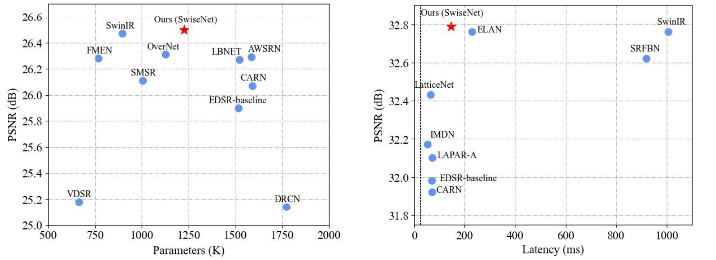
Lightweight Image Super-Resolution with Scale-wise Network

Xiaole Zhao Xinkun Wu

Southwest Jiaotong University, China

Problem



- Current CNN-based super-resolution (SR) network depends on a significant number of model parameters, which results in increased computational requirements and memory consumption during the training process.
- Most conventional upsampling modules utilize single-layer PixelShuffle or Bicubic, which leads to a loss of feature information. And the missing information is also crucial for reconstructing high quality image networks.

Contribution

- A lightweight recursive feature extractor that improves performance even in the most advanced models.
- A Scale-wise Upsample module (SUM) to retain multi-scale information that helps restore HR images accurately.
- SUM favors lightweight model design. Based on this, we construct a lightweight SwiseNet, using the multi-scale information fusion strategy to extract multi-scale context information.
- Replacing the upsampling operation with our SUM produces better SR results for the baseline model.

Approach

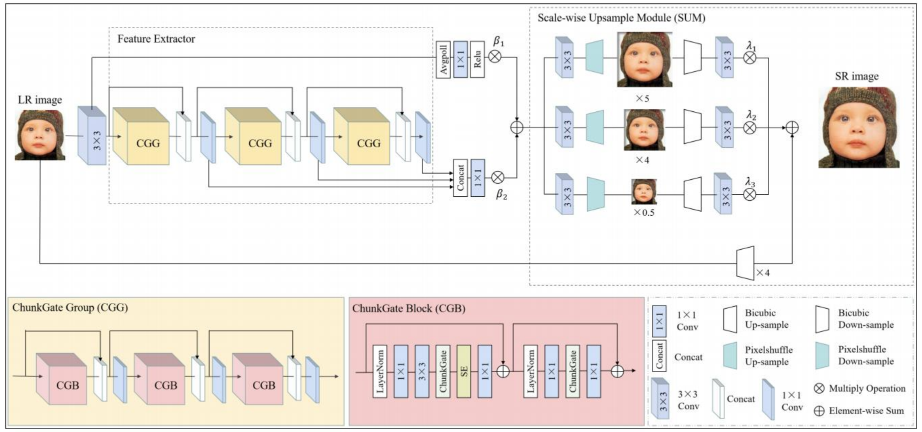


Figure 3: The overall structure of our proposed SwiseNet model. As the maximum scaling factor in this example is set to N = 4, the required scaling-factors are ×4, ×5 and ×0.5.

- To deal with all the mentioned challenges, we proposed a lightweight SR model that includes several continuous feature extractors and a novel upsampling module.
- Our proposed method, called multi-scale fusion, utilizes feature information from multiple scales simultaneously for up-sampling. This approach has the advantage of integrating information from different scales to provide a more comprehensive understanding of the image features, leading to improved performance and network robustness.
- After many ablation experiments, we have the following findings: multiple up-sampled information of different scales can make the model produce better SR results, but at the same time, it will sacrifice some computational efficiency.
- Based on these findings, we employed a scale transformation approach with the target scale as the reference in three different directions, including magnification, equivalence, and reduction. Subsequently, we adjusted them to the target scale, thereby providing additional information for the same pixel and enhancing the effectiveness of SR.

Experimental Results

Table 1: Average PSNR/SSIM for 2×, 3×, 4× SR. The best results are highlighted in red color and the second best is in blue.

