, as can be seen in model (9). Therefore, we also compare HapkeCNN with supervised approaches. We emphasize that such comparison is not fair to our technique; however, it reveals the strength of HapkeCNN.

Table X shows the performance of HapkeCNN and the supervised techniques applied to different datasets in terms of abundance RMSE. The comparisons confirm that HapkeCNN

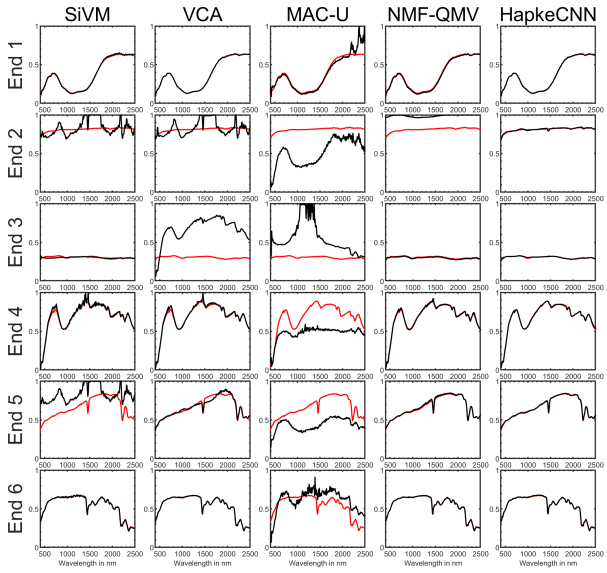


Fig. 8. Simulated dataset 1 (40dB) - Visual comparison of the endmembers obtained by different unmixing techniques. Red: ground truth endmembers; Black: estimated endmembers.

considerably outperforms the other techniques. In the case of Simulated dataset 1, Simulated dataset 2, and RELAB, HapkeCNN shows significant improvements in abundance RMSE compared to all methods. In the case of Ray tracing,

TABLE IV
RMSE (SIMULATED 2). THE BEST PERFORMANCES ARE SHOWN IN BOLD.

	FCLSU	PPNM	MLM	LR-NTF	MAC-U	UnDIP	NMF-QMV	HapkeCNN
20dB	21.57±3.07	23.62±5.05	22.71±5.83	25.07±5.48	15.3±3.73	10.08±0.65	12.54±1.4	**4.69±0.06**
30dB	18.74±1.72	20.24±1.71	19.92±2.12	20.23±1.78	13.07±1.38	10.3±0.48	7.71±1	**1.67±0.02**
40dB	14.68±1.62	18.63±2.36	17.05±2.49	18.16±2.63	14.54±1.32	8.48±1.6	3.13±1.26	**0.76±0.16**
50dB	11.53±2.26	17.9±3	16.02±2.15	17.12±4.13	13.49±1.9	7.24±0.15	3.36±1.88	**0.52±0.01**

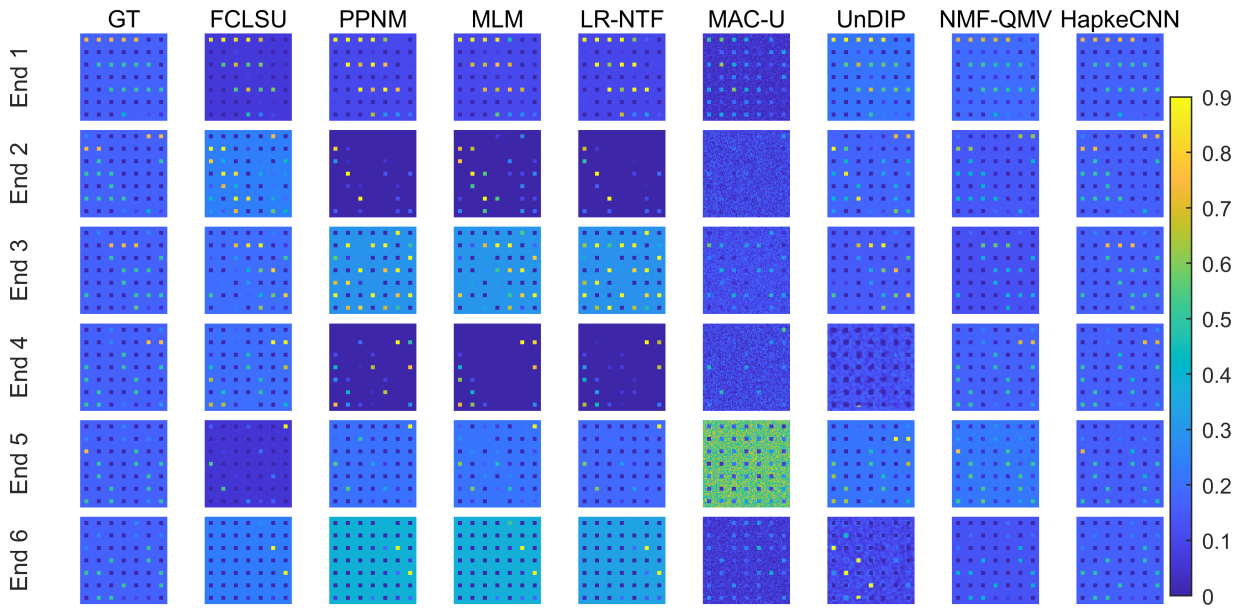


Fig. 9. Simulated dataset 2 (40dB) - Visual comparison of the abundance maps obtained by the different unmixing techniques.

