PoNet: A Universal Physical Optimization-based Spectral Super-resolution Network for Arbitrary Multispectral Images

2021.11.13

Reporter: Jiang He

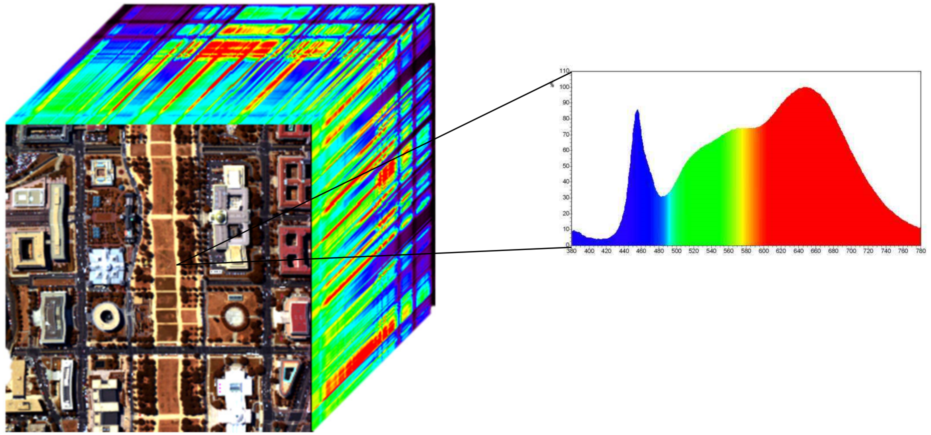
Supervisor: Qiangqiang Yuan

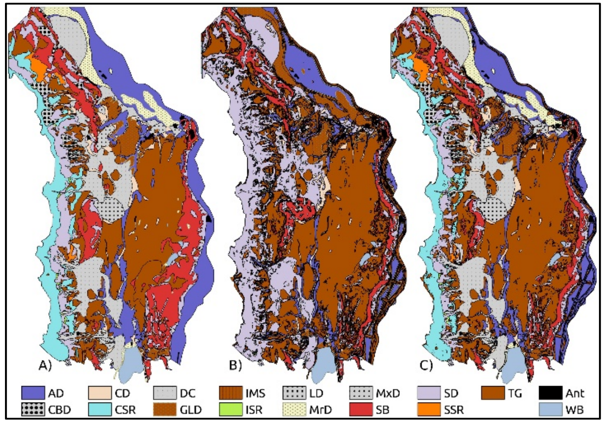
WUHAN UNIVERSITY

Outline

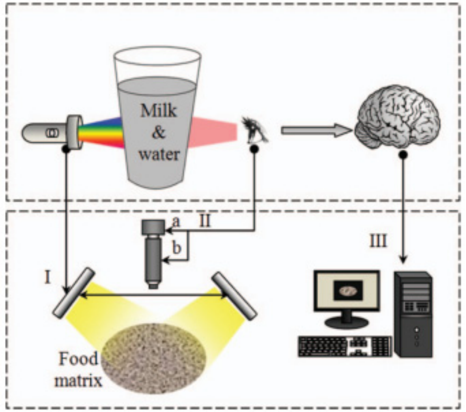
1 Background

- ☐ **Hyperspectral images are of great significance !**

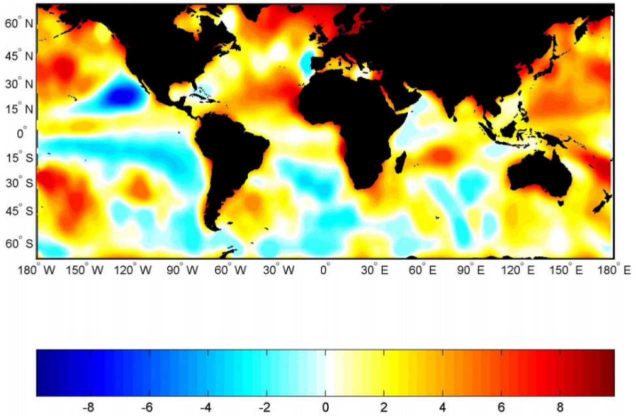




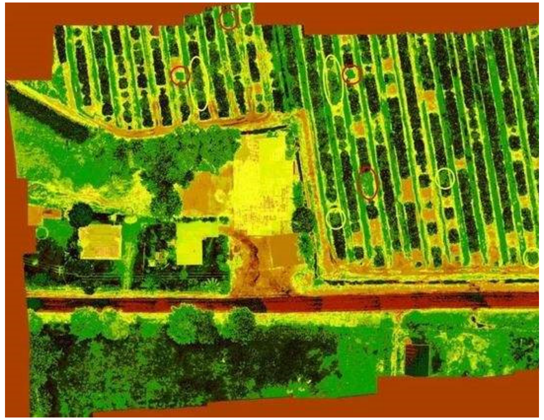
Geological analysis



Food Safety



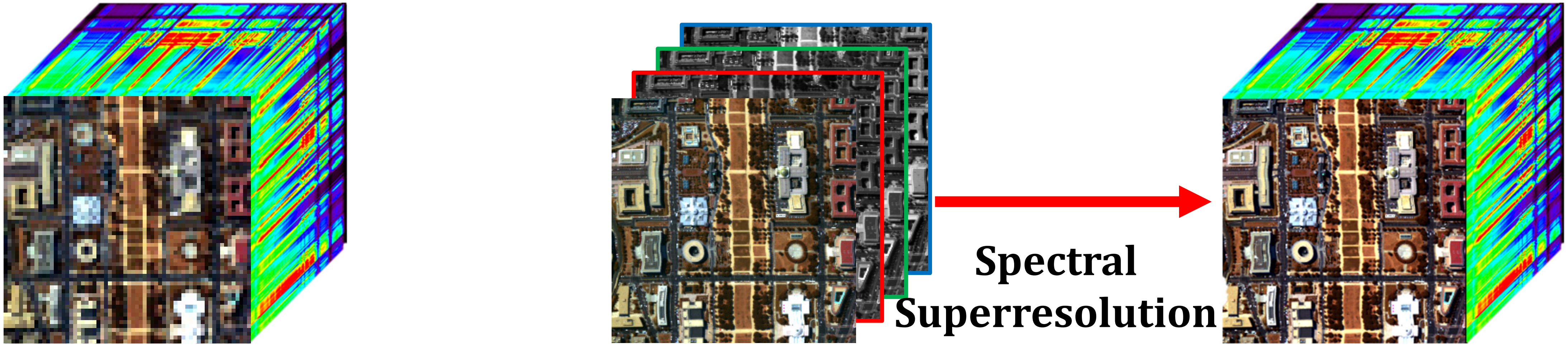
Ocean monitoring



Agricultural monitoring

1 Background

Spectral Super-resolution



Rich Spectral information

More Spatial Detail

High Spatial Resolution

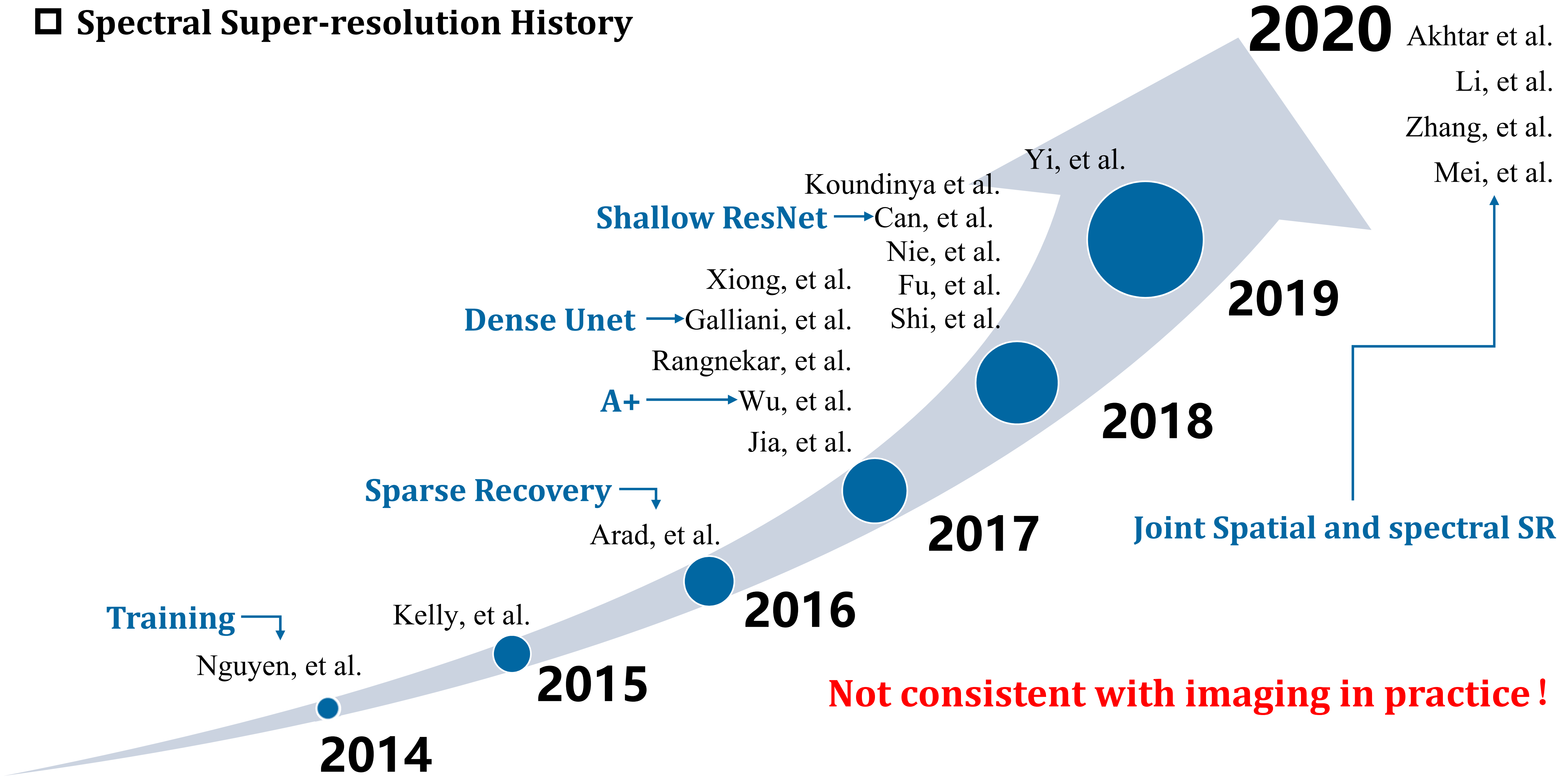
Low Spatial Resolution
High Acquisition Cost

