SSconv: Explicit Spectral-to-Spatial Convolution for Pansharpening

Yudong Wang, Liang-Jian Deng, Tian-Jing Zhang, Xiao Wu
University of Electronic Science and Technology of China, Chengdu, China

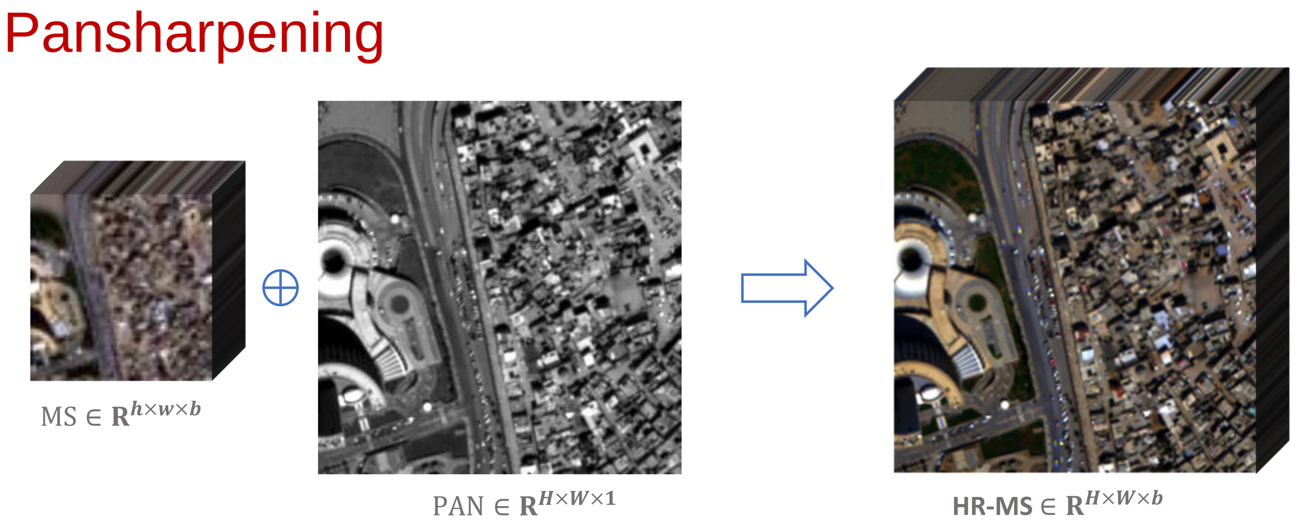
ACM multimedia
Chengdu, China OCT 20-24 2021

Introduction:

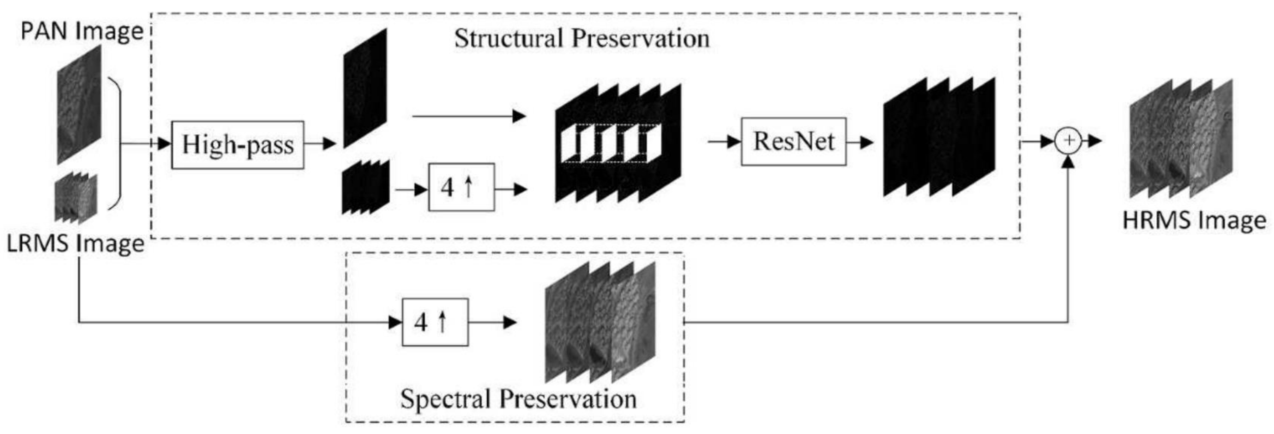
Pansharpening aims to fuse a high spatial resolution panchromatic (PAN) image and a low resolution multispectral (LR-MS) image to obtain a multispectral image with the same spatial resolution as the PAN image. Thanks to the flexible structure of convolution neural networks (CNNs), they have been successfully applied to the problem of pansharpening. However, most of the existing methods only simply feed the up-sampled LR-MS into the CNNs and ignore the spatial distortion caused by direct up-sampling. In this paper, we propose an explicit spectral-to-spatial convolution (SSconv) that aggregates spectral features into the spatial domain to perform the up-sampling operation, which can get better performance than the direct up-sampling. Furthermore, SSconv is embedded into a multiscale U-shaped convolution neural network (MUCNN) for fully utilizing the multispectral information of involved images. In particular, multiscale injection branch and mixed loss on crossscale levels are employed to fuse pixel-wise image information. Benefiting from the distortion-free property of SSconv, the proposed MUCNN can generate state-of-the-art performance with a simple structure, both on reduced-resolution and full-resolution datasets acquired from WorldView-3 and GaoFen-2

