frontiers

A blind convolutional deep autoencoder for spectral unmixing of hyperspectral images over waterbodies

Estefanía Alfaro Mejía [1,*], **Vidya Manian** [1], **Joseph Ortiz** [2], **and Roger Tokars** [3]

[1] *Artificial Inteligence Imaging Groupd (AIIG), University of Puerto Rico At Mayaguez, Departament of Electrical and Computer Engineering, Mayaguez, PR 00681, USA*
[2] *Kent State University, Departament of Earth Sciences, Kent Ohio, USA*
[3] *NASA Gleen Research Center 21000 Brookpark Rd, Cleveland, OH 44135, USA*

Correspondence*:
Corresponding Author
estefania.alfaro@upr.edu

ABSTRACT

The harmful algal blooms (HABs) have dangerous repercussions for biodiversity, the ecosystem, and public health. Automatic identification based on remote sensing hyperspectral image analysis provides a valuable mechanism for extracting the spectral signatures of HABs and their respective percentage in a region of interest. This paper proposes a new model called non-symmetrical autoencoder for spectral unmixing (NSAE-SU) to perform endmember extraction and fractional abundance estimation. The model is assessed in benchmark datasets, such as Jasper and Samson. Additionally, a case study of at HSI2 images acquired by NASA over Lake Erie in 2017 is proposed for extracting optical water types. The results using the NSAE-SU for the benchmark datasets improve the performance as indicated by the spectral angle distance (SAD) compared to three baseline algorithms. Better results are obtained: for the Samson dataset, the NSAE-SU model performs better for the water endmember extraction **0.060**, and soil **0.025**. The mean SAD performance is also better than the other three baseline algorithms; for the abundances maps, the NSAE-SU method performs better with an RMSE for water **0.091**, and soil **0.187**, compared to the ground truth. Also, for the Jasper Ridge dataset, NSAE-SU performs well for the tree **0.039** and road **0.068** endmember extraction and achieving better results also for the abundances maps of the water **0.1121** and soil **0.2316**. The NSAE-SU, can identify the presence of Chl-a over water bodies. Chl-a is an essential indicator of the presence of the macrophytes under different concentrations and cyanobacteria. The NSAE-SU achieves the performance of the SAD metric of **0.307**, compared with a reference ground truth of the spectral signatures of the Chl-a.

Keywords: Hyperspectral images, spectral unmixing, endmembers, abundances maps,image processing, Deep learning, Autoencoder, algal bloom.

1 INTRODUCTION

The increase in harmful algal blooms occuring in recent years is related to global warming and different human activities that contaminate and modify these water-bearing zones causing severe problems to the marine ecosystems, biodiversity, and collateral damage to the health of humans Guo et al. (2022). Human health is affected by the consumption of water coming mainly from lakes, which contains different kinds of algae, scum, and sediments. Hence, it is necessary to determine the proportions of these algae so that water quality management authorities can establish safe thresholds for consumption and recreative purposes.

The objective of this article is hence to identify the presence and the percentage or fractional abundance of algae, the composition of the different materials, and concentrations in a region of interest in a hyperspectral image. At present, multiple techniques have been developed to detect and quantify Chlorophyll-a (Chl-a), with the concentration being an indicator of the algae presence, Chl-a models can be divided into two branches: physics-based methods and data-driven methods Zhu et al. (2023). Physics-based methods simulate the behavior of Chl-a at the water bodies. On the other hand, the data-driven methods are related to analyzing previous data of the region of interest (ROI). This data can be weather variables, field samples, and images. The last approach constitutes Machine learning Chong et al. (2023) and Deep learning Park et al. (2022), methods which are applied mainly to hyperspectral images. Additional measurements such as temperature, wind speed, and fluorometric data samples are acquired from the same ROI as the images.

In order to perform Chl-a extraction, it is necessary to analyze the measurements; typically, much of these measurements are fluorometer samples taken at the field, weather variables, and hyperspectral images (HSI) acquired from an ROI. HSI measurements are recorded by sensors that record spectral signatures over hundreds of narrow contiguous bands ranging from 380 to 2500 nanometer wavelengths. However, the acquired wavelength range varies depending on the type of the sensor Xu et al. (2019)-Zhong et al. (2018). One of the most relevant features of the hyperspectral images is the high spectral content enabling the use of these images for specific applications involving spectral unmixing, classification, and region of interest-based segmentation.

Spectral Unmixing (SU) is an active field of research whose goal is to analyze the materials and compositions of an acquired hyperspectral scene; from the analysis of the reflectance of the HSI image, the pure spectral signature called endmember of each material and the proportions of the endmembers of the different materials present in each pixel in the HSI scene known as fractional abundances maps are estimated. Typically, SU has been explored by classical approaches such as Linear Mixing Models (LMM) Heylen et al. (2014), optimization approaches such as MESMA, which are based on extracting multiple materials in a scene resulting in the application of classical optimization Tane et al. (2018), and machine learning techniques, such as Support Vector Machine (SVM) Chunhui et al. (2018) - Wang et al. (2013) and Neural networks Qi et al. (2023) - Zhang et al. (2022).

The LMM is a baseline algorithm for the SU framework. LMM is based on a linear relationship between the endmembers or pure substances and their fractional abundances. Each pixel intensity can be considered as the linear combination of all materials that belong to the acquired scene; this approach represents an appropriate solution for the macroscopic analysis in which the object in analysis represents a large percentage of the acquired scene, such as soil, grass, and vegetation.

Typically, SU is addressed using classical approaches for endmember extraction, such as the multiple endmember spectral signature Yang et al. (2022) method and the LMM Imbiriba et al. (2018). In addition, the methods performed for the atmospheric correction, based on the varimax-rotated principal component method, Ortiz et al. (2019). Notwithstanding, the spectral unmixing is solved using machine learning

approaches, such as Neural networks Qi et al. (2023) - Zhang et al. (2022) to increase the accuracies obtained in the fractional abundance maps and endmember extraction. Most of the applications of SU methods are for sediment analysis from satellite images Waga et al. (2022).

On the other hand, geometrical approaches have been applied to solve the spectral unmixing problem, such as N-FINDR Winter (1999), Vertex Component Analysis (VCA) Nascimento and Dias (2005), and Fast Pure Pixel Index (FPPI) Das et al. (2019). These methods are based on iterative algorithms, which compute the determinant to maximize the volume estimation of a convex hull. In the ideal case, the endmembers or pure substances represent the vertices, and the mixed pixel is contained in the geometrical surface.

New strategies based on supervised and unsupervised machine learning and optimization techniques have been developed to improve the endmember extraction algorithm's accuracy Shah et al. (2020)-Xu et al. (2019). The majority of the unsupervised approaches are based on Autoencoders for endmember extraction and estimation of the fractional abundances maps Ranasinghe et al. (2020), Palsson et al. (2021) Hadi et al. (2022). The autoencoders used to address the SU framework are configured as non-symmetrical models where the encoder has more degrees of freedom in the design to add more layers. Commonly, the constraints for non-negativity (ANC) and sum to one (ASC) are included in the encoder. On the contrary, the decoder typically possesses one layer for conducting the endmember extraction and has additional layers added for non-negative regularization.

In order to perform endmembers extraction and fractional maps estimation, we propose a new method for the analysis of optical water types for detecting Chl-a based on an unsupervised deep learning approach. The method is composed of five stages, as follows: the input is the HSI2 image, the region of interest selection that enables the analysis of the water bodies by regions due to the large size of the images, the spectral derivatives computation for performing the sunglint correction, and a block for endmembers extraction using the model NSAE-SU for detecting the Chl-a in Lake Erie. Also, other endmembers, such as HABs, sediments, and surface scum, are detected.

This article introduces and assesses a new deep autoencoder called NSAE-SU for endmember extraction and estimation of fractional abundances maps. The model has been assessed using benchmark datasets, such as Jasper and Urban, with known ground truth for the endmembers and the fractional abundances maps. Additionally, to evaluate the performance of the proposed workflow illustrated in Fig. 2 has been used the HSI images at Lake Erie not possess ground truth for the endmembers and abundances maps. The workflow is applied to extract the optical water types or endmembers, and their abundances maps for each ROI selected at the HSI2 images. The main contributions of this article are as follows:

- An unsupervised deep learning model called NSAE-SU is proposed. The NSAE-SU can perform better the extraction of the endmembers, particularly for the water and soil, without removing bands at the image. In addition, the model is also robust for the extraction of abundance maps of water, soil, and tree.
- NSAE-SU is an unsupervised deep learning end-to-end model that can perform the endmember extraction and the fractional abundances map estimation, exploiting the spatial and spectral features of the hyperspectral images and addressing the problem of the mixed pixels.
- In addition, NSAE-SU performs the endmembers extraction and the fractional abundances maps over benchmark datasets such as Jasper and Samson, and is proposed a case study at Lake Erie over a water body in order to analyze the different Optical water types, then compared with spectral

signatures extracted experimentally from the Lakes, which have been used as a baseline to compare the endmembers.

The rest of this article is organized as follows: section 2 provides the background of hyperspectral images, optical water types, and mathematical foundations for the deep autoencoder. Section 3 describes the NSAE-SU method and its application. The experiments with hyperspectral images are described in section 4. Section 5 presents the metrics used to assess the NSAE-SU model. The analysis of the results and the selection of hyperparameters are explained in section 6 for the benchmark datasets and for HSI2 images. Finally, the conclusions are presented in Section 7.

2 HYPERSPECTRAL UNMIXING

2.1 Hyperspectral Images

Hyperspectral images (HSI) have hundreds of narrow bands, providing a continuous measurement for each pixel in a limited wavelength range; this range depends on the sensor type, one of the most popular are near infrared (nm) and middle infrared (nm). The measurement is performed from the emitted and reflected light in a scene Vivone (2023).

