Adaptive Weighted Attention Network with Camera Spectral Sensitivity Prior for Spectral Reconstruction from RGB Images

Jiaojiao Li, Chaoxiong Wu, Rui Song, Yunsong Li, Fei Liu

CVPR SEATTLE WASHINGTON JUNE 16-18 2020

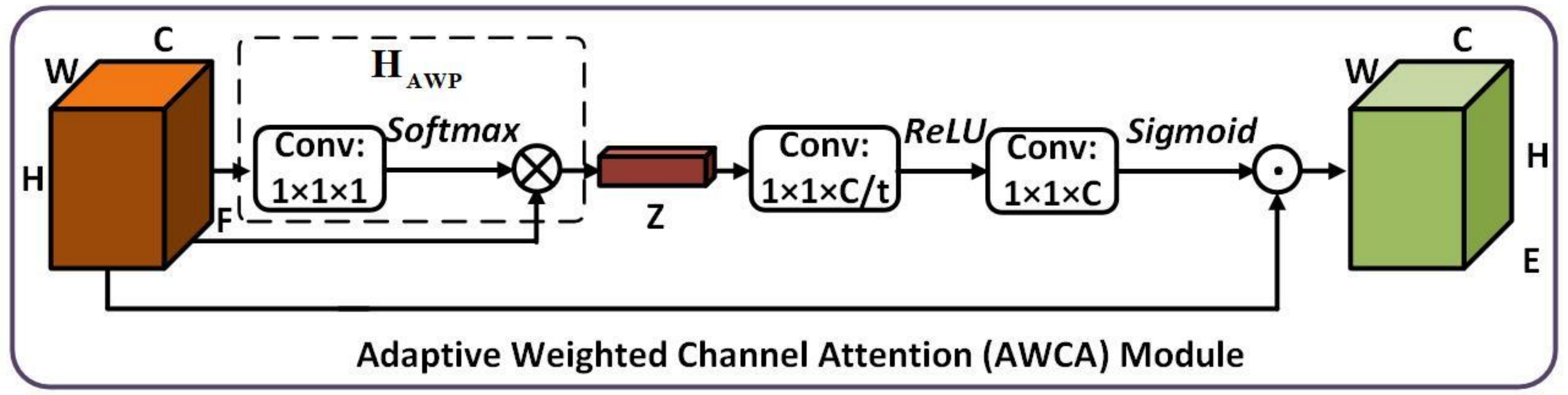
1. Motivation

• Recent promising effort for spectral reconstruction (SR) focuses on learning a complicated mapping through using a deeper and wider convolutional neural networks.
• Most of CNN-based SR methods neglect to explore the rich contextual information and dependencies among intermediate features.
• The existing CNN-based SR algorithms rarely consider to integrate the camera spectral sensitivity prior into SR.
• Therefore, these approaches fail to obtain more powerful representation ability of the network and better performance of SR.

2. Contributions

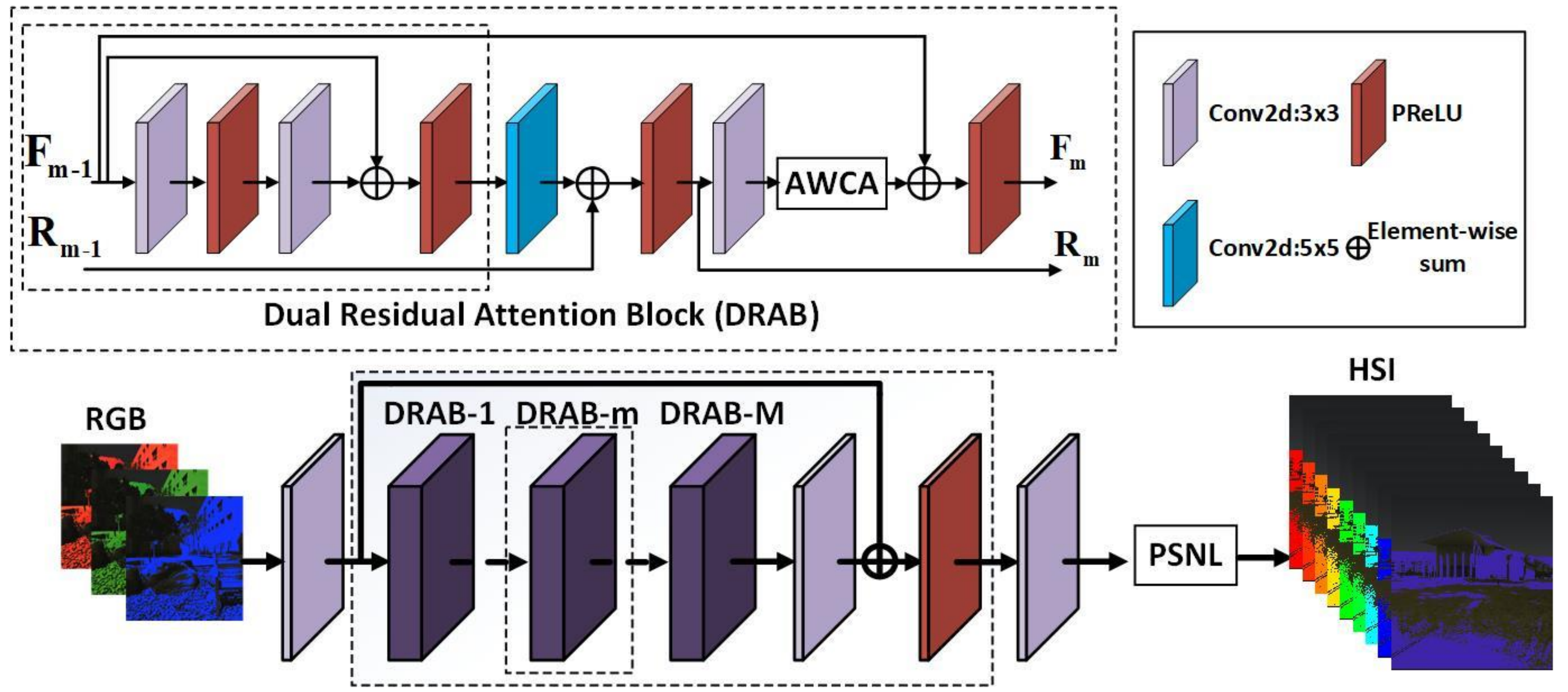
• We propose a novel deep adaptive weighted attention network (AWAN) for SR in the paper.
• An adaptive weighted channel attention (AWCA) module is presented to adaptively recalibrate channel-wise feature responses by exploiting the adaptive weighted feature statistics instead of average-pooled ones.
• We develop a patch-level second-order non-local (PSNL) module to capture long-range spatial contextual information via second-order non-local operations for more powerful feature representations.
• Based on the fact that the recovered RGB images can be created through applying the given camera spectral sensitivity function to reconstructed hyperspectral images, CSS prior is incorporated into the loss function to improve the performance of SR.