Introduction:

Pansharpening



CNN-based Method for Pansharpening, i.e., PanNet[1]

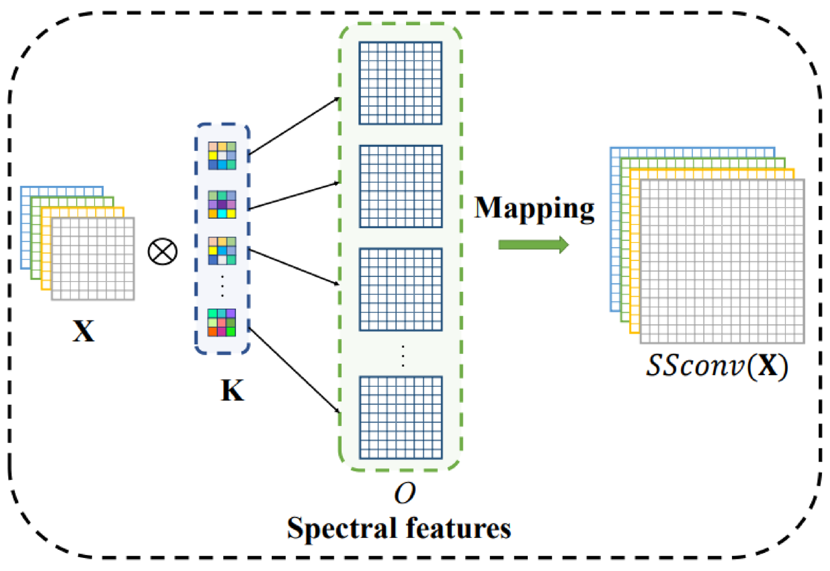


More development history of methods for Pan-sharpening can be find from [2]

[1] Junfeng Yang, Xueyang Fu, Yuwen Hu, Yue Huang, Xinghao Ding, and John Paisley. 2017. PanNet: A deep network architecture for pan-sharpening. In Proceedings of the IEEE international conference on computer vision. 5449–5457.
[2] Vivone, G., et al. "A New Benchmark Based on Recent Advances in Multispectral Pansharpening: Revisiting pansharpening with classical and emerging pansharpening methods. IEEE Geoscience and Remote Sensing Magazine.

