Efficient Information Modulation Network for Image Super-Resolution

Xiao Liu, Xiangyu Liao, Xiuya Shi, Linbo Qing and Chao Ren
College of Electronics and Information Engineering, Sichuan University, Chengdu, China

SICHUAN UNIVERSITY

Codes Wechat

Motivations and Contribution

Motivations:

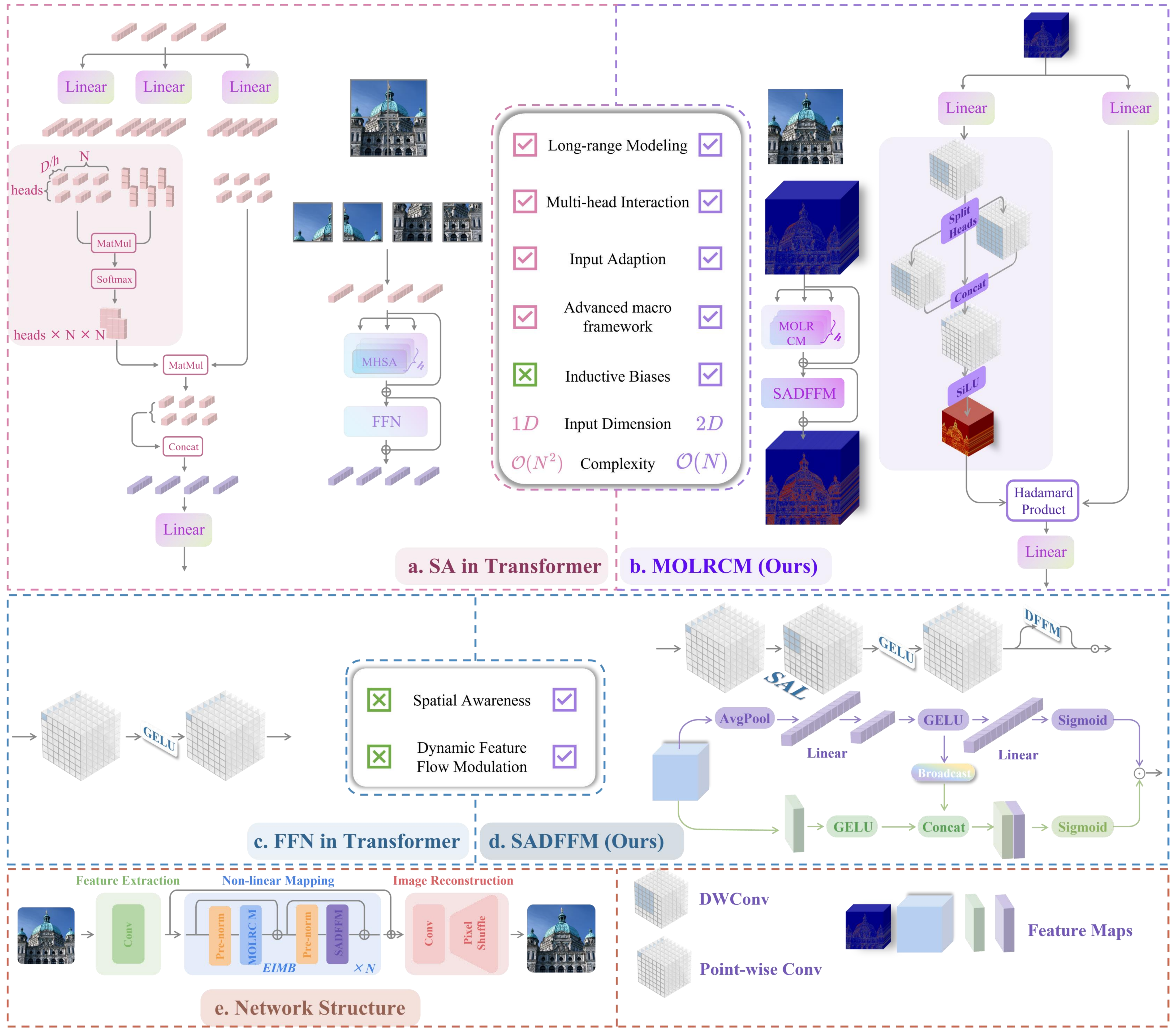
- **Insights 1: Self-attention is may not be all you need:** Recent researches have shown that the success of Transformers comes from their macro-level framework and advanced components, not their SA mechanism. Comparable results can be obtained by replacing SA with spatial pooling, shifting, MLP, fourier transform and constant matrix, all of which have spatial information aggregation capability like SA.
- **Insights 2: Large kernel is may be all you need:** Throughout the evolution of ConvNets, the usage of large kernel convolutions experienced fluctuations. Transformer has powerful long-range modeling capabilities, but the quadratic complexity makes it difficult to practically apply. Large kernel convolutions have regained attention due to their efficient implementation and the emergence of advanced computing hardware. Recently, a large amount of work has been devoted to exploring the potential of large kernel convolution, while demonstrating the effectiveness of large kernel convolution in terms of performance and computational complexity.