3.2. Adaptive Weighted Channel Attention (AWCA)



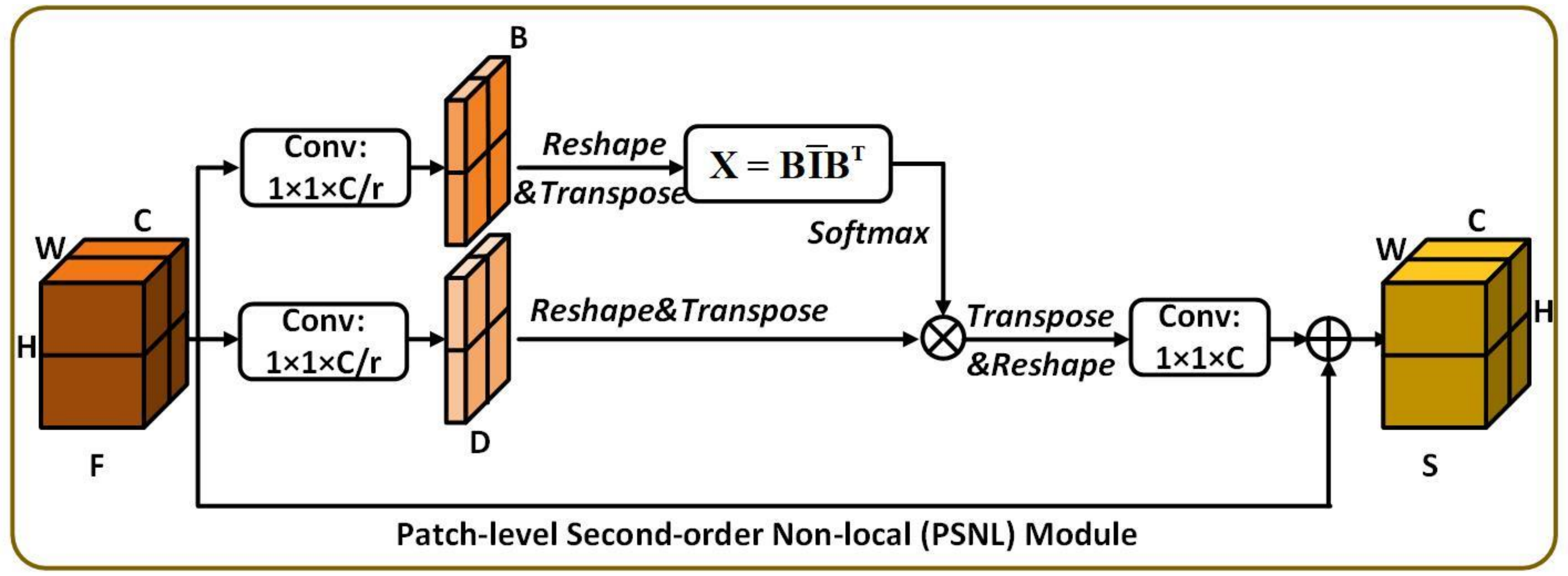
$$\mathbf{Z} = H_{AWP}(\mathbf{F}) \qquad V = \delta(W_2(\sigma(W_1(\mathbf{Z})))) \qquad \mathbf{e}_c = v_c \cdot \mathbf{f}_c$$

3.1. Adaptive Weighted Attention Network (AWAN)



• The backbone architecture is stacked with multiple dual residual attention blocks (DRAB).
• Each DRAB block consists of a fundamental residual module and additional paired convolutional operations with a large and small size kernels, in which the long and short skip connections form the dual residual learning.
• Such module can allow abundant low-frequency information of the original RGB images to be bypassed and utilized adequately.

