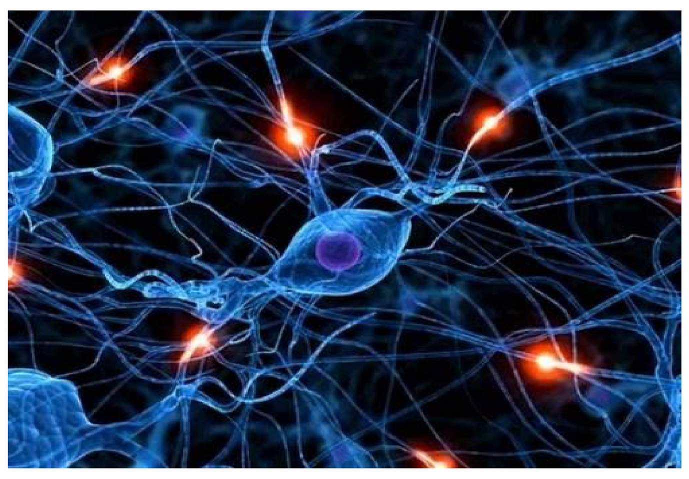


Lightweight Lateral Inhibition Network for Single MR Image Super-Resolution

Xiaole Zhao, Tao Zhang, Xueming Zou

zxlation@foxmail.com; taozhangjin@gmail.com; mark.zou@alltechmed.com

Brainformatics Lab
UESTC

Contributions

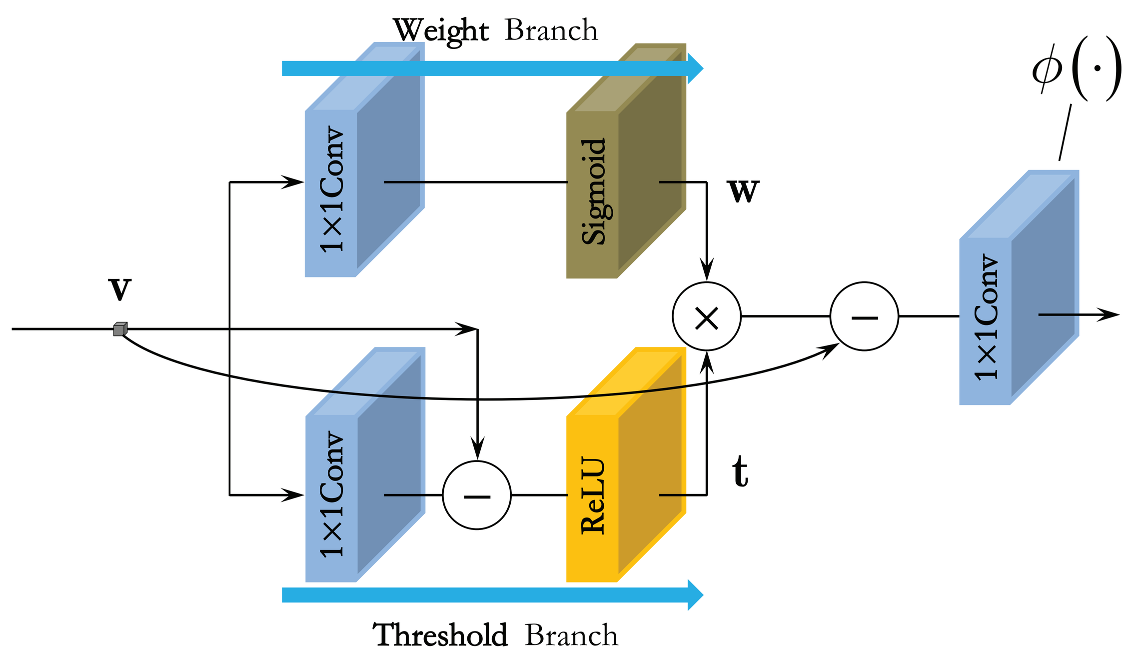
1. A lateral inhibition network (LIN) is proposed for single MR image SR. With a small amount of parameters and computational overhead, LIN can achieve accurate and fast SR reconstruction.
2. We conduct a inhibition unit (LIU) to impose inhibitory regulation on features explicitly, which is motivated by the lateral inhibition mechanism.
3. We propose to fuse the shallow features with different receptive field sizes, which increases the diversity of the extracted features and provide more effective evidence for nonlinear inference and image reconstruction.
4. We experimentally verify that combining the lateral inhibition mechanism with the proposed shallow feature extraction strategy contributes to improving the performance of deep models.

