remote sensing

MDPI

Article

Anomaly-Guided Double Autoencoders for Hyperspectral Unmixing

Hongyi Liu [1,*], Chenyang Zhang [1], Jianing Huang [1] and Zhihui Wei [2]

[1] School of Mathematics and Statistics, Nanjing University of Science and Technology, Nanjing 210094, China; hjning@njust.edu.cn (J.H.)
[2] School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China
* Correspondence: hyliu@njust.edu.cn

Abstract: Deep learning has emerged as a prevalent approach for hyperspectral unmixing. However, most existing unmixing methods employ a single network, resulting in moderate estimation errors and less meaningful endmembers and abundances. To address this imitation, this paper proposes a novel double autoencoders-based unmixing method, consisting of an endmember extraction network and an abundance estimation network. In the endmember network, to improve the spectral discrimination, a logarithm spectral angle distance (SAD), integrated with anomaly-guided weight, is developed as the loss function. Specifically, the logarithm function is used to boost the reliability of a pixel based on its high SAD similarity to other pixels. Moreover, the anomaly-guided weight mitigates the influence of outliers. As for the abundance network, a spectral convolutional autoencoder combined with the channel attention module is employed to exploit the spectral features. Additionally, the decoder weight is shared between the two networks to reduce computational complexity. Extensive comparative experiments with state-of-the-art unmixing methods demonstrate that the proposed method achieves superior performance in both endmember extraction and abundance estimation.

Keywords: hyperspectral unmixing; autoencoder network; logarithm SAD; anomaly detection; attention module

check for updates

Academic Editor: Salah Bourennane

Received: 18 January 2025
Revised: 14 February 2025
Accepted: 20 February 2025
Published: 25 February 2025

Citation: Liu, H.; Zhang, C.; Huang, J.; Wei, Z. Anomaly-Guided Double Autoencoders for Hyperspectral Unmixing. *Remote Sens.* **2025**, *17*, 800. https://doi.org/10.3390/rs17050800

Copyright: © 2025 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 sensor can obtain both of the spectral and spatial information of the target scene with a finer resolution, which promotes the development of quantitative remote sensing [1–3]. Nowadays, hyperspectral images are widely used in many geoscience fields including target detection, environmental monitoring, soil survey, and so on [4,5]. However, due to the limited spatial resolution of the hyperspectral sensor and complex real surfaces and light scattering, the hyperspectral image (HSI) often contains mixed pixels composed of multiple spectral features, which has become a challenging problem in hyperspectral image processing and applications [6–8]. Therefore, it is important to extract the basic features (endmembers) in HSI spectral pixels and calculate the corresponding fractions (abundances), which is called hyperspectral unmixing (HU) [9,10].

By considering the mixing model of spectra, linear and nonlinear spectral unmixing are taken into consideration. Most unmixing methods are based on the linear unmixing model (LMM), which assumes that the observed spectra of mixed pixels are a linear combination of pure spectra (endmembers) weighted by the corresponding mixing coefficients (abundances). On the other hand, unmixing methods can be categorized into model-based and deep learning-based (DL-based) methods. The traditional model-based methods include

geometry-based methods, non-negative matrix factorization (NMF), and statistics-based methods. Among these, the geometry-based unmixing methods first extract endmembers under the assumption that endmembers form a simplex in the spectral space. Subsequently, the estimation of abundances can be regarded as a regression problem. Two primary methods for endmember extraction are pixel purity (PP) and minimum volume (MV) methods. Classical PP methods include pixel purity index (PPI), vertex component analysis (VCA) [11], and so on. On the other hand, MV methods aim to obtain endmembers by searching the smallest simplex that encloses all pixel points, such as minimum volume simplex analysis (MVSA) [12]. For abundance estimation, the commonly used algorithm is fully constrained least squares (FCLS) [13], which leverages non-negative matrix factorization (NMF) to simultaneously extract the endmember and abundance matrices from hyperspectral images. However, the decomposition of non-negative matrix is inherently non-unique, resulting in an ill-posed inverse problem. To mitigate this issue, prior information is often incorporated into the endmember and abundance matrices to constrain the solution space [14]. Examples include minimum volume constrained non-negative matrix factorization (MVCNMF) [15], spatial group sparsity regularized non-negative matrix factorization (SGSNMF) [16], and $l_{1/2}$ non-negative matrix factorization ($l_{1/2}$-NMF) [17], etc. Statistical-based unmixing methods, such as the normal compositional model (NCM), beta compositional model (BCM), and hierarchical Bayesian models [18,19] treat abundances and endmembers as estimators derived from maximum a posterior (MAP) estimation. In addition, there is a conceptual link between the NMF and statistical methods [20,21], in which the fidelity term in the NMF model corresponds to a Gaussian likelihood distribution, and the regularization term represents the probability priors of endmembers and abundances, respectively.

Different from the model-based unmixing methods, the DL-based unmixing methods depend on the deep neural network (DNN), which directly learns hidden features by inputting a large number of training samples [22–26]. Autoencoder networks are an effective and simple unsupervised learning method, which can jointly learn endmembers and abundances from HSI by encoding and decoding. As a result, autoencoder [27–31] along with sparse autoencoder (SAE), stacked denoising autoencoder (SDAE) [32,33], and variational autoencoder (VAE) [34] has been extensively researched and demonstrated to be highly effective and convenient for unmixing. Specifically, network architectures and loss functions are the most concerned in the autoencoder-based unmixing networks. The fidelity term of the loss function is mainly based on the similarity measure between the observed and reconstructed HSI data, such as spectral angular distance (SAD) [34–36], mean squared error (MSE) [32,37], spectral information divergence (SID) [38], Wasserstein distance [39], and hyper-Laplace distribution [40]. Meanwhile, full connection (FC), activation function, and batch normalization (BN) are designed based on the motivations of the unmixing networks. EndNet [41] uses the shallow sparse autoencoder network with Kullback–Leibler (KL) divergence loss for unmixing, in which the SAD function is applied in the encoder layer to increase the nonlinearity and spectral similarity. The abundance orthogonal sparse regularization (OSPAEU) and the endmember initialization based on outlier removal were deeply exploited in [40]. To enhance spatial information utilization, various advanced techniques have been developed, including patch-based or neighborhood-based convolution networks (CNNs), deep image prior, transformer-based architectures. For instance, based on autoencoder network, hyperspectral patches were extracted and used to train the unmixing network, taking the spatial–contextual information into consideration (CNNAEU) [42]. In [43], multiple unmixing autoencoders were performed on a single pixel in a neighborhood by shared weights (MTAEU). Additionally, leveraging deep image prior, abundances were estimated by a deep network as explored in [44,45]. Moreover, the

self-attention mechanism was integrated into the networks to simultaneously guide the unmixing and reconstruction processes [46]. These diverse techniques have significantly advanced the field of hypersectral unmixing.

Despite these remarkable advancements, a common drawback persists. Autoencoder-based unmixing methods using a single network often succeed in generating high-quality endmembers but suffer from the lower precision of the estimated abundances. This issue arises because the unmixing process inherently involves a trade-off between the fidelity of endmembers and the accuracy of abundances. Balancing these two aspects remains a critical challenge in achieving optimal unmixing performance. To address this, a more effective approach is to employ two separate networks: one dedicated to learning and extracting endmembers, and the other focused on precisely estimating abundances [33,34,42,47–49]. This strategy not only alleviates the trade-off but also enhances the physical significance and interpretability of the endmembers. By decoupling these tasks, the method can achieve a more comprehensive and accurate unmixing outcome. Furthermore, it should be mentioned that endmembers can be viewed as a shared basis for the entire hyperspectral image, making them easier to learn. In contrast, abundances represent a unique mapping for each individual pixel, posing a substantial challenge for learning. Inspired by these fundamental differences between endmembers and abundances, we use two networks for hyperspectral unmixing. Moreover, to improve the robustness of the network, we introduce a new loss function that combines logarithm SAD with an anomaly-guided mechanism. This helps handle outliers and anomalies present in the data, thereby improving the overall stability and performance. The contributions of our work are summarized as follows:

1. Two autoencoders-based unmixing networks with different architectures and loss functions are designed to extract endmember and estimate abundance. In addition, to enhance the accuracy of endmember extraction, the weight of the AE-Net is shared from the EE-Net, ensuring a more consistent and precise unmixing.
2. Rather than the commonly used loss function, the logarithm SAD is adopted as a loss function in the EE-Net, which improves the network's learning ability by leveraging the guidance from "good" pixels with high spectral similarity. Additionally, an anomaly-based weight is integrated into the logarithm SAD loss to further enhance the robustness of the EE-Net.
3. The AE-Net is constructed using 1D spectral convolution and attention module. Each hidden layer is incorporated with an attention module, enabling the network to capture more discriminative spatial–spectral features, improving its ability to estimate abundances accurately.

