Deep Learning-Based Super-Resolution for Landsat-to-AVIRIS Hyperspectral Image Enhancement

Judy[a]
[a]Department, Institution, Address

Abstract. We present a two-stage deep learning approach for enhancing Landsat-8 multispectral imagery to AVIRIS-NG-quality hyperspectral resolution. The method consists of (1) spectral super-resolution that expands 7 Landsat bands to 340 hyperspectral bands using learned spectral correlations and attention mechanisms, and (2) spatial super-resolution that increases spatial resolution from 30m to 4m using residual channel attention networks (RCAN). Training is performed on synthetically generated paired data derived from AVIRIS-NG radiance imagery. The approach achieves [results to be added] on validation data, demonstrating the feasibility of deep learning-based hyperspectral reconstruction from limited multispectral observations.

Keywords: Hyperspectral imaging, Super-resolution, Deep learning, RCAN, Spectral unmixing, AVIRIS-NG, Landsat-8.

1 INTRODUCTION

Hyperspectral imaging provides detailed spectral information crucial for remote sensing applications including mineral identification, vegetation analysis, and environmental monitoring. However, hyperspectral sensors like AVIRIS-NG[1] are limited in spatial and temporal coverage due to aircraft-based deployment. Conversely, satellite multispectral sensors like Landsat-8 OLI[2] provide global coverage but with reduced spectral resolution (7-11 bands vs. hundreds of hyperspectral bands).

This work addresses the challenge of reconstructing high-resolution hyperspectral imagery from low-resolution multispectral observations using deep learning. Previous approaches have explored spectral unmixing,[3] sparse representations, and more recently, convolutional neural networks.[4,5] We build upon residual channel attention networks (RCAN)[6] and spectral attention mechanisms[7] to achieve both spectral and spatial super-resolution.

1.1 Contributions

Our key contributions are:

- A two-stage architecture separating spectral (7→340 bands) and spatial (30m→4m) super-resolution, enabling efficient training with limited data
- Novel use of spectral angle mapper (SAM) loss[8] combined with spectral gradient constraints for preserving spectral fidelity
- Synthetic training data generation from AVIRIS-NG imagery that accurately simulates Landsat-8 spectral response functions and spatial degradation
- Validation on real AVIRIS-NG scenes demonstrating practical applicability

2 RELATED WORK

2.1 Hyperspectral Super-Resolution

Traditional hyperspectral super-resolution methods rely on spectral unmixing[3] and sparse coding. Aiazzi et al.[9] proposed Laplacian pyramid-based fusion of hyperspectral and multispectral data. More recently, deep learning approaches have shown superior performance. Sidorov and Hardeberg[4] introduced deep hyperspectral priors for denoising and super-resolution. Xie et al.[5] used 3D CNNs for multispectral-hyperspectral fusion. For comprehensive coverage, we refer to the recent survey by Dian et al.[10]

2.2 Residual Learning for Image Super-Resolution

Residual learning has proven highly effective for image super-resolution. Zhang et al.[6] introduced residual channel attention networks (RCAN), achieving state-of-the-art results through residual-in-residual structure with channel attention mechanisms. Zhang et al.[11] proposed residual dense networks (RDN) with dense connections. Our spatial super-resolution stage builds directly on the RCAN architecture, adapting it for multi-band hyperspectral data.

2.3 *Spectral Attention and 1D Convolutions*

For spectral dimension processing, Shi et al.[7] demonstrated effectiveness of 1D convolutions along the spectral axis. Jiang et al.[12] combined spectral and spatial attention in a unified framework. Our Stage 1 network employs spectral residual blocks with 1D convolutions and spectral attention, motivated by these works.

2.4 *Loss Functions for Hyperspectral Data*

Spectral Angle Mapper (SAM)[8] is a standard metric in hyperspectral remote sensing that measures spectral similarity regardless of illumination differences. Recent work by Lanaras et al.[13] incorporated SAM and spectral gradient constraints into training objectives. We adopt SAM loss as our primary spectral fidelity metric, combined with L1 reconstruction loss.

