



Article

DHCAE: Deep Hybrid Convolutional Autoencoder Approach for Robust Supervised Hyperspectral Unmixing

Fazal Hadi, Jingxiang Yang, Matee Ullah, Irfan Ahmad, Ghulam Farooque and Liang Xiao *

School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China

* Correspondence: xiaoliang@mail.njust.edu.cn

Abstract: Hyperspectral unmixing (HSU) is a crucial method to determine the fractional abundance of the material (endmembers) in each pixel. Most spectral unmixing methods are affected by low signal-to-noise ratios because of noisy pixels and bands simultaneously, requiring robust HSU techniques that exploit both 3D (spectral–spatial dimension) and 2D (spatial dimension) domains. In this paper, we present a new method for robust supervised HSU based on a deep hybrid (3D and 2D) convolutional autoencoder (DHCAE) network. Most HSU methods adopt the 2D model for simplicity, whereas the performance of HSU depends on spectral and spatial information. The DHCAE network exploits spectral and spatial information of the remote sensing images for abundance map estimation. In addition, DHCAE uses dropout to regularize the network for smooth learning and to avoid overfitting. Quantitative and qualitative results confirm that our proposed DHCAE network achieved better hyperspectral unmixing performance on synthetic and three real hyperspectral images, i.e., Jasper Ridge, urban and Washington DC Mall datasets.

Keywords: hyperspectral unmixing (HSU); hybrid 3D; 2D-convolutional autoencoder (AE) network; supervised; spectral–spatial; remote sensing; abundances



Citation: Hadi, F.; Yang, J.; Ullah, M.; Ahmad, I.; Farooque, G.; Xiao, L. DHCAE: Deep Hybrid Convolutional Autoencoder Approach for Robust Supervised Hyperspectral Unmixing. *Remote Sens.* **2022**, *14*, 4433. <https://doi.org/10.3390/rs14184433>

Academic Editor: Pedro Melo-Pinto

Received: 6 July 2022

Accepted: 30 August 2022

Published: 6 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Hyperspectral image (HSI) analysis is an essential and critical technology in the field of aerial and satellite-based images (remote sensing). Hyperspectral data has been successfully applied in various fields, such as classification [1–5], environmental monitoring [6,7], and object recognition [8–10]. Because of the limited spatial resolution of HSI, many pixels are typically mixed by way of a number of materials, which degrade the overall performance of hyperspectral data processing [11,12]. Therefore, hyperspectral unmixing (HSU) has become an important technique to handle the mixed pixels issue. HSU aims to decompose the measured pixel spectra from remote sensing data into a set of pure spectral signatures, referred to as endmembers and their fractional abundances. At each pixel, the endmembers are normally assumed to signify the pure materials, and the abundances show the ratio of every endmember. HSU applications are in various fields, such as agriculture, geology [13], and environmental biology studies [14]. The issue of mixed pixels can be addressed via three distinct deep learning approaches, such as supervised, unsupervised and semi supervised. In this work, the supervised approach can be considered to estimate the abundance map as the endmembers are supposed to be known. They can be extracted from endmember extraction algorithms, such as pixel purity index (PPI) [15], N-FINDER [16], vertex component analysis (VCA) [17] and minimum volume simplex analysis (MVSA) [18]. In the unsupervised approach, different HSU methods have been proposed [19–24] to estimate both endmembers and their fractional abundances simultaneously. Semi supervised approaches consider that each mixed pixel in the observed image is represented as a mixture of more than one spectral signature known in a huge spectral library [25].

There exist two kinds of HSU models: the linear spectral mixing model (LSMM) and the nonlinear spectral mixing model (NLSMM) [26]. LSMM holds when the incident

light interacts with just one material [27], and the pixels are linearly expressed by the combination of endmembers. NLSMM holds when the light interacts with more than one material in the scene. The LSMM is broadly used in several applications due to its efficiency and simplicity [28]. Several methods, such as the sparse regression-based method [29], nonnegative matrix factorization (NMF) [30], and the geometrical-based method [31], have been used to solve linear spectral mixing problems. These linear spectral unmixing methods derive endmembers and their abundance from the hyperspectral image.

Over the last few decades, several supervised approaches to the linear spectral mixing problem have been proposed. One of the well-known supervised spectral mixing problems based on the fully constrained least square (FCLS) method was adopted to solve the mixing problem, it minimizes the error between the true spectrum and the linearly mixed spectrum, subject to the physical constraints that the abundances should be positive and sum to one. A similar method, mixture tuned matched filtering (MTMF) [32], has been adopted to extract abundance. Another recent unmixing algorithm, such as sparse unmixing by variable splitting and augmented Lagrangian (SUnSAL) [33], and its constrained alternative (CSUnSAL) have also been proposed to solve the optimization problem by taking advantage of the alternating direction method of multipliers (ADMM) [34]. Both SUnSAL and CSUnSAL apply an ℓ_1 regularization term on the fractional abundances. SUnSAL uses an ℓ_2 regularization term as the fidelity term while CSUnSAL assumes a constraint to enforce the data fidelity. SUnSAL was improved in [35] by the use of spatial information via total variation (TV) regularization on the abundances (SUnSAL-TV). However, this could result in issues with over smoothness and blurred abundance maps. Thus, collaborative sparse unmixing [29] applies $\ell_{2,1}$ regularization terms to promote the sparsity of abundance maps. The spectral variability can be represented using a perturbed linear mixing model (PLMM) [36] and extended linear mixing model (ELMM) [37] to estimate abundances. In [38], a new augmented linear mixing model (ALMM) was presented. First, the letter introduces the spectral variability dictionary and then explores data-driven dictionary learning to extract abundances in a second step.

In the past few decades, deep learning approaches have been employed in computer vision, object detection [39], and pattern recognition. These approaches automatically extract rich contents from the remote sensing input data and have been used within hyperspectral image classification. A literature review has shown that very few deep learning methods have been applied to hyperspectral unmixing, as they have for things such as image classification. In this work, we intend to fill in this gap by creating a new connection between the deep supervised learning approach and robust supervised HSU to overcome the issue of a mixed pixels.

Various deep learning approaches have been employed in hyperspectral unmixing with the success of artificial neural networks. Recent neural network methods used for the HSU problem are [40–47]. In [48], 1D and 3D methods utilizing a scattering transform framework (STF) were presented to extract features from HSIs, and then used k-nearest neighbor (k-NN) to extract abundance. More recently, autoencoder as a deep learning model based on the neural network has been used to address HSU problems. Autoencoder has gained tremendous popularity in the hyperspectral community. Two specific instances of autoencoder, de-noising and nonnegative sparse autoencoder (NNSAE), were utilized to address the HSU problem by simultaneously estimating endmembers and their fractional abundances from HSI [49,50]. Another unmixing method [51] has been employed to solve the HSU problem using a variational autoencoder and stacked NNSAE. In [52], a concatenated marginalized de-noising model and NNSAE were used to solve hyperspectral unmixing in noisy environments. In [53], the authors proposed a stacked NNSAE to minimize the impact of low signal-to-noise ratio. In addition, several other autoencoders, such as [54], have been developed to improve spectral unmixing performance using multitask-based learning and convolutional neural network (CNN) [55]. Another two-stream network [56], consisting of an endmember extraction and an unmixing network called a weakly supervised method, has been proposed for the spectral unmixing problem.

Recent research has made clear that using 2D-CNN or 3D-CNN exclusively has drawbacks, such as wasting band-related information or requiring the use of extremely complex techniques. Furthermore, they inhibit the deep learning techniques from obtaining exceptional accuracy. Their main justification is that HSI produces spectral-dimensional volumetric data. The spectral interpretations cannot be used to create useful maps with distinguishing features using the 2D-CNN approach alone. A deep 3D-CNN approach is also computationally expensive. When employed alone, it performs badly for materials of related features across many spectral bands. The approaches also require extra processing time to examine and comprehend the spectral–spatial data cubes. Additionally, the presence of various types of noise, i.e., noisy pixels and channels in the remote sensing data, badly degrades the overall performance of spectral unmixing. As the literature suggests, few of the existing unmixing methods have been proposed to achieve robustness in spectral dimension [57,58]. For these thoughts, we are motivated to propose a supervised robust HSU method that considers both noisy pixels and channels to enhance the robustness of spectral and spatial dimensions.

We propose a novel supervised end-to-end deep hybrid convolutional autoencoder (DHCAE) network for robust HSU in this paper. The proposed method can utilize spectral and spatial information to achieve abundances given the endmembers in the HSI. The main research contributions of our proposed approach are highlighted as threefold:

1. According to the best of the authors' knowledge, this is the first time that a robust HSU model using a deep hybrid convolutional autoencoder has been proposed to build an end-to-end framework. This framework can learn discriminative features from HSI to produce better unmixing performance.
2. We used 3D and 2D layer information in the proposed approach, utilizing spectral–spatial information to improve the hyperspectral unmixing performance.
3. The proposed method performance is evaluated on one synthetic and three real datasets. The results confirm that the DHCAE approach outperforms existing methods.

The remainder of this article is organized as follows. The proposed DHCAE network is described in Section 2. The results on the experiments conducted on synthetic and three real world remote sensing datasets are illustrated in Section 3. Section 4 offers the discussion; and finally, Section 5 concludes this article.

2. Methodology

2.1. Notation and Problem Formulation

The unmixing model for the proposed network follows the LSMM, and notations are as follows. Let $\mathbf{X} = [x_1, \dots, x_N] \in \mathbb{R}^{L \times N}$ is the observed HSIs with L bands, the letter $N = w \times h$ denotes the total number of pixels, $\mathbf{M} = [m_1, \dots, m_P] \in \mathbb{R}^{L \times P}$ denotes the endmember matrix with P pure spectra in each column and $\mathbf{A} = [a_1, \dots, a_N] \in \mathbb{R}^{P \times N}$ represents the abundance matrix. Finally, let $\mathbf{N} \in \mathbb{R}^{L \times N}$ be the residual matrix (e.g., additive noise and other errors). The spectrum of LSMM can be expressed as

$$x_i = \sum_{j=1}^P a_{i,j} m_j + n_i = M a_i + n_i \quad (1)$$

The Equation (1) can be considered as the complete observed image as expressed below:

$$\mathbf{X} = \mathbf{M}\mathbf{A} + \mathbf{N} \quad (2)$$

The physical constraints enforced on the abundances are the abundance non-negativity constraint (ANC), i.e., $a_{i,j} \geq 0$ while $\sum_{j=1}^P a_{i,j} = 1$ is the abundance sum-to-one constraint (ASC). In this work, given the HSIs (\mathbf{X}) and endmembers matrix (\mathbf{M}), abundances matrix (\mathbf{A}) is the solution target. It should be noted that the pure spectra are manually chosen from the original hyperspectral image by visual judgement and by looking at the spectral signature to overcome the spectral variability.

Due to the powerful learning capabilities, an autoencoder is constructed for the unmixing problem, as shown in Figure 1. An autoencoder generally aims to learn the low-dimensional representation of input data, which comprises two parts.