The HSI images are a hypercube or 3D representation with sizes given by $W \times H \times L$ when $W \times H$ means the spatial resolution and L corresponds to the number of bands. In addition, HSI possesses a high spectral resolution and low spatial resolution that is useful for studying the material composition of each pixel in an HSI by spectral unmixing.

2.2 Linear Mixing Models

The spectral signatures presented in an acquired scene are considered as endmembers, and their proportion in each pixel are the abundance maps. Typically, Linear Mixing Model (LMM) performs the endmember extraction and estimation of the fractional abundance maps based on the physical behavior of the interaction between the light and the endmembers, described as a linear function, as we see in the following equation 1.

$$\begin{aligned} Y_n &= M\alpha_n + \eta_n \\ \text{s.t} 1^T\alpha &= 1 \\ \alpha_n &\geq 0 \end{aligned} \tag{1}$$

where M represents the endmember matrix of dimension LXR, R is the endmembers, and α_n is the proportion of endmembers in each pixel, and η_n is the noise vector. The abundances maps are constrained by the sum-to-one (ASC) and non-negativity constraints (ANC).

2.3 Mathematical Foundations of Autoencoder

In order to perform the endmember extraction and the fractional abundance map estimation, a deep convolutional autoencoder is proposed as illustrated in fig. 1. An autoencoder is an unsupervised deep neural network that has learned the structure of the data and performs feature extraction due to a latent data

representation. This method does not require labels for spectral unmixing and the HSI data analysis can be performed without ground truth.

An encoder and decoder together constitutes the autoencoder model. The encoder is given by $f_e = E_{(xd)}$ and performs transformation of the input data into a hidden representation. Then, the decoder makes the reconstruction of the data $\hat{x_d} = D((f_e))$, subject to a loss function, given by the following equation:

$$\mathcal{L}\left(x_d, D(E(\hat{x}_d))\right) \tag{2}$$

The reconstructed data can be represented as a forward pass given by the equation. 3, being α_D, α_E, their respective activation functions at the hidden layers of the model, and $\mathbf{w_d}$, $\mathbf{w_e}$, are the weighted matrices for the decoder and encoder, respectively Goodfellow et al. (2016).

$$x_d = \alpha_D\left(\mathbf{w_d}\left(\alpha_e\left(\mathbf{w_e}\right)\right)\right) \tag{3}$$

However, in order to obtain an accurate reconstruction result in feature enhancement based on the learned distribution of the training data, it is necessary to apply a regularized function, given by the following equation 4, where λ is a tuning parameter, and $\mathcal{J}\left(f_e, \mathbf{w_e}, \mathbf{w_d}\right)$ is a penalty function.

$$\mathcal{L} = \mathcal{L}\left(x_d, \hat{x}_d\right) + \lambda\mathcal{J}(f_e, \mathbf{w_e}, \mathbf{w_d}) \tag{4}$$

In order to perform spectral unmixing analysis using autoencoder; it is necessary to impose the ASC and ANC constraints at the encoder configuration; this enforces the endmembers and abundance maps to be non-negative and not greater than one. The encoder encodes the input data in a latent space, performing convolutional operations, leaky ReLU activations, and dropout to prevent overfitting.

$$\hat{X}_d = [\hat{X}_1, \hat{X}_2, \hat{X}_3, \ldots, \hat{X}_L] \tag{5}$$

$$\hat{X}_d = W_d^N[\sigma^{(N-1)(1)}, \sigma^{(N-1)(2)}, \sigma^{(N-1)(3)}, \ldots, \sigma^{(N-1)(W\times H)}] \tag{6}$$

$$\hat{X}_d = W_{d}{}^{(N)}_{(W\times H)\times R} \times \sigma_{R\times(W\times H)} \tag{7}$$

Subsequently, the decoder reconstructs the data patch from the latent space; the data patch is given by the eq. 5, and the decoder can be rewritten as the equation 6 where W_d corresponds to the weight matrix of the decoder, and σ are the activation functions from the previous (N-1) layers. However, as it is required to perform R endmember extraction, the last layer has R neurons, resulting in the equation 7, which is similar to the equation 1.

3 PROPOSED METHOD

3.1 NSAE-SU Autoencoder

In order to perform the endmembers extraction and fractional map estimation, a convolutional model is proposed as depicted in figure 1, the model is a convolutional autoencoder, and the layers are arranged as described in table 1, the loss function used is the cross entropy given by the equation 8, where P is the number of patches. The input shape selected is patches of size 9×9, then four convolutional operations are applied with the following filter sizes $(3 \times 3, 128)$, $(3 \times 3 \times 64)$, $(3 \times 3 \times 32)$, $(1 \times 1 \times R)$ where R is

the number of endmembers. Between each convolutional operations a dropout operation is performed to prevent the overfitting of the data, except at the first convolutional operation as the input data is directly applied to the convolutional operation. Then, the ANS constraint for the abundances maps is performed by the softmax function.

$$\mathcal{L} = -\frac{1}{P}\sum_{i=1}^{P}(\alpha_i \log \hat{\alpha}_i) \tag{8}$$

Next, at the decoder a 2-dimensional convolutional operation is performed to reconstruct the hyperspectral images with filters of size $(7 \times 7 \times 198)$; the weights of the last layer are the extracted endmembers for each image.

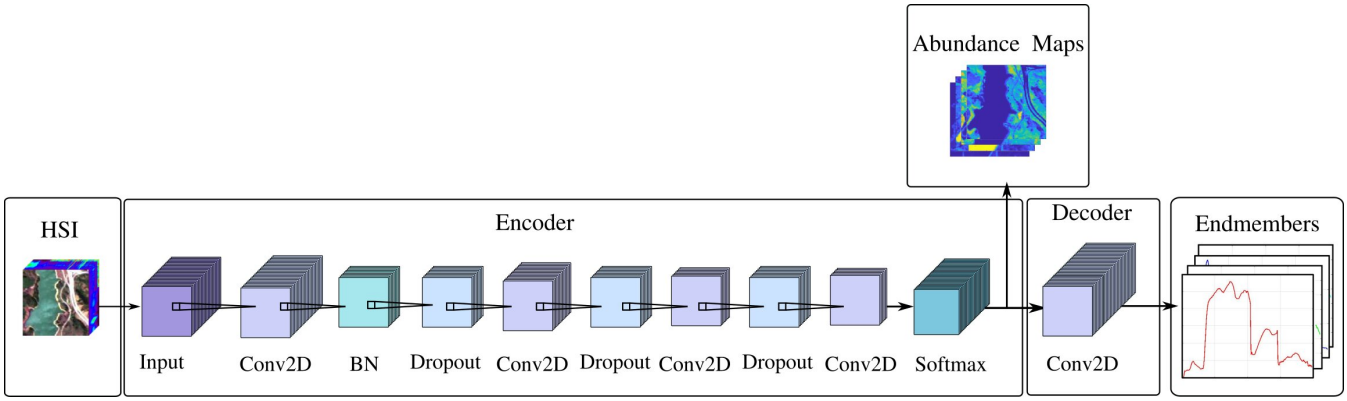


Figure 1. Architectural illustration of the NSAE-SU for performing the spectral unmixing, extracting the abundances maps, and endmembers. The input patch for the model corresponds to the extracted patches of the original image which is the input to the encoder, where convolutional operations, batch normalization (BN), flatten, dense, and softmax operations to extract the abundances maps are performed. Finally, the decoder performs a dense operation with a linear activation function for endmember extraction.

3.2 Hyperparameter configurations

The NSAE-SU model is an unsupervised deep learning model autoencoder. The NSAE-SU is programmed in Python using the TensorFlow libraries; the encoder and decoder integrate the model. Estimating the abundances map is done by the encoder, and endmember extraction is done by the decoder.Estimating the abundances map is done by the encoder, and endmember extraction is done by the decoder. The encoder uses the Leaky Relu activation function with a slope of 0.1 in four 2D convolutional processes to estimate the abundance map; after the first 2D convolution is applied, custom batch normalization is conducted by each batch, removing the gamma factor which is typically performed during a batch normalization operation. Then, a dropout with a rate of 0.03 is applied for the consecutive convolutional operations. This is performed in the ASC layer, a softmax operation is applied using a scale factor of 3; the details for each filter are described in Table. 1. After the abundance maps have been estimated, the decoder executes a 2D convolution operation, whose number of filters equals the number of image bands. A non-negative kernel is employed to add ASC requirements; since the endmembers must be greater than zero, a kernel constraint is applied to avoid the non-negativity. The optimizer employed is the RMSProp, and the learning rate is 0.0001, with 250 epochs total.

Parameter	Value
Input data	9×9
Number of filters Convolution 2D	128
Filters Convolution 2D	3×3
Number of filters Convolution 2D	64
Filters Convolution 2D	3×3
Number of filters Convolution 2D	32
Filters Convolution 2D	3×3
Number of filters Convolution 2D	16
Filters Convolution 2D	3×3
Number of filters Convolution 2D decoder	L
Scaling factor	3
Optimizer	RMSprop
Learning rate	0.0001
Batch size	20
Epochs	250

Table 1. Configuration of the NSAE-SU model and parameter settings for Jasper Ridge dataset.