Methods:

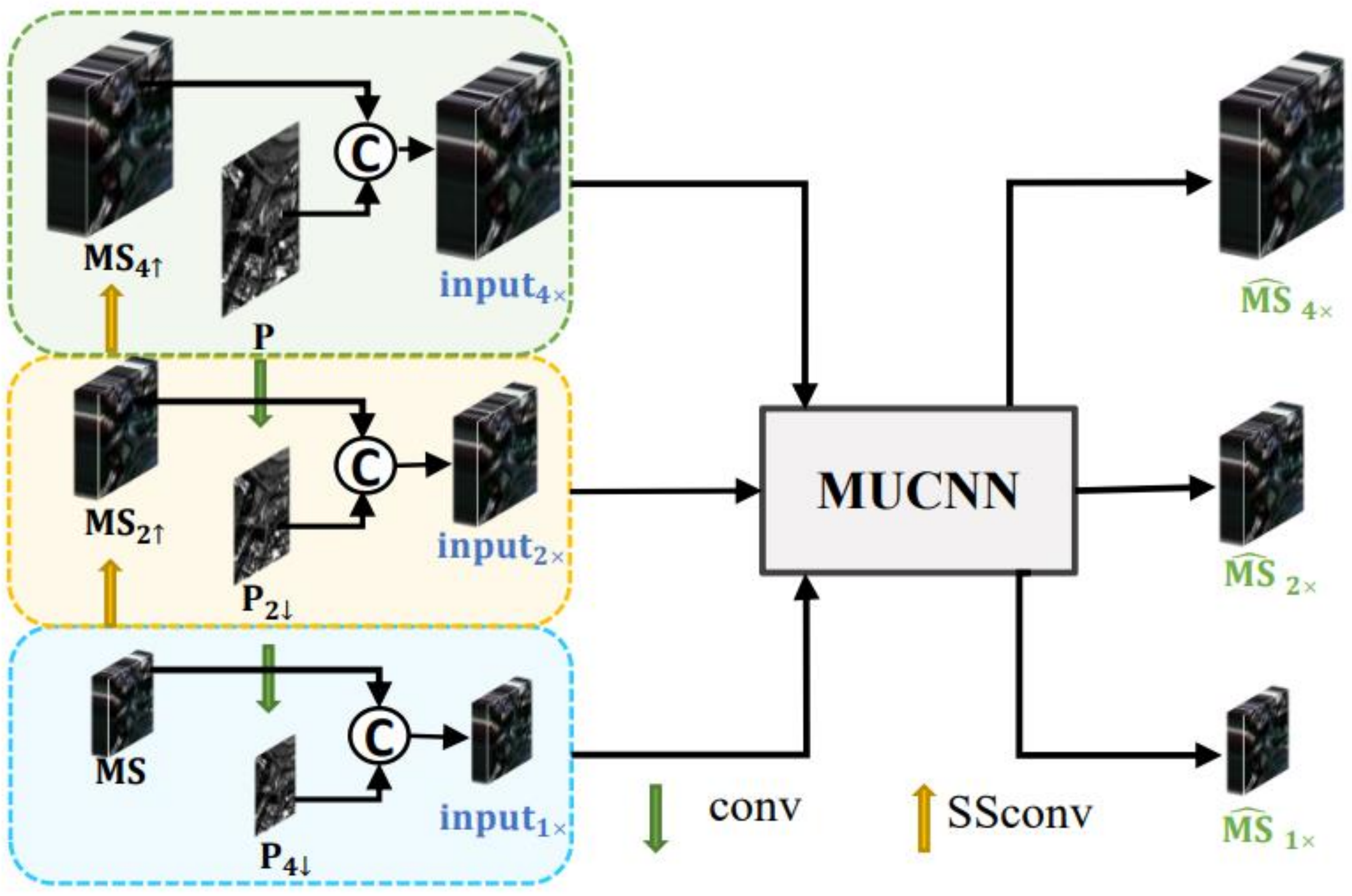
Spectral to Spatial Convolution (SSconv)



Mapping

$$SSconv(\mathbf{X})_{ri+c_1,rj+c_2,k} = O_{i,j,kr^2+c_1r+c_2}$$
$$(i = 0,1,\dots,w-1; j = 0,1,\dots,h-1;$$
$$c_1 = 0,1,\dots,r-1; c_2 = 0,1,\dots,r-1;$$
$$k = 0,1,\dots,b-1)$$

