GAUSS: Guided encoder - decoder Architecture for hyperspectral Unmixing with Spatial Smoothness

H.M.K.D. Wickramathilaka, D. Fernando, D. Jayasundara, D. Wickramasinghe, D.Y.L. Ranasinghe, G.M.R. I. Godaliyadda, M.P.B. Ekanayake, H.M.V.R. Herath, L. Ramanayake, N. Senarath and H.M.H.K. Weerasooriya

Department of Electrical and Electronic Engineering, University of Peradeniya, Peradeniya, Sri Lanka

ABSTRACT
This study introduces GAUSS (Guided encoder-decoder Architecture for hyperspectral Unmixing with Spatial Smoothness), a novel autoencoder-based architecture for hyperspectral unmixing (HU). GAUSS consists of an Approximation Network (AN), Unmixing Network (UN), and a Mixing Network (MN). The AN incorporates spatial context within a hyperspectral pixel’s neighborhood, while the UN utilizes a pseudo-ground truth mechanism to enhance abundance estimation. The MN provides estimated endmembers’ signatures. By incorporating UN-produced abundances, unlike the conventional AE model, GAUSS overcomes the single-layer constraint of the MN. Thereafter, a secondary training phase improves the accuracy of endmembers and abundance estimation using a reliable Signal Processing (SP) algorithm, resulting in superior HU performance. The results demonstrate the effectiveness of GAUSS on two Standard datasets and a Simulated dataset compared to the state-of-the-art SP and Deep Learning (DL) based methods. This signifies the benefit of integrating an SP algorithm in the training process, contributing to advancements in DL-based HU techniques.

Introduction
Hyperspectral (HS) imaging is extensively used in earth observation satellites for various applications (Khan et al., 2018): lithological mapping (Chen et al., 2019; Kruse et al., 2003; Liu et al., 2018; Lorenz et al., 2018; Ni et al., 2020), environmental monitoring (Moroni et al., 2013; Stuart et al., 2019; Zeng et al., 2019), and agricultural activities (Lu et al., 2020; Teke et al., 2013; Zhu et al., 2020), since it captures more spectral information than its counterparts: multispectral and RGB imaging. However, HS images (HSIs) are of poor spatial resolution and are corrupted by noise and interference from atmospheric alterations and instrumentation imperfections (Rasti et al., 2018; Zhou et al., 2022).

Herefore, identifying the composition of a topography requires information recovery algorithms, known as unmixing algorithms.

HS unmixing (HU) is a source separation technique Guan et al. (2022) that strives to extract the constituent source spectral signatures and estimate their fractional presence in each pixel, referred to as endmembers and abundances. Over the evolution of HU techniques, numerous unmixing algorithms have been proposed for both linear and non-linear modeling methods (Dobigeon et al., 2013; Gu et al., 2021; Halimi et al., 2011; Heylen et al., 2014; Zhao & Zhao, 2019) under signal processing (SP) and deep learning (DL) domains (Wang et al., 2021). When considering the SP approaches (Nascimento & Bioucas-Dias, 2012), have proposed an unsupervised hyperspectral unmixing (HU) method suitable for linear and highly mixed hyperspectral datasets. This method utilizes a statistical approach and incorporates a prior abundance fraction model based on a mixture of Dirichlet densities.

In the study done by (Das et al., 2017), the authors introduce a method that is both fast and accurate for the complete estimation of the number of endmembers in an HS image. To accomplish this, the authors rely on the convex properties of endmember abundances, indicating a mathematical approach to solving the problem.

Moreover, sparse hyperspectral unmixing methods aim to estimate the abundance of endmembers, assuming a spectral library as an overcomplete set of endmembers. (Das & Routray, 2019) introduce an efficient dictionary pruning approach for hyperspectral unmixing. Whereas in the work done by (Ma et al., 2017), the authors introduce a robust SU (RSU) method that employs the $\ell_{2,1}$ norm loss function, which enhances robustness to noise and outliers. Along the same line of methods which utilize spectral libraries (Das et al., 2019), introduces a semi-supervised, and parameter-
free algorithm for hyperspectral image unmixing that utilizes sparsity measures for library pruning.

Largely in the literature, SP unmixing benefits from various algorithmic approaches (Bioucas-Dias et al., 2012) such as geometric (Bioucas-Dias, 2009; Boardman, 1993; Gader et al., 2012; Li et al., 2016; Miao & Qi, 2007; Nascimento & Dias, 2005; Winter, 1999; Zare & Gader, 2007; Zhuang et al., 2019), statistical (Dobigeon et al., 2009; Nascimento & Bioucas-Dias, 2007; Parra et al., 1999), and sparse regression-based (Bioucas-Dias & Figueiredo, 2010; Iordache et al., 2012). Furthermore, Non-negative matrix factorization (NMF) (Lee & Seung, 2000) is used in HU to extract endmembers and abundances with matrix decomposition and adhere to constraints imposed by practical realities. The introduced variants on the fundamental NMF algorithm can promote sparsity of abundances (He et al., 2017; Qian et al., 2010), independence of endmembers (Ekanayake et al., 2021), spatial-spectral information, and piece-wise smoothness (Rathnayake et al., 2020; Sun et al., 2018; Yang et al., 2015; Zhou et al., 2020).

Moreover, DL has also been utilized for processing remote-sensing images. Recent research, as illustrated by (Hong et al., 2021), investigates the use of a multi-modal framework (MDL) to tackle the challenges of classifying materials in diverse scenes, with a focus on cross-modality learning (CML) in remote sensing (RS) image classification. This framework introduces various fusion strategies to combine information from different modalities, allowing for both pixel-wise classification and spatial information modeling through convolutional neural networks (CNNs).

In another study (Hong et al., 2021), propose a mini-batch GCN (miniGCN) for efficient large-scale GCN training in remote sensing (RS) problems. They explore three fusion strategies (additive, element-wise multiplicative, and concatenation) for combining CNNs and miniGCNs to extract diverse hyperspectral features.

On the other hand, the study done by (Li et al., 2023) considers a baseline network called LRR-Net, which combines the low-rank representation model (LRR) with deep learning techniques. The LRR-Net uses the alternating direction method of multipliers (ADMM) optimizer to efficiently solve the LRR model and incorporates the solution as prior knowledge for optimizing parameters in the deep network.

(Wu et al., 2023) utilize a U-Net for infrared small object detection. The proposed U-Net incorporates a small U-Net into a larger U-Net backbone, facilitating multi-level and multi-scale representation learning of objects. The model consists of two modules: the resolution-maintenance deep supervision (RM-DS) module and the interactive-cross attention (IC-A) module. The RM-DS uses Residual U-blocks to maintain multi-scale resolution features while learning global context information. Meanwhile, the IC-A encodes local context information between low-level details and high-level semantic features.

Furthermore, applications of DL are becoming increasingly popular in HU problems due to its computational power, contextual features, and the potential to extract information from unstructured data (Signoroni et al., 2019).

CNNs (Fang et al., 2022; Han et al., 2022; Palsson et al., 2020; Qi et al., 2020; Ranasinghe et al., 2020a; Wang et al., 2019; Zhang et al., 2018; Zhao et al., 2021, 2022), dense networks (Han et al., 2020; Hong et al., 2021a; Ozkan et al., 2018; Qu & Qi, 2018; Su et al., 2019), and recurrent networks (Zhao et al., 2021) are extensively used for semi-supervised or unsupervised DL unmixing (Benabdi et al., 2019), which use functional layers like pooling, regularization, and shaping. Nonetheless, most DL unmixing algorithms are structured as autoencoders (AEs) due to their resemblance to matrix decomposition used in blind source separation techniques and their unsupervised learning capabilities.

Even though AEs are widely used in DL unmixing, they have architectural limitations when used for blind unmixing. In particular,

- The encoder optimization depends on the decoder optimization due to the end-to-end structure.
- The convergence of the encoder is slower due to the absence of a direct steering mechanism other than the decoder.

Besides, according to the ablation studies in (Palsson et al., 2019; Qi et al., 2020; Ranasinghe et al., 2020a), instead of learning the associations between the pixels, the convolutional layers only distill abstract spatial patterns from the HSI into the abundance maps. Furthermore, due to the fact that the loss function uses the original HSI as its reference, there is no forcing mechanism to maintain the spatial correlation in abundance values. Moreover, in the end-to-end architecture of the AE for blind unmixing, the decoder will be restricted to a single-layer network with linear activations to preserve the similarity with matrix decomposition (Palsson et al., 2018; Ranasinghe et al., 2020b).

In this regard, we propose a three-network HU architecture (HUA), which undergoes two training phases. The “Guided encoder-decoder Architecture for hyperspectral Unmixing with Spatial Smoothness” (GAUSS) consists of three networks. The first two networks: the “approximation network” (AN) and “unmixing network” (HU), are encoders, and the “mixing network” (MN) is the final network that serves as the decoder of the HUA.

The spatial context of hyperspectral images is an important aspect which is considered by NNMF algorithms when it comes to abundance map generation. In these SP-based algorithms, regularizers have been utilized to incorporate the aforementioned spatial
context. However, DL approaches lack a method of injecting such spatial contextual information to the unmixing process. As one of the contributions of this paper, the AN was proposed to accomplish this task. Specifically, the functionality of the proposed AN is to improve the association between a pixel referred to as the center pixel and its neighborhood in order to incorporate contextual spatial information in the abundance maps. Furthermore, the estimation through the neighborhood has also enabled the spectral correlation and smoothness of the given local environment to be distilled into the unmixing process. While it is susceptible to the limitations of the given environment, the generation process guarantees functional continuity among the pixels, maintaining spectral smoothness and continuity.

