Paper Presentation

Synergic Feature Attention for Image Restoration

Chong Mou, Jian Zhang

School of Electronic and Computer Engineering, Shenzhen Graduate School

Peking University, China

6-11 June 2021 • Toronto, Ontario, Canada

CONTENTS

01 Background and Motivation

The Task of Image Restoration

High quality image $(\mathbf{x})$

Degradation

Corrupted image $(\mathbf{y})$

Restoration

Recovered image $(\hat{\mathbf{x}})$

Degradation process：

$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}$, where $\mathbf{A}$ is the degradation matrix, and $\mathbf{n}$ is the additive noise.

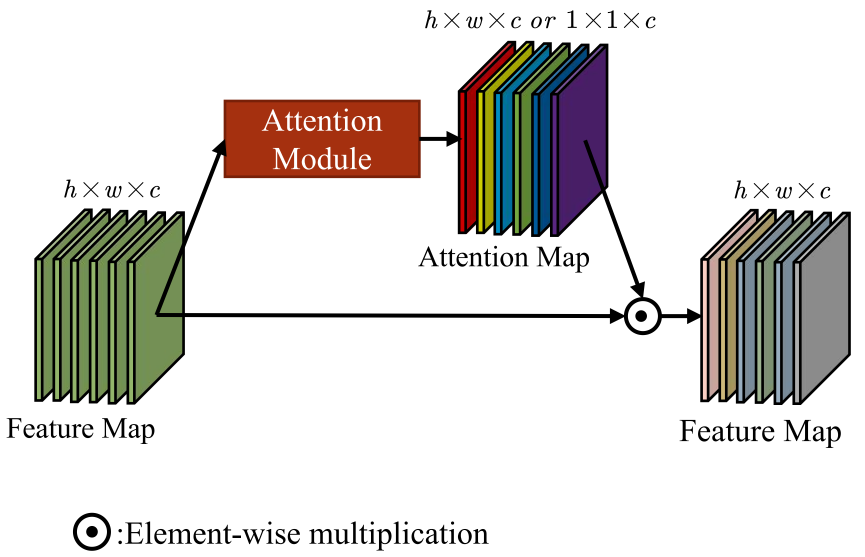
Image restoration：

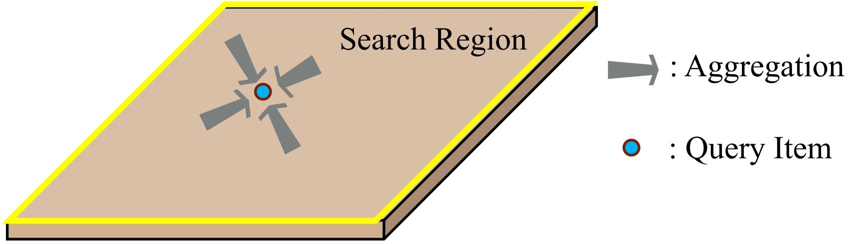
Image restoration aims to recover the latent high-quality image $\mathbf{x}$ from its degraded measurement $\mathbf{y}$.

01 Background and Motivation

Local and Non-local Attention Mechanisms for Image Restoration

Local





Non-local attention operation assumes that the similar image content would recur across the whole image and the local content can be recovered with the help of similar items from other places. The non-local operation can be generally defined as:

$$\hat{\mathbf{x}}_i = \frac{1}{z_i}\sum_{j\in\mathbb{Q}} \phi(\mathbf{y}_i,\mathbf{y}_j)G(\mathbf{y}_j)$$