Low Spectral Resolution

High Spectral Resolution

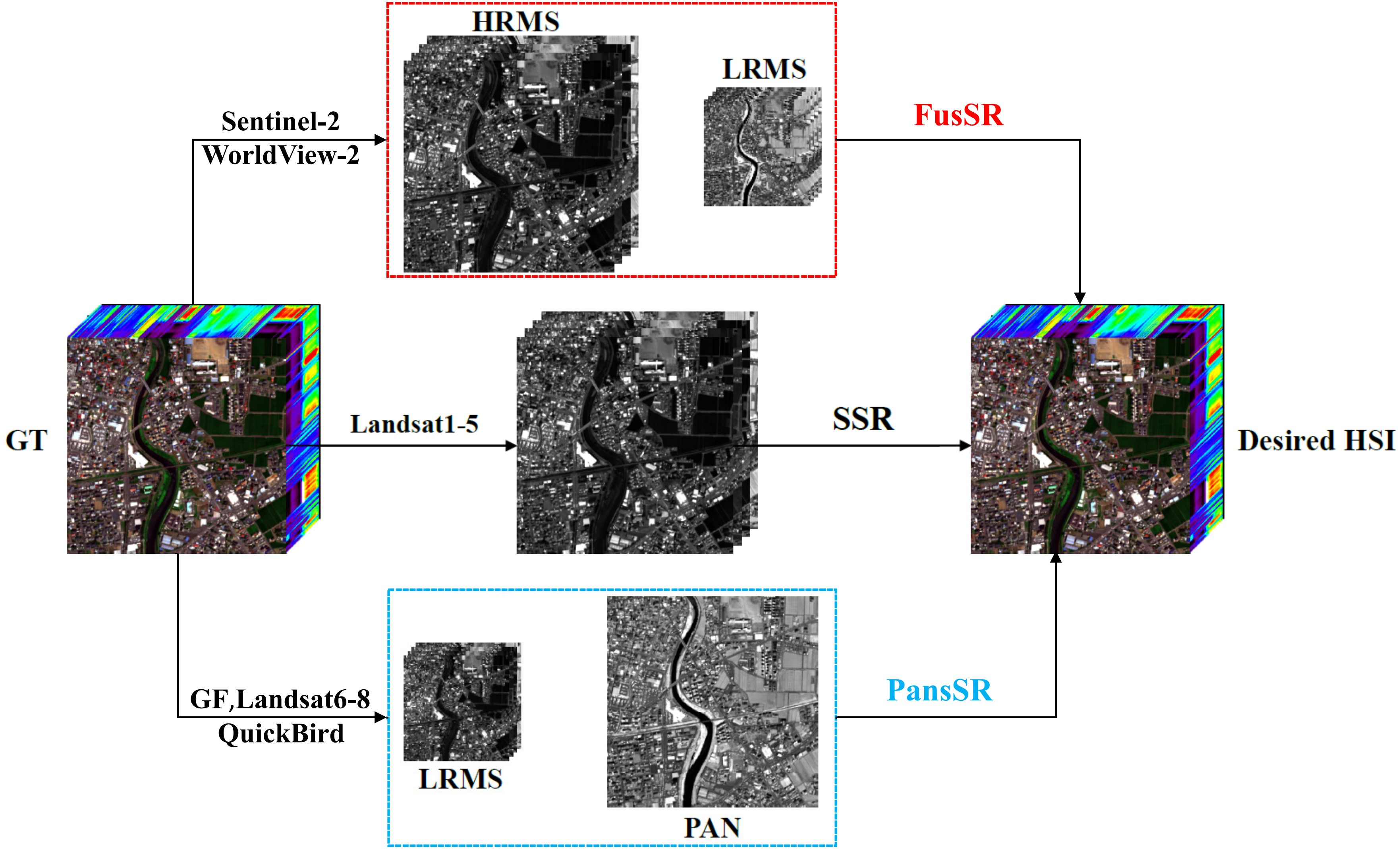
1 Background

Spectral Super-resolution History



Not consistent with imaging in practice！

Generalized Spectral Super-resolution



Outline

2 Method

Degradation Model

$$M_H = X\Phi_1$$
$$M_L = DX\Phi_2$$

Model Formulation

$$\hat{X} = \underset{X}{\operatorname{argmin}} \frac{1}{2}\|M_H - X\Phi_1\|_2^2 + \frac{1}{2}\|M_L - DX\Phi_2\|_2^2 + \lambda\mathcal{R}(X)$$

Half-Quadratic Splitting

$$\begin{cases} \hat{X}_{k+1} = (1-\epsilon\mu)X_k - \epsilon X_k\Phi_1\Phi_1^T - \epsilon D^T D X_k \Phi_2 \Phi_2^T + \epsilon M_H \Phi_1^T + \epsilon D^T M_L \Phi_2^T + \epsilon\mu Z_k \\ \hat{Z}_{k+1} = \operatorname{Prox}(X_{k+1}) = \underset{Z}{\operatorname{argmin}} \|Z - X_{k+1}\|_2^2 + \frac{\lambda}{\mu}\mathcal{R}(Z) \end{cases}$$

Model Requires Manual intervention

Deep unrolling

Inadequate expression of single prior

Deep unrolling

- X subproblem

$$\hat{X}_{k+1} = (1-\epsilon\mu)X_k - \epsilon X_k \Phi_1 \Phi_1^T - \epsilon D^T D X_k \Phi_2 \Phi_2^T + \epsilon M_H \Phi_1^T + \epsilon D^T M_L \Phi_2^T + \epsilon\mu Z_k$$

$$\boxed{\hat{X}_{k+1} = (1-\epsilon\mu)X_k + X_{k+1}^{\downarrow\uparrow} + M_{k+1}^{\uparrow} + \epsilon\mu Z_k}$$

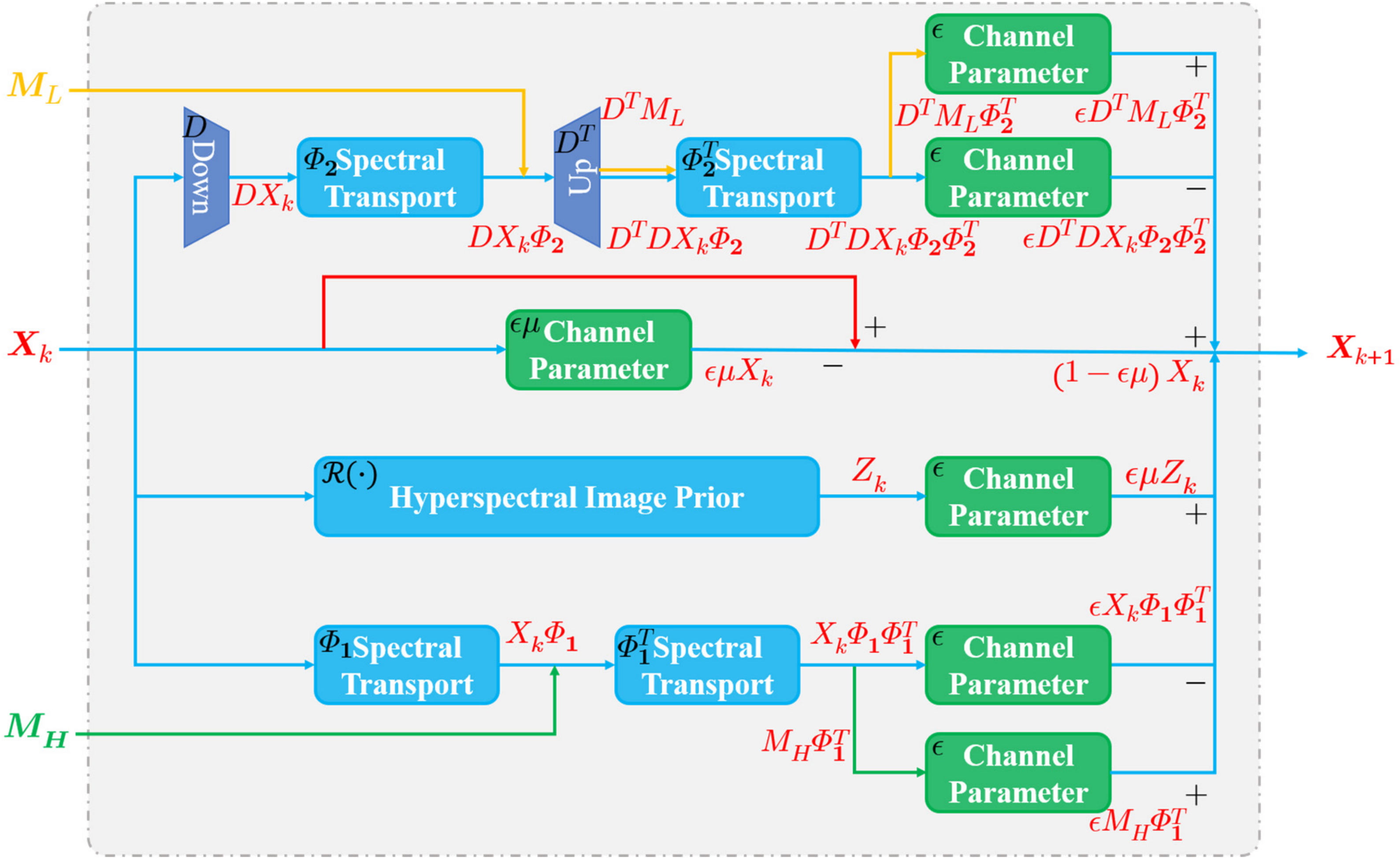
$$X_{k+1}^{\downarrow\uparrow} = -\epsilon X_k \Phi_1 \Phi_1^T - \epsilon D^T D X_k \Phi_2 \Phi_2^T$$

$$M_{k+1}^{\uparrow} = \epsilon M_H \Phi_1^T + \epsilon D^T M_L \Phi_2^T$$

- **Deep unrolling**
  - **X subproblem**

$$\hat{X}_{k+1} = (1-\epsilon\mu)X_k + X_{k+1}^{\downarrow\uparrow} + M_{k+1}^{\uparrow} + \epsilon\mu Z_k$$

$$X_{k+1}^{\downarrow\uparrow} = -\epsilon X_k \Phi_1 \Phi_1^T - \epsilon D^T D X_k \Phi_2 \Phi_2^T \qquad M_{k+1}^{\uparrow} = \epsilon M_H \Phi_1^T + \epsilon D^T M_L \Phi_2^T$$



