Dynamic Cross Feature Fusion for Remote Sensing Pansharpening

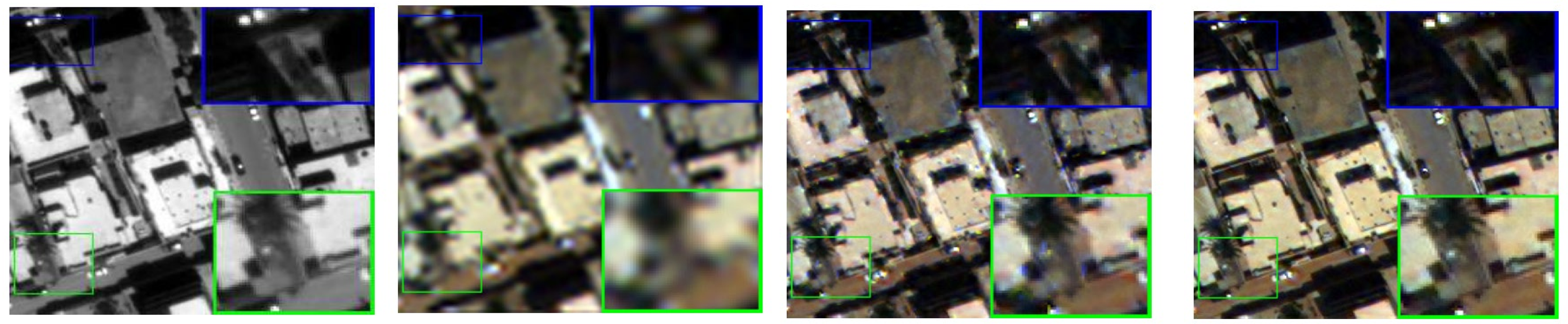
Xiao Wu, Ting-Zhu Huang*, Liang-Jian Deng*, Tian-Jing Zhang