3.3 General Pipeline for the HSI2 Lake Erie Images

This section presents the workflow developed to address the SU of the HSI2 hyperspectral image for endmember extraction and the estimation of abundance maps. The proposed workflow has five stages, as illustrated in Fig. 2, and are explained as follows: the first stage is the data representation of the hyperspectral image in hypercube format with sizes $W \times H \times L$. The second stage is the selection of ROIs; this procedure is necessary because of the high spatial resolution of the image HSI2 496×5000; the image is subdivided into rectangles of small areas in order to cover the entire image $[(W_1 \times H_1 \times L), \ldots, (W_n \times H_n \times L)]$, the areas of these rectangles are heuristically chosen. Once the ROIs are selected, the third stage performs the spectral derivative (SD) to remove the sunglint effects in the image; the SD is given by eq. 9, where x is the middle band, and k is the step-length, corresponds to the hypercube wavelength.

$$f^{`}(x) \approx \frac{f(x+k) - f(x-k)}{2k} \tag{9}$$

The data is represented in a 2D array once the spectral derivatives have been completed in order to extract the patches for the suggested model NSAE-SU, the batch sizes have the dimensions $9 \times 9 \times L$ are the input data at the model, and the last stage is used to visualize the abundance maps and the endmembers. This workflow is executed for each ROI of the HSI2 image.

The algorithms are executed in a Dell precision server 7920 Rack XCTO Base, Intel Xeon Gold, Graphic card 4GB Nvidia T1000, 1 TB SATA hard drive, 64GB RAM, performance optimized.

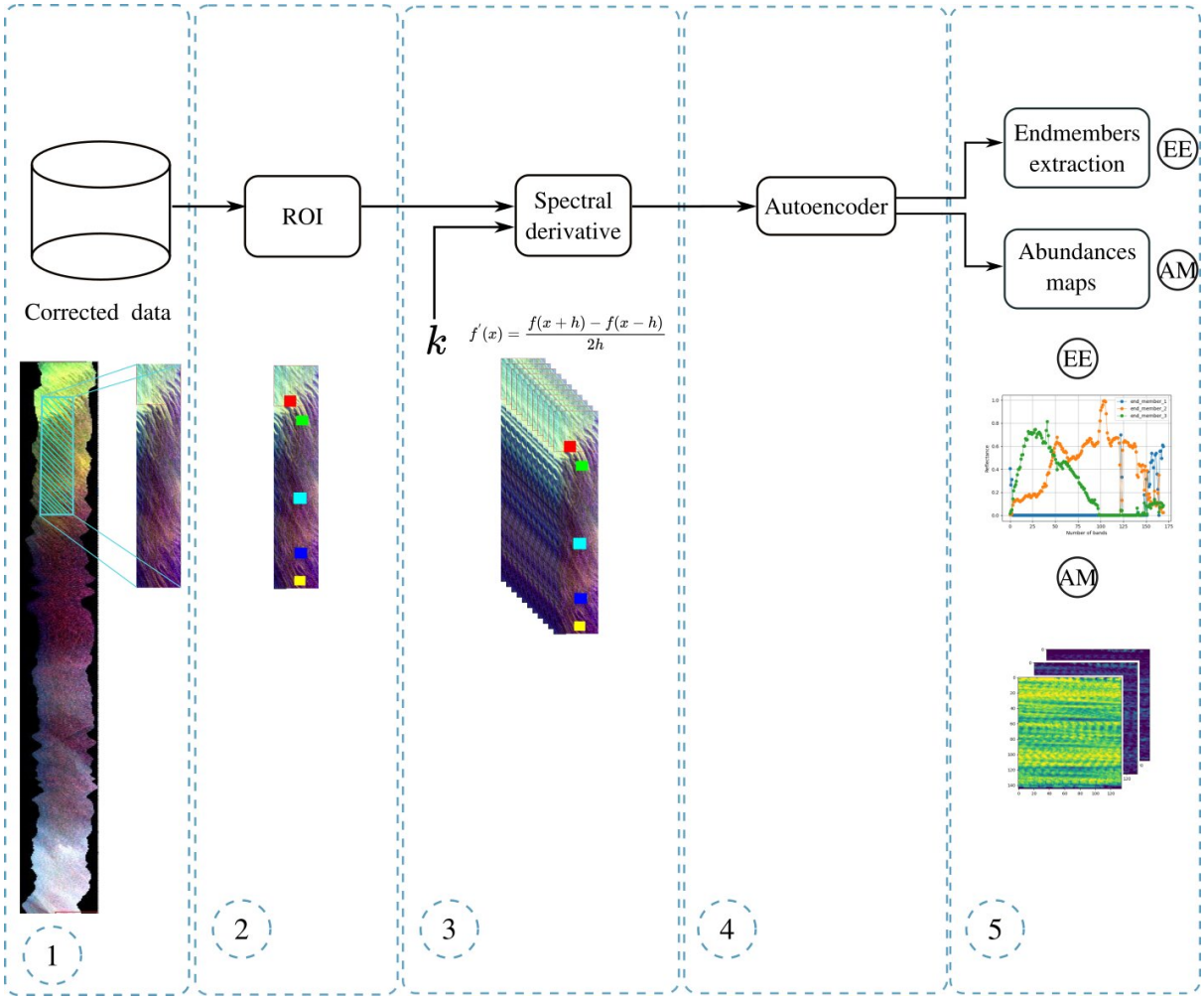


Figure 2. Proposed workflow for spectral umnixing analysis of the HSI2 image

Algorithm 1 Pseudo code of the workflow for performing the endmember extraction and fractional abundance map estimation for the HSI2 Lake Erie image.

Input: ROI's
Output: Endmembers extracted, fractional abundances maps.
$h \leftarrow 3$
for each roi in range(ROI's) **do**
  input_patches = extract_patches(ROI's)
  $f^{'}(x) = \frac{f(x+h)-f(x-h)}{2h}$
  endmember_NSAE-SU, abundances_NSAE-SU = NSAE-SU($f^{'}(x)$)
end for

4 EXPERIMENTAL ANALYSIS

The experiments for the proposed method are conducted on three different datasets of which two are benchmark datasets: Samson and Jasper. The other dataset corresponds to the HSI2 NASA flight campaign from 2017. The datasets used in the experiments are described as follows:

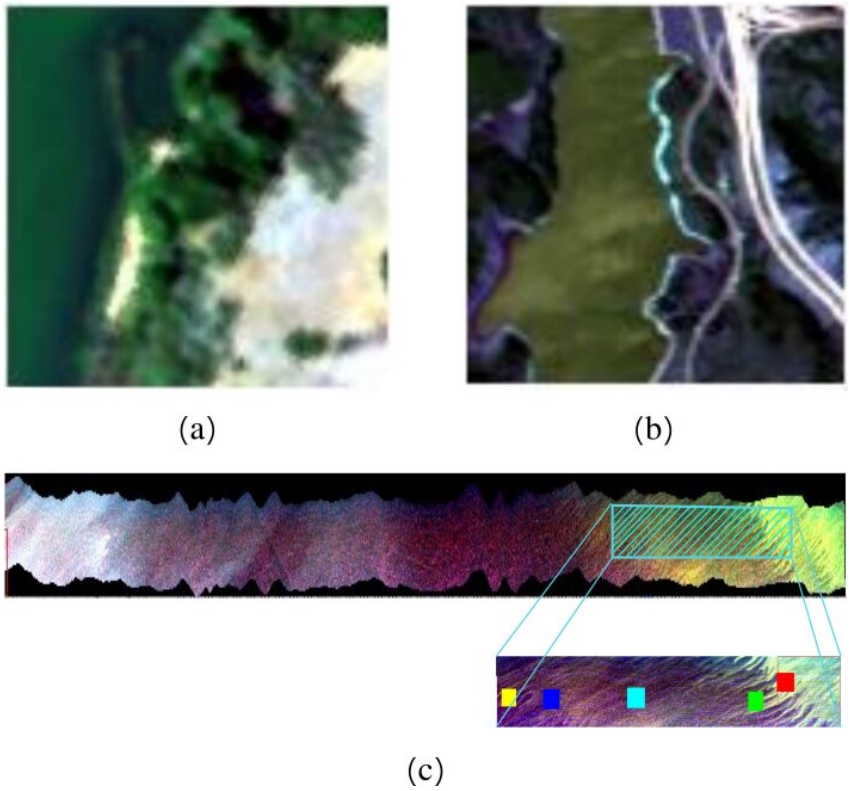


Figure 3. RGB images of the datasets used to evaluate the performance of the NSAE-SU model. Figure. (a) Samson, Figure. (b) Jasper Ridge, and (c) Lake Erie HSI2 image with ROIs highlighted.

1. Jasper: This dataset has 224 bands, out of which 194 are chosen after the noisy channel correction, and has a resolution of 100×100, the wavelength range of each recorded scene is $380 - 2500$ nm; Jasper dataset has four endmembers, as follows: Road, Soil, Water and Tree.
2. Samson: This dataset is a hyperspectral image with 156 bands and a spatial resolution of 95×95, the wavelength range recorded by each pixel of this scene is $401 - 889$ nm. This dataset has three Regions of Interest (ROI) selected from the original hyperspectral image, each of which has its respective ground truth. Samson datasets have three endmembers corresponding to Soil, Tree, and Water.
3. HSI2 Image of Lake Erie: This dataset is taken by NASA during the campaign flight of 2017. After the acquisition stage, these images were preprocessed for atmospheric correction. The image is called in this paper as Hyperspectral image (HSI2), and has a spatial resolution of $496 \times 5000 \times 170$. The HSI2 image do not possess ground truth for the endmembers and the abundances maps. Instead, the ground truth is extracted from optical water types acquired from field measurements explained in the proposed methods section.