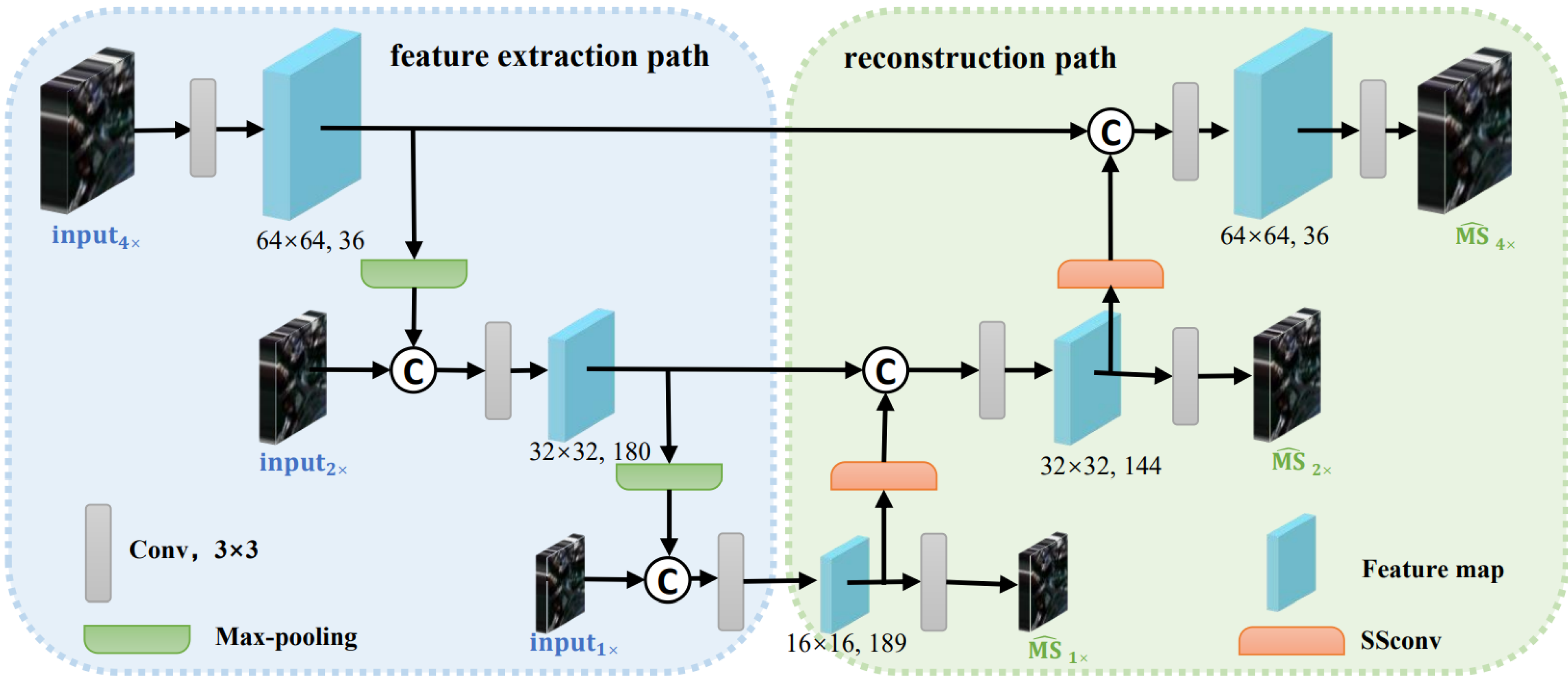
Multiscale injection branch and Mixed multiscale loss



Loss function

$$Loss(\Theta_{MUCNN}) = \frac{\lambda_1}{b}\sum_{i=1}^{b}\|\mathbf{GT}-\widehat{\mathbf{MS}}_{4\times}\|_F^2 + \frac{\lambda_2}{b}\sum_{i=1}^{b}\|\mathbf{GT}_{2\downarrow}-\widehat{\mathbf{MS}}_{2\times}\|_F^2 + \frac{\lambda_3}{b}\sum_{i=1}^{b}\|\mathbf{GT}_{4\downarrow}-\widehat{\mathbf{MS}}_{1\times}\|_F^2$$

