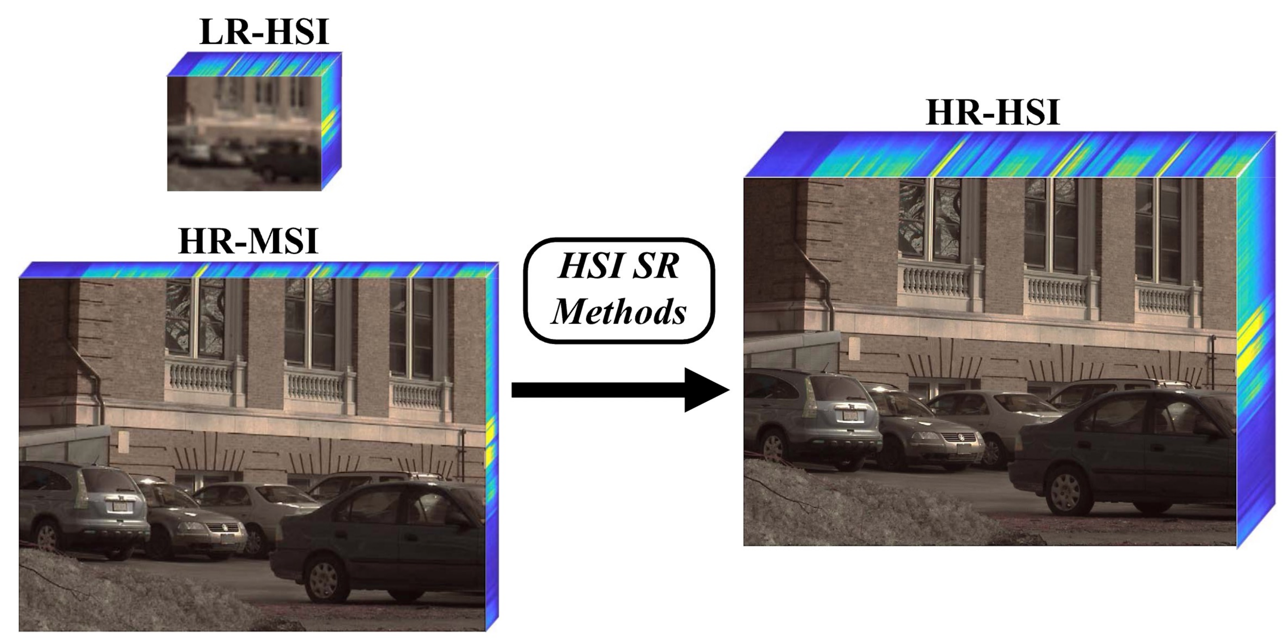
Hyperspectral Image Super-Resolution via Spatiospectral Attention and Frequency Domain Loss

Peter Phan
University of Toronto

Motivation

- Hyperspectral images are crucial for applications that require both spatial information and detailed spectral information. For example, in disease diagnosis, materials identification, and environmental monitoring.
- Effectiveness is limited due to low spatial resolution caused by physical constraints of sensors.
- High-resolution multispectral images (RGB) are often available alongside a corresponding low-resolution hyperspectral image. Various fusion-based methods have been explored to merge a high-resolution multispectral image with a low-resolution hyperspectral image to obtain a high-resolution hyperspectral image.



Related Work

- Zhang et al. introduced SSR-Net, a deep CNN fusion model incorporating a spatial edge loss and spectral edge loss. [1]
- Hu et al. introduced HSRnet, a deep CNN fusion network with separate spectral and spatial attention modules. [2]
- Xie et al. introduced MHF-Net, a model-based deep learning method. The MS/HS fusion model integrates generalization models of low-resolution images with low-rank prior knowledge of high-resolution hyperspectral images. The network is then constructed by unfolding the proximal gradient algorithm. This has advantages in interpretability and generalization. [3]