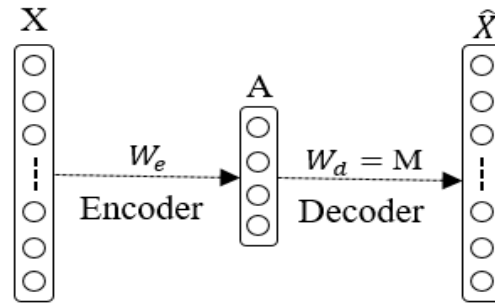


Figure 1. Supervised approach of HSU using autoencoder.

- (1) Encoder: The encoder part encodes the input data, $X \in \mathbb{R}^{L \times N}$ into a hidden representation, $A \in \mathbb{R}^{P \times N}$ (i.e., abundance) is denoted by

$$A = f(X) = \sigma_e(W_e X) \quad (3)$$

where σ_e denotes component-wise activation function of the encoder part, such as rectified linear unit (ReLU) or Softmax and W_e represents the weights of the encoder part.

- (2) Decoder: The decoder part reconstructs the input data, $\hat{X} \in \mathbb{R}^{L \times N}$ from the hidden representation, $A \in \mathbb{R}^{P \times N}$ (i.e., abundance) can be expressed as

$$\hat{X} = g(A) = \sigma_d(W_d A) \quad (4)$$

where W_d denotes weights of the decoder part, and \hat{X} represents the reconstructed input data. The autoencoder network is trained using backpropagation to minimize the loss function, $\mathcal{L}(X, g(f(X)))$.

2.2. Proposed DHCAE Network

This section presents the DHCAE network that exploits spectral–spatial information in a remote sensing hyperspectral data cube X , implementing 3D and 2D CNN models to achieve better unmixing performance. At this point, the observed HSIs data cube is denoted by $X \in \mathbb{R}^{H \times W \times L}$ where H and W correspond to the height and width of the observed image, respectively, and L represents the number of channels (spectral bands) around each pixel to extract its corresponding abundance vector. Therefore, first, we divide the HSIs data cube into small patches, $p \in \mathbb{R}^{S \times S \times L}$ where L indicates the spectral bands and $(S \times S)$ represents spatial window size. Thus, each 3D patch is fed into the encoder part of the DHCAE network to extract spectral–spatial information through 3D and 2D convolutional operators, and its abundance map (latent space) is obtained. The decoder part then takes the latent space as input to recreate the original input with the LSMM in Equation (1). Thus, we set the weights of the decoder part equal to the endmember matrix in the DHCAE network with no trainable parameters.

In a 2D convolutional autoencoder, the input patches are convolved with a 2D kernel that computes the sum of the dot product between the input patch and kernel. The kernel is striding over the input patch to cover the whole spatial dimension. In the 2D convolutional layer, the value at position (x, y) of the j^{th} feature map in the i^{th} layer is formulated as

$$v_{i,j}^{x,y} = \text{ReLU} \left(\sum_m \sum_{h=0}^{H_k-1} \sum_{w=0}^{W_k-1} w_{ijm}^{hw} v_{(i-1)m}^{(x+h)(y+w)} \right) + b_{i,j} \quad (5)$$

where $\text{ReLU}(\cdot)$ is the activation function, m indexes over the set of feature maps of the $(i-1)$ th layer, W_k and H_k are the width and height of the 2D kernel, respectively. w_{ijm}^{hw} is the

weight parameter at position (h, w) connected to the m^{th} feature map, $b_{i,j}$ is the bias term. The 3D convolutional operation of the convolutional layer can be formulated as

$$v_{i,j}^{x,y,z} = ReLU \left(\sum_m \sum_{h=0}^{H_k-1} \sum_{w=0}^{W_k-1} \sum_{d=0}^{D_k-1} w_{ijm}^{hwd} v_{(i-1)m}^{(x+h)(y+w)(z+w)} \right) + b_{i,j} \quad (6)$$

where D_k is the depth of the 3D convolution kernel along the spectral dimension. $v_{ij}^{x,y,z}$ is the weight at position $(x, y, \text{ and } z)$ of the kernel connected to the m^{th} feature map, and other parameters are the same as in Equation (5).

The 2D convolutional kernel cover only spatial dimension. On the other side, the 3D convolutional kernel can concurrently extract spectral and spatial features from HSI, but it consumes computation time. Therefore, to take advantage of both the 2D and 3D convolutional autoencoder, we proposed a DHCAE network for robust hyperspectral unmixing.

Figure 2 shows the proposed DHCAE network for robust hyperspectral unmixing. It is composed of three 3D convolutions (C1, C2, C3), two 2D convolutions (C4, C5) and two fully connected (F6, F7) layers regarding the encoder part and only one fully connected layer (O8) in the decoder part. Using ReLU and Softmax activation functions, the output of the F7 layer imposes ANC and ASC on the abundance map, respectively.

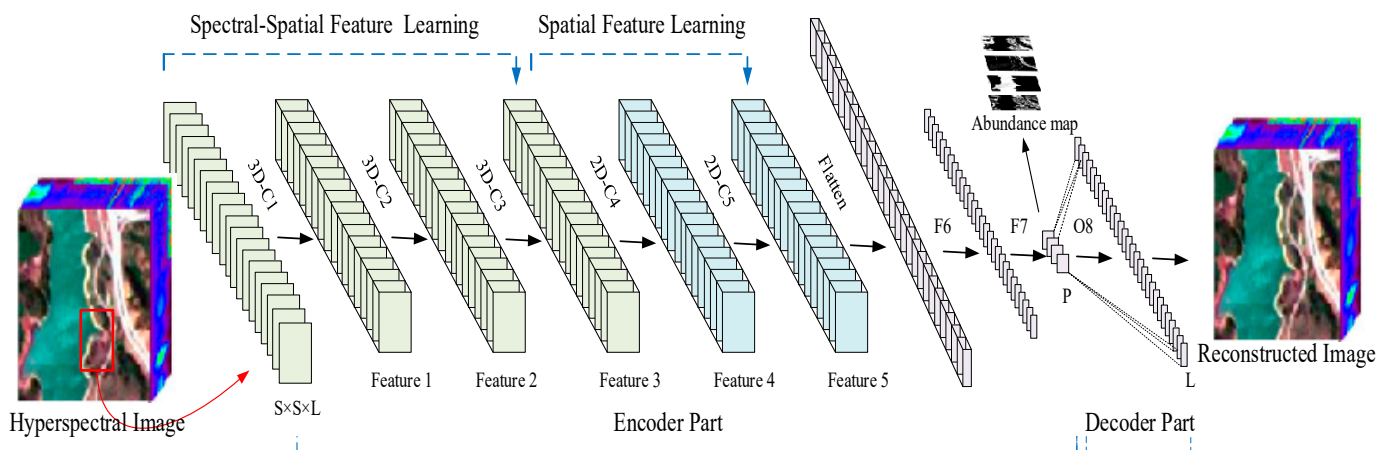


Figure 2. Flowchart of the proposed DHCAE Approach for robust hyperspectral unmixing (HSU).

In the DHCAE framework, the dimensions of the 3D convolution kernels are as follows: 3D_C1 = $8 \times 3 \times 3 \times 9 \times 1$ (i.e., $K_1^1 = K_2^1 = 3$, and $K_3^1 = 9$), 3D_C2 = $16 \times 3 \times 3 \times 7 \times 1$ (i.e., $K_1^2 = K_2^2 = 3$, and $K_3^2 = 7$) and 3D_C3 = $32 \times 3 \times 3 \times 5 \times 1$ (i.e., $K_1^3 = K_2^3 = 3$, and $K_3^3 = 5$). Whereas dimensions of the 2D convolution kernels are 2D_C4 = $64 \times 3 \times 3 \times 5760$ (i.e., $K_1^4 = K_2^4 = 3$) and 2D_C5 = $128 \times 3 \times 3 \times 128$ (i.e., $K_1^5 = K_2^5 = 3$). Where 64 is the number of kernels, 3×3 means the spatial dimension of 2D kernels, and 5760 is the number of the input feature map. Three 3D convolution layers are deployed to preserve the spectral information in the HSIs input data to increase the number of spectral–spatial feature maps. Before the flatten layer, two 2D convolution layers are applied to ensure the model can discriminate the spatial information within different spectral bands without any loss. A detailed summary of the proposed DHCAE network can be found in Table 1.

Table 1. Detailed summary of the proposed DHCAE network.

Layers	Output Shape	Activation	# Parameters
Input_1 (Input Layer)	(15, 15, 198, 1)	-	0
3D_C1 (Conv3D)	(13, 13, 190, 8)	ReLU	656
3D_C2 (Conv3D)	(11, 11, 184, 16)	ReLU	8080
3D_C3 (Conv3D)	(9, 9, 180, 32)	ReLU	23,072
Reshape_1 (Reshape)	(9, 9, 5760)	-	0
2D_C4 (Conv2D)	(7, 7, 64)	ReLU	3,317,824
2D_C5 (Conv2D)	(5, 5, 128)	ReLU	73,856
flatten_1 (Flatten)	(3200)	-	0
F6 (Dense)	(128)	-	409,728
dropout_1 (Dropout)	(128)	-	0
F7 (Dense)	(P)	ReLU + Softmax	516
O8 (Dense)	(L)	-	792
Total Trainable Parameters: 3,834,524			

3. Experiment and Analysis

The proposed DHCAE was implemented in a Keras and Tensorflow framework using python. The experiments were performed on a HP Notebook—15-da0001tx Intel Core i7-8550U CPU and a GPU with 4 GB of memory. We demonstrate the unmixing performance of the proposed approach on four datasets, one synthetic hyperspectral data and three real datasets. We compared the proposed approach (DHCAE) with six related unmixing methods, namely, FCLS [59], nonnegative matrix factorization quadratic minimum volume (NMF-QMV) [60], augmented linear mixing model (ALMM) [31], hyperspectral unmixing using deep image prior (UnDIP) [61], deep hyperspectral unmixing using transformer network (DHTN) [62] and deep convolutional autoencoder (DCAE) [63]. In terms of abundance estimation, DCAE performs well. It is easier to compare their capabilities in hyperspectral unmixing issues because DCAE and DHCAE networks have similar structures. For the DCAE network, the encoder part comprises four convolutional layers and two fully connected layers. The decoder part of the DCAE network is the same as our approach. For the quantitative assessment of the algorithm, we used three criteria, such as abundance overall root mean square error (aRMSE), reconstruction overall root mean square error (rRMSE), and average spectral angle mapper (aSAM). When the reference of the abundance maps ($A^s = [a_1^s, \dots, a_N^s] \in \mathbb{R}^{P \times N}$) is given, then aRMSE can be used to measure the estimated abundance maps ($A^e = [a_1^e, \dots, a_N^e] \in \mathbb{R}^{P \times N}$). Without the groundtruth of abundance maps, the other two metrics are used to measure the reconstruction error between the original HSIs ($X^o = [x_1^o, \dots, x_N^o] \in \mathbb{R}^{L \times N}$) and its reconstruction ($X^r = [x_1^r, \dots, x_N^r] \in \mathbb{R}^{L \times N}$). These are defined as follows:

$$\text{aRMSE} = \frac{1}{N} \sum_{i=1}^N \sqrt{\frac{1}{P} \sum_{j=1}^P (a_{i,j}^s - a_{i,j}^e)^2} \quad (7)$$

$$\text{rRMSE} = \frac{1}{N} \sum_{i=1}^N \sqrt{\frac{1}{L} \sum_{l=1}^L (x_{i,j}^o - x_{i,j}^r)^2} \quad (8)$$

$$\text{aSAM} = \frac{1}{N} \sum_{i=1}^N \arccos \left(\frac{x_i^{rT} x_i^o}{\|x_i^r\| \|x_i^o\|} \right) \quad (9)$$

3.1. Experiments on Synthetic Dataset

The synthetic data were generated using five spectral endmember references randomly chosen from the United States Geological Survey (USGS) digital spectral library [64]. Figure 3 shows that the five endmember references have 480 spectral bands. The size of the abundance map is 64×64 pixels with a maximum abundance purity of 0.8. We followed the procedures in [65,66].