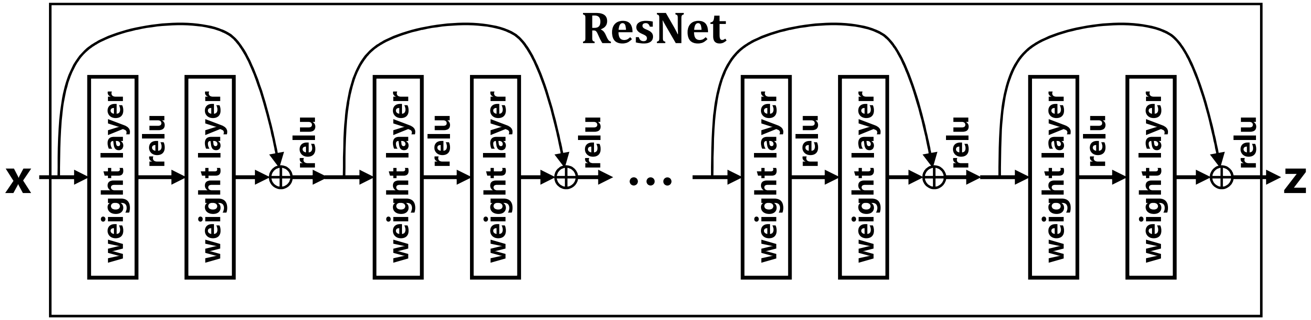
2 Method

- **Deep unrolling**
  - **Z subproblem**

$$\hat{Z}_{k+1} = \mathrm{Prox}(X_{k+1}) = \underset{Z}{\mathrm{argmin}} \|Z - X_{k+1}\|_2^2 + \frac{\lambda}{\mu}\mathcal{R}(Z)$$

$$\begin{aligned}\hat{Z}_{k} &= \mathrm{Prox}(X_{k}) = \underset{Z}{\mathrm{argmin}} \frac{\mu}{\lambda} \|Z - X_{k}\|_2^2 + \mathcal{R}(Z) \\ &= \underset{Z}{\mathrm{argmin}} \frac{1}{2(\sqrt{\lambda/2\mu})^2} \|Z - X_{k}\|_2^2 + \mathcal{R}(Z)\end{aligned}$$

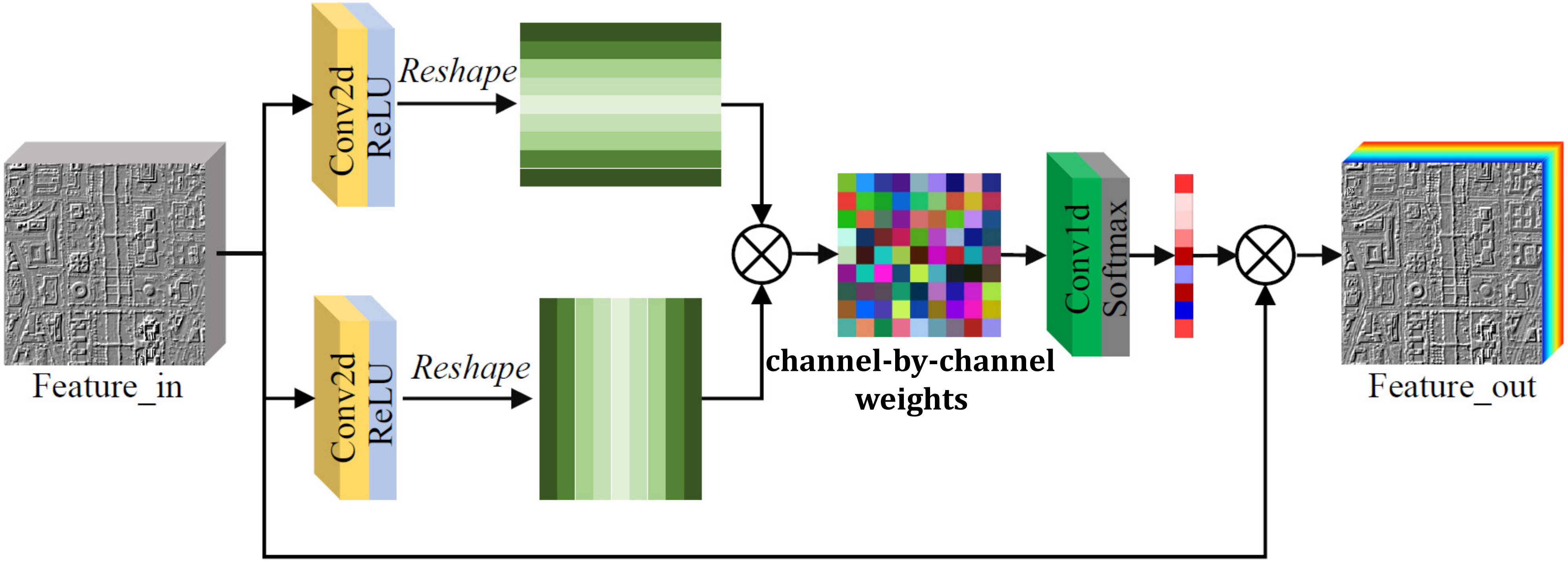
$$\hat{Z}_{k} = \mathrm{Prox}(X_{k}) = \mathrm{Denoiser}(X_{k}) = ResNet(X_{k})$$



Deep unrolling

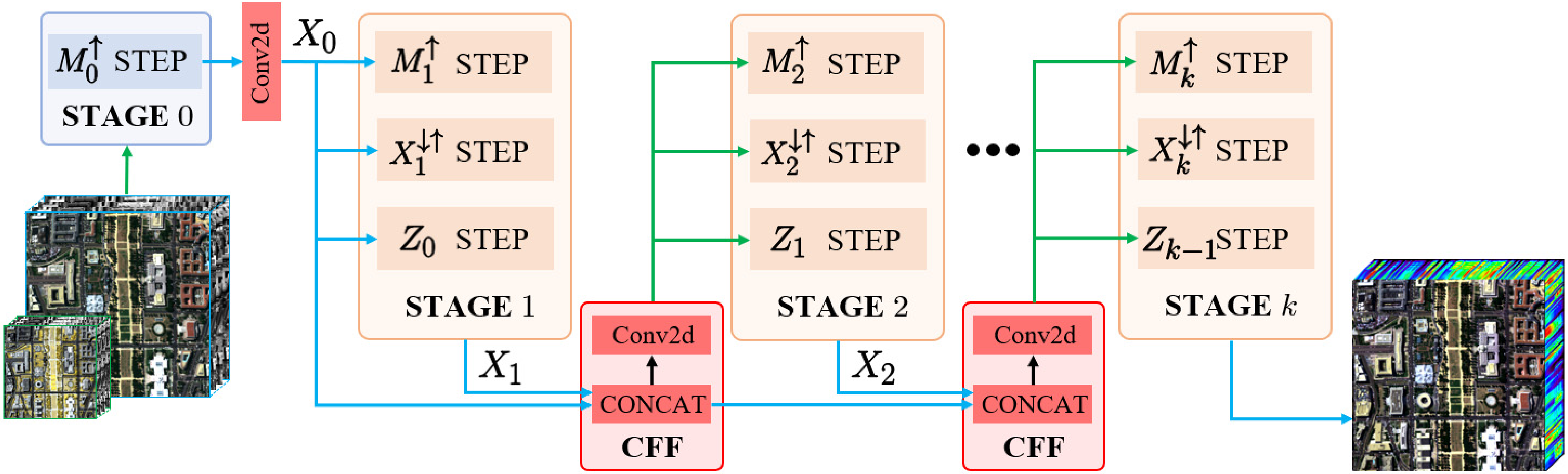
- Cross-Dimensional Channel Parameter Self-learning

More accurate weight extraction ⟷ Faster computational speed



2 Method

Physical Optimization-based Spectral Super-resolution Network (PoNet)