The rest of the paper is organized as follow. Section 2 gives a brief introduction of the autoencoder network for hyperspectral unmixing. Section 3 describes the details of the proposed unmixing ADAEU network, including the architecture and the loss function of the EE-Net and AE-Net. The experimental results and analysis are presented in Section 4. Finally, the conclusions are given in Section 5.

2. Autoencoder-Based Unmixing Network

Let $\mathbf{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_N] \in R^{B\times N}$ be the HSI with N pixels, each $\boldsymbol{x_i} \in R^{B\times 1}$, and B be the number of bands. Based on the LMM, $\mathbf{X}$ can be represented as

$$\mathbf{X} = \mathbf{MA} + \mathbf{U} \tag{1}$$

where $\mathbf{M} = [\boldsymbol{m}_1, \boldsymbol{m}_2, ..., \boldsymbol{m}_R] \in R^{B\times R}$ is the endmember matrix with R endmembers, $\mathbf{U} \in R^{B\times N}$ is the Gaussian noise, and $\mathbf{A} \in R^{R\times N}$ is a matrix composed of abundances, which often satisfy non-negative and sum-to-one constraints.

Autoencoder is an artificial neural network used in unsupervised learning, which is composed of input, hidden, and output layers. In the autoencoder network for unmixing, the spectral curve $\boldsymbol{x_i}$ of HSI is put into the network, and the encoder layer maps it to the hidden layer:

$$\boldsymbol{h}_i = \sigma(\mathbf{W}_e^T \boldsymbol{x}_i + \boldsymbol{b}_e) \tag{2}$$

where $\mathbf{W}_e$ and $\boldsymbol{b}_e$ are the weights and biases of the encoder layer, respectively, and σ is an activation function.

For the decoder layer in unmixing, a decoder weight $\mathbf{W}_d$ is multiplied with the hidden layer $\boldsymbol{h_i}$, resulting in the reconstructed spectrum $\hat{x}_i$ as follows:

$$\hat{\boldsymbol{x}}_i = \mathbf{W}_d^T \boldsymbol{h}_i \tag{3}$$

Next, the mean least square is commonly used as a loss function

$$L = \frac{1}{n} \sum_{i=1}^{n} \|\boldsymbol{x}_i - \hat{\boldsymbol{x}}_i\|_2^2 \tag{4}$$

After encoding and decoding, the obtained $\mathbf{W}_d$ and $\boldsymbol{h}_i$ are the estimated endmembers and abundance, respectively.

3. Proposed Unmixing Method

In this section, we will give the details of the proposed anomaly-guided double autoencoders unmixing method (ADAEU) step by step. Figure 1 shows the flowchart of the ADAEU unmixing networks. There are two autoencoder networks: the first stage is the endmember extraction network (EE-Net), then it is followed with the abundance estimation network (AE-Net). Moreover, in the AE-Net stage, the endmembers estimated by EE-Net are used as weights of the decoder layer, without being updated during the iteration.

3.1. Endmember Extract Network (EE-Net)

3.1.1. The Architecture of the EE-Net

The endmember extraction network (EE-Net) is shown in Figure 1, using a deep autoencoder network to extract effective features. The network includes two parts, the encoder and the decoder. The encoder layer is mainly composed of four FC layers and five LeakyReLU activation functions, and the decoder is composed of one FC layer without an activation function or bias.

Let $\boldsymbol{x} \in R^{B \times 1}$ denote the input hyperspectral data, then each hidden layer $\boldsymbol{h}_i$ can be represented as

$$\boldsymbol{h}_i = LeakyReLU(\mathbf{W_i} \boldsymbol{h}_{i-1} + b_i)(i = 1, 2, 3, 4) \tag{5}$$

where $\boldsymbol{h}_0$ is the input of EEN with $\boldsymbol{h}_0 = \boldsymbol{x}$, and $\boldsymbol{h}_2$, $\boldsymbol{h}_3$, $\boldsymbol{h}_4$ are the output of the i-th layer with size of $\frac{\mathbf{t_{i-1}}}{2} \times 1$, i=2, 3, 4. $\mathbf{W_i}$, b_i are the weights and biases of the i-th layer, respectively.

After $\boldsymbol{h}_4$, a batch normalization (BN) layer is performed to accelerate training and mitigate the vanishing gradients, resulting in the fifth layer $\boldsymbol{h}$ with

$$\boldsymbol{h} = LeakyReLU(BN(\boldsymbol{h}_4)) + \rho \tag{6}$$

where ρ is a random noisy term, which can not only accelerate the convergence speed of the network but also improve the accuracy of the endmember.

Furthermore, in order to ensure the sparsity and non-negativity of abundances, truncated ReLU is used as

$$\boldsymbol{z} = ReLU(\boldsymbol{h} - k) \tag{7}$$

where the initial value of the parameter k is 0.5, and it is updated during the network training.

Finally, z is normalized to obtain a rough estimated abundance $\widetilde{\mathbf{y}}$, which meets the sum-to-one constraint (ASC) as follows:

$$\widetilde{\mathbf{y}} = \frac{z}{\|z\|_1 + \varepsilon} \tag{8}$$

The reconstructed $\hat{\boldsymbol{x}}$ can be obtained by

$$\hat{\boldsymbol{x}} = \mathbf{W}\widetilde{\boldsymbol{y}} \tag{9}$$

where $\mathbf{W}$ is the estimated endmember matrix.

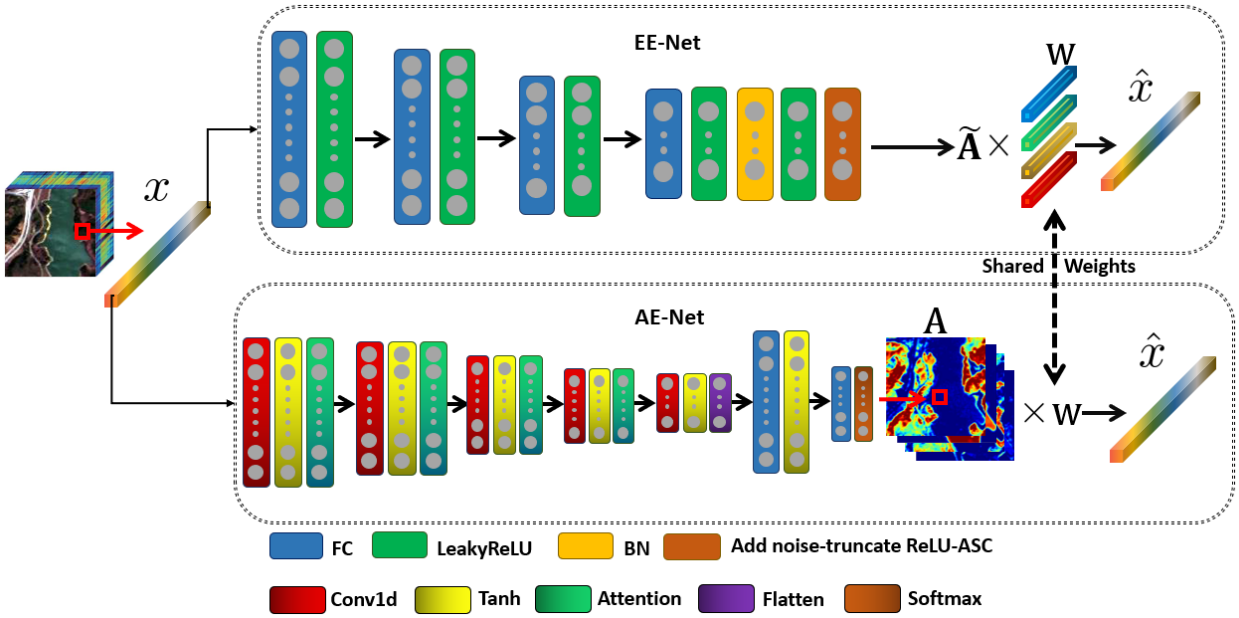

Figure 1. The architecture of the proposed ADAEU network, including the endmember extraction network (EE-Net) and abundance estimation network (AE-Net).

3.1.2. The Anomaly-Guided Loss Function of EE-Net

For EE-Net, the SAD function is commonly used as part of the loss function due to their effectiveness in measuring the similarity of the spectrum and its estimator. The definition of SAD is

$$SAD(\boldsymbol{x}_i, \hat{\boldsymbol{x}}_i) = arccos(\frac{\boldsymbol{x}_i^T \hat{\boldsymbol{x}}_i}{\|\boldsymbol{x}_i\|_2 \|\hat{\boldsymbol{x}}_i\|_2}) \tag{10}$$

where $\boldsymbol{x_i}$ and $\hat{\boldsymbol{x}}_i$ are the i-th original spectral signature and the reconstructed one.

