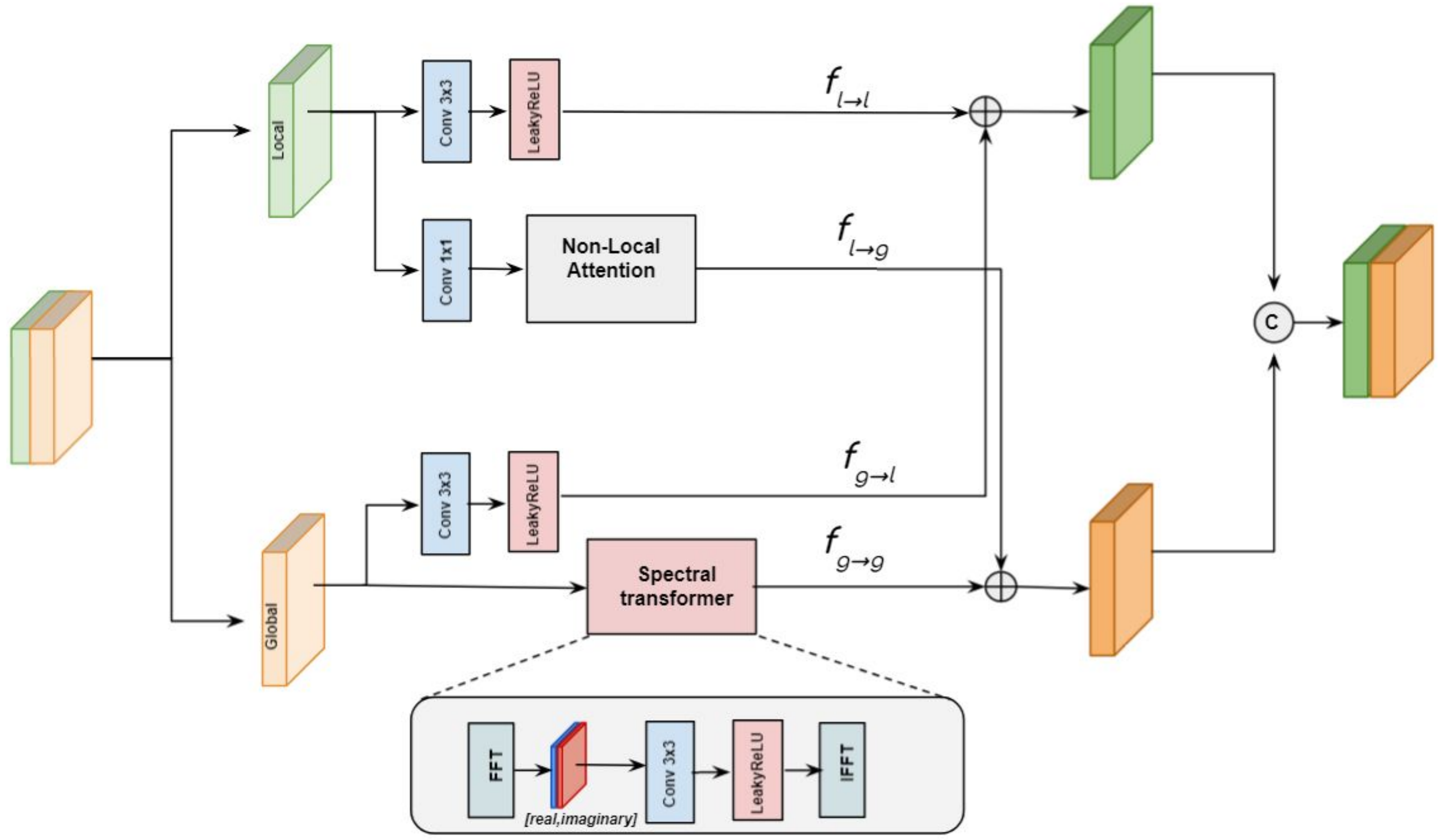
NL-FFC: Non-Local Fast Fourier Convolution for Image Super Resolution

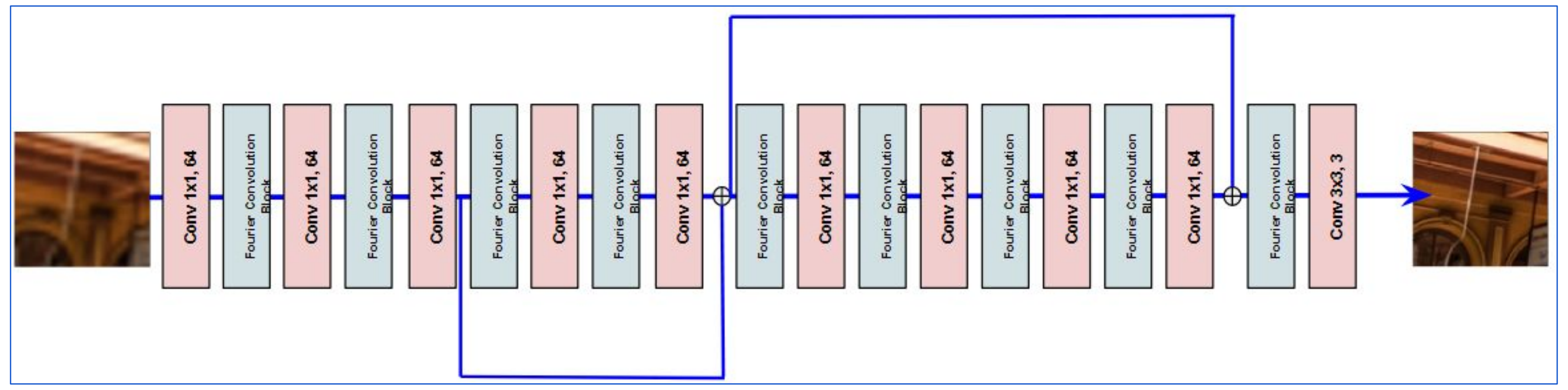
Abhishek Kumar Sinha, S. Manthira Moorthi, Debajyoti Dhar
SAC Ahmedabad

Objectives

The main objective is to study the effectiveness of Fourier features in learning long range dependencies for image super-resolution. The role of different features in the image quality is further analyzed to enhance the super resolution quality.

