Journal of Future Artificial Intelligence and Technologies

E-ISSN: 3048-3719

Research Article

Supervised-Blind Unmixing Comparison Technique with MiSiCNet and Deep Image Prior Model for Accurate Abundance Estimation in Hyperspectral Datasets

Bipasha Chakrabarti Banik *

JIS Institute of Advanced Studies and Research (JISIASR), Howrah, Kolkata, West Bengal 711112, India; e-mail : bcbipasha@gmail.com

* Corresponding Author : Bipasha Chakrabarti Banik

Abstract: Hyperspectral unmixing is a fundamental task in remote sensing that aims to decompose mixed spectral pixels into their constituent materials and estimate their respective abundances. Traditional methods often face limitations due to nonlinear mixing effects and the lack of pure pixels. This paper proposes a deep learning-based comparative framework that integrates modified versions of the Minimum Simplex Convolutional Network (MiSiCNet) and Unsupervised Deep Image Prior (UnDIP), referred to as Blind MiSiCNet and Supervised UnDIP, respectively, to achieve robust unmixing and abundance estimation in real-world scenarios. The proposed Blind MiSiCNet removes downsampling and upsampling layers and leverages spatial and geometric priors through convolutional layers with a minimum simplex volume constraint, producing reliable initial abundance maps even in the absence of pure pixels. The Supervised UnDIP variant further refines these estimates using the implicit regularization of convolutional neural networks, incorporating known endmembers to generate noise-free and spatially coherent abundance maps. Experimental results on the real-world Jasper Ridge dataset demonstrate that the proposed supervised and blind unmixing methods, evaluated comparatively, significantly outperform existing approaches regarding Spectral Angle Distance (SAD) and Root Mean Square Error (RMSE). The results also highlight improved noise robustness and better preservation of spatial structures.

Keywords: Abundance Estimation; Blind Unmixing; Deep Learning; Endmembers; Hyperspectral Image; Hyperspectral Unmixing; Spectral Angle Distance; Supervised Unmixing.

Received: April, 23rd 2025
Revised: June, 12th 2025
Accepted: June, 13th 2025
Published: June, 20th 2025
Curr. Ver.: June, 20th 2025

Copyright: © 2025 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY SA) license (https://creativecommons.org/licenses/by-sa/4.0/)

1. Introduction

A spectral pixel in a hyperspectral image typically comprises a combination of pure spectral signatures (referred to as endmembers) corresponding to the materials present within the pixel. Estimating the fractional contributions of these endmembers is known as spectral unmixing. In cases where materials are intimately mixed, and multiple reflections occur before the signal reaches the sensor, nonlinear unmixing becomes necessary. Conversely, linear unmixing has proven effective in remote sensing scenarios where hyperspectral images predominantly consist of spatially homogeneous material regions [1]–[3].

When pure pixels for each material are available in the scene, geometrical techniques can be employed for endmember extraction by leveraging the simplex structure of the data. Once endmembers are identified, abundance estimation can be performed by minimizing the reconstruction error between the observed and reconstructed spectra, subject to the abundance non-negativity constraint (ANC) and the abundance sum-to-one constraint (ASC) [4]. Because endmembers are assumed to reside at the vertices of the data simplex, methods such as Simplex Volume Maximization (SiVM) can be applied to identify them [5]. Additionally,

convex optimization techniques have been proposed to solve the Maximum Volume Inscribed Simplex (MVIS) problem, which aims to locate data points forming a simplex of maximal volume within the convex hull of the data [6].

Several alternative geometrical approaches include Vertex Component Analysis (VCA), which iteratively projects data in orthogonal directions to extract extreme points [7]; the Pixel Purity Index (PPI), which evaluates spectral extremeness based on random vector projections [8]; and N-FINDR, which seeks to find the set of pure pixels that form the largest possible simplex by expanding an initial simplex inside the data cloud [9]. In scenarios where pure pixels are absent, geometrical methods can still perform effectively if sufficient data points lie along the facets of the data simplex. In such cases, virtual endmembers at the simplex vertices can be estimated using methods like the Minimum Volume Enclosing Simplex (MVES), which does not rely on the pure-pixel assumption [10], [11]. The Maximum Volume Inscribed Ellipsoid (MVIE) approach identifies the largest ellipsoid within the data convex hull, using contact points on simplex facets as cues to infer virtual endmembers [12]. In blind unmixing scenarios—where both endmembers and abundances are simultaneously estimated—constrained or penalized least squares methods, commonly aligned with nonnegative matrix factorization, are widely used [13], [14]. These methods often incorporate geometric priors via penalty terms, such as the squared simplex volume [15] or Euclidean distances between endmembers and the data center [13]. Other geometrical penalties include the distance between estimated endmembers and simplex extremes identified through methods like VCA [16], and the total variation (TV) across endmembers [17]. Selecting an appropriate regularization parameter to balance the data fidelity and penalty terms remains a significant challenge. A parameter selection strategy for such geometrical penalties was introduced in [18]. Alternatively, the unmixing problem can be cast as a non-convex optimization (or non-concave maximization) task, where the goal is to find the minimum volume simplex subject to ANC and ASC [19], [20]. These non-convex methods are often highly sensitive to initialization.

Another class of techniques, sparse unmixing, estimates fractional abundances by formulating a sparse regression problem using a comprehensive dictionary of pure spectral signatures. Each observed spectrum is approximated as a sparse linear combination of dictionary elements. This results in a penalized least squares problem, where a sparsity-promoting regularizer is applied to the abundances [21], [22]. For example, SUnCNN is a recent deep learning-based method that leverages convolutional neural networks for sparse unmixing [23]. However, a major drawback of sparse unmixing is its reliance on external spectral libraries, which may compromise accuracy due to spectral mismatch. A variation known as bundle-based sparse unmixing addresses this by constructing endmember bundles from the data itself. Collaborative LASSO, a representative method in this category, has effectively managed endmember variability [24]–[26].

Deep learning has gained significant traction in hyperspectral image analysis in recent years, particularly for unmixing tasks. Autoencoder-based architectures are widely employed [27]–[31], where abundances are extracted by enforcing ANC and ASC constraints in the encoder, and endmembers are learned as decoder weights. Several variants have been proposed: SNSA uses stacked nonnegative sparse autoencoders for robustness to outliers [32]; DAEN combines a stacked and variational autoencoder for joint estimation [33]; a VAE-based model generates endmembers directly [34]; and uDAS introduces denoising and sparsity constraints in the decoder [35]. Other methods include sparse autoencoders incorporating Kullback–Leibler divergence, SAD, and sparsity terms [36].

However, most existing autoencoder-based and sparse learning models process pixels individually, ignoring critical spatial correlations across neighboring regions [37]. To address this, spatially-aware models have been introduced, including adaptive smoothing-based autoencoders [38], patch-wise or cube-wise convolutional autoencoders [39], and parallel spectral patch-based autoencoders [40]. Recently, a 3D convolutional autoencoder was proposed for supervised unmixing, assuming known endmembers [41]. While spatial convolutions improve endmember estimation by capturing spectral variability, they may lead to blurred abundance maps[42]. Unmixing using Deep Image Prior (UnDIP) [43] offers an alternative by leveraging an implicit regularization framework. It combines convolutional encoder-decoder architecture with fixed noise inputs to optimize reconstruction and spatial coherence without requiring ground truth labels [44]. However, standard UnDIP requires pre-extracted endmembers (typically via geometrical methods such as SiVM), limiting its application in pure-pixel-scarce environments.