However, there are two other aspects should also be taken into consideration in the loss function: (1) ensuring that the matching error of $\boldsymbol{x}_i$ and $\hat{\boldsymbol{x}}_i$ guides the energy or loss function to obtain the minimum value; and (2) reducing the impact of outliers on the loss function. Next, we are going to focus on these issues.

(1) For the first issue, the points with smaller matching errors are considered "good points", playing a predominant role in the endmember estimation and improving the network's matching performance. On the other hand, the points with large errors are taken as "bad points", and their contribution to the loss function should be reduced. Therefore, in the loss function, it is desired to assign larger weights to the "good points" with smaller errors to guide the network to learn more essential and accurate features. One practical way is to estimate the weights adaptively by the network itself, in which the weights are

inversely proportional to the errors. Since SAD represents the error, $\frac{1}{SAD}$ can be used as a weight for the derivative of the loss function, resulting in $\frac{1}{SAD}[\nabla_{x_i} SAD]$. It can be interpreted that the weight increases the proportion of "good points" with small errors when minimizing the loss function. In addition, it should be noted that $\frac{1}{SAD}[\nabla_{x_i} SAD]$ is the derivative of $log(SAD)$. Therefore, the $log(SAD)$ is used as a part of the loss function.

(2) For the second issue, to reduce the impact of the anomaly points, the outliers are detected firstly, then the anomaly-related weights are added to the loss function. Therefore, based on the Reed–Xiaoli (RX) anomaly detection method [50], the anomaly-based weighting coefficients can be defined as

$$\mathbf{c_x} = 1 - \frac{r_x}{max(r_x)} \tag{11}$$

where $r_x = (\boldsymbol{x} - \mu)^T \Sigma^{-1} (\boldsymbol{x} - \mu)$ is the Mahalanobis distance measuring the anomaly of the point x, μ and Σ are the mean and variance with $\mu = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{x}_i$ and $\Sigma = \frac{1}{N}\sum_{i=1}^{N} (\boldsymbol{x_i} - \mu)(\boldsymbol{x_i} - \mu)^T$, respectively.

From the definition (11), a smaller weight coefficient $\mathbf{c_x}$ corresponds to a larger r_x, which means the outliers are suppressed, improving the robustness of the network. Moreover, the anomaly weight is computed related to the neighborhood, thus effectively introducing spatial information into spectral identification.

Finally, combing the two data-related issues with the SAD function, the proposed anomaly-guided loss function of EE-Net is formulated as

$$L_{EE} = \frac{1}{N}\sum_{i=1}^{N} log(c_{x_i} SAD(\boldsymbol{x}_i, \hat{\boldsymbol{x}}_i)) + \mu_1 \|\rho\|_F + \mu_2 \|\mathbf{W}\|_F \tag{12}$$

where the first term is the fidelity term with a log(SAD) to measure the similarity of the spectral curves. The last two terms constrain the smoothness of ρ and $\mathbf{W}$ to prevent overfitting. μ_1 and μ_2 are regularization parameters.

3.2. Abundance Estimation Network (AE-Net)

The main architecture of AE-Net is also autoencoder based, including the one-dimensional spectral convolution, FC, attention and tanh function. To increase the generalization, a noise term is added to the network. Moreover, due to the strong redundancy of hyperspectral images, attention mechanisms are introduced to enable the network to learn more important features. Specifically, the encoding layer of the network is mainly composed of five one-dimensional convolutional layers, two full connection layers, and the tanh activation function and spectral attention layer. The decoding layer consists of a single full connected layer without an activation function or offset term. The weight of the decoding layer is determined by the estimated endmember obtained from the endmember extraction network, and it is not updated with the network training.

In formulation, the hidden nodes in the encoding layer are represented as

$$\begin{aligned} \boldsymbol{t_1} &= Tanh(Conv_1(\boldsymbol{x})) \\ \boldsymbol{t_i} &= Tanh(Conv_i(Attention_{i-1} \odot \boldsymbol{t_{i-1}}))(i = 2, 3, 4, 5) \end{aligned} \tag{13}$$

where $Tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$, $\boldsymbol{t}_i$ is the output of the i-th layer. The kernel size of the convolution operator is of 1×3 with stride 1. The spectral attention module refers to the spatial attention part of the Convolutional Block Attention Module (CBAM) [51]. For $\boldsymbol{t_i} \in R^{1 \times B_i \times C_i}$, the output of the attention module $\boldsymbol{At_i} \in R^{1 \times B_i \times 2}$ can be written as

$$\boldsymbol{At_i} = \sigma(f^{1 \times 7}[Avgpolling(\boldsymbol{t_i}); Maxpooling(\boldsymbol{t_i})]) \tag{14}$$

where average pooling $Avgpolling(\boldsymbol{h_i}) \in R^{1\times B_i\times 1}$ and maximum pooling $Maxpooling(\boldsymbol{h_i}) \in R^{1\times B_i\times 1}$ are concatenated along the channel dimension, and $f^{1\times 7}$ is a 1×7 convolutional operator with a two-dimensional input channel and one-dimensional output channel.

After that, the obtained features are represented as follows

$$\mathbf{z} = Tanh(Fc1(Flatten(\boldsymbol{t_5}))) \tag{15}$$

Then the final estimated abundance $\boldsymbol{y}$ is obtained as

$$\boldsymbol{y} = Softmax(Fc2(\boldsymbol{z})) \tag{16}$$

Finally, the decoder layer reconstructs the input as

$$\hat{\boldsymbol{x}} = (\mathbf{W} + \mathbf{W}_{\boldsymbol{err}})\boldsymbol{y} \tag{17}$$

where $\mathbf{W}$ is the estimated emdmember from the EE-Net, and $\mathbf{W_{err}}$ is the random noise used to increase the generalization of the network.

Consider that in AE-Net, each pixel is needed to estimate its abundance whether it is an outlier or not, so rather than log(SAD), the SAD function is employed as follows:

$$L_{AE} = \frac{1}{N}\sum_{i=1}^{N} SAD(\boldsymbol{x_i}, \hat{\boldsymbol{x}}_{\boldsymbol{i}}) \tag{18}$$

4. Experiments Results and Discussions

In this section, we compare the proposed ADAEU method with several state-of-the-art unmixing methods, including VCA [11], SGSNMF [16], CNNAEU [42], OSPAEU [40], MTAEU [43], ENDNET [41], PGMSU [34], and A2SAN [46]. Among them, VCA, associated with FCLS [13], is a pure endmember-based method, SGSNMF is a non-negative matrix decomposition method, and others are autoencoder-based unmixing methods. The parameters of the other methods refer to those of the original ones, and Table 1 gives the network parameters in the proposed method.

Table 1. The parameters of the ADAEU network.

Layer Names	Conv1	Conv2	Conv3	Conv4	Conv5	FC1	Fc2
Feature maps	2	4	8	12	16	120	N

Three datasets are used in the experiments (Figure 2). The first dataset is the Jasper Ridge dataset, which is captured by the airborne visible infrared imaging spectrometer (AVIRIS) sensor. It contains 512×614 pixels and 224 bands covering the 380–2500 nm wave length range. Bands 1–3, 108–112, 154–166, and 220–224 are removed due to water absorption and atmospheric effects, leaving 198 spectral bands for experiment. A 100×100 subimage is selected from the original Jasper Ridge dataset, and four main endmembers are observed: tree, water, soil, and road. The second dataset is the Urban dataset, which is obtained by the hyperspectral digital image collection experiment (HYDICE) sensor and contains 307×307 pixels and 210 bands covering the 400–2500 nm wave length range. Bands 1–4, 76, 87, 101–111, 136–153, and 198–210 are removed due to water absorption and atmospheric effects, leaving 162 spectral bands for experiments. In the Urban dataset, four main endmembers are observed: asphalt, grass, tree, and roof. The third dataset is the Samson dataset, which is acquired by the SAMSON sensor. It is a 95×95 subimage of a large HSI, and contains 156 bands covering the 401–889 nm wave

length range. In the Samson dataset, three main endmembers are observed: soil, tree, and water.



(a) (**b**) (**c**)

Figure 2. The datasets: (**a**) Jasper Ridge. (**b**) Urban. (**c**) Samson.