University of Electronic Science and Technology of China, Chendu, China

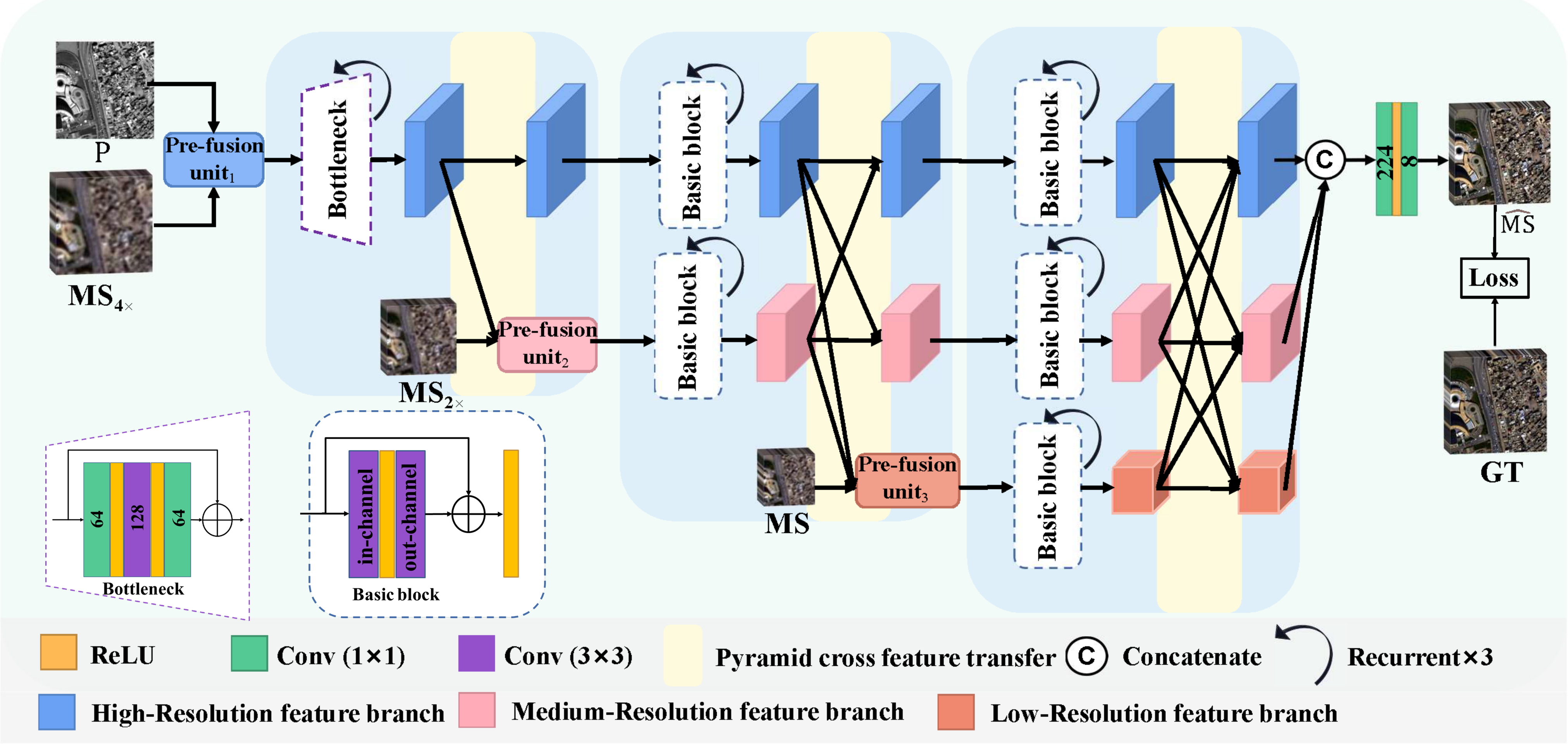
ICCV 2021 OCTOBER 11-17 VIRTUAL

1. Abstract

Deep Convolution Neural Networks have been adopted for pansharpening and achieved state-of-the-art performance. However, most of the existing works mainly focus on single-scale feature fusion, which leads to failure in fully considering relationships of information between high-level semantics and low-level features, despite the network is deep enough. In this paper, we propose a dynamic cross feature fusion network (DCFNet) for pansharpening. Specifically, DCFNet contains multiple parallel branches, including a high-resolution branch served as the backbone, and the low-resolution branches progressively supplemented into the backbone. Thus our DCFNet can represent the overall information well. In order to enhance the relationships of inter-branches, dynamic cross feature transfers are embedded into multiple branches to obtain high-resolution representations. Then contextualized features will be learned to improve the fusion of information. Experimental results indicate that DCFNet significantly outperforms the prior arts in both quantitative indicators and visual qualities.



2. Introduction



Motivation: high-resolution representation and better inter-branch information fusion

A family of pansharpening networks only adopts single-scale feature fusion to generate final HRMS;

Existing multiscale network always reduces the spatial resolution of features in the process of feature extraction, e.g. U-shaped network and pyramidal network.

Thus, we reform HRNet with proposed modules to achieve better inter-branch information fusions.

☐ **Loss function:** $\mathcal{L}oss = \frac{1}{n}\sum_{k=1}^{n}\left\|\mathcal{F}_{\Theta_{DCFNet}}\left(\mathbf{I}^{\{k\}}\right) - \mathbf{GT}^{\{k\}}\right\|_F^2$

☐ **Formulation on Dynamic Branch Fusion:** $O = \sum_i \frac{w_i}{\sum_j w_j + \epsilon} \cdot I_i$

Pre-fusion Units and Building blocks:

They play complete stages to deepen the network to extract better features.

Pyramid Cross Feature Transfer:

It maintains the high-resolution branch and aggregates features from high-to-low and low-to-high branches, then transfers the cross-scale features back to high-resolution branches.