In contrast, cycle-consistent models such as CyCUNet [45] introduce dual autoencoders and composite loss terms to promote spectral and spatial consistency. Yet, they often suffer from optimization complexity and parameter sensitivity. Nevertheless, UnDIP remains attractive due to its simplicity, robustness to noise, and general applicability to unsupervised cases. In particular, simplex volume minimization has shown promise for blind unmixing without pure pixels [46]. Despite these advances, there is still a clear methodological gap in developing unified deep learning models that can handle both supervised and blind unmixing structurally consistent and computationally efficient, especially under real-world noise and spatial variability.

We propose a comparative analysis of two purpose-built architectures to address this gap: a modified Deep Image Prior model (Supervised UnDIP) and a novel geometry-preserving network (Blind MiSiCNet). These models are designed to jointly estimate both abundance maps and the number of endmembers, using spatially regularized encoder–decoder frameworks tailored for their respective settings—supervised and blind. Supervised UnDIP introduces a minimal but effective encoder–decoder variant that preserves spatial coherence with low computational cost, while Blind MiSiCNet eliminates all downsampling operations and incorporates learnable skip connections to enhance feature continuity under uncertainty.

This study is the first to systematically contrast these two paradigms within a unified deep learning framework, highlighting their strengths in handling noisy, real-world hyperspectral datasets.

2. Literature Review: Previous Methods

Paura et al. [47] proposed hyperspectral unmixing, addressing the challenge of spectral mixing due to the low spatial resolution of hyperspectral images. Traditional approaches such as Nonnegative Matrix Factorization (NMF), Linear Mixture Modeling (LMM), and Autoencoder Networks have been widely used to estimate endmembers and their abundances. However, the diverse nature of hyperspectral datasets makes it difficult to determine the generalizability and effectiveness of these methods. This paper proposes a benchmarking methodology for evaluating hyperspectral unmixing algorithms to address this issue. This study aims to compare existing algorithms and identify the best-performing ones by conducting experiments on multiple hyperspectral datasets.

Karoui and Djerriri [48] addressed the challenge of shadow effects in hyperspectral remote sensing data, which can hinder automatic interpretation due to lower intensity and fuzzy boundaries in shaded areas. Shadow correction, typically consisting of detection and de-shadowing steps, is a widely used to mitigate this issue. Focusing solely on the de-shadowing step, this paper proposes a hyperspectral unmixing-based shadow correction method. The approach is evaluated using a real hyperspectral image, and experimental results demonstrate that the proposed method achieves effective de-shadowing, outperforming an existing unmixing-based method from the literature.

The paper [49] explored sparse unmixing, a spectral unmixing approach that models observed hyperspectral signatures using a linear sparse regression framework with a large spectral library. Spatial information has been incorporated through spatial regularization in sparse unmixing formulations to enhance unmixing accuracy. However, traditional spatial regularization sparse unmixing (SRSU) algorithms face challenges in selecting appropriate user-defined regularization parameters and often involve high computational costs. This paper proposes an adaptive spatial regularization sparse unmixing (ASRSU) strategy based on the joint maximum a posteriori (JMAP) estimation technique to address this issue. ASRSU jointly estimates the regularization parameters and abundances through an alternating iterative process, deriving their relationships from the JMAP model. Two ASRSU algorithms are introduced: adaptive total variation spatial regularization sparse unmixing and adaptive nonlocal means filtering sparse unmixing. Experimental results on three simulated datasets and two real hyperspectral images demonstrate that the proposed ASRSU algorithms can adaptively determine near-optimal regularization parameters, improving unmixing performance.

Li et al. [50] presented graph-based hyperspectral unmixing, where graph construction is used to model relationships between pixels for cooperative analysis rather than independently analyzing each pixel. Graph theory provides a powerful mathematical framework for structuring spatial and spectral relationships in hyperspectral imagery. The paper reviews fundamental graph concepts and explores various methods for constructing graphs in both

spatial and spectral domains. A general hyperspectral unmixing problem is formulated by incorporating graph regularization, which has applications in remote sensing and environmental monitoring. The Alternating Direction Method of Multipliers (ADMM) is introduced as an optimization tool for solving the unmixing problem. Experimental results on both synthetic and real hyperspectral data validate the effectiveness of the proposed approach.

Chi et al. [51] explored spectral unmixing for estimating and monitoring crop residue cover using airborne and space-based hyperspectral sensors. Crop residue is crucial in improving soil quality, water use efficiency, and erosion control, making its accurate assessment important for sustainable land management. Traditional hyperspectral indices, such as those based on band differences and ratios, are commonly used but limited by push-broom sensors like Hyperion's low signal-to-noise ratio (SNR). The study investigates spectral unmixing as an alternative approach to address this limitation. Additionally, it compares linear unmixing with manifold learning-based unmixing to account for nonlinearities in hyperspectral data. Experimental results demonstrate that manifold learning approaches provide more robust crop residue cover estimates than the Cellulose Absorption Index (CAI) and linear unmixing applied to airborne and Hyperion data.

Again, Li et al. [52] presented a superpixel-based approach for hyperspectral unmixing, leveraging both spatial and spectral correlations to improve unmixing performance. Traditional pixel-wise unmixing can be computationally expensive and fails to utilize spatial dependencies fully. To address this, the paper proposes using Simple Linear Iterative Clustering (SLIC) to group adjacent pixels into superpixels with similar spectral signatures. Unmixing is then performed within these superpixels using graph-based total variation regularization, which accounts for local heterogeneity while maintaining computational efficiency. Experimental results on synthetic and real hyperspectral data demonstrate the advantages of the proposed approach over conventional pixel-wise unmixing methods.

Sen et al. [53] addressed the role of dimensionality reduction in hyperspectral unmixing, a process that decomposes a mixed pixel's spectrum into constituent endmembers and their mixing proportions. Due to the high spectral dimensionality of hyperspectral images, directly applying unmixing methods leads to increased computational complexity and slower processing. To address this, band selection is introduced as a preprocessing step to retain the most relevant wavelengths while reducing data size. The study employs affinity propagation for band selection and constrained nonnegative matrix factorization (NMF) for unmixing. Experimental results indicate that the band selection-based unmixing approach achieves comparable performance to unmixing without band selection, demonstrating its potential for improving computational efficiency.

While the approaches mentioned earlier have achieved notable progress in hyperspectral unmixing, their performance may degrade in the presence of high noise levels, outliers, or when dealing with small image patches. To address these challenges, this paper explores an alternative framework by incorporating blind and supervised unmixing strategies based on deep learning.

3. Proposed Method

Hyperspectral unmixing presents significant challenges in real-world remote sensing tasks, especially when pure pixels are absent and spectral mixtures are nonlinear or contaminated with noise. Traditional methods, including the original Deep Image Prior (DIP) by Ulyanov et al. [44], were designed for general-purpose image restoration and do not natively support constraints such as ANC and ASC, which are fundamental in spectral unmixing.

Moreover, baseline deep learning-based unmixing methods, such as SUnCNN or standard autoencoders, often employ deep and complex architectures with multiple downsampling stages. While these methods perform well in idealized settings, they tend to degrade in performance when faced with spatial variability, limited labeled data, or the absence of pure endmembers. To overcome these limitations, we propose two lightweight yet task-specific architectures that explicitly address abundance estimation and endmember extraction using deep learning:

- Supervised UnDIP: an improved, constraint-aware variant of DIP tailored for supervised unmixing scenarios where endmembers are either known or can be extracted reliably.