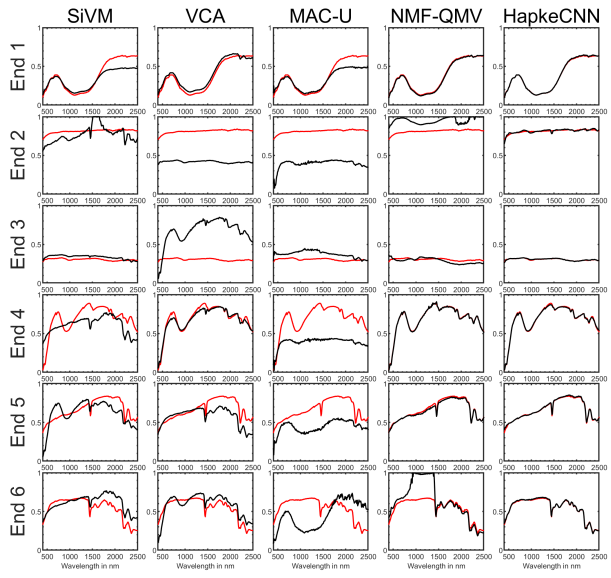


Fig. 10. Simulated dataset 2 (40dB) - Visual comparison of the endmembers obtained by the different unmixing techniques. Red: ground truth endmembers; Black: estimated endmembers.

TABLE V
SAD (SIMULATED 2). THE BEST PERFORMANCES ARE SHOWN IN BOLD.

	SiVM	VCA	MAC-U	NMF-QMV	HapkeCNN
20dB	9.49±0.27	13.26±2.44	12.49±2.11	10.85±2.08	**2.08±0.04**
30dB	6.8±0.47	9.25±1.27	9.07±1.32	4.6±1.06	**1.06±0.02**
40dB	6.49±0.32	7.2±1.58	7.22±1.55	3.14±1.51	**1.51±0.02**
50dB	5.8±0.01	5.33±1.14	5.31±1.13	3.13±1.76	**1.76±0.01**

only LR-NTF outperforms HapkeCNN for 30, 40, and 50 dB; however, the differences are not large.

E. Qualitative Analysis of Abundance Maps and Endmembers

We further applied the proposed method to a cropped image of the Cuprite dataset. This hyperspectral image contains 250×190 pixels (see Fig. 15(a) for true-color image and Fig. 15(b) for geological reference map). The reflectance values were acquired for 224 different spectral bands covering the wavelength between 389 nm to 2442 nm. After removing water absorption and noisy bands, 185 bands remain. Although in this scene, 12 materials are available, Alunite, Kaolinite1, Kaolinite2, Muscovite, Montmorillonite, Sphene, and Chalcedony are the most dominant ones. Since ground truth fractional abundances of this hyperspectral image is not available, we used supervised linear unmixing (SLU) to estimate abundances for the sake of visual comparison. We manually picked spectra of these materials from the hyperspectral image,

TABLE VI
RMSE (RAY TRACING). THE BEST PERFORMANCES ARE SHOWN IN BOLD.

	FCLSU	PPNM	MLM	LR-NTF	MAC-U	UnDIP	NMF-QMV	HapkeCNN
20dB	22.79±4.34	26.74±7.84	26.56±7.47	25.35±6.48	20.37±3.7	26.66±3.96	17.98±2.81	**13.08±0.06**
30dB	24.88±1.56	25.93±1.35	26.7±3.59	24.91±1.88	22.51±0.69	27.68±1.97	24.28±1.02	**13.31±0.73**
40dB	25.31±0.63	25.97±0.39	28.67±0.64	25.55±0.47	23.39±0.44	27.29±1.41	25.01±0.51	**13.07±0.01**
50dB	25.48±0.22	26.11±0.09	28.81±0.46	25.76±0.12	23.58±0.13	26.95±1.88	24.89±0.13	**13.07±0.01**

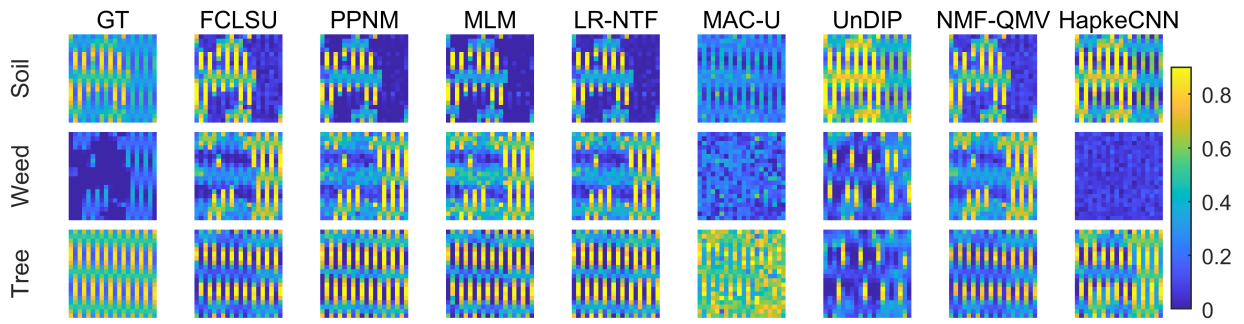


Fig. 11. Ray tracing dataset (40dB)- Visual comparison of the abundance maps obtained by the different unmixing techniques.

TABLE VII
SAD (RAY TRACING). THE BEST PERFORMANCES ARE SHOWN IN BOLD.