The autoencoder is decoupled into the encoder (UN) and the decoder (MN). This decoupling of the autoencoder enables the injection of a pseudo-ground truth into the training process. In order to provide a pseudo-ground truth, the HSI is segmented with a clustering algorithm, which is then converted to a one-hot encoded representation. The significance of the pseudo-ground truth is that it provides the spatial context and the fundamental texture for the UN’s abundance map. Therefore, one can think of the pseudo-ground truth as a gradient provider for the UN to optimize its weights, independent from the decoder gradients. This introduction of a pseudo-ground truth is another one of the contributions introduced in this paper. This is in contrast to the traditional optimization of the AEs’ unmixing procedure, which depends on the HS pixel or image reconstruction. Moreover, a single-layered decoder is present in most of the existing DL-based HU methods (Ozkan et al., 2019). However, the use of a pseudo-ground truth to train the UN before training the MN gives the decoder greater flexibility and makes it possible to add more layers to it. Additionally, because the UN was trained before the MN, the resulting abundance is guaranteed to be, at least, a rough estimate of the real abundance for the corresponding pixel. This allows the decoder of the network to generate endmembers that closely approximate the ground truth endmembers as the decoder is fed with the rich abundances provided by the encoder. Thus, the training of the decoder can be thought of as a fine-tuning of the model to find the optimal weights of the decoder.

Next, for improved performance of GAUSS, a secondary training phase has been introduced in which the model is retrained with a more refined pseudo-ground truth generated by a specifically selected SP algorithm (Qian et al., 2010) based on its unmixing performance. The output of the MN is the reconstructed pixel, with the output of the UN serving as its input. Hence, the functionality of the decoder is to find the layer weights which will eventually generate the spectral signature once the abundances are provided.

In summary, the primary contribution of this article is twofold.

1. The introduction of a unique deep learning architecture that enables the preservation of the local subspace structure’s spatial correlation while enhancing the neighboring pixels’ associations with the center pixel.

2. The introduction of a pseudo-ground truth at the end of the UN which decouples the encoder abundance map generation process from the decoder mixing for endmember generation. This allows the encoder and decoder to be optimized separately.

Furthermore, a summary of the proposed architecture and the training procedure is as follows.

- The GAUSS architecture has two encoders: the Approximation Network (AN) and the Unmixing Network (UN). The AN predicts the center pixel by utilizing the neighboring pixels, while the UN maps the center pixel to an optimal abundance using the pseudo-ground truth. The decoder, or Mixing Network (MN), reconstructs the center pixel from the bottleneck layer abundances, with MN weights providing endmember signatures. (See the Methodology Section for technical details.)

- The first training phase utilizes the one-hot encoded pseudo-ground truth to train the model. During the second training phase, the network is initialized with the weights and biases from the first training phase. Then, the model is trained with the abundances obtained by an SP algorithm, such that it aligns the model towards unmixing and is updated with more enriched endmember spectra provided by the SP algorithm.

The remainder of the paper is as follows: In the Materials Section, the concept behind the GAUSS architecture is delineated, followed by the Methodology Section which describes the implementation details of the HUA. In the next section, the details of the experiments and the results are discussed. Thereafter, an Ablation study and a Performance Analysis are provided. Finally, the conclusion is presented. Additionally, Table 1 provides a comprehensive summary of the abbreviations used in this study and their corresponding details.

**Materials**

**Notation**

Scalars are denoted by both lower ($x$) and uppercase ($X$) normal fonts. Boldface lowercase fonts ($\mathbf{x}$) are for 1-D arrays, and 1-D arrays are ordered as column vectors.
Then, boldface uppercase fonts (X) are used for both 2-D and 3-D arrays.

Consider a HSI, \( H \in \mathbb{R}^{w \times h \times B} \), with a width and a height of \( w \) and \( h \) in pixels, respectively, and \( B \) spectral bands. The HS datacube can be rearranged as a matrix of size \( B \times N \), where \( N \) is the total number of observations (pixels) and is the product of \( w \) and \( h \). Accordingly, the HSI can be denoted as \( Y = [y_1 \cdots y_j \cdots y_N] \in \mathbb{R}^{B \times N} \) and each column \( y_j \in \mathbb{R}^B \) is the spectrum of the \( j \)th pixel. The spectral signature of the pixel at the \((m, n)\) location of \( H \) corresponds to the \( j \)th column vector of \( X \), such that \( j = \{w(m - 1) + n|1 < m \leq h, 1 < n \leq w\} \) where \( m \) and \( n \) is the row and column position of the pixel, respectively.

**Mixture models**

The linear mixture model (LMM) depends on the assumption that the incident light is scattered from the macroscopic surface only once and is measured by the sensing device of the imaging spectrometer without scattering. For its simplicity, LMM is the most frequently used model for HSI in the literature (Heylen et al., 2014). The luxury of the LMM is that a linear combination of endmember spectra can be written for each observation as

\[
y_j = \sum_{k=1}^{K} S_{kj} a_k + e_j \tag{1}
\]

where \( S_{kj} \) is the fractional composition of the \( k \)th endmember in the \( j \)th pixel, \( a_k \in \mathbb{R}^B \) is the spectrum of the \( k \)th endmember of the HSI, \( e_j \in \mathbb{R}^B \) is an additive disruption due to noise and modeling errors, and \( K \) is the predetermined number of endmembers in the HSI. The elements in \( y_j \) and \( a_k \) are non-negative as the HSI is represented using the reflectance values because it inherits atmospheric corrections and sensor independencies. Due to physical implications, the abundance non-negativity constraint (ANC) \( S_{kj} \geq 0 \) and the abundance sum-to-one constraint (ASC) \( \sum_{k=1}^{K} S_{kj} = 1 \) are imposed in (1) to ensure that the fractional compositions are non-negative and the HSI only consists of the predetermined endmembers. Under the LMM, the HSI can be represented using the matrix notation as

\[
Y = A \times S + E \tag{2}
\]

where \( A \in \mathbb{R}^{B \times K} \) is the endmember matrix whose columns are the spectra of each of the \( K \) endmembers, \( S \in \mathbb{R}^{K \times N} \) is the abundance matrix with the fractional composition of each of the \( N \) pixels is organized as a column vector, and \( E \in \mathbb{R}^{B \times N} \) is the noise matrix. Accordingly, the extraction of the endmember spectra (A) and their abundances (S) are the functions of HU.

**Neighbourhood approximation**

While the endmember spectra contain information about the spectral response of the underlying materials in the HSI, the spatial attributes of that image are stored in the abundance matrix. Therefore, in literature, to increase spatial smoothness and spatial correlation of the fractional composition, several regularizaton functions have been introduced in the objective function (He et al., 2017; Iordache et al., 2012; Rathnayake et al., 2020; Sigurdsson et al., 2016; Sun et al., 2018; Yang et al., 2015; Zhou et al., 2020) to optimize HU. However, given that the spectral signature of a given HSI is the product of the endmember spectra and their fractional abundances, the spatial correlation and smoothness could be increased using the spectral signatures of the neighboring pixels.
If the center pixel can be approximated from its immediate neighborhood, then the spatial correlation and smoothness shall be established in spectral signatures, thereby in abundance values. Consider the pixel at the \((m, n)\) location of the HSI \((\mathbf{H})\) and the set \((\mathcal{S})\) of immediate neighborhood pixels. The set can be defined as, \(\mathcal{S} = \{ \mathbf{H}_{ij} | \mathbf{H}_{ij} \in \mathbb{R}_+^B, i \in \{m-1, m+1\}, j \in \{n-1, n+1\}, i, j \in \mathbb{Z}^+ \}\) and the cardinality of the set is equal to the number of neighborhood pixels considered during the construction of the set and is denoted by \(|\mathcal{S}|\).

Then, following the pixel-to-column position transition explained in Section 2.1, by rearranging the vectors in \(\mathcal{S}\) as a single column vector denoted by \(\mathbf{n}_i\), the following relationship can be formulated between the center pixel \((y_j)\) and its neighborhood.

\[
y_j \approx \mathbf{W}_i \mathbf{n}_j \tag{3}
\]

where \(\mathbf{W}_i \in \mathbb{R}^{B \times |\mathcal{S}|}\) approximates the centre pixel using its neighbourhood pixel vector \(\mathbf{n}_i\). To learn the spatial correlation across the HSI, \(\mathbf{W}_i\) can be trained using the following objective function,

\[
\arg\min_{\mathbf{W}_i} \| \mathbf{Y} - \mathbf{W}_i \mathbf{N} \|^2_F \tag{4}
\]

Here, \(\mathbf{N}\) is the matrix constructed by concatenating the neighbor vectors of each center pixel, and \(\mathbf{Y}\) is the matrix form of the HSI and \(\| \cdot \|_F\) is the Frobenius norm. Then, by performing HU on pixels generated from the trained \(\mathbf{W}_i\) in (4), the spatial correlation and smoothness could be brought forward to fractional abundances by the AN displayed in Figure 2.

**Controlled abundance estimation**

In general, DL unmixing is performed using the AE architecture with the pixel reconstruction error as the objective function that is optimized. However, a drawback of using the AE for unmixing is that it precludes direct optimization of abundances in contrast to non-DL methods. Moreover, a conventional AE architecture may suffer from the vanishing gradient problem as the network becomes deeper and deeper. This is because gradients descending through back-propagation may exponentially decrease when they reach the layers in the encoder. Moreover, the unmixing process is dependent on the decoder performance of the AE, because of the end-to-end connection between the encoder and the decoder.

Consequently, to avoid the dependence of the abundance optimization on the endmember signature optimization, the proposed work considers an additional objective function to train the network. First, a reference is created for the abundance maps to train the UN by segmenting the HSI with a clustering algorithm. Here, k-means clustering is utilized for this purpose, and then the result is converted to a one-hot encoded representation. This reference for the abundances is referred to as the pseudo-ground truth and is described in Section 3.4. Further, the injection of a pseudo-ground truth facilitated by the split architecture provides a mechanism to reinforce the abundance optimization. This allows for a more refined pseudo-ground truth from unsupervised SP algorithms, which subsequently leads to a better accuracy with further training.