- **Cross-Depth Feature Fusion (CFF)**

$$F_{k-1}^{C} = \mathrm{Concat}(X_0, X_1, X_2, \cdots X_{k-1})$$

$$X_k^{In} = \mathrm{ReLU}(W_{k-1}^{F} * F_{k-1}^{C} + b_{k-1}^{F})$$

Deep and Shallow features

Outline

01 Background

02 Method

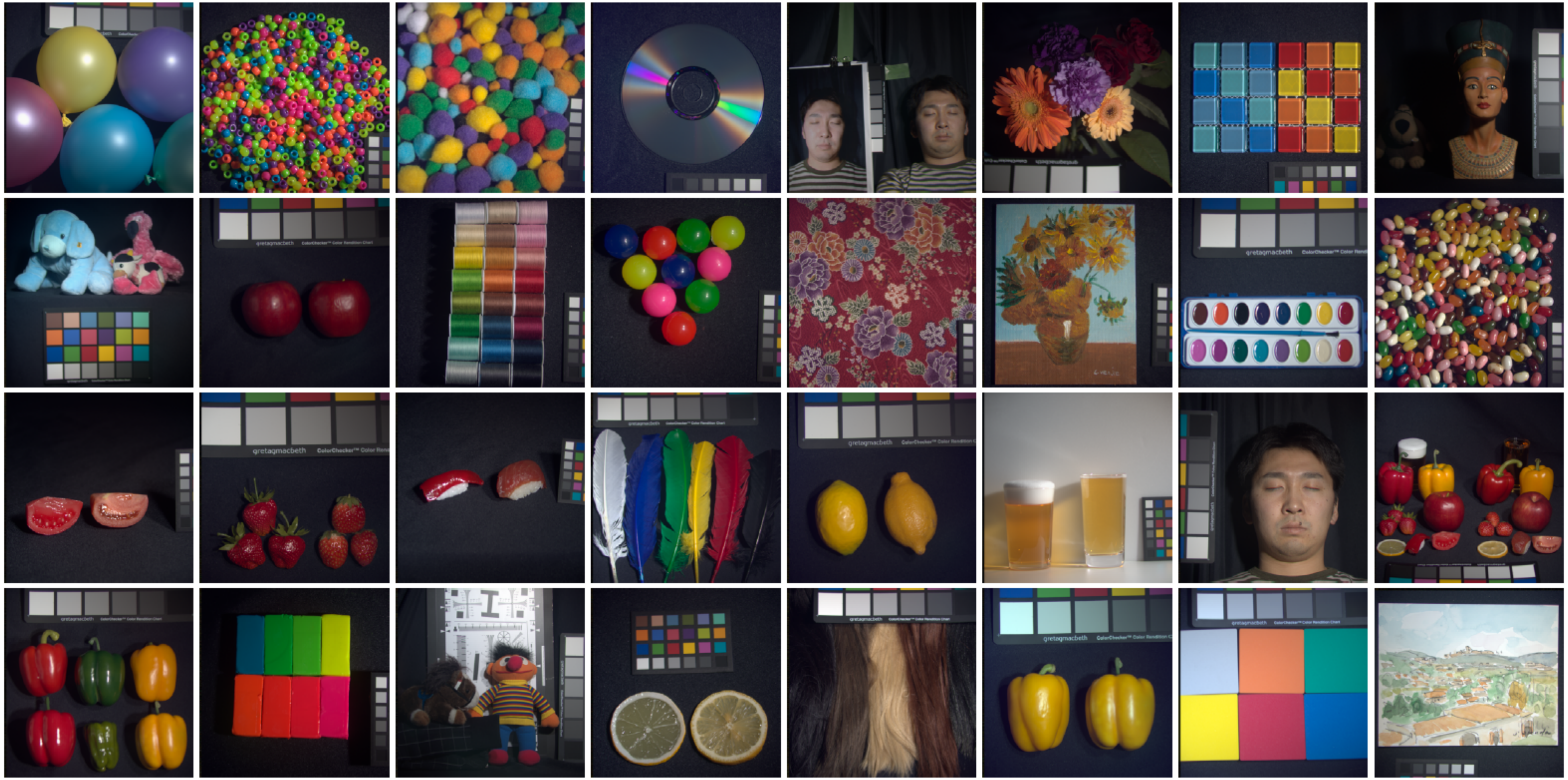
03 Experiment

04 Summary

3 Experiment

☐ Datasets

➢ **RgB2CAVE**



HSIs:
- 400-700nm
- 31 bands

Red:
- 400-550nm
- HR

Green:
- 450-700nm
- LR

Blue:
- 550-700nm
- HR

Datasets

- Sen2CHRIS

Xiongan dataset

Chikusei dataset

DC Mall dataset

HSIs:
- 406-1003nm
- 62 bands

HR MSIs:
- R G B NIR
- 4 bands

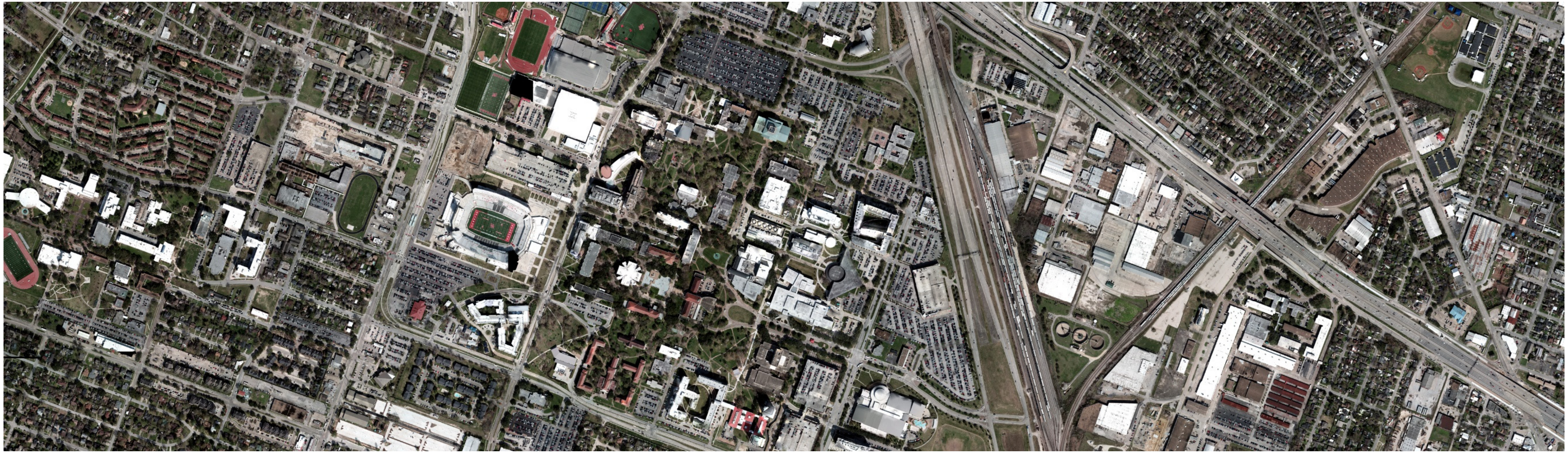
LR MSIs:
- Red Edge1 2 3
- 3 bands

3 Experiment

□ Datasets

➢ **RgB2CASI**

2018 IGARSS Data Fusion Contest



HSIs:

- **380-1050nm**
- **48 bands**

RGB:

- **380-1050nm**
- **LR Green**

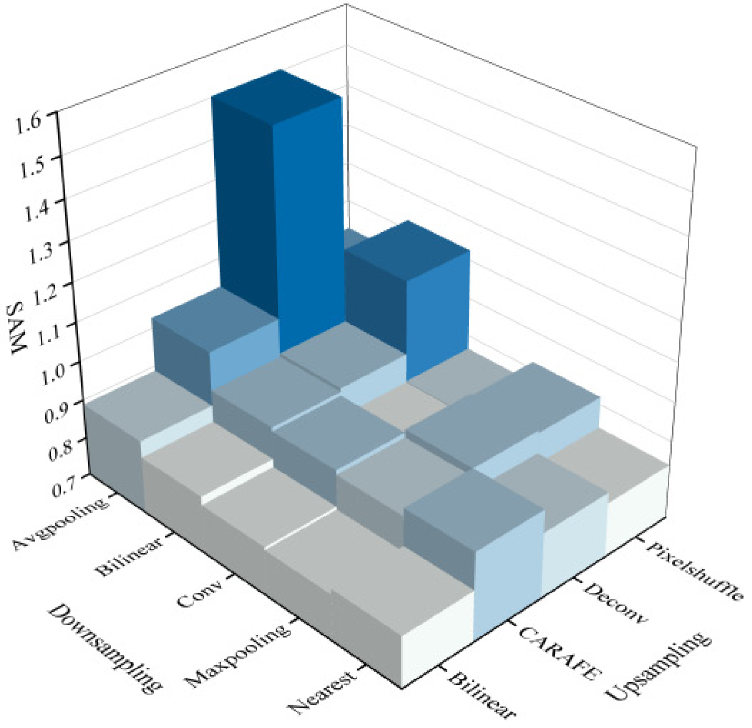
3 Experiment

Influence of Sampling Operator

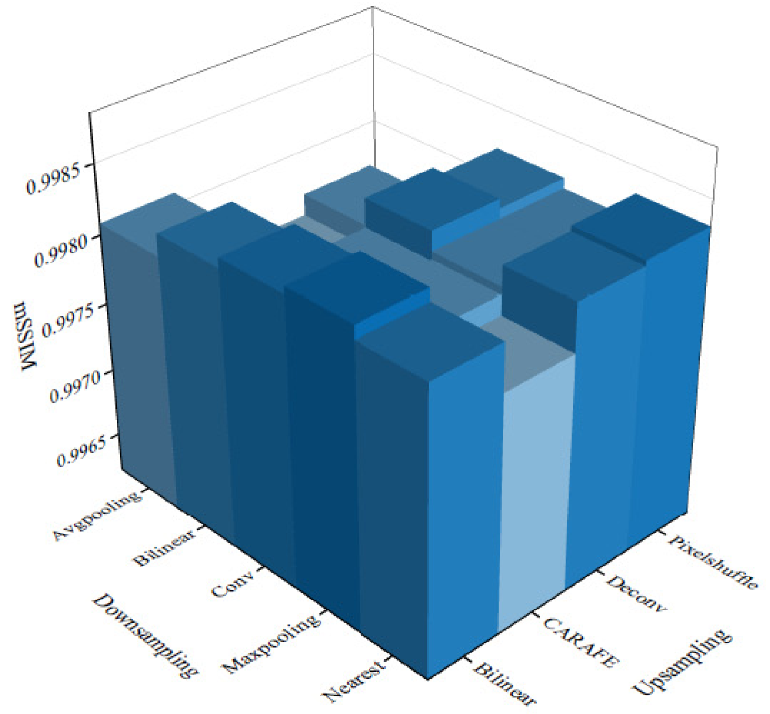
	SAM					mSSIM				
	Avg	Bili	Conv	Max	Near	Avg	Bili	Conv	Max	Near
Bili	**0.8883**	**0.8459**	**0.8241**	**0.8169**	**0.8315**	**0.99808**	**0.99824**	**0.99832**	**0.99837**	0.99827
CARAFE	1.0264	0.9274	0.9177	0.9056	0.9265	0.99742	0.99793	0.99807	0.99808	0.99794
Deconv	1.5033	0.9364	0.8624	0.9260	0.8674	0.99667	0.99808	0.99827	0.99813	0.99828
PixelS	1.0258	1.1251	0.8649	0.9196	0.8396	0.99726	0.99700	0.99821	0.99813	**0.99832**

ERGAS	Avg	Bili	Conv	Max	Near
Bili	**0.9828**	**0.9456**	**0.9307**	**0.9201**	**0.9328**
CARAFE	1.1534	1.0392	0.9971	1.0022	1.0342
Deconv	1.6843	1.0227	0.9897	1.0281	0.9782
PixelS	1.2234	1.2494	0.9734	1.0144	0.9518