Contributions:

- We present a novel approach, named EIMN, to achieve efficient SISR that leverages the potential of large kernel ConvNets and advanced Transformer macro framework. We rethink a new spatial information aggregation technology for integrating spatial features efficiently by introducing large kernel convolution operation to realize long-range correlations and input content adaptation.
- MOLRCM and SADFFM modules are designed based on the analysis of the generation process of SA and the sub-optimality of vanilla FFN. The former utilizes large kernel convolution modulation technology to encode long-range and multi-order spatial information as a weight matrix, and self-adaptively recalibrates value features. The latter introduces spatial awareness and locality, improves feature diversity, and dynamically regulates the flow of information between layers compared to vanilla FFN.

Method

Network Architecture:



Contact

Hey! I'm Xiao Liu, a Master at Sichuan University. My research interests include computer vision and image restoration. Should you have any question, please contact at liuxmail1220@gmail.com.

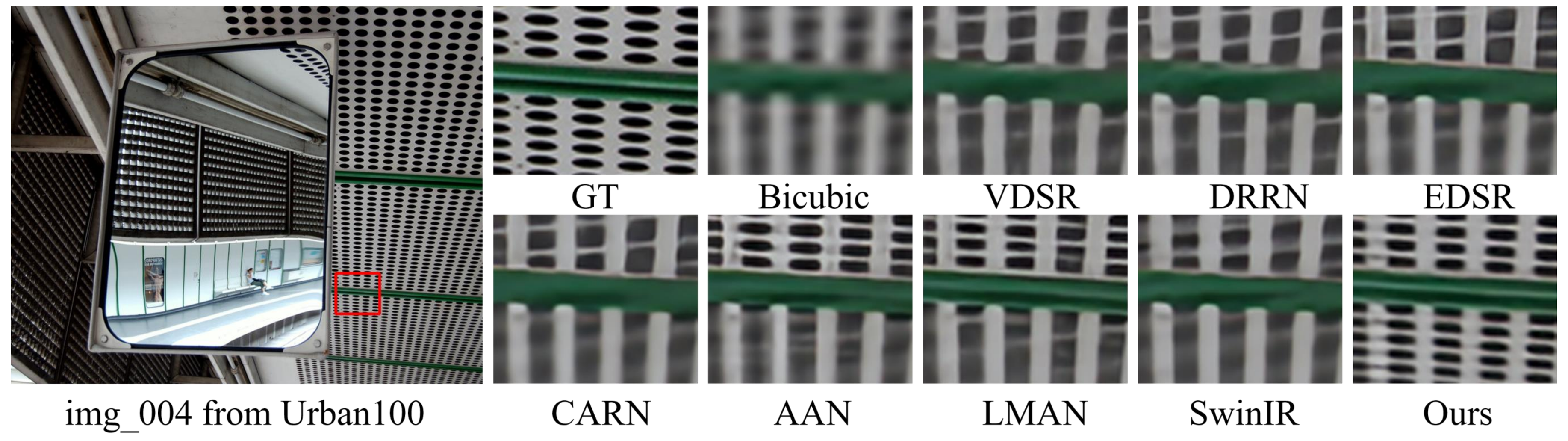
Results

Quantitative comparisons on the Cityscapes dataset:

Table 1: Quantitative comparison with SOTA methods on five popular benchmark datasets. Red text indicates the best and blue text indicates the second best PSNR/SSIM results, respectively. 'Multi-Adds' is calculated with a 1280 × 720 HR image.

