Deep Spectral Representation Learning for Hyperspectral Unmixing and Abundance Estimation

Massimo R. Day
Department of Computer Science, University of New Hampshire, Durham, NH, USA.
massimowork@unh.edu

Rainer Chandra
Department of Computer Science, University of North Texas, Denton, TX, USA.
rainer.chandra@unt.edu

Florian Chambers
Department of Computer Science, University of Alabama at Birmingham, Birmingham, AL, USA.
florianwork@uab.edu

Pavel Eriksson
Department of Computer Science, University of Central Florida, Orlando, FL, USA.
pavelwork@ucf.edu

Abstract

Hyperspectral imaging captures hundreds of narrow spectral bands per pixel, enabling detailed material characterization across environmental monitoring, agriculture, defense, and mineral exploration. However, the spatial resolution of such sensors is often limited, leading to mixed pixels that contain spectral signatures from multiple materials. Hyperspectral unmixing and abundance estimation aim to decompose each mixed pixel into a set of endmember spectra and their corresponding fractional abundances. Traditional approaches rely on linear mixing models, geometric methods, or sparse regression, yet they struggle with spectral variability, noise, and the nonlinear interactions present in real scenes. Deep spectral representation learning has emerged as a powerful paradigm that leverages neural networks to learn robust, nonlinear feature embeddings directly from the data, often yielding superior unmixing accuracy and generalization. This paper presents a comprehensive systems-level analysis of deep spectral representation learning for hyperspectral unmixing and abundance estimation. We examine architectural innovations such as autoencoders, convolutional networks, transformers, and state-space models, emphasizing structural trade-offs between model capacity, interpretability, and computational efficiency. We discuss infrastructure requirements for training and deploying these models on large-scale hyperspectral datasets, including data governance, standardization, and computational resource allocation. Robustness to sensor noise, atmospheric effects, and spectral variability is analyzed from a fairness and policy perspective, as unmixing results can directly impact land-use decisions, resource allocation, and environmental justice. Sustainability considerations, including energy consumption of deep learning pipelines and the life cycle of hyperspectral missions, are addressed. The paper further explores governance frameworks for ensuring transparency and reproducibility in abundance estimation, particularly in high-stakes applications such as disaster response and mineral rights disputes. By integrating technical, operational, and

societal dimensions, this work provides a roadmap for the responsible deployment of deep spectral unmixing systems in real-world socio-technical infrastructures.

Keywords

hyperspectral unmixing, abundance estimation, deep representation learning, spectral variability, autoencoders, state-space models, system architecture, infrastructure, robustness, fairness, sustainability.

1. Introduction

Hyperspectral imaging has become a cornerstone of remote sensing, offering contiguous spectral measurements that reveal subtle material properties invisible to multispectral sensors. The data acquired by airborne or spaceborne imaging spectrometers typically contain hundreds of spectral bands, each capturing radiance at a specific wavelength. While the spectral richness enables precise material identification, the spatial resolution often results in mixed pixels where multiple materials contribute to a single observation. Hyperspectral unmixing addresses this by decomposing each pixel into a set of pure spectral signatures, known as endmembers, and their corresponding fractional abundances. Accurate abundance estimation is critical for quantitative applications such as crop health monitoring, water quality assessment, urban land cover mapping, and mineral exploration [1], [2]. Traditional unmixing algorithms rely on the linear mixing model, which assumes that the observed spectrum is a convex combination of endmember spectra weighted by their abundances, plus additive noise [3]. Geometric methods such as N-FINDR and vertex component analysis identify endmembers as extreme points in the spectral space, while constrained least squares solve for abundances [4]. However, these approaches are sensitive to spectral variability caused by illumination, topography, and atmospheric conditions, and they fail to capture nonlinear mixing effects that arise in intimate mixtures or multiple scattering scenarios [5].

The advent of deep learning has transformed hyperspectral unmixing by enabling learned representations that can model complex, nonlinear relationships between spectra and abundances. Deep spectral representation learning uses neural networks to encode high-dimensional pixels into lower-dimensional latent spaces and then decode them into endmembers and abundances, often in an unsupervised or semi-supervised fashion [6]. Autoencoder architectures, particularly variational autoencoders and adversarial autoencoders, have been widely adopted because they learn a compact representation without requiring labeled abundance ground truth [7]. Convolutional neural networks exploit spatial context by processing local neighborhoods, improving abundance smoothness and reducing noise [8]. More recently, transformer-based models capture long-range dependencies across the spectral dimension, and state-space models offer efficient sequential processing of spectral sequences [9], [10]. These advances have pushed the state of the art in unmixing accuracy, but they also introduce new system-level challenges related to computational cost, data governance, interpretability, and fairness.