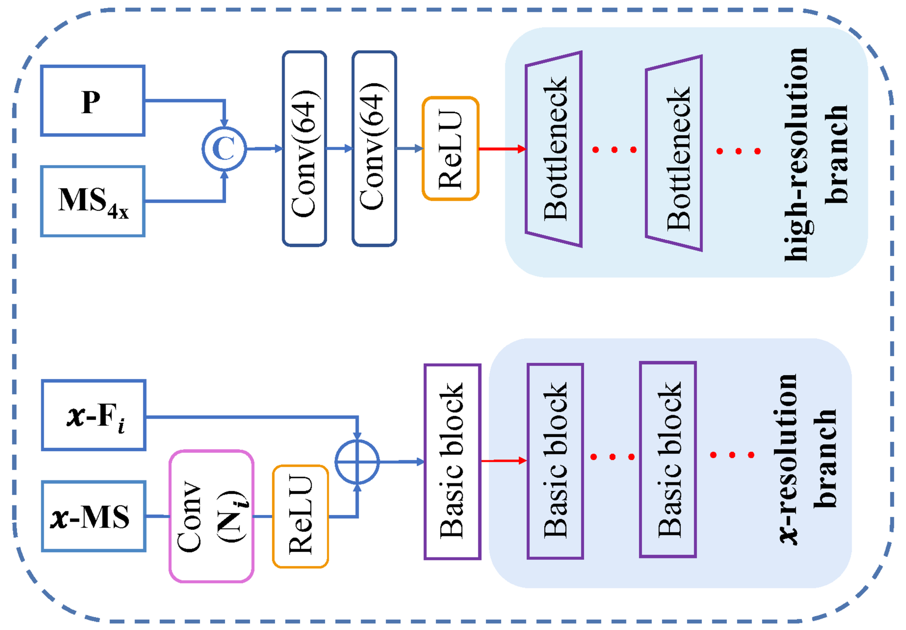
Dynamic Branch Fusion:

It mitigates the unequal effects of different resolutions on the final result.

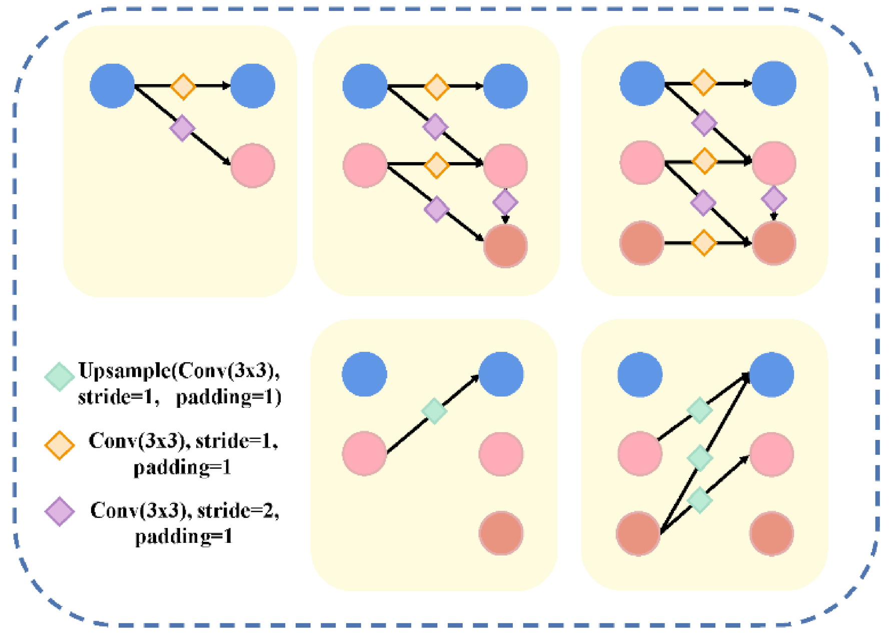
Diverse Structure Deformation:

DCFNet has superior foundations for feature extraction and fusion.