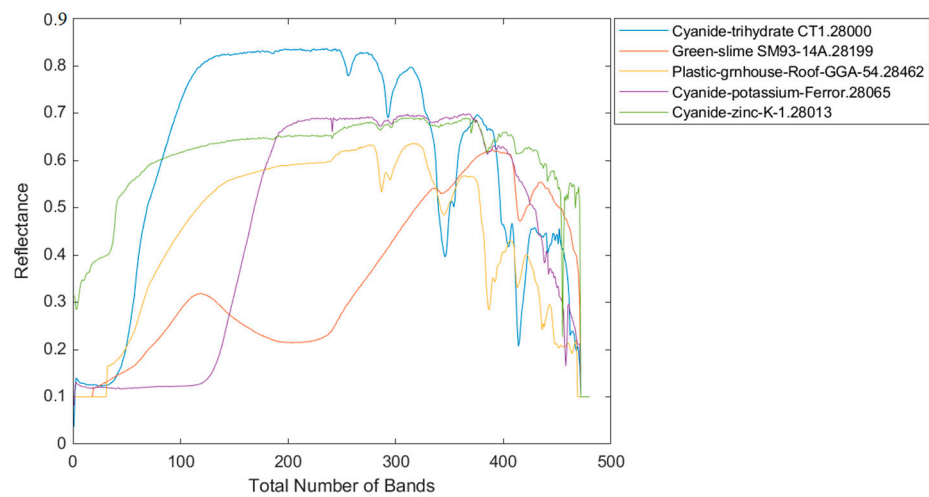


Figure 3. Five selected endmembers from the USGS library.

The extracted abundance map of different materials in the synthetic dataset is shown in Figure 4. The first column is the true abundance map of distinct spectral signatures, and the rest of the columns are the extracted abundance maps by various unmixing methods, respectively. DHTN and DCAE achieve desirable abundance maps; however, the outcomes are still unsatisfactory. In comparison with the previous state-of-the-art approaches, our proposed approach extracts abundance maps that are typically similar to the real abundance maps. All the extracted abundances of the five spectral signatures are satisfied and stable.

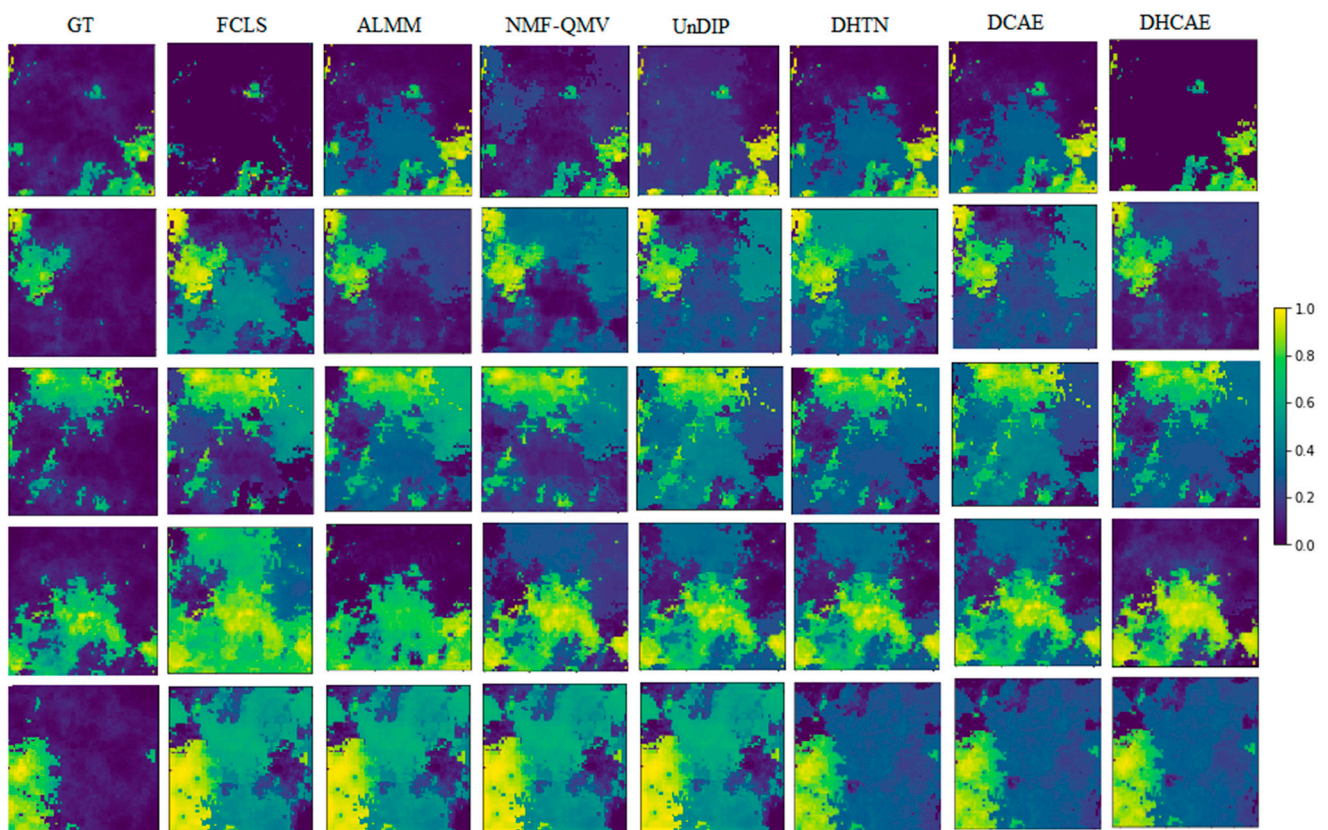


Figure 4. Synthetic data. Ground truth (GT) map in the first column and extracted abundance maps comparison by different spectral unmixing methods.

To quantitatively validate the robustness of our proposed approach in the simulated experiments. We compared the performance of various unmixing algorithms on synthetic data with three distinct types of noise. These three types of noises are as follows: just band noise added, only pixel noise added, and both band noise and pixel noise added.

For the added band noise, to each band was added Gaussian noises with four various levels of SNR, i.e., 05, 15, 25 and 35 dB, to test the performance of the proposed approach under different levels of noise. Figure 5 shows the score of aSAM and aRMSE for all the methods in the various noise levels. It can be seen that our proposed approach has good robustness against band noise of various SNRs than other unmixing methods.

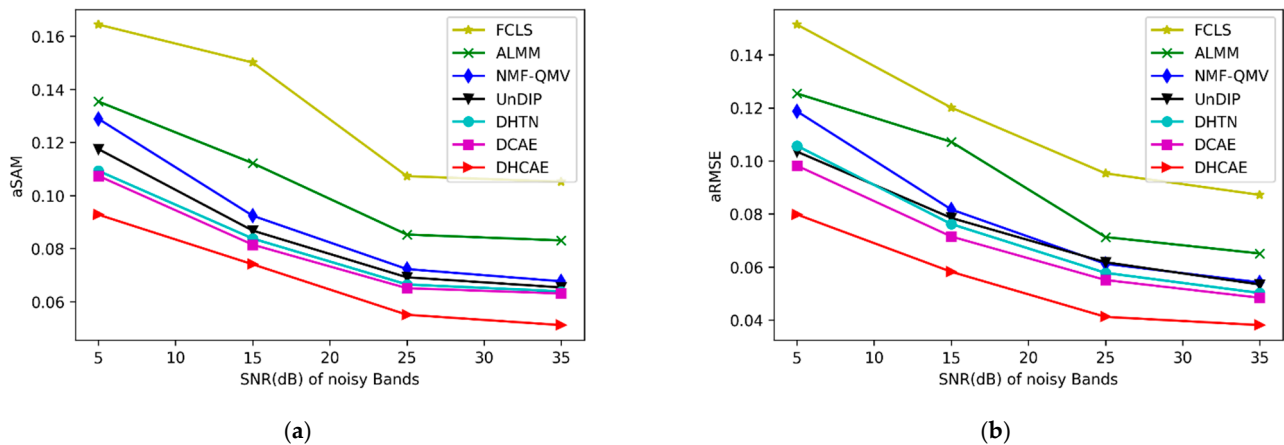


Figure 5. Robustness evaluation of the unmixing methods with various levels of band noise. (a) aSAM. (b) aRMSE.

For the added pixel noise, we also added Gaussian noises to each pixel on the synthetic image with different SNR values from 05 to 35 dB. Figure 6 shows the aSAM and aRMSE values for all the methods in various noise levels. The figure shows that our proposed approach performs well and is more robust against pixel noise.

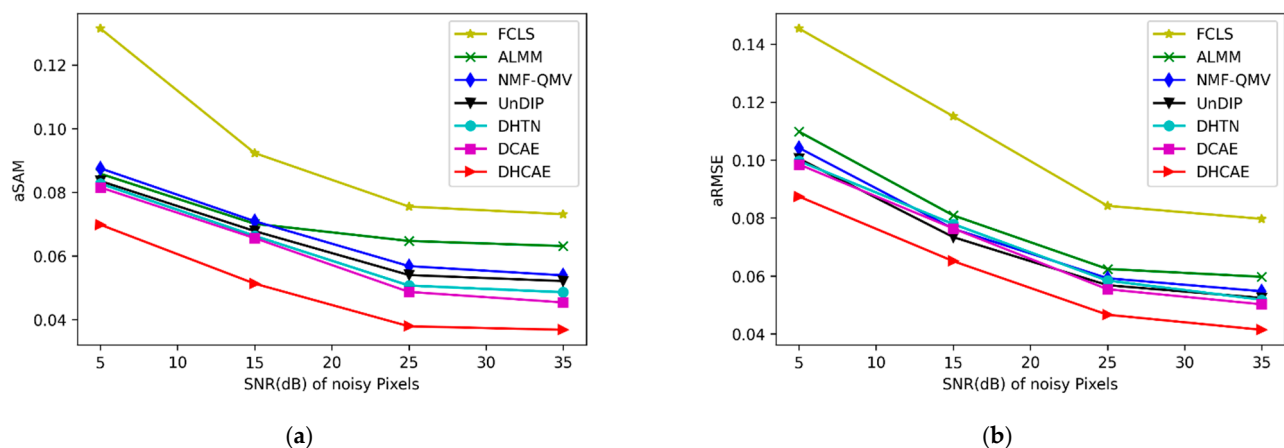


Figure 6. Robustness evaluation of the unmixing methods with various levels of pixel noise. (a) aSAM. (b) aRMSE.