2.5 *Training with Synthetic Data*

Limited availability of paired training data motivates synthetic data generation. Xie et al.[5] and Zhu et al.[14] successfully trained networks on synthetically degraded hyperspectral imagery. We follow this paradigm, generating synthetic Landsat observations from AVIRIS-NG data through accurate spectral integration and spatial degradation.

3 METHODOLOGY

3.1 *Problem Formulation*

Let $\mathbf{X}_L \in \mathbb{R}^{H\times W\times 7}$ represent a Landsat-8 multispectral image with 7 bands at 30m spatial resolution, and $\mathbf{Y}_A \in \mathbb{R}^{7.5H\times 7.5W\times 340}$ represent the corresponding AVIRIS-NG hyperspectral image with 340 bands at 4m resolution. Our goal is to learn a mapping $f : \mathbf{X}_L \to \mathbf{Y}_A$ that reconstructs high-resolution hyperspectral imagery from low-resolution multispectral observations.

3.2 *Two-Stage Architecture*

We decompose the problem into two stages:

1. **Stage 1 - Spectral Super-Resolution**: $f_{\text{spec}} : \mathbb{R}^{H\times W\times 7} \to \mathbb{R}^{H\times W\times 340}$
2. **Stage 2 - Spatial Super-Resolution**: $f_{\text{spat}} : \mathbb{R}^{H\times W\times 340} \to \mathbb{R}^{7.5H\times 7.5W\times 340}$

This decomposition enables efficient training and provides interpretability by separating spectral and spatial reconstruction.

3.3 *Stage 1: Spectral Super-Resolution Network*

3.3.1 *Architecture*

The spectral SR network (SpectralSRNet) consists of:

- Initial spectral expansion via 2D convolutions (7→340 bands)
- $N_r = 8$ spectral residual blocks operating on flattened spatial dimensions
- Spectral attention module for band weighting
- Skip connection for stable training

Each spectral residual block uses 1D convolutions along the spectral dimension after reshaping $(B, C, H, W) \to (B, C, HW)$. This enables learning spectral correlations while remaining computationally efficient.

3.3.2 *Spectral Attention*

The spectral attention module computes importance weights for each of the 340 output bands:

$$\mathbf{w} = \sigma(\mathbf{W}_2 \cdot \text{ReLU}(\mathbf{W}_1 \cdot \text{GAP}(\mathbf{F}))) \tag{1}$$

where GAP is global average pooling, $\mathbf{W}_1 \in \mathbb{R}^{C/r\times C}$ and $\mathbf{W}_2 \in \mathbb{R}^{C\times C/r}$ are learned weights with reduction ratio $r = 4$, and σ is the sigmoid function.

3.4 *Stage 2: Spatial Super-Resolution Network*

3.4.1 *RCAN-Based Architecture*

The spatial SR network adapts RCAN[6] for hyperspectral data:

- Shallow feature extraction: Conv(340, 64, 3×3)
- $N_g = 4$ residual groups, each containing $N_b = 4$ residual channel attention blocks (RCAB)
- Sub-pixel convolution upsampling[15] with 3×2 PixelShuffle for 8× upsampling
- Bicubic interpolation adjustment for exact 7.5× scale
- Residual learning with bicubic baseline

3.4.2 *Residual Channel Attention Block*

Each RCAB contains:

$$\mathbf{F}_{out} = \mathbf{F}_{in} + \mathbf{CA}(\text{Conv}_2(\text{ReLU}(\text{Conv}_1(\mathbf{F}_{in})))) \tag{2}$$

where $\mathbf{CA}$ is the channel attention operation identical to spectral attention but applied to spatial feature maps.

