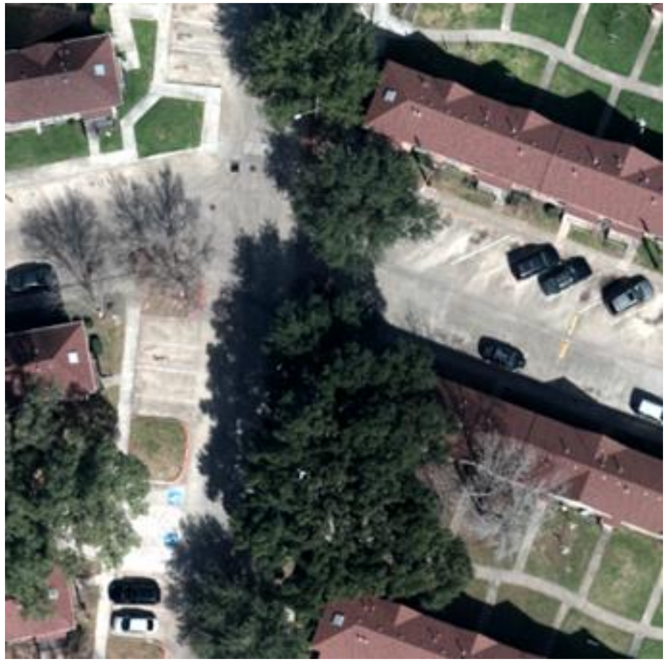
SpectraMorph: Structured Latent Learning for Self-Supervised Hyperspectral Super-Resolution

Ritik Shah and Marco F. Duarte

ICASSP, MAY 6 2026

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

Hyperspectral Image Fusion

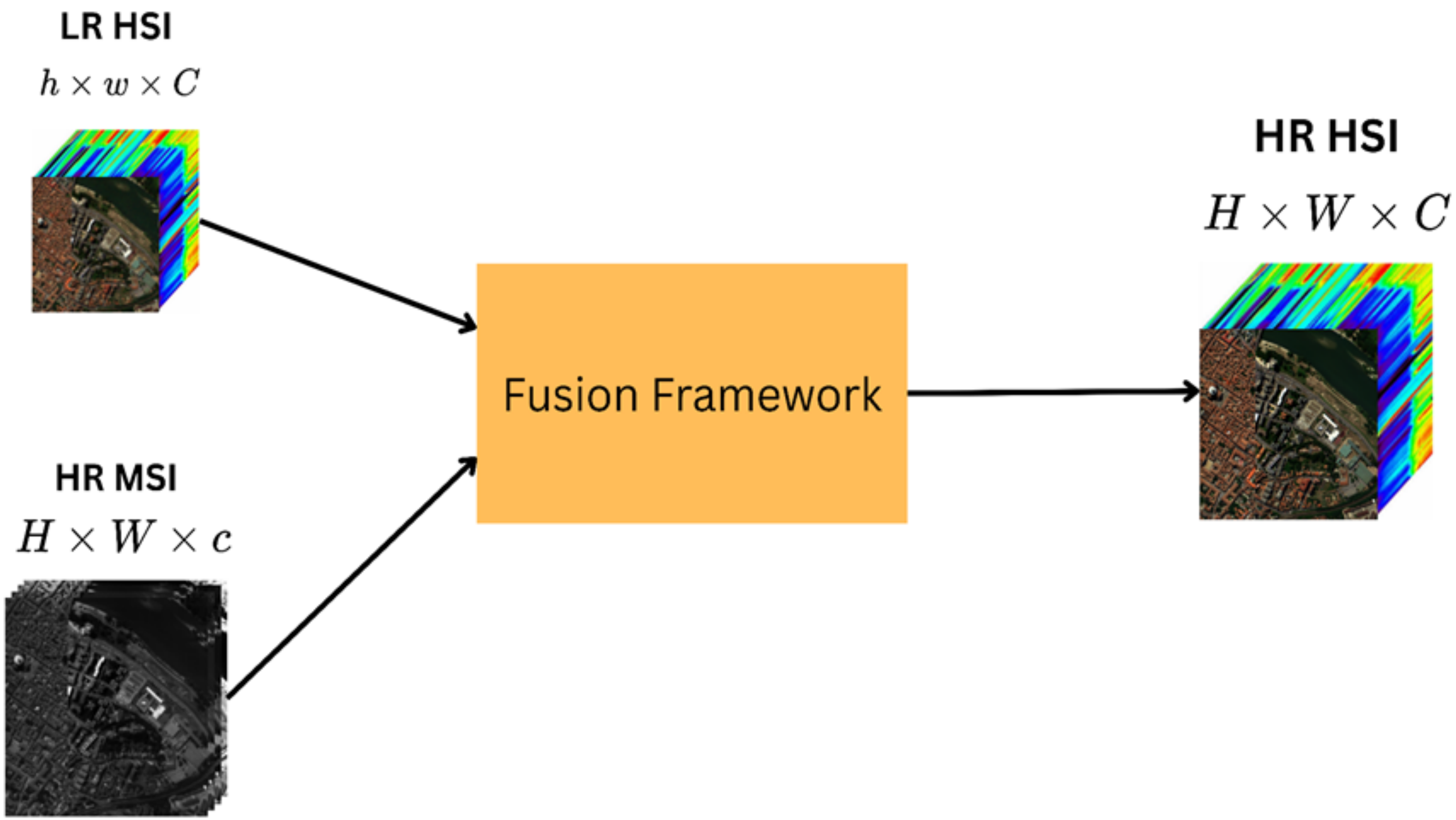


RGB High-Resolution Image
(HR-MSI)