	SiVM	VCA	MAC-U	NMF-QMV	HapkeCNN
20dB	**6.49±0.56**	7.79±1.09	6.09±0.54	8.96±0.88	10.29±0.14
30dB	6.98±0.39	7.08±0.48	6.94±0.39	**6.61±0.34**	10.01±1.08
40dB	6.66±0.13	6.71±0.03	6.64±0.14	**6.43±0.1**	10.36±0.02
50dB	6.6±0.07	6.7±0.01	6.59±0.07	**6.35±0.04**	10.36±0.02

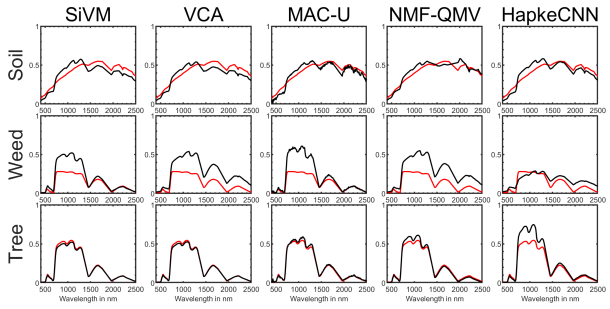


Fig. 12. Ray tracing dataset (40dB) - Visual comparison of the endmembers obtained by the different unmixing techniques. Red: ground truth endmembers; Black: estimated endmembers.

and we used FCLSU to estimate fractional abundance maps. It is interesting to observe that the abundance maps obtained by HapkeCNN (see the second row of Fig. 16) highly resembles the geological reference map (see Fig. 15(b)) and the ones obtained by SLU (see the first row of Fig. 16) .

F. Sensitivity Analysis of HapkeCNN

In this section, we investigate the sensitivity of HapkeCNN to the parameters and hyperparameters of the network. Additionally, the impact of different initializations is reported. All the experiments are carried out on the simulated dataset 2, with 30 dB SNR. The results are the mean values of five experiments, and error bars show the standard deviations.

1) Regularization Parameters: HapkeCNN uses two regularization parameters in the loss function i. e., λ and α. Figs. (18) (a) and (b) depict the performance of HapkeCNN w.r.t. λ and α, respectively, in terms of both SAD and RMSE. HapkeCNN shows the lowest SAD and RMSE for $\lambda = 10^{-1}$ and $\alpha = 10^{-4}$. $\lambda \geq 1$ and $\alpha \geq 10^{-3}$ considerably increase the errors.

2) Learning Rate: The learning rate of the adam optimizer is another hyperparameter that may affect the performance of HapkeCNN. However, it is often independent of the data and can be tuned and set after designing a network. As can be seen in Fig. (18) (c), the network provides the minimum errors for $LR = 10^{-3}$ in terms of both SAD and RMSE.

3) Number of Filters: Here, we compared the performance of HapkeCNN w.r.t. number of filters selected for the convolutional layers. The results are shown Fig. (18) (d) for 32, 64, 128, and 256 filters. It can be seen that increasing the number of filters affects the error bars and makes the method more robust to randomness.

4) Kernel Size: Fig.18(e) shows the abundance RMSE (in percent) and SAD (in degree) of HapkeCNN for different size of Kernels i.e., 1×1, 3×3, and 5×5, selected for the convolutional layers. As can be seen, a kernel size of 5×5 provides RMSE of 2% with a very high standard deviation. The kernel size of 3×3 provides lower SAD but slightly higher RMSE than 1×1 for this dataset. Fig.18(f) demonstrate the loss function value w.r.t. the number of iterations. We should note that the input of the network is random noise, and therefore the errors in the early steps are very high.

5) Endmember Initialization: In blind unmixing, the endmember initialization often plays a significant role in the final results. It is usually done by a geometrical approach, such as VCA and SiVM. Since these methods properly work in pure-pixel scenarios, they will provide endmembers in no-pure pixel scenarios that may already be close to the real ones. We evaluate the performance of HapkeCNN w.r.t. different ini-

TABLE VIII
RMSE (RELAB). THE BEST PERFORMANCES ARE SHOWN IN BOLD.

	FCLSU	PPNM	MLM	LR-NTF	MAC-U	UnDIP	NMF-QMV	HapkeCNN
20dB	20.73±1.64	21.97±6.71	20.11±8.86	20.83±7.56	21.59±5.06	29.58±4.62	20.32±1.49	**1.94±0.09**
30dB	15.77±1.08	16.89±4.21	11.91±3.67	14.85±2.17	14.13±3.04	17.62±1.79	8.86±1	**1.39±0.23**
40dB	20.41±2.03	21.04±2.55	20.03±4.5	22.68±3.71	18.35±2.21	16.92±0.21	6.72±0.38	**1.58±0.17**
50dB	20.76±2.98	20.66±3.97	19.51±6.71	24.79±4.38	17.62±3.45	16.82±0.06	6.2±0.28	**1.54±0.14**