The fractional abundances of each pixel can be considered as a probability distribution, where the abundance values for the corresponding endmembers are the individual probabilities. This definition allows us to reconsider the unmixing process as a classification problem, using categorical cross-entropy as the objective function. The UN algorithm learns hidden representations of hyperspectral images (HSIs) by mapping spectral signatures onto the probability space using the pseudo-ground truth.

**Expanding decoder flexibility**

AEs for HU are omnipresent because they resemble the matrix decomposition used in blind source separation methods. Since the encoder task is designed to transform the pixel or image of the HS dataset into a latent representation, there is no limitation on the number of layers in the encoder. However, unlike the encoder, the decoder of the AE does not emulate a process; instead, it performs the matrix multiplication between the abundances and endmembers. Therefore, the endmembers of the HS data are extracted by the weights of the decoder. Hence, the decoder architecture is bound to a single-layer decoder to ensure that the output of the encoder will be the abundances.

By adopting the pseudo-ground truth mechanism, the UN is trained to produce abundance values at its output in our proposed architecture independent of the decoder architectural form. Therefore, since it is guaranteed that the input of the MN is the abundances of the center pixel, the single-layered decoder constraint can be released without using recurrent layers as in (Zhao et al., 2021). Since the output of the decoder is compared against the pixel signature, the decoder requirement is to find the optimal combination of weights that will map the abundances to spectral signatures. Furthermore, the decoder can model the non-linear mixing process with the proposed change.

Therefore, in the GAUSS architecture, the pixel’s neighborhood is fed to the AN, and its output is compared against the actual center pixel and the input to the UN. The UN will generate the abundances of the center pixel at its output which is subsequently used as the input to the MN. Since the decoder
architecture contains more than one layer, the endmembers can no longer be extracted by the weights of the decoder. Therefore, after the network is trained, an identity matrix \( \mathbf{I} \in \mathbb{R}^{K \times K} \) is fed to the MN. Each column of the identity matrix represents a pixel with a pure abundance and the number of rows is equal to the number of endmembers. Thus, when the matrix is fed, it is ensured that the decoder output corresponds to the pure endmembers. Thus, this process provides the required endmember signatures.

**Methodology**

**Approximation network**

This section elaborates on the proposed architecture for HU with spatial smoothness. The proposed method is primarily based on the traditional AE model used in standard HU practice under DL. However, the proposed method can be viewed as three separate networks (as shown in Figure 1), instead of the standard AE model in HU due to the method of training adapted in this work. The architecture of the approximation network is illustrated in Figure 2(a). The AN is used to impose spatial smoothness on the HSI as discussed in Section 2.3. The network approximates the centre pixels \( \mathbf{Y} \in \mathbb{R}^{B \times S} \) using their neighbourhood pixels \( \mathbf{N} \in \mathbb{R}^{B \times S} \) given by

\[
\hat{\mathbf{Y}} = f_n(\mathbf{N})
\]

with \( f_n : \mathbb{R}^{B \times S} \rightarrow \mathbb{R}^{B \times N} \). The network learns the parameters for approximation by minimizing the average reconstruction error between the input \( \mathbf{Y} \) and the reconstructed spectral signatures \( \hat{\mathbf{Y}} \) from the neighborhood pixels. Therefore, after the learning process is terminated, the smoothness matrix \( (\mathbf{W}_u) \) in (4) and the spectral signatures with improved spatial characteristics can be realized by

Smoothness transitions: \( f_u(\mathbf{N}) \Rightarrow \mathbf{W}_u \mathbf{N} \) \hspace{1cm} (6)

Smoothed pixels: \( \hat{\mathbf{Y}} \).

Next, to learn the hidden representation of the HSI, the UN is trained, which transforms the smoothed spectral signatures \( (\hat{\mathbf{Y}}) \) from the AN to a low-dimensional representation \( (\hat{\mathbf{L}} \in \mathbb{R}_+^{C \times N}) \) given by

\[
\hat{\mathbf{L}} = f_e(\hat{\mathbf{Y}}),
\]

with \( f_e : \mathbb{R}^{B \times N} \rightarrow \mathbb{R}^{K \times N} \). The output from the UN is fed to the MN, which decompresses the hidden representation to reconstruct the smoothed spectral signatures at its output. The decoder function, therefore, can be represented as

\[
\hat{\mathbf{Y}} = f_d(\hat{\mathbf{L}})
\]

where \( f_d : \mathbb{R}^{K \times N} \rightarrow \mathbb{R}^{B \times N} \), and \( \hat{\mathbf{Y}} \) is the output of the decoder network. After the successful termination of the learning process for HU, the unmixing results can be extracted by:

Abundance estimation: \( \hat{\mathbf{L}}^* \Rightarrow \hat{\mathbf{S}} \)

Endmember estimation: \( f_e(\mathbf{I}_K) \Rightarrow \hat{\mathbf{A}} \) \hspace{1cm} (10)

where \( \mathbf{I}_K \in \mathbb{R}_+^{K \times K} \) is the identity matrix and * denotes the optimal solution at the end of the training process.

Figure 1. Architecture of the proposed framework (GAUSS) that consists of the approximation network (AN), the unmixing network (UN) and the mixing network (MN).
Encoder

In this work, the encoder combines the AN and the UN to improve spatial characteristics and learn hidden representations. For each pixel of the HSI, we first consider its immediate neighborhood and reconstruct the center pixel from the set of neighborhood pixels. Then, the output of the AN is sent to the UN to learn the unmixing process and estimate abundances. The architecture of the encoder is given in Figure 2(b), and a fully connected layer (FCNL) architecture is used to construct the encoder.

In our work, the fractional abundances are considered probabilities, and learning hidden representations of the HSI is similar to performing a multi-class classification with the pseudo-ground truths. In our work, the encoder is designed to use spatial information and optimize the estimation of fractional abundances without interference from the decoder optimization. To optimize the AN, we use the reconstruction error of the center pixel from its neighborhood as given by

$$ J_{\text{smooth}} = \frac{1}{N} \| Y - \tilde{Y} \|_F^2 $$

where $\| \cdot \|_F$ is the Frobenius norm. Then, to abide by the ANC and ASC, we apply the following scaling and standardization to the final layer of the UN defined as follows,

$$ \hat{l}_{kj} = \frac{e^{-z_{kj}}}{\sum_{k=1}^N e^{-z_{kj}}} $$

where $z_{kj}$ is the output of the $k$th node at the final layer of the encoder prior activation and $\hat{l}_{kj}$ is the estimated fractional abundance for the $k$th endmember of the $j$th column in $L$. The standardization to probability distributions given in (12) can be realized through Softmax activation function of the output layer of the encoder. Then, to optimize the unmixing process of...
the encoder, the categorical cross-entropy is used as below,

$$\mathcal{J}_{\text{unmixing}} = -\frac{1}{N} \sum_{j,k=1}^{N,K} l_{kj} \log \hat{l}_{kj} + (1 - l_{kj}) \log(1 - \hat{l}_{kj})$$  \hspace{1cm} (13)

where \(l_{kj}\) are abundance values for \(k\) th endmember of the \(j\) th column of the pseudo-ground truth used to expedite optimization of the UN.

**Decoder**

An MLP-based network is adopted for the decoder in our proposed scheme, with its structure reported in Figure 2(c). In this work, the decoder is designed to reconstruct the input to the UN, which comprises smoothed pixels generated from the AN of the proposed architecture. As discussed in Section 3.2, with a pseudo-ground truth mechanism for abundance estimation, the decoder architecture is not restricted to a single-layered linear network; therefore, it can learn non-linear relationships as well.

The decoder of the proposed network learns relationships between abundances and endmember signatures during the mixing process because of the introduced extension in the layers. Nevertheless, by extending the decoder layers, the extraction of the endmember signatures cannot be done through the network weights because the unmixing process is now not limited to a single layer but rather laid out over a single network with multiple layers. In contrast, because the decoder is trained using abundances produced by the encoder network, it can reconstruct the spectral signature for any abundance vector in the low-dimensional space that the encoder has learnt. Thus, after the conclusion of weight training, the spectral information of the appropriate endmember may be obtained by feeding this one-hot encoded vector to the decoder. Hence, by feeding this one-hot encoded vector to the decoder at the end of weight training, the spectral information of the corresponding endmember can be extracted. For example, to find the endmember signature of the \(k\) th endmember, we will construct a one-hot vector \(\mathbf{o}_k\) such that \(\mathbf{o}_k = [0, \ldots, 0, 1, 0 \ldots, 0]\) where the unity entry is at the \(k\) th position of the vector. The vector format implies the existence of a pure pixel containing only that endmember. Hence, the resulting spectral signature of the MN is simply the corresponding endmember signature. Then, by concatenating these one-hot vectors of endmember, we can construct an identity matrix, when given to the decoder, which will output the endmember signatures of the HSI in a single attempt as presented in (10) for endmember estimation.