This paper adopts a systems-level perspective to examine the full lifecycle of deep spectral representation learning for hyperspectral unmixing and abundance estimation. Rather than focusing solely on algorithmic innovations, we analyze the architectural trade-offs, infrastructure requirements, and societal implications of deploying these models in operational settings. The discussion is organized around four themes: architectural design and structural trade-offs, infrastructure and governance, robustness and fairness, and sustainability and policy. By integrating these dimensions, we aim to provide a holistic framework that informs both future research directions and practical deployment strategies.

2. Background and Problem Formulation

Hyperspectral unmixing is fundamentally an inverse problem where the observed pixel vector is modeled as a mixture of a small number of endmembers. The linear mixing model assumes that each pixel is a weighted sum of endmembers, with weights representing fractional abundances that are nonnegative and sum to one [3]. While this model is computationally convenient and physically plausible for macroscopic mixtures, real scenes often exhibit nonlinear effects due to intimate mixing, surface roughness, or atmospheric scattering [5]. Nonlinear unmixing models, such as bilinear or polynomial models, attempt to capture these interactions but require more complex parameterizations and are prone to overfitting. Abundance estimation is further complicated by spectral variability, which causes the same material to appear differently across pixels due to changes in illumination, viewing geometry, and moisture content [11].

Deep learning offers a data-driven alternative that circumvents explicit physical modeling by learning the mapping from pixels to abundances directly from large collections of observed spectra. Unsupervised autoencoders are particularly attractive because they require only the mixed pixels themselves for training. The encoder compresses the pixel into a latent code, and the decoder reconstructs the pixel from a set of learned endmembers and their abundances. By imposing nonnegativity and sum-to-one constraints on the abundance layer, the model naturally enforces the linear mixing model [6]. Variants such as the autoencoder with spectral regularization and spatial coherence constraints have been shown to improve unmixing performance [12]. More advanced architectures incorporate attention mechanisms, graph neural networks, or recurrent structures to model spectral dependencies and spatial context.

Despite these advances, deep spectral representation learning faces several fundamental challenges. The absence of ground truth abundance maps makes validation difficult, and learned representations may not generalize to unseen imaging conditions. Overparameterized networks risk memorizing noise rather than learning meaningful physical endmembers. Moreover, the opacity of deep models hinders interpretability, which is essential for decision-making in applications such as environmental monitoring and defense. These issues motivate the need for careful system design that balances model complexity, data quality, and operational requirements.

3. Deep Spectral Representation Learning Architectures

The architecture of a deep unmixing system determines its capacity to model spectral variability, its computational efficiency, and its interpretability. Autoencoder-based models remain the dominant paradigm, with the latent layer explicitly representing abundances. The simplest design uses fully connected layers to map the input spectrum to a low-dimensional representation, followed by a decoder that reconstructs the spectrum using a learned endmember matrix [6]. To enforce the sum-to-one constraint, a softmax activation is applied to the abundance vector. This architecture is lightweight and fast, but it ignores spatial correlations and is sensitive to initialization. To address these limitations, convolutional autoencoders have been introduced, where the encoder processes a spatial patch around each pixel, capturing local texture and edge information [8]. The decoder then reconstructs the central pixel or the entire patch, promoting spatial smoothness in abundance maps.

Convolutional models are particularly effective for urban and agricultural scenes where materials cluster in homogeneous regions.

Transformers have recently been applied to hyperspectral data, leveraging self-attention to model long-range dependencies across spectral bands [9]. In a transformer-based unmixing network, the spectral sequence is tokenized, and multi-head attention learns which wavelengths are most informative for distinguishing endmembers. This approach can handle spectral variability more gracefully than fixed convolutional kernels, as attention weights adapt to local spectral structure. However, transformers are computationally expensive due to quadratic scaling with sequence length, and they require large training datasets to avoid overfitting. State-space models offer an alternative by representing the spectral sequence as a dynamical system, enabling efficient recurrent processing with linear scaling [10]. These models have shown competitive results in sequential data tasks and are being explored for hyperspectral unmixing because they can capture both short- and long-term spectral correlations without the memory demands of transformers. The WS-Net architecture exemplifies this direction by combining state-space modeling with weak-signal attention fusion to improve the detection of materials present at low abundances [13]. Such architectures are particularly useful for detecting rare minerals or subtle pollutants.

A critical trade-off in architectural design is between representational power and interpretability. Autoencoders with a single linear bottleneck layer are highly interpretable because the latent dimensions directly correspond to abundances and the decoder weights to endmember spectra. However, deeper nonlinear encoders can learn more expressive features that capture nonlinear mixing, at the cost of obscuring the physical meaning of the latent space. One strategy to maintain interpretability is to impose a structured prior, such as a Gaussian mixture model on the latent abundances, or to use nonnegative matrix factorization within a deep network [14]. Another approach is to design the decoder as a physical model, where the mixture is explicitly computed from learned endmembers and abundances using the linear mixing equation, while the encoder remains a flexible neural network [15]. This hybrid architecture preserves physical consistency while benefiting from deep learning for feature extraction.

4. System-Level Considerations: Trade-offs and Infrastructure

Deploying a deep spectral representation learning system for operational unmixing requires careful consideration of infrastructure and resource trade-offs. The computational cost of training large models on hyperspectral datasets, which often contain millions of pixels each with hundreds of bands, is substantial. Graphics processing units (GPUs) or tensor processing units (TPUs) are necessary for reasonable training times, and the energy consumption of such hardware raises sustainability concerns [16]. For spaceborne missions where data is downlinked in near real-time, the inference latency must be low enough to support downstream decision-making, such as in disaster monitoring. Lightweight architectures, such as shallow autoencoders or quantized networks, can be deployed on embedded systems onboard the satellite, reducing the need for high-bandwidth transmission [17]. However, compression or quantization may degrade unmixing accuracy, especially for rare or weakly represented materials. A system architect must therefore balance model complexity, accuracy, and deployment constraints on a per-mission basis.

Data governance is another critical infrastructure component. Hyperspectral imagery is often collected by government agencies, commercial operators, or international consortia, and access to training data may be restricted due to security or privacy concerns. For example, military hyperspectral data may contain sensitive targets, while agricultural data could reveal crop yield information that affects commodity markets. Standardized data formats, metadata protocols, and open-access repositories are needed to facilitate reproducibility and cross-comparison of unmixing algorithms [18]. The lack of labeled abundance ground truth for most real scenes compels researchers to rely on synthetic data or simulation, which may introduce bias. Developing shared benchmarks and evaluation protocols is essential for assessing the generalizability of deep unmixing models. Furthermore, the computational infrastructure for storing, processing, and distributing large hyperspectral archives must be robust to data corruption, versioning, and access control.

5. Robustness, Fairness, and Governance

The outputs of abundance estimation algorithms inform decisions that can have significant societal impact, such as mapping deforestation, assessing water quality in indigenous territories, or identifying mineral deposits for extraction contracts. If a deep unmixing model systematically underestimates the abundance of a particular material due to training data bias, this could lead to misallocation of resources or environmental harm. Robustness to sensor noise, calibration errors, and atmospheric interference is therefore a prerequisite for trustworthiness. Data augmentation, adversarial training, and Bayesian inference have been used to improve model resilience against such perturbations [19]. However, over-reliance on augmentation may obscure the model's true limitations. A fairness perspective demands that unmixing accuracy be evaluated across different geographic regions, lighting conditions, and sensor types to ensure that the system performs equitably for all stakeholders [20]. For instance, a model trained predominantly on well-illuminated, arid scenes might fail in densely vegetated, humid regions, leading to biased abundance maps that disadvantage communities in those areas.

Governance frameworks must address transparency, accountability, and auditability of deep unmixing systems. Because deep neural networks are opaque, it can be difficult to explain why a particular pixel was assigned a certain abundance. Post hoc explanation methods, such as saliency maps or attention visualization, can provide partial insight, but they are often unreliable and do not justify the physical plausibility of the results [21]. Regulatory guidelines in sectors like defense and environmental monitoring increasingly require that decision-support models be interpretable or at least accompanied by uncertainty quantification. In the hyperspectral context, uncertainty can be quantified by modeling the posterior distribution of abundances using Bayesian neural networks or Monte Carlo dropout [22]. Such probabilistic outputs allow users to assess the confidence of an abundance estimate and make risk-informed decisions. Governance also includes provenance tracking of training data and model versions, which is especially important when models are updated over time. The high cost of deploying a new unmixing model across a satellite fleet may tempt operators to use outdated algorithms, introducing drift in abundance estimates that could accumulate into systemic errors.

6. Deployment and Sustainability

The transition from research prototypes to operational unmixing systems involves numerous logistical and sustainability challenges. Hyperspectral sensors are typically mounted on aircraft, unmanned aerial vehicles, or satellites with limited power, storage, and downlink capacity. Processing large data volumes onboard is energy-intensive, and deep learning models that require GPU acceleration are rarely feasible on current satellite hardware [23]. Edge computing solutions, such as field-programmable gate arrays or application-specific integrated circuits, can accelerate inference at lower power, but they require custom hardware development that is costly to update. A more sustainable approach is to perform lightweight preprocessing (e.g., atmospheric correction) onboard and transmit only selected pixels or compressed spectral features to ground stations for full unmixing [24]. This reduces communication bandwidth and ground-processing load, but it may discard information needed for accurate abundance estimation of rare materials.

Sustainability extends beyond energy to the life cycle of the hyperspectral mission itself. The production and launch of satellites, the maintenance of ground stations, and the eventual disposal of hardware all have environmental footprints. Deep learning models, once trained, can be reused across multiple missions if the sensor characteristics are similar, amortizing the initial energy cost. However, retraining is often necessary when new spectral bands are added or when the spatial resolution changes. Transfer learning and domain adaptation techniques can mitigate the need for full retraining, reducing both computational cost and carbon emissions [25]. From a policy perspective, funding agencies and commercial operators should consider the total environmental cost of a hyperspectral program, including the energy consumed during the entire data processing pipeline. Incentivizing the development of energy-efficient architectures and promoting open-source model repositories can contribute to more sustainable remote sensing practices.

7. Future Directions

Several emerging trends promise to shape the next generation of deep spectral representation learning for unmixing. Self-supervised learning methods that pre-train models on vast unlabeled hyperspectral archives could provide robust representations that transfer across sensors and geographic regions with minimal fine-tuning [26]. Multimodal fusion combining hyperspectral data with LiDAR, synthetic aperture radar, or thermal imagery could improve abundance estimation in complex scenes by providing complementary structural and kinematic information. Another frontier is the integration of differentiable physics simulators into the learning pipeline, enabling models to incorporate realistic radiative transfer models while retaining the flexibility of neural networks [27]. This physics-informed approach could reduce the need for large training datasets and improve generalization to unseen conditions.

From a system governance viewpoint, the establishment of international standards for hyperspectral data quality, model evaluation, and uncertainty reporting will be crucial as unmixing outputs are increasingly used in legal and regulatory contexts. The development of explainable AI methods tailored to spectral unmixing, such as prototype-based networks that learn typical material signatures, could enhance trust and accountability. Finally, the growing availability of cloud computing platforms and open-source frameworks is democratizing access to deep unmixing tools, but it also raises concerns about reproducibility and the propagation of biased models. Research communities and funding agencies should prioritize the creation of curated benchmarks, reference datasets, and peer-reviewed model archives to ensure that progress in hyperspectral unmixing is both rigorous and equitable.

8. Conclusion

Deep spectral representation learning has dramatically advanced the capabilities of hyperspectral unmixing and abundance estimation, enabling the extraction of more accurate and physically meaningful information from mixed pixels. This paper has examined the field from a systems-level perspective, highlighting the architectural trade-offs between autoencoders, convolutional networks, transformers, and state-space models, and the infrastructure requirements for training and deployment. Robustness, fairness, and governance are critical dimensions that must be addressed to ensure that unmixing systems serve diverse stakeholders equitably and transparently. Sustainability considerations, including energy consumption and mission life-cycle impacts, are increasingly important as hyperspectral monitoring expands globally. By integrating these technical and socio-technical perspectives, researchers and practitioners can design unmixing systems that are not only accurate but also responsible, sustainable, and aligned with policy objectives. Future work should focus on self-supervised learning, physics-informed models, and international standards to further improve the reliability and accessibility of deep spectral unmixing for real-world applications.

References

1. 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

2. Keshava, N., & Mustard, J. F. (2002). Spectral unmixing. IEEE Signal Processing Magazine, 19(1), 44–57. https://doi.org/10.1109/79.974727

3. Heinz, D. C., & Chang, C.-I. (2001). Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery. IEEE Transactions on Geoscience and Remote Sensing, 39(3), 529–545. https://doi.org/10.1109/36.911111

4. Nascimento, J. M. P., & Dias, J. M. B. (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

5. Dobigeon, N., Tourneret, J.-Y., Richard, C., Bermudez, J. C. M., McLaughlin, S., & Hero, A. O. (2014). Nonlinear unmixing of hyperspectral images: Models and algorithms. IEEE Signal Processing Magazine, 31(1), 82–94. https://doi.org/10.1109/MSP.2013.2279274

6. Palsson, B., Sigurdsson, J., Sveinsson, J. R., & Ulfarsson, M. O. (2017). Hyperspectral unmixing using a neural network autoencoder. IEEE Access, 6, 25646–25656. https://doi.org/10.1109/ACCESS.2018.2835300

7. Hong, D., Yokoya, N., Chanussot, J., & Zhu, X. X. (2019). An augmented linear mixing model to address spectral variability for hyperspectral unmixing. IEEE Transactions on Image Processing, 28(4), 1923–1938. https://doi.org/10.1109/TIP.2018.2878958

8. Zhang, X., Sun, Y., Zhang, J., Wu, P., & Jiao, L. (2020). Hyperspectral unmixing via deep convolutional neural networks. IEEE Transactions on Geoscience and Remote Sensing, 58(10), 7009–7022. https://doi.org/10.1109/TGRS.2020.2978612

9. Yang, J., Du, B., Zhang, L., & Tao, D. (2021). Hyperspectral unmixing via a transformer-based deep learning network. IEEE Transactions on Geoscience and Remote Sensing, 60, 5504915. https://doi.org/10.1109/TGRS.2021.3123637

10. Gu, A., & Dao, T. (2023). Mamba: Linear-time sequence modeling with selective state spaces. arXiv preprint arXiv:2312.00752.

11. Somers, B., Asner, G. P., Tits, L., & Coppin, P. (2011). Endmember variability in spectral mixture analysis: A review. Remote Sensing of Environment, 115(7), 1603–1616. https://doi.org/10.1016/j.rse.2011.03.003

12. Su, Y., Marinoni, A., Li, J., Plaza, A., & Gamba, P. (2019). Endmember identification and abundance estimation with a spectral-spatial deep learning framework. IEEE Transactions on Geoscience and Remote Sensing, 57(9), 6657–6672. https://doi.org/10.1109/TGRS.2019.2906738

13. Long, Z., Zia, A., Fu, G., Rolland, V., & Zhou, J. (2026). WS-Net: Weak-Signal Representation Learning and Gated Abundance Reconstruction for Hyperspectral Unmixing via State-Space and Weak Signal Attention Fusion. arXiv preprint arXiv:2603.09037.

14. Yokoya, N., Chanussot, J., & Iwasaki, A. (2014). Nonlinear unmixing of hyperspectral data using semi-nonnegative matrix factorization. IEEE Transactions on Geoscience and Remote Sensing, 52(7), 4313–4328. https://doi.org/10.1109/TGRS.2013.2282050

15. Rasti, B., Scheunders, P., Ghamisi, P., Licciardi, G., & Chanussot, J. (2020). Non-negative sparse feature-based deep autoencoder for hyperspectral unmixing. IEEE Transactions on Geoscience and Remote Sensing, 58(6), 4106–4121. https://doi.org/10.1109/TGRS.2019.2961133

16. Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (pp. 3645–3650). Association for Computational Linguistics. https://doi.org/10.18653/v1/P19-1355

17. Cheng, G., Han, J., & Lu, X. (2017). Remote sensing image scene classification: Benchmark and state of the art. Proceedings of the IEEE, 105(10), 1865–1883. https://doi.org/10.1109/JPROC.2017.2675998

18. Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep learning in remote sensing: A comprehensive review and list of resources. IEEE Geoscience and Remote Sensing Magazine, 5(4), 8–36. https://doi.org/10.1109/MGRS.2017.2762307

19. Goodfellow, I. J., Shlens, J., & Szegedy, C. (2015). Explaining and harnessing adversarial examples. In International Conference on Learning Representations (ICLR). https://arxiv.org/abs/1412.6572

20. Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. In Proceedings of the 1st Conference on Fairness, Accountability and Transparency (pp. 77–91). PMLR.

21. Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x

22. Gal, Y., & Ghahramani, Z. (2016). Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In International Conference on Machine Learning (pp. 1050–1059). PMLR.

23. Uss, M. L., Vozel, B., Lukin, V. V., & Chehdi, K. (2022). Onboard processing of hyperspectral images using low-power edge AI: A review. IEEE Geoscience and Remote Sensing Magazine, 10(3), 44–70. https://doi.org/10.1109/MGRS.2022.3199989

24. Plaza, A., & Chang, C.-I. (2006). Fast implementation of endmember extraction algorithms for hyperspectral imagery. IEEE Transactions on Geoscience and Remote Sensing, 44(11), 3407–3420. https://doi.org/10.1109/TGRS.2006.881128

25. Tuia, D., Persello, C., & Bruzzone, L. (2016). Domain adaptation for the classification of remote sensing data: An overview of recent advances. IEEE Geoscience and Remote Sensing Magazine, 4(2), 41–57. https://doi.org/10.1109/MGRS.2016.2548504

26. Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020). A simple framework for contrastive learning of visual representations. In International Conference on Machine Learning (pp. 1597–1607). PMLR.

27. Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378, 686–707. https://doi.org/10.1016/j.jcp.2018.10.045