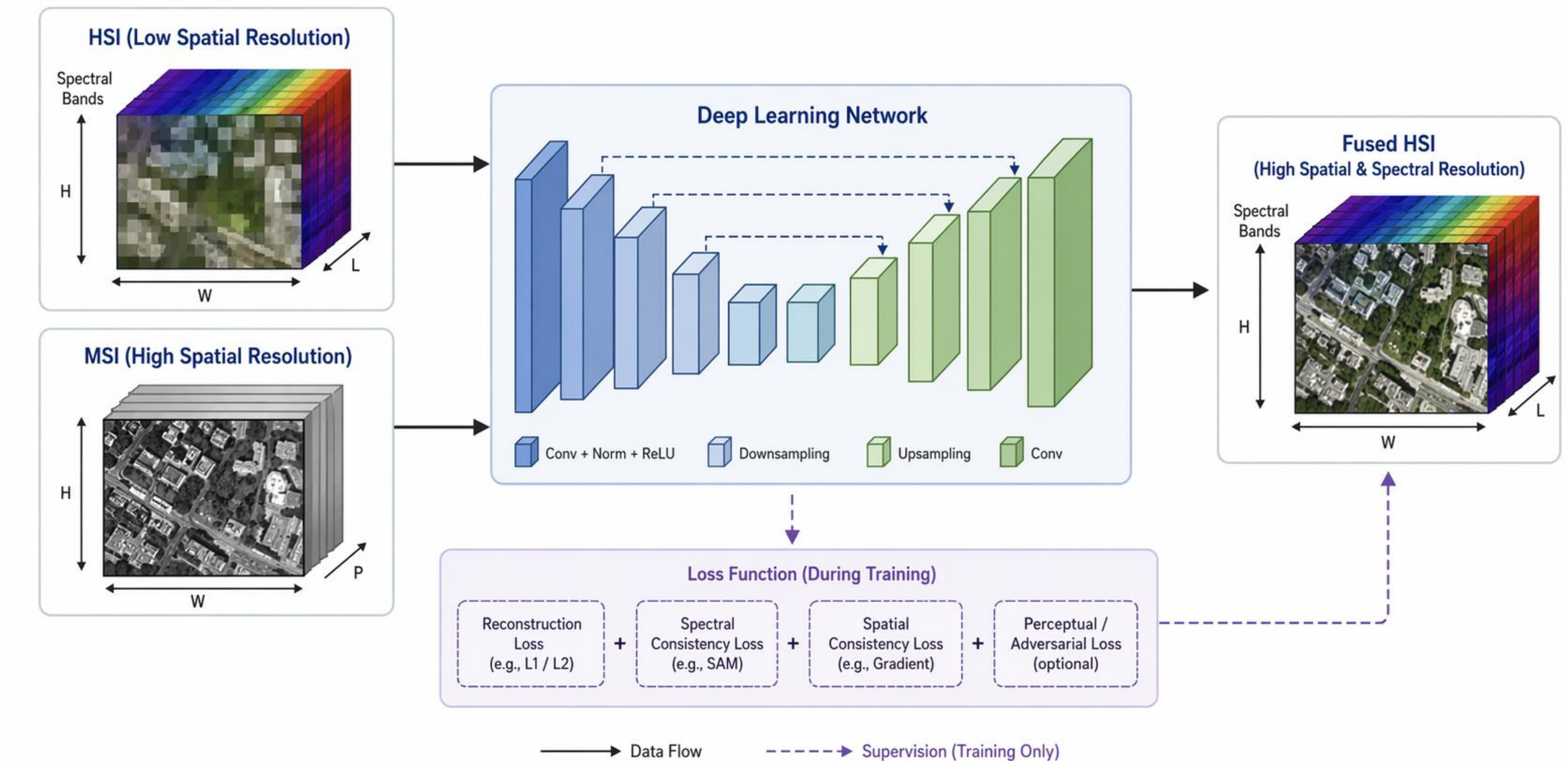
Hyperspectral Low-Resolution Image
(LR-HSI)

University of Houston Dataset
(2018 IEEE GRSS Data Fusion Contest)