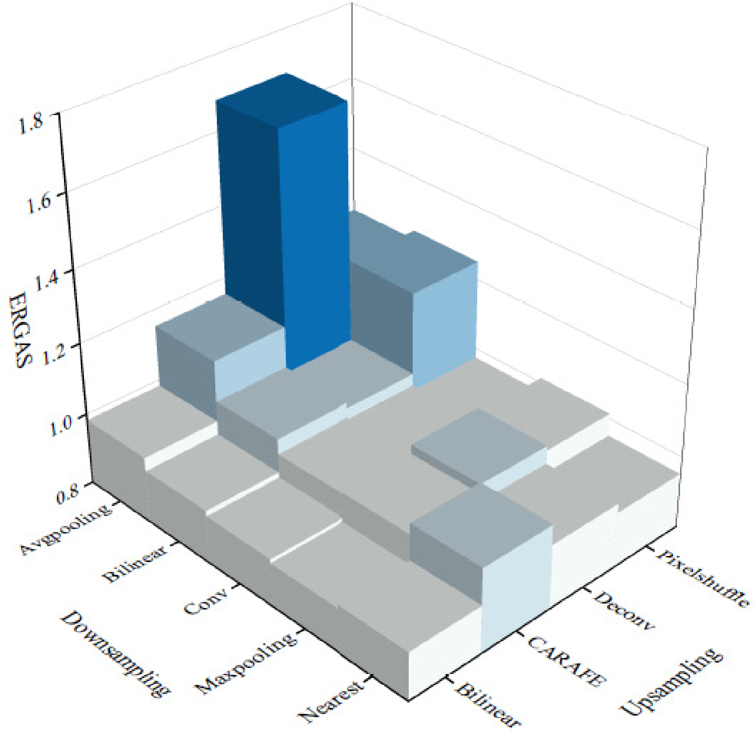
Influence of Sampling Operator



(a) SAM



(b) mSSIM



(c) ERGAS

Min Max

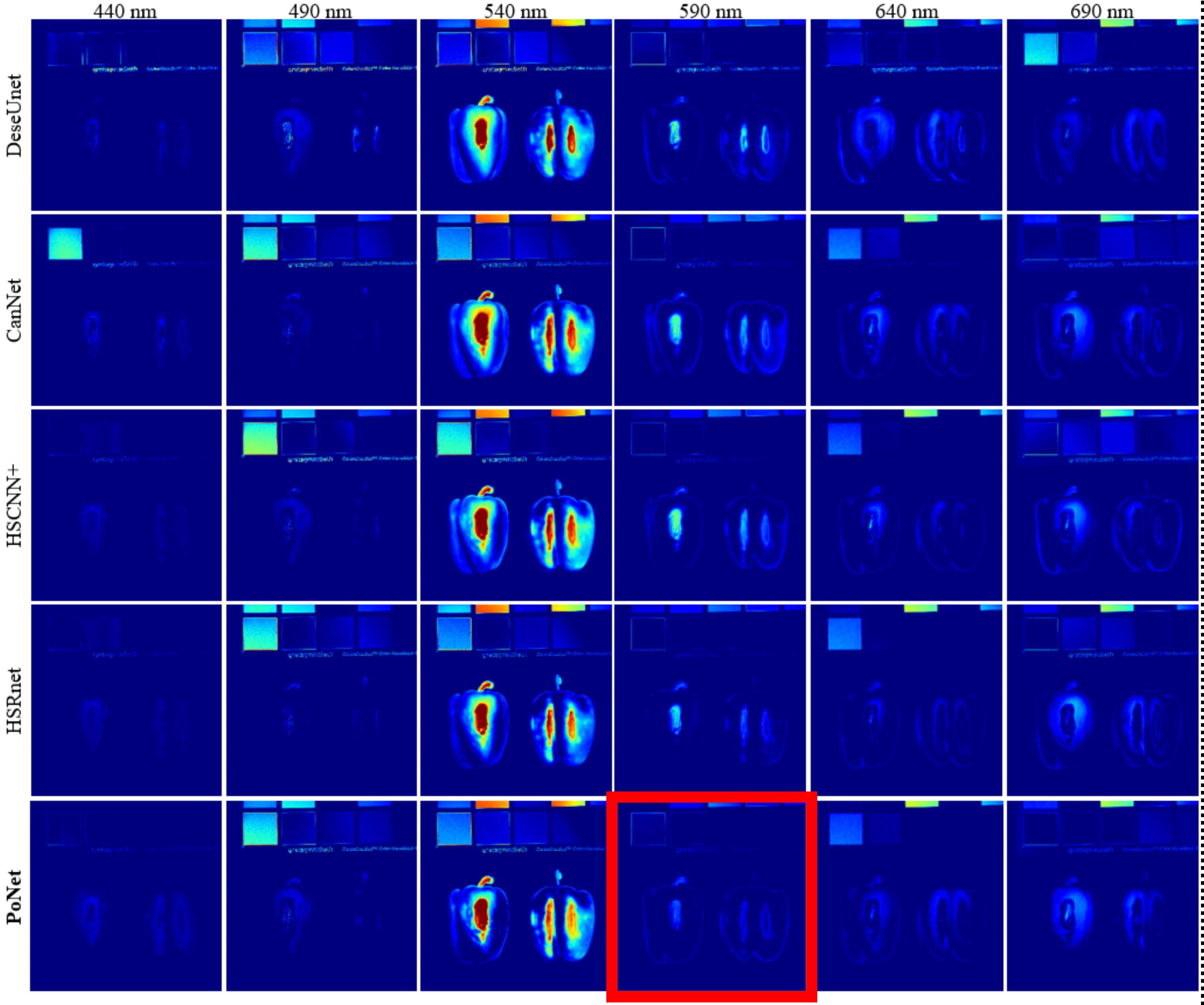
RgB2CAVE

- **Quantitative results**

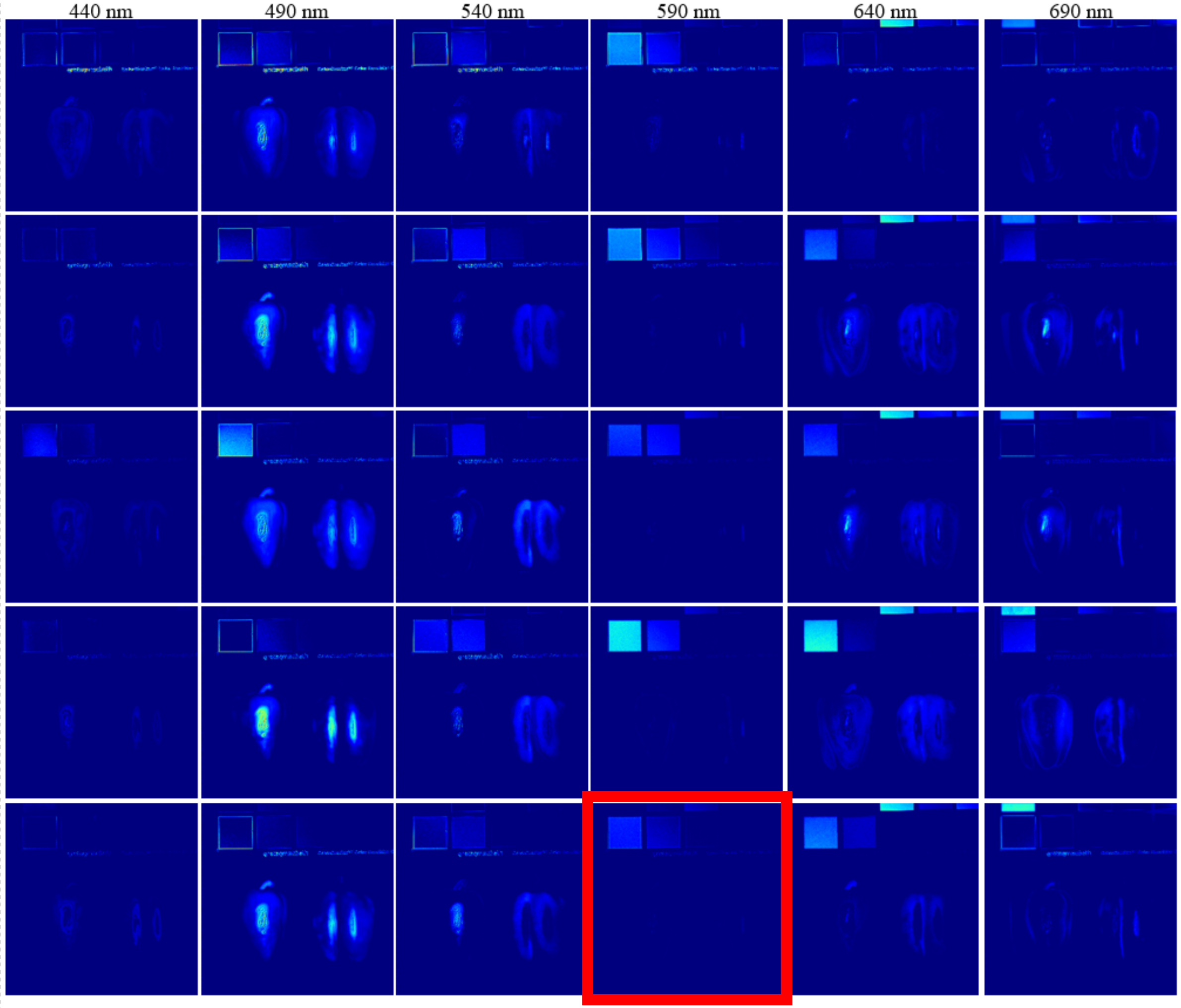
	SSR				FusSR				
	CC	mPSNR	mSSIM	SAM	CC	mPSNR	mSSIM	SAM	ERGAS
DenseU	0.9762	29.2668	0.9336	11.4231	0.9911	32.0511	0.9549	8.1459	8.9606
CanNet	0.9742	29.0431	0.9469	11.2296	0.9917	32.9176	0.9695	8.4988	8.6572
HSCNN+	0.9737	30.1048	0.9514	11.3278	0.9920	33.0453	0.9698	8.9057	8.7036
HSRnet	**0.9755**	30.4968	**0.9551**	11.0565	0.9929	33.6890	0.9720	8.0627	8.0939
PoNet	0.9748	**30.5074**	0.9528	**10.8340**	**0.9933**	**34.2215**	**0.9744**	**7.9154**	**7.6652**

