Multi-scale Residual Network for Image Super-Resolution

Juncheng Li, Faming Fang, Kangfu Mei, Guixu Zhang

East China Normal University, Jiangxi Normal University

ECCV 2018 European Conference on Computer Vision

8 – 14 September 2018 | Munich, Germany

Background

SISR aims to reconstruct a high-resolution (HR) image from a low-resolution (LR) image, which is an ill-posed problem since the mapping between LR and HR has multiple solutions.

Currently, convolutional neural networks (CNNs) have indicated that they can provide remarkable performance in the SISR problem. More and more models tend to construct deeper and more complex network structures, which means training these models consumes more resources, time, and tricks.

We also found most existing SR models have the following problems:

(a) Hard to Reproduce.

(b) Inadequate of Features Utilization.

(c) Poor Scalability.

Contribution

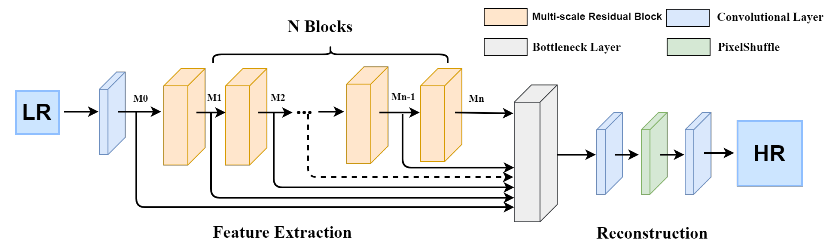
☞We propose a novel multi-scale residual block (MSRB), which can not only detect the image features adaptively, but also achieve feature fusion at different scales.

☞We extend our work to computer vision tasks and the results exceed most of state-of-the-art methods in SISR without deep network structure. MSRB can also be used for feature extraction in other restoration tasks which show promising results.

☞We propose a simple architecture for hierarchical features fusion (HFFS) and image reconstruction. It can be easily extended to any upscaling factors.

Method

☞ Multi-scale Residual Network

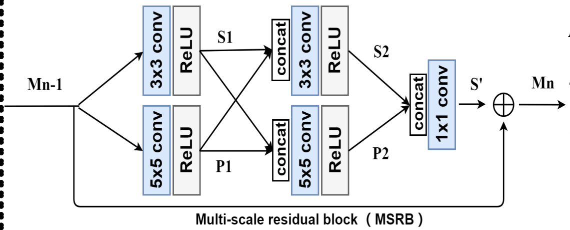


Our ultimate goal is to learn an end-to-end mapping function F between the I^{LR} and the I^{HR}. The loss function of our model can be defined as:

$$\hat{\theta} = arg\min_{\theta} \min \frac{1}{N}\sum_{i=1}^{N} \mathcal{L}^{SR}\left(F_\theta(I_i^{LR}), I_i^{HR}\right) \quad (1)$$

$$\mathcal{L}^{SR}\left(F_\theta(I_i^{LR}), I_i^{HR}\right) = \| F_\theta(I_i^{LR}) - I_i^{HR} \|_1 \quad (2)$$