Hyperspectral Image Fusion



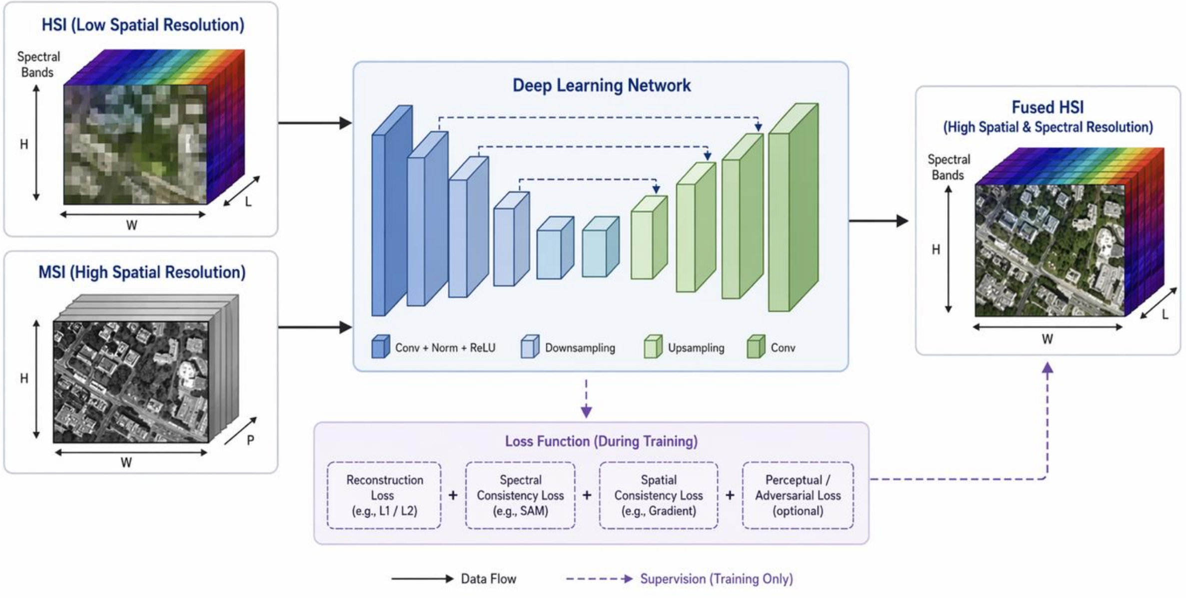
Prior State of the Art



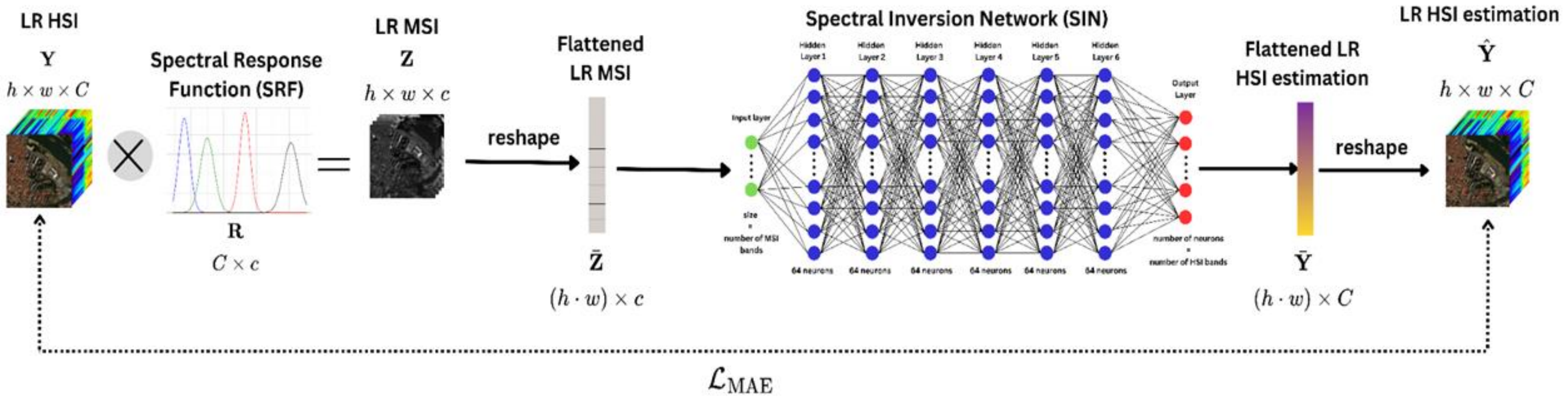
Prior State of the Art

SEVERAL ISSUES

- Supervised deep learning methods: Require ground truth HR HSI - impractical for real sensors
- Unsupervised/Self-supervised deep learning methods: Depend on accurate PSF or its estimation - often unreliable to obtain or estimate from data alone
- Explicit fusion methods: Sensitive to temporal misalignment - motion or illumination changes cause artifacts in output
- Performance depends on the blurring used: Some prior methods are sensitive to the blurring type present in the LR HSI
- Large complex architectures



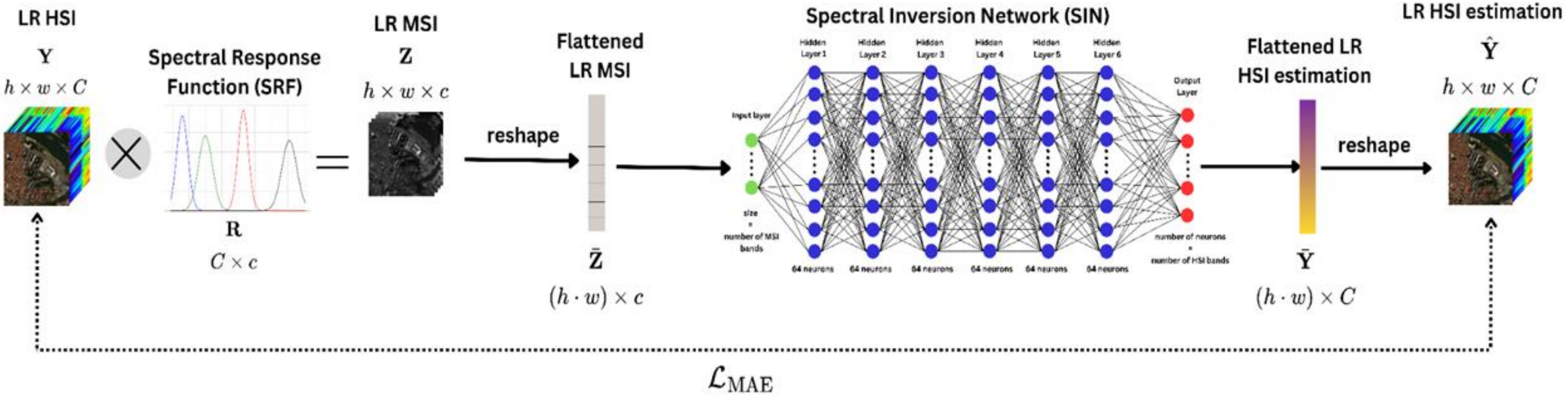
SpectraLift: Self-supervised HSI Image Fusion



- Spectral Inversion Network (SIN) learns mapping from multispectral to hyperspectral pixels
- Self-supervised learning: create LR-MSI from LR-HSI, learn inverse mapping
- Apply SIN to HR-MSI to obtain pixel-wise reconstruction of HR-HSI
- No spatial or temporal alignment issues

[Shah and Duarte, WHISPERS 2025]

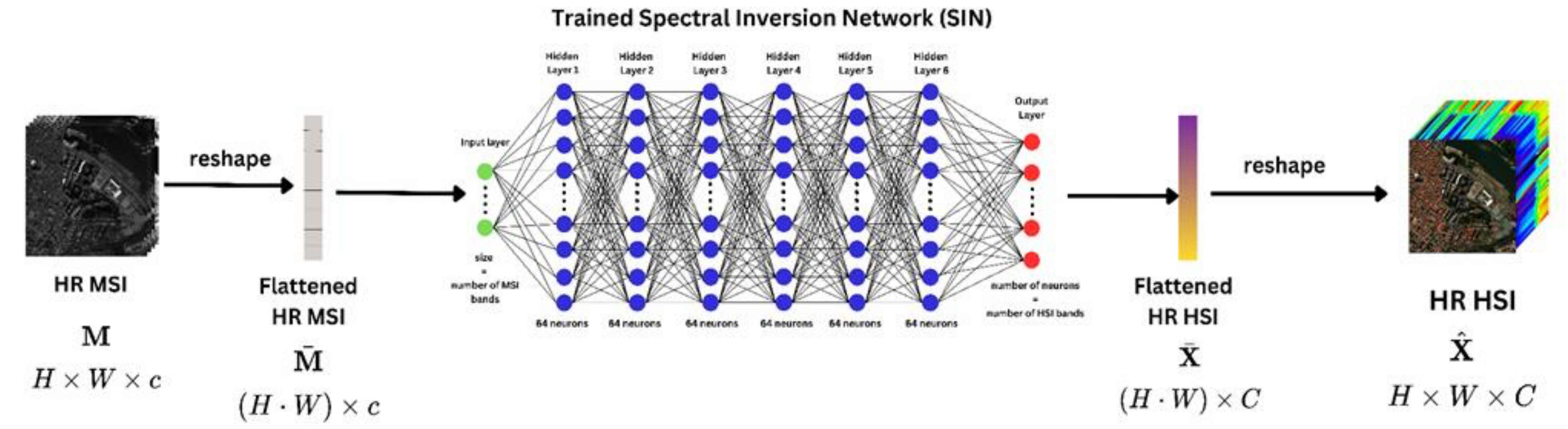
SpectraLift: Self-supervised HSI Image Fusion



