Dual-Granularity Semantic Guided Sparse Routing Diffusion Model for General Pansharpening

Supplementary Material

A. Experimental Results on GaoFen-2 Dataset

This supplementary material presents experimental results using images acquired from the GaoFen-2 (GF-2) satellite. These images belong to the PanCollection dataset [1]. The experiments are conducted under both reduced-resolution and full-resolution configurations [6].

The quantitative evaluations are shown in Tables 1-2, which reveals that SGDiff outperforms comparison methods in the majority of metrics. In Figure 1, we present the qualitative assessment of the pansharpened images at reduced resolution, where SGDiff effectively captures fine details, particularly along the edges of the waves and the vegetation on the beach. The difference map, predominantly blue, indicates minimal error at reduced resolution. In comparison, BIMPAN [2] and CrossDiff [8] exhibit slight blurring along the white line. Although TMDiff [7] performs better, it still falls short of SGDiff in preserving texture details. Furthermore, the experimental results at full resolution, shown in Figure 2, also emphasize the exceptional spectral and spatial fidelity of SGDiff.

Table 1. Evaluation indexes of different methods on GaoFen-2 (GF-2) dataset at reduced resolution.

Method	$Q_8\uparrow$	$SAM\downarrow$	$ERGAS\downarrow$	$SCC\uparrow$
C-GSA [4]	0.8962±0.0306	1.7064±0.3329	1.6492±0.3830	0.9443±0.0176
MTF [5]	0.8904±0.0251	1.6795±0.3559	1.6197±0.3720	0.9386±0.0191
CrossDiff [8]	0.9724±0.0091	0.9369±0.1601	0.8553±0.1332	0.9836±0.0024
LGPConv [9]	0.9658±0.0095	1.0229±0.1929	0.9711±0.1845	0.9790±0.0041
BIMPAN [2]	0.9386±0.0095	1.5001±0.3047	1.2400±0.2102	0.9689±0.0065
U2Net [3]	0.9826±0.0068	0.7354±0.1301	0.6889±0.1142	0.9900 ±0.0018
TMDiff [7]	0.9822±0.0068	0.7253±0.1413	0.6604±0.1225	0.9901±0.0022
Proposed SGDiff	**0.9844±0.0067**	**0.6888±0.1221**	**0.6157±0.0974**	**0.9915±0.0015**

Table 2. Evaluation indexes of different methods on GaoFen-2 (GF-2) dataset at full resolution.

Method	$D_\lambda\downarrow$	$D_s\downarrow$	$HQNR\uparrow$
C-GSA [4]	0.0666±0.0269	0.1397±0.0259	0.8034±0.0416
MTF [5]	0.0346±0.0126	0.1429±0.0277	0.8277±0.0334
CrossDiff [8]	0.0208±0.0114	0.0551±0.0103	0.9252±0.0106
LGPConv [9]	0.0261±0.0153	0.0798±0.0099	0.8962±0.0164
BIMPAN [2]	0.0322±0.0154	0.0512±0.0139	0.9183±0.0188
U2Net [3]	**0.0139±0.0164**	0.0531±0.0196	0.9262±0.01995
TMDiff [7]	0.0194±0.0148	0.0323±0.0054	0.9488±0.0140
Proposed SGDiff	0.0168±0.0101	**0.0240±0.0061**	**0.9596±0.0093**