5 PERFORMANCE METRICS

In order to evaluate the performance of the proposed algorithm, the extracted endmembers and fractional abundances are evaluated separately. For the endmember extraction, the spectral angle distance is given by eq. 10, where $\hat{\mathbf{m}}_\mathbf{i}$ represents the endmembers extracted for the model, and $\mathbf{m}_\mathbf{i}$ are the ground truth endmembers.

$$SAD = \frac{1}{R}\sum_{i=1}^{R} \arccos\left(\frac{\langle \hat{\mathbf{m}}_\mathbf{i}, \mathbf{m}_\mathbf{i}\rangle}{\|\hat{\mathbf{m}}_\mathbf{i}\|_2 \|\mathbf{m}_\mathbf{i}\|_2}\right) \tag{10}$$

For the abundances, maps are used to mean square error, given by the eq. 11 where $\hat{\alpha}_{\mathbf{i}}$ represents the abundances of all pixels for i endmember, and $\alpha_{\mathbf{i}}$ are the reference abundance maps.

$$MSE = \frac{1}{R}\sum_{i=1}^{R} \|\alpha_{\mathbf{i}} - \hat{\alpha}_{\mathbf{i}}\| \tag{11}$$

6 RESULTS AND DISCUSSION

This section presents and discusses the results of applying the general workflow and the NSAE-SU for spectral unmixing analysis of the benchmark datasets and HSI2 Lake Erie hyperspectral images.

6.1 Endmembers extraction and abundance maps estimation from benchmark datasets: Samson and Jasper

In order to measure the performance achieved for the NSAE-SU at the endmembers extraction (EE) and for the abundance maps estimation (AM), it is assessed with Samson and Jasper Ridge benchmark datasets. The Samson dataset is also compared with three methods, two of them based on Deep learning, such as CNNAEU and UnDIP, and a classical method based on geometrical approaches. The endmembers extracted in Samson dataset for the three baseline algorithms and our model NSAE-SU are illustrated in Figure.4. The comparison is shown for the endmembers produced from the baseline algorithms and the ground truth; each spectral signature is scaled between 0 and 1 to allow for comparison.

As described in the experimental analysis section, Samsom possesses three endmembers, as follows water, soil, and road, the measure used to assess the endmembers' fidelity is the SAD, described in the section of performance metrics in the Table. 2, is shown the performance obtained for the baseline algorithms and the NSAE-SU. Our algorithm achieved better performance for the water endmember extraction, obtaining 0.060 and 0.025 for the soil, and we achieved the best mean SAD. On the other hand, the abundance maps are evaluated for the RMSE metrics. In this case, our algorithm achieves better results for the water, 0.091, and soil, 0.187, compared to the other baseline algorithms.

Method/ Materials	NSAE-SU	UnDIP	CNNAEU	VCA
Water	**0.060**	0.130	0.113	0.200
Tree	0.029	**0.022**	0.041	0.055
Soil	**0.025**	0.040	0.048	1.839
Mean SAD	**0.038**	0.064	0.067	2.095

Table 2. Comparison between the models UnDIP Yu et al. (2022), CNNAEU Yu et al. (2022), VCA Ranasinghe et al. (2020), and our NSAE-SU model for performing the endmember extraction for the Samson dataset using the spectral angle distance metric.

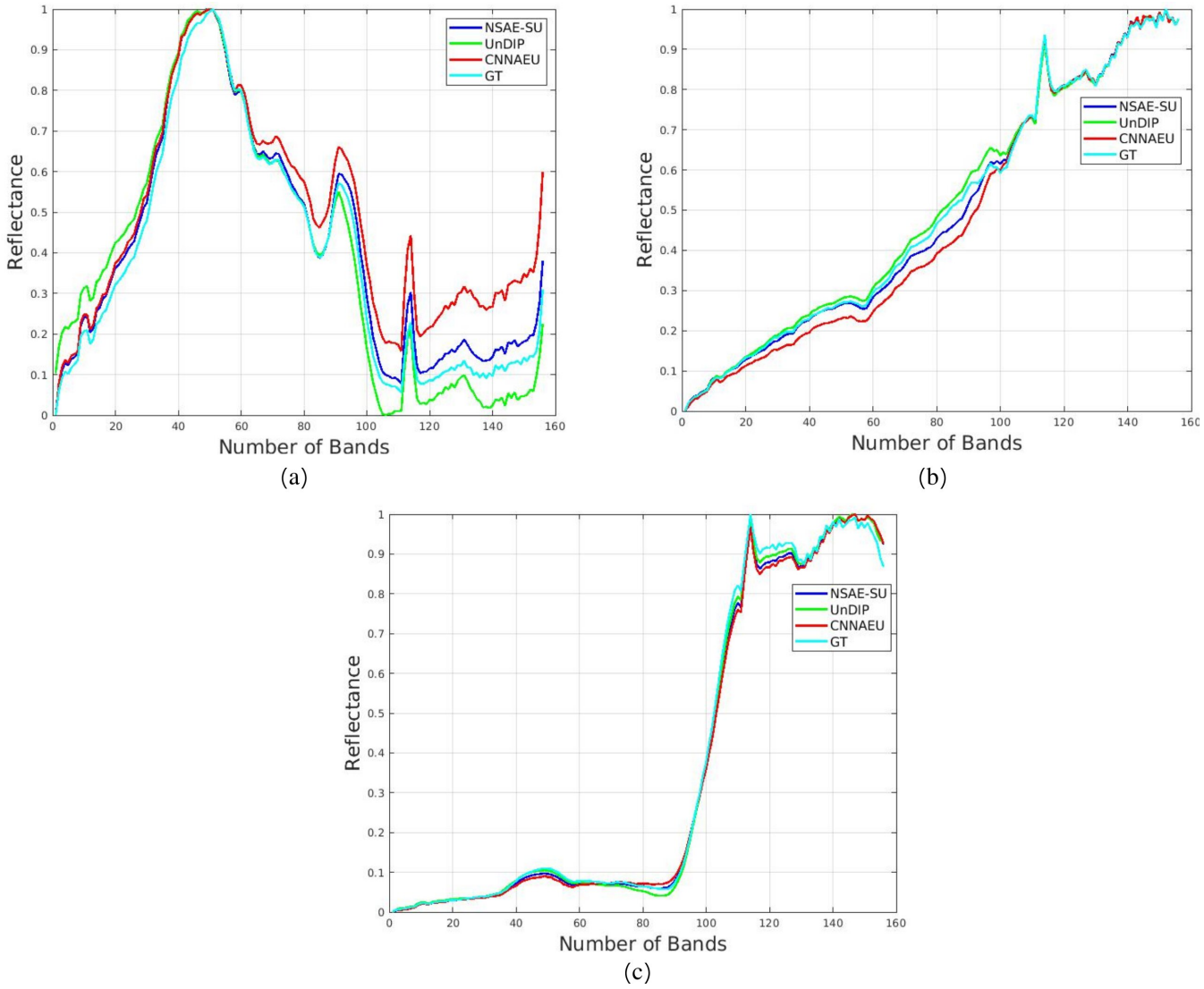


Figure 4. Extracted endmembers for Samson dataset using the techniques CNNAE, UnDIP, and our NSAE-SU model compared with the ground truth. The colors for the endmembers are red-CNNAE, green-UnDIP, blue-NSAE-SU, and cyan -ground truth. (a) water spectral signature, (b) soil spectral signature, and (c) tree spectral signature.

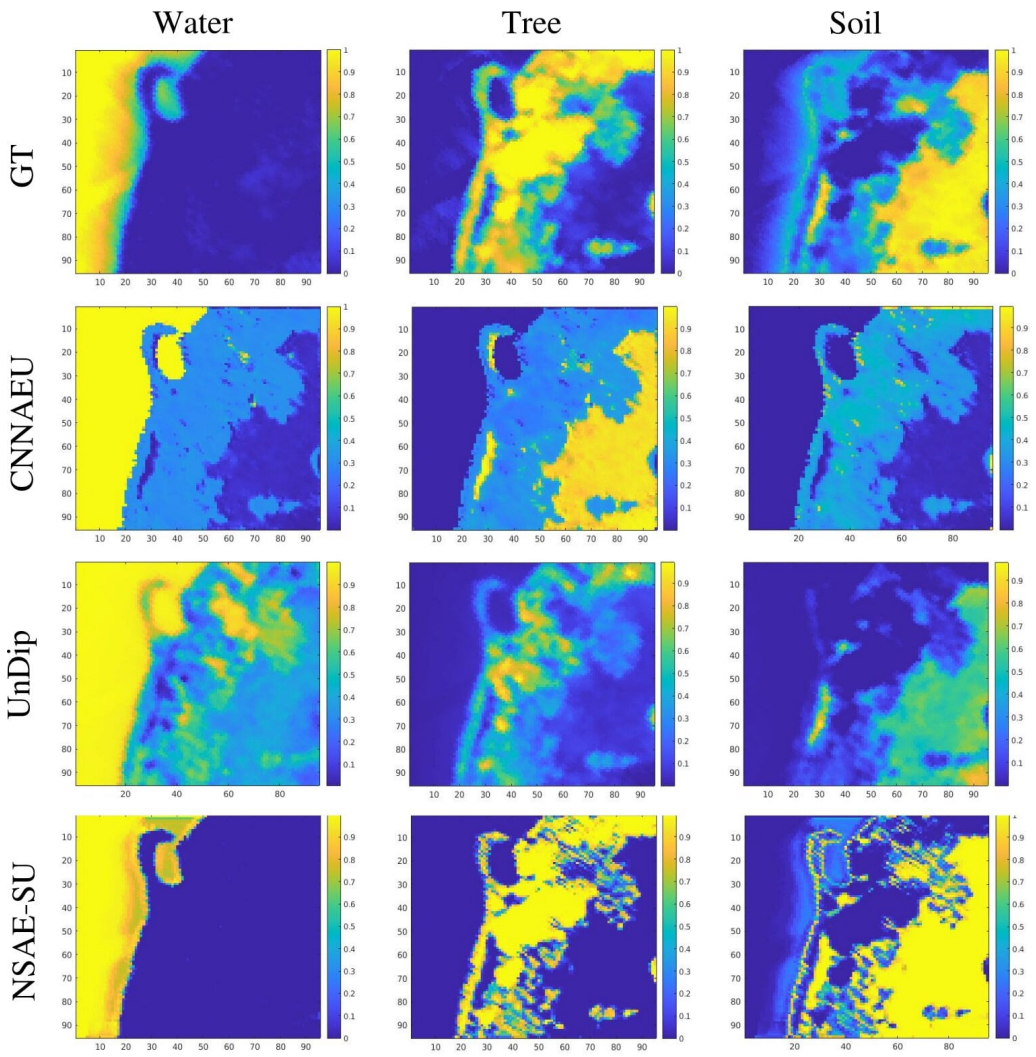


Figure 5. Comparison of the abundance maps extracted from Samson between with the ground Truth (GT) or reference maps and the following models: CNNAEU Palsson et al. (2021), UnDip Rasti et al. (2022), and NSAE-SU.