- Light calibration: Requires neither PSF nor ground truth, only requires the MSI's SRF
- Resolution agnostic: Per-pixel spectral inversion naturally handles any spatial resolution - can train on LR and infer on HR
- Blur agnostic: Performs strongly across various blur types
- Compact: Small MLP (SIN) learns a physically intuitive spectral inversion, not explicit fusion

[Shah and Duarte, WHISPERS 2025]

SpectraLift: Self-supervised HSI Image Fusion



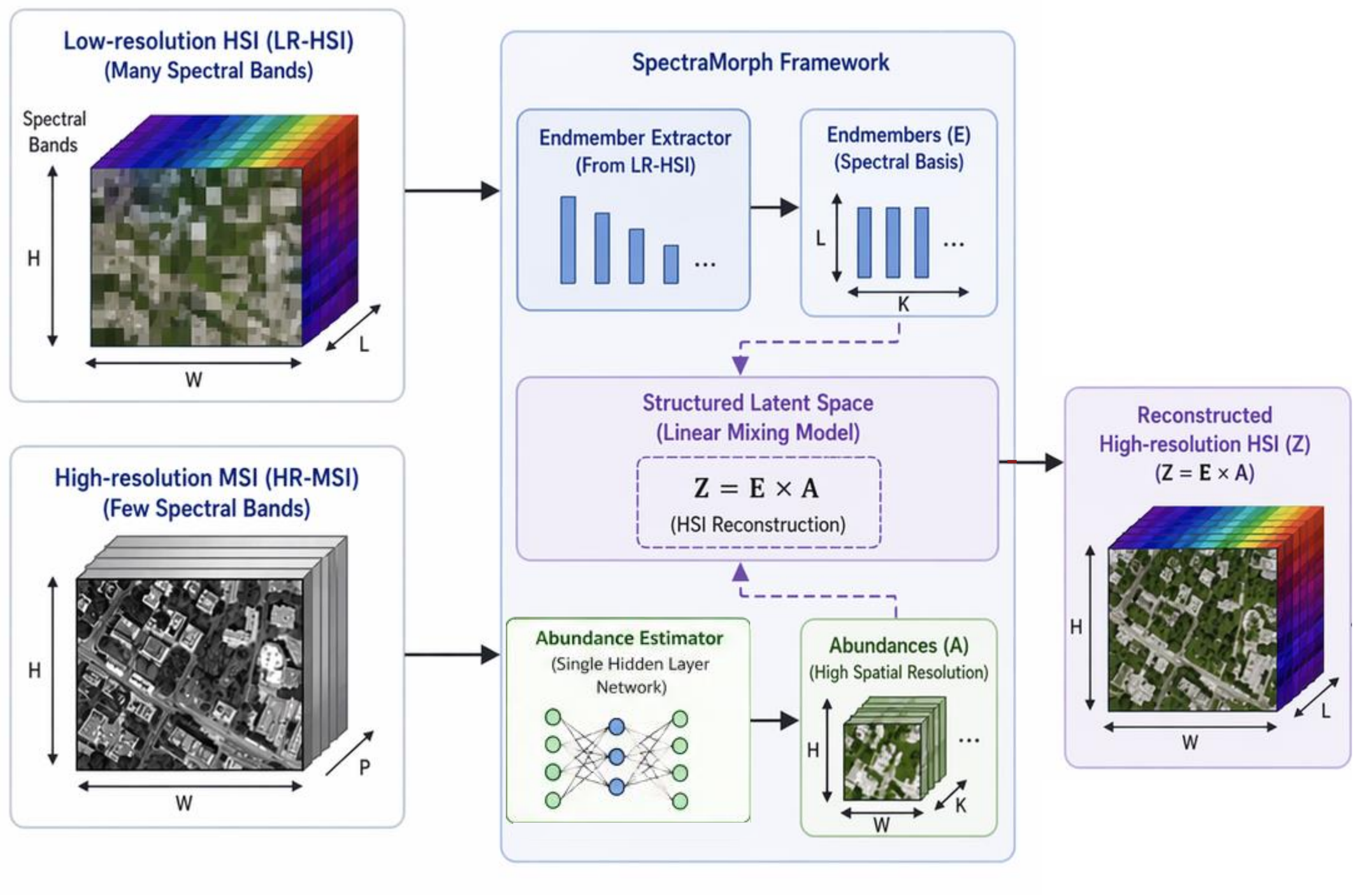
- Spectral Inversion Network (SIN) learns mapping from multispectral to hyperspectral pixels
- Self-supervised learning: create LR-MSI from LR-HSI, learn inverse mapping
- Apply SIN to HR-MSI to obtain pixel-wise reconstruction of HR-HSI
- No spatial or temporal alignment issues

[Shah and Duarte, WHISPERS 2025]