3.5 *Synthetic Training Data Generation*

3.5.1 *Spectral Degradation*

We simulate Landsat-8 observations from AVIRIS-NG radiance data using Gaussian-approximated spectral response functions (SRFs):

$$L_i = \int_\lambda R(\lambda) \cdot \text{SRF}_i(\lambda) d\lambda \tag{3}$$

where $R(\lambda)$ is AVIRIS radiance, $\text{SRF}_i(\lambda)$ is the Landsat band i response, and L_i is the integrated Landsat radiance. We use band centers from Roy et al.:[2] 443nm (Coastal), 482nm (Blue), 562nm (Green), 655nm (Red), 865nm (NIR), 1609nm (SWIR-1), 2201nm (SWIR-2).

3.5.2 *Spatial Degradation*

Spatial degradation simulates 30m Landsat resolution from 4m AVIRIS:

1. Apply Gaussian PSF with $\sigma = 2.0$ pixels to simulate sensor blur
2. Downsample by factor 7.5 using area averaging
3. Bicubic upsample back to original spatial dimensions for pixel-wise correspondence

3.5.3 *Data Validation*

Patches with invalid pixels (data ignore value -9999) are rejected. We apply contamination checks to exclude patches affected by Gaussian blur spreading invalid values. Final dataset consists of 36 patches (256×256 @ 4m resolution) from 3 AVIRIS-NG scenes over Pasadena, CA.

3.6 *Loss Functions*

3.6.1 *Stage 1 Loss*

For spectral super-resolution:

$$\mathcal{L}_{\text{spec}} = \lambda_1 \mathcal{L}_{\text{L1}} + \lambda_2 \mathcal{L}_{\text{SAM}} + \lambda_3 \mathcal{L}_{\text{SG}} \tag{4}$$

where:

- $\mathcal{L}_{\text{L1}} = \|\mathbf{Y}_{\text{pred}} - \mathbf{Y}_{\text{true}}\|_1$ (reconstruction)
- $\mathcal{L}_{\text{SAM}} = \arccos\left(\frac{\mathbf{y}_p \cdot \mathbf{y}_t}{\|\mathbf{y}_p\|\|\mathbf{y}_t\|}\right)$ (spectral fidelity)
- $\mathcal{L}_{\text{SG}} = \|\nabla_\lambda \mathbf{Y}_{\text{pred}} - \nabla_\lambda \mathbf{Y}_{\text{true}}\|_1$ (spectral smoothness)

We use weights $\lambda_1 = 1.0, \lambda_2 = 0.1, \lambda_3 = 0.1$.

3.6.2 Stage 2 Loss

For spatial super-resolution:

$$\mathcal{L}_{\text{spat}} = \lambda_1 \mathcal{L}_{\text{L1}} + \lambda_2 \mathcal{L}_{\text{SAM}} \tag{5}$$

with weights $\lambda_1 = 1.0, \lambda_2 = 0.05$ (reduced SAM weight to emphasize spatial fidelity).

4 EXPERIMENTAL SETUP

4.1 Dataset

AVIRIS-NG Data: We use radiance data (not reflectance) from AVIRIS-NG flights over Pasadena, CA (June 2019). After excluding water absorption bands, 340 bands remain spanning 380-2510nm.

Training Set: 36 patches (256×256 @ 4m) from 3 scenes, split 80/20 train/validation.

Normalization: Robust percentile-based normalization (1st-99th percentile) to [0,1] range.