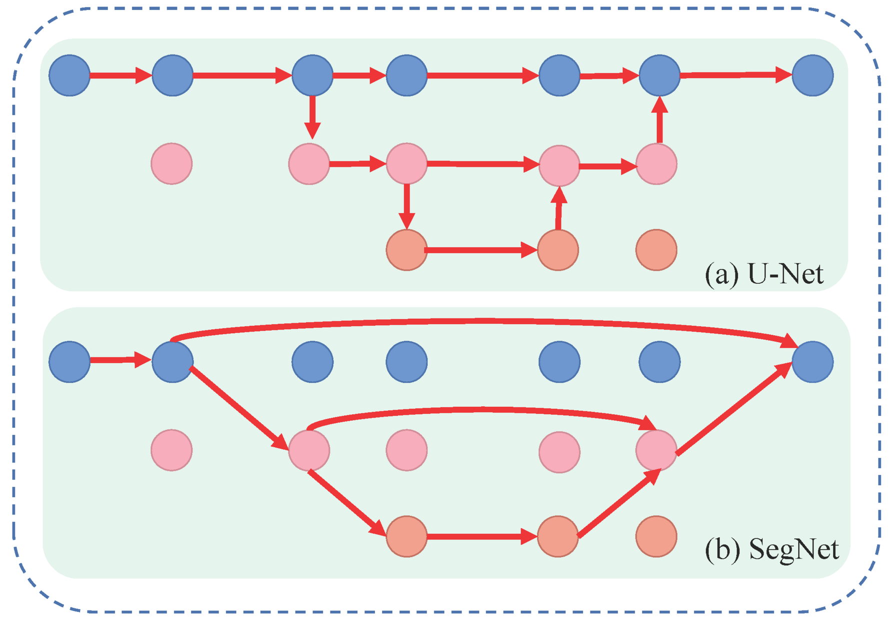
Pre-fusion Units and Building Blocks



Pyramid Cross Feature Transfer



Diverse structural deformation



3. Results

☐ **Pansharpening：**

1) **Datasets:**

8-band data: WorldView-3 (WV3)

1) reduced-resolution examples

2) full-resolution examples

4-band data: QuickBird (QB), GaoFen-2 (GF-2)