- Blind MiSiCNet: a geometry-aware convolutional architecture designed for fully unsupervised unmixing without pure pixels.

Both models integrate customized skip connections, noise-based input priors, and efficient convolutional pipelines that facilitate accurate spectral reconstruction while maintaining spatial fidelity and computational efficiency.

As illustrated in Figure 1 and Figure 2, the proposed framework consists of two encoder–decoder-based architectures. Both models take a random noise input of the same spatial dimension as the hyperspectral image and aim to estimate the abundance maps and the number of endmembers.

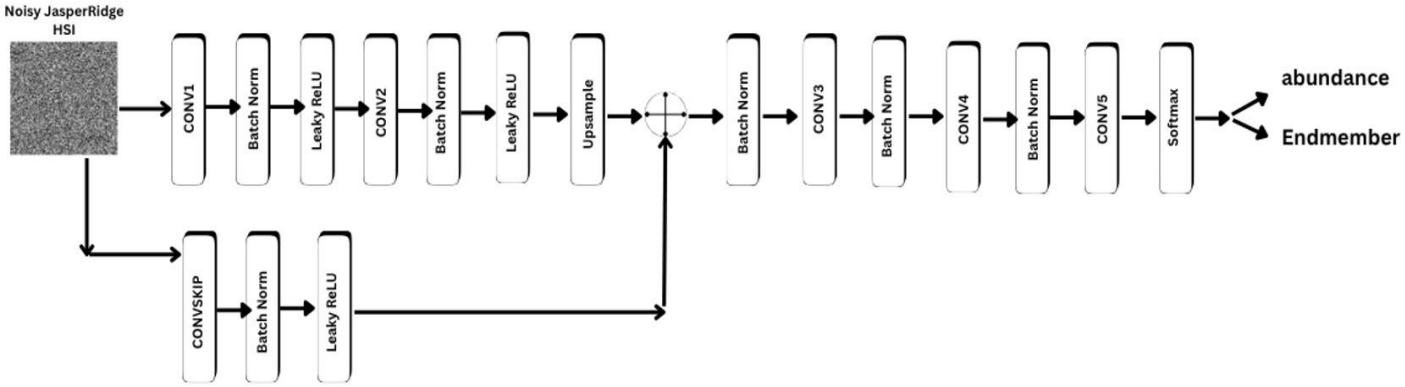

Figure 1. Proposed UnDIP model architecture for Supervised Unmixing

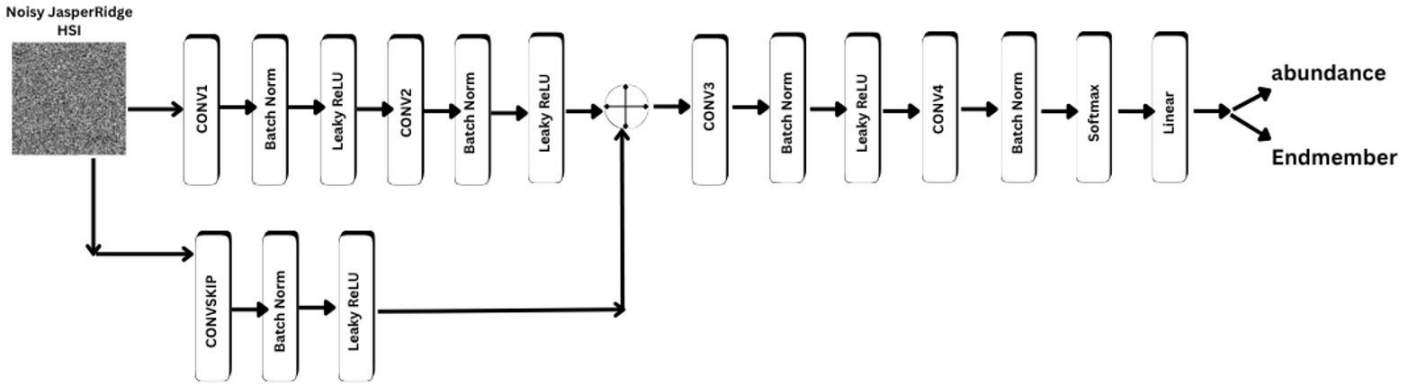

Figure 2. Proposed MiSiCNet model architecture for Blind Unmixing

Supervised UnDIP (Figure 1) adopts a shallow hourglass-shaped encoder-decoder structure. It incorporates a downsampling block to extract compact features, followed by an upsampling block for spatial reconstruction. A skip connection directly merges low-level features from the encoder into the decoder to preserve local structures. While the proposed Supervised UnDIP model is inspired by the original Deep Image Prior framework [43], it differs substantially in architecture and objective. Unlike the original UnDIP, which performs unsupervised abundance estimation using externally extracted endmembers, our version integrates supervision by incorporating known or estimated endmembers during training.

Furthermore, we introduce a lightweight encoder-decoder with a single skip connection and minimal downsampling, allowing simultaneous estimation of abundance and the number of endmembers, with improved convergence and spatial regularity. Blind MiSiCNet (Figure 2) maintains full-resolution spatial processing throughout its convolutional layers. It removes all downsampling and upsampling stages and employs reflection padding and a specialized ConvSkip residual connection to enhance feature flow and maintain structural coherence. Both models include a final softmax layer to enforce abundance constraints (ANC and ASC). The output branches simultaneously predict the abundance distributions and the estimated number of endmembers per pixel.

The Supervised UnDIP architecture leverages the core idea of Deep Image Prior but is redesigned to be more efficient and task-specific for abundance estimation. This model is

particularly effective for supervised settings where endmember spectra are known or estimated separately and training convergence and spatial coherence are critical. Key features include:

- Encoder–decoder with minimal downsampling: Only one downsampling block is applied using stride-based convolution to reduce spatial resolution and extract deep features, followed by a bilinear upsampling block to restore the resolution.
- Skip connection: A single skip connection is used to merge encoder features into the decoder, supporting the preservation of spatial detail and improving gradient flow.
- Layer components consist of Convolutional layers (Conv1–Conv5) for hierarchical feature extraction. Batch Normalization (BN) after each convolution to stabilize learning. Leaky ReLU is the activation function that prevents dying neurons. Softmax at the final layer to enforce abundance non-negativity and sum-to-one properties.

The MiSiCNet model is designed to operate without assuming that pure pixels or labeled endmembers are available. MiSiCNet is ideal for blind unmixing cases where endmembers' number and spectral signatures must be inferred directly from mixed pixels. Its architectural design ensures spatial consistency and robustness under data uncertainty:

- No downsampling or upsampling: The entire network operates at the input resolution, allowing it to capture and preserve detailed spatial structures in abundance maps.
- ConvSkip-based skip connection: A learnable ConvSkip block forms a residual connection instead of a simple concatenation. This allows direct propagation of low-level features and assists in preserving fine-grained spectral information.
- Layer components consost of four sequential convolutional layers (Conv1–Conv4) followed by a linear projection. Reflection padding in each convolutional layer to prevent spatial distortions near boundaries. Batch Normalization and Leaky ReLU are applied consistently for stable and nonlinear transformations. A softmax layer, followed by a linear layer, enforces constraints and maps to abundance and endmember outputs.