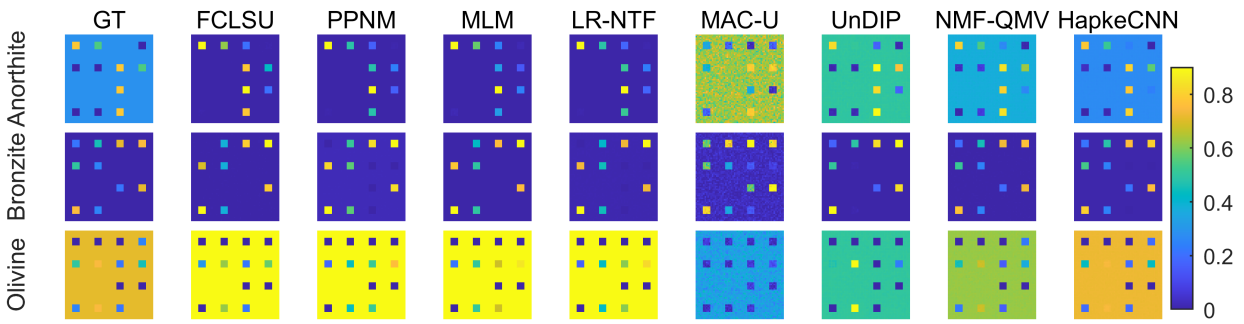


Fig. 13. Relab dataset (40dB) - Visual comparison of the abundance maps obtained by the different unmixing techniques.

TABLE IX
SAD (RELAB). THE BEST PERFORMANCES ARE SHOWN IN BOLD.

	SiVM	VCA	MAC-U	NMF-QMV	HapkeCNN
20dB	10.8±1.91	16.12±3.61	10.61±1.95	10.68±1.14	**1.67±0.1**
30dB	6.61±0.94	13.04±0.67	6.29±0.73	3.7±0.97	**1.57±0.07**
40dB	6.94±0.63	13.4±0.03	6.81±0.48	2.29±0.26	**1.61±0.05**
50dB	7.25±0.85	13.5±0.01	7.06±0.65	2.06±0.28	**1.58±0.05**

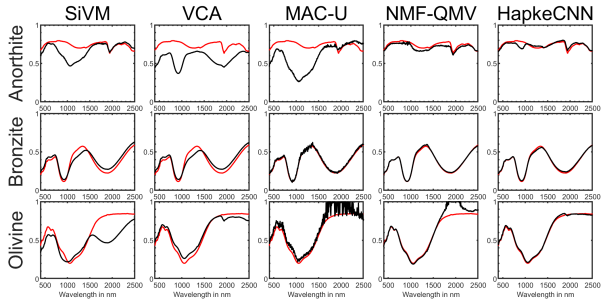


Fig. 14. RELAB dataset (40dB) - Visual comparison of the endmembers obtained by the different unmixing techniques. Red: ground truth endmembers; Black: estimated endmembers.

TABLE X
RMSE (SUPERVISED). THE BEST PERFORMANCES ARE SHOWN IN BOLD.

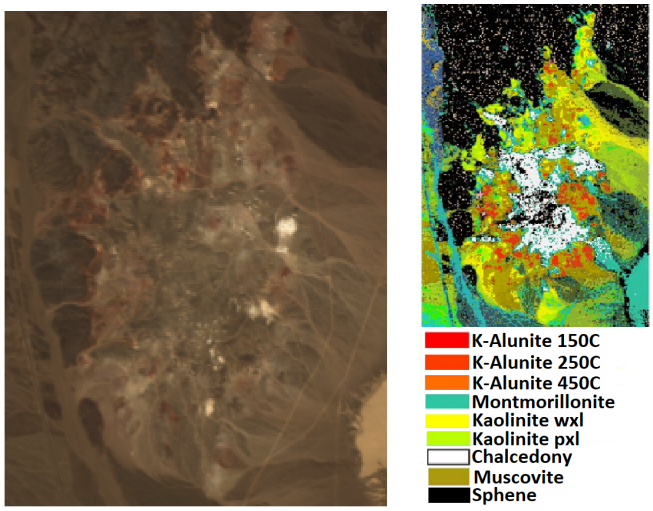
	FCLSU	PPNM	MLM	LR-NTF	HapkeCNN
Simulated dataset 1					
20dB	13.89±0.01	14.79±0.02	13.86±0.01	13.21±0.01	**4.83±0.09**
30dB	14.02±0.01	14.31±0.01	13.96±0.01	13.26±0.01	**1.75±0.02**
40dB	14.06±0.01	14.33±0.01	14±0.01	13.29±0.01	**0.8±0.03**
50dB	14.06±0.01	14.35±0.01	14.01±0.01	13.28±0.01	**0.63±0.02**
Simulated dataset 2					
20dB	13.98±0.01	14.92±0.02	13.92±0.01	13.3±0.02	**4.69±0.06**
30dB	14.1±0.01	14.46±0.01	14.03±0.01	13.35±0.01	**1.67±0.02**
40dB	14.14±0.01	14.49±0.01	14.08±0.01	13.38±0.01	**0.76±0.16**
50dB	14.14±0.01	14.5±0.01	14.08±0.01	13.37±0.01	**0.52±0.01**
Ray Tracing dataset					
20dB	17.82±0.04	17.28±0.19	17.95±0.03	13.1±0.11	**13.08±0.06**
30dB	17.81±0.02	16.29±0.05	17.93±0.02	**12.52±0.04**	13.31±0.73
40dB	17.8±0.01	16.2±0.02	17.94±0.02	**12.45±0.02**	13.07±0.01
50dB	17.8±0.01	16.21±0.01	17.94±0.01	**12.44±0.01**	13.07±0.01
RELAB dataset					
20dB	10.45±0.02	10.8±0.02	8.28±0.02	8.62±0.02	**1.94±0.09**
30dB	10.39±0.01	10.76±0.01	7.97±0.01	8.42±0.01	**1.39±0.23**
40dB	10.39±0.01	10.76±0.01	7.93±0.01	8.41±0.01	**1.58±0.17**
50dB	10.39±0.01	10.76±0.01	7.92±0.01	8.41±0.01	**1.54±0.14**

tializations, i.e., VCA, SiVM, and Random Pixels (randomly selected pixels from the observed data). Fig. 19 compares the obtained RMSE and SAD for different initialization. The comparisons confirm that SiVM provides the lowest errors. Additionally, the Random Pixels initialization leads to poor abundance estimation in terms of RMSE.