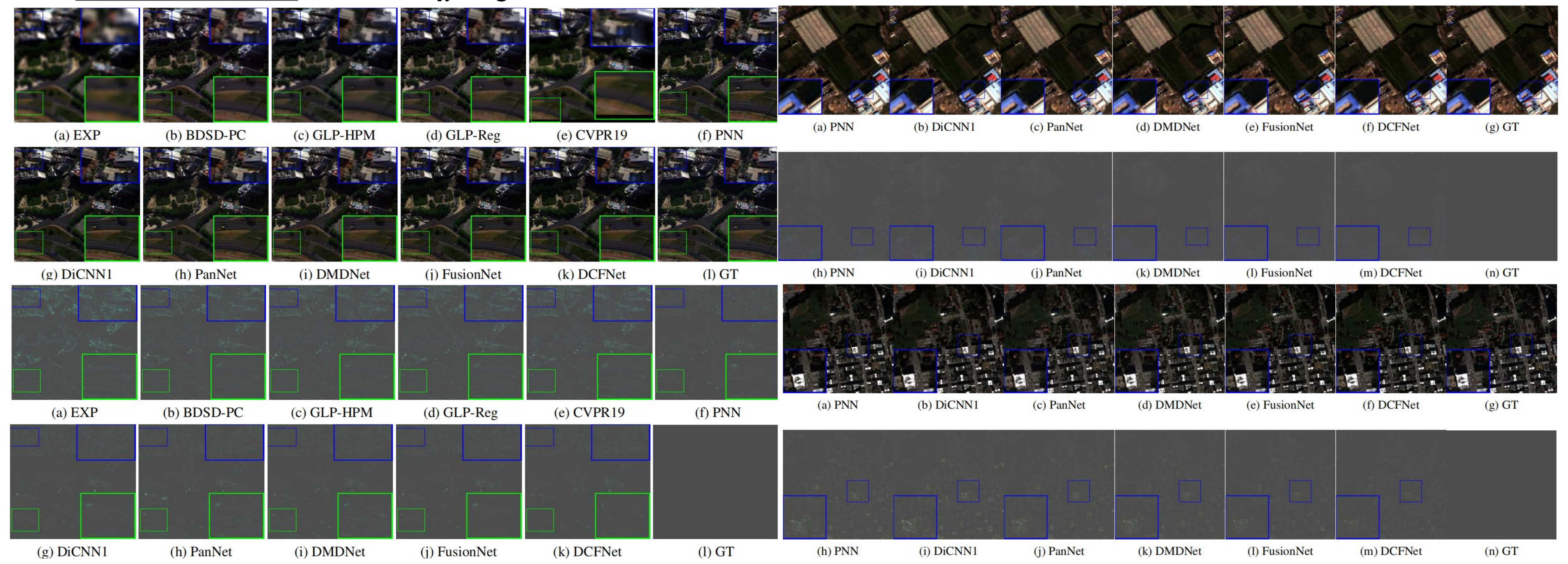
2) **Metrics:**

☐ Reduced-resolution: SAM, ERGAS, SCC, Q8

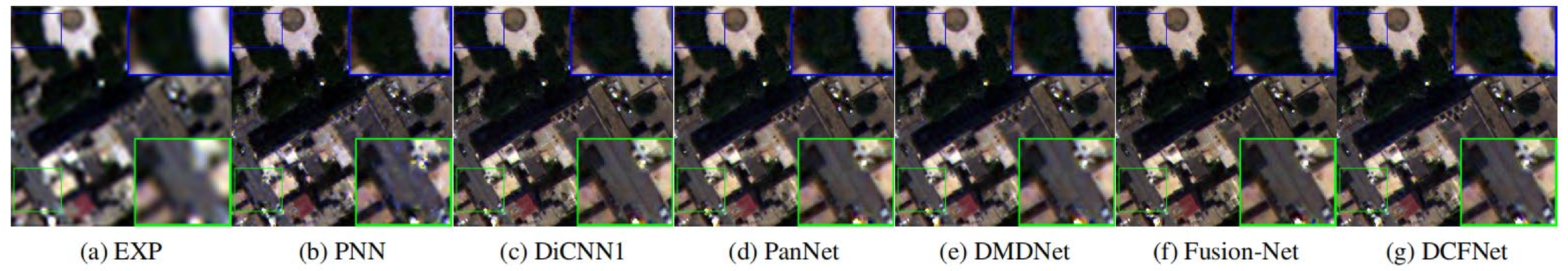
☐ Full-resolution: QNR, D_λ, D_S

Table 1: Quantitative metrics the compared CNN-based methods on **1258 reduced-resolution test datasets (WV3).**

Method	SAM (± std)	ERGAS (± std)	Q8 (± std)	SCC (± std)
PNN [13]	4.401 ± 1.329	3.228 ± 1.004	0.888 ± 0.112	0.921 ± 0.046
DiCNN1 [10]	3.980 ± 1.318	2.736 ± 1.015	0.909 ± 0.111	0.951 ± 0.047
PanNet [28]	4.092 ± 1.273	2.952 ± 0.977	0.894 ± 0.117	0.949 ± 0.046
DMDNet [6]	3.971 ± 1.248	2.857 ± 0.966	0.900 ± 0.114	0.952 ± 0.044
FusionNet [4]	3.743 ± 1.225	2.567 ± 0.944	0.913 ± 0.112	0.958 ± 0.045
DCFNet	**3.377 ± 1.200**	**2.257 ± 0.910**	**0.926 ± 0.107**	**0.967 ± 0.043**
Ideal value	0	0	1	1



Visual comparisons on **Tripoli dataset (WV3), GF-2 and QB** with natural colors/absolute error maps.



Visual comparisons on **Tripoli-OS dataset (WV3) at the original scale.**

Table3: 50 full-resolution samples (WV3).

Method	QNR (± std)	D_λ (± std)	D_s (± std)
PNN [13]	0.946 ± 0.022	0.023 ± 0.014	0.032 ± 0.012
DiCNN1 [10]	0.939 ± 0.024	0.026 ± 0.016	0.035 ± 0.011
PanNet [28]	0.948 ± 0.017	0.029 ± 0.011	**0.022 ± 0.009**
DMDNet [6]	0.945 ± 0.020	0.024 ± 0.012	0.030 ± 0.013
FusionNet [4]	0.941 ± 0.022	0.024 ± 0.013	0.031 ± 0.013
DCFNet	**0.956 ± 0.013**	**0.022 ± 0.009**	**0.022 ± 0.006**
Ideal value	1	0	0

Table 5: network generalization on Stockholm dataset (WorldView2).