4.1. Assessment Indexes

Two indexes, spectral angle distance (SAD) and root mean square error (RMSE), are employed to give a quantitative assessment. The mean SAD is defined as

$$SAD = \frac{1}{R}\sum_{r=1}^{R} arccos\left(\frac{m_r^T \hat{m}_r}{\|m_r\|_2 \|\hat{m}_r\|_2}\right) \tag{19}$$

where $\boldsymbol{m_r}$ and $\hat{\boldsymbol{m}}_{\boldsymbol{r}}$ are the true endmember and the extracted endmember, respectively. R is the number of endmembers.

The RMSE measures the error between the abundance map $\boldsymbol{A_r}$ and the estimated one $\hat{\boldsymbol{A}}_{\boldsymbol{r}}$ as

$$RMSE(\boldsymbol{A}) = \sqrt{\frac{1}{R}\|\boldsymbol{A_r} - \hat{\boldsymbol{A}}_{\boldsymbol{r}}\|_F^2} \tag{20}$$

4.2. Experiment with Jasper Ridge Dataset

Table 2 shows the SAD and RMSE indexes on the Jasper Ridge datasets. Generally, VCA exhibits the poorest performance, likely due to the unrealistic pure pixel assumption. Similarly, SGSNMF shows high SAD and RMSE values, which can be attributed to its square-minimization loss function without scale invariance. Compared with the classical unmixing methods, the learning-based methods significantly improve the unmixing performance, particularly in endmember extraction. Specifically, CNNAEU, OSPAEU and the proposed ADAEU achieve SAD values below 0.1, showing high precision in endmember estimation. Among them, OSPAEU, leveraging the orthogonal constraints, ranks third in endmember estimation accuracy. Although ENDNET has the lowest SAD for the first endmember *Tree*, its performance degrades significantly for the remaining endmembers compared to ADAEU. Considering the spectral variability, PGMSU performs suboptimally in both RMSE and SAD metrics, especially exhibiting the lowest SAD for the third endmember *Soil*. A2SAN delivers moderate performance in both SAD and RMSE. Notably, the proposed ADAEU method far surpasses all others. It achieves an average SAD of 0.0364 and RMSE of 0.0421, which are 0.0247 and 0.049 lower than the suboptimal PGMSU.

Figures 3 and 4 present the visualizations of the endmember and abundance estimation results on the Jasper Ridge dataset. Figure 3 shows that all unmixing methods are capable of approximating endmembers of *tree* and *water* with a high degree of accuracy. However, estimating the last two endmembers (i.e., *soil* and *road*), especially endmember *road*, remains challenging. For instance, ENDNET achieves near-optimal estimation for the first two endmembers but incurs significant errors for the fourth endmember. In comparison, the proposed method exhibits nearly overlapped curves with the real ones, particularly

for the first, second, and the last endmembers, outperforming the suboptimal PGMSU method. Regarding abundance estimation, as shown in Figure 4, VCA and SGSNMF also exhibit significant abundance estimation errors for endmembers *tree* and *water*. CNNAEU, OSPAEU, and MTAEU perform worse. The limitations of CNNAEU may stem from its patch-based method, which introduces neighboring pixels into the abundance estimation. The performance of OSPAEU is likely hampered by its strict assumption of abundance orthogonality, while MTAEU is degraded by its multitask framework. Additionally, PGMSU performs closely to the proposed methods but has large abundance estimation errors for the first and second endmembers. Although A2SAN accurately estimates the abundance map of the first endmember, it suffers from large errors for other endmembers. In contrast, the proposed ADAEU yields accurate abundance estimations, closely approximating the ground truth.

Table 2. Quantitative results for the Jasper Ridge dataset. The best results are shown in bold, and the second-best results are in blue.

Methods		VCA	SGSNMF	CNNAEU	OSPAEU	MTAEU	ENDNET	PGMSU	A2SAN	ADAEU
SAD	Tree	0.1409	0.1320	0.0883	0.0872	0.0477	**0.0301**	0.0622	0.1295	0.0406
	Water	0.2851	0.2078	0.0382	0.0640	0.0647	0.0338	0.0138	0.1452	**0.0286**
	Soil	0.1278	0.1665	0.0928	0.0671	0.1300	0.0815	**0.0194**	0.1134	0.0535
	Road	0.0889	0.1359	0.1526	0.0581	0.2950	0.3415	0.0238	0.1002	**0.0228**
Mean SAD		0.1607	0.1606	0.0930	0.0691	0.1348	0.1217	0.0611	0.1221	**0.0364**
RMSE	Tree	0.1593	0.1290	0.2224	0.2667	0.1856	0.0818	0.1031	0.0970	**0.0360**
	Water	0.2081	0.1784	0.2027	0.1011	0.0481	0.1225	0.1179	0.1065	**0.0203**
	Soil	0.1386	0.1491	0.3655	0.2767	0.2665	0.1574	0.0834	0.1437	**0.0672**
	Road	0.1152	0.1336	0.2485	0.1222	0.2542	0.2105	0.0599	0.1305	**0.0447**
Mean RMSE		0.1521	0.1475	0.2598	0.1916	0.1886	0.1430	0.0911	0.1194	**0.0421**

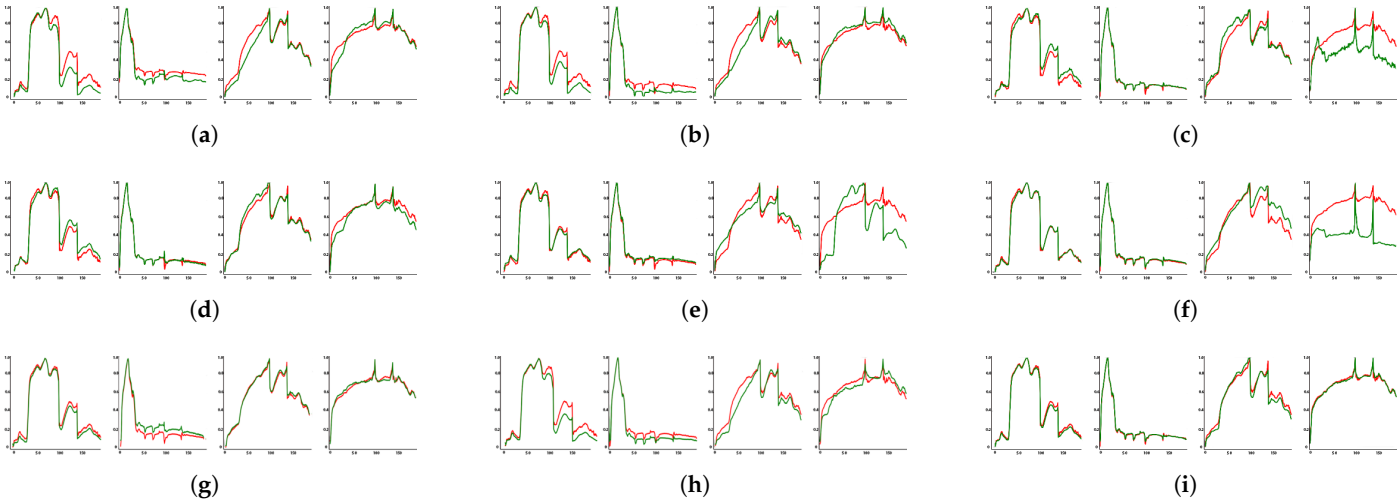

Figure 3. Endmembers on Jasper Ridge dataset: (**a**) VCA. (**b**) SGSNMF. (**c**) CNNAEU. (**d**) OSPAEU. (**e**) MTAEU. (**f**) ENDNET. (**g**) PGMSU. (**h**) A2SAN. (**i**) ADAEU.