RgB2CAVE

➢ SSR



➢ FusSR



Sen2CHRIS

Quantitative results

SSR

Models	Xiongan				Washington DC Mall				Chikusei			
	CC	mPSNR	mSSIM	SAM	CC	mPSNR	mSSIM	SAM	CC	mPSNR	mSSIM	SAM
DenseU	0.9776	31.0294	0.9544	4.3396	0.9882	41.4252	0.9805	2.2813	0.9877	39.0021	0.9773	4.6744
CanNet	0.9945	37.5414	0.9830	2.9026	0.9984	51.4529	0.9979	1.0816	0.9937	43.4565	0.9872	4.3289
HSCNN+	0.9924	35.7680	0.9846	3.4531	0.9984	51.6680	0.9979	1.0519	0.9941	42.6957	0.9880	4.3493
HSRnet	**0.9963**	37.4468	**0.9897**	3.4239	0.9986	52.1912	0.9982	0.9935	**0.9947**	**44.0038**	**0.9889**	**4.2829**
PoNet	0.9952	**39.2589**	0.9879	**2.1122**	**0.9987**	**52.9510**	**0.9984**	**0.9724**	0.9939	43.8969	0.9878	4.3337

FusSR

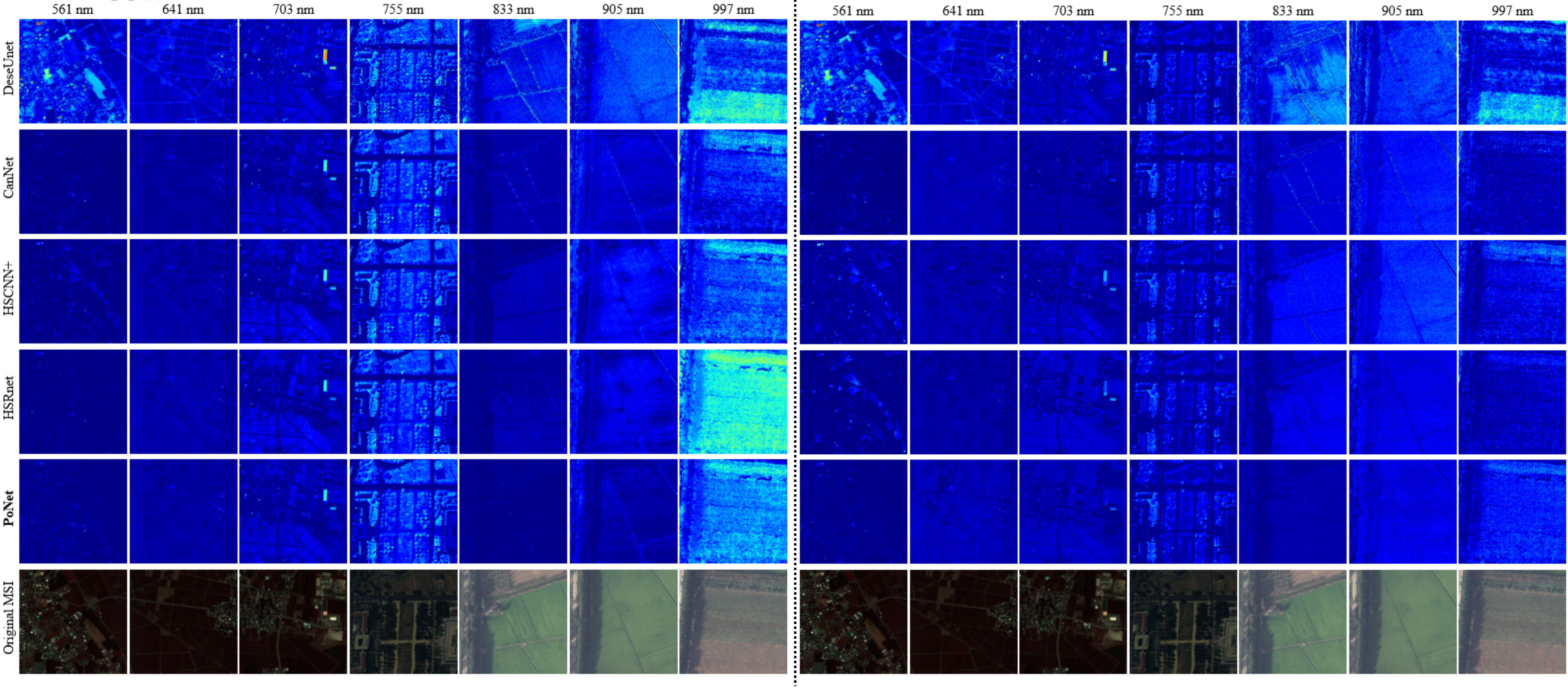
Models	Xiongan					Washington DC Mall					Chikusei				
	CC	mPSNR	mSSIM	SAM	ERGAS	CC	mPSNR	mSSIM	SAM	ERGAS	CC	mPSNR	mSSIM	SAM	ERGAS
DenseU	0.9801	31.2499	0.9557	3.7330	3.5526	0.9883	41.4672	0.9809	2.2485	3.2449	0.9883	39.2332	0.9773	4.6973	8.1293
CanNet	0.9958	39.7155	0.9836	1.6726	1.5151	0.9984	51.3461	0.9979	1.0409	1.1562	0.9961	44.4941	0.9862	4.1710	6.2716
HSCNN+	0.9953	38.2452	0.9889	1.9854	1.8595	0.9981	51.2922	0.9978	1.0624	1.2053	0.9953	42.6871	0.9885	4.2578	6.4048
HSRnet	0.9972	40.2616	0.9924	1.7292	1.5238	0.9987	52.7957	0.9983	0.9238	1.0271	0.9968	44.4520	**0.9902**	4.0105	6.1612
PoNet	**0.9974**	**41.9316**	**0.9928**	**1.4185**	**1.2398**	**0.9989**	**53.7637**	**0.9986**	**0.8865**	**0.9739**	**0.9969**	**45.1265**	0.9901	**3.9189**	**6.0190**

3 Experiment

Sen2CHRIS

SSR

FusSR



3 Experiment

RgB2CASI

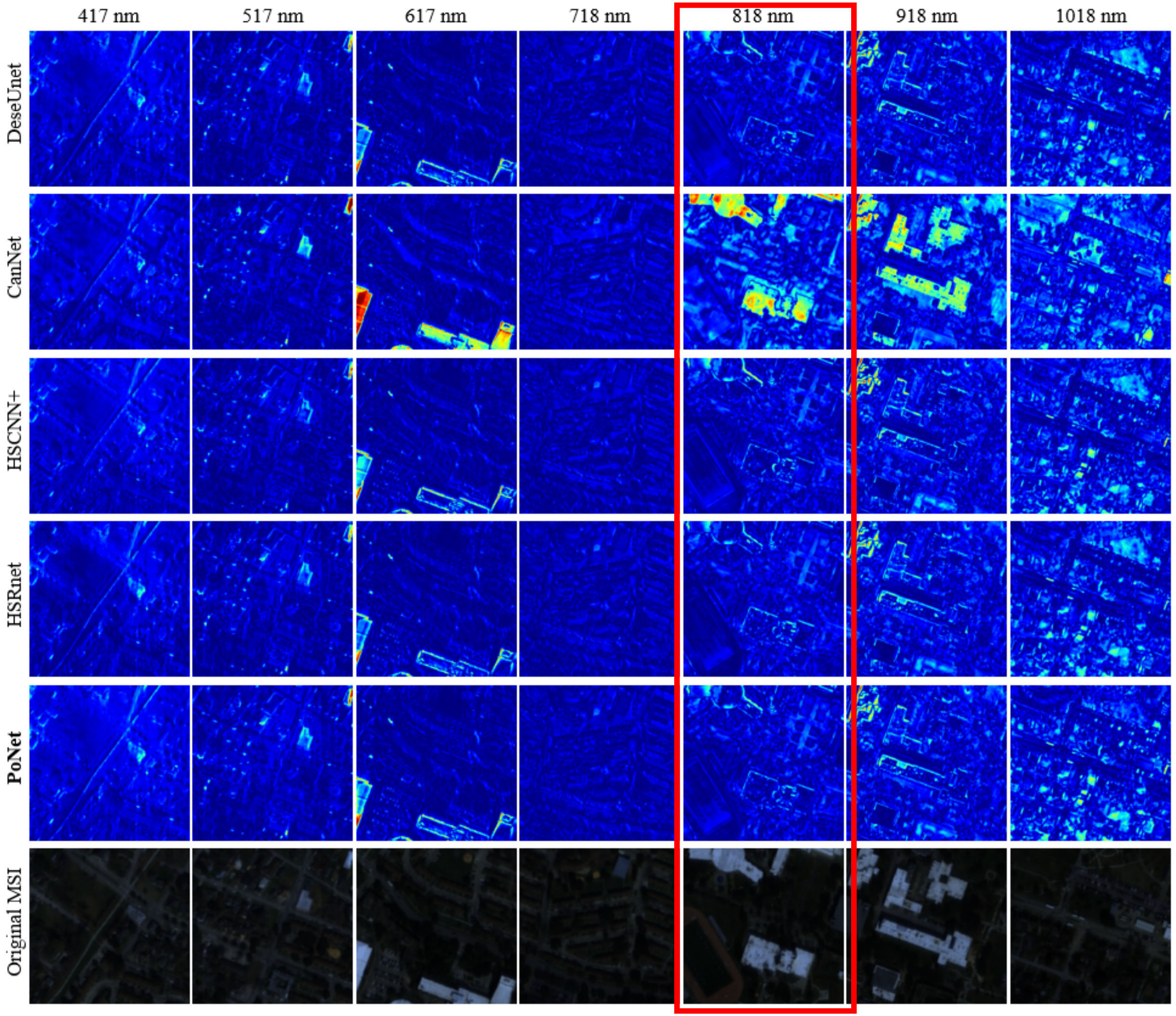
Quantitative results

	SSR				FusSR				
	CC	mPSNR	mSSIM	SAM	CC	mPSNR	mSSIM	SAM	ERGAS
DenseU	**0.8010**	29.7784	<u>0.8424</u>	14.7484	0.8433	30.4219	0.8631	13.7883	11.8762
CanNet	0.6266	27.0725	0.7956	15.8461	0.7775	28.2414	0.8418	14.0059	14.0456
HSCNN+	0.7874	29.5541	0.8395	14.7601	0.8415	30.3498	0.8652	<u>13.7815</u>	<u>11.8079</u>
HSRnet	<u>0.7996</u>	**29.8161**	**0.8446**	<u>14.6122</u>	<u>0.8454</u>	**30.5858**	<u>0.8657</u>	13.8958	11.8192
PoNet	0.7951	<u>29.5439</u>	0.8420	**13.4246**	**0.8479**	<u>30.5309</u>	**0.8683**	**12.5997**	**9.2775**