G. Discussion

The overall results confirm the significant advantage of HapkeCNN compared to the other techniques. Here, we summarize the major outcomes of the experimental results.

Most of the proposed nonlinear techniques (such as PPNM, MLM, LRNTF) in the literature are supervised (i.e., the endmembers are assumed to be known). Those techniques rely on the performance of the endmember extraction technique used, which often fails in non-pure pixel scenarios. Alternatively, linear techniques can be applied for unmixing by exploiting a nonlinear projection (such as FCLSU, NMF-QMV and UnDIP). This group suffers from translating noise and errors from one space to another. A nonlinear projection might magnify such an error, which is demonstrated in Fig. 20. The figure shows endmember six from the simulated dataset 2, estimated in albedo space and after backprojection into reflectance space. The endmember extracted/estimated in



(a) (b)

Fig. 15. Cuprite image: (a) True-color image (Red: 654 nm, Green: 550 nm, Blue: 455 nm); (b) Ground truth mineral map.

albedo space by the competing techniques closely follows the ground truth; however, after the nonlinear projection into the reflectance space, the errors considerably enlarge. This leads to misinterpretations of the materials and/or the absorption bands, which is not desirable in hyperspectral analysis.

On the other hand, HapkeCNN addresses the issues mentioned above. HapkeCNN simultaneously estimates both endmembers and abundances using a CNN architecture. The proposed loss function utilizes the hyperspectral model (9), which is based on reflectance. Therefore, the endmembers are estimated directly in the reflectance domain, exploiting the albedo domain's linearization. Moreover, HapkeCNN uses a TV-based minimum volume penalty in the loss function, addressing non-pure pixel scenarios. As a result, HapkeCNN outperforms the other nonlinear techniques even in supervised scenarios, which confirms the effectiveness of HapkeCNN and the importance of blind unmixing for nonlinear problems. Additionally, HapkeCNN provides low standard deviations even for low SNR, which confirms its robustness with respect to noise.

We should note that the main advantages of the Hapke model compared to bilinear models are that the parameter estimated by this model is physically interpretable, and that this mixing model considers higher-order interactions in intimate mixtures. Additionally, interpreting abundances using the proposed algorithm is easy since HapkeCNN estimates the endmembers in the reflectance domain. It is worth mentioning that the proposed method is extendable to other nonlinear models, as long as the nonlinear function R is known and invertible.

1) Processing Time: Table XI compares the processing time of the unmixing algorithms applied to different datasets in seconds. The reported processing times are mean values over five experiments. HapkeCNN and UnDIP were implemented in Python (3.9), and the results were obtained using a GPU (graphical processing unit) server with an NVQRTX8000 NVIDIA PNY Quadro RTX8000 (48 GB), AMD EPYC 7343 16-Core Processor (max 3.2 GHz), 128GB of memory, and a 64-bit Linux Operating System (CentOS 7). The other techniques were implemented in Matlab 2020b and ran on an Intel Core i7-8700K CPU (max 3.2 GHz) with six cores. As can be seen from Table XI, FCLSU is the fastest algorithm. However, we can see that with an increasing number of pixels (i.e., in the case of Cuprite), the conventional unmixing algorithms become less efficient compared with the DL-based algorithms, i.e., UnDIP and HapkeCNN.

TABLE XI
PROCESSING TIME (IN SECONDS) OF THE UNMIXING TECHNIQUES APPLIED TO DIFFERENT DATASETS.

	FCLSU	PPNM	MLM	LR-NTF	MAC-U	UnDIP	NMF-QMV	HapkeCNN
Ray Tracing	**0.06**	1.35	1.18	1.71	38.29	13.94	1.64	47.33
RELAB	**0.48**	10.78	6.06	12.95	126.86	10.53	12.27	43.73
Cuprite	**9.73**	347.27	305.86	1631.5	316.90	79.34	564.53	231.57

IV. Conclusion

We proposed a blind nonlinear unmixing technique for intimate mixtures using the Hapke model and CNN. HapkeCNN estimates both endmembers and abundances simultaneously by relying on the linear mixture assumption of the single-scattering albedos. The proposed loss function containing three terms: a quadratic term to capture the nonlinearity, a reconstruction term, and a penalty term to exploit the geometrical information. We used two real and two simulated datasets to evaluate the proposed method. The performance of HapkeCNN was compared with other nonlinear methods on all datasets. The experimental results showed that HapkeCNN considerably outperforms the other techniques. We also compared HapkeCNN with supervised methods for which the ground truth endmembers were provided. The experimental results demonstrated that HapkeCNN outperforms the other techniques even in supervised scenarios. Moreover, our experiments confirmed that HapkeCNN successfully estimates virtual endmembers in the absence of pure pixels, and that it is robust to noise.

Acknowledgment

The work of B. Koirala was funded by the Research Foundation-Flanders - project G031921N. The authors would like to thank Prof. Ben Somers for providing the Ray tracing dataset.