Method/ Materials	NSAE-SU	UnDIP	CNNAEU	VCA
Water	**0.091**	0.426	0.202	1.111
Tree	**0.172**	0.252	**0.172**	0.245
Soil	**0.187**	0.267	0.198	1.284
Mean RMSE	**0.150**	0.315	0.190	0.879

Table 3. Comparison between the models: UnDIP Yu et al. (2022), CNNAEU Yu et al. (2022), VCA Ranasinghe et al. (2020), and our NSAE-SU model for performing the abundance map estimation for the Samson dataset using the root mean square error metric.

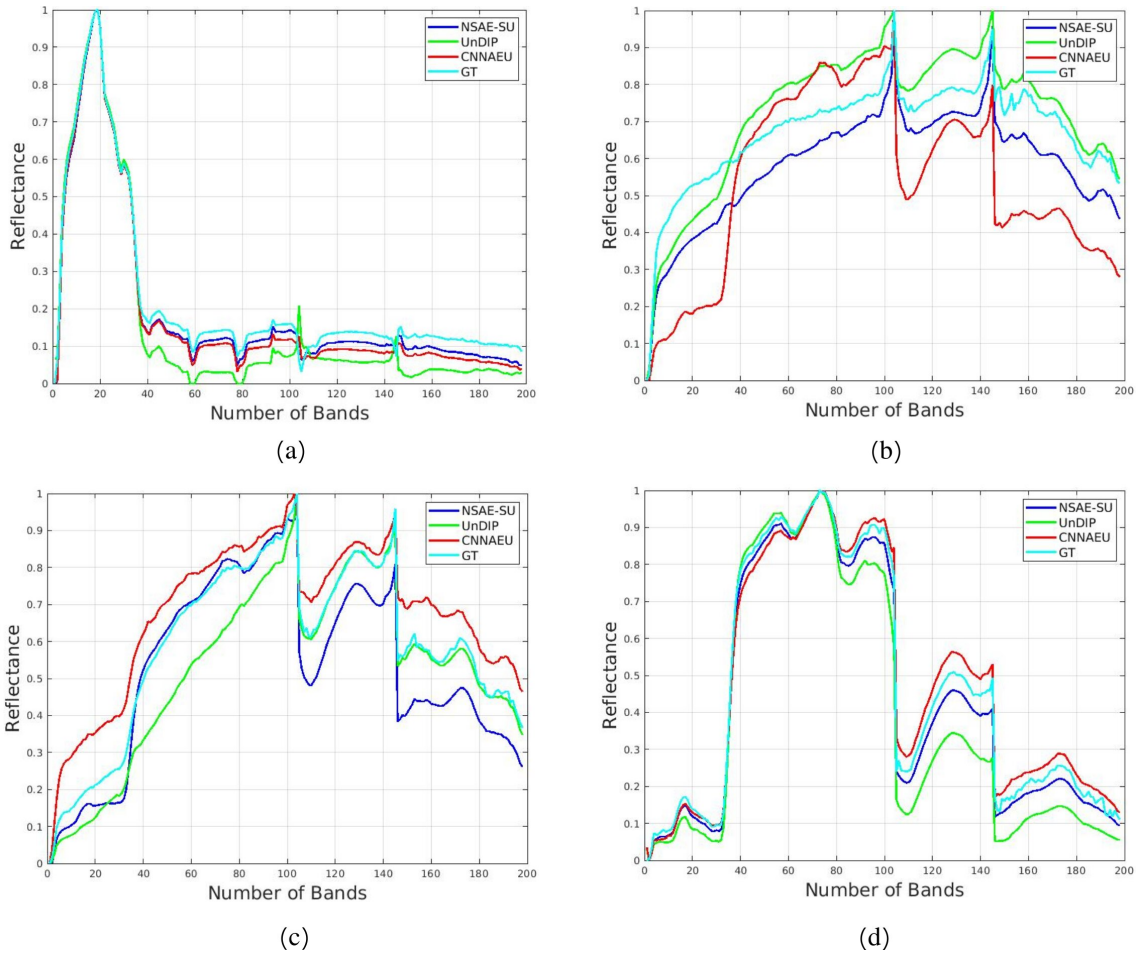


Figure 6. Extracted endmembers for the Jasper dataset performed by the techniques CNNAE, UnDIP, and our NSAE-SU model and compared with the ground truth. The colors of the endmembers are red-CNNAE, green-UnDIP, blue-NSAE-SU, and cyan for the ground truth. (a) water spectral signature, (b) road spectral signature, (c) soil spectral signature, and (d) tree spectral signature.

Method/ Materials	NSAE-SU	UnDIP	CNNAEU	VCA
Water	0.077	0.252	**0.061**	0.139
Tree	**0.039**	0.149	0.060	0.405
Soil	0.118	**0.114**	0.140	1.535
Road	**0.068**	0.086	0.134	0.530
Mean SAD	**0.076**	0.150	0.099	0.652

Table 4. Comparison between the models: UnDIP Yu et al. (2022), CNNAEU Yu et al. (2022), VCA Ranasinghe et al. (2020), and our NSAE-SU model for performing the endmember extraction for the Jasper dataset using the spectral angle distance metric.

The second benchmark dataset studied is Jasper Ridge; this dataset has four endmembers, and the ground truth for the abundances maps is given in Table. 4. This table gives the performance for the endmember

extraction measured by the spectral angle distance; the NSAE-SU achieved better results for the tree endmember, for which the SAD is 0.039, 0.068 for road. The water endmember is the second best result with 0.077. The NSAE-SU method also performs well for the mean SAD, it exhibit better results than the baseline algorithms with a mean of 0.076. For the abundance maps, better results are obtained for the water abundance map with RMSE of 0.1121, and 0.2316 for the soil (see Table. 5).

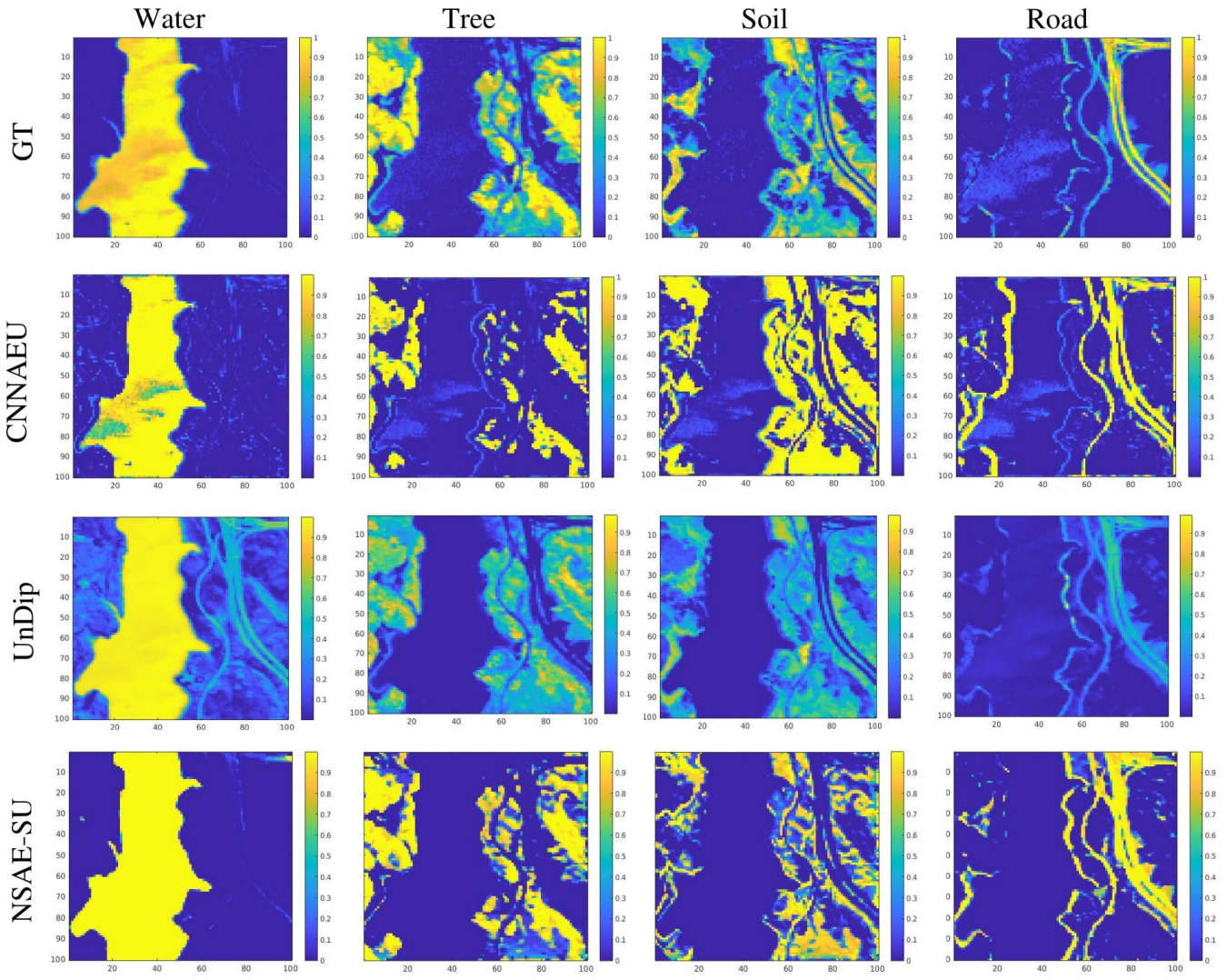


Figure 7. Comparison of the abundance maps extracted from Jasper with the ground Truth (GT) or reference maps and the following models CNNAEU Palsson et al. (2021), UnDip Rasti et al. (2022), and NSAE-SU.