References

[1] Zhang, Huang, Wang, Li. SSR-NET: Spatial–spectral reconstruction network for hyperspectral and multispectral image fusion. IEEE Transactions on Geoscience and Remote Sensing, 2020
[2] Hu, Huang, Deng, Jiang, Vivone, Chanussot, Hyperspectral image super-resolution via deep spatiospectral attention convolutional neural networks, IEEE Transactions on Neural Networks and Learning Systems, 2021
[3] Xie, Zhou, Zhao, Xu, Meng. MHF-Net: An interpretable deep network for multispectral and hyperspectral image fusion. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020
[4] Park, Lee, Grossberg, Nayer. Multispectral Imaging Using Multiplexed Illumination. IEEE 11th International Conference on Computer Vision, 2007

Methods

- HSRnet fails to capture high-frequency features.
- **Dataset:** CAVE hyperspectral image dataset [4]
- **Training:** All models were trained on the same dataset and hyperparameters (200 epochs, Adam optimizer with 1e-4 learning rate, NVIDIA RTX 4060, approx. 1-1.5 hours training time per model)
- We experiment with various frequency domain-based adjustments to HSR-Net to improve frequency preservation.
  1. **High-Frequency Domain Loss Term**: punish loss of high-frequency details.

$$\mathcal{L}_{hf} = \frac{1}{HWC} \sum_{i=1}^{HWC} \|\mathrm{HighPass}(I_{\mathrm{GT}})_i - \mathrm{HighPass}(f(I_{\text{LR-HSI}}, I_{\text{HR-MSI}}))_i\|_2^2$$

  2. **Frequency Domain Fusion (FD-HSRnet)**: apply HSRnet to learn residuals in both the frequency domain and spatial domain.