To train the decoder network, we diverge from the conventional practice of using the mean squared error between the encoder input \(\mathbf{Y}\) and the decoder output of each pixel \(\hat{\mathbf{Y}}\) defined as

$$\mathcal{J}_{\text{reconstruction}} = \frac{1}{NB} \| \mathbf{Y} - \hat{\mathbf{Y}} \|_F^2$$  \hspace{1cm} (14)

for the HSI. Rather, we adopt the pixel spectral information divergence (pSID) proposed by (Chang, 1999) as the reconstruction error to optimize the mixing process, which is defined by,

$$\mathcal{J}_{\text{mixing}} = D(\hat{\mathbf{Y}} \| \mathbf{Y}) + D(\mathbf{Y} \| \hat{\mathbf{Y}})$$

$$= \frac{1}{N} \sum_{j,k=1}^{N,K} (q_{kj} \log(q_{kj}) + \hat{q}_{kj} \log(\hat{q}_{kj}))$$  \hspace{1cm} (15)

where \(q_{kj} = \frac{\hat{y}_{kj}}{||\hat{y}_{kj}||_1}\) and \(\hat{q}_{kj} = \frac{\hat{y}_{kj}}{||\hat{y}_{kj}||_1}\) are the probability distribution vectors of the spectral signature of the \(j\) th pixel of \(\mathbf{Y}\) and \(\hat{\mathbf{Y}}\), respectively. Accordingly, \(q_{kj}\) and \(\hat{q}_{kj}\) represent the probability values of the \(b\) th band of the \(j\) th pixel. Then, \(\| \cdot \|_1\) is 1-norm of the corresponding spectral signature.

In (15), \(D(\cdot \| \cdot)\) is called the relative entropy between the smoothed pixels at the UN input and reconstructed pixels at the decoder output. The pSID defined by (15) can be used to measure the spectral similarity between corresponding pixel vectors in \(\hat{\mathbf{Y}}\) and \(\hat{\mathbf{Y}}\).

**Pseudo-ground truth preparation for abundances**

In the GAUSS architecture, a one-hot encoded representation is used as the pseudo-ground truth to guide the UN of the encoder. To construct the one-hot encoded representation of the HSI, we first segment the HSI using the k-means clustering algorithm on the spectral signatures.

At the second stage of the training, we consider the abundance result of an SP algorithm with reliable unmixing performance as a pseudo-ground truth for HUA. This is analogous to using unsupervised HU algorithms for endmember weight initialization. The output of the UN will then be compared against the new pseudo-ground truth. The pseudo-ground truth produced by the high-performing SP algorithm enables the refinement of the abundance optimization. Accordingly, we use the abundance results of the L1/2 NMF algorithm to maintain consistency across all the datasets.

**Implementation and training of the network**

The HUA was developed in TensorFlow v2 using the Google Colaboratory with default settings for the optimization parameters. The detailed architecture of the HUA is given in Table 2. The Adam optimizer was
used to optimize the three networks. For each HSI dataset, a batch size of 32 was used for the input-output dataset preparation.

In the first training stage, each of the three networks was trained for 250 epochs. Then, in the second training stage, both the unmixing network and the mixing network were trained for 250 epochs each (The approximation network was not trained in the second stage).

Experiments and discussion

Performance criteria

The performance of the abundance estimation is measured by the average root mean square error (aRMSE), average abundance information divergence (aAID), and average abundance angle distance (aAAD). Then, the accuracy of endmember estimation from the HU algorithm is evaluated using the average spectral angle distance (aSAD) and the average spectral information divergence (aSID), which are defined as follows,

\[
aRMSE = \sqrt{\frac{1}{NK} \| S - \hat{S} \|_{F}^{2}}
\]

\[
aAID = \frac{1}{N} \sum_{i=j}^{N} \cos^{-1} \left( \frac{\langle s_{j}, s_{j} \rangle}{\| s_{j} \| \cdot \| s_{j} \|} \right)
\]

\[
aAID = \frac{1}{N} \sum_{j=1}^{NK} \left\{ \lambda_{jk} \log \left( \frac{\lambda_{jk}}{\hat{\lambda}_{jk}} \right) + \hat{\lambda}_{jk} \log \left( \frac{\lambda_{jk}}{\hat{\lambda}_{jk}} \right) \right\}
\]

\[
aAID = \frac{1}{N} \sum_{k=1}^{K} \cos^{-1} \left( \frac{\langle a_{k}, \hat{a}_{k} \rangle}{\| a_{k} \| \cdot \| \hat{a}_{k} \|} \right)
\]

\[
aAID = \frac{1}{K} \sum_{k=1}^{K} \sum_{b=1}^{B} \left\{ p_{bk} \log \left( \frac{p_{bk}}{\hat{p}_{bk}} \right) + \hat{p}_{bk} \log \left( \frac{\hat{p}_{bk}}{p_{bk}} \right) \right\}
\]

where \( S \) and \( \hat{S} \) represents the ground truth and estimated abundances. Then, \( s_{j} \) and \( \hat{s}_{j} \) are the ground truth and estimated abundances at \( j \) th column or pixel, and \( s_{kj} \) and \( \hat{s}_{kj} \) are the ground truth and estimated abundance value for \( k \) th endmember of the \( j \) th pixel. Next, \( a_{k} \) and \( \hat{a}_{k} \) are ground truth and extracted endmember for the \( k \) th source, and \( p_{bk} \) and \( \hat{p}_{bk} \) are probability values for the \( b \) th spectral band of the \( k \) th ground truth and the extracted endmembers, respectively. The probability vector for the \( k \) th endmember is computed as \( p_{bk} = \frac{a_{kb}}{\| a_{k} \|} \) and \( \hat{p}_{bk} = \frac{\hat{a}_{kb}}{\| \hat{a}_{k} \|} \) for the ground truth and extracted signatures, respectively.

Experimental setting

To evaluate the performance of GAUSS in a complex environment, an algorithm that is capable of generating simulated datasets with a sufficient amount of complexity is designed. Compared with the simulated datasets in the literature, the dimensions were selected to be 100 × 100 with a spectral resolution of 198 bands.

For the generation of abundances, initially, a zeros tensor of size 100 × 100 × 4 is segmented into sixteen 25 × 25 × 4 tensors. Next, the algorithm randomly selects one from the segmented tensors and initiates the assignment of abundance fractions. This is done through a superpixel of size 3 × 3. Thereafter, a random selection of a center pixel is done and checked for occupancy. If the pixel is not already occupied, then the possibility of it being on the outer edge is investigated. If the pixel is not a part of the outer edge, the superpixel of the above-mentioned dimension is constructed and the abundance fractions are assigned while satisfying the sum to one constraint. If the selected center pixel is at an edge, then necessary alterations are made to the superpixel to be adapted according to the nature of the edge. However, following this procedure will not make the dataset realistic, as it contains pure regions and sparse sections as remotely sensed HSIs. To achieve this, some of the 25 × 25 pixels are allotted to be pure or to maintain a constant abundance fraction variation. This process creates sharp edges. To resolve this issue, initially, the sharp edges are extended irregularly, and then the entire image is smoothened through mean filtering. The generated image is shown in Figure 3(a).
Next, to generate the simulated image, endmember assignment is carried out. For this, initially, four endmembers have been chosen from the United States Geological Survey (USGS)\(^1\) endmember library (Kokaly et al., 2017). The spectral signatures of “Ilmenite, Montmorillonite, Limestone, and Tree” are chosen as the candidate endmembers. Then, as the spectral signatures extracted contained bands that had not been calibrated, a preprocessing step is carried out to remove the uncalibrated bands. Furthermore, as the spectral signatures of different elements are acquired from different sensors, there can be a slight deviation in the wavelengths present in each spectral signature. To overcome this issue, the cubic interpolation method is used to generate spectral signatures containing the same wavelengths. As the next step, every endmember spectrum is multiplied by the corresponding abundance fraction, and then, those are summed to generate a simulated data tensor of size \(1 \times 1 \times 198\). Eventually, the previously described process has been carried out for the entire pixel space of the simulated dataset.

Moreover, two real HS datasets\(^2\) Samson and Jasper-Ridge, were used for quantitative evaluation purposes. A precise ground truth for the endmembers and abundances for each dataset is available on each website. The Samson dataset consists of three endmembers with 156 spectral bands covering 401 nm to 889 nm wavelengths, while Jasper-Ridge comprises four sources with 198 spectral bands ranging from 380 nm to 2500 nm.

To evaluate the proposed architecture, nine different hyperspectral unmixing algorithms were used. These algorithms included deep learning architectures such as CNNAEU (Palsson et al., 2020), DAEN (Su et al., 2019), DAEU (Palsson et al., 2018), uDAS (Qu & Qi, 2018), and EGU-Net (Hong et al., 2021b), as well as signal processing algorithms that leverage different NMF variants and other geometric frameworks, such as \(L^{1/2}\) (Qian et al., 2010), R-CoNMF (Li et al., 2016), MLNMF (Rajabi & Ghassemian, 2014), and VCA-FCLS (Nascimento & Dias, 2005). These algorithms were tested on the real hyperspectral datasets as well as the simulated dataset incorporating the optimum parameters suggested by the authors in their work. The RGB images of the datasets are illustrated in Figure 3.

The best-performing algorithms are ranked up to the third, and the ranks are in superscript.

**Experiments on simulated data**

Most of the conventional unmixing algorithms used in this study are geometry-based methods. These methods were implemented based on the assumption that these algorithms would unmix ideal HSIs. The method proposed for generating a simulated dataset in this study also aligns with the assumption used in the above-mentioned algorithms. The presence of pure or marginally pure pixels, favoring unmixing.

---

\(^1\)https://www.usgs.gov/labs/spectroscopy-lab/science/spectral-library
\(^2\)http://lesun.weebly.com/hyperspectral-data-set.html:
algorithms that exploit the underlying geometric frameworks, is reflected in these results.

However, the results of the proposed architecture are exceptional in abundance estimation. Notably, the estimated abundances of the architecture have proven to be superior to those obtained from the pseudo-ground truth abundances generated by the $L^{1/2}$ NMF algorithm. This outcome serves as evidence that the supervisory inputs, which are utilized in the architecture, have effectively facilitated accurate abundance estimation, even in complex environments. Regarding endmember extraction, SP-based $L^{1/2}$ NMF and RCo-NMF deliver the best results in terms of aSID and aSAD error matrices, respectively.