Method/ Materials	NSAE-SU	UnDIP	CNNAEU	VCA
Water	**0.112**	0.201	0.183	2.212
Tree	0.172	**0.160**	0.199	0.380
Soil	**0.232**	0.132	0.294	1.754
Road	0.192	**0.109**	0.308	0.264
Mean RMSE	0.177	0.150	0.246	1.152

Table 5. Comparison between the models UnDIP Yu et al. (2022), CNNAEU Yu et al. (2022), VCA Ranasinghe et al. (2020), and the NSAE-SU model for performing the abundance map estimation for the Jasper dataset using the root mean square error metric.

6.2 Endmembers extraction and abundance maps estimation for HSI2 Lake Erie Hyperspectral image

The NSAE-SU model is used to perform the endmember extraction and the computation of the fractional abundance maps for the HSI2 image in order to establish the different spectral signatures that the image exhibits. The experiment is conducted on the five ROIs selected previously and classified in Manian et al. (2022). The selected ROIs are called as follows: red ROI with an area of 25.730 pixels, green ROI with an area of 25.344 pixels, cyan ROI with an area of 19.430, blue ROI with an area of 17.856 pixels, and yellow ROI with an area of 19.296 pixels, as depicted in Figure 3. Once the ROIs are selected, a spectral derivative is used to perform the atmospheric correction, using $k = 3$, defined heuristically.

The data is then distributed in a patch size selection of (9, 9, 170); the number of patches changes due to the difference in the area of the ROIs, as shown in Table. 6 . Subsequently, each ROI is analyzed, and the abundance maps and endmembers are extracted, as shown in the Figure. 8.

ROI	Area	Patch size
Red	25.730	(23068, 9, 9, 170)
Green	25.344	(22680, 9, 9, 170)
Cyan	19.340	(17125, 9, 9, 170)
Blue	17.856	(15640, 9, 9, 170)
Yellow	19.296	(17000, 9, 9, 170)

Table 6. Description of the number of patch sizes which are input to the NSAE-SU model. Each ROI is stacked with size $9 \times 9 \times 170$ for each area.

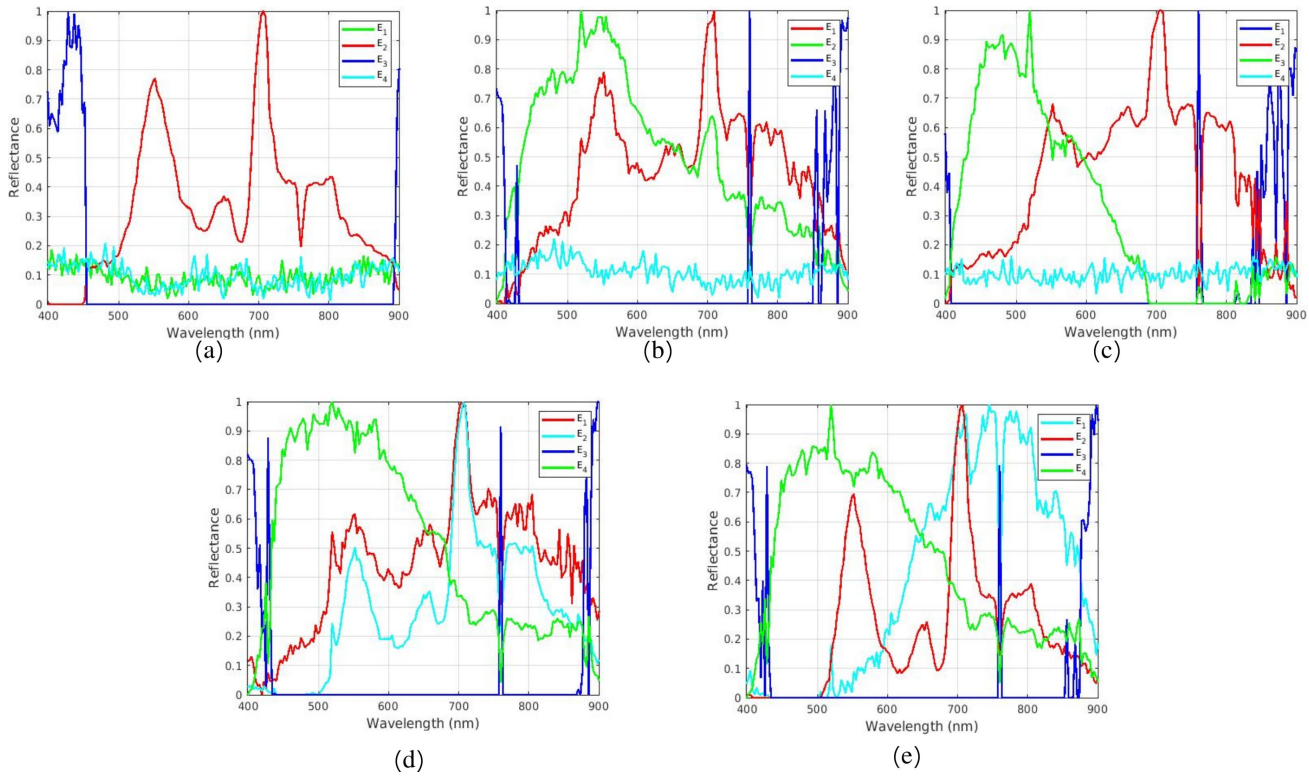


Figure 8. Endmembers extracted from the HSI2 Lake Erie image using the NSAE-SU for each ROI, as follows, (a) red ROI, (b) green ROI, (c) cyan Roi, (d) blue ROI, and (e) yellow ROI.

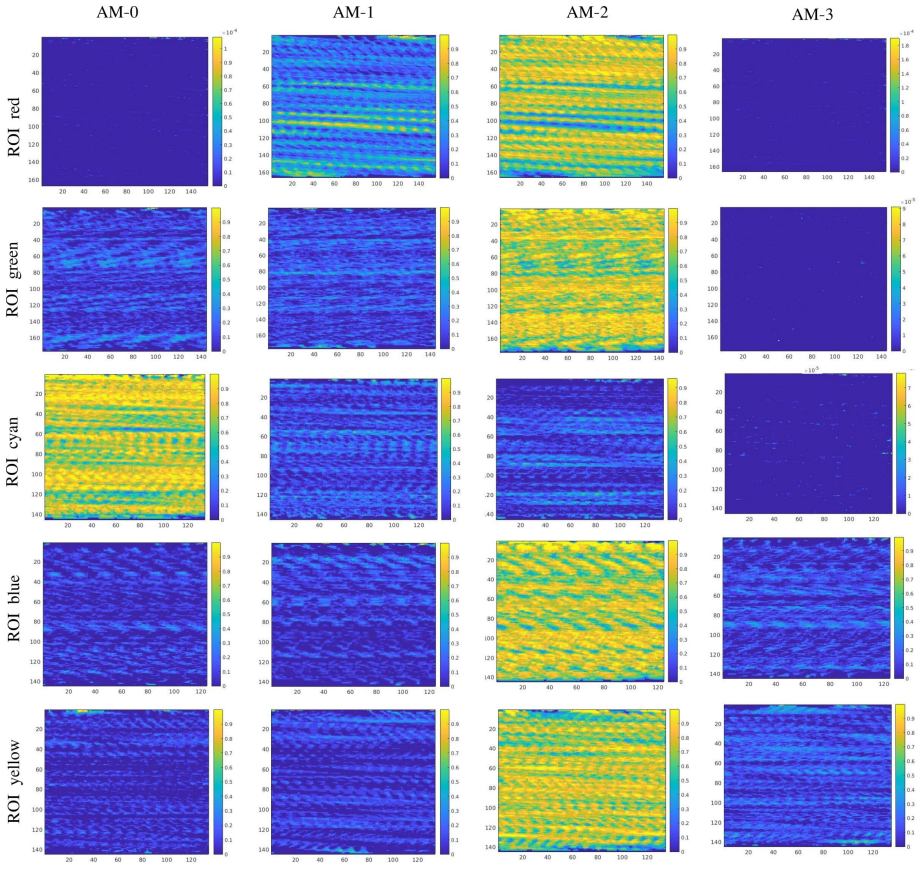
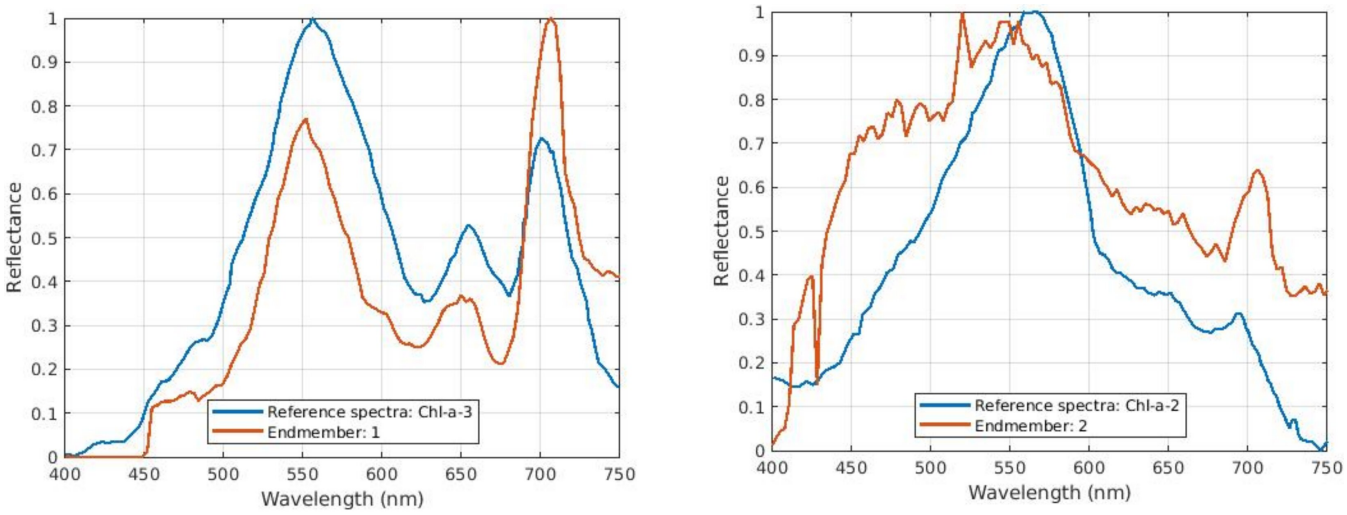