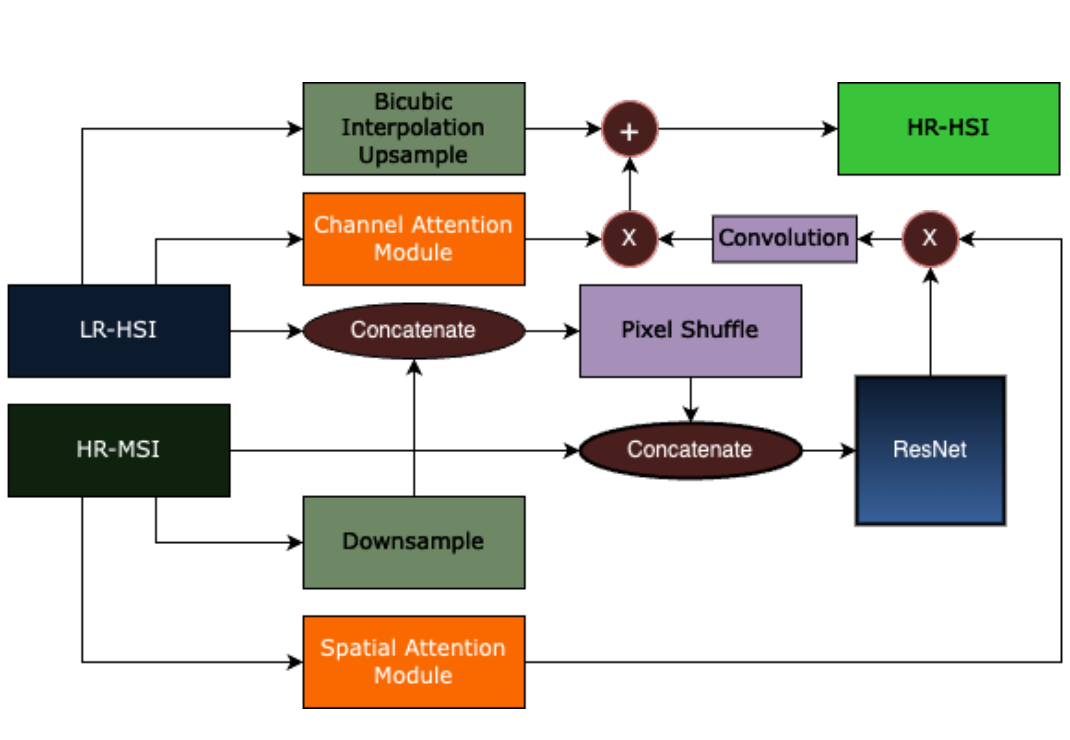


Figure 1: Base HSR-Net

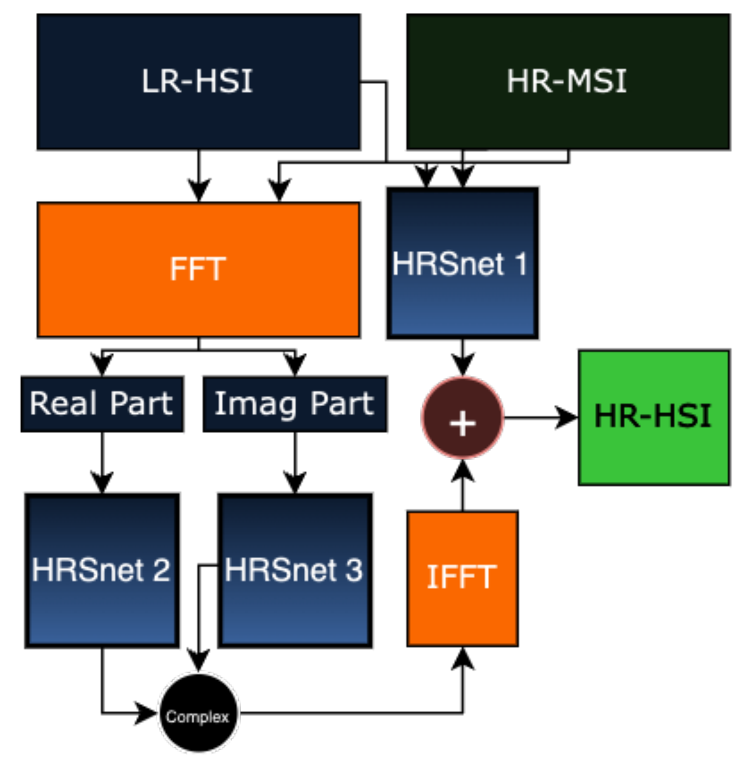


Figure 2: Frequency Domain Fusion

Experimental Results

- HSRnet trained with a high-frequency domain loss term has greater performance than base HSRnet in all metrics, in both the raw images and high-pass filtered images.

Model	PSNR	SSIM	MSE	SAM	ERGAS
Bicubic Interpolation	33.72	0.93	0.00046	0.065	3.83
HSRnet	33.73	0.930	0.00046	0.065	3.82
HSRnet + 0.05$\mathcal{L}_{hf}$	34.17	0.941	0.00041	0.061	3.66
HSRnet + 0.10$\mathcal{L}_{hf}$	34.69	0.951	0.00036	0.058	3.49
HSRnet + 0.15$\mathcal{L}_{hf}$	35.00	0.959	0.00033	0.055	3.36
HSRnet + 0.20$\mathcal{L}_{hf}$	**35.20**	**0.963**	**0.00032**	**0.054**	**3.29**
FD-HSRnet	33.73	0.93	0.00046	0.065	3.82

Table 1: Average Performance Across 11 Testing Images of the CAVE Dataset

Model	PSNR	SSIM	MSE	SAM	ERGAS
Bicubic Interpolation	18.95	0.109	0.051	1.337	38984
HSRnet	18.95	0.109	0.051	1.336	38984
HSRnet + 0.05$\mathcal{L}_{hf}$	20.02	0.213	0.040	1.142	35949
HSRnet + 0.10$\mathcal{L}_{hf}$	20.48	0.256	0.036	1.072	34642
HSRnet + 0.15$\mathcal{L}_{hf}$	21.32	0.360	0.030	0.954	31310
HSRnet + 0.20$\mathcal{L}_{hf}$	**21.55**	**0.386**	**0.028**	**0.925**	**30819**
FD-HSRnet	18.95	0.109	0.051	1.337	38984