**Experiments on real data**

**Samson dataset**

It is imperative to understand the Samson dataset to analyze the performance of the proposed method. It has three principal endmembers, and the boundaries of these endmembers are almost non-overlapping.

When considering Table 3, our architecture for HU performs the best in terms of abundance estimation. Moreover, using the $L^{1/2}$ NMF abundances as the pseudo-ground truth, the architecture has produced better results than the $L^{1/2}$ NMF algorithm. A possible reason for this could be that the proposed method had benefited from the first training phase of the network, which may have trained GAUSS to learn various properties of the latent space of the HSI.

While the proposed algorithm excels in terms of spectral angle distance (SAD), the uDAS method outperforms it in spectral information divergence (SID). Additionally, the endmember extraction performance of GAUSS yields better results compared to the $L^{1/2}$ NMF algorithm, which generated the pseudo-ground truths for GAUSS. Therefore, as the proposed architecture surpasses the $L^{1/2}$ NMF algorithm in both tasks of abundance estimation and endmember

**Figure 4.** Abundance maps obtained for the simulated, Jasper-Ridge, and Samson datasets using GAUSS, and their ground truths.

---

Table 3. Unmixing performance comparison for the three datasets: simulated, Samson and Jasper-Ridge

|                      | Proposed method | Deep learning methods | Signal processing methods |
|----------------------|-----------------|-----------------------|---------------------------|
|                      | GAUSS           | CNNAEU                | DAEN                      | DAEU                      | EGUnet                   | uDAS                      | $L^{1/2}$ NMF | R-CoNMF | MLNMF | VCA FCLS |
| Simulated            |                 |                       |                           |                           |                          |                          |              |         |       |
| aRMSE                | 0.1312          | 0.2453                | 0.2550                    | 0.2371                    | 0.2783                   | 0.2349                    | 0.2419        | 0.1618  | 0.2896 | 0.2874  |
| aAAD                 | 0.3414          | 0.8532                | 0.7386                    | 0.8297                    | 0.6844                   | 0.8430                    | 0.6909        | 0.4256  | 0.8240 | 0.8168  |
| aAID                 | 0.2496          | 3.4171                | 1.5346                   | 2.8457                    | 2.3981                   | 7.1713                    | 2.5311        | 0.9626  | 3.3080 | 3.9947  |
| aSAD                 | 0.2459          | 0.1316                | 0.0301                    | 0.0756                    | 0.1059                   | 0.3111                    | 0.0008        | 0.0269  | 0.0984 | 0.0009  |
| aSID                 | 0.2739          | 0.3764                | 0.5461                    | 0.3246                    | 0.2332                   | 0.4308                    | 0.3436        | 0.0003  | 0.2467 | 0.3433  |
| Samson               |                 |                       |                           |                           |                          |                          |               |         |       |        |
| aRMSE                | 0.0067          | 0.3724                | 0.3909                    | 0.3902                    | 0.1573                   | 0.3027                    | 0.2689        | 0.3330  | 0.3424 | 0.3996  |
| aAAD                 | 0.0724          | 0.9073                | 0.7774                    | 0.8384                    | 0.2518                   | 0.5793                    | 0.4991        | 1.0979  | 0.6533 | 0.8825  |
| aAID                 | 4.4736          | 7.6865                | 6.7775                    | 5.5561                    | 1.0453                   | 4.7794                    | 4.0999        | 10.7096 | 6.4476 | 8.0004  |
| aSAD                 | 0.0211          | 0.1769                | 0.1413                    | 0.1155                    | 0.0603                   | 0.0697                   | 0.0811        | 0.2451  | 0.0897 | 0.1364  |
| aSID                 | 0.0584          | 1.5040                | 1.5217                    | 2.4887                    | 0.0486                   | 0.0328                   | 0.1953        | 0.4725  | 0.4758 | 0.6746  |
| Jasper-Ridge         |                 |                       |                           |                           |                          |                          |               |         |       |        |
| aRMSE                | 0.0083          | 0.2888                | 0.2227                    | 0.2623                    | 0.4651                   | 0.1190                   | 0.3393        | 0.3570  | 0.4135 | 0.4903  |
| aAAD                 | 0.1489          | 0.7186                | 0.5001                    | 0.6239                    | 1.2312                   | 0.2177                   | 0.8437        | 1.0048  | 1.0811 | 1.2592  |
| aAID                 | 12.4845         | 6.8717                | 4.5497                    | 7.4464                    | 15.9190                  | 1.5840                   | 9.6208        | 8.6153  | 11.3338| 19.8816 |
| aSAD                 | 0.0965          | 0.1552                | 0.1328                    | 0.1634                    | 0.1447                   | 0.1062                   | 0.1677        | 0.1422  | 0.2241 | 0.2092  |
| aSID                 | 0.1487          | 0.2635                | 0.2092                    | 0.2710                    | 1.3000                   | 0.0441                   | 0.8390        | 0.2914  | 1.1527 | 1.3342  |
extraction performances, the performance improvement can be attributed to the integration of the best attributes of the weights and biases from the initial training phase and the supervisory input from the L^{1/2} NMF algorithm provided in the secondary training phase. This validates the combination of these two domains, achieved through the proposed architecture.

The evaluation of the endmember extraction results demonstrates that the proposed algorithm outperforms the latest CyCU-Net (Gao et al., 2021) and MSNet (Tao et al., 2022) architectures when applied to the Samson dataset. Specifically, when comparing the achieved average root mean square error (aRMSE), the CyCU-Net and MSNet architectures yield values of 0.4164 and 0.0526, respectively. These findings highlight the superior performance of the proposed architecture for endmember extraction tasks in hyperspectral imaging.

Jasper-ridge dataset
The Jasper Ridge dataset has four endmembers, and their boundaries are not as non-overlapping as in the Samson dataset. GAUSS has achieved the best results in abundance estimation, with the exception of the aAID error metric. This can be attributed to the fact that the supervisory input provided during the secondary training phase offers a more effective latent structure for the algorithm.

The extraction of the road endmember in this dataset presents challenges due to its scarcity compared to other endmembers and its strong correlation with the dirt endmember, despite having high reflectance values. Consequently, accurately identifying and separating the road endmember from the spectral data becomes a complex task. Based on Figure 5, the proposed architecture in this study has extracted an endmember signature for the road that closely matches its ground truth. Similarly, the uDAS algorithm has also demonstrated successful prediction of the road endmember. The achievement of GAUSS can be attributed to the AN and the pseudo-ground truth. The AN appears to have effectively decorrelated the pure endmembers. As a result, the decoder has exhibited enhanced accuracy in extracting the road endmember compared to the other algorithms.

Upon evaluating the results of the endmember extraction error metrics (refer Table 3), it is evident that GAUSS demonstrates superior performance when utilizing the Jasper-Ridge dataset, outperforming both the latest CyCU-Net and MSNet architectures. Specifically, the CyCU-Net and MSNet architectures achieved considerably higher aRMSE values of 0.2876 and 0.0868, respectively. In contrast, GAUSS was able to achieve a significantly lower aRMSE of 0.0483. These findings unequivocally highlight the superiority of the proposed architecture over the CyCU-Net and MSNet architectures for endmember extraction tasks in hyperspectral imaging.

According to Table 3, the method with the least reconstruction error for each dataset has the best-performing figures for unmixing results. Therefore, though the authors have chosen the L^{1/2} NMF algorithm as the conventional algorithm to construct a pseudo-ground truth for the training of the UN, the selection of a suitable unsupervised algorithm can also be based on the reconstruction error of the HSI for that algorithm. However, in this work, the L^{1/2} NMF algorithm was used to maintain consistency in the comparison to remove the effect of using different algorithms in constructing the pseudo-ground truth.

The average error metric variances after five consecutive trials are shown in Table 4. Moreover, the abundance maps obtained for the simulated, Jasper-Ridge, and Samson datasets using GAUSS, and their ground truths are shown in Figure 4. Likewise, the abundance maps obtained for the Samson datasets using ground truths, GAUSS, and other state-of-the-art algorithms are depicted in Figure 6. Additionally, Figure 7 illustrates the extracted endmember signatures obtained for the Samson datasets using ground truths, GAUSS, and other state-of-the-art algorithms. Also note that all the experiments have been performed in a computer with a 2.30 GHz Intel Dual Core CPU and 12 GB of memory.

### Table 4. Average error metric variances after five consecutive trials.

| Metric | Simulated | Samson | Jasper Ridge |
|--------|-----------|--------|--------------|
| aRMSE  | 4.3 × 10^{-5} | 6.4 × 10^{-5} | 7.9 × 10^{-5} |
| aAID   | 2.1 × 10^{-4} | 2.0 × 10^{-3} | 2.0 × 10^{-4} |
| aAID   | 9.6 × 10^{-5} | 6.2 × 10^{-5} | 4.2 × 10^{-5} |
| aSAD   | 6.4 × 10^{-5} | 8.1 × 10^{-5} | 1.2 × 10^{-4} |
| aSID   | 7.3 × 10^{-4} | 5.0 × 10^{-5} | 2.1 × 10^{-4} |

### Ablation study and performance analysis

An ablation study has been conducted to further illustrate the effectiveness of GAUSS architectures on HU. The Samson data set is used for this study, and the findings are tabulated in Table 5. Furthermore, it is evident that all quantitative error metrics show the effectiveness of the proposed method. Here, AE refers to the Autoencoder, AN
refers to the Approximation Network, and pGT refers to the pseudo-ground truth. The presented network configurations are as follows – AE: Pure autoencoder-based unmixing method, AE+AN: Autoencoder coupled with approximation network for the neighborhood association, AE+pGT: Autoencoder with pseudo-ground truth initialization at the bottleneck, AE+AN+pGT: Autoencoder with neighborhood association through approximation network and pseudo-ground truth at the bottleneck. Due to the two-stage training with the injection of the two pseudo ground truths, GAUSS has provided the best results when considering all the other combinations. This is evident from the results tabulated in Table 5.

