# BAM: Bilateral Activation Mechanism for Image Fusion

Zi-Rong Jin, Liang-Jian Deng, Tian-Jing Zhang, Xiao-Xu Jin

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

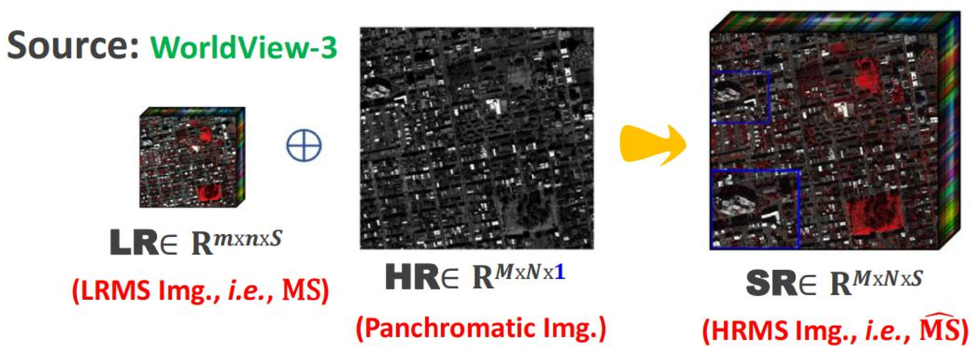
ACM multimedia Chengdu, China OCT 20-24 2021

## 1. ABSTRACT

As the conventional activation functions such as ReLU, LeakyReLU, and PReLU, the negative parts in feature maps are simply truncated or linearized, which may result in unflexible structure and undesired information distortion. In this paper, we propose a simple but effective Bilateral Activation Mechanism (BAM) which could be applied to the activation function to offer an efficient feature extraction model. Based on BAM, the Bilateral ReLU Residual Block (BRRB) that still sufficiently keeps the nonlinear characteristic of ReLU is constructed to separate the feature maps into two parts,i.e., the positive and negative components, then adaptively represent and extract the features by two independent convolution layers. Besides, our mechanism will not increase any extra parameters or computational burden in the network. We finally embed the BRRB into a basic ResNet architecture, called BRResNet, it is easy to obtain state-of-the-art performance in two image fusion tasks, i.e., pansharpening and hyperspectral image super-resolution (HISR). Additionally, deeper analysis and ablation study demonstrate the effectiveness of BAM, the lightweight property of the network, etc.

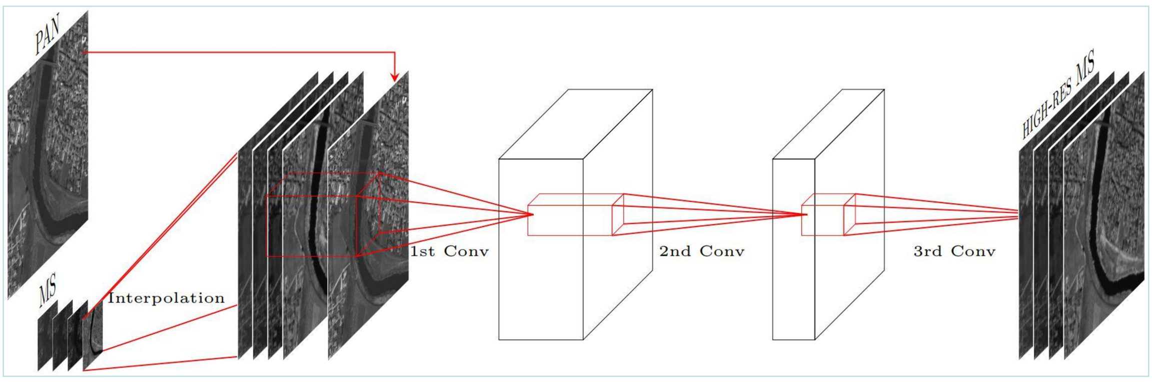
## 2. INTRODUCTION

- **Pansharpening**




- **Hyperspectral Image Super-Resolution (HISR)**


- **CNNs-based Approaches for Pansharpening, e.g., PNN [1]**




The current main improvement direction is to change the network structure, such as deepening of depth, increasing width, and multi-scale convolution operations