Multiscale U-shaped Convolution Neural Network (MUCNN)



Experiments:

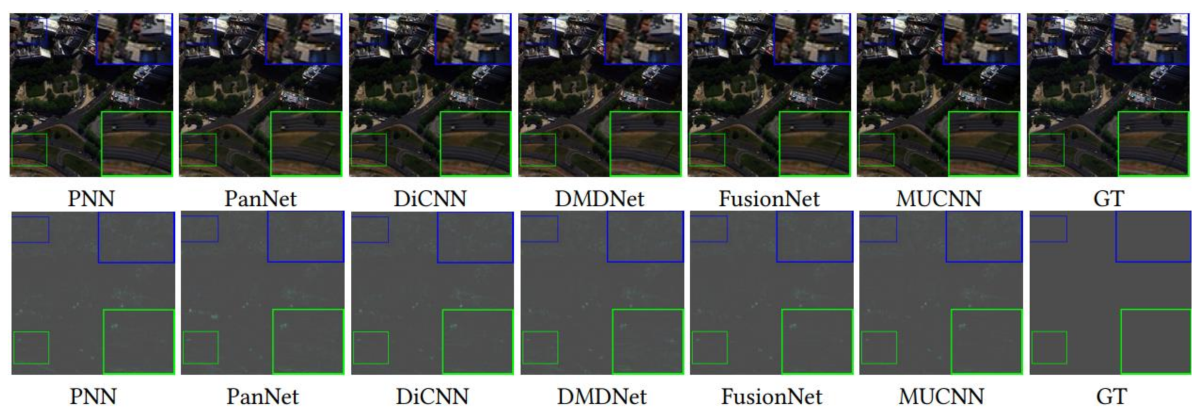
Datasets:

- 8-band data: WorldView-3 (WV3)
  1) reduced-resolution examples
  2) full-resolution examples
- 4-band data: GaoFen2 (GF2)

Metrics：

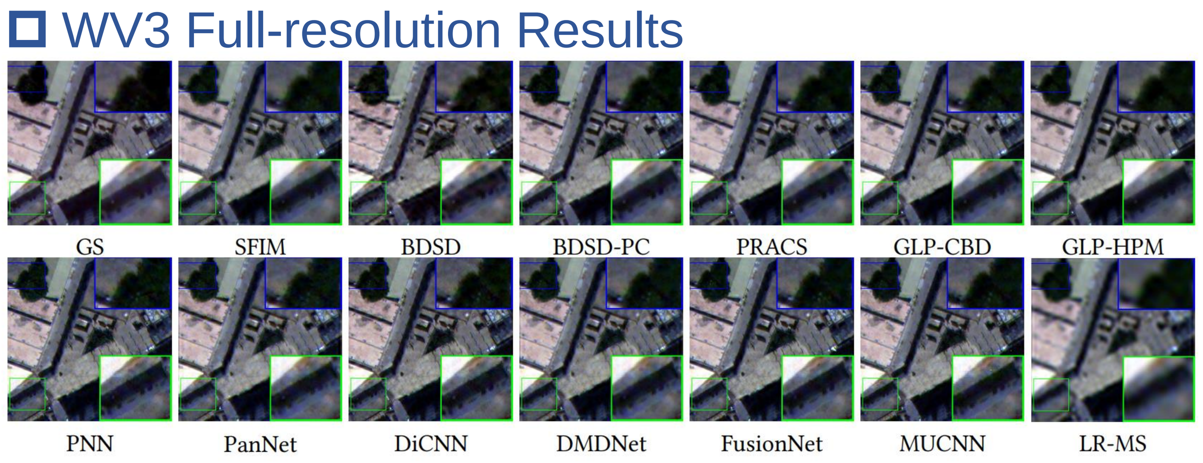
- Reduced-resolution:
  SAM, ERGAS, SCC, Q8
- Full-resolution:
  QNR, D_λ, D_S