References

[1] 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 Geoscience and Remote Sensing Magazine*, vol. 1, no. 2, pp. 6–36, 2013.

[2] 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.

[3] Wei He, Hongyan Zhang, and Liangpei Zhang, "Sparsity-regularized robust non-negative matrix factorization for hyperspectral unmixing," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 9, pp. 4267–4279, 2016.

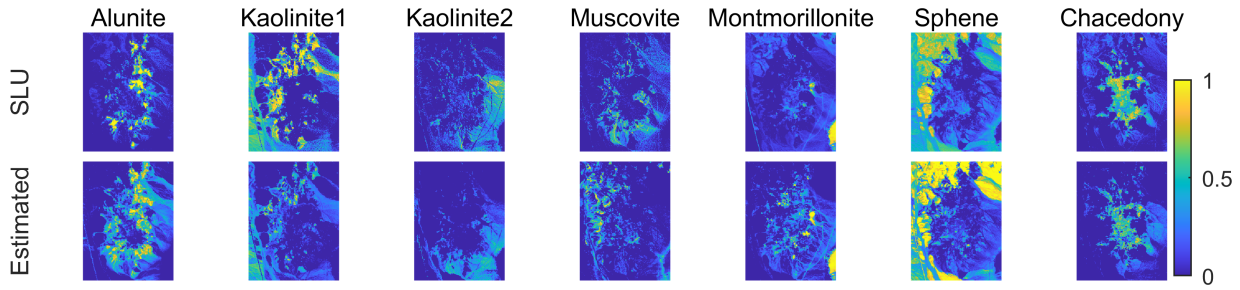


Fig. 16. Cuprite dataset - Visual comparison of the abundance maps obtained by the proposed method.

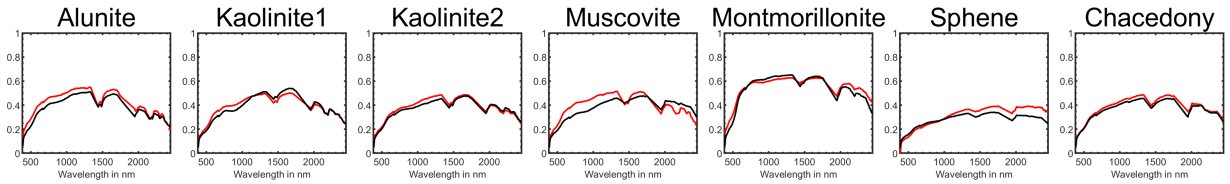


Fig. 17. Cuprite dataset - Visual comparison of the endmembers obtained by the proposed method. Red: ground truth endmembers; Black: estimated endmembers.

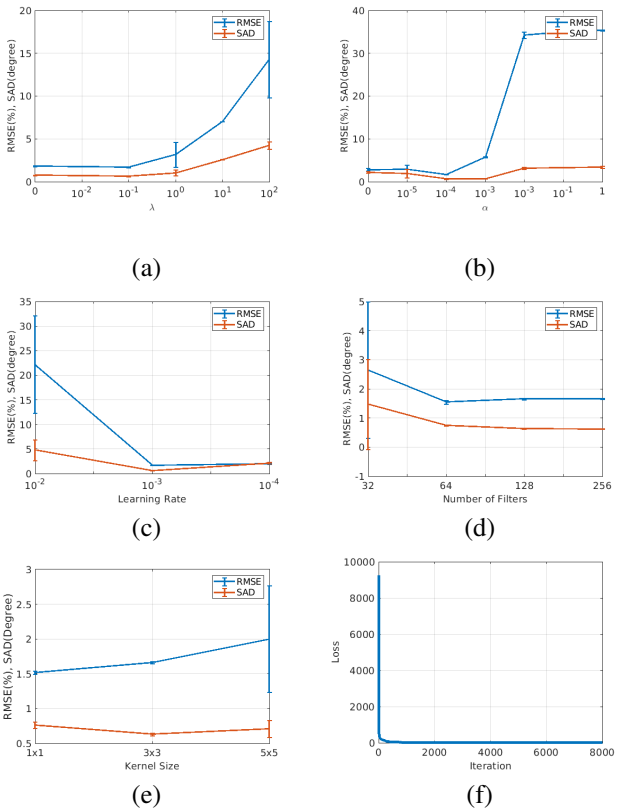


Fig. 18. Sensitivity of HapkeCNN to the parameters/ hyperparameters of the network. The performance is given in terms of SAD in degree and RMSE in percentage. The experiments were performed on the simulated dataset 2 (30 dB).

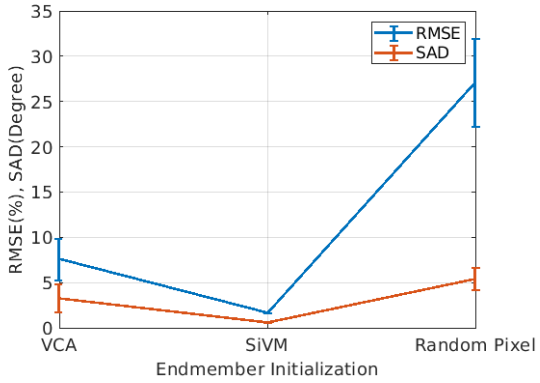


Fig. 19. Comparison of different endmember initializations for HapkeCNN. The performance is given in terms of SAD in degree and RMSE in percentage. The experiments were performed on the simulated dataset 2 (30 dB).