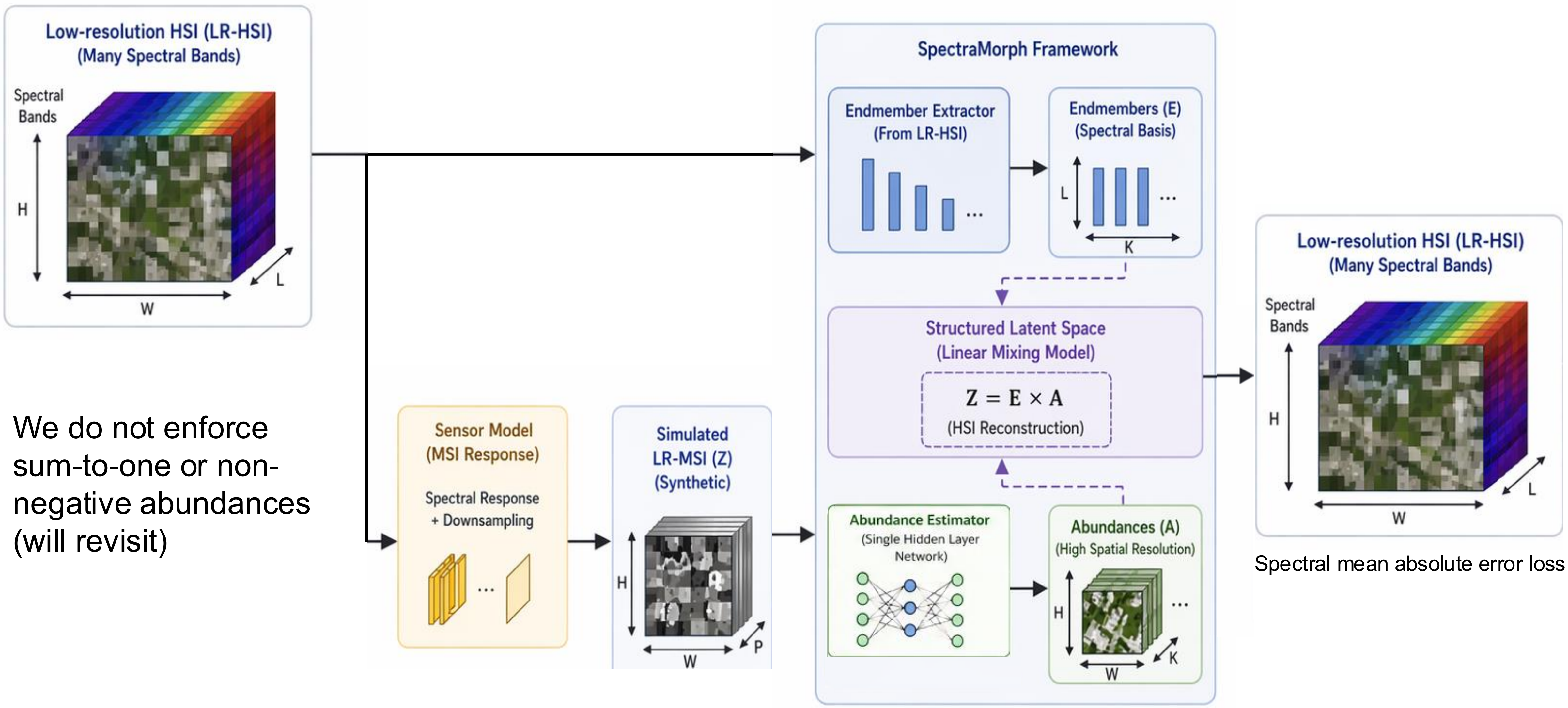
SpectraMorph: Self-Supervised HSI Image Fusion

ENDMEMBER UNMIXING: A SOUND PHYSICAL MODEL

- Endmember identification provides a strong physics-guided model for HSI pixels/spectral signatures
- Replace SpectraLift's blackbox inversion network w/ lightweight model (single hidden layer, 1024 neurons) that predicts endmember abundances from MSI pixel values
- Self-supervised training: compute endmembers with off-the-shelf methods (e.g., NNMF) applied on LR-HSI