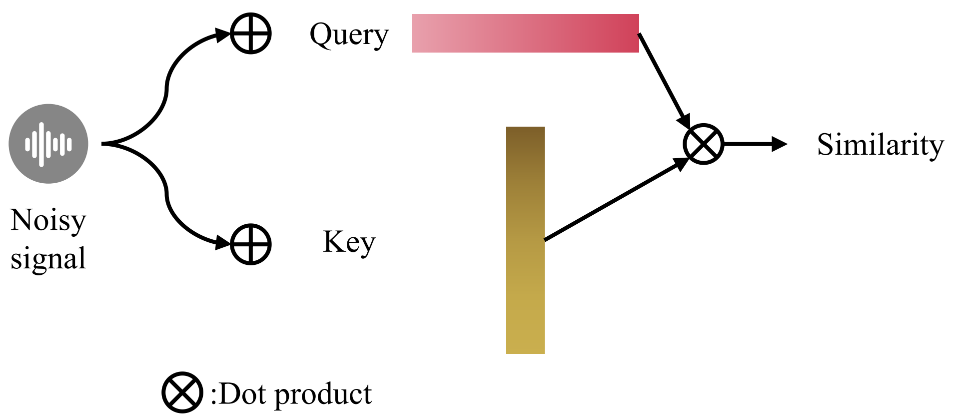
Non-local

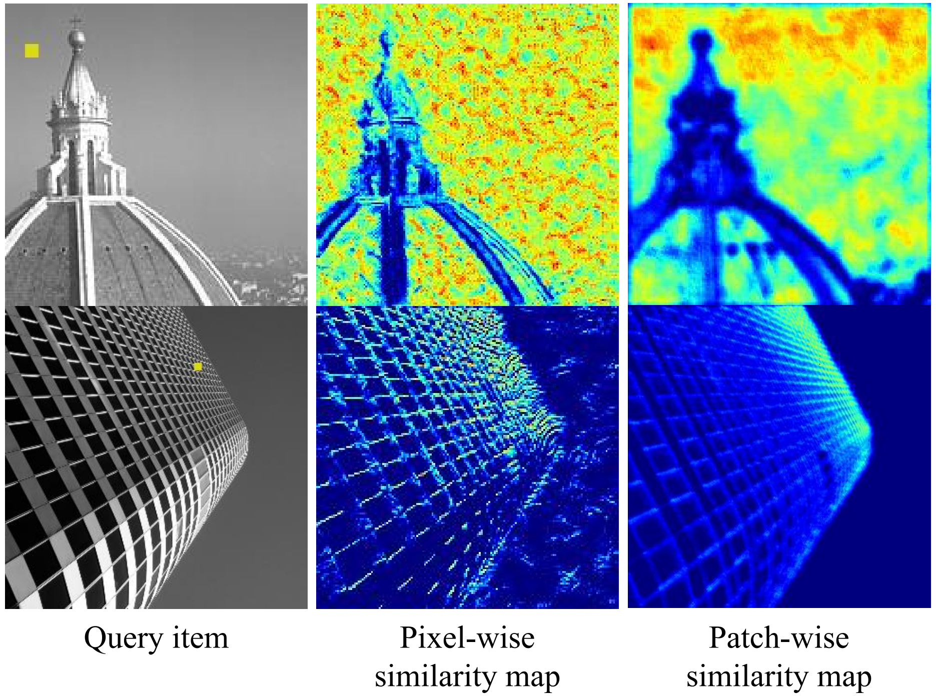
How to make a trade-off ?

Local and non-local attentions are both effective methods in the domain of image restoration (IR). However, most existing image restoration methods use these two strategies indiscriminately, and how to make a trade-off between local and non-local attention operations has hardly been studied.