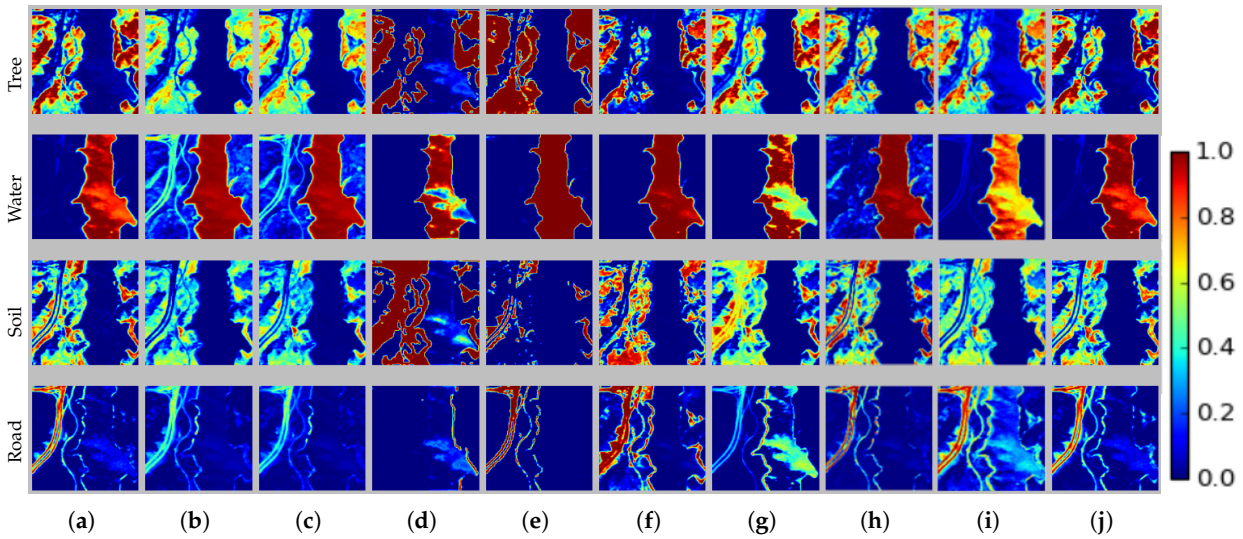

Figure 4. Abundance maps on the Jasper Ridge dataset: (**a**) GT. (**b**) VCA. (**c**) SGSNMF. (**d**) CNNAEU. (**e**) OSPAEU. (**f**) MTAEU. (**g**) ENDNET. (**h**) PGMSU. (**i**) A2SAN. (**j**) ADAEU.

4.3. Experiment with Urban Dataset

Table 3 exhibits the assessment indexes for endmember and abundance on the Urban dataset. Due to the complexity of the Urban imaging scenario, the unmixing results are markedly inferior to those of the other two datasets as evidenced by higher SAD and RMSE indexes. For instance, the traditional methods exhibit relatively large SAD values exceeding 0.35 and RMSE values above 0.28. On the other hand, the DL-based methods significantly reduce these metrics, achieving SAD and RMSE values as low as 0.1. In detail, CNNAEU demonstrates a strong ability for endmember extraction, attaining the smallest SAD for the first endmember *asphalt* and the second-lowest average SAD overall. Nevertheless, it struggles with abundance estimation, showing larger errors in this regard. PGMSU performs worse on the Urban dataset compared to its results on the Jasper Ridge image, with higher SAD and RMSE values. A2SAN achieves suboptimal performance in abundance estimation as indicated by its RMSE values. The proposed ADAEU stands out, achieving a significantly lower SAD of 0.0357 and RMSE of 0.0655, making a notable improvement over the other methods.

Table 3. Quantitative results for the Urban dataset. The best results are shown in bold, and the second-best results are in blue.

Methods		VCA	SGSNMF	CNNAEU	OSPAEU	MTAEU	ENDNET	PGMSU	A2SAN	ADAEU
SAD	Asphalt	0.2572	0.1844	**0.0526**	0.2665	0.0779	0.0809	0.0863	0.1113	0.0740
	Grass	0.9983	0.9591	0.0413	0.1260	0.0688	0.0699	0.1113	0.0494	**0.0314**
	Tree	0.3689	0.0931	0.0389	0.0400	0.0403	0.0303	0.0439	0.0557	**0.0251**
	Roof	0.0889	0.1988	0.0363	0.2532	0.0666	0.0586	0.0564	0.1337	**0.0122**
Mean SAD		0.4492	0.3589	0.0423	0.1714	0.0639	0.0599	0.0745	0.0875	**0.0357**
RMSE	Asphalt	0.2396	0.2790	0.2652	0.2396	0.1517	0.1561	0.1440	0.1145	**0.0843**
	Grass	0.4413	0.4405	0.3200	0.3090	0.1498	0.1308	0.1673	0.1014	**0.0749**
	Tree	0.2924	0.2568	0.2368	0.2187	0.0822	0.1246	0.1692	0.0982	**0.0490**
	Roof	0.2156	0.1795	0.1623	0.1248	0.0899	0.1013	0.1137	0.0870	**0.0539**
Mean RMSE		0.2972	0.2889	0.2461	0.2230	0.1226	0.1282	0.1485	0.1003	**0.0655**

In terms of spectral curves and abundance maps as shown in Figures 5 and 6, it can be seen that for the second and fourth endmembers, VCA and SGSNMF have a large deviation from the real ones. Among the DL-based methods, CNNAEU, MATAEU, ENDNET, PGMSU, and ADAEU exhibit high similarity to the ground truth in their spectral curves. The pro-

posed ADAEU consistently takes the first place with the smallest errors. The same conclusion can be drawn from Figure 6 regarding abundance estimation, further validating that the proposed ADAEU is highly effective in achieving precise estimation for both endmembers and abundances.

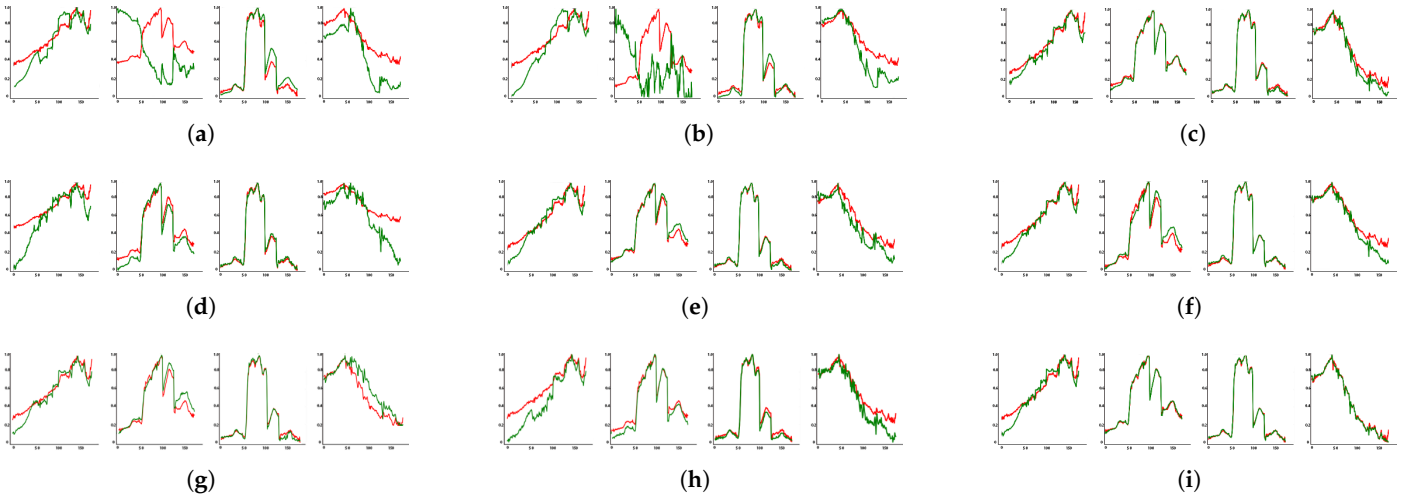

Figure 5. Endmembers on the Urban dataset: (**a**) VCA. (**b**) SGSNMF. (**c**) CNNAEU. (**d**) OSPAEU. (**e**) MTAEU. (**f**) ENDNET. (**g**) PGMSU. (**h**) A2SAN. (**i**) ADAEU.

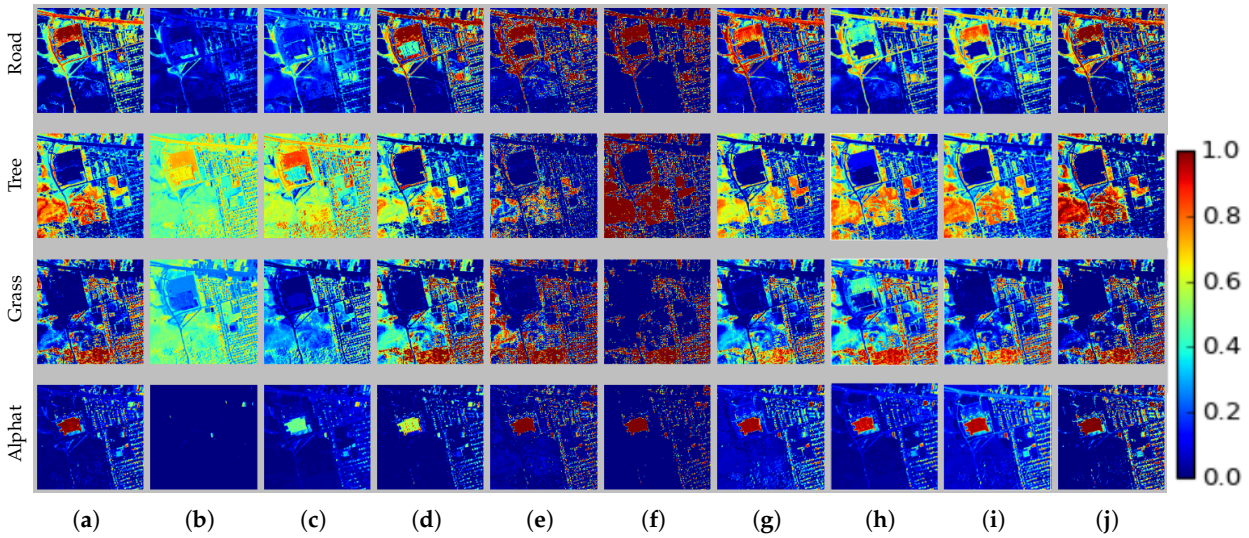

Figure 6. Abundance maps on the Urban dataset: (**a**) GT. (**b**) DMAXD. (**c**) SGSNMF. (**d**) CNNAEU. (**e**) OSPAEU. (**f**) MTAEU. (**g**) ENDNET. (**h**) PGMSU. (**i**) A2SAN. (**j**) ADAEU.