To investigate the robustness of the proposed approach we also performed experiments where we simultaneously added pixel noise and band noise. Each pixel and band were added with Gaussian noise with different SNR levels of 05, 15, 25, and 35 dB, to the synthetic data set. As illustrated in Figure 7, our proposed approach is more robust to noise with different pixel noise and band noise levels. The performance of the other competitors reduces when the number of noisy pixels and bands increases, which means that the

other competitors are easily corrupted by noisy pixels and noisy bands compared with the proposed approach.

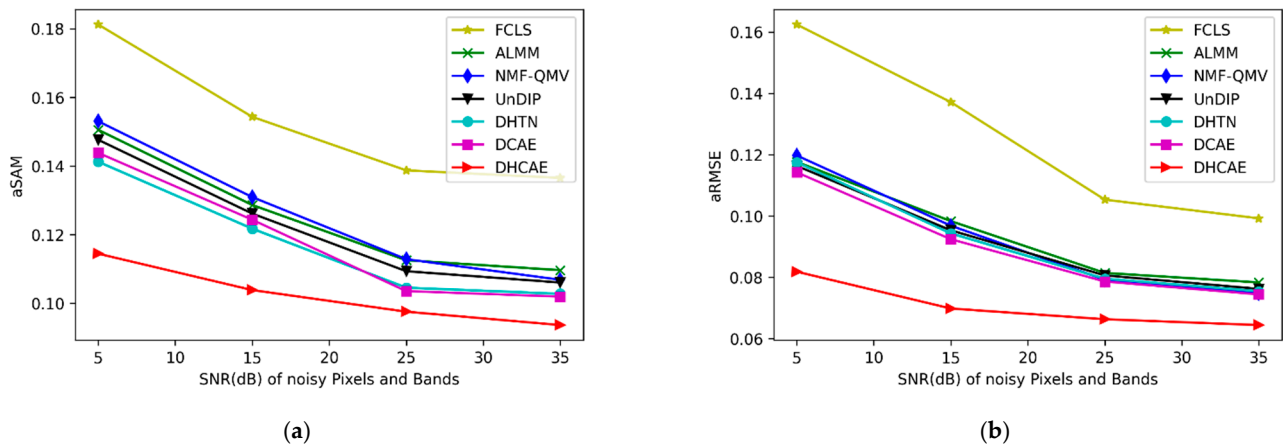


Figure 7. Robustness evaluation of the unmixing methods with various pixel and band noise levels. (a) aSAM. (b) aRMSE.

3.2. Experiments on Jasper Ridge Dataset

The first real HSIs scene used in our experiments was the Jasper Ridge dataset captured by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), as shown in Figure 8a. The image of size 100×100 pixels was measured in 224 spectral bands covering the wavelength range from 0.38 to $2.5 \mu\text{m}$. Tree, Water, Dirt, and Road are the four main endmembers observed in this dataset. In all 224 bands, there exist some noisy bands (2–4, 220–224) and blank bands (1, 108–112, 154–166). Figure 8b–e shows the noisy bands seriously corrupted by noise. The presence of noisy bands in the HSIs reduces the performance of the HSU methods. Therefore, to test the robustness of our proposed approach, we conducted experiments on data without noisy bands (i.e., bands = 198) and with noisy bands (i.e., bands = 224).

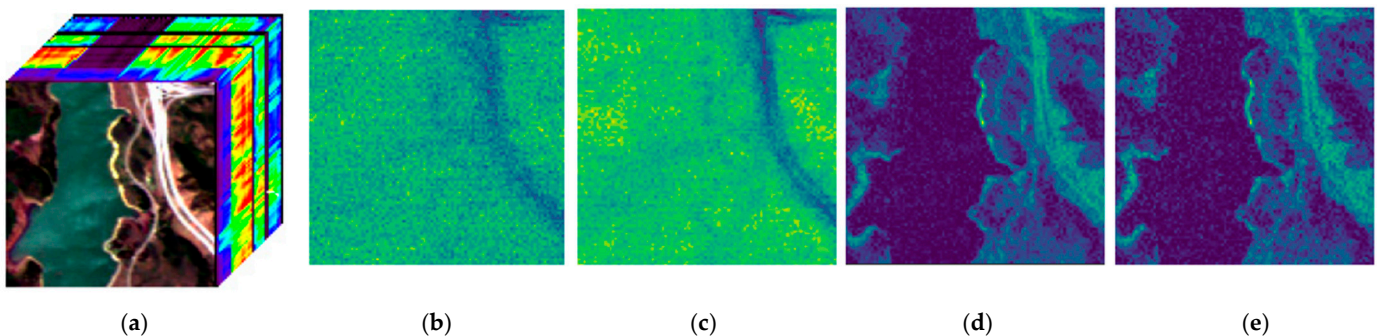


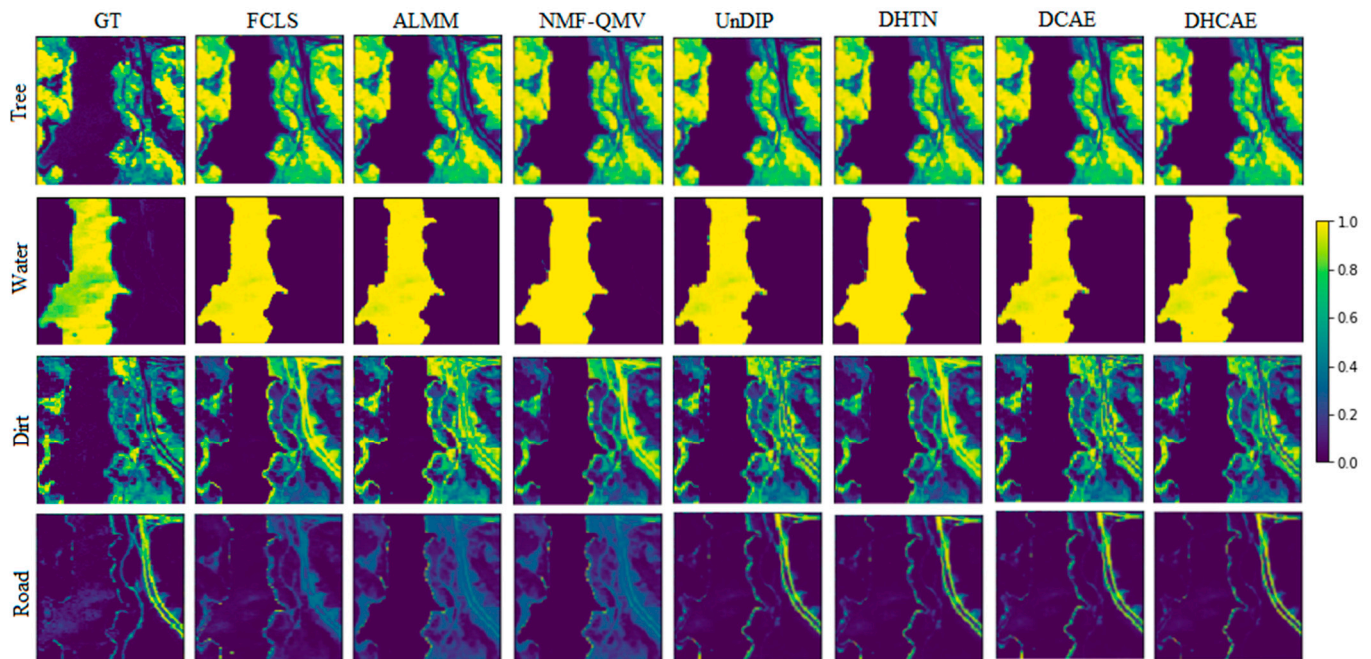
Figure 8. Hyperspectral image of (a) the Jasper Ridge data set. (b) Band 2. (c) Band 3. (d) band 223. (e) Band 224.

3.2.1. Results without Noisy Bands

Table 2 lists the quantitative results for the Jasper Ridge dataset without noisy bands, where the proposed DHCAE network achieves the best performance compared with other competitors. The proposed method obtains the rRMSE value of 0.0068 and aSAM value of 0.0314. On this dataset, the second-best result was obtained by DCAE. Figure 9 shows the extracted abundances for different materials. The extracted abundances, excluding our proposed approach, are shown in Figure 9. The DCAE method also achieved accurate abundance estimation for all endmembers.

Table 2. Quantitative performance on the Jasper Ridge dataset without noisy bands by various approaches. The best result is shown in bold.

Algorithm	FCLS	ALMM	NMF-QMV	UnDIP	DHTN	DCAE	DHCAE
rRMSE	0.0216	0.0515	0.0157	0.0136	0.0143	0.0124	0.0068
aSAM	0.0816	0.3378	0.0623	0.0497	0.0514	0.0508	0.0314

**Figure 9.** Jasper Ridge data. Ground truth (GT) map in the first column and extracted abundance maps without noisy bands comparison by different spectral unmixing methods.

3.2.2. Results with Noisy Bands

We also tested the unmixing performance of our proposed approach on the Jasper Ridge dataset with noisy bands containing 224 bands. Table 3 represents the rRMSE and aSAM values of various unmixing approaches. The results show that the rRMSE and aSAM value achieved by the DHCAE network is better than other unmixing methods. Figure 10 illustrates the qualitative results of extracted abundances for different spectral signatures. According to the qualitative results, our proposed approach achieves an abundance map with good robustness to noise compared with other unmixing approaches.

Table 3. Quantitative performance on the Jasper Ridge dataset with noisy bands by various approaches. The best result is shown in bold.