Figure 9. Abundance maps extracted from the selected ROIs for the HSI2 image, where AM-0, AM-1, AM-2, and AM-3 represent the number of the abundance map associated with the endmembers.

It is essential to compare the recovered endmembers with baseline spectral signatures to conduct the analysis; for this case study, we propose two comparison methods. The first is a comparison method to analyze the presence of Chl-a spectral signature, which is provided in the paper Ficek et al. (2011) with eight curves that possess the following Chl-a concentrations, respectively for each curve: $0.020mgL^{-1}$, $0.038mgL^{-1}$, $0.052mgL^{-1}$, $0.112mgL^{-1}$, $0.276mgL^{-1}$, $0.742mgL^{-1}$, $0.966mgL^{-1}$, $1.660mgL^{-1}$. Each concentration curve for this analysis will be called Chl-a-1, Chl-a-2, Chl-a-3, Chl-a-4, Chl-a-5, and Chl-a-6, respectively. The second method performs the analysis of the spectral signatures as extracted from Liang et al. (2017) of Cyanobacteria scum, Nymphoides, Potamogeton, and different proportions of Chl-a as follows: suspended solid (SS) concentrations $266.2mgL^{-1}$, Chl-a $0.0083mgL^{-1}$, SS $228.7mgL^{-1}$, Chl-a $0.0077mgL^{-1}$, SS $127.7mgL^{-1}$, Chl-a $0.0034mgL^{-1}$, SS $65.9mgL^{-1}$,Chl-a $0.0023mgL^{-1}$, SS $28.8mgL^{-1}$, Chl-a $0.0024mgL^{-1}$, SS $21.2mgL^{-1}$, Chl-a $0.0057mgL^{-1}$. Each concentration curve of Chl-a is called: Chl-a-11, Chl-a-22, Chl-a-33, Chl-a-44, Chl-a-55, and Chl-a-66, respectively to distinguish between the labels provided for the first comparison. The SAD metric is employed, the reference spectra are the spectral signatures offered in papers Ficek et al. (2011), Liang et al. (2017), and the endmember estimation is carried out by NSAE-SU model to execute the comparison.



(a) (b)

Figure 10. Comparison between the spectra provided in Ficek et al. (2011), as ground truth, and the spectra extracted by the NSAE-SU model illustrating the best match using the SAD metric for the red and green ROI, respectively. (a) The best match obtained for the red ROI is the Chl-a-3 spectra and the endmember 1 extracted from NSAE-SU. (b) The best match obtained for the green roi is the Chl-a-2 spectra and the endmember 1 extracted from NSAE-SU.

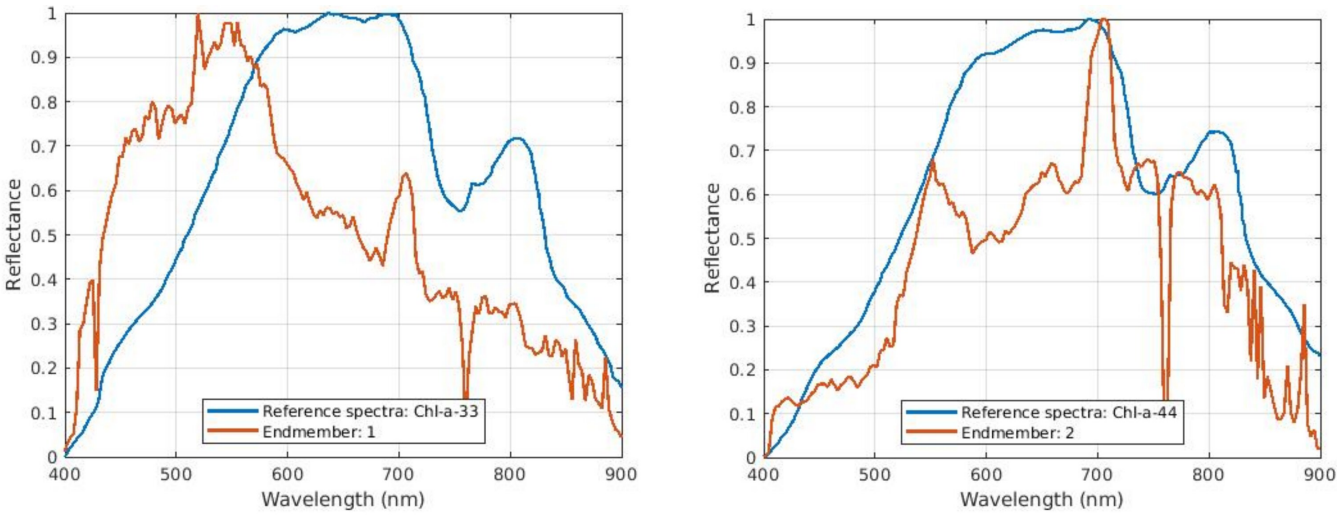
In Figure. 10 (a), the comparison between the reference spectra provided by Ficek et al. (2011), and the spectra obtained for NSAE-SU for the red ROI is shown. The best approximation for the SAD metric is 0.369, which corresponds to Chl-a-3, that possessed $0.052mgL^{-1}$ of the content of Chl-a with the endmember 1 obtained from NSAE-SU. Figure .10(b) analyzes the green ROI whose best match according to the spectral angle distance is 0.311, corresponding to the combination of Chl-a-2 and the second endmember extracted from NSAE-SU. The SAD metric obtained for each region of interest using the spectra of Ficek et al. (2011) as ground truth, and the spectra obtained using the NSAE-SU method as the estimated endmembers are summarized in Table .7, for the best endmembers by each ROI. The comparative analysis for the endmembers obtained from NSAE-SU and the ground truth provided by Liang et al. (2017)

are assessed by the SAD metric for each ROI; the results are summarized in Table. 8, the term endmember is used to refer to the spectral signatures extracted from the model NSAE-SU.

The concentration of Chl-a-33 compared to endmember 1 for the green ROI, Chl-a-44 compared to endmember 1 for the cyan ROI, and Chl-a-33 compared to endmember 2 for the yellow ROI provides the best curve fitting for the ROIs using the SAD metric. In Figure. 10(b) the best curve fitting according to the SAD metric for green and cyan ROIs are depicted. The best concentration of Chl-a for Chl-a-33 is Chl-a $0.0034 mgL^{-1}$, and for Chl-a-44 is Chl-a $0.0023 mgL^{-1}$, respectively.

	Red Roi	Green Roi	Cyan Roi	Blue Roi	Yellow Roi
Chl-a-1	0.771	0.334	**0.382**	0.339	0.364
Chl-a-2	0.738	**0.311**	0.416	**0.307**	**0.328**
Chl-a-3	**0.369**	0.352	0.754	0.463	0.456
Chl-a-4	0.453	0.417	0.798	0.507	0.490
Chl-a-5	0.550	0.408	0.742	0.476	0.453
Chl-a-6	0.605	0.567	0.896	0.629	0.601

Table 7. Results for the SAD metric comparison of the spectral signatures of different Chl-a concentrations extracted from Ficek et al. (2011), with the endmembers obtained from NSAE-SU.



(a) (b)

Figure 11. Comparison between the spectra provided from Liang et al. (2017), as ground truth, and the spectra extracted by the NSAE-SU model according to the SAD metric for the green and cyan ROI, respectively. (a) The best match for the green ROI is between the Chl-a-33 spectra and the endmember 1 extracted from NSAE-SU. (b) The best match for the cyan roi is between the Chl-a-44 spectra and the endmember 2 extracted from NSAE-SU.

	Red Roi	Green Roi	Cyan Roi	Blue Roi	Yellow Roi
Cyanobacterial scums	0,741	0,639	0,739	0,545	0,798
Nymphoides peltatum	0,845	0,746	0,834	0,635	0,885
Potamogeton crispus	0,786	0,695	0,762	0,587	0,799
Chl-a-11	0,540	0,447	0,447	0,547	0,754
Chl-a-22	0,654	0,617	0,640	0,733	0,857
Chl-a-33	**0,448**	**0,307**	0,319	0,376	**0,658**
Chl-a-44	0,477	0,321	**0,307**	**0,356**	0,669
Chl-a-55	0,616	0,549	0,547	0,650	0,822
Chl-a-66	0,619	0,551	0,563	0,659	0,832

Table 8. Results for the SAD metric comparing the spectral signatures of different Chl-a and macrophytes under different concentrations extracted from Liang et al. (2017), with the endmembers obtained from NSAE-SU.

