Key Ideas

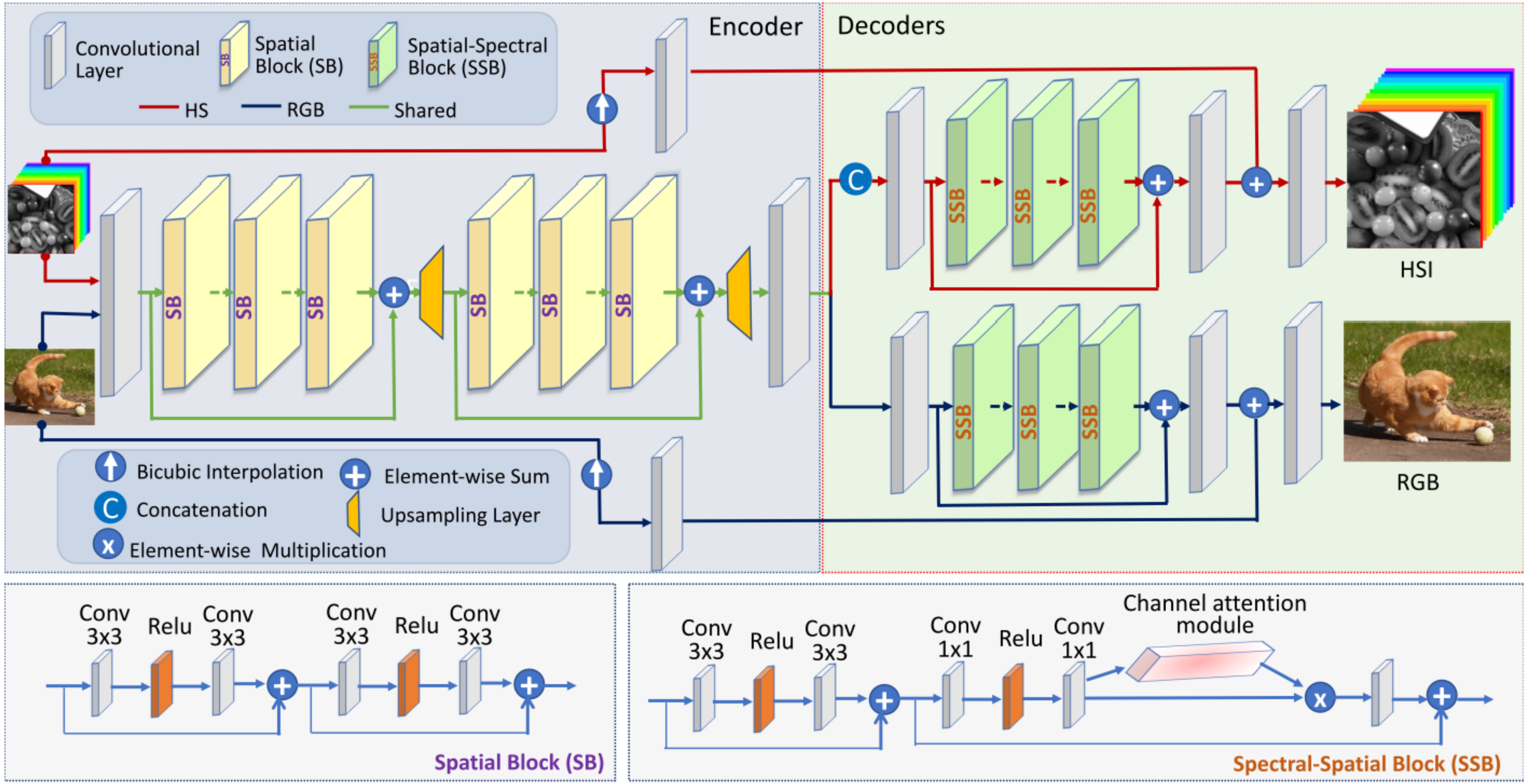
1. Basic

- RGBI SR and HIS SR **share common goals** in integrating information from neighbouring spatial regions during the learning;
- Difference in spectral band numbers -> propose a novel spatial-spectral neural network to solve them in a **multi-tasking framework**;
- The parameter distribution induced by the RGBI SR task can serve as an **effective regularization** for HIS SR task.