Activation function, e.g., ReLU, as the important tool which plays a role in activating the nonlinear fitting ability of CNNs. Thus we present a framework from the direction of the activation mechanism extension, expecting to explore and utilize the features that can not be activated while retaining the nonlinearity

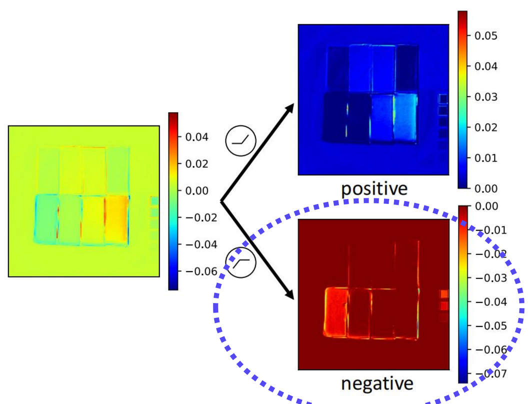
- **More development history of CNN can be found from [2]**

[1] G. Masi, D. Cozzolino, L. Verdoliva, and G. Scarpa, "Pansharpening by convolutional neural networks," Remote Sensing, vol. 8, pp. 594, 2016.

[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.

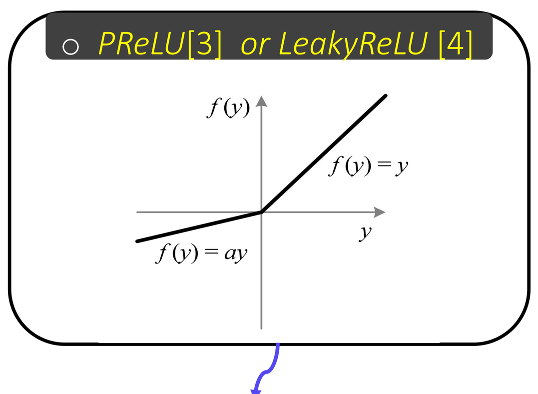
## 3. MOTIVATION

- **An Observation**




Still existing image residuals, which can be further utilized in next step for better output!

- **Limitation**

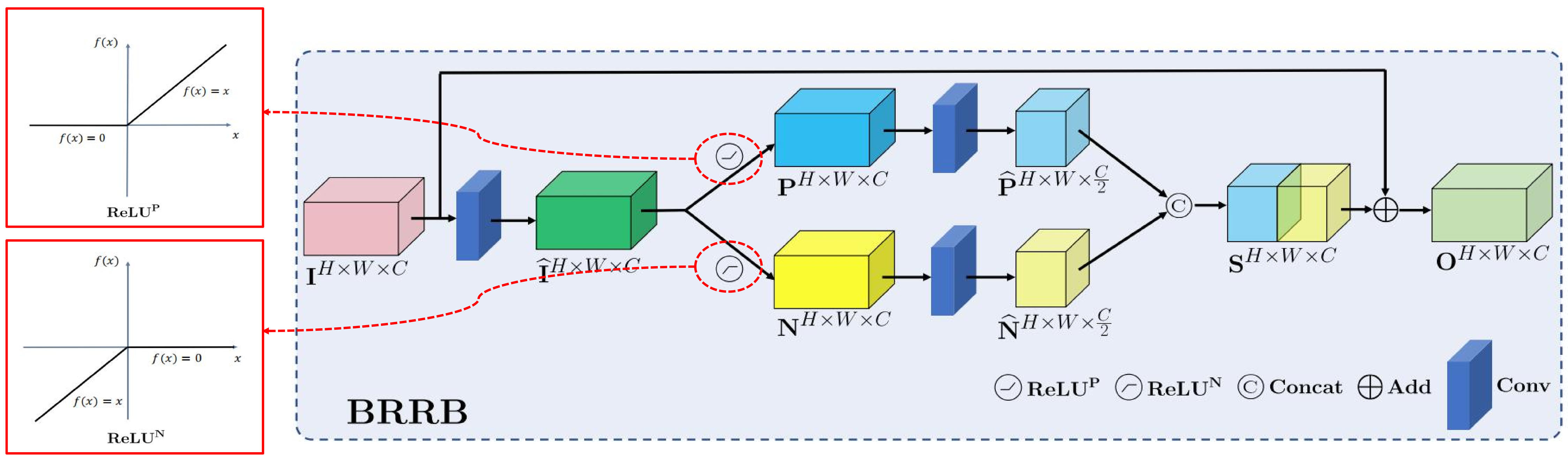



Feature extraction of the negative part is strengthened by weakening the nonlinearity