Algorithm	FCLS	ALMM	NMF-QMV	UnDIP	DHTN	DCAE	DHCAE
rRMSE	0.0752	0.0945	0.0663	0.0523	0.0529	0.0496	0.0415
aSAM	0.2267	0.5116	0.1528	0.1102	0.1127	0.1083	0.0706

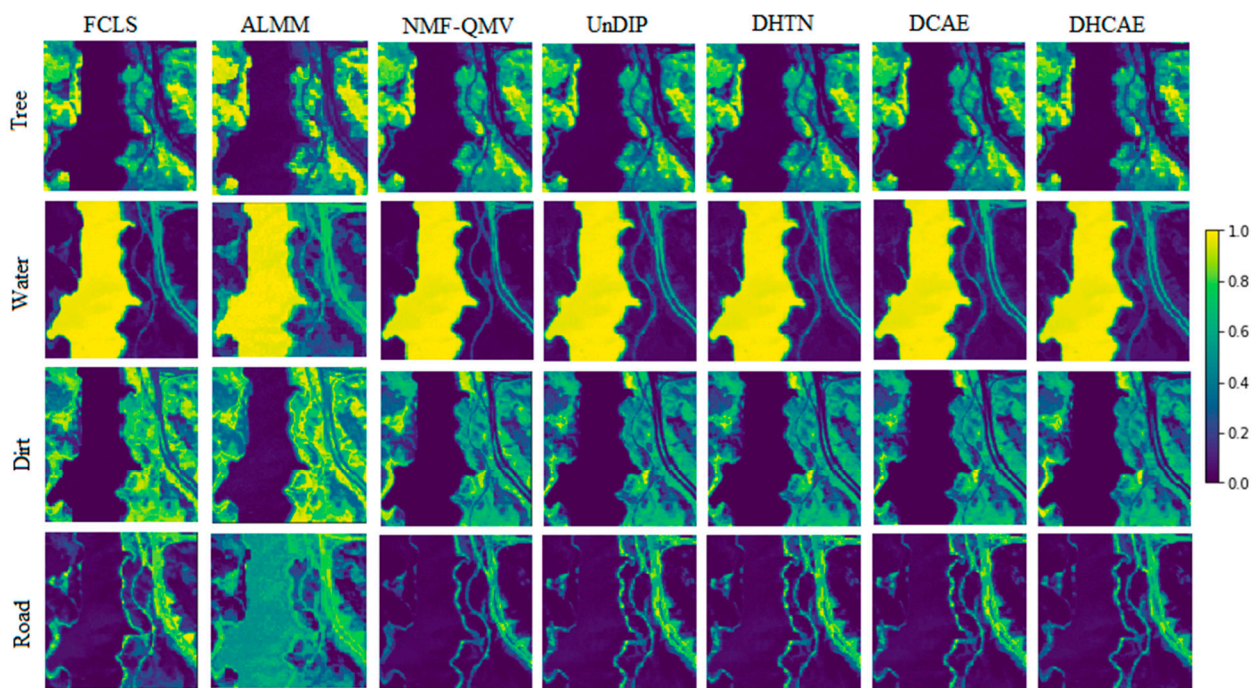


Figure 10. Jasper Ridge data. Comparison of extracted abundance maps with noisy bands by different spectral unmixing methods.

3.3. Experiments on Urban Dataset

The second real dataset was collected by the Hyperspectral Digital Imagery Collection Experiment (HYDICE) sensor in October 1995. Figure 11a shows the observed image of 307×307 pixels. Four kinds of different materials are included in the scene: asphalt, grass, trees and roofs. The original dataset has 210 spectral bands ranging from 400 to 2500 nm. Some noisy bands exist in this dataset (1–4, 76, 87, 101–111, 136–153, and 198–210) due to the atmospheric effects and water absorption. Figure 11b–e shows some noisy bands of this dataset. It can be seen that bands (138, 149, 207, and 209) are corrupted by noise. The existence of noisy bands degrades the performance of spectral unmixing methods, but the noisy bands may also contain important information. Therefore, we performed experiments on both data with removed noise band (i.e., bands = 162) and containing noise band (i.e., bands = 210) to demonstrate the robustness of our approach.

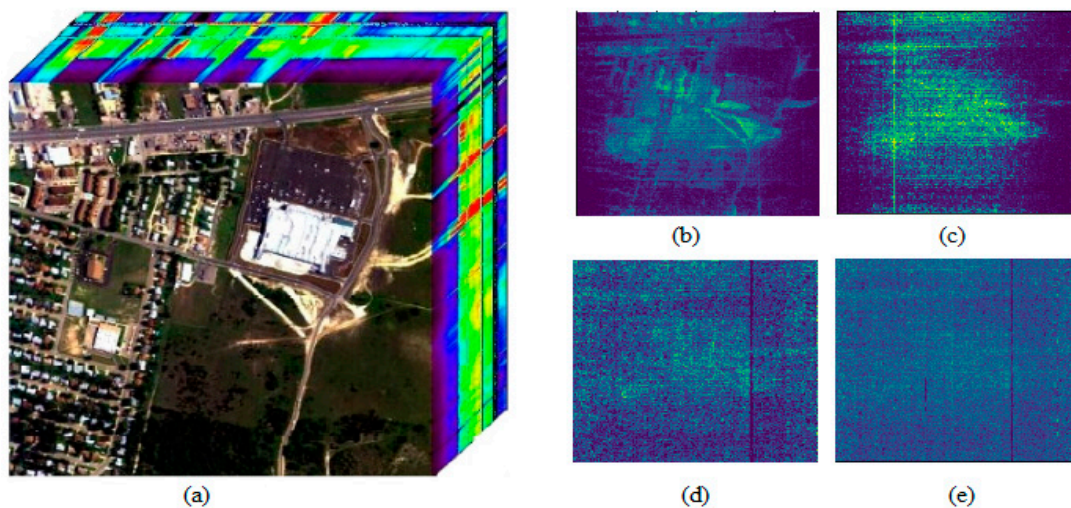


Figure 11. Hyperspectral dataset. (a) Urban. (b) Band 138. (c) Band 149. (d) Band 208. (e) Band 209.

3.3.1. Results with Removing Noisy Bands

Table 4 presents the quantitative comparison of rRMSE and aSAM results based on the urban datasets. According to Table 4, the proposed DHCAE network provides superior unmixing results to the other competitors. The value of rRMSE and aSAM in the proposed method are 0.0115 and 0.0331 respectively, DCAE can achieve the second-best results on this dataset. Figure 12 depicts the extracted abundances for different materials (endmembers) in the urban dataset. It can be observed that our proposed approach extracts abundance maps that are more separable and closer to the ground truth abundance maps than those provided by the state-of-the-art competitors.

Table 4. Quantitative performance on the urban dataset without noisy bands by various approaches. The best result is shown in bold.

Algorithm	FCLS	ALMM	NMF-QMV	UnDIP	DHTN	DCAE	DHCAE
rRMSE	0.0528	0.0484	0.0318	0.0286	0.0276	0.0249	0.0115
aSAM	0.1512	0.2602	0.0927	0.0603	0.0627	0.0525	0.0331

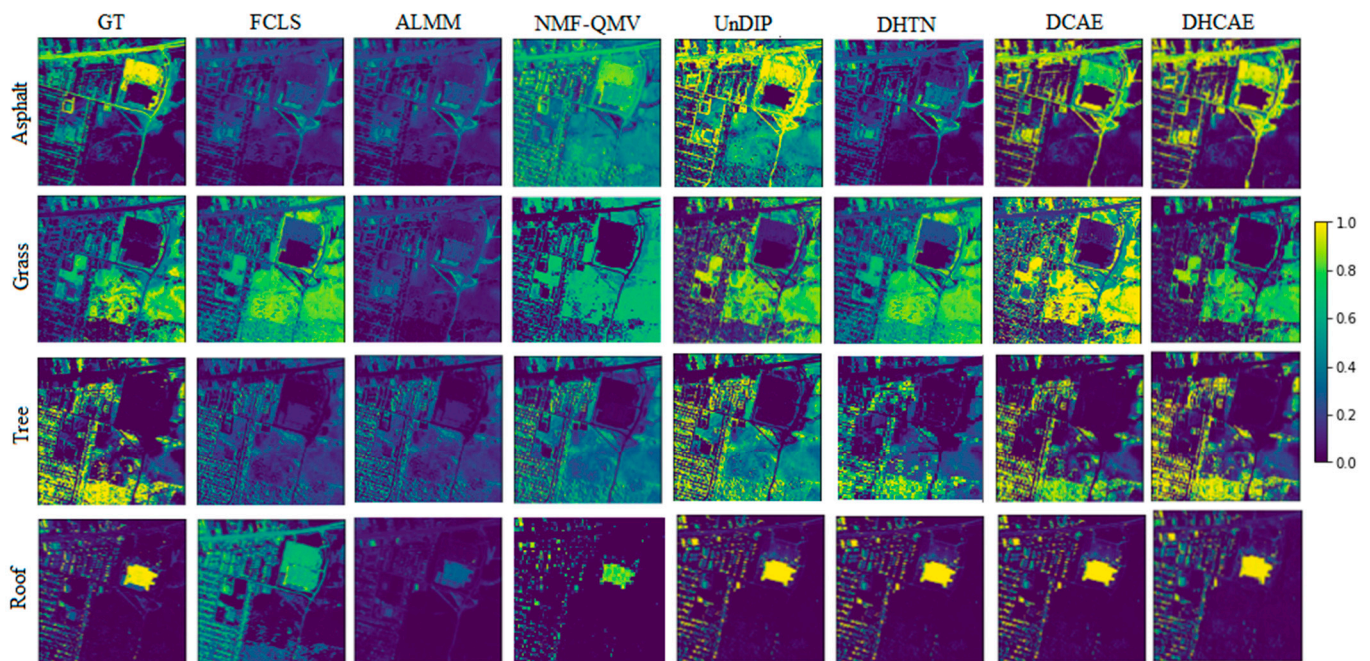


Figure 12. Urban data. Ground truth (GT) map in the first column and extracted abundance maps with remove noisy bands comparison by different spectral unmixing methods.

3.3.2. Results Containing Noisy Bands

We also investigated the robustness of our proposed approach on the urban dataset containing noisy bands, i.e., 210. Table 5 lists the rRMSE and aSAM values yielded by our proposed approach and the other six unmixing methods. As seen in Table 5, the DHCAE network achieved a better result than other competitors. Figure 13 displays the estimated abundance map of various unmixing approaches. When comparing Figures 12 and 13, it can be easily seen that the FCLS and ALMM approaches are still corrupted by noise. We can also observe that the results achieved by our proposed approach are more robust to noise than those achieved by other unmixing competitors.

Table 5. Quantitative performance on the urban dataset containing noisy bands by various approaches. The best result is shown in bold.