2. Furthermore

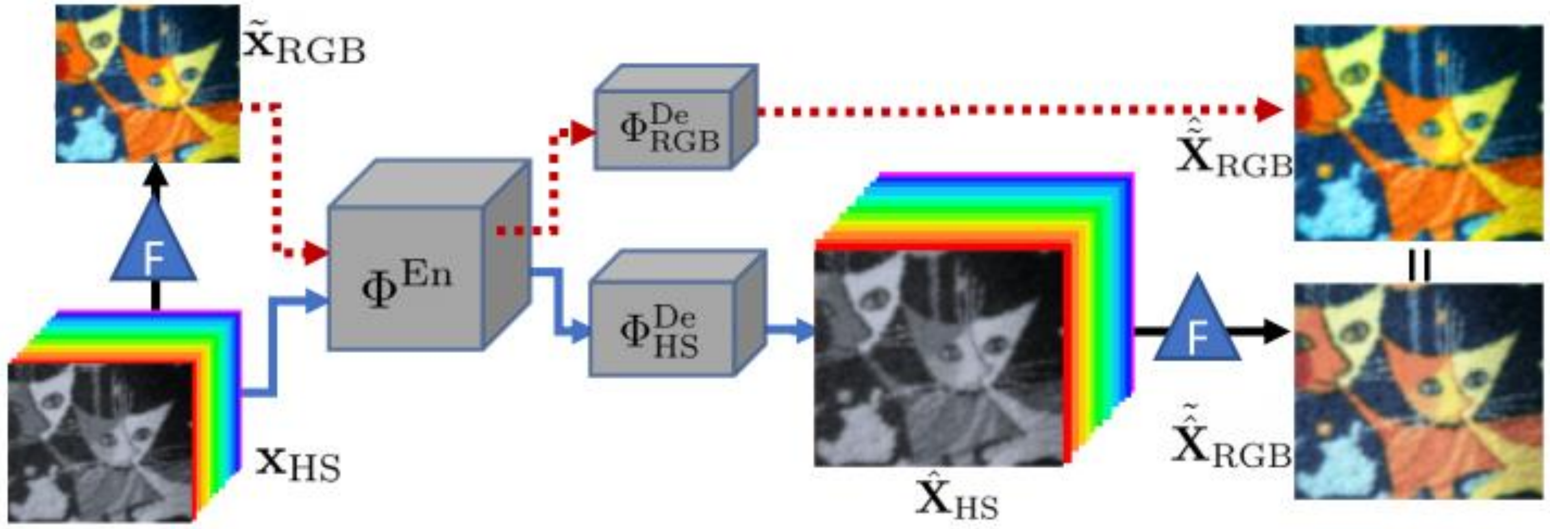
- Extend to Semi-supervised learning: (1)convert LR HSIs into LR RGB images and pass those through the **trained RGBI SR network;** (2) pass the LR HSIs **through our HSI SR network** to get the super-resolved HSIs and convert them to RGBIs;
(3) enforce the consistency between the two versions of super-resolved RGBIs.

Models



Convolutional Layer
Spatial Block (SB)
Spatial-Spectral Block (SSB)
HS
RGB
Shared
Encoder
Decoders
SB
SSB
HSI
RGB
Bicubic Interpolation
Element-wise Sum
Concatenation
Upsampling Layer
Element-wise Multiplication
Conv 3x3
Relu
Conv 3x3
Spatial Block (SB)
Conv 1x1
Channel attention module
Spectral-Spatial Block (SSB)

Models



$$\mathcal{L}^{\text{Total}} = \mathcal{L}^{\text{HS}}(\mathbf{X}_{\text{HS}}, \hat{\mathbf{X}}_{\text{HS}}) + \mathcal{L}^{\text{RGB}}(\mathbf{X}_{\text{RGB}}, \hat{\mathbf{X}}_{\text{RGB}}) \\ + \mathcal{L}^{\text{SSL}}(\hat{\tilde{X}}_{\text{RGB}}, \tilde{\hat{\mathbf{X}}}_{\text{RGB}}).$$

Data

Low HSI Datasets:

1) CAVE:
- 31 bands ranging from 400 nm to 700 nm at a step of 10 nm;
- 32 images of 512 x 512 pixels;
- 20(5/15) for training and 10 for testing.