01 Background and Motivation

Patch-wise Non-local Operation--A More Stable Non-local Strategy

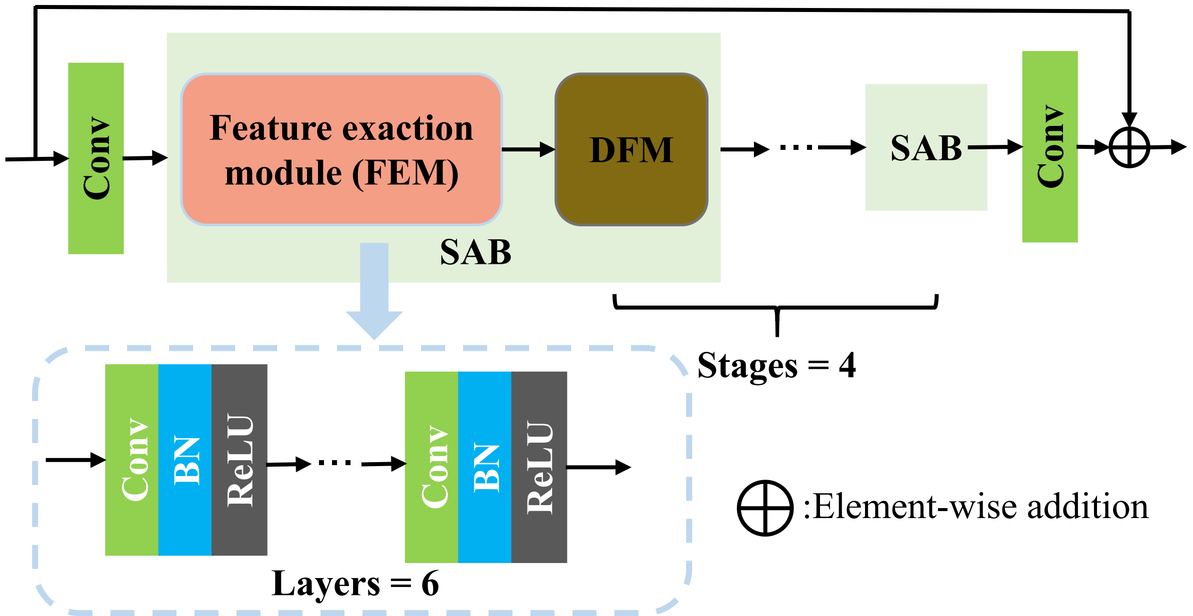




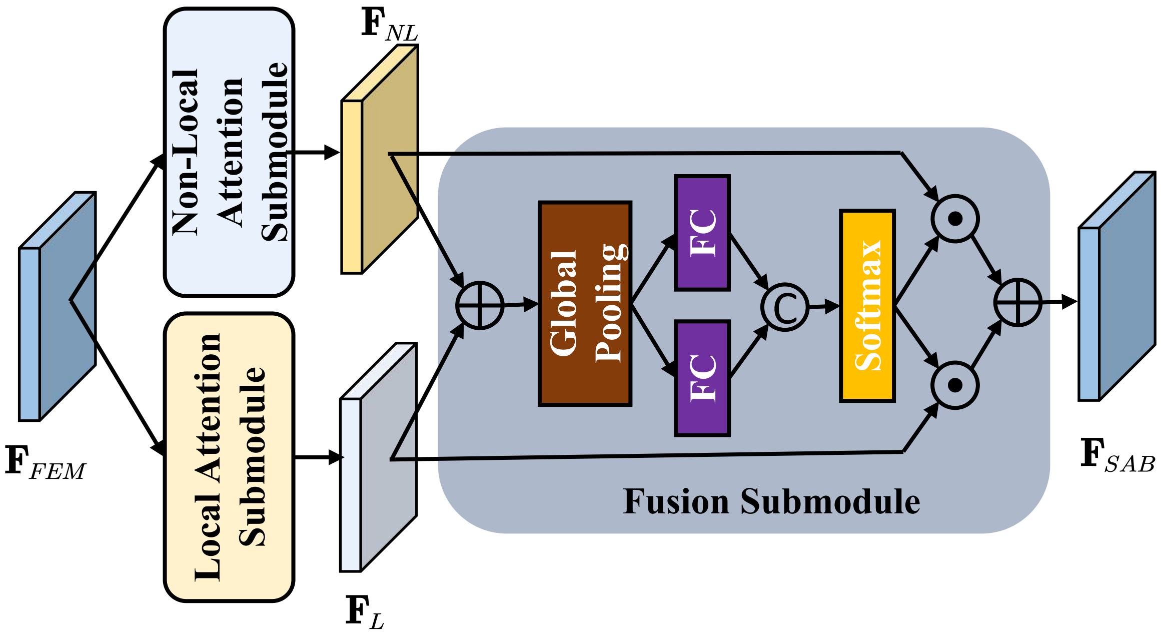
In image restoration tasks, images are corrupted, thus the methods performed non-local operations on pixel-level are easily influenced by noisy signals within the corrupted images. In comparison, the patch-wise method has a larger receptive field during the compution of similarity.