Algorithm	FCLS	ALMM	NMF-QMV	UnDIP	DHTN	DCAE	DHCAE
rRMSE	0.1275	0.0987	0.0859	0.0738	0.0723	0.0698	0.0609
aSAM	0.4528	0.3181	0.1624	0.1005	0.997	0.0981	0.0942

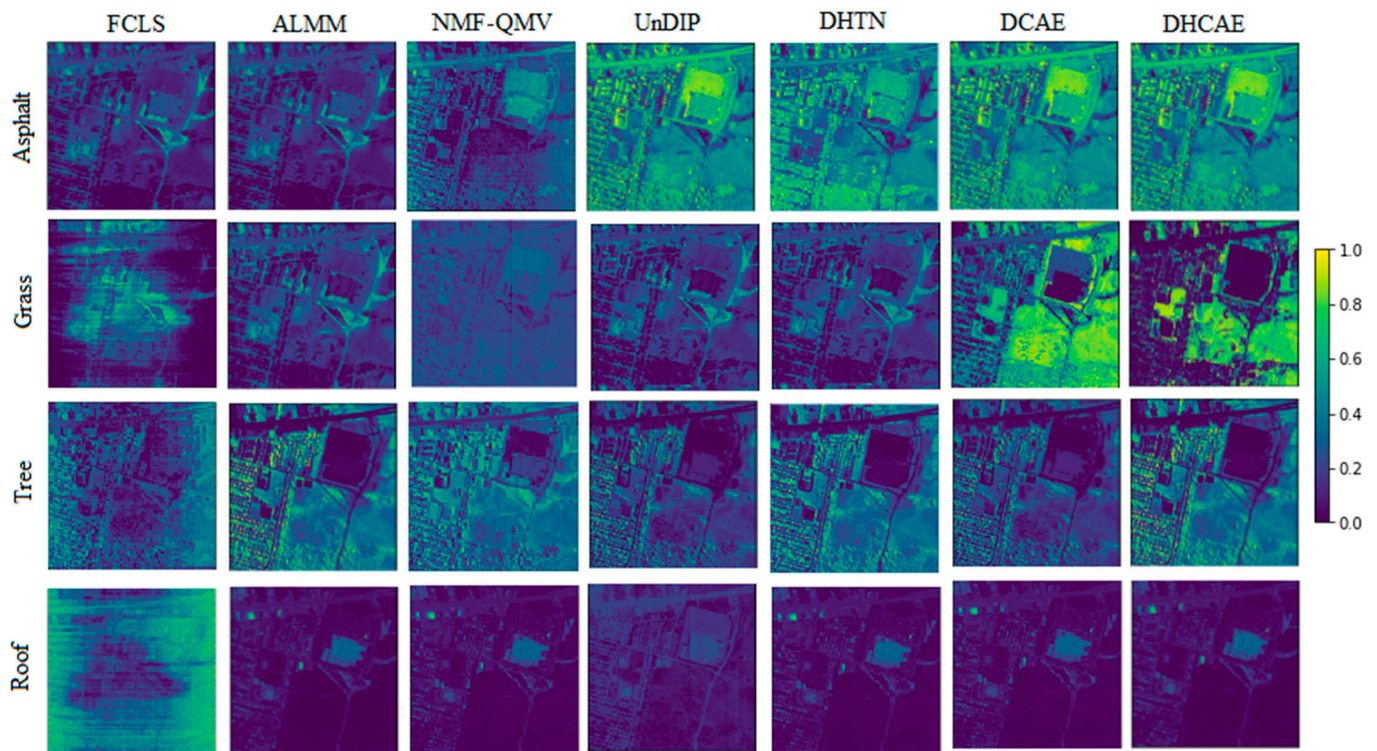


Figure 13. Urban data. Comparison of extracted abundance maps with noisy bands by different spectral unmixing methods.

3.4. Experiments on Washington DC Dataset

The third set of real hyperspectral data was collected by the HYDICE sensor. The observed image comprises 1280×307 pixels with 210 channels ranging from 0.4 to 2.4 μm . Noise and water vapor channels were removed. We investigated a cropped image of size 319×292 pixels with 191 channels. According to [67], six endmembers are included in the Washington DC Mall scene: Grass, Tree, Roof, Road, Water and Trail. The cropped image and six endmembers are shown in Figure 14.

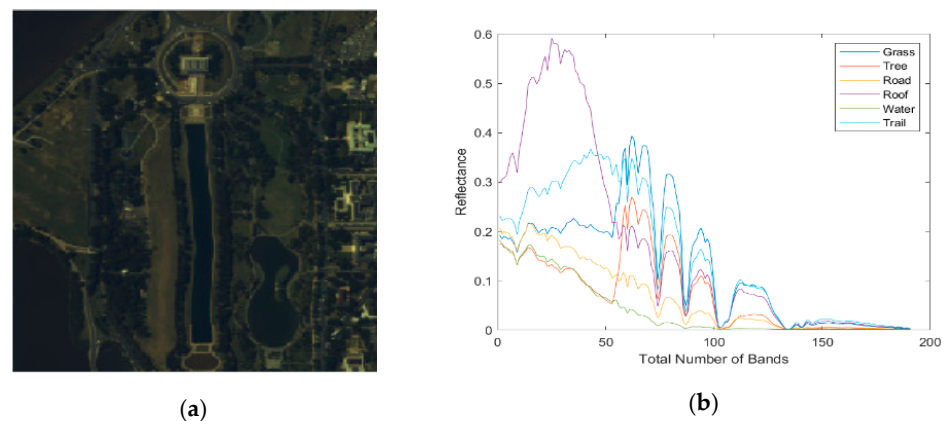


Figure 14. Washington DC Mall hyperspectral image. (a) Color image. (b) Endmembers.

Results on Washington DC Dataset

Table 6 presents the quantitative assessment results for the Washington DC Mall datasets. According to the results in Table 6, the proposed unmixing method provides the best results compared with other unmixing methods. The second-best results are achieved by DHTN and DCAE in term of rRMSE and aSAM. Figure 15 shows the qualitative results of estimated abundance maps for six different spectral signatures. The results clearly indicate that the proposed DHCAE network provides a clear and smoother abundance map.

Table 6. Quantitative performance on the Washington DC Mall dataset by various approaches. The best result is shown in bold.

Algorithm	FCLS	ALMM	NMF-QMV	UnDIP	DHTN	DCAE	DHCAE
rRMSE	0.0713	0.0548	0.0523	0.0474	0.0452	0.0426	0.0357
aSAM	0.1773	0.0967	0.0895	0.0832	0.0786	0.0759	0.0623

3.5. Parameters Setting

In this section, we selected a random search approach to find suitable values for hyperparameters employed in the proposed model. The learning rates were investigated in the range [0.001, 0.003, 0.004, 0.005]; based on the results, the optimal learning rate is 0.005. Similarly, the Adam optimizer was selected to optimize the model among different optimizers. Several experiments were conducted to assess the impact of different dropout values, and the optimal dropout value is 20%. We tried different spatial window sizes [13 × 13, 15 × 15, 17 × 17, 19 × 19] to learn adequate spatial information, and a spatial window size of (15 × 15) was selected. The network was trained for 100 epochs and a minibatch of 100.

The regularization parameters influence the performance of the different unmixing approaches. We discovered that the following parameter values produced the highest results: the parameters in ALMM were to be set as $\alpha = \beta = 0.002$, $\gamma = \eta = 0.005$ and the number of basis vectors $L = 100$. The learning rate was set to 0.001 for the UnDIP; the number of iterations was 3000. The DCAE was optimized using Adam optimizer with a learning rate of 0.005. The number of epochs and batch size were set to 100 each.

3.6. Effects of Spatial Window Size

The effect of spatial window size ($S \times S$) over the unmixing performance of our proposed approach is shown in Figure 16. It is clear that the best estimation was achieved with a spatial window size of 15 × 15. Therefore, we set 15 × 15 as the input size for all the hyperspectral datasets.

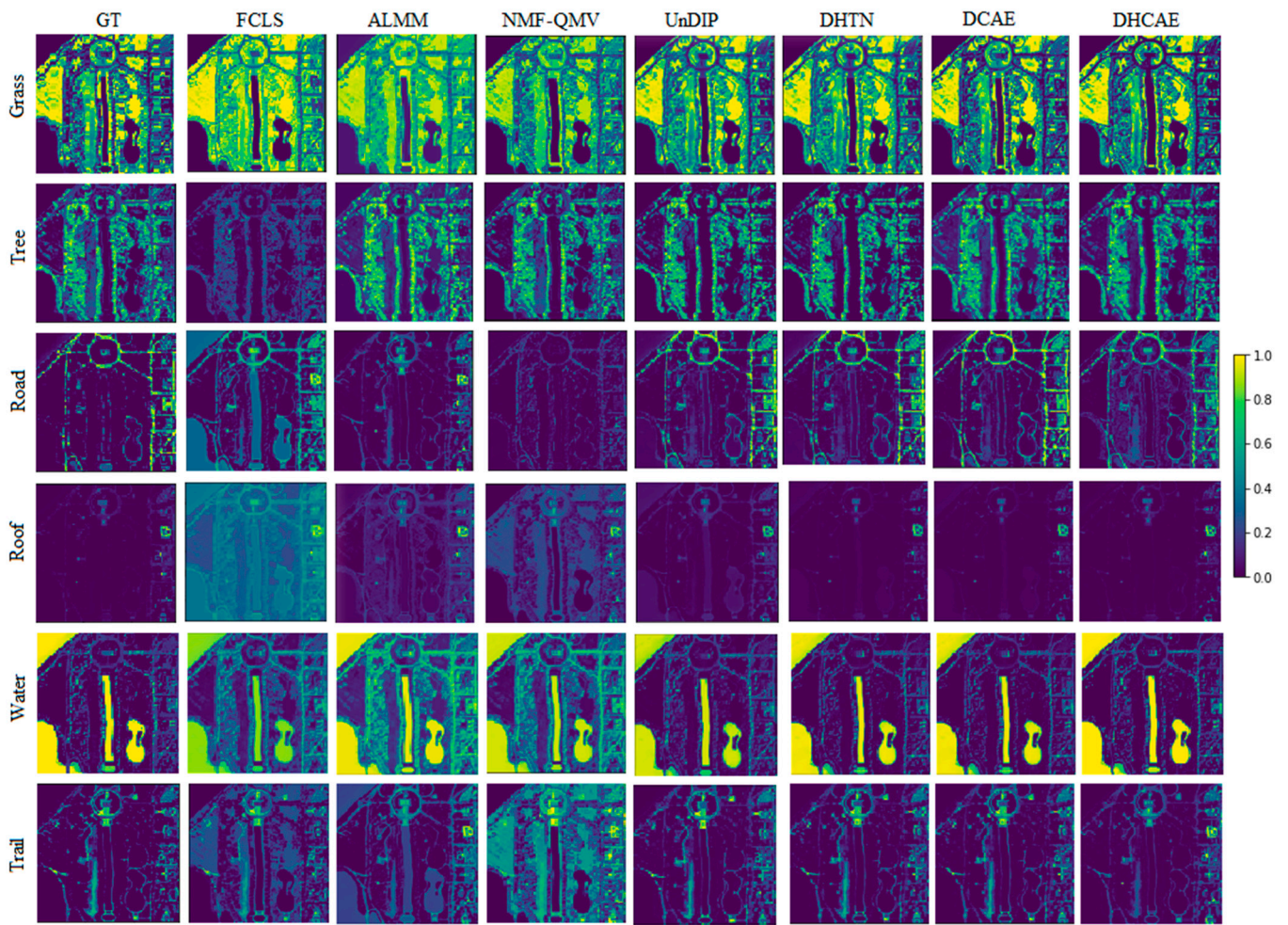


Figure 15. Washington DC Mall data. Ground truth (GT) map in the first column and extracted abundance maps of different spectral unmixing methods.