4.4. Experiment with Samson Dataset

For the Samson dataset, SAD and RMSE are presented in Table 4. It seems that the SAD values are considerably lower compared to those of the other two datasets. Most methods yield SAD values less than 0.1, with the exception of VCA and PGMSU for the third endmember *water*. ENDNET shows better endmember extraction capabilities, achieving a SAD of 0.0338, which is the second-best value. In contrast, the proposed ADAEU achieves the lowest SAD and RMSE for each endmember and abundance, highlighting its strong endmember identification and abundance estimation abilities.

Figures 7 and 8 show the endmembers and abundances on Samson dataset. For the endmember extraction, these methods exhibit higher performance in endmember curves matching compared to the other two datasets. The main difference is in the third endmember *water*. Meanwhile, the proposed method achieves the best endmember estimation with nearly

overlapped curves. This could be due to the EE-Net being weighted by anomaly pixels. However, for the abundance as shown in Figure 8, other methods display varying degrees of deviation from the ground truth across different endmembers. For instance, the first three methods and PGMSU exhibit a large deviation from the ground truth. A2SAN also performs well, similar to its performance on Urban dataset, ranking second place in terms of abundance maps. The abundance maps generated by our ADAEU unmixing method closely align with the real ones, demonstrating its exceptional effectiveness in abundance estimation.

Table 4. Quantitative results for the Samson dataset. The best results are shown in bold, and the second-best results are in blue.

Methods		VCA	SGSNMF	CNNAEU	OSPAEU	MTAEU	ENDNET	PGMSU	A2SAN	ADAEU
SAD	Soil	0.0268	0.0097	0.0570	0.1038	0.0379	0.0107	0.0186	0.0261	**0.0096**
	Tree	0.0409	0.0497	0.0409	0.0965	0.0553	0.0410	0.0329	0.0402	**0.0302**
	Water	0.4325	0.2150	0.1260	0.0957	0.0531	0.0498	0.3506	0.1755	**0.0156**
Mean SAD		0.1667	0.0915	0.0746	0.0953	0.0488	0.0338	0.1341	0.0806	**0.0184**
RMSE	Soil	0.2502	0.1793	0.2119	0.2404	0.1318	0.1385	0.1812	0.0773	**0.0331**
	Tree	0.2682	0.2483	0.3089	0.2140	0.0859	0.0850	0.1923	0.0497	**0.0252**
	Water	0.4277	0.3590	0.2294	0.1384	0.0768	0.1021	0.3071	0.0755	**0.0115**
Mean RMSE		0.3154	0.2572	0.2467	0.1976	0.0981	0.1085	0.2268	0.0662	**0.0233**

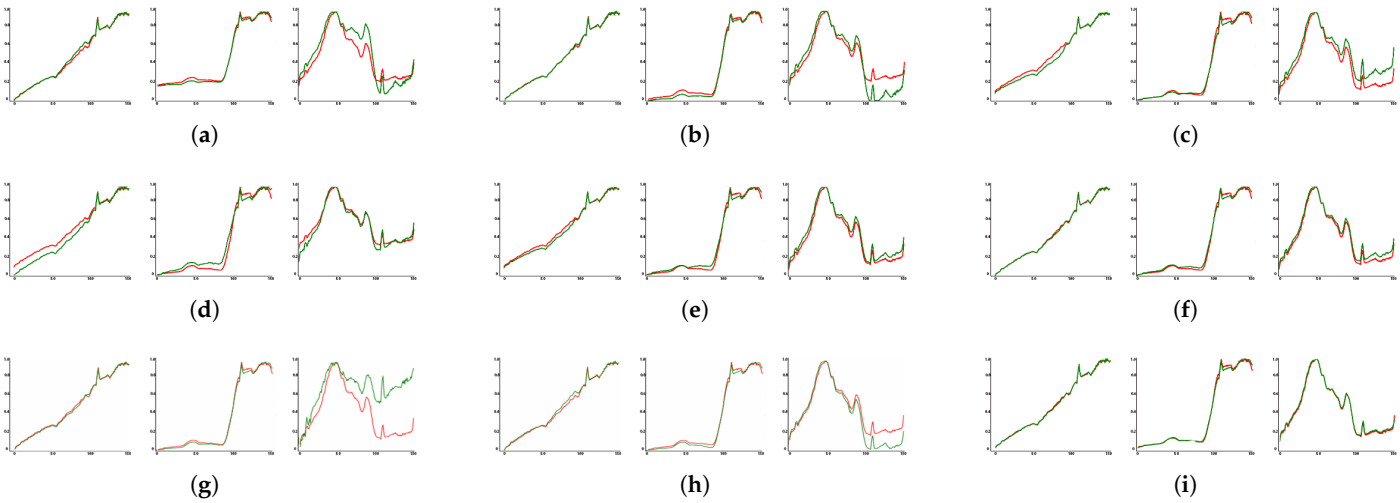

Figure 7. Endmembers on the Samson dataset: (**a**) VCA. (**b**) SGSNMF. (**c**) CNNAEU. (**d**) OSPAEU. (**e**) MTAEU. (**f**) ENDNET. (**g**) PGMSU. (**h**) A2SAN. (**i**) ADAEU.

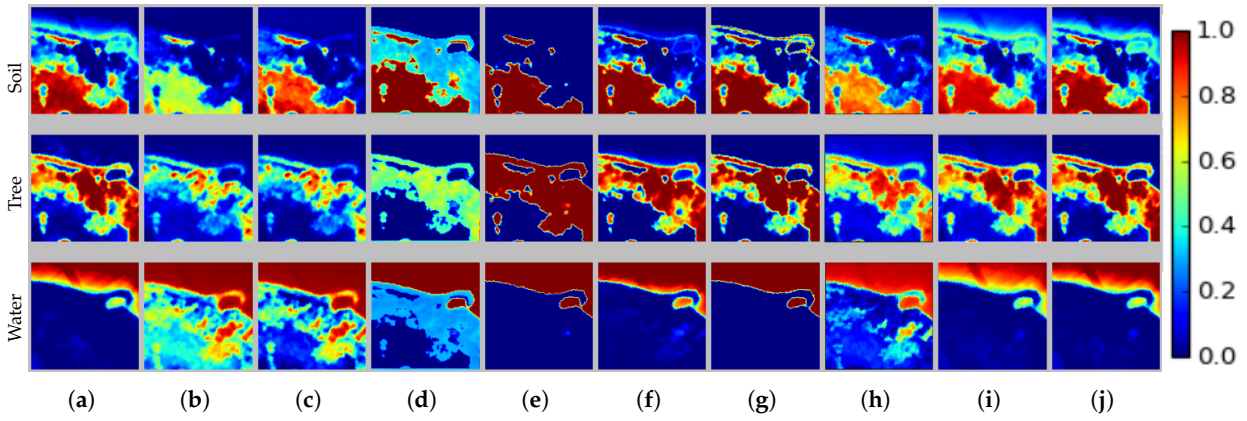

Figure 8. Abundance maps on the Samson dataset: (**a**) GT. (**b**) VCA. (**c**) SGSNMF. (**d**) CNNAEU. (**e**) OSPAEU. (**f**) MTAEU. (**g**) ENDNET. (**h**) PGMSU. (**i**) A2SAN. (**j**) ADAEU.

4.5. Ablation Study

In this section, we perform several ablation experiments to evaluate the effectiveness of each component of the proposed ADAEU unmixing method. To maintain the clarity and focus of the paper, we illustrate the Urban and Jasper Ridge datasets, as they are highly representative.