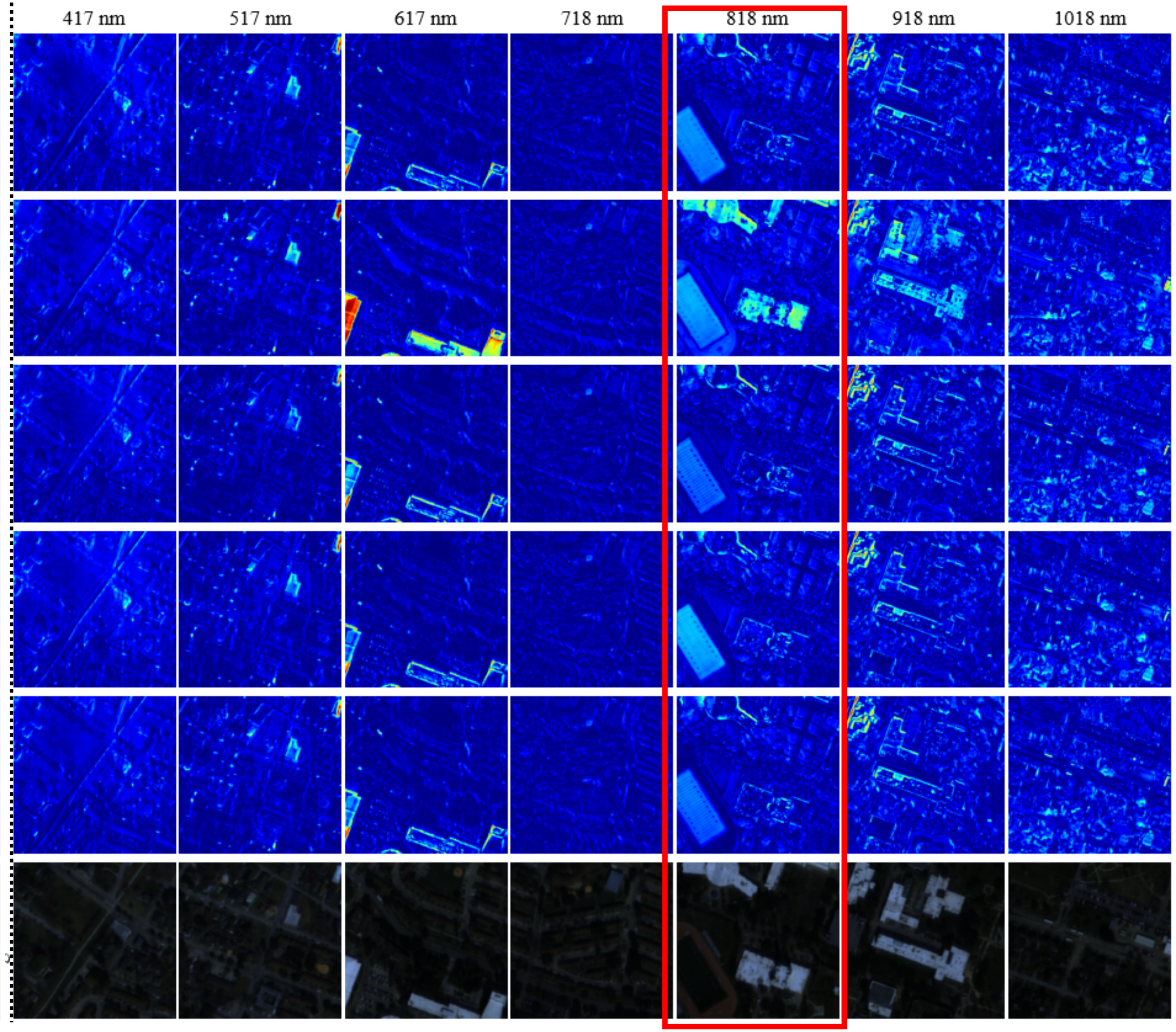
3 Experiment

RgB2CASI

- SSR



- FusSR



PansSR

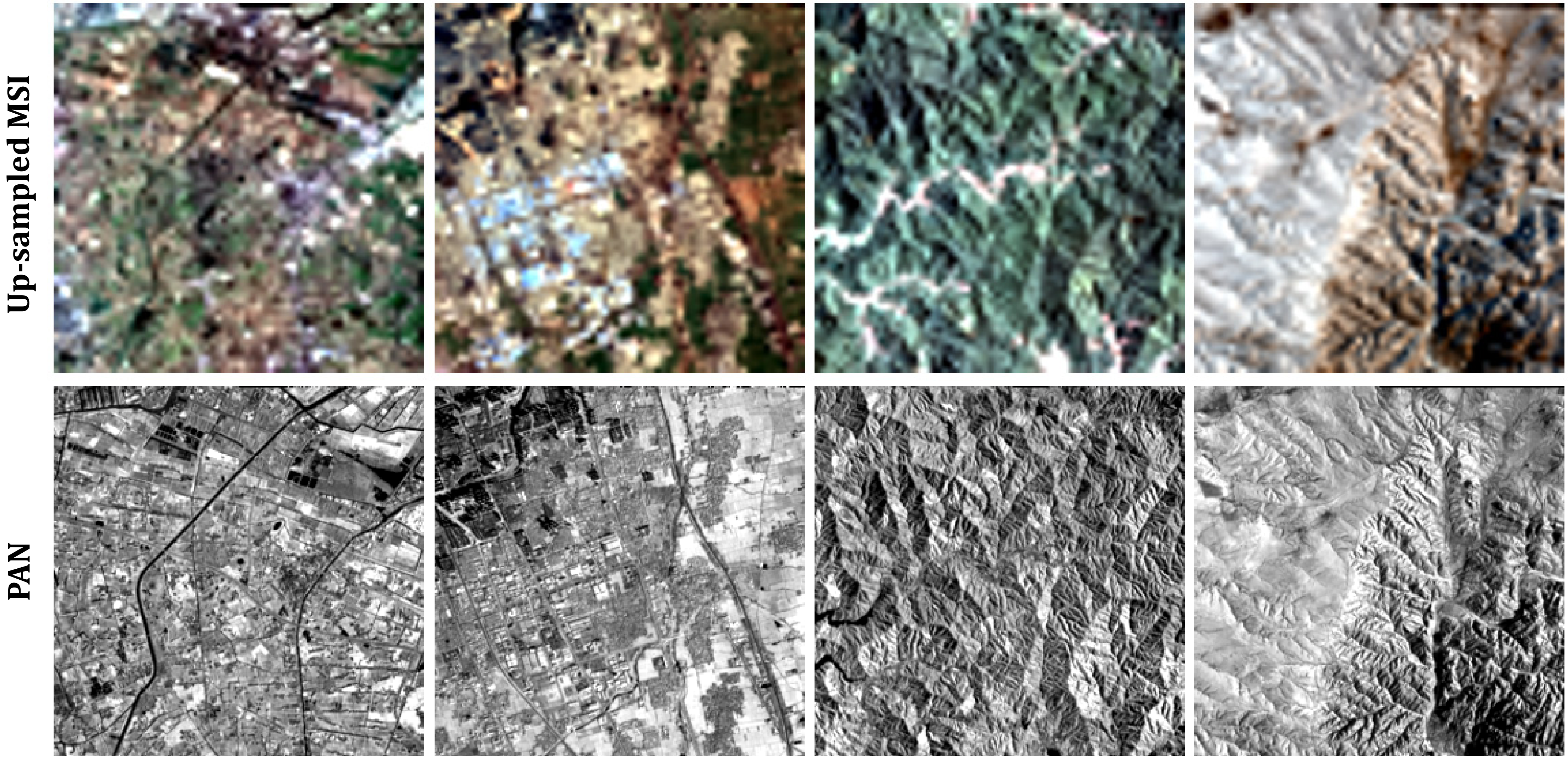
Quantitative results

Images	CC	mPSNR	mSSIM	SAM	ERGAS
1	0.9284	44.2446	0.9828	4.2478	3.3479
2	0.8972	48.5929	0.9914	2.2095	2.0790
3	0.9325	45.8955	0.9849	3.9832	3.0857
4	0.9071	44.3540	0.9522	6.1078	4.6244
5	0.8770	49.3000	0.9937	3.0700	4.9934
6	0.8315	45.0522	0.9870	2.0565	2.8733
7	0.9691	43.9667	0.9767	2.8755	1.8548
AVG	0.9284	44.2446	0.9828	4.2478	3.3479