Table 6 provides an overview of the training time required for the proposed networks, along with the corresponding root mean square error (RMSE) values for the reconstructed center pixel and the original center pixel. According to the table, the Samson dataset exhibited the fastest convergence towards an optimal solution. In contrast, both the Jasper Ridge and
Synthetic datasets require more computational time to achieve comparable RMSE values compared to the Samson dataset.

Additionally, this study investigates the impact of increasing the number of immediate neighborhood pixels on the proposed architecture for the Samson dataset, and the results are summarized in Table 7. These findings clearly indicate that as the number of neighboring pixels increased, the values of aRMSE, aAAD, and aAID decreased. However, aSAD and aSID displayed a different pattern compared to the other examined error metrics.

Furthermore, when considering a linear mixture model that utilizes the ground truth abundances and endmembers to reconstruct the image for the Samson dataset, the RMSE between the original image and the reconstructed image is non-zero (RMSE: 0.07678).

Table 5. Ablation study for GAUSS on the Samson Dataset.

|               | aRMSE   | aAAD    | aAID    | aSAD    | aSID    |
|---------------|---------|---------|---------|---------|---------|
| AE            | 0.4150  | 0.7826  | 6.4259  | 0.1416  | 1.5271  |
| AE + AN       | 0.4386  | 0.7470  | 6.1648  | 0.1113  | 1.2928  |
| AE + pGT      | 0.4186  | 0.6975  | 5.4374  | 0.0857  | 0.8212  |
| AE + AN + pGT | 0.1989  | 0.6183  | 4.1830  | 0.0218  | 0.0657  |

Table 6. Performance analysis for the computation time and RMSE loss for each dataset.

|                | Simulated | Samson   | Jasper   |
|----------------|-----------|----------|----------|
| Time(s)        | 145.53    | 86.89    | 101.84   |
| RMSE           | 0.0367    | 0.0221   | 0.0836   |

Table 7. Ablation study for the number of neighbouring pixels taken into account.

|                  | 24 Pixels | 48 Pixels | 80 Pixels |
|------------------|-----------|-----------|-----------|
| aRMSE            | 0.5209    | 0.4719    | 0.4532    |
| aAAD             | 0.9551    | 0.9253    | 0.8319    |
| aAID             | 10.2043   | 8.9533    | 7.3562    |
| aSAD             | 0.5169    | 0.5138    | 0.5228    |
| aSID             | 0.8378    | 0.8384    | 0.1023    |

Table 8. Ablation study for the performance of complex non-linear unmixing scenarios.

|                  | 1 layer   | 2 layers (GAUSS) | 3 layers | 5 layers |
|------------------|-----------|------------------|----------|----------|
| aRMSE            | 0.4983    | 0.0467           | 0.4874   | 0.4855   |
| aAAD             | 0.8101    | 0.0724           | 0.8075   | 0.8075   |
| aAID             | 6.9717    | 4.4736           | 7.7449   | 7.0333   |
| aSAD             | 0.2287    | 0.0211           | 0.1748   | 0.1671   |
| aSID             | 0.2002    | 0.0584           | 0.7609   | 0.7543   |

Table 9. Ablation study for the performance of GAUSS under different levels of noise.

|                  | Original Image | 20 dB SNR | 10 dB SNR | 5 dB SNR |
|------------------|----------------|-----------|-----------|----------|
| aRMSE            | 0.0467         | 0.4178    | 0.4190    | 0.4683   |
| aAAD             | 0.0724         | 0.6948    | 0.6987    | 0.8428   |
| aAID             | 4.4736         | 5.3201    | 5.2977    | 7.3963   |
| aSAD             | 0.0211         | 0.2160    | 0.2757    | 0.2277   |
| aSID             | 0.0584         | 0.1205    | 0.2025    | 0.7242   |

Figure 7. Extracted endmember signatures obtained for the Samson datasets using ground truths, GAUSS, and other state-of-the-art algorithms.
This indicates that the Samson dataset contains non-linearities. In Table 8, the variation of the error metrics of the model with the increasing number of layers of the decoder is presented. Based on the RMSE, AAAD, and aAID, values, the best results are found when utilizing two layers for the decoder architecture. Thus, it can be seen that GAUSS, with its two-layer decoder, performs best when considering the unmixing of a non-linear dataset.

Moreover, an ablation study was performed to assess GAUSS’ performance when handling noisy HSIs. Here, different levels of noise were added to the Samson dataset as suggested in (Hong et al., 2018), and this noisy dataset was unmixed using GAUSS. The results of this study are tabulated in Table 9. When considering the results, it can be seen that with the addition of noise, the performance of GAUSS has degraded.

Conclusion
The conventional optimization of an autoencoder for HU often seems to lack optimal results for abundances and endmembers. This paper proposes a novel autoencoder-based HU algorithm to surpass the state-of-the-art methods for HU in terms of abundance and endmembers estimation accuracy. Initially, for this, a blind-spot network has been introduced to improve the spatial context of the pixels. Thereafter, a vanilla autoencoder with two training phases has been introduced. The enriched supervisory endmember spectra and abundances are provided as a guiding mechanism for the architecture through a reliable SP algorithm. The performance of the proposed architecture has been evaluated for two standard datasets and a simulated dataset. Notably, the results obtained by the proposed methodology surpass those achieved by existing unmixing algorithms in both the DL and SP domains, establishing its superiority. For future work, the authors hope to make improvements to the model such that the model will address the non-linearities in HS datasets more effectively. Also, the robustness of the model to noise is another aspect that the authors wish to work on. Moreover, another future direction of pursuit is that of extending the depth of the network by using the pseudo-ground truth mechanism to create deep and rich features that could be used to unmix HSIs more effectively.

Disclosure statement
No potential conflict of interest was reported by the author(s).

Data availability statement
The authors confirm that the data supporting the findings of this study are available within the article [and/or] its supplementary materials.

References
Benabadj, S. I., Karoui, M. S., Djerriki, K., Boukerch, I., Farhi, N., & Boughlama, M. A. (2019). Unsupervised hyperspectral band selection by combination of unmixing and sequential clustering techniques. European Journal of Remote Sensing, 52(1), 30–39. https://doi.org/10.1080/22797254.2018.1549511
Bioucas-Dias, J. M. (2009). A variable splitting augmented Lagrangian approach to linear spectral unmixing. Proceedings of the 2009 First workshop on hyperspectral image and signal processing: Evolution in remote sensing, 26-28 August 2009, Grenoble, France, (pp. 1–4). IEEE.
Bioucas-Dias, J. M., & Figueiredo, M. A. (2010). Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing. Proceedings of the 2010 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, 4-16 June 2010, Reykjavik, Iceland, (pp. 1–4). IEEE.
Bioucas-Dias, J. M., Plaza, A., Dobigeon, N., Parente, M., Du, Q., Gader, P., & Chanussot, J. (2012). Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 5(2), 354–379. https://doi.org/10.1109/JSTARS.2012.2194696
Boardman, J. (1993). Automating spectral unmixing of aviris data using convex geometry concepts. summaries of the 4th annual jpl airborne geosciences workshop. Jet Propulsion Laboratory Pasadena, California.
Chang, C.-I. (1999). Spectral information divergence for hyperspectral image analysis. In IEEE 1999 International Geoscience and Remote Sensing Symposium . IGARSS’99, Hamburg, Germany. (Cat. No. 99CH36293), (Vol. 1, pp. 509–511). IEEE.
Chen, X., Chen, J., & Pan, J. (2019). Sulfate mineral mapping with hyperspectral imagery, a case study of the rodalquilar area, se spain. Proceedings of the 2019 10th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS), 24-26 September 2019, Amsterdam, Netherlands, (pp. 1–5). IEEE.
Das, S., & Routray, A. (2019). Covariance similarity approach for semiblind unmixing of hyperspectral image. IEEE Geoscience and Remote Sensing Letters, 16 (6), 937–941. https://doi.org/10.1109/LGRLS.2018.2888580
Das, S., Routray, A., & Deb, A. (2019). Sparsity measure based library aided unmixing of hyperspectral image. IET Image Processing, 13(12), 2077–2083. https://doi.org/10.1049/iet-imp.2018.5426
Das, S., Routray, A., & Deb, A. K. (2017). Convex set based abundance constrained unmixing of hyperspectral image. Proceedings of the 2017 14th IEEE India Council International Conference (INDICON), 15-17 December 2017, Roorkee, India, (pp. 1–5).
Dobigeon, N., Moussaoui, S., Coulon, M., Tourneret, J.-Y., & Hero, A. O. (2009).Joint bayesian endmember extraction and linear unmixing for hyperspectral imagery. IEEE Transactions on Signal Processing, 57(11), 4355–4368. https://doi.org/10.1109/TSP.2009.2025797
Dobigeon, N., Tourneret, J.-Y., Richard, C., Bermudez, J. C. M., McLaughlin, S., & Hero, A. O. (2013). Nonlinear unmixing of hyperspectral images: Models and algorithms. IEEE Signal Processing Magazine, 31(1), 82–94. https://doi.org/10.1109/MSP.2013.2279274
Ekanayake, E., Weerasooriya, H., Ranasinghe, D., Herath, S., Rathnayake, B., Godaliyadda, G., Ekanayake, M., & Herath, H. (2021). Constrained nonnegative matrix factorization for blind hyperspectral unmixing incorporating endmember independence. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, 11853–11869. https://doi.org/10.1109/JSTARS.2021.3126664

Fang, Y., Wang, Y., Xu, L., Zhuo, R., Wong, A., & Clausi, D. A. (2022). Bcun: bayesian fully convolutional network for hyperspectral spectral unmixing. Proceedings of the IEEE Transactions on Geoscience and Remote Sensing, (pp. 1–1).