A side-by-side architectural and design comparison is presented in Table 1, summarizing key differences such as convolutional depth, downsampling strategy, skip connection type, and convergence behavior.

Table 1. Comparative analysis of MiSiCNet and UnDIP architectures.

Facet	MiSiCNet	UnDIP
Network Type	Convolutional encoder–decoder with ConvSkip	Shallow encoder–decoder with skip connection
Input	Random noise (same shape as HSI)	Random noise (same shape as HSI)
Convolutional Layers	4 Conv layers + 1 ConvSkip	5 Conv layers (2×3-layer blocks + 1 final)
Skip Connection	One ConvSkip residual connection	One residual-style skip connection
Downsampling	None	One block with stride convolution
Upsampling	Not required	Bilinear upsampling
Padding	Reflection padding	Zero or replication padding
Activation	Leaky ReLU (all layers) + softmax	Leaky ReLU (all layers) + softmax
Output Objective	Abundance + Endmembers	Abundance + Endmembers
Design Motivation	Preserve spatial fidelity, structural continuity	Faster convergence, task-specific supervision
Convergence Behavior	Moderate	Faster than MiSiCNet and baseline DIP

4. Results and Analysis

4.1. Dataset Description

In this work, we considered widely used real hyperspectral datasets such as Jasper Ridge HSI. The Jasper Ridge dataset, popularly used for unmixing, contains 512 × 614 pixels and 224 spectral bands. As per ground truth, the dataset consists of four endmembers: road, soil, water, and tree. Each pixel is recorded at 224 channels ranging from 380 nm to 2500 nm. The spectral resolution is up to 9.46nm.

4.2. Experimental Results

Both models were implemented using PyTorch and trained with identical optimization settings. A summary of the hyperparameter setup is provided in Table 2, including learning rate, number of epochs, optimizer, and padding strategies. The models were trained on the entire image with a batch size of 1 for 3000–5000 epochs or until convergence.

Table 2. Hyperparameters used in UnDIP and MiSiCNet models.

Hyperparameters Used	UnDIP	MiSiCNet
Input Type	Random noise (same size as image)	Random noise (same size as image)
Number of Convolutional Layers	5 (2 blocks × 3 layers + final Conv)	4 Conv layers + 1 ConvSkip
Skip Connection	1 skip connection	1 skip connection (ConvSkip)
Downsampling	1 block (moderate resolution reduction)	None (to preserve spatial resolution)
Upsampling	Bilinear upsampling	Not required
Filter Sizes	3×3	3×3
Padding	Zero padding or replication padding	Reflection padding
Batch Normalization	After every convolutional layer	After every convolutional layer
Activation Function	Leaky ReLU (all layers except final), Softmax (final)	Leaky ReLU (all layers except final), Softmax (final)
Learning Rate	0.001	0.001
Optimizer	Adam optimizer	Adam optimizer
Loss Function	MSE or customized reconstruction loss	MSE or customized reconstruction loss
Batch Size	1 (entire image used in training)	1 (entire image used in training)
Number of Epochs	3000–5000	3000–5000

4.3. Experimental Results

Figure 3 illustrates the estimated abundance maps for four endmembers, i.e., tree, water, dirt, and road, generated using the MiSiCNet and UnDIP models. Each map shows the spatial distribution of material proportions across the hyperspectral image. For Tree Abundance MiSiCNet, it captures the tree regions well but shows slight overestimation in adjacent areas, particularly in the lower-left and central regions, indicating spectral mixing. UnDIP delivers sharper and more confined tree abundance with better suppression in non-tree regions, implying better generalization and spectral separation. So, UnDIP provides more accurate localization of tree areas with reduced spatial ambiguity. For water abundance, both MiSiCNet and UnDIP produce similar and highly confident abundance values for water bodies (clearly seen as bright yellow regions). Slight differences can be observed at the water-land boundaries, where MiSiCNet shows more blending. Both models effectively detect water, but UnDIP may preserve boundary precision better. For dirt abundance, MiSiCNet shows a strong and widespread response for dirt but includes false positives in some vegetative regions.UnDIP yields more refined dirt detection with clearer boundaries and better noise suppression, particularly in the top-center area.UnDIP improves specificity for dirt regions and avoids cross-material leakage. For road abundance, MiSiCNet detects road regions but with scattered activation and slightly noisy representation. UnDIP enhances road visibility by producing a clearer, continuous structure, particularly along curved paths, which aligns more closely with the expected ground truth. From this comparative analysis of the abundance maps, UnDIP demonstrates superior performance over MiSiCNet regarding spatial precision, material-specific localization, and suppression of false responses. This makes UnDIP more effective, especially for complex land-cover materials like roads and trees where boundary clarity and accurate material unmixing are critical.

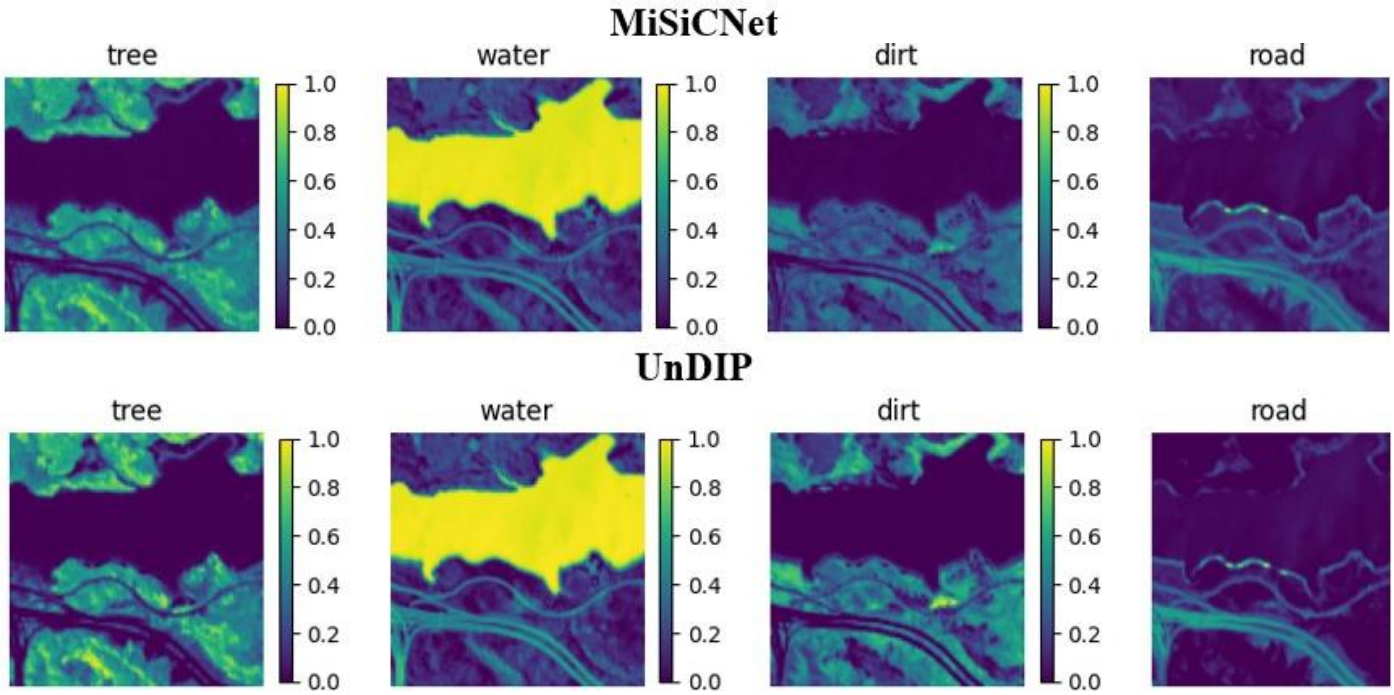

Figure 3. Abundance Maps obtained from UnDIP and MiSiCNet Model

Figure 4 and Figure 5 represent the abundance maps extracted using the MiSiCNet and UnDIP models for four land cover materials: tree, water, dirt, and road. These maps demonstrate the spatial distribution of each material across the hyperspectral scene. The UnDIP model presents sharper and more spatially consistent abundance patterns compared.