WV3 Reduced-resolution Results



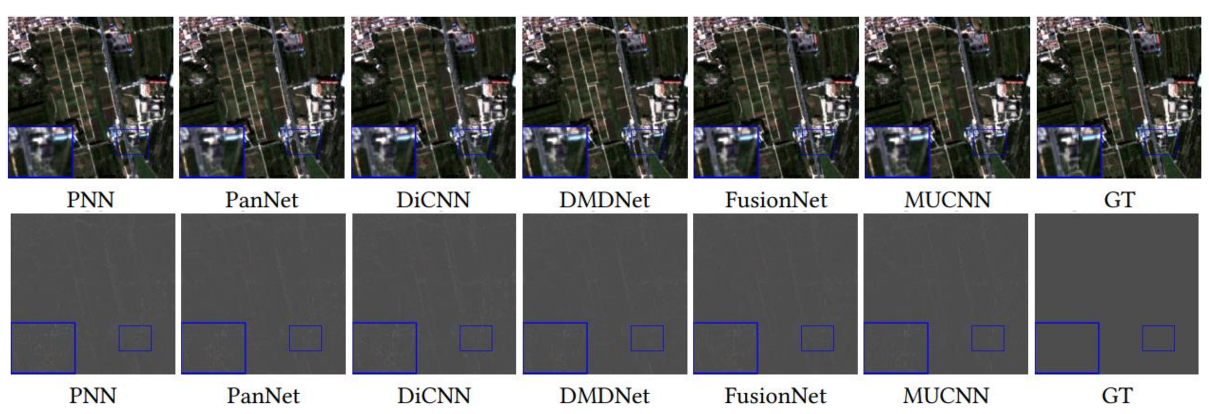
Method	*SAM*	*ERGAS*	*SCC*	*Q8*	*QAVE*
GS [18]	5.698 ± 2.008	5.282 ± 2.187	0.873 ± 0.071	0.766 ± 0.139	0.768 ± 0.146
SFIM [19]	5.452 ± 1.903	5.200 ± 6.574	0.866 ± 0.067	0.798 ± 0.122	0.811 ± 0.130
BDSD [11]	7.000 ± 2.853	5.167 ± 2.248	0.871 ± 0.080	0.813 ± 0.123	0.817 ± 0.126
BDSD-PC [28]	5.425 ± 1.972	4.246 ± 1.860	0.891 ± 0.069	0.853 ± 0.116	0.852 ± 0.124
PRACS [5]	5.286 ± 1.958	4.163 ± 1.775	0.890 ± 0.070	0.854 ± 0.114	0.849 ± 0.123
GLP-CBD [4]	5.286 ± 1.958	4.163 ± 1.775	0.890 ± 0.070	0.854 ± 0.114	0.849 ± 0.123
GLP-HPM [32]	5.604 ± 1.974	4.764 ± 1.935	0.873 ± 0.065	0.817 ± 0.128	0.810 ± 0.135
PNN [21]	4.002 ± 1.329	2.728 ± 1.004	0.952 ± 0.046	0.908 ± 0.112	0.911 ± 0.114
PanNet [36]	4.092 ± 1.273	2.952 ± 0.978	0.949 ± 0.046	0.894 ± 0.117	0.907 ± 0.118
DiCNN [13]	3.981 ± 1.318	2.737 ± 1.016	0.952 ± 0.046	0.910 ± 0.112	0.911 ± 0.115
DMDNet [9]	3.971 ± 1.248	2.857 ± 0.966	0.953 ± 0.045	0.900 ± 0.114	0.913 ± 0.115
FusionNet [8]	3.744 ± 1.226	2.568 ± 0.994	0.958 ± 0.045	0.914 ± 0.112	0.914 ± 0.117
MUCNN	**3.495 ± 1.254**	**2.425 ± 0.956**	**0.963 ± 0.044**	**0.923 ± 0.109**	**0.921 ± 0.114**
Ideal value	**0**	**0**	**1**	**1**	**1**