SpectraMorph: Self-Supervised Training

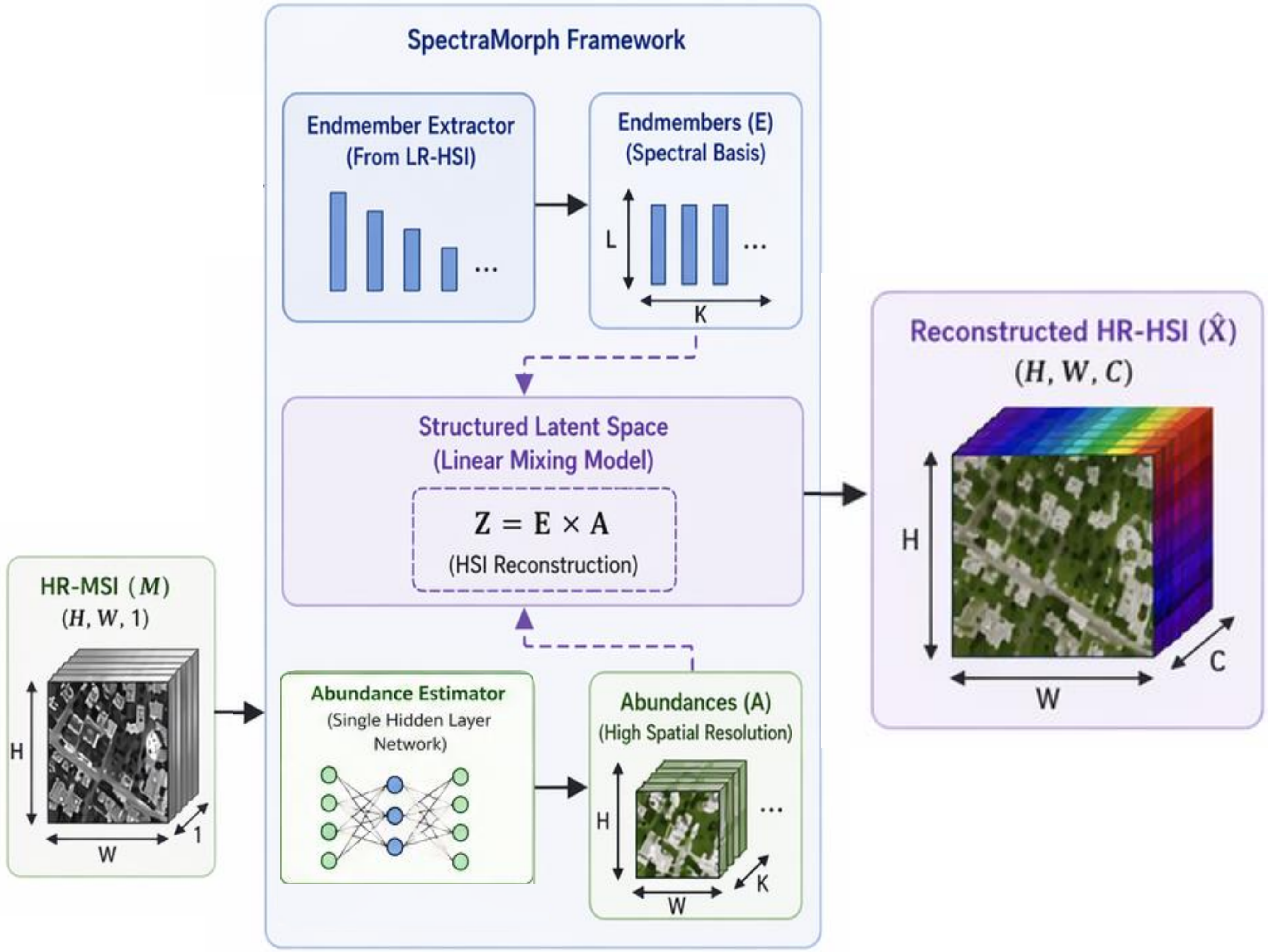


We do not enforce sum-to-one or non-negative abundances (will revisit)

Spectral mean absolute error loss

SpectraMorph: Self-Supervised Training

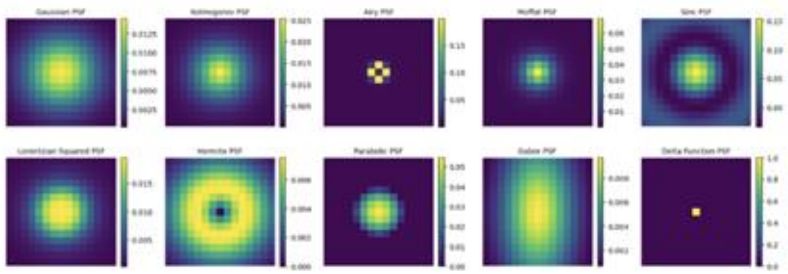
As with SpectraLift, apply learned model on each individual pixel to map HR-MSI pixels to HR-HIS pixels

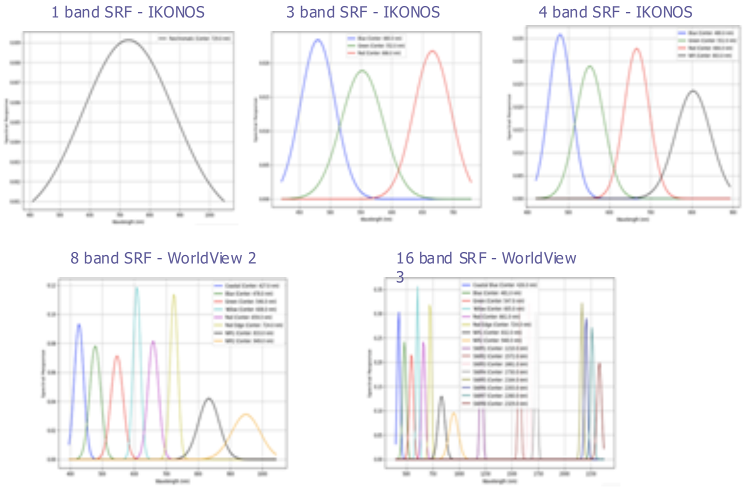


Evaluation Strategy with Synthetic Data

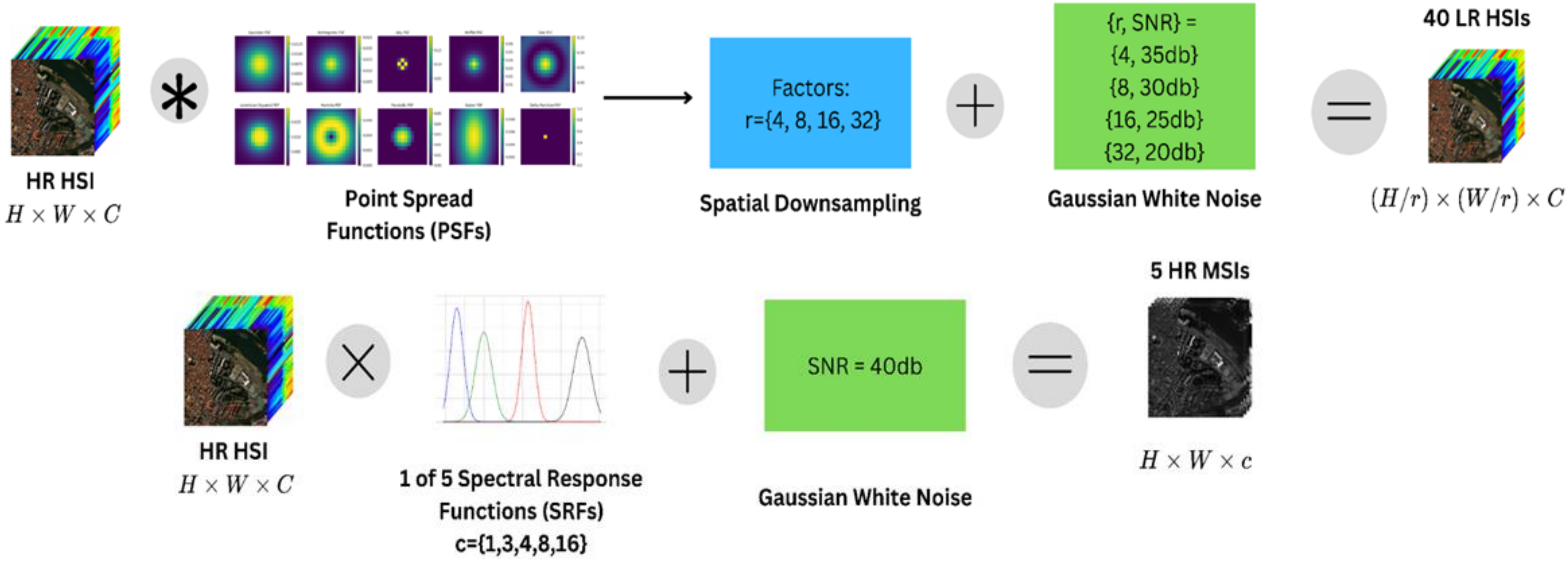
Library of HSI Fusion setups: **HyperBench**