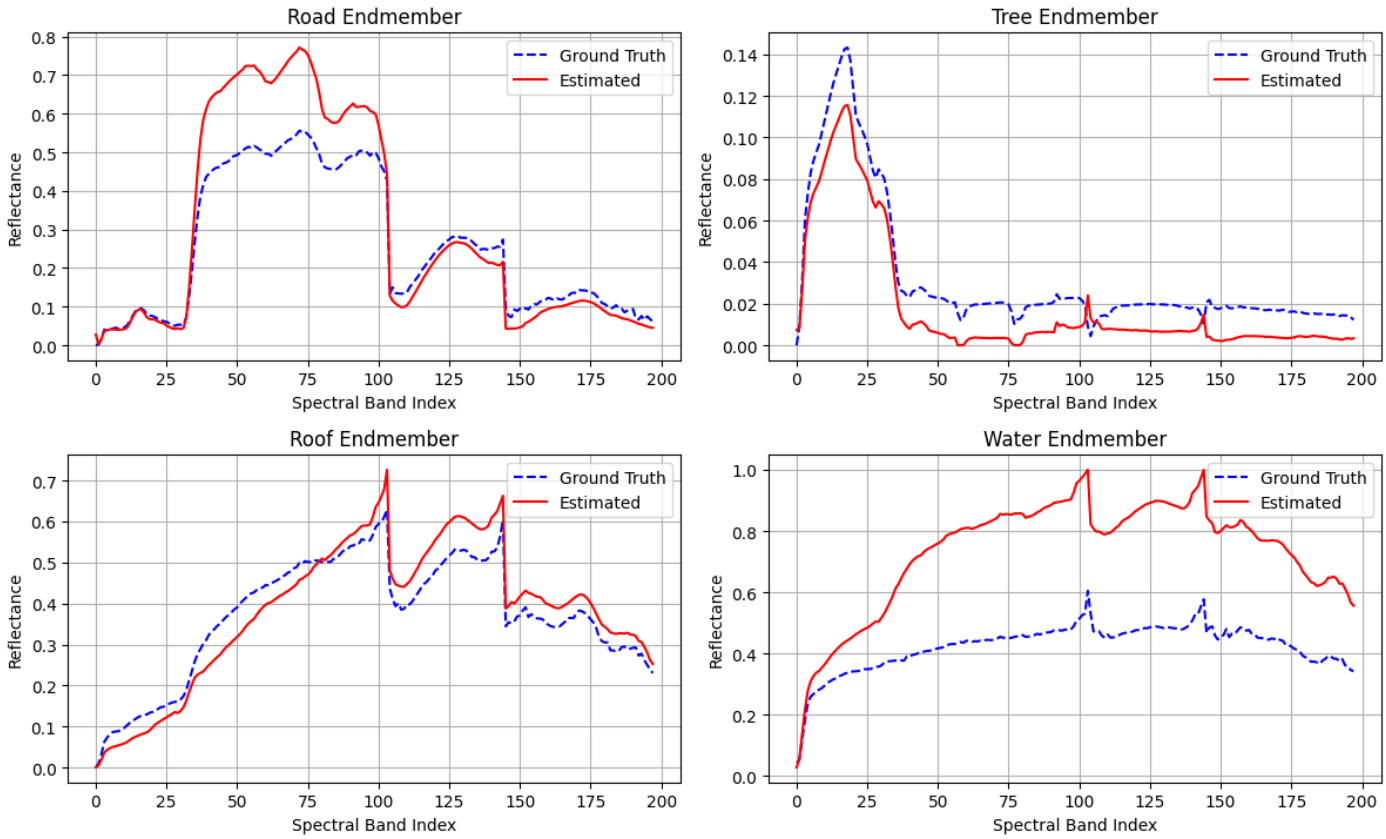

Figure 4. JasperRidge dataset - The visual comparison of the endmembers was obtained using the UnDIP Model. Red: the estimated endmembers; Blue: the ground truth endmembers.

to MiSiCNet. The water body is well localized, with minimal mixing in neighboring pixels. The road and dirt maps show finer spatial transitions, indicating that UnDIP better preserves sub-pixel variability and spatial structure. The abundance maps generated by MiSiCNet are smoother and show distinct separation between classes. For instance, the tree and water classes are delineated, and the abundance values range sharply from 0 to 1, indicating strong material presence. However, some regions in the dirt and road maps show blurring at boundaries, potentially due to over-smoothing in feature encoding. In conclusion, UnDIP outperforms MiSiCNet as UnDIP provides more precise material localization, which is a crucial factor in real-world remote sensing where materials may overlap.