Method	SAM	ERGAS	Q8	SCC
EXP [1]	7.8500	9.6793	0.6540	0.4505
BDSD-PC [20]	7.0953	6.3233	0.8819	0.8578
GLP-HPM [2, 24]	7.2988	6.9965	0.8527	0.8355
CVPR19 [5]	7.1098	6.5434	0.8752	0.8457
GLP-Reg [2, 23]	7.1195	6.4998	0.8776	0.8453
PNN [13]	6.8624	5.6259	0.8642	0.8539
DiCNN1 [10]	6.8159	5.9773	0.8802	0.8797
PanNet [28]	6.3916	5.6302	0.8897	0.8895
DMDNet [6]	**6.1986**	5.5692	0.8903	0.8965
FusionNet [4]	6.2784	5.5499	0.8969	0.8897
DCFNet	6.6871	**5.1682**	**0.9175**	**0.9125**
Ideal value	0	0	1	1

Table 6: Ablation study on some fusion operations on Tripoli dataset.

Method	SAM	ERGAS	Q8	SCC
w/o DFB	3.893	2.836	0.971	0.959
w/o PCFT	4.001	2.852	**0.972**	0.959
DCFNet	**3.852**	**2.825**	**0.972**	**0.960**
Ideal value	0	0	1	1

Table 4: Quantitative metrics of the compared CNN-based methods for the GF-2 testing dataset (81 samples) and the QB testing dataset (48 samples).

Method	SAM (± std)	ERGAS (± std)	Q8 (± std)	SCC (± std)
Guangzhou (GF-2)				
PNN [13]	1.659 ± 0.360	1.570 ± 0.324	0.927 ± 0.020	0.928 ± 0.020
DiCNN1 [10]	1.494 ± 0.381	1.320 ± 0.354	0.944 ± 0.021	0.945 ± 0.022
PanNet [28]	1.395 ± 0.326	1.223 ± 0.282	0.946 ± 0.022	0.955 ± 0.012
DMDNet [6]	1.297 ± 0.315	1.128 ± 0.266	0.952 ± 0.021	0.964 ± 0.010
FusionNet [4]	1.179 ± 0.271	1.002 ± 0.227	0.962 ± 0.016	0.971 ± 0.007
DCFNet	**0.994 ± 0.185**	**0.811 ± 0.144**	**0.971 ± 0.016**	**0.982 ± 0.004**
Indianapolis dataset (QB)				
PNN [13]	5.799 ± 0.947	5.571 ± 0.458	0.857 ± 0.148	0.902 ± 0.048
DiCNN1 [10]	5.307 ± 0.995	5.231 ± 0.541	0.882 ± 0.143	0.922 ± 0.050
PanNet [28]	5.314 ± 1.017	5.162 ± 0.681	0.883 ± 0.139	0.929 ± 0.058
DMDNet [6]	5.119 ± 0.939	4.737 ± 0.648	0.890 ± 0.146	0.134 ± 0.065
FusionNet [4]	4.540 ± 0.778	4.050 ± 0.266	0.910 ± 0.136	0.954 ± 0.045
DCFNet	**4.342 ± 0.719**	**3.749 ± 0.266**	**0.920 ± 0.129**	**0.961 ± 0.046**
Ideal value	0	0	1	1

4. Conclusions

- **We introduce a novel inter-branch dynamic feature fusion with dynamic Pre-fusion Units and Pyramid Cross Feature.**
- **Propose a dynamic cross feature fusion network (DCFNet), which significantly improves the effects of feature fusion in pansharpening.**
- **Proposed DCFNet achieves state-of-the-art performance and have a reliable generalization capability in pansharpening.**