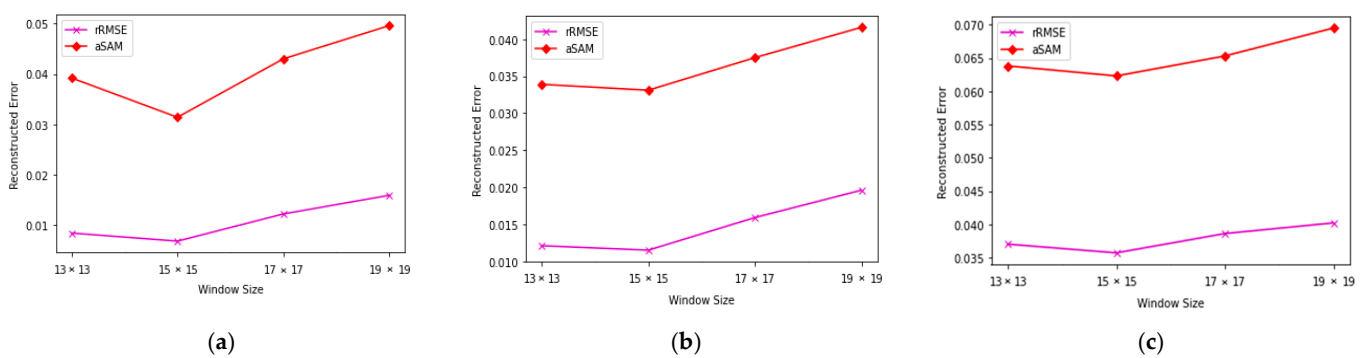


Figure 16. Analysis of spatial window size on the proposed approach. (a) Jasper. (b) Urban. (c) Washington DC Mall.

3.7. Comparative Analysis

To validate the effectiveness of the proposed DHCAE method, we compared it to two existing methods: 3D-CNN and 2D-CNN. Figure 17 depicts the outcomes of the three methods. According to Figure 17, the proposed DHCAE method achieved the best performance in terms of rRMSE and aSAM for each hyperspectral dataset. The proposed method is based on a hierarchical representation of spectral-spatial 3D-CNN followed by a spatial 2D-CNN that is complimentary to it. When compared to hybrid 3D and 2D

convolutions, it is clear that 3D or 2D convolutions alone are incapable of representing the highly discriminative feature.

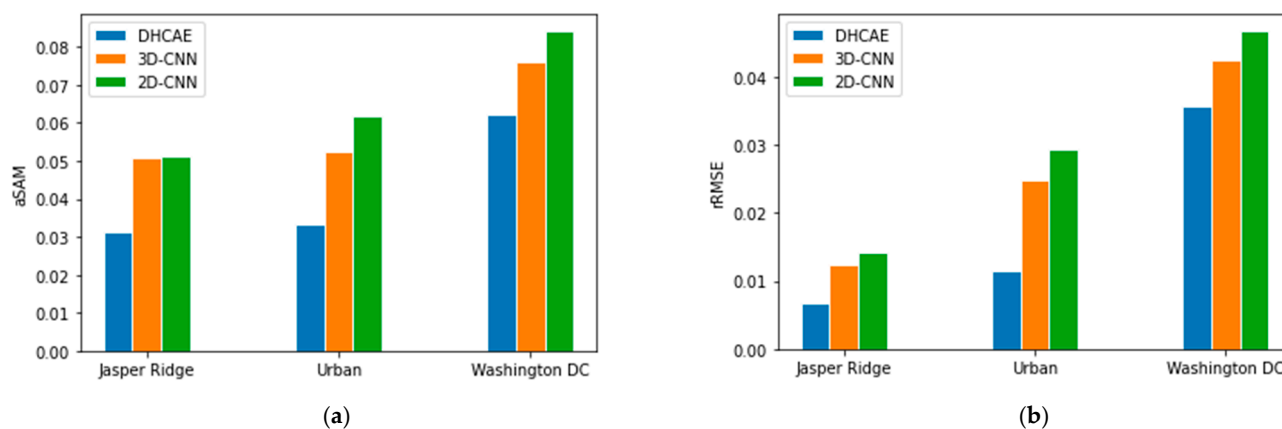


Figure 17. Comparison between DHCAE and other methods. (a) aSAM. (b) rRMSE.

3.8. Computational Time

We discussed the computational time of our proposed DHCAE network and compared it to the computational time of the previous state-of-the-art unmixing approaches. The average computational times in seconds for all the spectral unmixing approaches on three datasets are reported in Table 7. Using the advantage of autoencoder for the unmixing problem consumes computational time. However, the DHCAE network can parallelize on graphical processing units, undoubtedly leading to less computational time.

Table 7. The average time on three datasets.

Algorithm	Computational Time (in Seconds)			
	Synthetic	Jasper	Urban	Washington DC
FCLS	905	240	516	401
ALMM	1410	221	1312	768
UnDIP	1386	475	1253	863
DHTN	1347	452	1231	845
DCAE	1330	427	1210	824
DHCAE	1270	312	1037	693

4. Discussion

The proposed DHCAE utilizes 3D-CNN and 2D-CNN, which are regarded as excellent frameworks for various computer vision tasks, including HSU. However, the proposed DHCAE differs significantly from existing methods. The DHCAE divides spectral and spatial features into two successive modules: 3D-CNN and 2D-CNN. The 3D-CNN module aims to learn robust representations of spectral and spatial features, while the 2D-CNN module learns spatial features. This sequence enables the extraction of more discriminative features and reduces the possibility of information loss. The experimental results confirm that the proposed DHCAE method significantly outperformed the other methods. Its capability to incorporate the spectral and spatial information globally as seen visually from the abundance maps of the three real world and one synthetic remote sensing datasets. The results also show that DHCAE is very robust to noise due to the implicit application of a regularizer in the network. The worst results were obtained by FCLS and ALMM, which are also demonstrated to be very sensitive to noise throughout the experiments. The NMF-QMV unmixing method obtained moderate results and proved more robust to noise. The other approaches (UnDIP, DHTN, DCAE and DHCAE) achieved similar unmixing results to the corresponding ground truth abundance maps. The unmixing results from the UnDIP

and DHTN approaches are better than those from the FCLS, ALMM, and NMF-QMV, demonstrating that spatial information is more critical for HSU. The DCAE model directly extracts the spectral–spatial features from the hyperspectral image but still spends much computational time compared with the DHCAE method. However, in all experiments, the overall computational complexity of the proposed method is lower than the DCAE method.

5. Conclusions

In this paper, we present an end-to-end DHCAE network for robust supervised hyperspectral unmixing. The proposed method combines spectral–spatial and spatial information in the form of 3D and 2D convolutions, respectively. We analyzed experimental results on synthetic scene and three real world hyperspectral scenes. We found that our proposed unmixing method has more robustness against noisy bands and noisy pixels and also obtains accurate abundance map estimation than other unmixing methods. Furthermore, our proposed unmixing approach is less computationally complex than the 3D-CNN models. However, in our future research, we would like to optimize the model architecture and provide a generic spectral–spatial blind remote sensing image unmixing using multi-modal data (e.g., multispectral data, Lidar). In addition, due to the recent increase in the use of transformers in computer vision applications, the question of how to covertly include both the CNN structure and the transformer structure is crucial for future research.

Author Contributions: F.H. Conceptualization and methodology, F.H. performed the experiments. M.U. and I.A. verified the experiments and provided feedback during the work. F.H. wrote the original draft. J.Y. improved the manuscript grammatically, G.F. and L.X. revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the National Natural Science Foundation of China (Grants No. 61871226, 61571230).

Data Availability Statement: The Hyperspectral image datasets used in this study are freely available at <http://lesun.weebly.com/hyperspectral-data-set.html> (accessed on 5 July 2022).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhang, F.; Du, B.; Zhang, L. Scene classification via a gradient boosting random convolutional network framework. *IEEE Trans. Geosci. Remote Sens.* **2015**, *54*, 1793–1802. [[CrossRef](#)]
2. Maggiori, E.; Charpiat, G.; Tarabalka, Y.; Alliez, P. Recurrent neural networks to correct satellite image classification maps. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 4962–4971. [[CrossRef](#)]
3. Goodman, J.A.; Ustin, S.L. Classification of benthic composition in a coral reef environment using spectral unmixing. *J. Appl. Remote Sens.* **2007**, *1*, 011501.
4. Fauvel, M.; Tarabalka, Y.; Benediktsson, J.A.; Chanussot, J.; Tilton, J.C. Advances in spectral-spatial classification of hyperspectral images. *Proc. IEEE* **2012**, *101*, 652–675. [[CrossRef](#)]
5. Villa, A.; Chanussot, J.; Benediktsson, J.A.; Jutten, C. Spectral unmixing for the classification of hyperspectral images at a finer spatial resolution. *IEEE J. Sel. Top. Signal Processing* **2010**, *5*, 521–533. [[CrossRef](#)]
6. Spangler, L.H.; Dobeck, L.M.; Repasky, K.S.; Nehrir, A.R.; Humphries, S.D.; Barr, J.L. A shallow subsurface controlled release facility in Bozeman, Montana, USA, for testing near surface CO₂ detection techniques and transport models. *Environ. Earth Sci.* **2010**, *60*, 227–239. [[CrossRef](#)]
7. Plaza, A.; Du, Q.; Bioucas-Dias, J.M.; Jia, X.; Kruse, F.A. Foreword to the special issue on spectral unmixing of remotely sensed data. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 4103–4110. [[CrossRef](#)]
8. Ozturk, S.; Esin, Y.E.; Artan, Y. Object detection in rural areas using hyperspectral imaging. In *Image and Signal Processing for Remote Sensing XXI*; SPIE: Bellingham, WA, USA, 2015; Volume 9643, pp. 725–731.
9. Huang, X.; Zhang, L. An adaptive mean-shift analysis approach for object extraction and classification from urban hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 4173–4185. [[CrossRef](#)]
10. Valero, S.; Salembier, P.; Chanussot, J. Object recognition in hyperspectral images using binary partition tree representation. *Pattern Recognit. Lett.* **2015**, *56*, 45–51. [[CrossRef](#)]
11. Hong, D.; Yokoya, N.; Chanussot, J.; Xu, J.; Zhu, X.X. Joint and progressive subspace analysis (JPSA) with spatial–spectral manifold alignment for semisupervised hyperspectral dimensionality reduction. *IEEE Trans. Cybern.* **2020**, *51*, 3602–3615. [[CrossRef](#)]
12. Keshava, N.; Mustard, J.F. Spectral unmixing. *IEEE Signal Processing Mag.* **2002**, *19*, 44–57. [[CrossRef](#)]