4.5.1. Impact of Loss Function in EE-Net

In order to verify the effectiveness of the loss function in EE-Net, ablation experiments are conducted on the Urban scenario as shown in Table 5. The experiments compare the unmixing results under four scenarios: with or without anomaly detection weights (AD or NoAD) and with or without the logarithm function (log(SAD) or SAD). A comparison between column 1 and column 2, as well as between column 3 and column 4, clearly demonstrates that employing log(SAD) into the loss function leads to a reduction in SAD values. Quantitatively, this improvement corresponds to an approximate decrease of 0.01 in mSAD, indicating enhanced unmixing performance. In addition, the impact of anomaly detection weight (AD) is evident from the results. In the absence of AD, the SAD values increase by approximately 0.001, underscoring the importance of AD in mitigating the influence of outliers. Overall, the integration of AD and log(SAD) in the loss function achieves the best performance metrics, effectively validating the superiority of the proposed loss function.

Table 5. Ablation study regarding loss function in EE-Net: impact of AD and logarithm on the Urban dataset (SAD). The best results are shown in bold, and the second-best are in blue.

	NoAD + SAD	**NoAD + log(SAD)**	**AD + SAD**	**AD + log(SAD)**
Asphalt	**0.0716**	0.0805	0.0870	0.0740
Grass	0.0409	0.0334	0.0400	**0.0314**
Tree	0.0341	0.0264	0.0320	**0.0251**
Roof	0.0947	0.0125	0.0243	**0.0122**
Average	0.0603	0.0382	0.0458	**0.0357**

4.5.2. Impact of Attention Mechanism in AE-Net

For the abundance network, we conduct the ablation experiments on Jasper Ridge dataset to evaluate the role of the attention mechanism. Specifically, we compare the performance of the network with or without attention mechanism (AM or NoAM). The results, as shown in Table 6, demonstrate that the inclusion of the attention mechanism significantly improves the accuracy of the abundance estimation. When the attention mechanism is employed, the RMSE values are consistently lower, with reductions of approximately 0.001 across all endmembers. This improvement highlights the effectiveness of the used attention mechanism.

Table 6. Ablation study for AE-Net: impact of attention mechanism (AM) on the Jasper Ridge dataset (RMSE). The best results are shown in bold.

	NoAM	**AM**
Tree	0.0223	**0.0213**
Water	0.0195	**0.0185**
Tree	0.0423	**0.0413**
Road	0.0349	**0.0336**
Average	0.0298	**0.0287**

4.5.3. Impact of Architecture of ADAEU Network

The proposed ADAEU uses two networks in an orderly manner, namely, the EE-Net first extracts the endmembers then puts them into the AE-Net to estimate the abundances. Here, in order to verify the rationality of the method, we conduct parallel experiments on two subnetworks. That is, the autoencoding network shares decoding layer weights while performing endmember extraction and abundance estimation, with the network architecture remaining unchanged. Then, the loss function is formulated as

$$L = \lambda_1 L_{EE} + \lambda_2 L_{AE} + \lambda_3 \|\boldsymbol{y_1} - \boldsymbol{y_2}\|_F^2 \tag{21}$$

where L_{EE} and L_{AE} are the loss function of networks EE-Net and AE-Net, respectively. The third term constrains the consistency of the abundance estimated by EE-Net and AE-Net.

Table 7 gives the results corresponding to different values of λ_1, λ_2, λ_3, where $\lambda_1 \in \{0, 0.1, 1\}$, and $\lambda_2 \in \{0, 1, 10\}$, $\lambda_3 \in \{0, 0.01, 0.1\}$. As can be seen in Table 7, the best endmember estimation occurs when $\lambda_1 = 0.1$, $\lambda_2 = 0$. Moreover, with $\lambda_1 \geq \lambda_2$, the endmember can be well estimated. However, for the AE-Net, there are no obvious rules governing the optimal parameter combinations for achieving the best performance. Interestingly, it is possible to obtain a very high SAD and a low RMSE simultaneously, such as $\lambda_1 = 0$, $\lambda_2 = 1$, $\lambda_3 = 0.1$. In this case, the network seems to be more inclined towards training for the best abundance. However, in this case, both the RMSE and SAD are relatively large.

Overall, the parallel use of EE-Net and AE-Net results in inferior unmixing compared to their individual use. This is primarily due to the mutual inheritance of each networks' shortcomings, which leads to error propagation and performance degradation. For instance, if EE-Net extracts endmembers with deviation, the errors are passed on to AE-Net, causing inaccuracies in the abundance estimation. In contrast, when the networks are operated independently, EE-Net can focus on optimizing the endmember extraction, and the AE-Net can accurately calculate the abundance without being affected by erroneous inputs. This approach leverages the strengths of each network, significantly improving unmixing accuracy.

Table 7. Ablation study for parallel network architecture: impact of different parameters on the Urban dataset (SAD). The best results are in bold.

λ_1	λ_2	λ_3	SAD	RMSE
0	1	0	0.4300	0.1955
0	1	0.01	0.4397	0.2022
0	1	0.1	0.4586	0.2074
0	10	0.01	0.4397	0.1894
0	10	0.1	0.4259	0.1941
0.1	0	0.01	**0.0432**	0.1103
0.1	0	0.1	0.1061	0.1648
0.1	1	0	0.1562	0.1089
0.1	1	0.01	0.1737	0.1147
0.1	1	0.1	0.2242	0.1764
0.1	10	0	0.3088	0.1559
0.1	10	0.01	0.3101	0.1587
0.1	10	0.1	0.3377	0.1718
1	0	0.01	0.0433	0.1098
1	0	0.1	0.0521	0.1117
1	1	0	0.0526	0.1072
1	1	0.01	0.0457	0.1048
1	1	0.1	0.0450	0.1031
1	10	0	0.4098	0.1087
1	10	0.01	0.1590	0.1078
1	10	0.1	0.1711	**0.1008**

5. Conclusions

This paper proposes a new hyperspectral image unmixing method based on two-stream autoencoder networks. Unlike general autoencoder network-based methods, we first use a network for endmember extraction, and then put the extracted endmember into another network with shared weight. Moreover, by modifying the loss function, log (SAD) is used to replace SAD, increasing the proportion of "good points" in the loss function, and adding anomaly detection coefficients to reduce the impact of anomalies on the network. For abundance networks, attention mechanisms enable the network to extract more important features. The unmixing results show that compared with other methods, the unmixing results of this method are closer to real data, especially in terms of abundance estimation, which has significant improvements. In addition, although the proposed unmixing method achieves good performance, the spectral and spatial information are not fully exploited in our work. Therefore, more information, such as spectral similarity in a spatial homogeneous neighborhood, Laplace graph, 3D convolution, and patch combined with 2D convolution, etc., can be considered in the EE-Net to incorporate spatial information in our future works.

Author Contributions: Conceptualization, H.L. and C.Z.; methodology, H.L., C.Z. and Z.W.; software, C.Z.; validation, C.Z. and J.H.; writing—original draft preparation, C.Z. and H.L.; writing—review and editing, H.L. and J.H.; visualization, H.L., C.Z. and J.H.; supervision, H.L. and Z.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China under Grant 61971223, and in part by Key Laboratory of Analysis of Mathematical Theory and Modeling of Complex Systems, the Ministry of Industry and Information Technology.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

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

References

1. Yang, H.; Yu, H.; Hong, D.; Xu, Z.; Wang, Y.; Song, M. Hyperspectral image classification based on multi-level spectral-spatial transformer network. In Proceedings of the 2022 12th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS), Rome, Italy, 13–16 September 2022 ; pp. 1–4.
2. Yu, C.; Zhou, S.; Song, M.; Chang, C.I. Semisupervised hyperspectral band selection based on dual-constrained low-rank representation. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 5503005. [CrossRef]
3. Zhang, Q.; Dong, Y.; Yuan, Q.; Song, M.; Yu, H. Combined deep priors with low-rank tensor factorization for hyperspectral image restoration. *IEEE Geosci. Remote Sens. Lett.* **2023**, *20*, 5500205. [CrossRef]
4. Su, H.; Wu, Z.; Zhang, H.; Du, Q. Hyperspectral anomaly detection: A survey. *IEee Geosci. Remote Sens. Mag.* **2021**, *10*, 64–90. [CrossRef]
5. Shang, W.; Jouni, M.; Wu, Z.; Xu, Y.; Dalla Mura, M.; Wei, Z. Hyperspectral anomaly detection based on regularized background abundance tensor decomposition. *Remote Sens.* **2023**, *15*, 1679. [CrossRef]
6. Xu, Y.; Wu, Z.; Chanussot, J.; Wei, Z. Nonlocal patch tensor sparse representation for hyperspectral image super-resolution. *IEEE Trans. Image Process.* **2019**, *28*, 3034–3047. [CrossRef]
7. Liu, Q.; Xiao, L.; Yang, J.; Wei, Z. Multilevel superpixel structured graph U-Nets for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 5516115. [CrossRef]
8. Ye, F.; Wu, Z.; Xu, Y.; Liu, H.; Wei, Z. Bayesian hyperspectral image super-resolution in the presence of spectral variability. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5545613. [CrossRef]
9. Chen, J.; Zhao, M.; Wang, X.; Richard, C.; Rahardja, S. Integration of physics-based and data-driven models for hyperspectral image unmixing: A summary of current methods. *IEEE Signal Process. Mag.* **2023**, *40*, 61–74. [CrossRef]
10. Palsson, B.; Sveinsson, J.R.; Ulfarsson, M.O. Blind hyperspectral unmixing using autoencoders: A critical comparison. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 1340–1372. [CrossRef]