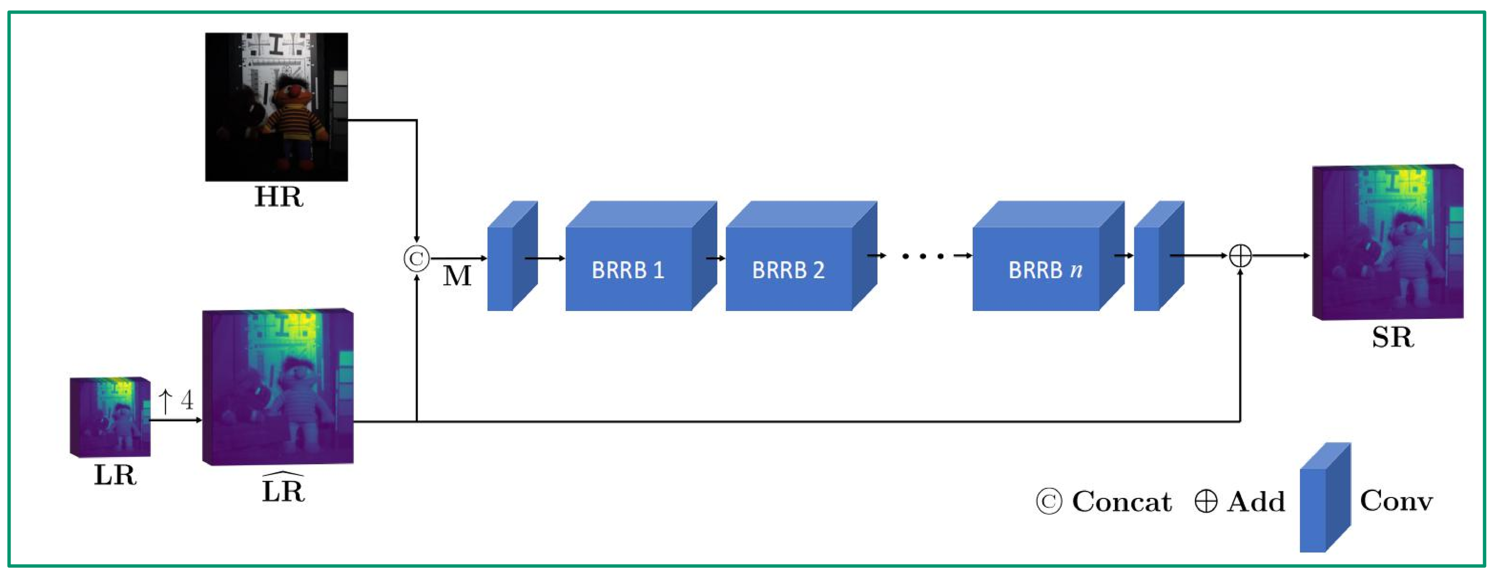
[1] K.M He, et al., Surpassing human-level performance on imagenet classification, ICCV, 2015

[2] Andrew, et al., Rectifier nonlinearities improve neural network acoustic models, ICML, 2013

### ○ Bilateral ReLU Residual Block (BRRB)




### ○ BRResNet




### ○ Loss Fuction

$$\mathcal{L}(\theta) = \frac{1}{N}\sum_{i=1}^{N}\left\| \mathcal{F}_{\theta}(\widehat{\mathbf{LR}}^{(i)}, \mathbf{HR}^{(i)}) + \widehat{\mathbf{LR}} - \widehat{\mathbf{GT}}^{(i)} \right\|_F^2$$

### ○ Contribution

As a mechanism, BAM provides a more efficient feature extraction mode without increase the computational burden. Also, it has many variants and can be used as a substitution to replace any structure like "Activation + Convolution", giving us more flexibility in designing the network structure.