Gader, P., Dranishnikov, D., Zare, A., & Chanussot, J. (2012). A sparsity promoting bilinear unmixing model. Proceedings of the 2012 4th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Shanghai, China, (pp. 1–4). IEEE.

Gao, L., Han, Z., Hong, D., & Chanussot, J. (2021). Cycu-net: Cycle-consistency unmixing network by learning cascaded autoencoders. Proceedings of the IEEE Transactions on Geoscience and Remote Sensing, (pp. 1–14).

Guan, Q., Xu, T., Feng, S., Yu, F., & Song, K. (2022). Optimal segmentation and improved abundance estimation for superpixel-based hyperspectral unmixing. European Journal of Remote Sensing, 55(1), 485–506. https://doi.org/10.1080/22797254.2022.2125447

Gu, J., Yang, B., & Wang, B. (2021). Nonlinear unmixing for hyperspectral images via kernel-transformed bilinear mixing models. Proceedings of the IEEE Transactions on Geoscience and Remote Sensing.

Halimi, A., Altmann, Y., Dobigeon, N., & Tourneret, J.-Y. (2011). Nonlinear unmixing of hyperspectral images using a generalized bilinear model. IEEE Transactions on Geoscience and Remote Sensing, 49(11), 4153–4162. https://doi.org/10.1109/TGRS.2010.2098414

Han, Z., Hong, D., Gao, L., Zhang, B., & Chanussot, J. (2020). Deep half-siamese networks for hyperspectral unmixing. IEEE Geoscience and Remote Sensing Letters, 18(11), 1996–2000. https://doi.org/10.1109/LGRS.2020.3011941

Han, Z., Hong, D., Gao, L., Zhang, B., Huang, M., & Chanussot, J. (2022). Autonas: Automatic neural architecture search for hyperspectral unmixing. IEEE Transactions on Geoscience and Remote Sensing, 60, 1–14. https://doi.org/10.1109/TGRS.2022.3186480

Heylen, R., Parente, M., & Gader, P. (2014). A review of nonlinear hyperspectral unmixing methods. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7(6), 1844–1868. https://doi.org/10.1109/JSTARS.2014.2320576

He, W., Zhang, H., & Zhang, L. (2017). Total variation regularized reweighted sparse nonnegative matrix factorization for hyperspectral unmixing. IEEE Transactions on Geoscience and Remote Sensing, 55(7), 3909–3921. https://doi.org/10.1109/TGRS.2017.2683719

Hong, D., Gao, L., Yao, J., Yokoya, N., Chanussot, J., Heiden, U., & Zhang, B. (2021a). Endmember-guided unmixing network (egu-net): A general deep learning framework for self-supervised hyperspectral unmixing. Proceedings of the IEEE Transactions on Neural Networks and Learning Systems.

Hong, D., Gao, L., Yao, J., Yokoya, N., Chanussot, J., Heiden, U., & Zhang, B. (2021b). Endmember-guided unmixing network (egu-net): A general deep learning framework for self-supervised hyperspectral unmixing. Proceedings of the IEEE Transactions on Neural Networks and Learning Systems.

Hong, D., Gao, L., Yao, J., Yokoya, N., Chanussot, J., Du, Q., & Zhang, B. (2021). More diverse means better: Multimodal deep learning meets remote-sensing imagery classification. IEEE Transactions on Geoscience and Remote Sensing, 59(7), 5966–5978. https://doi.org/10.1109/TGRS.2020.3015157

Hong, D., Gao, L., Yokoya, N., Yao, J., Chanussot, J., Du, Q., & Zhang, B. (2021). More diverse means better: Multimodal deep learning meets remote-sensing imagery classification. IEEE Transactions on Geoscience and Remote Sensing, 59(7), 4340–4354. https://doi.org/10.1109/TGRS.2020.3016682

Hong, D., Yokoya, N., Chanussot, J., & Zhu, X. X. (2018). An augmented linear mixing model to address spectral variability for hyperspectral unmixing. CoRR, abs/1810.12000.

Iordache, M.-D., Bioucas-Dias, J., & Plaza, A. (2012). Total variation spatial regularization for sparse hyperspectral unmixing. Proceedings of the IEEE Transactions on Geoscience and Remote Sensing.

Iordache, M.-D., Bioucas-Dias, J. M., & Plaza, A. (2012). Total variation spatial regularization for sparse hyperspectral unmixing. IEEE Transactions on Geoscience and Remote Sensing, 50(11), 4484–4502. https://doi.org/10.1109/TGRS.2012.2191590

Khan, M. J., Khan, H. S., Yousaf, A., Khurshid, K., & Abbas, A. (2018). Modern trends in hyperspectral image analysis: A review. Institute of Electrical and Electronics Engineers Access, 6, 14118–14129. https://doi.org/10.1109/ACCESS.2018.2812999

Kokaly, R. F., Clark, R. N., Swayze, G. A., Livo, K. E., Hoefen, T. M., Pearson, N. C., Wise, R. A., Benzel, W. M., Lowers, H. A., Driscoll, R. L., & Klein, A. J. (2017). USGS spectral library version 7. Report 1035, U.S. Geological Survey.

Kruse, F. A., Boardman, J. W., & Huntington, J. F. (2003). Comparison of airborne hyperspectral data and eo-1 hyperion for mineral mapping. IEEE Transactions on Geoscience and Remote Sensing, 41(6), 1388–1400. https://doi.org/10.1109/TGRS.2003.812908

Lee, D. D., & Seung, H. S. (2000). Algorithms for non-negative matrix factorization. In Todd, K. L., Thomas, G. D., & Volker, T. (Eds.), Advances in neural information processing systems (pp. 535–541). Bradford Books.

Li, J., Bioucas-Dias, J. M., Plaza, A., & Liu, L. (2016). Robust collaborative nonnegative matrix factorization for hyperspectral unmixing. IEEE Transactions on Geoscience and Remote Sensing, 54(10), 6076–6090. https://doi.org/10.1109/TGRS.2016.2580702

Liu, L., Feng, J., Han, L., Zhou, J., Xu, X., Liu, R., & Li, R. (2018). Mineral mapping using spaceborne tiangong-1 hyperspectral imagery and aster data: A case study of alteration detection in support of regional geological survey at jintan-miaolianquan area, beishan, Gansu province, china. Geological Journal, 53(S2), 372–383. https://doi.org/10.1002/gj.3260

Li, C., Zhang, B., Hong, D., Yao, J., & Chanussot, J. (2023). Lrr-net: An interpretable deep unfolding network for hyperspectral anomaly detection. IEEE Transactions on Geoscience and Remote Sensing, 61, 1–12. https://doi.org/10.1109/TGRS.2023.3279834

Lorenz, S., Kirsch, M., Zimmermann, R., Tusa, L., Möckel, R., Chamberland, M., & Gloaguen, R. (2018). Long-wave hyperspectral imaging for lithological mapping: A case study. Proceedings of the IGARSS 2018-2018
Lu, B., Dao, P. D., Liu, J., He, Y., & Shang, J. (2020). Recent advances of hyperspectral imaging technology and applications in agriculture. Remote Sensing, 12(16), 2659. https://doi.org/10.3390/rs12162659

Ma, Y., Li, C., Mei, X., Liu, C., & Ma, J. (2017). Robust sparse hyperspectral unmixing with $\ell_2,1$ norm. IEEE Transactions on Geoscience and Remote Sensing, 55(3), 1227–1239. https://doi.org/10.1109/TGRS.2016.2616161

Miao, L., & Qi, H. (2007). Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization. IEEE Transactions on Geoscience and Remote Sensing, 45(3), 765–777. https://doi.org/10.1109/TGRS.2006.888466

Moroni, M., Lupo, E., Marra, E., & Cenedese, A. (2013). Hyperspectral image analysis in environmental monitoring: Setup of a new tunable filter platform. Procedia Environmental Sciences, 19, 885–894. https://doi.org/10.1016/j.proenv.2013.06.098

Nascimento, J. M., & Bioucas-Dias, J. M. (2007). Hyperspectral unmixing algorithm via dependent component analysis. Proceedings of the 2007 IEEE International Geoscience and Remote Sensing Symposium, Barcelona, Spain, (pp. 4033–4036). IEEE.

Nascimento, J. M. P., & Bioucas-Dias, J. M. (2012). Hyperspectral unmixing based on mixtures of Dirichlet components. IEEE Transactions on Geoscience and Remote Sensing, 50(3), 863–878. https://doi.org/10.1109/TGRS.2011.2163941

Nascimento, J. M., & Dias, J. M. (2005). Vertex component analysis: A fast algorithm to unmix hyperspectral data. IEEE Transactions on Geoscience and Remote Sensing, 43 (4), 898–910. https://doi.org/10.1109/TGRS.2005.844293

Ni, L., Xu, H., & Zhou, X. (2020). Mineral identification and mapping by synthesis of hyperspectral vnr/swir and multispectral tir remotely sensed data with different classifiers. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, 3155–3163. https://doi.org/10.1109/JSTARS.2020.2999057

Ozkan, S., Kaya, B., & Akar, G. B. (2018). Endnet: Sparse autoencoder network for endmember extraction and hyperspectral unmixing. IEEE Transactions on Geoscience and Remote Sensing, 57(1), 482–496. https://doi.org/10.1109/TGRS.2018.2856299

Ozkan, S., Kaya, B., & Akar, G. B. (2019). Endnet: Sparse autoencoder network for endmember extraction and hyperspectral unmixing. IEEE Transactions on Geoscience and Remote Sensing, 57(1), 482–496. https://doi.org/10.1109/TGRS.2018.2856299

Palsson, B., Sigurdsson, J., Sveinsson, J. R., & Ulfarsson, M. O. (2018). Hyperspectral unmixing using a neural network autoencoder. Institute of Electrical and Electronics Engineers Access, 6, 25646–25656. https://doi.org/10.1109/ACCESS.2018.2818280

Palsson, B., Ulfarsson, M. O., & Sveinsson, J. R. (2019). Convolutional autoencoder for spatial-spectral hyperspectral unmixing. Proceedings of the IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, (pp. 357–360). IEEE.