13. Chi, J.; Crawford, M.M. Spectral unmixing-based crop residue estimation using hyperspectral remote sensing data: A case study at Purdue university. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 2531–2539. [[CrossRef](#)]
14. Hedegaard, M.; Matthäus, C.; Hassing, S.; Krafft, C.; Diem, M.; Popp, J. Spectral unmixing and clustering algorithms for assessment of single cells by Raman microscopic imaging. *Theor. Chem. Acc.* **2011**, *130*, 1249–1260. [[CrossRef](#)]
15. Boardman, J.W. Automating spectral unmixing of AVIRIS data using convex geometry concepts. In Proceedings of the JPL, Summaries of the 4th Annual JPL Airborne Geoscience Workshop. Volume 1: AVIRIS Workshop, Washington, DC, USA, 25–29 October 1993.
16. Winter, M.E. N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data. In *Imaging Spectrometry V*; SPIE: Bellingham, WA, USA, 1999; Volume 3753, pp. 266–275.
17. Nascimento, J.M.; Dias, J.M. Vertex component analysis: A fast algorithm to unmix hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 898–910. [[CrossRef](#)]
18. Li, J.; Agathos, A.; Zaharie, D.; Bioucas-Dias, J.M.; Plaza, A.; Li, X. Minimum volume simplex analysis: A fast algorithm for linear hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 5067–5082.
19. Su, Y.; Li, J.; Plaza, A.; Marinoni, A.; Gamba, P.; Chakravorty, S. DAEN: Deep autoencoder networks for hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 4309–4321. [[CrossRef](#)]
20. Wang, M.; Zhao, M.; Chen, J.; Rahardja, S. Nonlinear unmixing of hyperspectral data via deep autoencoder networks. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 1467–1471. [[CrossRef](#)]
21. Févotte, C.; Dobigeon, N. Nonlinear hyperspectral unmixing with robust nonnegative matrix factorization. *IEEE Trans. Image Processing* **2015**, *24*, 4810–4819. [[CrossRef](#)]
22. Ozkan, S.; Kaya, B.; Akar, G.B. Endnet: Sparse autoencoder network for endmember extraction and hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2018**, *57*, 482–496. [[CrossRef](#)]
23. Ranasinghe, Y.; Herath, S.; Weerasooriya, K.; Ekanayake, M.; Godaliyadda, R.; Ekanayake, P.; Herath, V. Convolutional autoencoder for blind hyperspectral image unmixing. In Proceedings of the 2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS), Rupnagar, India, 26–28 November 2020; pp. 174–179.
24. Rasti, B.; Koirala, B. SUnCNN: Sparse unmixing using unsupervised convolutional neural network. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 5508205. [[CrossRef](#)]
25. Iordache, M.-D.; Plaza, A.; Bioucas-Dias, J. On the use of spectral libraries to perform sparse unmixing of hyperspectral data. In Proceedings of the 2010 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, Reykjavik, Iceland, 14–16 June 2010; pp. 1–4.
26. Zhang, X.; Sun, Y.; Zhang, J.; Wu, P.; Jiao, L. Hyperspectral unmixing via deep convolutional neural networks. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 1755–1759. [[CrossRef](#)]
27. Bioucas-Dias, J.M.; Plaza, A.; Dobigeon, N.; Parente, M.; Du, Q.; Gader, P.; Chanussot, J. Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *5*, 354–379. [[CrossRef](#)]
28. Dobigeon, N.; Moussaoui, S.; Coulon, M.; Tourneret, J.-Y.; Hero, A.O. Joint Bayesian endmember extraction and linear unmixing for hyperspectral imagery. *IEEE Trans. Signal Process.* **2009**, *57*, 4355–4368. [[CrossRef](#)]
29. Iordache, M.-D.; Bioucas-Dias, J.M.; Plaza, A. Collaborative sparse regression for hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2013**, *52*, 341–354. [[CrossRef](#)]
30. Pauca, V.P.; Piper, J.; Plemmons, R.J. Nonnegative matrix factorization for spectral data analysis. *Linear Algebra Its Appl.* **2006**, *416*, 29–47. [[CrossRef](#)]
31. Bioucas-Dias, J.M. A variable splitting augmented Lagrangian approach to linear spectral unmixing. In Proceedings of the 2009 First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, Grenoble, France, 26–28 August 2009; pp. 1–4.
32. Dópido, I.; Plaza, A. Unmixing prior to supervised classification of urban hyperspectral images. In Proceedings of the 2011 Joint Urban Remote Sensing Event, Munich, Germany, 11–13 April 2011; pp. 97–100.
33. Bioucas-Dias, J.M.; Figueiredo, M.A. Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing. In Proceedings of the 2010 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, Reykjavik, Iceland, 14–16 June 2010; pp. 1–4.
34. Boyd, S.; Parikh, N.; Chu, E.; Peleato, B.; Eckstein, J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.* **2011**, *3*, 1–122.
35. Iordache, M.-D.; Bioucas-Dias, J.M.; Plaza, A. Total variation spatial regularization for sparse hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 4484–4502. [[CrossRef](#)]
36. Thouvenin, P.-A.; Dobigeon, N.; Tourneret, J.-Y. Hyperspectral unmixing with spectral variability using a perturbed linear mixing model. *IEEE Trans. Signal Process.* **2015**, *64*, 525–538. [[CrossRef](#)]
37. Drumetz, L.; Veganzones, M.-A.; Henrot, S.; Phlypo, R.; Chanussot, J.; Jutten, C. Blind hyperspectral unmixing using an extended linear mixing model to address spectral variability. *IEEE Trans. Image Process.* **2016**, *25*, 3890–3905. [[CrossRef](#)]
38. Hong, D.; Yokoya, N.; Chanussot, J.; Zhu, X.X. An augmented linear mixing model to address spectral variability for hyperspectral unmixing. *IEEE Trans. Image Process.* **2018**, *28*, 1923–1938. [[CrossRef](#)]

39. Wang, Q.; Wan, J.; Yuan, Y. Locality constraint distance metric learning for traffic congestion detection. *Pattern Recognit.* **2018**, *75*, 272–281. [[CrossRef](#)]
40. Palsson, B.; Sigurdsson, J.; Sveinsson, J.R.; Ulfarsson, M.O. Hyperspectral unmixing using a neural network autoencoder. *IEEE Access* **2018**, *6*, 25646–25656. [[CrossRef](#)]
41. Qu, Y.; Qi, H. uDAS: An untied denoising autoencoder with sparsity for spectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2018**, *57*, 1698–1712. [[CrossRef](#)]
42. Gao, L.; Han, Z.; Hong, D.; Zhang, B.; Chanussot, J. CyCU-Net: Cycle-consistency unmixing network by learning cascaded autoencoders. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 5503914. [[CrossRef](#)]
43. Qi, L.; Li, J.; Wang, Y.; Lei, M.; Gao, X. Deep spectral convolution network for hyperspectral image unmixing with spectral library. *Signal Process.* **2020**, *176*, 107672. [[CrossRef](#)]
44. Dou, Z.; Gao, K.; Zhang, X.; Wang, H.; Wang, J. Blind hyperspectral unmixing using dual branch deep autoencoder with orthogonal sparse prior. In Proceedings of the ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 4–8 May 2020; pp. 2428–2432.
45. Sigurdsson, J.; Ulfarsson, M.O.; Sveinsson, J.R. Blind Hyperspectral Unmixing Using Total Variation and ℓ_p Sparse Regularization. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 6371–6384. [[CrossRef](#)]
46. Imbiriba, T.; Borsoi, R.A.; Bermudez, J.C.M. Low-rank tensor modeling for hyperspectral unmixing accounting for spectral variability. *IEEE Trans. Geosci. Remote Sens.* **2019**, *58*, 1833–1842. [[CrossRef](#)]
47. Zhang, S.; Zhang, G.; Deng, C.; Li, J.; Wang, S.; Wang, J.; Plaza, A. Spectral-Spatial Weighted Sparse Nonnegative Tensor Factorization for Hyperspectral Unmixing. In Proceedings of the IGARSS 2020–2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September 2020–2 October 2020; pp. 2177–2180.
48. Zeng, Y.; Ritz, C.; Zhao, J.; Lan, J. Scattering transform framework for unmixing of hyperspectral data. *Remote Sens.* **2019**, *11*, 2868. [[CrossRef](#)]
49. Qu, Y.; Guo, R.; Qi, H. Spectral unmixing through part-based non-negative constraint denoising autoencoder. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; pp. 209–212.
50. Su, Y.; Marinoni, A.; Li, J.; Plaza, A.; Gamba, P. Nonnegative sparse autoencoder for robust endmember extraction from remotely sensed hyperspectral images. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; pp. 205–208.
51. Su, Y.; Li, J.; Plaza, A.; Marinoni, A.; Gamba, P.; Huang, Y. Deep auto-encoder network for hyperspectral image unmixing. In Proceedings of the IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 6400–6403.
52. Guo, R.; Wang, W.; Qi, H. Hyperspectral image unmixing using autoencoder cascade. In Proceedings of the 2015 7th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Tokyo, Japan, 2–5 June 2015; pp. 1–4.
53. Su, Y.; Marinoni, A.; Li, J.; Plaza, J.; Gamba, P. Stacked nonnegative sparse autoencoders for robust hyperspectral unmixing. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 1427–1431. [[CrossRef](#)]
54. Palsson, B.; Sveinsson, J.R.; Ulfarsson, M.O. Spectral-spatial hyperspectral unmixing using multitask learning. *IEEE Access* **2019**, *7*, 148861–148872. [[CrossRef](#)]
55. Palsson, B.; Ulfarsson, M.O.; Sveinsson, J.R. Convolutional autoencoder for spectral-spatial hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 535–549. [[CrossRef](#)]
56. Hong, D.; Chanussot, J.; Yokoya, N.; Heiden, U.; Heldens, W.; Zhu, X.X. WU-Net: A weakly-supervised unmixing network for remotely sensed hyperspectral imagery. In Proceedings of the IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July 2019–2 August 2019; pp. 373–376.
57. Li, M.; Zhu, F.; Guo, A.J. A Robust Multilinear Mixing Model with l_2, l_1 norm for Unmixing Hyperspectral Images. In Proceedings of the 2020 IEEE International Conference on Visual Communications and Image Processing (VCIP), Macau, China, 1–4 December 2020; pp. 193–196.
58. Zhu, F.; Halimi, A.; Honeine, P.; Chen, B.; Zheng, N. Correntropy maximization via ADMM: Application to robust hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 4944–4955. [[CrossRef](#)]
59. Heinz, D.C. Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 529–545. [[CrossRef](#)]
60. Zhuang, L.; Lin, C.-H.; Figueiredo, M.A.; Bioucas-Dias, J.M. Regularization parameter selection in minimum volume hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 9858–9877. [[CrossRef](#)]
61. Rasti, B.; Koirala, B.; Scheunders, P.; Ghamisi, P. UnDIP: Hyperspectral unmixing using deep image prior. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 5504615. [[CrossRef](#)]
62. Ghosh, P.; Roy, S.K.; Koirala, B.; Rasti, B.; Scheunders, P. Deep hyperspectral unmixing using transformer network. *arXiv* **2022**, arXiv:2203.17076.
63. Khajehrayeni, F.; Ghassemian, H. Hyperspectral unmixing using deep convolutional autoencoders in a supervised scenario. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 567–576. [[CrossRef](#)]
64. Clark, R.N.; Swayze, G.A.; Wise, R.A.; Livo, K.E.; Hoefen, T.M.; Kokaly, R.F.; Sutley, S.J. *USGS Digital Spectral Library Splib06a*; 2327-638X; US Geological Survey: Reston, VA, USA, 2007.

-
65. Huang, R.; Li, X.; Zhao, L. Spectral–spatial robust nonnegative matrix factorization for hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 8235–8254. [[CrossRef](#)]
 66. Li, X.; Huang, R.; Zhao, L. Correntropy-based spatial-spectral robust sparsity-regularized hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 1453–1471. [[CrossRef](#)]
 67. Rasti, B.; Koirala, B.; Scheunders, P.; Chanussot, J. Misticnet: Minimum simplex convolutional network for deep hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5522815. [[CrossRef](#)]