- Automate generation of degraded (LR, MSI) data from HR-HSI ground truth
- Easily illustrate degradation details and resulting images
- Automate computation of performance measures (RMSE, PSNR, SSIM, UIQI, ERGAS, SAM)
- All that is needed: integration of HSI fusion method (wrapper) and hyperspectral datacube (MATLAB)





HyperBench Data Degradation Strategy



Experiments use 80 LR HSI/HR MSI input pairs:
{r,c} = {4,4}, {8,4}, {16,4}, {32,4}, {8,1}, {8,3}, {8,8}, {8,16}
For each of the 10 PSFs

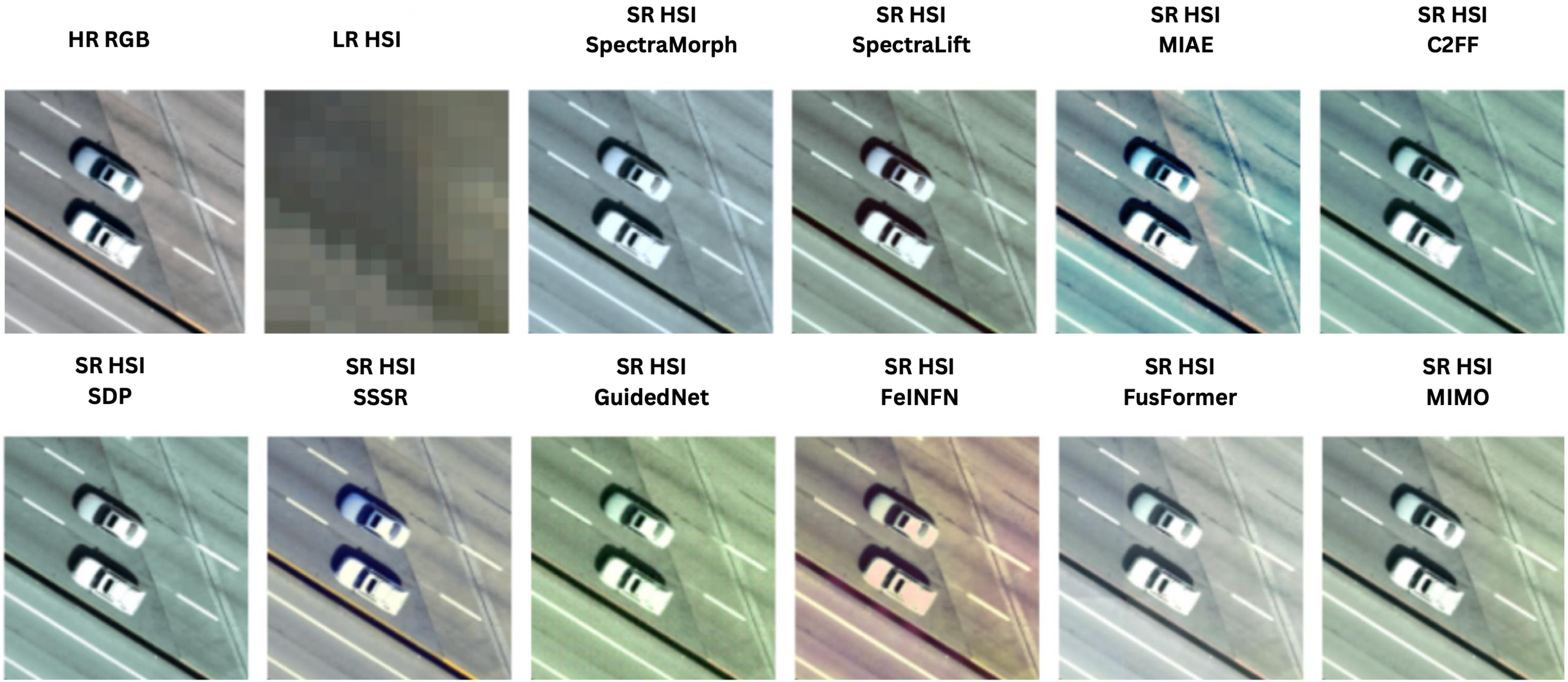
Evaluation Strategy with Synthetic Data

Table 1: Quality measures for the Washington DC Mall dataset: Mean value of 80 LR-HSI/HR-MSI configurations (**best in bold**, *second best in italics*; supervised methods are shaded).

Method	RMSE ↓	PSNR ↑	SSIM ↑	UIQI ↑	ERGAS ↓	SAM ↓	Time (s) ↓	Params (M) ↓	FLOPs (G) ↓	GPU Mem (MB) ↓
MIAE	0.03611	30.28	0.934	0.968	5.33	5.21	212.69	**0.0218**	8.46	707.54
C2FF	0.02681	34.74	0.960	0.975	4.73	3.53	78.54	0.0979	*7.29*	948.88
SDP	0.03029	31.99	0.939	0.907	16.59	4.29	799.75	6.4546	511.52	5370.37
SSSR	0.05024	27.98	0.908	0.938	7.49	5.83	*67.24*	*0.0334*	**0.000067**	*286.31*
SpectraLift	0.02344	**35.96**	0.967	0.969	5.87	*3.22*	91.20	0.0336	26.28	*286.31*
GuidedNet	0.03494	29.98	0.919	0.872	30.30	4.81	398.07	6.7111	178.41	**81.52**
FeINFN	0.02558	32.66	0.966	0.956	7.83	3.79	3018.80	3.7021	251.70	982.47
FusFormer	0.02555	32.55	0.876	0.688	17.34	3.89	11110.07	0.1883	946.28	6328.36
MIMO-SST	*0.02193*	35.21	*0.969*	*0.982*	**2.92**	3.29	376.34	2.1879	56.37	393.74
SpectraMorph	**0.01943**	*35.74*	**0.976**	**0.984**	*3.38*	**2.63**	**42.96**	0.0383	30.73	300.97
Oracle	0.01778	36.44	0.979	0.985	3.32	2.21	-	-	-	-