02 Network Architecture

Global Architecture of our proposed SAT-Net



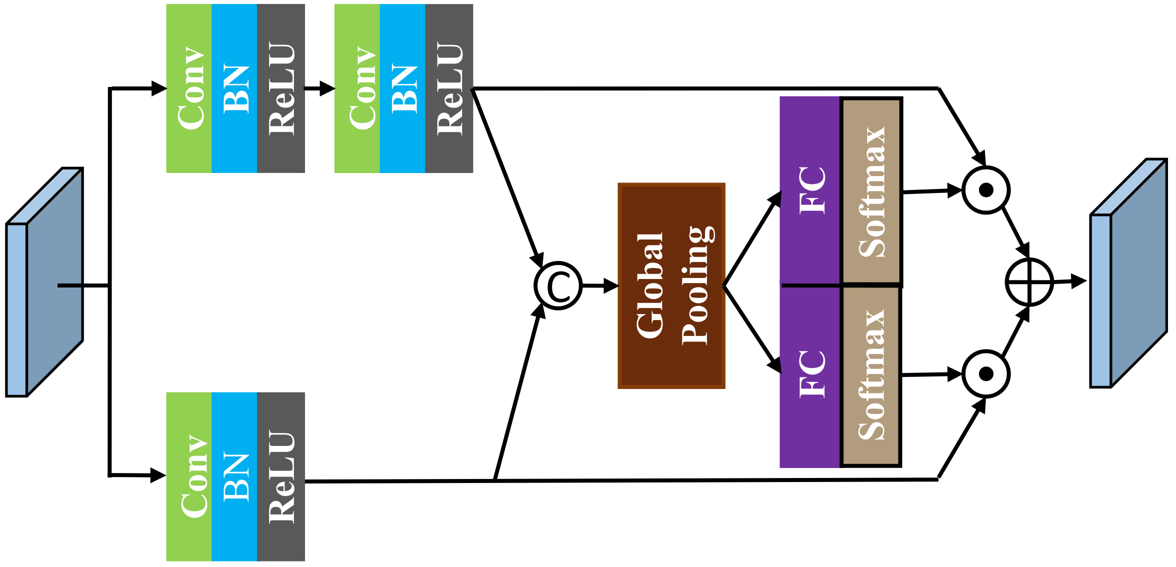
Details of our proposed Dual-branch fusion module (DFM)