## 5. EXPERIMENT SETTING

**□ HISR:**

1) **Datasets:**

- □ CAVE dataset
- □ HAVARD dataset

2) **Metrics:** SAM, PSNR, SSIM, ERGAS

**□ Pansharpening:**

1) **Datasets:**

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

2) **Metrics:**

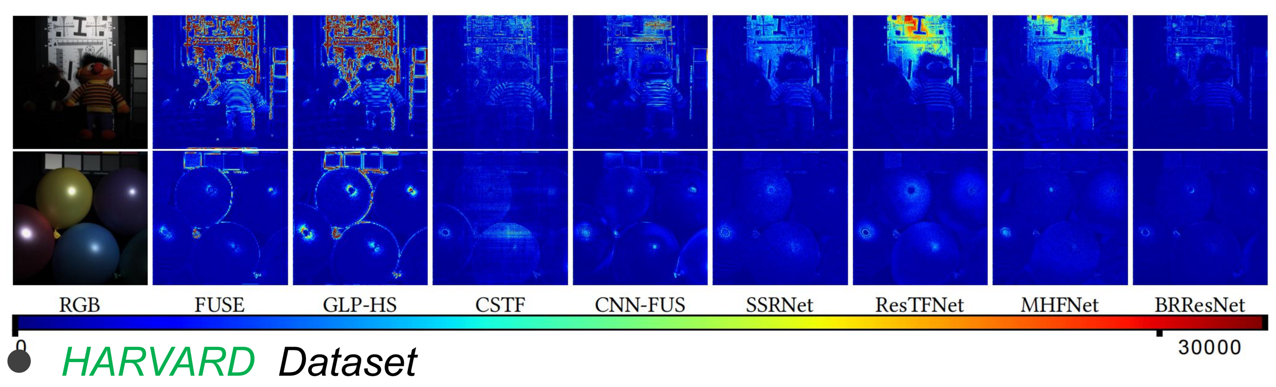
- □ Reduced-resolution: SAM, ERGAS, SCC, Q8
- □ Full-resolution: QNR, $D_\lambda$, $D_S$

## 7. CONCLUSIONS

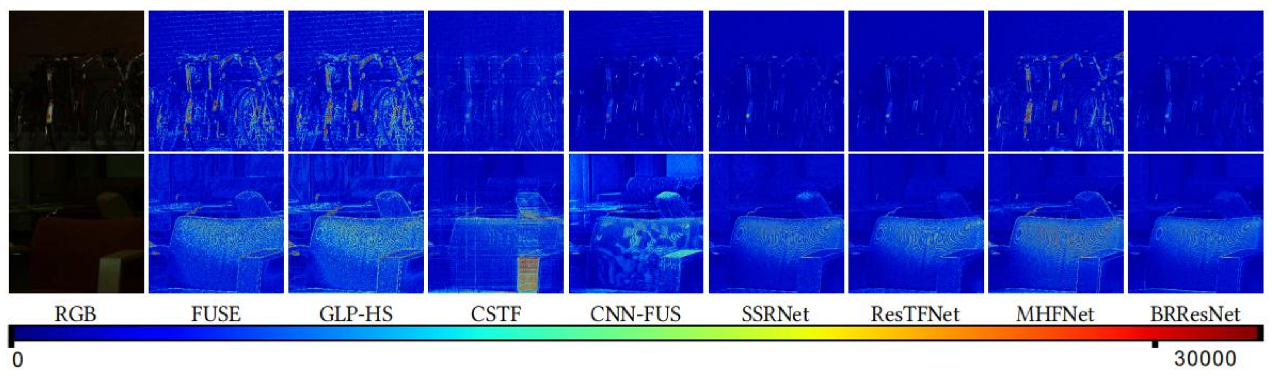
- **We introduce a simple but effective Bilateral Activation Mechanism (BAM) that not only retains the nonlinearity of the activation function but also avoids information distortion caused by inactivation.**
- **A network with residual structure using BAM with ReLU (BRResNet) is proposed, which significantly improves the efficiency of feature extraction in image fusion tasks.**
- **Proposed BRResNet achieves state-of-the-art performance in two fusion tasks. Especially, the given BRResNet holds a large margin among other CNNs-based methods in terms of the parameters, thus can be viewed as a lightweight network.**