☞ Multi-scale Residual Block



$$S_1 = \sigma(w_{3x3}^1 * M_{n-1} + b^1) \quad (3)$$

$$P_1 = \sigma(w_{5x5}^1 * M_{n-1} + b^1) \quad (4)$$

$$S_2 = \sigma(w_{3x3}^2 * [S_1, P_1] + b^2) \quad (5)$$

$$P_2 = \sigma(w_{5x5}^2 * [P_1, S_1] + b^2) \quad (6)$$

$$S' = \sigma(w_{1x1}^3 * [S_2, P_2] + b^3) \quad (7)$$

$$M_n = S' + M_{n-1} \quad (8)$$

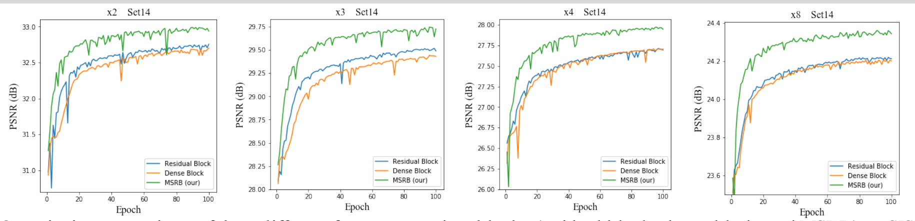
☞Hierarchical Feature Fusion Structure (HFFS)

We introduce a bottleneck layer which is essential for a convolutional layer with 1x1 kernel. The output of hierarchical feature fusion structure (HFFS) can be formulated as:

$$F_{LR} = \omega * [M_0, M_1, M_2, \ldots, M_N] + b \quad (9)$$

where M_0 is the output of the first convolutional layer, M_i $(i \neq 0)$ represents the output of the i^{th} MSRB, and $[M_0, M_1, M_2, \ldots, M_N]$ denotes the concatenation operation.

Quantitative comparison



Quantitative comparison of three different feature extraction blocks (residual block, dense block, and MSRB) on SISR.