02 Network Architecture

The architecture of local attention submodule in SAT-Net

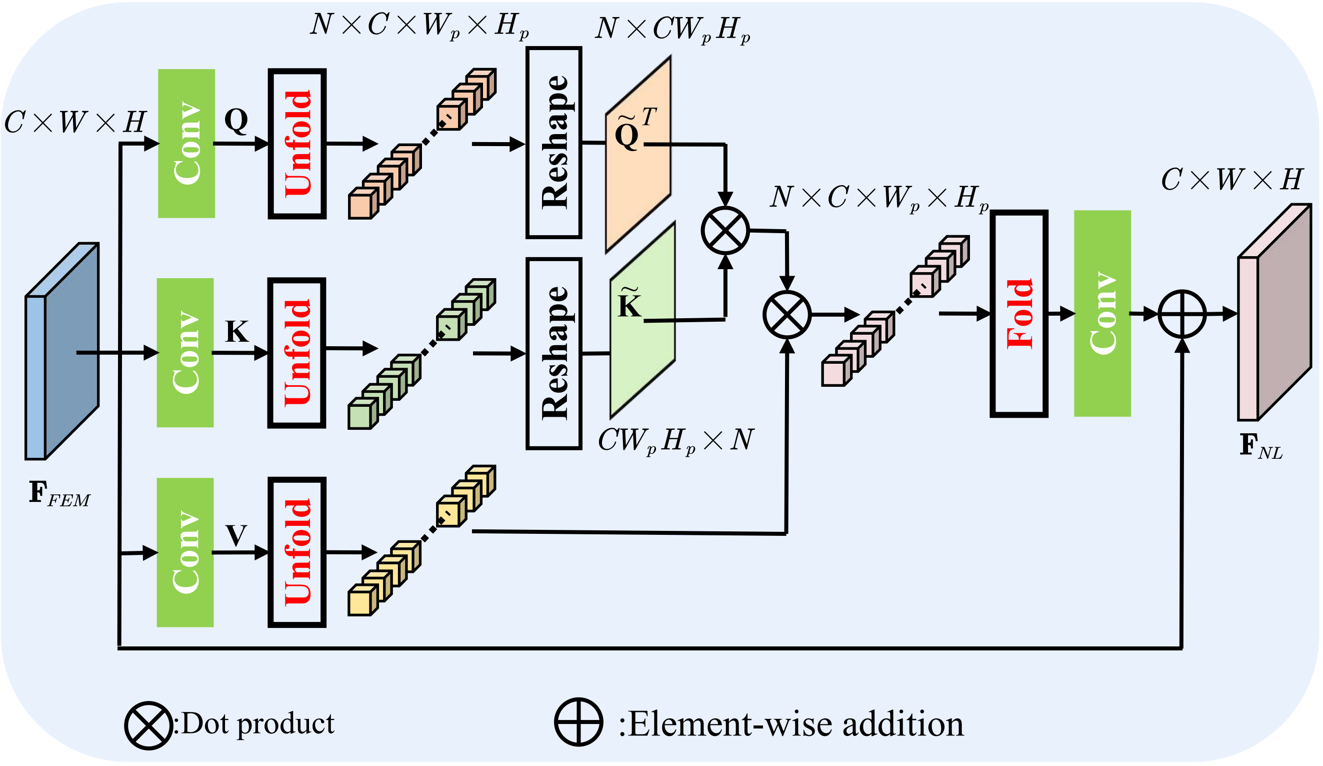
The local attention submodule applies channel attention mechanism, which is inspired by [1].



⊙ :Element-wise multiplication © :Concatenate ⊕ :Element-wise addition

[1] Li, Xiang, et al. "Selective kernel networks." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2019.

Details of our proposed patch-wise non-local attention submodule



Synthetic and Real Image Denoising

Table 1. Quantitative results (PSNR and SSIM) about gray-scale image denoising. Best results are **highlighted**.

Dataset	σ	DnCNN [8]	FFDNet [9]	ADNet [22]	N3Net [17]	NLRN [16]	SAT-Net (Ours)
Urban100	15	32.68/0.9255	32.43/0.9273	32.87/0.9304	33.08/0.9333	33.45/0.9354	**33.60/0.9376**
	25	29.97/0.8797	29.92/0.8887	30.24/0.8920	30.19/0.8925	30.94/0.9018	**31.17/0.9062**
	50	26.28/0.7874	26.52/0.8057	26.64/0.8072	26.82/0.8184	27.49/0.8279	**27.76/0.8373**