## 6. RESULTS

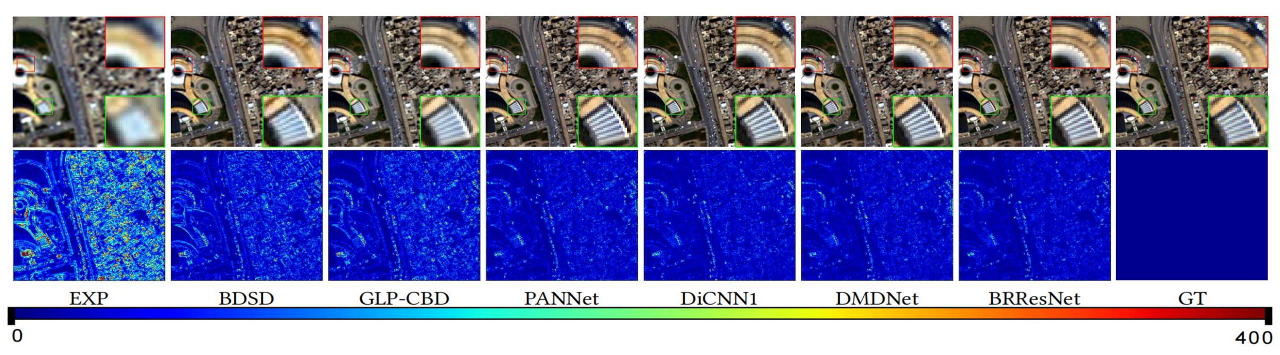
- *CAVE* Dataset




- *HARVARD* Dataset




- *WV3 Reduced-Resolution* Dataset




- *WV3 Full-Resolution* Dataset

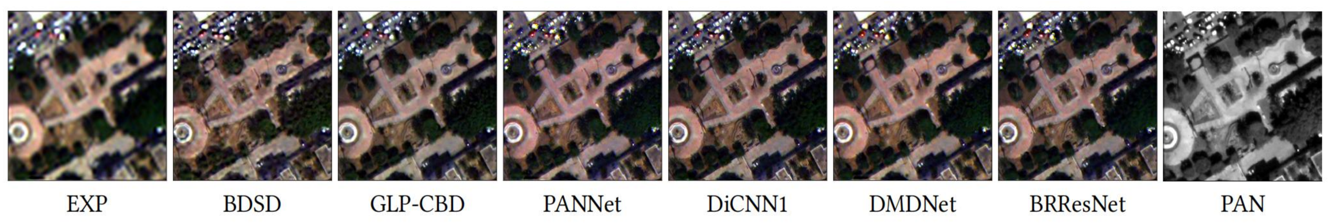



**Table 1:** Average quantitative comparisons on 11 CAVE examples (Red: the best; Blue: the second best).

| Method | *PSNR* | *SAM* | *ERGAS* | *SSIM* |
|---|---|---|---|---|
| FUSE | 39.72 ± 3.52 | 5.83 ± 2.02 | 4.18 ± 3.08 | 0.975 ± 0.018 |
| GLP-HS | 37.81 ± 3.06 | 5.36 ± 1.78 | 4.66 ± 2.71 | 0.972 ± 0.015 |
| CSTF | 42.14 ± 3.04 | 9.92 ± 4.11 | 3.08 ± 1.56 | 0.964 ± 0.027 |
| CNN-FUSE | 42.66 ± 3.46 | 6.44 ± 2.31 | 2.95 ± 2.24 | 0.982 ± 0.007 |
| SSRNet | 45.28 ± 3.13 | 4.72 ± 1.76 | 2.06 ± 1.30 | 0.990 ± 0.004 |
| ResTFNet | 45.35 ± 3.68 | 3.76 ± 1.31 | 1.98 ± 1.62 | 0.993 ± 0.003 |
| MHFNet | 46.32 ± 2.76 | 4.33 ± 1.48 | 1.74 ± 1.44 | 0.992 ± 0.006 |
| BRResNet (Our) | 47.85 ± 3.56 | 2.96 ± 0.89 | 1.50 ± 1.18 | 0.995 ± 0.003 |
| Ideal value | ∞ | 0 | 0 | 1 |