Method	Scale	#Params(K)	Multi-Adds(G)	Set5	Set14	BSDS100	Urban100	Manga109
EDSR-baseline [20]	×2	1370	316	37.99/0.9604	33.57/0.9175	32.16/0.8994	31.98/0.9272	38.54/0.9769
SRFBN-S [18]	×2	282	574.4	37.78/0.9597	33.35/0.9156	32.00/0.8970	31.41/0.9207	38.06/0.9757
SMSR [37]	×2	985	131.6	38.00/0.9601	33.64/0.9179	32.17/0.8990	32.19/0.9284	38.76/0.9771
A2N [3]	×2	1036	247.5	38.06/0.9608	33.75/0.9194	32.22/0.9002	32.43/0.9311	38.87/0.9769
LMAN [35]	×2	1531	347.1	38.08/0.9608	33.80/0.9023	32.22/0.9001	32.42/0.9302	38.92/0.9772
SwinIR [19]	×2	878	195.6	38.14/0.9611	33.86/0.9206	32.31/0.9012	32.76/0.9340	39.12/0.9783
B-GSCN 10 [22]	×2	1490	343	38.04/0.9606	33.64/0.9182	32.19/0.8999	32.19/0.9293	38.64/0.9771
DRSDN [4]	×2	1055	243.1	38.06/0.9607	33.65/0.9189	32.23/0.9003	32.40/0.9308	-
FPNet [8]	×2	1615	-	38.13/0.9619	33.83/0.9198	32.29/0.9018	32.04/0.9278	-
PILN [29]	×2	580	-	38.08/0.9607	33.72/0.9181	32.23/0.9003	32.38/0.9306	38.92/0.9771
NGswin [5]	×2	998	140.4	38.05/0.9610	33.79/0.9199	32.27/0.9008	32.53/0.9324	38.97/0.9777
EIMN-A(Ours)	×2	860	186.3	**38.26/0.9619**	**34.12/0.9222**	**32.40/0.9034**	**33.15/0.9373**	**39.48/0.9788**
EIMN(Ours)	×2	981	212.7	**38.26/0.9620**	**34.14/0.9227**	**32.41/0.9034**	**33.23/0.9381**	**39.42/0.9786**
EDSR-baseline [20]	×3	1555	160	34.37/0.9270	30.28/0.8417	29.09/0.8052	28.15/0.8527	33.45/0.9439
SRFBN-S [18]	×3	375	686.4	34.20/0.9255	30.10/0.8372	28.96/0.8010	27.66/0.8415	33.02/0.9404
SMSR [37]	×3	993	67.8	34.40/0.9270	30.33/0.8412	29.10/0.8050	28.25/0.8536	33.68/0.9445
A2N [3]	×3	1036	1175	34.47/0.9279	30.44/0.8437	29.14/0.8059	28.41/0.8570	33.78/0.9458
LMAN [35]	×3	1718	173.8	34.56/0.9286	30.46/0.8439	29.17/0.8067	28.47/0.8576	34.00/0.9470
SwinIR [19]	×3	886	872	34.60/0.9289	30.54/0.8463	29.20/0.8082	28.66/0.8624	33.98/0.9097
B-GSCN 10 [22]	×3	1510	154	34.30/0.9271	30.35/0.8425	29.11/0.8035	28.20/0.8535	33.54/0.9445
DRSDN [4]	×3	1071	109.8	34.48/0.9282	30.41/0.8445	29.17/0.8072	28.45/0.8589	-
FPNet [8]	×3	1615	-	34.48/0.9285	30.53/0.8454	29.20/0.8086	28.19/0.8534	-
PILN [29]	×3	588	-	34.39/0.9269	30.34/0.8415	29.08/0.8048	28.09/0.8500	33.68/0.9446
NGswin [5]	×3	1007	66.6	34.52/0.9282	30.53/0.8456	29.19/0.8078	28.52/0.8603	33.89/0.9470
EIMN-A(Ours)	×3	868	83.58	**34.70/0.9299**	**30.65/0.8481**	**29.31/0.8121**	**28.87/0.8660**	**34.45/0.9492**
EIMN(Ours)	×3	990	95.2	**34.76/0.9304**	**30.70/0.8490**	**29.33/0.8127**	**29.05/0.8698**	**34.60/0.9502**
EDSR-baseline [20]	×4	1518	114	32.09/0.8938	28.58/0.7813	27.57/0.7357	26.04/0.7849	30.35/0.9067
SRFBN-S [18]	×4	483	852.9	31.98/0.8923	28.45/0.7779	27.44/0.7313	25.71/0.7719	29.91/0.9008
SMSR [37]	×4	1006	41.6	32.12/0.8932	28.55/0.7808	27.55/0.7351	26.11/0.7868	30.54/0.9085
A2N [3]	×4	1047	72.4	3230/0.8966	28.71/0.7842	27.61/0.7374	26.27/0.7920	30.67/0.9110
LMAN [35]	×4	1673	122.0	32.40/0.8974	28.72/0.7842	27.66/0.7388	26.36/0.7934	30.84/0.9129
SwinIR [19]	×4	897	49.6	32.44/0.8976	28.77/0.7858	27.69/0.7406	26.47/0.7980	30.92/0.9151
B-GSCN 10 [22]	×4	1530	88	32.18/0.8950	28.60/0.7821	27.59/0.7364	26.12/0.7872	30.50/0.9080
DRSDN [4]	×4	1095	63.1	32.28/0.8962	28.64/0.7836	27.64/0.7388	26.30/0.7933	-
FPNet [8]	×4	1615	-	32.32/0.8962	28.78/0.7856	27.66/0.7394	26.09/0.7850	-
PILN [29]	×4	600	-	32.22/0.8949	28.62/0.7813	27.59/0.7365	26.19/0.7878	30.54/0.9086
NGswin [5]	×4	1019	36.4	32.33/0.8963	28.78/0.7859	27.66/0.7396	26.45/0.7963	30.80/0.9128
EIMN-A(Ours)	×4	880	47.78	**32.53/0.8993**	**28.89/0.7882**	**27.79/0.7447**	**26.68/0.8027**	**31.22/0.9148**
EIMN(Ours)	×4	1002	54.37	**32.63/0.9008**	**28.94/0.7897**	**27.82/0.7458**	**26.88/0.8084**	**31.52/0.9183**

Qualitative comparisons on the Cityscapes dataset:



Visual analysis: