Lightweight Real-Time Image Super-Resolution Network for 4K Images

Ganzorig Gankhuyag*, Kihwan Yoon*, Jinman Park, Haeng Seon Son, Kyoungwon Min
Korea Electronics Technology Institute

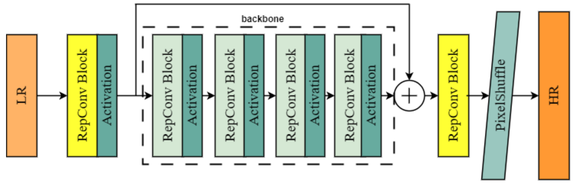
JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

Previous Problem & Contribution:

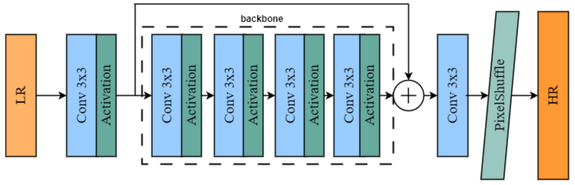
- Previous Problem
  - Previous SISR networks have high computational complexity with insufficient accuracy
- Contribution
  - We proposed the lightweight real-time image super resolution (LRSRN) network structure that simultaneously achieves high accuracy and real-time speed.
    - low computational complexity and high accuracy compared to traditional SISR methods.
  - We employed a reparameterized convolution (RepConv) layer, which enhances image quality while maintaining model size and inference speed.

Network Architecture

- We applied RepConv to each convolution layer, which is a more efficient method



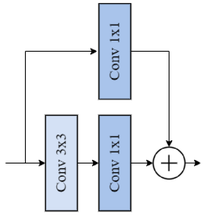
(a) The Training mode of proposed network



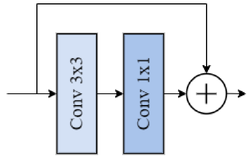
(b) The inference mode of proposed network

Reparameterized Block:

- We applied RepConv to each convolution layer, which is a more efficient method.
- We applied an advanced version of the RepConv block when In/Out channels are not equals.

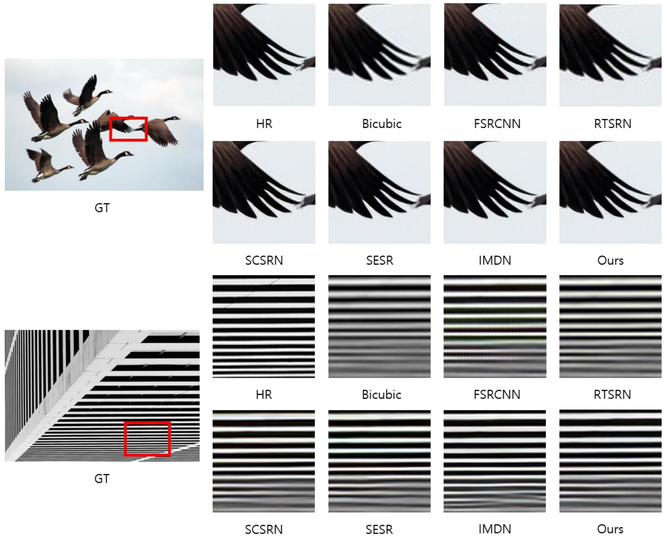


(a) In/Out channels are equals



(b) In/Out channels are not equals

Qualitative Results



Quantitative Results

Scale	Network	Set5			Set14			B100			Urban100			DIV2K		
		PSNR	SSIM	PSNR (Y)	PSNR	SSIM	PSNR (Y)	PSNR	SSIM	PSNR (Y)	PSNR	SSIM	PSNR (Y)	PSNR	SSIM	PSNR (Y)
X2	Bicubic	29.96	0.8676	31.86	27.38	0.8121	28.58	27.67	0.8169	28.05	24.98	0.8069	25.53	29.85	0.8662	30.79
	FSRCNN [13]	31.36	0.8892	33.41	28.38	0.8335	28.81	28.57	0.8379	29.01	26.44	0.8440	27.16	30.39	0.8849	31.95
	SESR [5]	31.71	0.8944	33.98	28.78	0.8423	30.29	28.94	0.8461	29.37	27.38	0.8641	28.20	31.40	0.8923	32.47
	IMDN [20]	32.24	0.9022	34.32	29.34	0.8525	30.66	29.26	0.8521	29.65	28.38	0.8821	29.18	31.98	0.9006	32.90
	RTSRN [40]	30.33	0.8713	32.31	27.72	0.8196	29.06	28.02	0.8171	28.44	25.41	0.8171	26.05	30.20	0.8719	31.23
	SCSRN [16]	31.72	0.8952	33.99	28.78	0.8426	30.29	28.92	0.8466	29.35	27.34	0.8640	28.15	31.41	0.8925	32.46
	Proposed work	31.84	0.8969	33.92	28.84	0.8436	30.26	28.93	0.8470	29.34	27.35	0.8641	28.11	31.44	0.8933	32.41
X3	Bicubic	27.28	0.7962	28.84	25.16	0.7190	26.11	25.43	0.7098	25.76	22.71	0.7040	23.13	27.42	0.7918	28.16
	FSRCNN [13]	28.41	0.8284	30.03	25.96	0.7436	27.06	26.03	0.7314	26.39	23.63	0.7430	24.12	28.17	0.8127	28.96
	SESR [5]	29.05	0.8429	30.84	26.39	0.7590	27.51	26.40	0.7432	26.73	24.37	0.7713	24.91	28.62	0.8240	29.42
	IMDN [20]	29.68	0.8559	31.32	26.86	0.7714	27.84	26.67	0.7514	26.96	25.13	0.7961	25.63	29.13	0.8357	29.79
	RTSRN [40]	27.55	0.7997	29.16	25.4	0.7263	26.43	25.62	0.7196	25.98	22.97	0.7137	23.44	27.63	0.7974	28.43
	SCSRN [16]	29.06	0.8431	30.84	26.41	0.7597	27.53	26.40	0.7449	26.74	24.41	0.7728	24.94	28.65	0.8253	29.44
	Proposed work	29.13	0.8459	30.75	26.47	0.7606	27.51	26.40	0.7443	26.72	24.40	0.7713	24.88	28.69	0.8255	29.41

Qualitative results comparison on benchmark datasets.

Ablation Study

Backbone	Channels	Patch Sizes	RepConv	Fine-Tune	NTIRE2023 val PSNR	NTIRE2023 val SSIM	Inference Time (ms)	Scores
5	64	96	X	X	31.897	0.9291	26.62	4.47
4	64	96	X	X	31.920	0.9296	22.19	4.98
4	32	96	X	X	31.909	0.9295	9.77	7.44
4	32	192	X	X	31.900	0.9294	9.77	7.40
4	32	192	O	X	32.784	0.9382	9.77	13.66
4	32	192	O	DIV2K [1]	32.812	0.9386	9.77	13.92
4	32	192	O	Proposed work	32.831	0.9388	9.77	14.11

Ablation study results on DIV2K val dataset.

Conclusion

- Our proposed network employs a reparameterization method that enhances the quality of super-resolved images without affecting the inference time performance.
- Our proposed network achieves a PSNR of 30.15 dB and an inference time of 4.75 ms on an RTX 3090Ti device on scale X3 dataset

Paper and Code

Code link

Paper link