11. 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]
12. Li, J.; Bioucas-Dias, J.M. Minimum volume simplex analysis: A fast algorithm to unmix hyperspectral data. In Proceedings of the IGARSS 2008–2008 IEEE International Geoscience and Remote Sensing Symposium, Boston, MA, USA, 7–11 July 2008; Volume 3, p. III-250.
13. Heinz, D.C.; Chang, C.I. 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]
14. Yang, Z.; Zhou, G.; Xie, S.; Ding, S.; Yang, J.M.; Zhang, J. Blind spectral unmixing based on sparse nonnegative matrix factorization. *IEEE Trans. Image Process.* **2010**, *20*, 1112–1125. [CrossRef]
15. Miao, L.; Qi, H. Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 765–777. [CrossRef]
16. Wang, X.; Zhong, Y.; Zhang, L.; Xu, Y. Spatial group sparsity regularized nonnegative matrix factorization for hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 6287–6304. [CrossRef]
17. Qian, Y.; Jia, S.; Zhou, J.; Robles-Kelly, A. Hyperspectral unmixing via *L*_{1/2} sparsity-constrained nonnegative matrix factorization. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 4282–4297. [CrossRef]
18. Liu, H.; Lu, Y.; Wu, Z.; Du, Q.; Chanussot, J.; Wei, Z. Bayesian unmixing of hyperspectral image sequence with composite priors for abundance and endmember variability. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 5503515. [CrossRef]
19. Lu, Y.; Liu, H.; Wu, Z.; Wei, Z. Multi-temporal hyperspectral images unmixing by mixed distribution considering smooth variation of abundance. In Proceedings of the IGARSS 2020–2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020; pp. 2185–2188.
20. Li, C.; Xie, H.B.; Mengersen, K.; Fan, X.; Da Xu, R.Y.; Sisson, S.A.; Van Huffel, S. Bayesian nonnegative matrix factorization with Dirichlet process mixtures. *IEEE Trans. Signal Process.* **2020**, *68*, 3860–3870. [CrossRef]
21. Chen, P.; Nelson, J.D.; Tourneret, J.Y. Toward a sparse Bayesian Markov random field approach to hyperspectral unmixing and classification. *IEEE Trans. Image Process.* **2016**, *26*, 426–438. [CrossRef] [PubMed]
22. 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]
23. Fang, Y.; Wang, Y.; Xu, L.; Zhuo, R.; Wong, A.; Clausi, D.A. BCUN: Bayesian fully convolutional neural network for hyperspectral spectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5523714. [CrossRef]
24. Wang, J.; Xu, J.; Chong, Q.; Liu, Z.; Yan, W.; Xing, H.; Xing, Q.; Ni, M. Ssanet: An adaptive spectral–spatial attention autoencoder network for hyperspectral unmixing. *Remote Sens.* **2023**, *15*, 2070. [CrossRef]
25. Zhao, M.; Chen, J.; Dobigeon, N. AE-RED: A hyperspectral unmixing framework powered by deep autoencoder and regularization by denoising. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 5512115. [CrossRef]
26. Liu, J.; Wu, Z.; Xiao, L.; Wu, X.J. Model inspired autoencoder for unsupervised hyperspectral image super-resolution. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5522412. [CrossRef]
27. Li, H.; Borsoi, R.A.; Imbiriba, T.; Closas, P.; Bermudez, J.C.; Erdoğmuş, D. Model-based deep autoencoder networks for nonlinear hyperspectral unmixing. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 5506105. [CrossRef]
28. Xu, X.; Song, X.; Li, T.; Shi, Z.; Pan, B. Deep autoencoder for hyperspectral unmixing via global-local smoothing. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5524216. [CrossRef]
29. Qu, Y.; Qi, H. uDAS: An untied denoising autoencoder with sparsity for spectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2018**, *57*, 1698–1712. [CrossRef]
30. Qi, L.; Gao, F.; Dong, J.; Gao, X.; Du, Q. SSCU-Net: Spatial–Spectral collaborative unmixing network for hyperspectral images. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5407515. [CrossRef]
31. Yao, J.; Hong, D.; Xu, L.; Meng, D.; Chanussot, J.; Xu, Z. Sparsity-enhanced convolutional decomposition: A novel tensor-based paradigm for blind hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 5505014. [CrossRef]
32. Su, Y.; Li, J.; Plaza, A.; Marinoni, A.; Gamba, P.; Chakravortty, S. DAEN: Deep autoencoder networks for hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 4309–4321. [CrossRef]
33. 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]
34. Shi, S.; Zhao, M.; Zhang, L.; Altmann, Y.; Chen, J. Probabilistic generative model for hyperspectral unmixing accounting for endmember variability. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5516915. [CrossRef]
35. Ghosh, P.; Roy, S.K.; Koirala, B.; Rasti, B.; Scheunders, P. Hyperspectral unmixing using transformer network. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5535116. [CrossRef]
36. Yang, Z.; Xu, M.; Liu, S.; Sheng, H.; Zheng, H. Spatial-spectral attention bilateral network for hyperspectral unmixing. *IEEE Geosci. Remote Sens. Lett.* **2023**, *20*, 5507505. [CrossRef]

37. Yan, B.; Wu, Z.; Liu, H.; Xu, Y.; Wei, Z. Hyperspectral unmixing via wavelet based autoencoder network. In Proceedings of the 2019 10th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS), Amsterdam, The Netherlands, 24–26 September 2019; pp. 1–5.
38. Palsson, F.; Sigurdsson, J.; Sveinsson, J.R.; Ulfarsson, M.O. Neural network hyperspectral unmixing with spectral information divergence objective. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; pp. 755–758.
39. Min, A.; Guo, Z.; Li, H.; Peng, J. JMnet: Joint metric neural network for hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 5505412. [CrossRef]
40. Dou, Z.; Gao, K.; Zhang, X.; Wang, H.; Wang, J. Hyperspectral unmixing using orthogonal sparse prior-based autoencoder with hyper-Laplacian loss and data-driven outlier detection. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 6550–6564. [CrossRef]
41. 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]
42. 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]
43. Pálsson, B.; Sveinsson, J.R.; Ulfarsson, M.O. Multitask learning for spatial-spectral hyperspectral unmixing. In Proceedings of the 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019.
44. Rasti, B.; Koirala, B.; Scheunders, P.; Ghamisi, P. UnDIP: Hyperspectral unmixing using deep image prior. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 5504615. [CrossRef]
45. Ulyanov, D.; Vedaldi, A.; Lempitsky, V. Deep image prior. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 9446–9454.
46. Tao, X.; Paoletti, M.E.; Wu, Z.; Haut, J.M.; Ren, P.; Plaza, A. An abundance-guided attention network for hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 5505414. [CrossRef]
47. Jin, Q.; Ma, Y.; Mei, X.; Ma, J. TANet: An unsupervised two-stream autoencoder network for hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 5506215. [CrossRef]
48. Hong, D.; Gao, L.; Yao, J.; Yokoya, N.; Chanussot, J.; Heiden, U.; Zhang, B. Endmember-guided unmixing network (EGU-Net): A general deep learning framework for self-supervised hyperspectral unmixing. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *33*, 6518–6531. [CrossRef]
49. Zheng, H.; Li, Z.; Sun, C.; Zhang, H.; Liu, H.; Wei, Z. Blind unmixing using dispersion model-based autoencoder to address spectral variability. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 5518214. [CrossRef]
50. Reed, I.; Yu, X. Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution. *IEEE Trans. Acoust. Speech, Signal Process.* **1990**, *38*, 1760–1770. [CrossRef]
51. Woo, S.; Park, J.; Lee, J.Y.; Kweon, I.S. Cbam: Convolutional block attention module. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 3–19.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.