Table 2: Average High-Frequency Performance Across 11 Testing Images of the CAVE Dataset

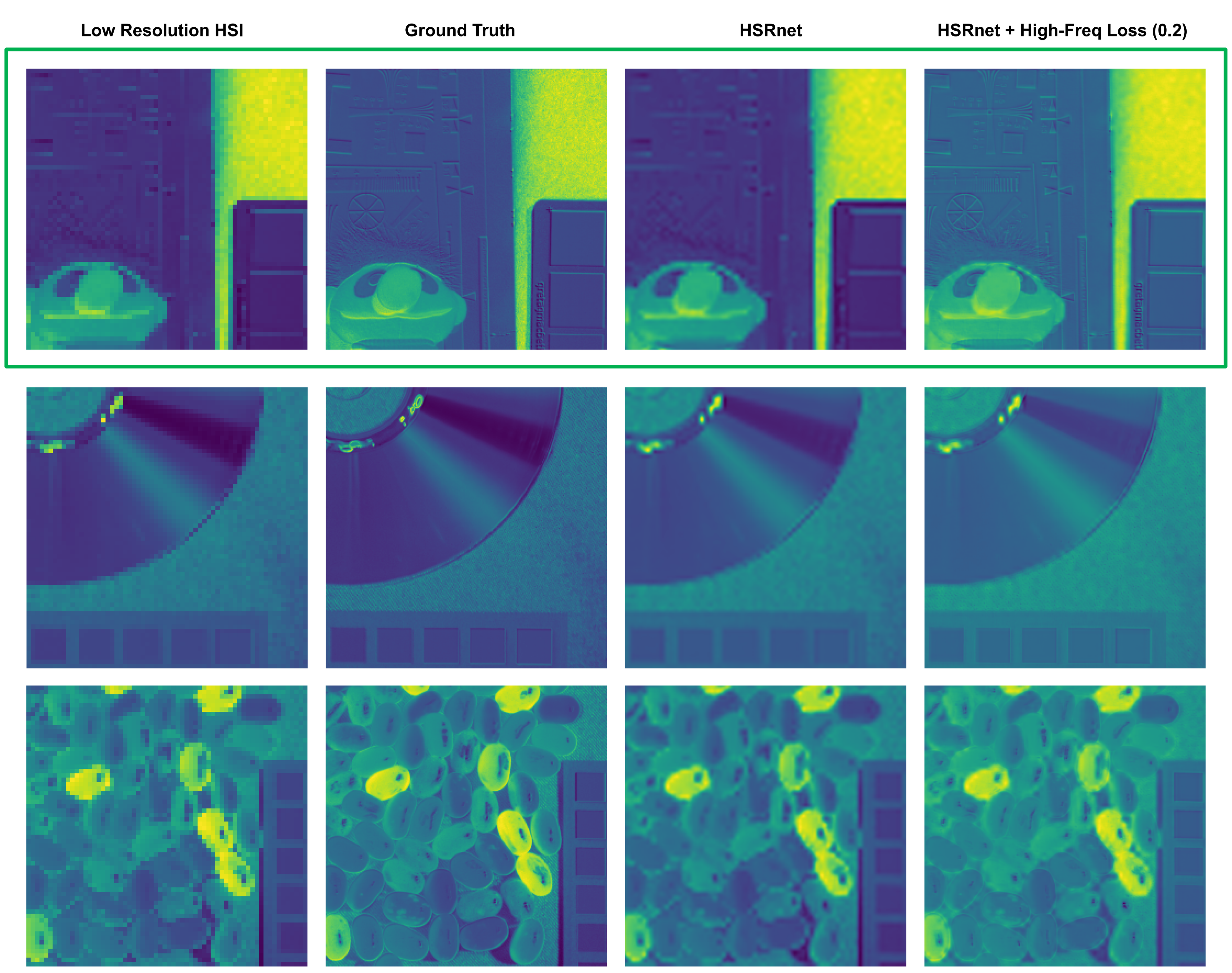


Figure 3: Experimental Results on Unseen Images