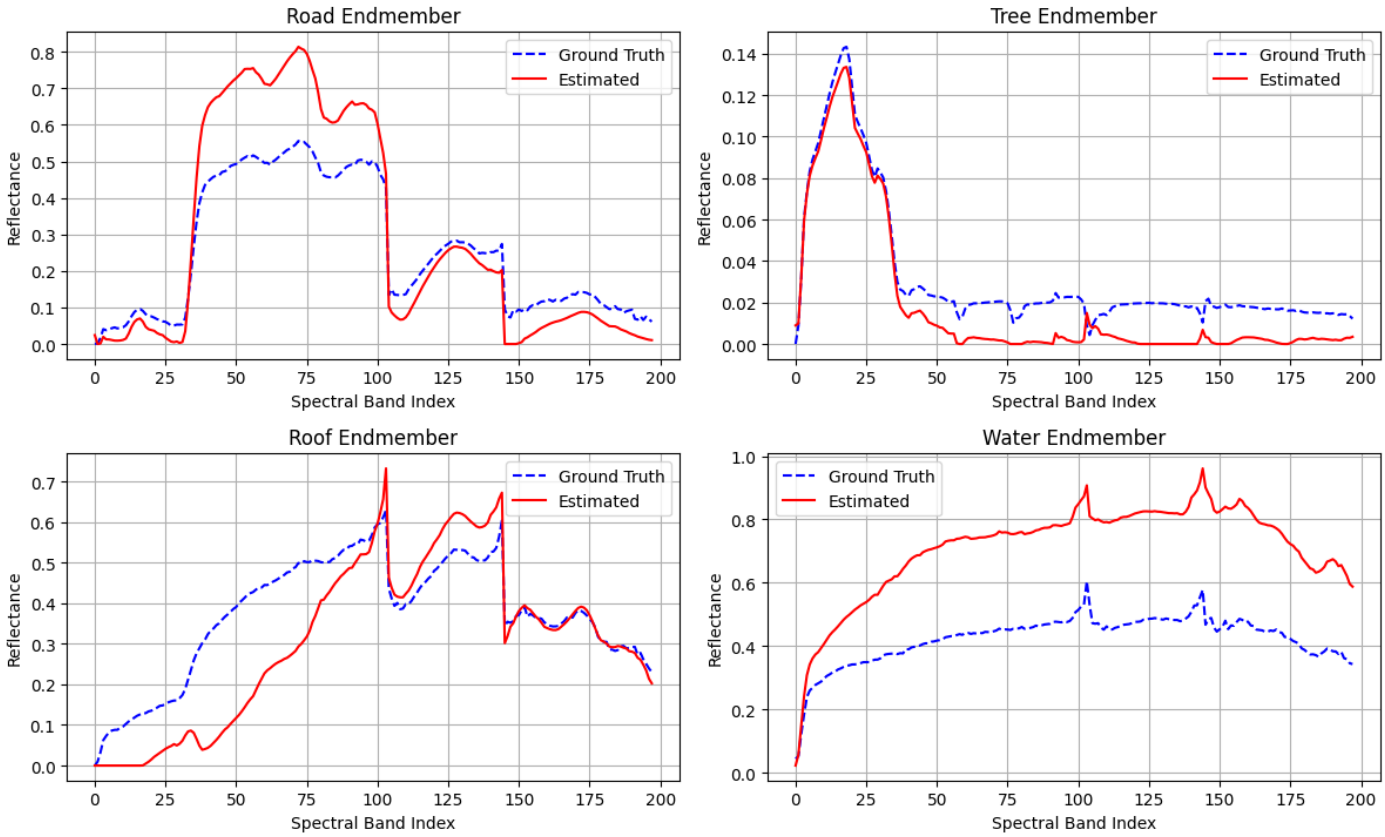

Figure 5. JasperRidge dataset - The visual comparison of the endmembers was obtained using the MiSiCNet Model. Red: the estimated endmembers; Blue: the ground truth. Endmembers.

4.3. Measurable metrics

To objectively assess the performance of MiSiCNet and UnDIP for hyperspectral unmixing for the input well-known HSI JasperRidge Image, three widely-used metrics were computed: Root Mean Squared Error (RMSE), Spectral Angle Distance (SAD), and Spectral Angle Mapper (SAM). The results for four representative classes, i.e., road, tree, roof, and water (four endmembers of JasperRidge image) are summarized in Table 3.

RMSE evaluates the reconstruction error between the estimated abundance maps and the ground truth abundances. Since the hyperspectral data were not normalized, the magnitude of RMSE is influenced by the scale and dynamic range of the abundance values for each class. UnDIP performs superior in 3 out of 4 classes (Road, Roof, and Water) with notably lower RMSE values. MiSiCNet performs better only for the Tree class, with a lower RMSE of 18.84 compared to 20.25 from UnDIP. This suggests MiSiCNet's advantage in modeling the abundance of spectrally subtle vegetation regions. SAD measures the angular deviation between estimated and true endmember vectors, assessing spectral similarity.UnDIP outperforms MiSiCNet across all four classes in terms of SAD. The significant SAD margin in the road and roof classes reflects UnDIP's capacity to model high-contrast materials with complex spectral variability. SAM evaluates the mean angular error over all pixels and is more robust to illumination differences and noise.UnDIP once again leads in 3 out of 4 classes.

MiSiCNet performs better only for water, achieving a lower SAM (2.98°) than UnDIP (4.91°), indicating better estimation of flat spectral regions such as water.

Table 3. Quantitative comparison of UnDIP and MiSiCNet models across four endmember classes using RMSE, SAD, and SAM metrics on the Jasper Ridge dataset.

Metric	Class	UnDIP	MiSiCNet	Better Model
RMSE	Road	15.91	17.45	UnDIP
	Tree	20.25	18.84	MiSiCNet
	Roof	13.46	18.36	UnDIP
	Water	10.90	13.48	UnDIP
SAD	Road	8.51	41.32	UnDIP
	Tree	14.21	18.49	UnDIP
	Roof	6.57	18.53	UnDIP
	Water	4.91	2.98	MiSiCNet
SAM	Road	8.51°	11.30°	UnDIP
	Tree	14.21°	18.46°	UnDIP
	Roof	6.57°	18.48°	UnDIP
	Water	4.91°	2.98°	MiSiCNet

The analysis shows that UnDIP consistently provides more accurate spectral unmixing, with lower SAD and SAM values across most materials, suggesting better spectral fidelity and angular consistency with ground truth. The RMSE values further reinforce UnDIP's advantage in abundance estimation and spatial unmixing accuracy. MiSiCNet shows strength in handling smooth spectra (e.g., tree, water) but tends to over-smooth or underestimate spectral features for high-contrast or structurally complex materials like roads and roof. The lower SAM for Water in MiSiCNet could be attributed to its architectural bias toward smoother representations, which work well on uniformly reflective surfaces like water bodies. Overall, UnDIP outperforms MiSiCNet in 9 out of the 12 metric-class combinations, proving to be a more reliable and versatile model for hyperspectral unmixing. It delivers higher spectral accuracy, better angular preservation, and lower abundance estimation error, making it highly effective for urban and natural scenes.

5. Conclusions

This paper presented a comparative evaluation of two advanced deep learning architectures (MiSiCNet and UnDIP) for hyperspectral image unmixing using the Jasper Ridge dataset. Both models were assessed regarding endmember spectral signature accuracy and abundance estimation performance, utilizing quantitative metrics such as RMSE, SAD, and SAM. The results reveal that UnDIP consistently outperforms MiSiCNet across a majority of the evaluated criteria. Specifically, UnDIP demonstrated superior accuracy in 3 out of 4 classes for both RMSE and SAM and outperformed MiSiCNet in all four classes for SAD. These metrics indicate that UnDIP more effectively reconstructs the abundance maps and better preserves the spectral characteristics of the true endmembers. The improved angular similarity (lower SAD and SAM) also emphasizes UnDIP's robustness in capturing spectral features under varying material complexity and noise levels.Visual inspection of the abundance maps further supports this finding. The abundance maps generated by UnDIP show higher spatial consistency and more accurate material localization, particularly in heterogeneous regions such as road networks and rooftop structures. While MiSiCNet performed reasonably well for classes with smoother or more homogeneous spectral profiles (e.g., tree and water), it exhibited limitations in preserving fine structural details. It showed higher spectral distortion in urban materials like roof and road. Endmember signature comparison also corroborates these findings. UnDIP-produced signatures were closer to the ground truth across the spectral range for most materials, indicating its ability to learn high-fidelity representations. MiSiCNet, in contrast, showed signs of over-smoothing and spectral deviation, especially in high-contrast spectral bands.

In conclusion, UnDIP is a more effective and generalizable architecture for hyperspectral unmixing, offering enhanced performance in spectral and spatial domains. Its ability to

accurately estimate endmembers and their abundances makes it particularly suitable for real-world remote sensing applications, such as environmental monitoring, land-cover classification, and urban planning. Future work may further explore hybrid architectures that combine UnDIP's robustness with MiSiCNet's regularization strengths to boost performance across diverse datasets.

Funding: This research received no external funding.

Data Availability Statement: The hyperspectral image data used in this study (Jasper Ridge dataset) is publicly available from (https://lesun.weebly.com/hyperspectral-data-set.html).