3 Experiment

PansSR

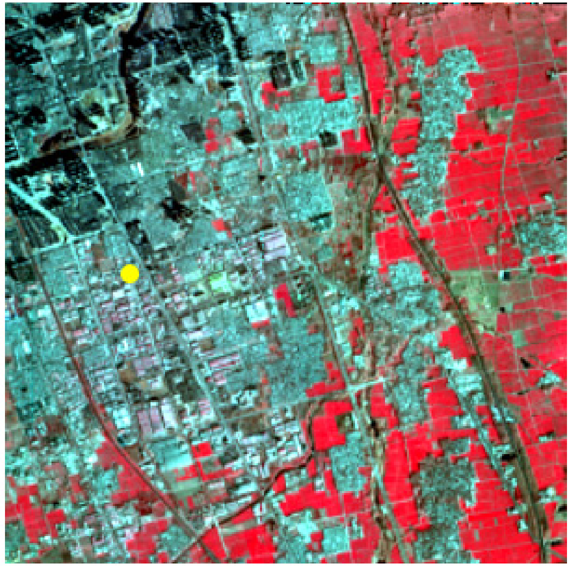
- Input data

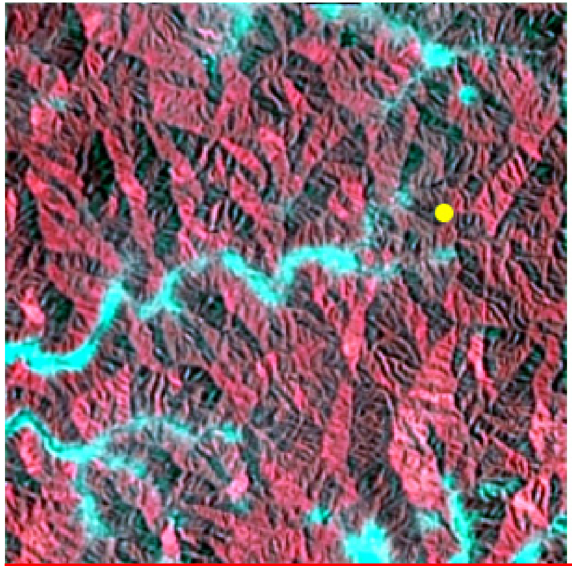


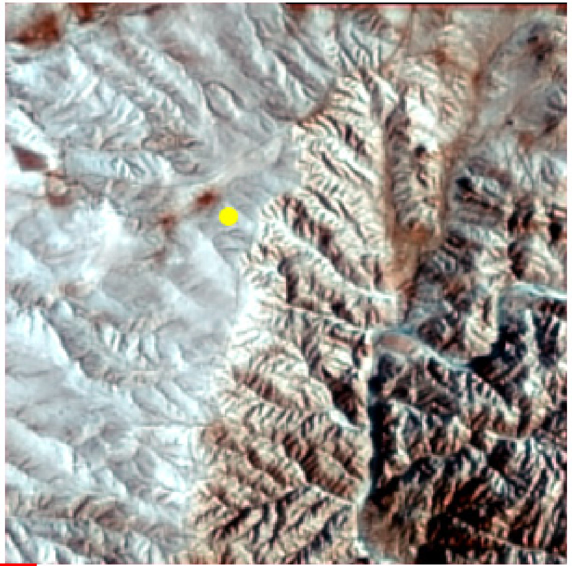
3 Experiment

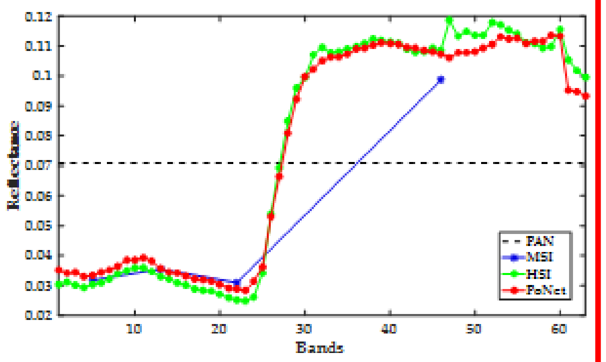
PansSR

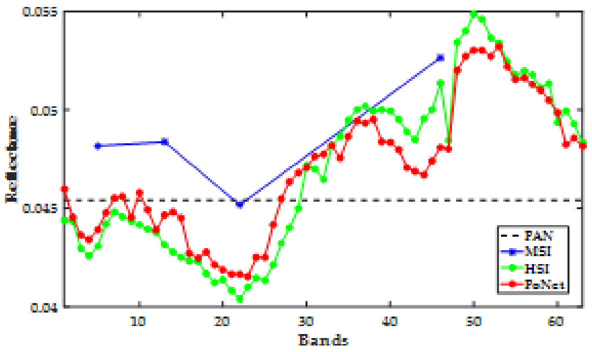
➢ Results

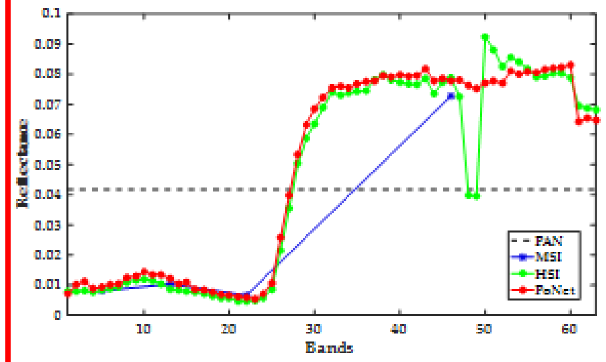


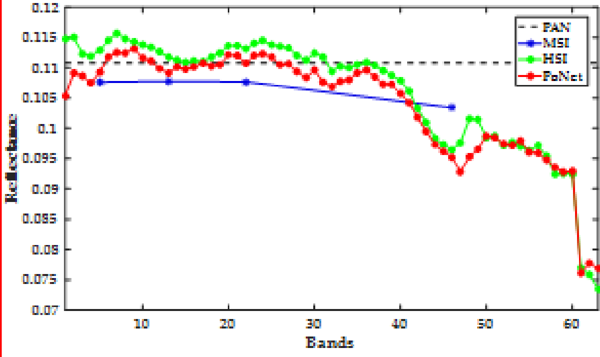








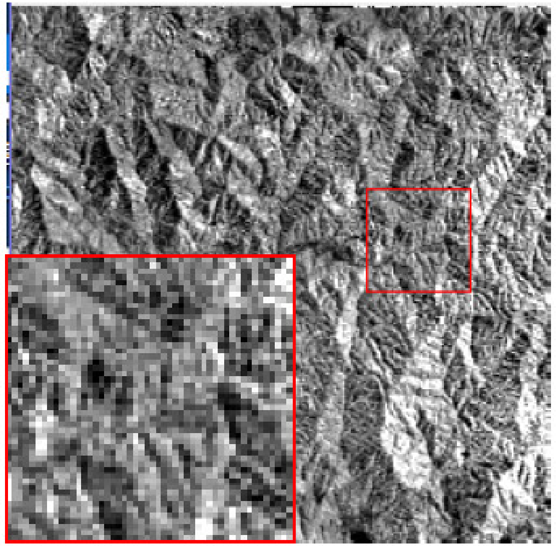




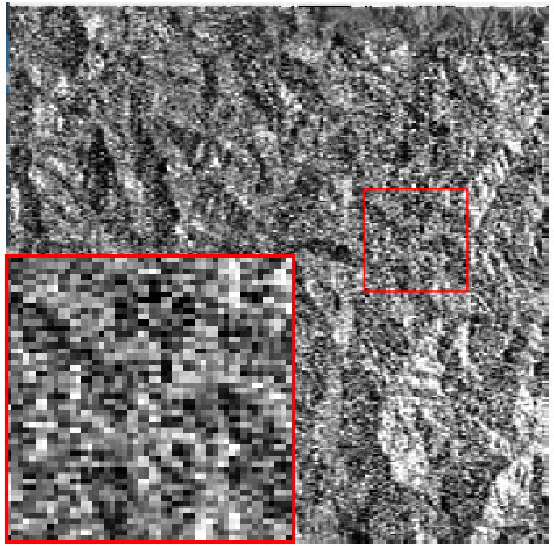
3 Experiment

□ PansSR

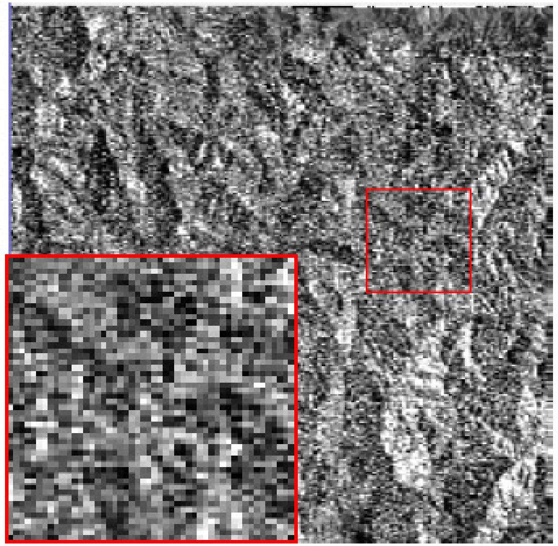
➢ Ground Truth



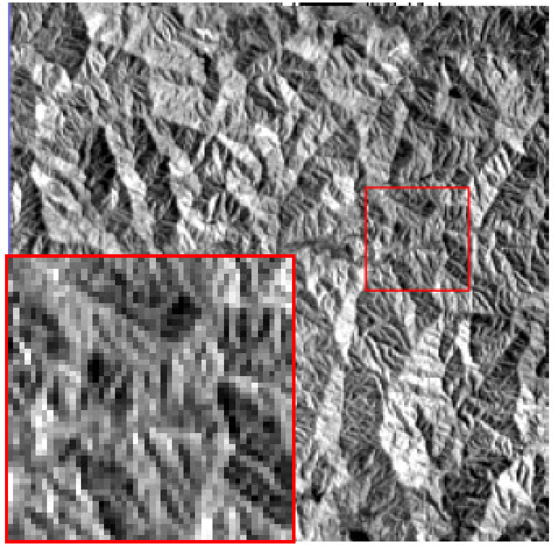
(a): Band 47 of ground truth



(b): Band 48 of ground truth



(c): Band 49 of ground truth



(d): Band 50 of ground truth

3 Experiment

☐ Ablation Study

Models	Optimization Stage	Down and Up	CDCA	CDFF	CC	mPSNR	mSSIM	SAM	ERGAS
HSCNN+	×	×	×	×	0.9920	33.0453	0.9698	8.9057	8.7036
HSRnet	√	×	×	×	0.9929	33.6890	0.9720	8.0627	8.0939
PoNet w/o CDCA	√	√	×	×	0.9930	33.8864	0.9734	8.1364	8.0202
PoNet w/o CDFF	√	√	√	×	0.9931	34.0456	0.9743	7.9463	7.7565
PoNet	√	√	√	√	**0.9933**	**34.2215**	**0.9744**	**7.9154**	**7.6652**

Outline

4 Summary

- This paper defines the generalized spectral super-resolution for arbitrary multispectral images including data with multiple spatial resolutions, namely, SSR, FusSR, and PansSR.

- Unfolding optimization algorithm considering physical degradation to deep learning gives CNN the important physical interpretability, which provides a great help to recover hyperspectral information. Besides, to learn parameters channel-to-channel adaptively, as well as boost computation, cross-dimensional channel attention is proposed. We also employed both deep and shallow features to perform better spectral enhancement as well as good spatial fidelity.

- Datasets involving from natural images to multi-scale remote sensing images, namely, RgB2CAVE, Sen2CHRIS, RgB2CASI, and GF2Hyper, to evaluate the method performance in three multispectral acquisitions are built. Quantitative and visual comparisons proved the superiority of the proposed PoNet. Furthermore, sampling operator discussion and ablation study are also shown to verify the effectiveness of each strategy.

Thanks!

2021.11.13

Homepage

https://jianghe96.github.io/

Reporter: Jiang He

WUHAN UNIVERSITY