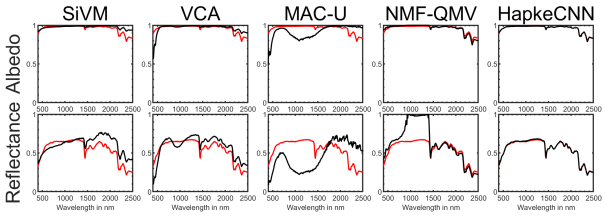


Fig. 20. Endmember 6 of simulated dataset 2 (40dB) - Visual comparison of the albedo and reflectance endmembers by the different unmixing techniques. Red: ground truth endmembers; Black: estimated endmembers.

[4] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354–379, April 2012.

[5] N. Dobigeon, J. Tourneret, C. Richard, J. C. M. Bermudez, S. McLaughlin, and A. O. Hero, "Nonlinear unmixing of hyperspectral images: Models and algorithms," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 82–94, 2014.

[6] W. Fan, B. Hu, J. Miller, and M. Li, "Comparative study between a new nonlinear model and common linear model for analysing laboratory simulated-forest hyperspectral data," *International Journal of Remote Sensing*, vol. 30, no. 11, pp. 2951–2962, 2009.

[7] R. Heylen, M. Parente, and P. Gader, "A review of nonlinear hyperspectral unmixing methods," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 6, pp. 1844–1868, 2014.

[8] Y. Altmann, A. Halimi, N. Dobigeon, and J.Y. Tourneret, "Supervised nonlinear spectral unmixing using a postnonlinear mixing model for hyperspectral imagery," *IEEE Transactions on Image Processing*, vol. 21, no. 6, pp. 3017–3025, 2012.

[9] A. Halimi, Y. Altmann, N. Dobigeon, and J. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4153–4162, 2011.

[10] L. Gao, Z. Wang, L. Zhuang, H. Yu, B. Zhang, and J. Chanussot, "Using low-rank representation of abundance maps and nonnegative tensor factorization for hyperspectral nonlinear unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–17, 2022.

[11] R. Heylen and P. Scheunders, "A multilinear mixing model for nonlinear spectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 1, pp. 240–251, Jan 2016.

[12] A. Marinoni and P. Gamba, "A novel approach for efficient p-linear hyperspectral unmixing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 6, pp. 1156–1168, Sep. 2015.

[13] A. Marinoni, J. Plaza, A. Plaza, and P. Gamba, "Nonlinear hyperspectral unmixing using nonlinearity order estimation and polytope decomposition," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 6, pp. 2644–2654, June 2015.

[14] A. Marinoni, A. Plaza, and P. Gamba, "Harmonic mixture modeling for efficient nonlinear hyperspectral unmixing," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 9, pp. 4247–4256, Sep. 2016.

[15] B. Rasti, D. Hong, R. Hang, P. Ghamisi, X. Kang, J. Chanussot, and J. A. Benediktsson, "Feature extraction for hyperspectral imagery: The evolution from shallow to deep (overview and toolbox)," *IEEE Geoscience and Remote Sensing Magazine*, pp. 0–0, 2020.

[16] P. Ghamisi, N. Yokoya, J. Li, W. Liao, S. Liu, J. Plaza, B. Rasti, and A. Plaza, "Advances in hyperspectral image and signal processing: A comprehensive overview of the state of the art," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 4, pp. 37–78, 2017.

[17] J. Broadwater, R. Chellappa, A. Banerjee, and P. Burlina, "Kernel fully constrained least squares abundance estimates," in *2007 IEEE International Geoscience and Remote Sensing Symposium*, 2007, pp. 4041–4044.

[18] X. Wu, X. Li, and L. Zhao, "A kernel spatial complexity-based nonlinear unmixing method of hyperspectral imagery," Berlin, Heidelberg, 2010, LSMS/ICSEE'10, p. 451–458, Springer-Verlag.

[19] M. Brown, H.G. Lewis, and S.R. Gunn, "Linear spectral mixture models and support vector machines for remote sensing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 5, pp. 2346–2360, 2000.

[20] B. Hapke, R. Nelson, and W. Smythe, "The opposition effect of the moon: Coherent backscatter and shadow hiding," *Icarus*, vol. 133, no. 1, pp. 89 – 97, 1998.

[21] B. Hapke, "Bidirectional reflectance spectroscopy: 1. theory," *Journal of Geophysical research*, vol. 86, pp. 3039–3054, 1981.

[22] M. Zhao, L. Yan, and J. Chen, "Lstm-dnn based autoencoder network for nonlinear hyperspectral image unmixing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 2, pp. 295–309, 2021.

[23] M. Zhao, M. Wang, J. Chen, and S. Rahardja, "Hyperspectral unmixing for additive nonlinear models with a 3-d-cnn autoencoder network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–15, 2022.

[24] K. T. Shahid and I. D. Schizas, "Spatial-aware hyperspectral nonlinear unmixing autoencoder with endmember number estimation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 20–41, 2022.

[25] H. Li, R. A. Borsoi, T. Imbiriba, P. Closas, J. C. M. Bermudez, and D. Erdoğmuş, "Model-based deep autoencoder networks for nonlinear hyperspectral unmixing," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.

[26] K. T. Shahid and I. D. Schizas, "Unsupervised hyperspectral unmixing via nonlinear autoencoders," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–13, 2022.

[27] Q. Lyu and X. Fu, "Identifiability-guaranteed simplex-structured post-nonlinear mixture learning via autoencoder," *IEEE Transactions on Signal Processing*, vol. 69, pp. 4921–4936, 2021.