4.2 Training Details

Stage 1:

- Model: SpectralSRNet (7M parameters)
- Optimizer: Adam ($\beta_1 = 0.9, \beta_2 = 0.999$)
- Learning rate: 10^{-4} with ReduceLROnPlateau (factor=0.5, patience=10)
- Batch size: 4 (CPU) / 8 (GPU)
- Epochs: 100
- Hardware: NVIDIA TITAN V / Lambda Labs A100

Stage 2:

- Model: LightweightSpatialSRNet (130K parameters)
- Same optimizer and learning rate schedule
- Training on Stage 1 outputs

4.3 Evaluation Metrics

- **PSNR**: Peak signal-to-noise ratio (dB)
- **SSIM**: Structural similarity index
- **SAM**: Spectral angle mapper (degrees)
- **RMSE**: Root mean squared error per band
- **Spectral curves**: Visual comparison of reconstructed vs. ground truth spectra

5 RESULTS

[Results to be added after training completes]

5.1 Stage 1: Spectral Super-Resolution

5.2 Stage 2: Spatial Super-Resolution

5.3 Ablation Studies

5.4 Qualitative Results

6 DISCUSSION

6.1 Limitations

6.2 Future Work

7 CONCLUSION

[To be written]

Acknowledgments

This work was supported by [funding source]. AVIRIS-NG data courtesy of NASA/JPL.

References

1 R. O. Green, M. L. Eastwood, C. M. Sarture, *et al.*, "Aviris-ng: Results from the first year of operation," in *2015 AGU Fall Meeting*, AGU (2015).

2 D. P. Roy, M. A. Wulder, T. R. Loveland, *et al.*, "Landsat-8: Science and product vision for terrestrial global change research," *Remote Sensing of Environment* **145**, 154–172 (2014).

3 W. Dong, F. Fu, G. Shi, *et al.*, "Hyperspectral image super-resolution via non-negative structured sparse representation," *IEEE Transactions on Image Processing* **25**(5), 2337–2352 (2016).

4 O. Sidorov and J. Y. Hardeberg, "Deep hyperspectral prior: Single-image denoising, inpainting, super-resolution," in *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, 0–0 (2019).

5 Q. Xie, M. Zhou, Q. Zhao, *et al.*, "Multispectral and hyperspectral image fusion by ms/hs fusion net," *IEEE Transactions on Geoscience and Remote Sensing* **57**(2), 1041–1054 (2018).

6 Y. Zhang, K. Li, K. Li, *et al.*, "Image super-resolution using very deep residual channel attention networks," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 286–301 (2018).

7 Z. Shi, C. Chen, Z. Xiong, *et al.*, "Spectral super-resolution via deep residual learning," *IEEE Transactions on Computational Imaging* **5**(3), 456–468 (2019).

8 F. A. Kruse, A. Lefkoff, J. Boardman, *et al.*, "The spectral image processing system (sips)—interactive visualization and analysis of imaging spectrometer data," *Remote Sensing of Environment* **44**(2-3), 145–163 (1993).

9 B. Aiazzi, L. Alparone, S. Baronti, *et al.*, "Mtf-tailored multiscale fusion of high-resolution ms and pan imagery," *Photogrammetric Engineering & Remote Sensing* **72**(5), 591–596 (2006).

10 R. Dian, S. Li, L. Fang, *et al.*, "Learning a low tensor-train rank representation for hyperspectral image super-resolution," *IEEE Transactions on Neural Networks and Learning Systems* **30**(9), 2672–2683 (2022).

11 Y. Zhang, Y. Tian, Y. Kong, *et al.*, "Residual dense network for image super-resolution," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2472–2481 (2018).

12 J. Jiang, H. Sun, X. Liu, *et al.*, "Deep spectral-spatial network for hyperspectral image super-resolution," *ISPRS Journal of Photogrammetry and Remote Sensing* **170**, 104–117 (2020).

13 C. Lanaras, E. Baltsavias, and K. Schindler, "Hyperspectral super-resolution with spectral unmixing constraints," *IEEE Transactions on Image Processing* **30**, 1723–1736 (2021).

14 Z. Zhu, J. Hou, J. Chen, *et al.*, "Hyperspectral image super-resolution via deep progressive zero-centric residual learning," *IEEE Transactions on Image Processing* **30**, 1423–1438 (2021).

15 W. Shi, J. Caballero, F. Huszár, *et al.*, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1874–1883 (2016).