**Table 2:** Average quantitative comparisons on 10 Harvard examples.

| Method | *PSNR* | *SAM* | *ERGAS* | *SSIM* |
|---|---|---|---|---|
| FUSE | 42.06 ± 2.94 | 3.23 ± 0.91 | 3.14 ± 1.52 | 0.977 ± 0.009 |
| GLP-HS | 40.14 ± 3.22 | 3.52 ± 0.96 | 3.74 ± 1.44 | 0.966 ± 0.012 |
| CSTF | 42.97 ± 3.33 | 3.30 ± 1.25 | 2.43 ± 1.06 | 0.972 ± 0.021 |
| CNN-FUSE | 43.61 ± 4.73 | 3.32 ± 1.17 | 2.78 ± 1.64 | 0.978 ± 0.016 |
| SSRNet | 44.40 ± 3.49 | 2.61 ± 0.72 | 2.39 ± 1.02 | 0.985 ± 0.007 |
| ResTFNet | 44.47 ± 4.04 | 2.56 ± 0.68 | 2.21 ± 0.87 | 0.985 ± 0.008 |
| MHFNet | 43.10 ± 3.94 | 2.76 ± 0.77 | 3.28 ± 1.54 | 0.977 ± 0.009 |
| BRResNet (Our) | 45.74 ± 3.86 | 2.39 ± 0.66 | 1.94 ± 0.69 | 0.986 ± 0.009 |
| Ideal value | ∞ | 0 | 0 | 1 |

**Table 3:** Average quantitative comparisons on 1258 reduced resolution WV3 examples.

| Method | *SAM* | *ERGAS* | *SCC* | *Q8* |
|---|---|---|---|---|
| BDSD | 6.9997 ± 2.8530 | 5.1670 ± 2.2475 | 0.8712 ± 0.0798 | 0.8126 ± 0.1234 |
| GLP-CBD | 5.2861 ± 1.9582 | 4.1627 ± 1.7748 | 0.8904 ± 0.0698 | 0.8540 ± 0.1144 |
| PanNet | 4.0921 ± 1.2733 | 2.9524 ± 0.9778 | 0.9495 ± 0.0461 | 0.8942 ± 0.1170 |
| DiCNN1 | 3.9805 ± 1.3181 | 2.7367 ± 1.0156 | 0.9517 ± 0.0472 | 0.9097 ± 0.1117 |
| DMDNet | 3.9714 ± 1.2482 | 2.8572 ± 0.9663 | 0.9527 ± 0.0447 | 0.9000 ± 0.1142 |
| BRResNet (Our) | 3.5881 ± 1.2185 | 2.4618 ± 0.9306 | 0.9612 ± 0.0444 | 0.9183 ± 0.1099 |
| Ideal value | 0 | 0 | 1 | 1 |

**Table 4:** Average quantitative comparisons on 36 full resolution WV3 examples.

| Method | *QNR* | $D_\lambda$ | $D_S$ |
|---|---|---|---|
| BDSD | 0.9368 ± 0.0416 | 0.0170 ± 0.0137 | 0.0473 ± 0.0320 |
| GLP-CBD | 0.9107 ± 0.0518 | 0.0323 ± 0.0243 | 0.0597 ± 0.0325 |
| PanNet | 0.9605 ± 0.01551 | 0.0215 ± 0.0098 | 0.0184 ± 0.0074 |
| DiCNN1 | 0.9454 ± 0.0268 | 0.0181 ± 0.0135 | 0.0374 ± 0.0159 |
| DMDNet | 0.9595 ± 0.0155 | 0.0201 ± 0.0098 | 0.0209 ± 0.0073 |
| BRResNet | 0.9671 ± 0.0095 | 0.0179 ± 0.0063 | 0.0152 ± 0.0050 |
| Ideal value | 1 | 0 | 0 |