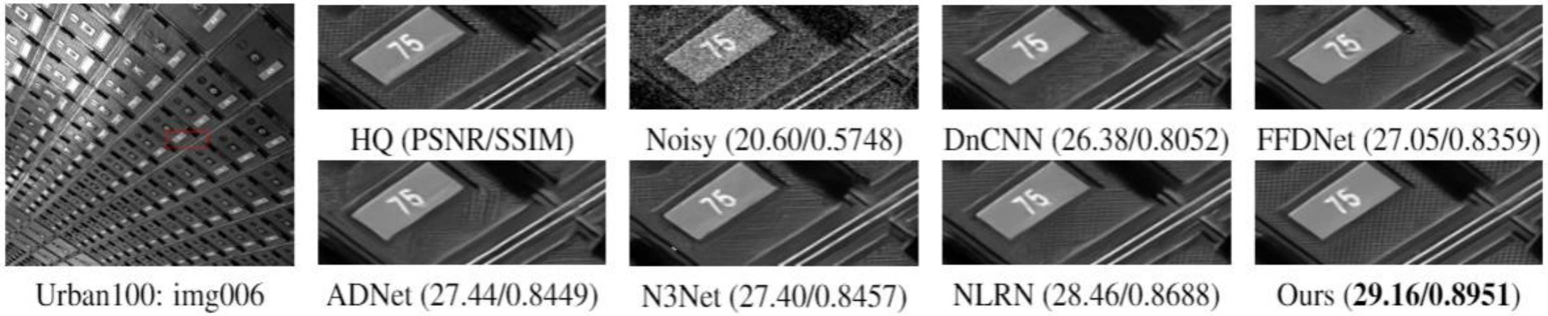


Table 3. The quantitative results on the DND benchmark.

Algorithm	Blind/Non-blind	PSNR	SSIM
BM3D [5]	Non-blind	34.51	0.851
CDnCNN-B [8]	Blind	32.43	0.790
FFDNet [9]	Non-blind	37.61	0.914
TWSC [28]	Blind	37.94	0.940
CBDNet [11]	Blind	38.06	0.942
SAT-Net (Ours)	Blind	**39.07**	**0.949**

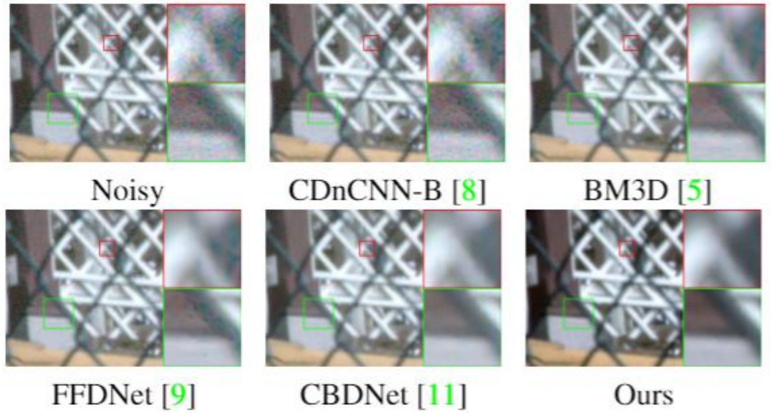


Fig. 5. Visual comparison of real image denoising application of various methods. These noisy images are corrupted by realistic noise from DND [29] dataset.

03 Experiments

Compression Artifact Reduction

Table 2. Quantitative results (PSNR and SSIM) of compression artifact reduction. Best results are **highlighted**.

Dataset	q	JPEG	ARCNN [24]	TNRD [25]	DnCNN [8]	RNAN [14]	SAT-Net (Ours)
Classic5	10	27.82/0.7800	29.04/0.7929	29.28/0.7992	29.40/0.8026	29.96/0.8178	**29.96/0.8180**
	20	30.12/0.8541	31.16/0.8517	31.47/0.8576	31.63/0.8610	32.11/0.8693	**32.18/0.8695**
	30	31.48/0.8844	32.52/0.8806	32.74/0.8837	32.91/0.8861	33.38/0.8924	**33.48/0.8929**

HQ (Classic5: Barbara) | JPEG (q=10) 25.78/0.7621 | ARCNN [24] 26.84/0.7980 | TNRD [25] 27.23/0.8099 | DnCNN [8] 27.59/0.8161 | RNAN [14] 28.86/0.8516 | Ours **29.04/0.8549**

Fig. 4. Visual comparison of image compression artifact reduction application of various methods with JPEG quality q = 10.

Ablation Study

This section mainly analyzes the effectiveness of the local attention submodule (LA) and non-local attention submodule (NLA) in our proposed SAT-Net. We first replace our patch-wise NLA with non-local neural networks [2] as the baseline. Then we separately train our SAT-Net without LA or NLA. These cases are trained with the same strategy and process for synthetic image denoising.

Table 4. Analysis of different components in SAT-Net. PSNR values are evaluated on Urban100 ($\sigma = 25$).

Mode	Baseline	w/o LA	w/o NLA	SAT-Net
PSNR	30.75	31.04	30.18	**31.17**

[2] Wang, Xiaolong, et al. "Non-local neural networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

04

Conclusion

1. Patch-wise Non-local Attention

We propose an effective and robust patch-wise non-local method to establish a more reliable long-range dependence during image restoration.