Explicit Visual Inhibition

A famous computing model for simulating visual inhibition is the Hartline-Ratliff Equation [1].

$$\hat{v}_i = \phi\left(v_i - \sum_{j \neq i} w_{ij} \cdot \max(0, v_j - t_{ij})\right), \qquad (1)$$

where $\hat{v}_i$ denotes the i-th element of the reversed feature map $\hat{\mathbf{v}}$, and v_i is the i-th element of input feature $\mathbf{v}$. w_{ij} indicates the inhibition coefficient of the j-th neuron on the i-th neuron, while t_{ij} denotes the threshold that the j-th neuron must reach to inhibit the i-th neuron. $\phi(\cdot)$ is a linear correction function, which we add to the original Hartline-Ratliff Equation. The symbols, i.e., $\otimes$ and $\ominus$, in the following figure are element-wise operations.



Inhibition tail (IT) of a lateral inhibition unit (LIU): The upper branch imitates the weight tensor **w** in the Hartline-Ratliff Equation, while the lower branch simulates the threshold tensor **t**.

References

[1] Hartline, H., Ratliff, F.: Studies on Excitation and Inhibition in the Retina. The Rockefeller University Press, New York, (1974)

[2] Manjón, J. V., Coupé, P., Buades, A., Fonov, V., Collins, D. L., and Robles, M.: Non-local MRI upsampling. Medical Image Analysis, **14**(6), 784-792, (2010)

[3] Dong, C., Loy, C. C., He, K., and Tang, X.: Image super-resolution using deep convolutional networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, **38**(2), 295-307, (2016)

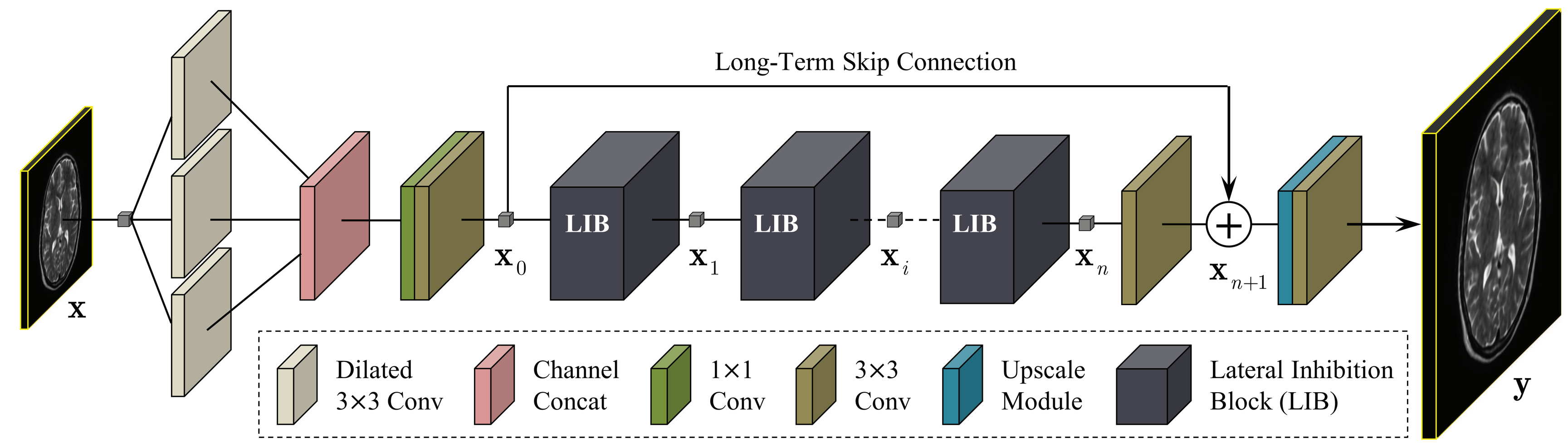
[4] Kim, J., Lee, J. K., Lee, K. M.: Accurate image super-resolution using very deep convolutional networks. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1646-1654 (2016)

[5] Hyun, C. M., Kim, H. P., Lee, S. M., Lee, S., and Seo, J. K.: Deep learning for undersampled MRI reconstruction. Physics in Medicine and Biology, **63**(13), 135007, (2018)

[6] Shi, J., Li, Z., Ying, S., Wang, C., Liu, Q., Zhang, Q., and Yan, P.: MR image super-resolution via wide residual networks with fixed skip connection. IEEE Journal of Biomedical and Health Informatics, **23**(3), 1129-1140, (2018)