WV3 Full-resolution Results



Method	*QNR*	D_λ	D_s
GS [18]	0.9026 ± 0.0453	0.0172 ± 0.0195	0.0821 ± 0.0322
SFIM [19]	0.9346 ± 0.0453	0.0216 ± 0.0210	0.0452 ± 0.0212
BDSD [11]	0.9352 ± 0.0389	0.0171 ± 0.0116	0.0488 ± 0.0309
BDSD-PC [28]	0.9166 ± 0.0495	0.0193 ± 0.0190	0.0660 ± 0.0357
PRACS [5]	0.9149 ± 0.0448	0.0174 ± 0.0165	0.0694 ± 0.0329
GLP-CBD [4]	0.9195 ± 0.0504	0.0278 ± 0.0242	0.0550 ± 0.0321
GLP-HPM [32]	0.8939 ± 0.0621	0.0470 ± 0.0322	0.0633 ± 0.0380
PNN [21]	0.9591 ± 0.0260	0.0163 ± 0.0149	0.0251 ± 0.0139
PanNet [36]	0.9581 ± 0.0199	0.0224 ± 0.0108	**0.0201 ± 0.0111**
DiCNN [13]	0.9460 ± 0.0325	0.0165 ± 0.0160	0.0385 ± 0.0201
DMDNet [9]	0.9460 ± 0.0196	0.0187 ± 0.0093	0.0214 ± 0.0122
fusionNet [8]	0.9559 ± 0.0276	0.0178 ± 0.0151	0.0269 ± 0.0161
MUCNN	**0.9629 ± 0.0215**	**0.0128 ± 0.0140**	0.0247 ± 0.0102
Ideal value	**1**	**0**	**0**

GF2 Reduced-resolution Results



Method	*SAM*	*ERGAS*	*SCC*	*Q8*	*QAVE*
GS [18]	2.975 ± 1.111	2.966 ± 1.010	0.852 ± 0.062	0.787 ± 0.076	0.797 ± 0.076
SFIM [19]	2.297 ± 0.637	2.189 ± 0.695	0.861 ± 0.054	0.865 ± 0.040	0.876 ± 0.037
BDSD [11]	2.307 ± 0.669	2.070 ± 0.610	0.877 ± 0.052	0.876 ± 0.042	0.885 ± 0.020
BDSD-PC [28]	2.304 ± 0.643	2.075 ± 0.604	0.878 ± 0.051	0.878 ± 0.040	0.887 ± 0.039
PRACS [5]	2.311 ± 0.597	2.169 ± 0.599	0.867 ± 0.050	0.872 ± 0.035	0.876 ± 0.034
GLP-CBD [4]	2.274 ± 0.733	2.046 ± 0.620	0.873 ± 0.053	0.877 ± 0.041	0.880 ± 0.040
GLP-HPM [32]	0.552 ± 0.777	2.299 ± 0.713	0.867 ± 0.054	0.852 ± 0.045	0.852 ± 0.044
PNN [21]	1.460 ± 0.361	1.271 ± 0.324	0.948 ± 0.021	0.947 ± 0.020	0.949 ± 0.017
PanNet [36]	1.395 ± 0.326	1.224 ± 0.283	0.956 ± 0.012	0.947 ± 0.022	0.957 ± 0.015
DiCNN [13]	1.495 ± 0.381	1.320 ± 0.354	0.946 ± 0.022	0.945 ± 0.021	0.947 ± 0.018
DMDNet [9]	1.297 ± 0.316	1.128 ± 0.267	0.964 ± 0.010	0.953 ± 0.022	0.963 ± 0.014
FusionNet [8]	1.219 ± 0.292	1.037 ± 0.256	0.968 ± 0.010	0.962 ± 0.017	0.964 ± 0.015
MUCNN	**1.100 ± 0.274**	**0.937 ± 0.234**	**0.975 ± 0.008**	**0.970 ± 0.013**	**0.970 ± 0.013**
Ideal value	**0**	**0**	**1**	**1**	**1**

Conclusions:

We proposed the SSconv to better use spectral information to up-sample Multi-Spectral images. Additionally, we proposed Multiscale U-shaped Convolution Neural Network to promote the performance of SSconv. We got the state-of-the-art performances in three Datasets.