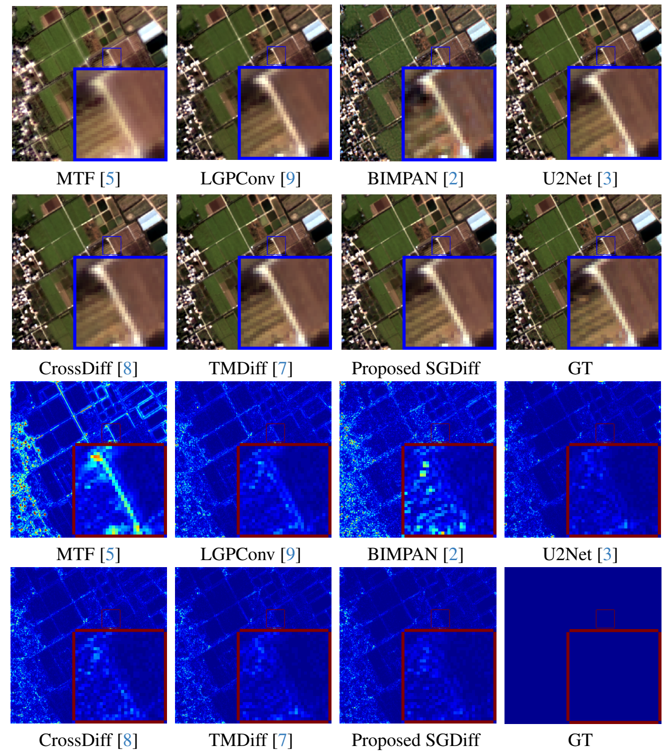

Figure 1. Visual comparison on GaoFen-2 (GF-2) dataset at reduced resolution.

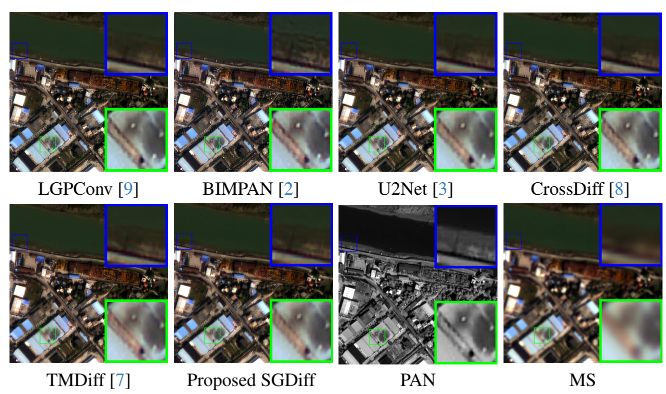

Figure 2. Visual comparison on GaoFen-2 (GF-2) dataset at full resolution.

References

[1] Liang-Jian Deng, Gemine Vivone, Mercedes E Paoletti, Giuseppe Scarpa, Jiang He, Yongjun Zhang, Jocelyn Chanussot, and Antonio Plaza. Machine learning in pansharpening: A benchmark, from shallow to deep networks. *IEEE Geoscience and Remote Sensing Magazine*, 10(3):279–315, 2022. 1

[2] Junming Hou, Qi Cao, Ran Ran, Che Liu, Junling Li, and Liang-jian Deng. Bidomain modeling paradigm for pansharpening. In *Proceedings of the 31st ACM International Conference on Multimedia*, pages 347–357, 2023. 1

[3] Siran Peng, Chenhao Guo, Xiao Wu, and Liang-Jian Deng. U2net: A general framework with spatial-spectral-integrated double u-net for image fusion. In *Proceedings of the 31st ACM International Conference on Multimedia*, page 3219–3227, New York, NY, USA, 2023. Association for Computing Machinery. 1

[4] Rocco Restaino, Mauro Dalla Mura, Gemine Vivone, and Jocelyn Chanussot. Context-adaptive pansharpening based on image segmentation. *IEEE Transactions on Geoscience and Remote Sensing*, 55(2):753–766, 2016. 1

[5] Gemine Vivone, Rocco Restaino, and Jocelyn Chanussot. Full scale regression-based injection coefficients for panchromatic sharpening. *IEEE Transactions on Image Processing*, 27(7): 3418–3431, 2018. 1

[6] Lucien Wald, Thierry Ranchin, and Marc Mangolini. Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images. *Photogrammetric engineering and remote sensing*, 63(6):691–699, 1997. 1

[7] Yinghui Xing, Litao Qu, Shizhou Zhang, Jiapeng Feng, Xiuwei Zhang, and Yanning Zhang. Empower generalizability for pansharpening through text-modulated diffusion model. *IEEE Transactions on Geoscience and Remote Sensing*, 62:1–12, 2024. 1

[8] Yinghui Xing, Litao Qu, Shizhou Zhang, Kai Zhang, Yanning Zhang, and Lorenzo Bruzzone. Crossdiff: Exploring self-supervised representation of pansharpening via cross-predictive diffusion model. *IEEE Transactions on Image Processing*, 2024. 1

[9] Chen-Yu Zhao, Tian-Jing Zhang, Ran Ran, Zhi-Xuan Chen, and Liang-Jian Deng. Lgpconv: learnable gaussian perturbation convolution for lightweight pansharpening. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, pages 4647–4655, 2023. 1