7 CONCLUSIONS

An unsupervised deep learning model for the endmember extraction and fractional abundance map estimation is presented. The NSAE-SU model performs well for the benchmark datasets, such as Samson and Jasper, and for the study case of the HSI2 image over Lake Erie. The model is able to identify the endmembers for the water, soil, and road, and the abundances maps for water, road, and trees better than the baseline algorithms. Additionally, the spectral signatures extracted using the NSAE-SU model over Lake Erie hyperspectral image is analyzed to determine the presence of Chl-a. The $9x9$ patch size is determined to be the ideal configuration, and the best hyperparameter settings for the model are listed in the Table. 1. Additionally, the models operate at a nominal speed of approximately 3 hours and 45 minutes.

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

EAM, VM, JO, RT: conceptualization, EAM, VM, JO, RF: methodology, EAM: software, EAM: validation, EAM: formal analysis, EAM, VM, JO, RF: investigation, EAM: data curation, EAM:writing—original draft preparation, EAM, VM:writing—review and editing, VM: project administration, VM:funding acquisition; .All authors contributed to the article and approved the submitted version.

FUNDING

This research was funded by NASA, grant number 80NSSC21M0155. The APC is funded by 80NSSC21M0155. Opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NASA.

ACKNOWLEDGMENTS

We would like to acknowledge the contribution of Mr. Carlos Julian Delgado Munoz for the discussions on the cluster configuration.

DATA AVAILABILITY STATEMENT

The Jasper and Samson hyperspectral images datasets, are available at: https://rslab.ut.ac.ir/data (accessed on 15 April 2023), and the Lake Erie hyperspectral image datasets are available at: https://oceandata.sci.gsfc.nasa.gov/directdataaccess/HSI-HABS-RAW/NASA_GRC_2015/20150727/Proc/ (accessed on 14 April 2023).

REFERENCES

Chong, J. W. R., Khoo, K. S., Chew, K. W., Vo, D.-V. N., Balakrishnan, D., Banat, F., et al. (2023). Microalgae identification: Future of image processing and digital algorithm. *Bioresource Technology* 369, 128418. doi:10.1016/j.biortech.2022.128418

Chunhui, Z., Bing, G., Lejun, Z., and Xiaoqing, W. (2018). Classification of hyperspectral imagery based on spectral gradient, SVM and spatial random forest. *Infrared Physics &amp Technology* 95, 61–69. doi:10.1016/j.infrared.2018.10.012

Das, S., Chakraborty, S., Routray, A., and Deb, A. K. (2019). Fast linear unmixing of hyperspectral image by slow feature analysis and simplex volume ratio approach. In *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium* (IEEE). doi:10.1109/igarss.2019.8898127

Ficek, D., Zapadka, T., and Dera, J. (2011). Remote sensing reflectance of pomeranian lakes and the baltic.the study was partially financed by MNiSW (ministry of science and higher education) as a research project n n306 066434 in the years 2008–2011. *Oceanologia* 53, 959–970. doi:10.5697/oc.53-4.959

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning* (MIT Press). `http://www.deeplearningbook.org`

Guo, H., Tian, S., Huang, J. J., Zhu, X., Wang, B., and Zhang, Z. (2022). Performance of deep learning in mapping water quality of lake simcoe with long-term landsat archive. *ISPRS Journal of Photogrammetry and Remote Sensing* 183, 451–469. doi:10.1016/j.isprsjprs.2021.11.023

Hadi, F., Yang, J., Ullah, M., Ahmad, I., Farooque, G., and Xiao, L. (2022). DHCAE: Deep hybrid convolutional autoencoder approach for robust supervised hyperspectral unmixing. *Remote Sensing* 14, 4433. doi:10.3390/rs14184433

Heylen, R., Parente, M., and Gader, P. (2014). A review of nonlinear hyperspectral unmixing methods. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7, 1844–1868. doi:10.1109/jstars.2014.2320576

Imbiriba, T., Borsoi, R. A., and Bermudez, J. C. M. (2018). Generalized linear mixing model accounting for endmember variability. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE). doi:10.1109/icassp.2018.8462214

Liang, Q., Zhang, Y., Ma, R., Loiselle, S., Li, J., and Hu, M. (2017). A modis-based novel method to distinguish surface cyanobacterial scums and aquatic macrophytes in lake taihu. *Remote Sensing* 9. doi:10.3390/rs9020133

Manian, V., Alfaro-Mejía, E., and Tokars, R. P. (2022). Hyperspectral image labeling and classification using an ensemble semi-supervised machine learning approach. *Sensors* 22. doi:10.3390/s22041623

Nascimento, J. and Dias, J. (2005). Vertex component analysis: a fast algorithm to unmix hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing* 43, 898–910. doi:10.1109/tgrs.2005.844293

Ortiz, J. D., Avouris, D. M., Schiller, S. J., Luvall, J. C., Lekki, J. D., Tokars, R. P., et al. (2019). Evaluating visible derivative spectroscopy by varimax-rotated, principal component analysis of aerial hyperspectral images from the western basin of lake erie. *Journal of Great Lakes Research* 45, 522–535. doi:10.1016/j.jglr.2019.03.005

Palsson, B., Ulfarsson, M. O., and Sveinsson, J. R. (2021). Convolutional autoencoder for spectral–spatial hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing* 59, 535–549. doi:10.1109/tgrs.2020.2992743

Park, J., Baek, J., Kim, J., You, K., and Kim, K. (2022). Deep learning-based algal detection model development considering field application. *Water* 14, 1275. doi:10.3390/w14081275

Qi, L., Chen, Z., Gao, F., Dong, J., Gao, X., and Du, Q. (2023). Multiview spatial–spectral two-stream network for hyperspectral image unmixing. *IEEE Transactions on Geoscience and Remote Sensing* 61, 1–16. doi:10.1109/tgrs.2023.3237556

Ranasinghe, Y., Herath, S., Weerasooriya, K., Ekanayake, M., Godaliyadda, R., Ekanayake, P., et al. (2020). Convolutional autoencoder for blind hyperspectral image unmixing. In *2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS)* (IEEE). doi:10.1109/iciis51140.2020.9342727

Rasti, B., Koirala, B., Scheunders, P., and Ghamisi, P. (2022). UnDIP: Hyperspectral unmixing using deep image prior. *IEEE Transactions on Geoscience and Remote Sensing* 60, 1–15. doi:10.1109/tgrs.2021.3067802

Shah, D., Zaveri, T., Trivedi, Y. N., and Plaza, A. (2020). Entropy-based convex set optimization for spatial–spectral endmember extraction from hyperspectral images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 13, 4200–4213. doi:10.1109/jstars.2020.3008939

Tane, Z., Roberts, D., Veraverbeke, S., Casas, Á., Ramirez, C., and Ustin, S. (2018). Evaluating endmember and band selection techniques for multiple endmember spectral mixture analysis using post-fire imaging spectroscopy. *Remote Sensing* 10, 389. doi:10.3390/rs10030389

Vivone, G. (2023). Multispectral and hyperspectral image fusion in remote sensing: A survey. *Information Fusion* 89, 405–417. doi:10.1016/j.inffus.2022.08.032

Waga, H., Eicken, H., Light, B., and Fukamachi, Y. (2022). A neural network-based method for satellite-based mapping of sediment-laden sea ice in the arctic. *Remote Sensing of Environment* 270, 112861. doi:10.1016/j.rse.2021.112861

Wang, L., Liu, D., Wang, Q., and Wang, Y. (2013). Spectral unmixing model based on least squares support vector machine with unmixing residue constraints. *IEEE Geoscience and Remote Sensing Letters* 10, 1592–1596. doi:10.1109/lgrs.2013.2262371

Winter, M. E. (1999). &lttitle&gtn-FINDR: an algorithm for fast autonomous spectral end-member determination in hyperspectral data&lt/title&gt. In *SPIE Proceedings*, eds. M. R. Descour and S. S. Shen (SPIE). doi:10.1117/12.366289

Xu, Y., Du, B., Zhang, L., Cerra, D., Pato, M., Carmona, E., et al. (2019). Advanced multi-sensor optical remote sensing for urban land use and land cover classification: Outcome of the 2018 IEEE GRSS data fusion contest. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 12, 1709–1724. doi:10.1109/jstars.2019.2911113

Yang, X., Chu, Q., Wang, L., and Yu, M. (2022). Water body super-resolution mapping based on multiple endmember spectral mixture analysis and multiscale spatio-temporal dependence. *Remote Sensing* 14, 2050. doi:10.3390/rs14092050

Yu, Y., Ma, Y., Mei, X., Fan, F., Huang, J., and Li, H. (2022). Multi-stage convolutional autoencoder network for hyperspectral unmixing. *International Journal of Applied Earth Observation and Geoinformation* 113, 102981. doi:https://doi.org/10.1016/j.jag.2022.102981

Zhang, G., Mei, S., Xie, B., Feng, Y., and Du, Q. (2022). Spectral variability augmented two-stream network for hyperspectral sparse unmixing. *IEEE Geoscience and Remote Sensing Letters* 19, 1–5. doi:10.1109/lgrs.2022.3214843

Zhong, Y., Wang, X., Xu, Y., Wang, S., Jia, T., Hu, X., et al. (2018). Mini-UAV-borne hyperspectral remote sensing: From observation and processing to applications. *IEEE Geoscience and Remote Sensing Magazine* 6, 46–62. doi:10.1109/mgrs.2018.2867592

Zhu, X., Guo, H., Huang, J. J., Tian, S., and Zhang, Z. (2023). A hybrid decomposition and machine learning model for forecasting chlorophyll-a and total nitrogen concentration in coastal waters. *Journal of Hydrology* 619, 129207. doi:10.1016/j.jhydrol.2023.129207