Acknowledgments: I would like to express my sincere gratitude to Prof. Ajay Kr. Ray, Supervisor, for his invaluable guidance, continuous support, and insightful feedback throughout the course of this research. I am deeply thankful to Dr. Samiran Das, Co-Guide, for his technical advice, constructive suggestions, and encouragement, which have been instrumental in the successful completion of this work. I also acknowledge the administrative and technical assistance provided by our institution, and the resources made available for conducting this study.

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

References

[1] J. M. Bioucas-Dias *et al.*, "Hyperspectral Unmixing Overview: Geometrical, Statistical, and Sparse Regression-Based Approaches," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2012, doi: 10.1109/JSTARS.2012.2194696.

[2] N. Dobigeon, J.-Y. Tourneret, C. Richard, J. C. M. Bermudez, S. McLaughlin, and A. O. Hero, "Nonlinear Unmixing of Hyperspectral Images: Models and Algorithms," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 82–94, Jan. 2014, doi: 10.1109/MSP.2013.2279274.

[3] B. Rasti *et al.*, "Feature Extraction for Hyperspectral Imagery: The Evolution From Shallow to Deep: Overview and Toolbox," *IEEE Geosci. Remote Sens. Mag.*, vol. 8, no. 4, pp. 60–88, Dec. 2020, doi: 10.1109/MGRS.2020.2979764.

[4] D. C. Heinz and Chein-I-Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 3, pp. 529–545, Mar. 2001, doi: 10.1109/36.911111.

[5] R. Heylen, D. Burazerovic, and P. Scheunders, "Fully Constrained Least Squares Spectral Unmixing by Simplex Projection," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4112–4122, Nov. 2011, doi: 10.1109/TGRS.2011.2155070.

[6] Tsung-Han Chan, Chong-Yung Chi, Yu-Min Huang, and Wing-Kin Ma, "A Convex Analysis-Based Minimum-Volume Enclosing Simplex Algorithm for Hyperspectral Unmixing," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4418–4432, Nov. 2009, doi: 10.1109/TSP.2009.2025802.

[7] J. M. P. Nascimento and J. M. B. Dias, "Vertex Component Analysis: A~Fast Algorithm to Extract Endmembers Spectra from Hyperspectral Data," in *Pattern Recognition and Image Analysis*, 2003, pp. 626–635. doi: 10.1007/978-3-540-44871-6_73.

[8] J. W. Boardman, F. A. Kruse, and R. O. Green, "Mapping Target Signatures via Partial Unmixing of AVIRIS Data: In Summaries," in *Fifth JPL Airborne Earth Science Workshop*, 1996, pp. 23–26.

[9] M. E. Winter, "N-FINDR: an algorithm for fast autonomous spectral endmember determination in hyperspectral data," in *Proceedings Volume 3753, Imaging Spectrometry V*, Oct. 1999, pp. 266–275. doi: 10.1117/12.366289.

[10] W. E. Full, R. Ehrlich, and J. E. Klovan, "EXTENDED QMODELPobjective definition of external end members in the analysis of mixtures," *J. Int. Assoc. Math. Geol.*, vol. 13, no. 4, pp. 331–344, Aug. 1981, doi: 10.1007/BF01031518.

[11] M. D. Craig, "Minimum-volume transforms for remotely sensed data," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 3, pp. 542–552, May 1994, doi: 10.1109/36.297973.

[12] C.-H. Lin, R. Wu, W.-K. Ma, C.-Y. Chi, and Y. Wang, "Maximum Volume Inscribed Ellipsoid: A New Simplex-Structured Matrix Factorization Framework via Facet Enumeration and Convex Optimization," *ArXiv*. Aug. 09, 2017. doi: 10.48550/arXiv.1708.02883.

[13] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Collaborative nonnegative matrix factorization for remotely sensed hyperspectral unmixing," in *2012 IEEE International Geoscience and Remote Sensing Symposium*, Jul. 2012, pp. 3078–3081. doi: 10.1109/IGARSS.2012.6350775.

[14] N. Dobigeon, S. Moussaoui, M. Coulon, J.-Y. Tourneret, and A. O. Hero, "Joint Bayesian Endmember Extraction and Linear Unmixing for Hyperspectral Imagery," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4355–4368, Nov. 2009, doi: 10.1109/TSP.2009.2025797.

[15] L. Miao and H. Qi, "Endmember Extraction From Highly Mixed Data Using Minimum Volume Constrained Nonnegative Matrix Factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 765–777, Mar. 2007, doi: 10.1109/TGRS.2006.888466.

[16] J. Li, J. M. Bioucas-Dias, A. Plaza, and L. Liu, "Robust Collaborative Nonnegative Matrix Factorization for Hyperspectral Unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6076–6090, Oct. 2016, doi: 10.1109/TGRS.2016.2580702.

[17] M. Berman, H. Kiiveri, R. Lagerstrom, A. Ernst, R. Dunne, and J. F. Huntington, "ICE: a statistical approach to identifying endmembers in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 10, pp. 2085–2095, Oct. 2004, doi: 10.1109/TGRS.2004.835299.

[18] L. Zhuang, C.-H. Lin, M. A. T. Figueiredo, and J. M. Bioucas-Dias, "Regularization Parameter Selection in Minimum Volume Hyperspectral Unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 9858–9877, Dec. 2019, doi: 10.1109/TGRS.2019.2929776.

[19] J. Li and J. M. Bioucas-Dias, "Minimum Volume Simplex Analysis: A Fast Algorithm to Unmix Hyperspectral Data," in *IGARSS 2008 - 2008 IEEE International Geoscience and Remote Sensing Symposium*, 2008, pp. III-250-III–253. doi: 10.1109/IGARSS.2008.4779330.

[20] J. M. Bioucas-Dias, "A variable splitting augmented Lagrangian approach to linear spectral unmixing," in *2009 First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, Aug. 2009, pp. 1–4. doi: 10.1109/WHISPERS.2009.5289072.

[21] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Sparse Unmixing of Hyperspectral Data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2014–2039, Jun. 2011, doi: 10.1109/TGRS.2010.2098413.

[22] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Total Variation Spatial Regularization for Sparse Hyperspectral Unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4484–4502, Nov. 2012, doi: 10.1109/TGRS.2012.2191590.

[23] B. Rasti and B. Koirala, "SUnCNN: Sparse Unmixing Using Unsupervised Convolutional Neural Network," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022, doi: 10.1109/LGRS.2021.3100992.

[24] L. Meier, S. Van De Geer, and P. Bühlmann, "The Group Lasso for Logistic Regression," *J. R. Stat. Soc. Ser. B Stat. Methodol.*, vol. 70, no. 1, pp. 53–71, Feb. 2008, doi: 10.1111/j.1467-9868.2007.00627.x.

[25] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Collaborative Sparse Regression for Hyperspectral Unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 341–354, Jan. 2014, doi: 10.1109/TGRS.2013.2240001.

[26] L. Drumetz, T. R. Meyer, J. Chanussot, A. L. Bertozzi, and C. Jutten, "Hyperspectral Image Unmixing With Endmember Bundles and Group Sparsity Inducing Mixed Norms," *IEEE Trans. Image Process.*, vol. 28, no. 7, pp. 3435–3450, Jul. 2019, doi: 10.1109/TIP.2019.2897254.