Non-Local Fast Fourier Convolution





The input features are split into the set of local and global features to learn local-to-local, local-to-global, global-to-local, and global-to-global mappings. Since the local features do not require long-range dependency, they are directly mapped to new local features using a vanilla convolution layer. For local-to-global mapping, it explores the global dependency for each query pixel using non-local attention model. The global-to-global mapping requires that the features are first transformed to frequency domain to widen the receptive field, and then transformation is applied to generate new global features. The new learnt features are mapped back to spatial domain to generate new set of global features. The split ratio is a hper-pramater and best one is found through ablation study.

Interesting Observations

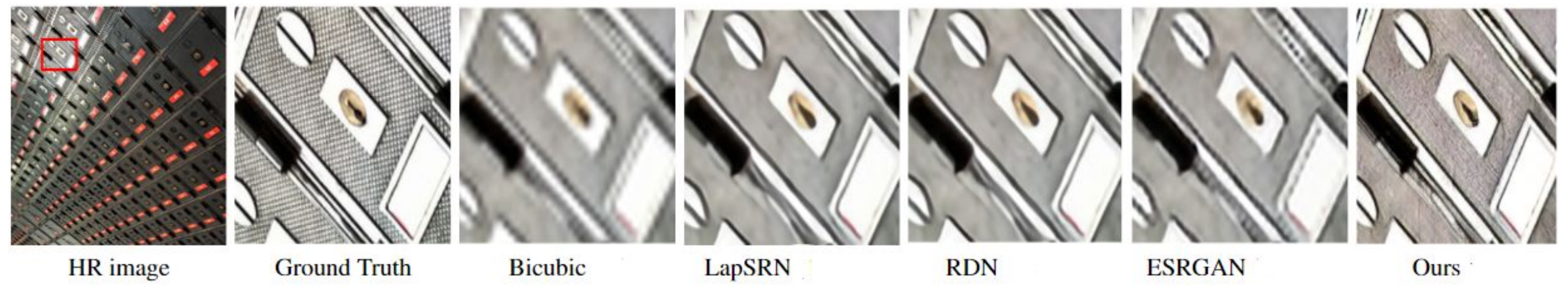
Fraction of Global Features (α)	PSNR (dB)	SSIM	NIQE
0.2	28.21	0.8676	6.14
0.3	28.97	0.8719	5.88
0.4	29.68	0.8801	5.04
0.5	32.76	0.9018	4.89
0.6	31.60	0.8961	5.01
0.7	30.16	0.8921	5.69

Experimental results show that the method performs when there is a balance between global and local features. It is observed that excessive global features cause over-smoothing whereas excessive local features causes unbalanced contrast and staggered edges.

Method	Non-Local Attention	Adversarial Learning	PSNR (dB)	SSIM	NIQE
NL-FFC	✔	✔	26.81	0.8651	4.89
FFC	X	✔	25.89	0.8524	5.10
NL-FFC	✔	X	32.760	0.9018	5.97
FFC	X	X	30.14	0.8876	6.14

The role of Non-Local Attention in the global-to-global mapping is evaluated in terms of both distortion and perceptual quality. For a given training condition (with or without adversarial learning), Non-Local Attention always improves the perceptual quality and reduces distortion.

Results & Conclusions



	DBPN	EDSR	SAN	CSNLN	Ours
Params	10M	43M	16M	3M	0.423M
FLOPS(G)	5209.4	1338.8	3835.9	2245.2	423.2

Complexity Analysis

The proposed method and previous works have already shown remarkable improvement in image quality by using long range features. However, these approaches are still limited by the computational complexity. The future scope of this work involves efficient global features learning algorithms for real-time super resolution.

References

1. Lu Chi, Borui Jiang, and Yandong Mu, Fast fourier convolution, NeurIPS, 2021
2. M. Haris, Gregory, Shakhnarovich. and Norimichi Ukita, Deep Back Projection Networks for image super resolution, CVPR, 2017
3. Bee Lim, Sanghyun Sun, Heewon Kim, Seungjan Nah, Kyoung Mu Le, Enhanced deep resiual network for single image super resolution, CVPRW, 2017
4. Tao Dai, Jianrui Cai, Yongbing Zhang, Shu-Tao Xia,, and Lei Zhang. Second-order attention network for single image super resolution, CVPR, 2019
5. Ian J. Goodfellow, Jean Pouget Abedie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville,, and Yoshua Bengio. Generative adversarial nets, NIPS 2014
6. Yiqun Mei, Yuchen Fan, Yuqian Zhou, Lichao Huang, Thomas S. Huang, and Humphrey Shi Ukita, Image super- resolution with cross-scale non-local attention and exhaus-tive self-exemplars mining, CVPR 2020