3.3. Patch-level Second-order Non-local (PSNL)



$$\mathbf{X}_k = \mathbf{B}_k \bar{\mathbf{I}} \mathbf{B}_k^T \qquad \mathbf{U}_k = softmax(\mathbf{X}_k)\mathbf{D}_k \qquad \mathbf{S}_k = \phi(\mathbf{U}_k) + \mathbf{F}_k$$

4.1. Quantitative Results

Method	MRAE	Runtime/s	Compute Platform
AWAN+	**0.03010**	0.56	NVIDIA 2080Ti
2nd method	0.03076	16	NVIDIA 1080Ti
3rd method	0.03231	3.748	NVIDIA Titan XP
4th method	0.03476	~1	——
5th method	0.03516	0.7	Tesla K80

Method	MRAE	Runtime/s	Compute Platform
1st method	0.06201	3.748	NVIDIA Titan XP
2nd method	0.06213	16	NVIDIA 1080Ti
AWAN+	**0.06217**	0.56	NVIDIA 2080Ti
4th method	0.06515	~30	NVIDIA Titan XP
5th method	0.06733	——	NVIDIA 2080Ti

Table. The quantitative results of official test set for NTIRE2020 “Clean” and “Real World” tracks.

• Our entries obtain the 1st ranking on the official test set of “Clean” track and the 3rd place only 1.59106e-4 more than the 1st on the “Real World” track in the NTIRE 2020 Spectral Reconstruction Challenge.

4.2. Visual Results

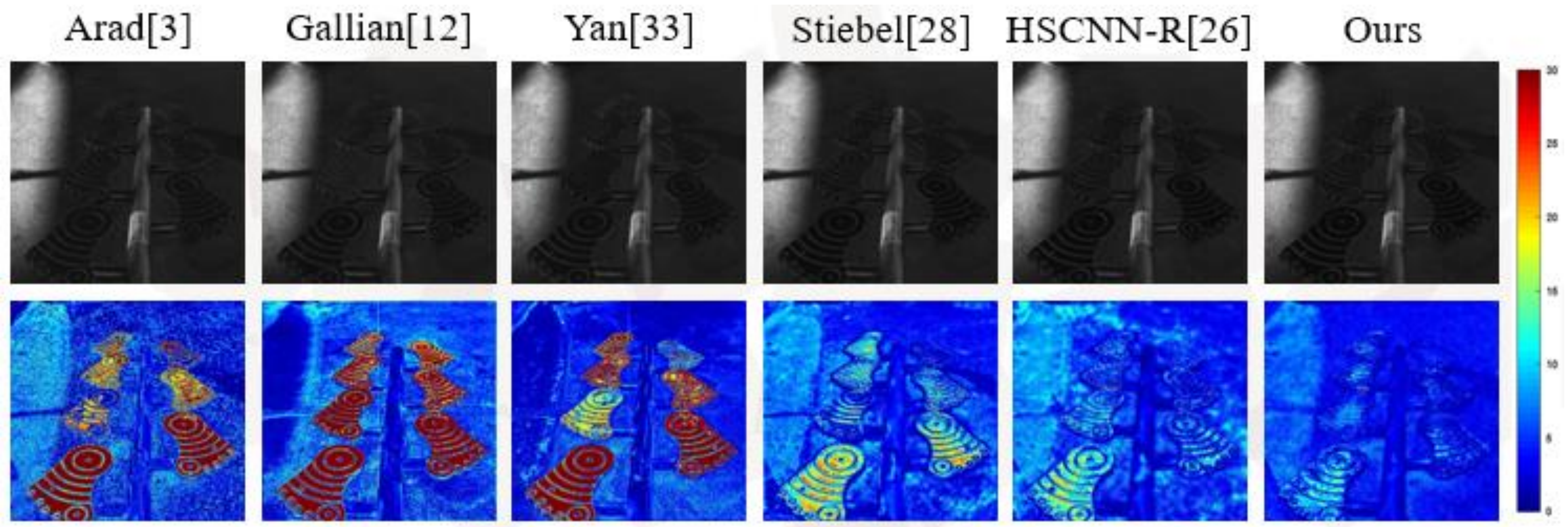


Fig. The visual results of the 18-th band and the reconstruction error images of a hyperspectral image chosen from validation set of NTIRE2020 “Clean” track.

• From these figures, we can see that our approach yields better recovery results and higher reconstruction fidelity than other methods. Also, the results of our proposed method are more accurate, which are closer to the ground truth.

5. Links

• Code available at:
https://github.com/Deep-imagelab/AWAN