Network Architecture

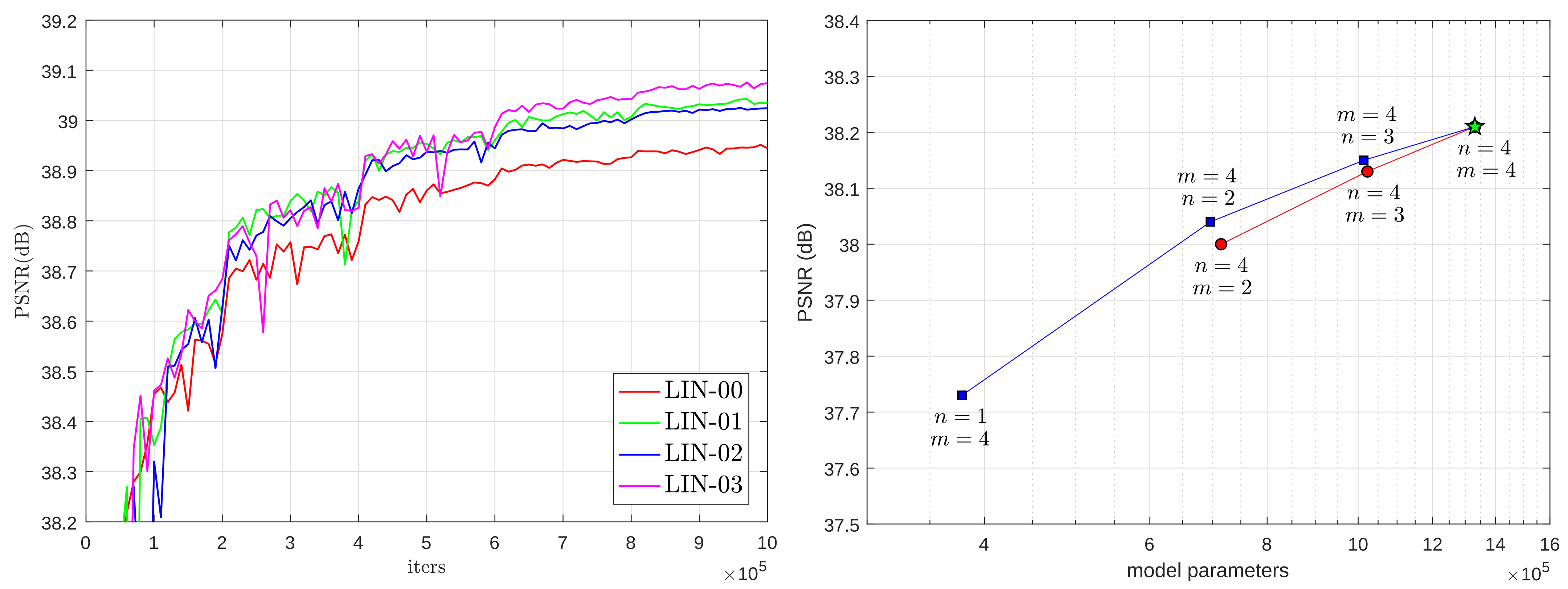
The network architecture is shown in the following figure. Feature extraction is composed of a set of parallel dilated 3×3 conv layers followed by a 1×1 conv and a 3×3 conv layer. The outputs of these dilation convolutions are concatenated together along the channel direction. The extracted shallow feature is fed into the nonlinear mapping part of the network, which consists of a series of cascaded lateral inhibition blocks (LIBs). Here, residual learning is adopted to stabilize model training. Image reconstruction includes a upscale module, which is usually followed by a 3×3 conv layer.



The overall network structure: The overall architecture of our LIN network. The dilated 3×3 convolutions for feature extraction have different dilation rates to collect features in the receptive fields with different sizes.

Ablation Investigation and Model Analysis

We keep the backbone of the network unchanged and adjust FE and IT accordingly. For FE, the comparative case is a single 3×3 conv layer that is denoted as "0", and a group of dilated 3×3 conv layers is denoted as "1". As for IT, "0" stands for removing the IT from a LIU, the opposite is denoted as "1". The valid curves of these different configurations are shown in the left figure. We also analysis the impact of the number of LIB and LIU on model performance. We fix n at 4 and set $m = 2, 3, 4$, and fix m at 4 and set $n = 1, 2, 3, 4$. The right figure shows the testing results of these configurations.