2. Mixed Attention Mechanism

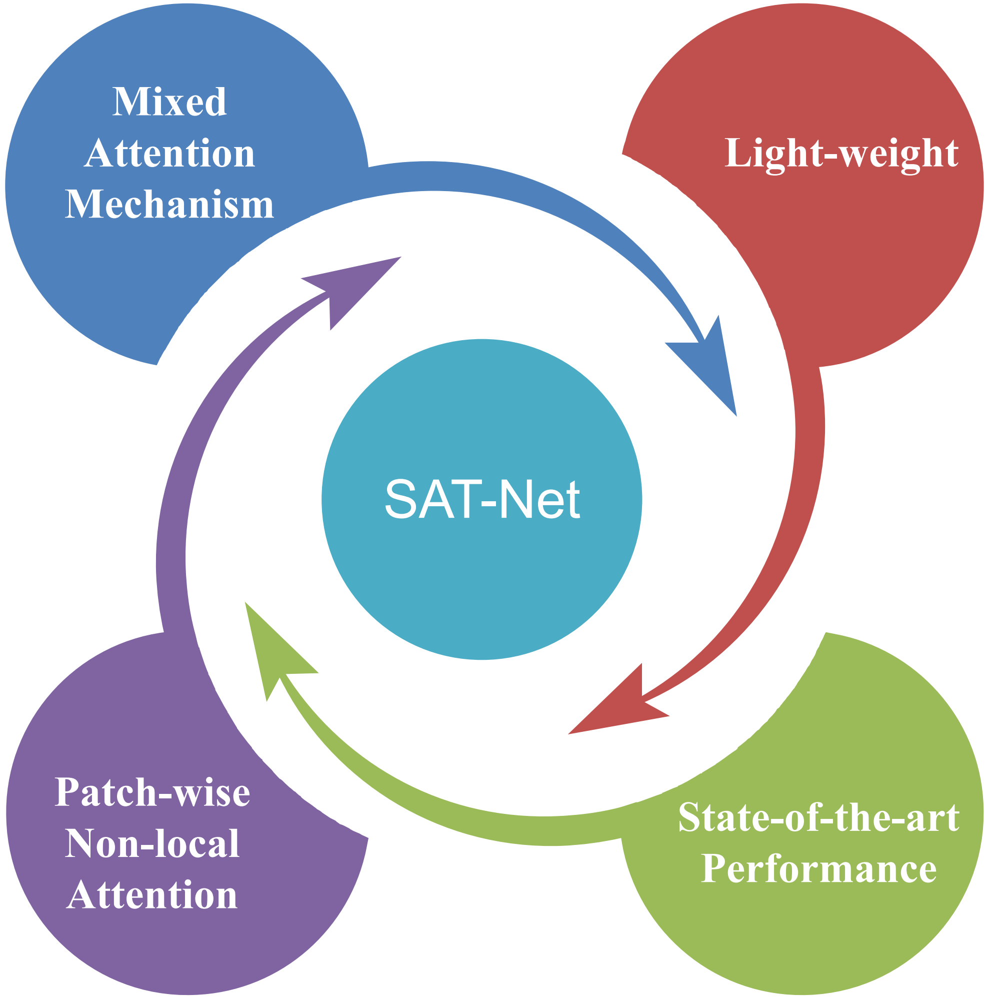
In this paper, we present an adaptive way to combine local and non-local attention operations to restore complex textures and repetitive details distinguishingly.

3. Light-weight

Our proposed SAT-Net can achieve attractive performance while maintaining fewer parameters.

4. State-of-the-art Performance

Experimental results on synthetic image denoising, real image denoising, and compression artifact reduction tasks show that our SAT-Net can achieve state-of-the-art performance under objective and subjective evaluations.



THANK YOU!

zhangjian.sz@pku.edu.cn

北京大學 PEKING UNIVERSITY

VILLA Visual-Information Intelligent Learning LAB