Palsson, B., Ulfarsson, M. O., & Sveinsson, J. R. (2020). Convolutional autoencoder for spectral-spatial hyperspectral unmixing. IEEE Transactions on Geoscience and Remote Sensing, 59(1), 535–549. https://doi.org/10.1109/TGRS.2020.2992743

Parra, L., Spence, C., Sajda, P., Ziehe, A., & Müller, K.-R. (1999). Unmixing hyperspectral data. Advances in Neural Information Processing Systems, 12, 12. https://proceedings.neurips.cc/paper_files/paper/1999/file/798cd7d4ee7138d49b8828958048130a-Paper.pdf

Qian, Y., Jia, S., Zhou, J., & Robles-Kelly, A. (2010). L1/2 sparsity constrained nonnegative matrix factorization for hyperspectral unmixing. Proceedings of the 2010 International conference on digital image computing: techniques and applications, Sydney, Australia, (pp. 447–453). IEEE.

Qi, L., Li, J., Wang, Y., Lei, M., & Gao, X. (2020). Deep spectral convolution network for hyperspectral image unmixing with spectral library. Signal Processing, 176, 107672. https://doi.org/10.1016/j.sigpro.2020.107672

Qu, Y., & Qi, H. (2018). Udas: An untied denoising autoencoder with sparsity for spectral unmixing. IEEE Transactions on Geoscience and Remote Sensing, 57(3), 1698–1712. https://doi.org/10.1109/TGRS.2018.2868690

Rajabi, R., & Ghasssemian, H. (2014). Spectral unmixing of hyperspectral imagery using multilayer nfm. IEEE Geoscience and Remote Sensing Letters, 12(1), 38–42. https://doi.org/10.1109/LGRS.2014.2325874

Ranasinghe, Y., Herath, S., Weerasooriya, K., Ekanyake, M., Godaliyadda, R., Ekanyake, P., & Herath, V. (2020a). Convolutional autoencoder for blind hyperspectral image unmixing. Proceedings of the 2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS), IIT Ropar, India, (pp. 174–179). IEEE.

Ranasinghe, Y., Herath, S., Weerasooriya, K., Ekanyake, M., Godaliyadda, R., Ekanyake, P., & Herath, V. (2020b). Convolutional autoencoder for blind hyperspectral image unmixing. Proceedings of the 2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS), IIT Ropar, India, (pp. 174–179).

Rasti, B., Scheunpers, D., Ghamisi, P., Licciardi, G., & Chaussod, J. (2018). Noise reduction in hyperspectral imagery: Overview and application. Remote Sensing, 10 (3), 482. https://doi.org/10.3390/rs10030482

Rathnayake, B., Ekanyake, E., Weerakoon, K., Godaliyadda, G., Ekanyake, M., & Herath, H. (2020). Graph-based blind hyperspectral unmixing via nonnegative matrix factorization. IEEE Transactions on Geoscience and Remote Sensing, 58(9), 6391–6409. https://doi.org/10.1109/TGRS.2020.2976799

Signoroni, A., Savardi, M., Baronio, A., & Benini, S. (2019). Deep learning meets hyperspectral image analysis: A multidisciplinary review. Journal of Imaging, 5(5), 52. https://doi.org/10.3390/imaging5050052

Sigurdsson, J., Ulfarsson, M., & Sveinsson, J. (2016). Blind hyperspectral unmixing using total variation and $\ell_q$ sparse regularization. IEEE Transactions on Geoscience and Remote Sensing, 54(11), 6371–6384. https://doi.org/10.1109/TGRS.2016.2582824

Stuart, M. B., McGonigle, A. J., & Willmott, J. R. (2019). Hyperspectral imaging in environmental monitoring: A review of recent developments and technological advances in compact field deployable systems. Sensors, 19(14), 3071. https://doi.org/10.3390/s19143071

Su, Y., Li, J., Plaza, A., Marinoni, A., Gamba, P., & Chakravorty, S. (2019). Daen: Deep autoencoder networks for hyperspectral unmixing. IEEE Transactions on Geoscience and Remote Sensing, 57(7), 4309–4321. https://doi.org/10.1109/TGRS.2018.2890633

Sun, X., Peng, Q., Zhang, B., Gao, L., & Yang, L. (2018). Constraint non-negative matrix factorization with sparseness and piece wise smoothness for hyperspectral
unmixing. Proceedings of the 2018 9th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 23–26 September 2018, Amsterdam, The Netherlands, (pp. 1–5). IEEE.

Tao, C., Meng, Y., Li, J., Yang, B., Hu, F., Li, Y., Cui, C., & Zhang, W. (2022). Msnet: Multispectral semantic segmentation network for remote sensing images. GScience & Remote Sensing, 59(1), 1177–1198. https://doi.org/10.1080/15481603.2022.2101728

Teke, M., Deveci, H. S., Haliloglou, O., Gürbüz, S. Z., & Sakarya, U. (2013). A short survey of hyperspectral remote sensing applications in agriculture. Proceedings of the 2013 6th international conference on recent advances in space technologies (RAST), Istanbul, Turkey, (pp. 171–176). IEEE.

Wang, L., Wang, S., Jia, X., & Bi, T. (2021). A novel hyperspectral unmixing method based on least squares twin support vector machines. European Journal of Remote Sensing, 54(1), 72–85. https://doi.org/10.1080/22797254.2021.1877572

Wang, M., Zhao, M., Chen, J., & Rahardja, S. (2019). Nonlinear unmixing of hyperspectral data via deep autoencoder networks. IEEE Geoscience and Remote Sensing Letters, 16(9), 1467–1471. https://doi.org/10.1109/LGRS.2019.2900733

Winter, M. E. (1999). N-findr: An algorithm for fast autonomous spectral end-member determination in hyperspectral data. In Imaging spectrometry V (Vol. 3753, pp. 266–275). International Society for Optics and Photonics.

Wu, X., Hong, D., & Chanussot, J. (2023). Uiu-net: U-net in u-net for infrared small object detection. IEEE Transactions on Image Processing, 32, 364–376. https://doi.org/10.1109/TIP.2022.3228497

Yang, S., Zhang, X., Yao, Y., Cheng, S., & Jiao, L. (2015). Geometric nonnegative matrix factorization (gnmf) for hyperspectral unmixing. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8(6), 2696–2703. https://doi.org/10.1109/JSTARS.2015.2417574

Zare, A., & Gader, P. (2007). Sparsity promoting iterated constrained endmember detection in hyperspectral imagery. IEEE Geoscience and Remote Sensing Letters, 4 (3), 446–450. https://doi.org/10.1109/LGRS.2007.895727

Zeng, D., Zhang, S., Chen, F., & Wang, Y. (2019). Multi-scale CNN based garbage detection of airborne hyper-spectral data. Institute of Electrical and Electronics Engineers Access, 7, 104514–104527. https://doi.org/10.1109/ACCESS.2019.2932117

Zhang, X., Sun, Y., Zhang, J., Wu, P., & Jiao, L. (2018). Hyperspectral unmixing via deep convolutional neural networks. IEEE Geoscience and Remote Sensing Letters, 15 (11), 1755–1759. https://doi.org/10.1109/LGRS.2018.2857804

Zhao, M., Shi, S., Chen, J., & Dobigeon, N. (2022). A 3d-cnn framework for hyperspectral unmixing with spectral variability. Proceedings of the IEEE Transactions on Geoscience and Remote Sensing.

Zhao, M., Wang, M., Chen, J., & Rahardja, S. (2021). Hyperspectral unmixing for additive nonlinear models with a 3-d-cnn autoencoder network. Proceedings of the IEEE Transactions on Geoscience and Remote Sensing.

Zhao, M., Yan, L., & Chen, J. (2021). Lstm-dnn based autoencoder network for nonlinear hyperspectral image unmixing. IEEE Journal of Selected Topics in Signal Processing, 15(2), 295–309. https://doi.org/10.1109/JSTSP.2021.3052361

Zhao, H., & Zhao, X. (2019). Nonlinear unmixing of minerals based on the log and continuum removal model. European Journal of Remote Sensing, 52(1), 277–293. https://doi.org/10.1080/22797254.2019.1610199

Zhou, J., Sun, W., Meng, X., Yang, G., Ren, K., & Peng, J. (2022). Generalized linear spectral mixing model for spatial–temporal–spectral fusion. IEEE Transactions on Geoscience and Remote Sensing, 60, 1–16. https://doi.org/10.1109/TGRS.2022.3188501

Zhou, L., Zhang, X., Wang, J., Bai, X., Tong, L., Zhang, L., Zhou, J., & Hancock, E. (2020). Subspace structure regularized nonnegative matrix factorization for hyperspectral unmixing. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, 4257–4270. https://doi.org/10.1109/JSTARS.2020.3011257

Zhuang, L., Lin, C.-H., Figueiredo, M. A., & Bioucas-Dias, J. M. (2019). Regularization parameter selection in minimum volume hyperspectral unmixing. IEEE Transactions on Geoscience and Remote Sensing, 57(12), 9858–9877. https://doi.org/10.1109/TGRS.2019.2929776

Zhu, M., Huang, D., Hu, X.-J., Tong, W.-H., Han, B.-L., Tian, J.-P., & Luo, H.-B. (2020). Application of hyperspectral technology in detection of agricultural products and food: A review. Food Science & Nutrition, 8(10), 5206–5214. https://doi.org/10.1002/fsn3.1852