[27] L. Qi, J. Li, Y. Wang, Y. Huang, and X. Gao, "Spectral–Spatial-Weighted Multiview Collaborative Sparse Unmixing for Hyperspectral Images," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 12, pp. 8766–8779, Dec. 2020, doi: 10.1109/TGRS.2020.2990476.

[28] S. Zhang *et al.*, "Spectral-Spatial Hyperspectral Unmixing Using Nonnegative Matrix Factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–13, 2022, doi: 10.1109/TGRS.2021.3074364.

[29] M. M. Elkholy, M. Mostafa, H. M. Ebied, and M. F. Tolba, "Hyperspectral unmixing using deep convolutional autoencoder," *Int. J. Remote Sens.*, vol. 41, no. 12, pp. 4799–4819, Jun. 2020, doi: 10.1080/01431161.2020.1724346.

[30] Y. Su *et al.*, "DAAN: A Deep Autoencoder-Based Augmented Network for Blind Multilinear Hyperspectral Unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, pp. 1–15, 2024, doi: 10.1109/TGRS.2024.3381632.

[31] X. Xu, X. Song, T. Li, Z. Shi, and B. Pan, "Deep Autoencoder for Hyperspectral Unmixing via Global-Local Smoothing," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–16, 2022, doi: 10.1109/TGRS.2022.3152782.

[32] Y. Su, A. Marinoni, J. Li, J. Plaza, and P. Gamba, "Stacked Nonnegative Sparse Autoencoders for Robust Hyperspectral Unmixing," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 9, pp. 1427–1431, Sep. 2018, doi: 10.1109/LGRS.2018.2841400.

[33] Y. Su, J. Li, A. Plaza, A. Marinoni, P. Gamba, and S. Chakravortty, "DAEN: Deep Autoencoder Networks for Hyperspectral Unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4309–4321, Jul. 2019, doi: 10.1109/TGRS.2018.2890633.

[34] R. A. Borsoi, T. Imbiriba, and J. C. M. Bermudez, "Deep Generative Endmember Modeling: An Application to Unsupervised Spectral Unmixing," *IEEE Trans. Comput. Imaging*, vol. 6, pp. 374–384, 2020, doi: 10.1109/TCI.2019.2948726.

[35] Y. Qu and H. Qi, "uDAS: An Untied Denoising Autoencoder With Sparsity for Spectral Unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 3, pp. 1698–1712, Mar. 2019, doi: 10.1109/TGRS.2018.2868690.

[36] S. Ozkan, B. Kaya, and G. B. Akar, "EndNet: Sparse AutoEncoder Network for Endmember Extraction and Hyperspectral Unmixing," *ArXiv*. Aug. 06, 2017. doi: 10.48550/arXiv.1708.01894.

[37] B. Palsson, J. Sigurdsson, J. R. Sveinsson, and M. O. Ulfarsson, "Hyperspectral Unmixing Using a Neural Network Autoencoder," *IEEE Access*, vol. 6, pp. 25646–25656, 2018, doi: 10.1109/ACCESS.2018.2818280.

[38] Z. Hua, X. Li, Q. Qiu, and L. Zhao, "Autoencoder Network for Hyperspectral Unmixing With Adaptive Abundance Smoothing," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 9, pp. 1640–1644, Sep. 2021, doi: 10.1109/LGRS.2020.3005999.

[39] X. Zhang, Y. Sun, J. Zhang, P. Wu, and L. Jiao, "Hyperspectral Unmixing via Deep Convolutional Neural Networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 11, pp. 1755–1759, Nov. 2018, doi: 10.1109/LGRS.2018.2857804.

[40] B. Palsson, J. R. Sveinsson, and M. O. Ulfarsson, "Spectral-Spatial Hyperspectral Unmixing Using Multitask Learning," *IEEE Access*, vol. 7, pp. 148861–148872, 2019, doi: 10.1109/ACCESS.2019.2944072.

[41] F. Khajehrayeni and H. Ghassemian, "Hyperspectral Unmixing Using Deep Convolutional Autoencoders in a Supervised Scenario," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 13, pp. 567–576, 2020, doi: 10.1109/JSTARS.2020.2966512.

[42] B. Palsson, M. O. Ulfarsson, and J. R. Sveinsson, "Convolutional Autoencoder for Spectral–Spatial Hyperspectral Unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 1, pp. 535–549, Jan. 2021, doi: 10.1109/TGRS.2020.2992743.

[43] B. Rasti, B. Koirala, P. Scheunders, and P. Ghamisi, "UnDIP: Hyperspectral Unmixing Using Deep Image Prior," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–15, 2022, doi: 10.1109/TGRS.2021.3067802.

[44] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep Image Prior," *Int. J. Comput. Vis.*, vol. 128, no. 7, pp. 1867–1888, Jul. 2020, doi: 10.1007/s11263-020-01303-4.

[45] L. Gao, Z. Han, D. Hong, B. Zhang, and J. Chanussot, "CyCU-Net: Cycle-Consistency Unmixing Network by Learning Cascaded Autoencoders," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–14, 2022, doi: 10.1109/TGRS.2021.3064958.

[46] Chia-Hsiang Lin, Wing-Kin Ma, Wei-Chiang Li, Chong-Yung Chi, and A. Ambikapathi, "Identifiability of the Simplex Volume Minimization Criterion for Blind Hyperspectral Unmixing: The No-Pure-Pixel Case," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 10, pp. 5530–5546, Oct. 2015, doi: 10.1109/TGRS.2015.2424719.

[47] V. Paura and V. Marcinkevičius, "Benchmark for Hyperspectral Unmixing Algorithm Evaluation," *Informatica*, pp. 285–315, Jun. 2023, doi: 10.15388/23-INFOR522.

[48] M. S. Karoui and K. Djerriri, "A New Unmixing-Based Approach for Shadow Correction of Hyperspectral Remote Sensing Data," in *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, Jul. 2018, pp. 2725–2728. doi: 10.1109/IGARSS.2018.8517483.

[49] R. Feng, Y. Zhong, and L. Zhang, "Adaptive Spatial Regularization Sparse Unmixing Strategy Based on Joint MAP for Hyperspectral Remote Sensing Imagery," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 9, no. 12, pp. 5791–5805, Dec. 2016, doi: 10.1109/JSTARS.2016.2570947.

[50] Z. Li, J. Chen, and S. Rahardja, "Graph Construction for Hyperspectral Data Unmixing," in *Hyperspectral Imaging in Agriculture, Food and Environment*, InTech, 2018. doi: 10.5772/intechopen.73158.

[51] J. Chi and M. M. Crawford, "Spectral Unmixing-Based Crop Residue Estimation Using Hyperspectral Remote Sensing Data: A Case Study at Purdue University," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 7, no. 6, pp. 2531–2539, Jun. 2014, doi: 10.1109/JSTARS.2014.2319585.

[52] Z. Li, J. Chen, and S. Rahardja, "Superpixel construction for hyperspectral unmixing," in *2018 26th European Signal Processing Conference (EUSIPCO)*, Sep. 2018, pp. 647–651. doi: 10.23919/EUSIPCO.2018.8553223.

[53] S. Jia, Z. Ji, and Y. Qian, "Band selection based hyperspectral unmixing," in *2009 IEEE International Workshop on Imaging Systems and Techniques*, May 2009, pp. 303–306. doi: 10.1109/IST.2009.5071654.