Method	Scale	#Params[K]	#FLOPs[G]	Set5	Set14	BSD100	Urban100
				PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
Bicubic	×2	-	-	33.66 / 0.9299	30.24 / 0.8688	29.56 / 0.8431	26.88 / 0.8403
SRCNN[8]		57	52.7	36.66 / 0.9542	32.45 / 0.9067	31.36 / 0.8879	29.50 / 0.8946
CARN[2]		1592	222.8	37.76 / 0.9590	33.52 / 0.9166	32.09 / 0.8978	31.92 / 0.9256
SMSR[29]		985	224.1	38.00 / 0.9601	33.64 / 0.9179	32.17 / 0.8990	32.19 / 0.9284
LBNET[11]		731	153.2	38.05 / 0.9607	33.65 / 0.9177	32.16 / 0.8994	32.30 / 0.9291
FMEN[10]		748	172.0	38.10 / 0.9609	33.75 / 0.9192	32.26 / 0.9007	32.41 / 0.9311
SwinIR[19]		878	195.6	38.14 / 0.9610	33.86 / 0.9206	32.31 / 0.9012	32.76 / 0.9338
SwiseNet(w/o SUM)		1035	212.7	38.10 / 0.9605	33.85 / 0.9207	32.22 / 0.9008	32.64 / 0.9326
SwiseNet		1077	236.5	38.18 / 0.9610	33.94 / 0.9217	32.28 / 0.9012	32.79 / 0.9339
Bicubic	×3	-	-	33.39 / 0.8682	27.55 / 0.7742	27.21 / 0.7385	24.46 / 0.7349
SRCNN[8]		57	52.7	32.75 / 0.9090	29.30 / 0.8215	28.41 / 0.7863	26.24 / 0.7989
CARN[2]		1592	118.8	34.29 / 0.9255	30.29 / 0.8407	29.06 / 0.8034	28.06 / 0.8493
SMSR[29]		993	100.5	34.40 / 0.9270	30.33 / 0.8412	29.10 / 0.8050	28.25 / 0.8536
LBNET[11]		736	68.4	34.47 / 0.9277	30.38 / 0.8417	29.13 / 0.8061	28.42 / 0.8559
FMEN[10]		757	77.2	34.45 / 0.9275	30.40 / 0.8435	29.17 / 0.8063	28.33 / 0.8562
SwinIR[19]		886	87.2	34.60 / 0.9289	30.54 / 0.8463	29.19 / 0.8082	28.66 / 0.8624
SwiseNet(w/o SUM)		1044	112.8	34.43 / 0.9268	30.44 / 0.8442	29.11 / 0.8066	28.45 / 0.8590
SwiseNet(Ours)		1068	128.0	34.61 / 0.9289	30.52 / 0.8453	29.20 / 0.8079	28.63 / 0.8615
Bicubic	×4	-	-	28.42 / 0.8104	26.00 / 0.7027	25.96 / 0.6675	23.14 / 0.6577
SRCNN[8]		57	52.7	30.48 / 0.8628	27.50 / 0.7513	26.90 / 0.7101	24.52 / 0.7221
CARN[2]		1592	90.9	32.13 / 0.8937	28.60 / 0.7806	27.58 / 0.7349	26.07 / 0.7837
SMSR[29]		1006	57.2	32.12 / 0.8932	28.55 / 0.7808	27.55 / 0.7351	26.11 / 0.7868
LBNET[11]		742	38.9	32.29 / 0.8960	28.68 / 0.7832	27.62 / 0.7382	26.27 / 0.7906
FMEN[10]		769	44.2	32.24 / 0.8955	28.70 / 0.7839	27.63 / 0.7379	26.28 / 0.7908
SwinIR[19]		897	49.6	32.42 / 0.8976	28.77 / 0.7858	27.68 / 0.7406	26.47 / 0.7980
SwiseNet(w/o SUM)		1056	75.3	32.31 / 0.8955	28.65 / 0.7829	27.57 / 0.7363	26.35 / 0.7928
SwiseNet(Ours)		1227	88.6	32.43 / 0.8979	28.72 / 0.7844	27.68 / 0.7396	26.48 / 0.7969