Optimizing Hyperspectral Soil Data Analysis: Developing Encoder–Decoder Architectures for Efficient Dimensionality Reduction and Reconstruction

Prajakta Kadukar, Computer Science
Mentor: Nakul Gopalan, Saurav Kumar
Arizona State University

How can deep learning encoder–decoder architectures optimize hyperspectral soil data for efficient dimensionality reduction and accurate reconstruction, while preserving essential spectral information?

Background

Hyperspectral imaging (HSI) captures hundreds of spectral bands (400–2500 nm), revealing detailed information about soil composition that standard RGB imaging cannot. However, its high dimensionality creates computational and storage challenges, known as the *"curse of dimensionality."*

Traditional techniques like PCA lose critical spectral details. This research explores encoder–decoder deep learning models to compress HSI efficiently while maintaining the information needed for soil carbon and environmental monitoring.

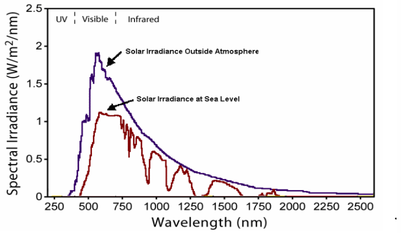

Figure 1. Hyperspectral range showing visible, VNIR, and SWIR bands

Figure 3. Collected soil samples used for hyperspectral imaging.

Research Objective

To design and train a self-supervised encoder–decoder model that compresses hyperspectral soil images into low-dimensional latent representations and reconstructs them with minimal loss.

The project aims to improve soil analysis, carbon mapping, and sustainable land-use monitoring through faster and more accurate hyperspectral data processing.

Methodology

- Data Collection: 110 soil samples captured using VNIR and SWIR cameras.
- Preprocessing: Noise reduction, normalization, calibration, and data splitting.
- Model Development: Encoder–decoder network trained in PyTorch/TensorFlow using Mean Squared Error loss; hyperparameters tuned (learning rate, batch size, latent dimension).
- Latent Analysis: PCA and t-SNE used to interpret learned features and relate them to soil properties.

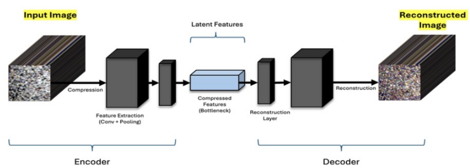

Figure 2. Encoder–decoder architecture for hyperspectral image compression and reconstruction.

Related Work

Classical dimensionality reduction methods like PCA often lose important spectral detail in hyperspectral data [1],[2].

Recent encoder–decoder-based models have shown better spectral preservation and improved reconstruction performance [3],[4].

This supports using deep learning-based compression for soil and environmental applications.

Preliminary Results

- loaded Indian Pines HSI cube (145×145×200)
- trained pixel-wise classifier using 16×16 patches
- achieved ~99.8% validation accuracy (99.8% = percent of validation pixels correctly classified)
- generated per-pixel vegetation classification map (Fig 4)
- confusion matrix confirms strong class separation (Fig 5)

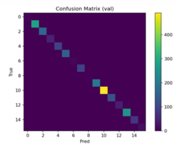

Figure 4. Confusion matrix illustrating model accuracy across the 16 vegetation classes.

Figure 5. Pixel-wise vegetation classification map generated for the Indian Pines data; colors = vegetation classes.

Future Work

- High-quality reconstructed images with low reconstruction error.
- Efficient data compression preserving spectral data.
- Insights into spectral–soil correlations through latent-space analysis.
- Contributions to scalable, sustainable hyperspectral data processing for environmental applications

Acknowledgements

Special thanks to Professor Nakul Gopalan and Professor Saurav Kumar for their mentorship, and to the Fulton Undergraduate Research Initiative (FURI) and ASU Research Computing for resources and support.

References

[1] C.-I. Chang, *Hyperspectral Data Processing*, Wiley, 2013.
[2] I. Fodor, *A Survey of Dimension Reduction Techniques*, 2002.
[3] Z. Zhang et al., "Encoder–Decoder Architecture for Hyperspectral Classification," *IEEE JSTARS*, 2021.
[4] G. Jaiswal et al., "Integration of hyperspectral imaging and autoencoders," *Computer Science Review*, 2023.

FURI

Ira A. Fulton Schools of Engineering
Arizona State University