Results

Algorithm	Scale	Set5 PSNR/SSIM	Set14 PSNR/SSIM	BSDS100 PSNR/SSIM	Urban100 PSNR/SSIM	Manga109 PSNR/SSIM
Bicubic	x2	33.69/0.9284	30.34/0.8675	29.57/0.8434	26.88/0.8438	30.82/0.9332
A+ [23]	x2	36.60/0.9542	32.42/0.9059	31.24/0.8870	29.25/0.8955	35.37/0.9663
SelfExSR [20]	x2	36.60/0.9537	32.46/0.9051	31.20/0.8863	29.55/0.8983	35.82/0.9671
SRCNN [1]	x2	36.71/0.9536	32.32/0.9052	31.36/0.8880	29.54/0.8962	35.74/0.9661
ESPCN [2]	x2	37.00/0.9559	32.75/0.9098	31.51/0.8939	29.87/0.9065	36.21/0.9694
FSRCNN [3]	x2	37.06/0.9554	32.76/0.9078	31.53/0.8912	29.88/0.9024	36.67/0.9694
VDSR [4]	x2	37.53/0.9583	33.05/0.9107	31.92/0.8965	30.79/0.9157	37.22/0.9729
DRCN [5]	x2	37.63/0.9584	33.06/0.9108	31.85/0.8947	30.76/0.9147	37.63/0.9723
LapSRN [6]	x2	37.52/0.9581	33.08/0.9109	31.80/0.8949	30.41/0.9112	37.27/0.9855
EDSR [9]	x2	38.11/0.9601	33.92/0.9195	32.32/0.9013	-/-	-/-
MSRN(our)	x2	38.08/0.9605	33.74/0.9170	32.23/0.9013	32.22/0.9326	38.82/0.9868
Bicubic	x3	30.41/0.8655	27.64/0.7722	27.21/0.7344	24.46/0.7411	26.96/0.8555
A+ [23]	x3	32.63/0.9085	29.25/0.8194	28.31/0.7828	26.05/0.8019	29.93/0.9089
SelfExSR [20]	x3	32.66/0.9089	29.34/0.8222	28.30/0.7839	26.45/0.8124	27.57/0.7997
SRCNN [1]	x3	32.47/0.9067	29.23/0.8201	28.31/0.7832	26.25/0.8028	30.59/0.9107
ESPCN [2]	x3	33.02/0.9135	29.49/0.8271	28.50/0.7937	26.41/0.8161	30.79/0.9181
FSRCNN [3]	x3	33.20/0.9149	29.54/0.8277	28.55/0.7945	26.48/0.8175	30.98/0.9212
VDSR [4]	x3	33.68/0.9201	29.86/0.8312	28.83/0.7966	27.15/0.8315	32.01/0.9310
DRCN [5]	x3	33.85/0.9215	29.89/0.8317	28.81/0.7954	27.16/0.8311	32.31/0.9328
LapSRN [6]	x3	33.82/0.9207	29.89/0.8304	28.82/0.7950	27.07/0.8298	32.21/0.9318
EDSR [9]	x3	34.65/0.9282	30.52/0.8462	29.25/0.8093	-/-	-/-
MSRN(our)	x3	34.38/0.9262	30.34/0.8395	29.08/0.8041	28.08/0.8554	33.44/0.9427
Bicubic	x4	28.43/0.8022	26.10/0.6936	25.97/0.6517	23.14/0.6599	24.91/0.7826
A+ [23]	x4	30.33/0.8565	27.44/0.7450	26.83/0.6999	24.34/0.7211	27.03/0.8439
SelfExSR [20]	x4	30.34/0.8593	27.55/0.7511	26.84/0.7032	24.83/0.7403	27.83/0.8598
SRCNN [1]	x4	30.50/0.8573	27.62/0.7453	26.91/0.6994	24.53/0.7236	27.66/0.8505
ESPCN [2]	x4	30.66/0.8646	27.71/0.7562	26.98/0.7124	24.60/0.7360	27.70/0.8560
FSRCNN [3]	x4	30.73/0.8601	27.71/0.7488	26.98/0.7029	24.62/0.7272	27.90/0.8517
VDSR [4]	x4	31.36/0.8796	28.11/0.7624	27.29/0.7167	25.18/0.7543	28.83/0.8809
DRCN [5]	x4	31.56/0.8810	28.15/0.7627	27.24/0.7150	25.15/0.7530	28.98/0.8816
LapSRN [6]	x4	31.54/0.8811	28.19/0.7635	27.32/0.7162	25.21/0.7564	29.09/0.8845
EDSR [9]	x4	32.46/0.8968	28.80/0.7876	27.71/0.7420	-/-	-/-
MSRN(our)	x4	32.07/0.8903	28.60/0.7751	27.52/0.7273	26.04/0.7896	30.17/0.9034
Bicubic	x8	24.40/0.6045	23.19/0.5110	23.67/0.4808	20.74/0.4841	21.46/0.6138
A+ [23]	x8	25.53/0.6548	23.99/0.5535	24.21/0.5156	21.37/0.5193	22.39/0.6454
SelfExSR [20]	x8	25.49/0.6733	24.02/0.5650	24.19/0.5146	21.81/0.5536	22.99/0.6907
SRCNN [1]	x8	25.34/0.6471	23.86/0.5443	24.14/0.5043	21.29/0.5133	22.46/0.6606
ESPCN [2]	x8	25.75/0.6738	24.21/0.5109	24.37/0.5277	21.59/0.5420	22.83/0.6715
FSRCNN [3]	x8	25.42/0.6440	23.94/0.5482	24.21/0.5112	21.32/0.5090	22.39/0.6357
VDSR [4]	x8	25.73/0.6743	23.20/0.5110	24.34/0.5169	21.48/0.5289	22.73/0.6688
DRCN [5]	x8	25.93/0.6743	24.25/0.5510	24.49/0.5168	21.71/0.5289	23.20/0.6686
LapSRN [6]	x8	26.15/0.7028	24.45/0.5792	24.54/0.5293	21.81/0.5555	23.39/0.7068
MSRN(our)	x8	26.59/0.7254	24.88/0.5961	24.70/0.5410	22.37/0.5977	24.28/0.7517

Algorithm	Feature extraction	Filters	Layers	Depth	Parameters	Updates	Channel
EDSR [9]	**32 blocks**	**256**	**69**	**69**	**43M**	1×10^6	RGB
MSRN (our)	8 blocks	64	44	28	6.3M	4×10^5	Y