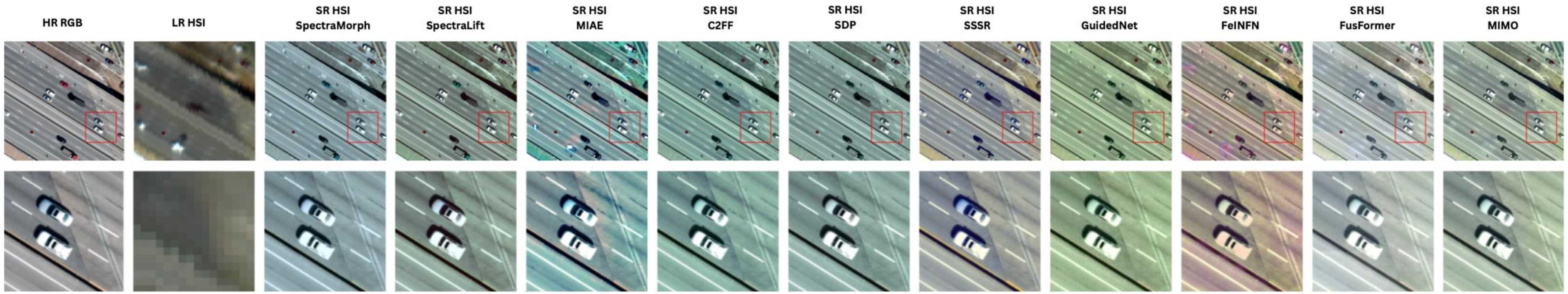
Results are representative of all datasets tested (DC, Pavia, Botswana) available in full paper

Evaluation Strategy: Datasets w/o Ground Truth

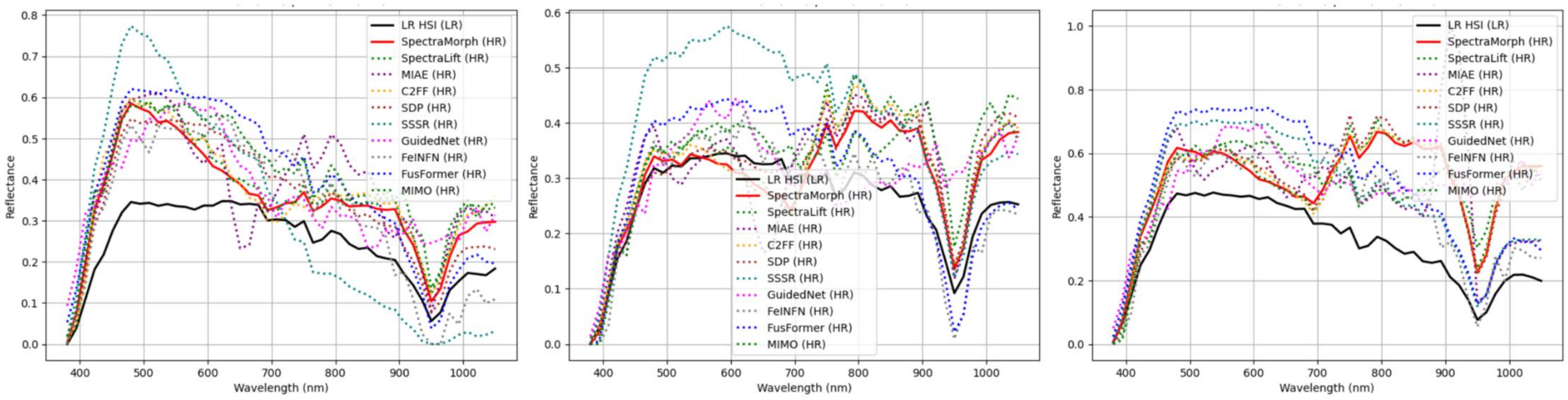


UH Dataset (2018 IEEE GRSS Data Fusion Contest)PSF, SRF unknown; IKONOS RGB SRF used, each method that needs PSF obtains estimate

Evaluation Strategy: Datasets w/o Ground Truth



(a) UH SR images for test scene 1. Second row shows a zoomed-in crop of the region marked in red.



(b) UH spectra for test scene 1. Left: Car, Center: Bare earth, Right: Highway.

UH Dataset
(2018 IEEE GRSS Data Fusion Contest)
PSF, SRF unknown; IKONOS RGB SRF used, each method that needs PSF obtains estimate

Ablation Study (Pavia Center, Gaussian PSF)

Variant	RMSE ↓	PSNR ↑	SSIM ↑	UIQI ↑	ERGAS ↓	SAM ↓	Params (M) ↓	FLOPs (G) ↓
Baseline	0.0217	34.41	**0.967**	**0.995**	**2.00**	**3.86**	0.029	45.98
Constraint: NN	0.0689	27.62	0.881	0.905	5.52	12.40	0.029	45.98
Constraint: NN and $\sum = 1$	0.0705	23.12	0.858	0.958	7.37	7.60	0.029	46.02
No hidden layers in LEN	0.0262	32.77	0.957	0.991	2.40	4.72	**0.0002**	**1.70**
64 Neurons in LEN	0.0226	34.00	0.965	0.994	2.09	4.04	0.002	4.23
128 Neurons in LEN	0.0226	34.10	0.965	0.994	2.06	3.96	0.004	7.01
256 Neurons in LEN	0.0221	34.30	0.966	0.994	2.03	3.91	0.007	12.58
512 Neurons in LEN	0.0218	34.36	**0.967**	**0.995**	2.01	3.87	0.014	23.71
2048 Neurons in LEN	**0.0216**	**34.45**	**0.967**	**0.995**	**2.00**	**3.86**	0.058	90.52
2 hidden layers in LEN	0.0218	34.40	**0.967**	0.994	2.01	3.89	1.079	1690.20
MAE $\rightarrow$ MSE loss	0.0220	34.26	0.966	0.994	2.03	3.95	0.029	45.98
ReLU $\rightarrow$ Leaky ReLU	0.0218	34.42	**0.967**	0.994	**2.00**	3.89	0.029	45.98
ReLU $\rightarrow$ GeLU	0.0229	34.02	0.964	0.993	2.09	4.07	0.029	49.19
ReLU $\rightarrow$ tanh	0.0231	33.98	0.964	0.994	2.10	4.08	0.029	45.98
$K \rightarrow K+4$	0.0217	34.41	**0.967**	**0.995**	**2.00**	3.89	0.033	53.04
$K \rightarrow K-4$	0.0227	33.90	0.966	0.994	2.10	4.07	0.025	38.92

QR: Full paper, Replicability code, HyperBench

SpectraMorph code on GitHub

SpectraMorph paper (extended version) on arXiv

HyperBench code on GitHub