[28] Mou Wang, Min Zhao, Jie Chen, and Susanto Rahardja, "Nonlinear unmixing of hyperspectral data via deep autoencoder networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 9, pp. 1467–1471, 2019.

[29] Maofeng Tang, Ying Qu, and Hairong Qi, "Hyperspectral nonlinear unmixing via generative adversarial network," in *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*, 2020, pp. 2404–2407.

[30] Yuanchao Su, Xiang Xu, Jun Li, Hairong Qi, Paolo Gamba, and Antonio Plaza, "Deep autoencoders with multitask learning for bilinear hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 10, pp. 8615–8629, 2021.

[31] B. Hapke, *Theory of Reflectance and Emittance Spectroscopy*, Cambridge University Press, 2 edition, 2012.

[32] S. Chandrasekhar, *Radiative Transfer*, Dover Books on Intermediate and Advanced Mathematics. Dover Publications, 1960.

[33] I. J. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, Cambridge, MA, USA, 2016, http://www.deeplearningbook.org.

[34] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

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

[36] B. Rasti, B. Koirala, P. Scheunders, and P. Ghamisi, "UnDIP: Hyperspectral unmixing using deep image prior," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–15, 2021.

[37] B. Rasti and B. Koirala, "Suncnn: Sparse unmixing using unsupervised convolutional neural network," *IEEE Geoscience and Remote Sensing Letters*, pp. 1–5, 2021.

[38] M. Berman, H. Kiiveri, R. Lagerstrom, A. Ernst, R. Dunne, and J.F. Huntington, "Ice: a statistical approach to identifying endmembers in hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 10, pp. 2085–2095, 2004.

[39] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Collaborative nonnegative matrix factorization for remotely sensed hyperspectral unmixing," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, July 2012, pp. 3078–3081.

[40] B. Somers, L. Tits, and P. Coppin, "Quantifying nonlinear spectral mixing in vegetated areas: Computer simulation model validation and first results," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 1956–1965, 2014.

[41] J. Stuckens, B. Somers, S. Delalieux, W. Verstraeten, and P. Coppin, "The impact of common assumptions on canopy radiative transfer simulations: A case study in citrus orchards," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 110, no. 1, pp. 1–21, 2009.

[42] J. F. Mustard and C. M. Pieters, "Photometric phase functions of common geologic minerals and applications to quantitative analysis of mineral mixture reflectance spectra," *Journal of Geophysical Research*, vol. 94, pp. 13619–13634, 1989.

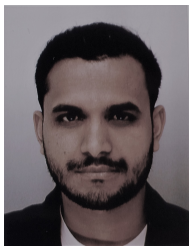
[43] J. Nascimento and J. Bioucas-Dias, "Vertex component analysis: A˜fast algorithm to extract endmembers spectra from hyperspectral data," in *Pattern Recognition and Image Analysis*, Francisco José Perales, Aurélio J. C. Campilho, Nicolás Pérez de la Blanca, and Alberto Sanfeliu, Eds., Berlin, Heidelberg, 2003, pp. 626–635, Springer Berlin Heidelberg.

[44] 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.

[45] 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.

Behnood Rasti (M'12–SM'19) received the B.Sc. and M.Sc. degrees both in electronics- electrical engineering from the Electrical Engineer- ing Department, University of Guilan, Rasht, Iran, in 2006 and 2009, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Iceland, Reykjavik, Iceland, in 2014. In 2015 and 2016, he worked as a Post-Doctoral Researcher with Electrical and Computer Engineering Department, University of Iceland. From 2016 to 2019, he has been a Lecturer with the Center of Engineering Technology and Applied Sciences, Department of Electrical and Computer Engineering, University of Iceland. Dr. Rasti was a Humboldt research fellow in 2020 and 2021. He is currently a Principal Research Associate with Helmholtz-Zentrum Dresden-Rossendorf (HZDR). His research interests include machine learning, signal and image processing, Earth observation, and remote sensing.

Dr. Rasti was the Valedictorian as an M.Sc. Student in 2009. He won the Doctoral Grant of The University of Iceland Research Fund "The Eimskip University fund," and the "Alexander von Humboldt Research Fellowship Grant" in 2013 and 2019, respectively. He serves as an Associate Editor for the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS (GRSL).



Bikram Koirala (Member, IEEE) received the M.S. degree in geomatics engineering from the University of Stuttgart, Stuttgart, Germany, in 2016, and the Ph.D. degree in development of advanced hyperspectral unmixing methods from the University of Antwerp, Antwerp, Belgium, in 2021.

In 2017, he joined the Vision Lab, Department of Physics, University of Antwerp, as a Ph.D. Researcher, where he is currently a Post-Doctoral Researcher. His research interests include machine learning and hyperspectral image processing.

Paul Scheunders (M'98) received the B.S. degree and the Ph.D. degree in physics, with work in the field of statistical mechanics, from the University of Antwerp, Antwerp, Belgium, in 1983 and 1990, respectively. In 1991, he became a research associate with the Vision Lab, Department of Physics, University of Antwerp, where he is currently a full professor. His current research interest includes remote sensing and hyperspectral image processing. He has published over 200 papers in international journals and proceedings in the field of image processing, pattern recognition, and remote sensing. Paul Scheunders is Associate Editor of the IEEE Transactions on Geoscience and Remote Sensing and has served as a program committee member in numerous international conferences. He is a senior member of the IEEE Geoscience and Remote Sensing Society.