2) NTIRE 2020:
- 31 bands ranging from 400 nm to 700 nm at a step of 10 nm;
- 480 images, 400(100/300) images for training and 80 images for test.

3) Harvard dataset:
- 31 bands as well but range from 420 nm to 720 nm;
- 50 images in total, use 40(6/34) for training and 10 for test.

Low RGBI Dataset:
- Down sampling by a factor of x2 from DIV2K Dataset;
- 137, 430image patches of 64 × 64 pixels;

Scaling factor ×4 and ×8.
For the case of ×4, we crop the images into patches of 64 × 64 pixels without overlapping to collect the training data.
For ×8, we use patches of 128 × 128 pixels;

Results

X4:

#(Mini-Batches)	0	1	2	3	4	5	6	8	10
RMSE ↓	0.01451	0.01357	0.01329	0.01309	0.01308	0.01305	0.01315	0.01315	0.01317

Table 1: Performance as a function of the number of mini-batches for RGBI SR loss.

	Components		CAVE			Harvard			NTIRE		
Methods	RGBSR	SSL	RMSE ↓	MPSNR ↑	ERGAS ↓	RMSE ↓	MPSNR↑	ERGAS ↓	RMSE ↓	MPSNR↑	ERGAS ↓
Ours			0.0144	40.8385	4.0345	0.0146	40.4666	3.1712	0.0154	38.3149	2.2069
Ours	✓		0.0118	42.3575	3.0128	0.0134	40.7579	3.0769	0.0150	38.7229	2.1189
Ours	✓	✓	**0.0114**	**42.7645**	**3.3346**	**0.0132**	**40.9317**	**3.0128**	**0.0150**	**38.9642**	**2.065**
Bicubic	-	-	0.0185	38.7380	5.2719	0.0167	38.8975	3.8069	0.0235	34.7401	3.1901
GDRRN [36]	-	-	0.0246	36.2775	7.0043	0.0160	38.6953	4.3031	0.0197	36.0793	2.8175
3DFCNN [38]	-	-	0.0173	38.3928	6.7055	0.0157	39.3441	3.6172	0.0208	35.6630	2.8246
SSPSR [27]	-	-	0.0144	40.9131	4.0406	0.0142	40.3209	3.2274	0.01636	38.0740	2.2539
MCNet [35]	-	-	0.0146	40.7385	4.1659	0.01468	40.1873	3.26059	0.0168	38.0248	2.2834

Results

X8:

Methods	Components RGBSR	SSL	CAVE RMSE ↓	CAVE MPSNR ↑	CAVE ERGAS ↓	Harvard RMSE ↓	Harvard MPSNR↑	Harvard ERGAS ↓	NTIRE RMSE ↓	NTIRE MPSNR↑	NTIRE ERGAS ↓
Ours			0.0241	35.8976	7.1154	0.0221	36.6527	4.8522	0.0232	32.8287	4.0434
Ours	✓		0.0215	37.1387	6.1442	0.0205	37.1859	4.5575	0.0269	33.3306	3.8548
Ours	✓	✓	**0.0206**	**37.3532**	**6.0027**	**0.0201**	**37.3546**	**4.5448**	**0.0263**	**33.4557**	**3.8437**
Bicubic	-	-	0.0304	34.2221	8.4350	0.0249	35.7409	5.4772	0.0396	29.9589	5.4594
GDRRN [36]	-	-	0.0347	32.9363	9.8554	0.0238	35.6441	5.7287	0.0359	30.6723	5.1265
3DFCNN [38]	-	-	0.0292	32.9024	16.7265	0.0237	36.0551	5.2192	0.3857	9.1753	6.1624
SSPSR [27]	-	-	0.0248	35.8896	7.0394	0.0228	36.4563	4.9978	0.0326	31.7896	4.4952
MCNet [35]	-	-	0.0280	34.3116	10.2985	0.0234	36.3921	5.0572	0.0327	31.9629	4.4169