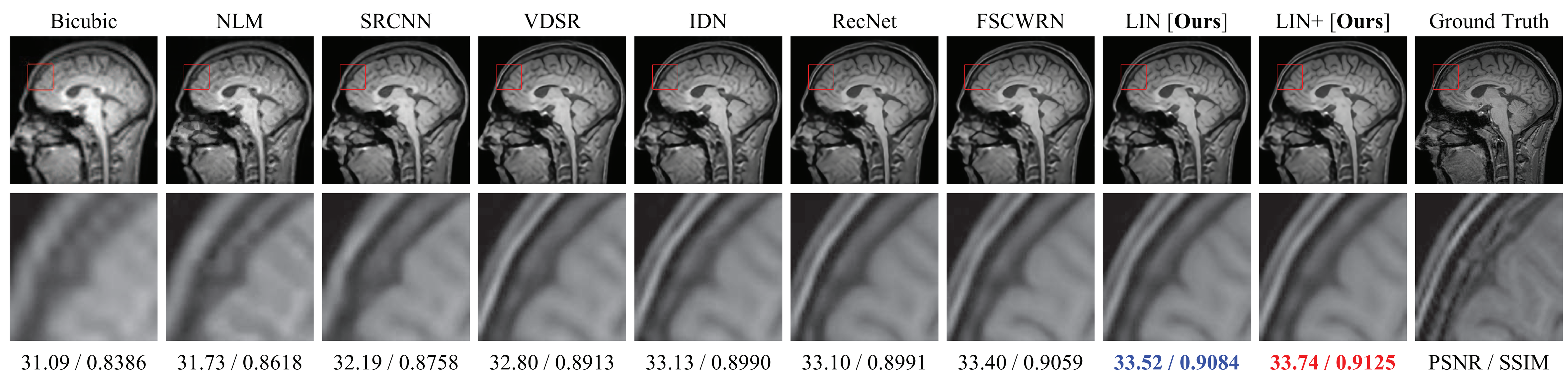


Alation study and model analysis: The valid curves of different configurations of the network for studying the feature extraction and visual inhibition mechanism (left). The testing performance comparison between the models with different numbers of LIB and LIU (right).

Quantitative and Qualitative Evaluation

Quantitative comparison between typical SR models. The maximum values in each comparative cell are marked in red and the second ones are marked in blue (PSNR (dB) / SSIM).

method	SR	param	bicubic downsampling			k-space truncation		
			PD	T1	T2	PD	T1	T2
Bicubic	×2	/	35.04 / 0.9664	33.80 / 0.9525	33.44 / 0.9589	34.65 / 0.9625	33.38 / 0.9460	33.06 / 0.9541
NLM [2]	×2	/	37.26 / 0.9773	35.80 / 0.9685	35.58 / 0.9722	36.18 / 0.9707	34.71 / 0.9581	34.56 / 0.9641
SRCNN [3]	×2	24.5K	38.96 / 0.9836	37.12 / 0.9761	37.32 / 0.9796	38.23 / 0.9802	36.52 / 0.9705	37.04 / 0.9773
VDSR [4]	×2	0.67M	39.97 / 0.9861	37.67 / 0.9783	38.65 / 0.9836	39.89 / 0.9850	37.58 / 0.9760	38.74 / 0.9823
RecNet [5]	×2	1.33M	40.43 / 0.9873	37.86 / 0.9792	39.13 / 0.9848	40.10 / 0.9857	37.54 / 0.9764	39.03 / 0.9832
FSCWRN [6]	×2	3.50M	40.72 / 0.9880	37.98 / 0.9797	39.44 / 0.9855	40.91 / 0.9876	38.04 / 0.9786	39.82 / 0.9851
LIN [Ours]	×2	1.33M	40.84 / 0.9883	38.04 / 0.9798	39.50 / 0.9857	41.11 / 0.9880	38.21 / 0.9793	40.02 / 0.9855
LIN+ [Ours]	×2	1.33M	41.03 / 0.9886	38.19 / 0.9803	39.62 / 0.9860	41.31 / 0.9886	38.40 / 0.9801	40.18 / 0.9859
Bicubic	×3	/	31.20 / 0.9230	30.15 / 0.8900	29.80 / 0.9093	30.88 / 0.9167	29.79 / 0.8793	29.50 / 0.9016
NLM [2]	×3	/	32.81 / 0.9436	31.74 / 0.9216	31.28 / 0.9330	32.02 / 0.9324	30.83 / 0.9027	30.57 / 0.9197
SRCNN [3]	×3	24.5K	33.60 / 0.9516	32.17 / 0.9276	32.20 / 0.9440	32.90 / 0.9432	31.72 / 0.9187	31.80 / 0.9381
VDSR [4]	×3	0.67M	34.66 / 0.9599	32.91 / 0.9378	33.47 / 0.9559	34.27 / 0.9555	32.57 / 0.9304	33.23 / 0.9515
RecNet [5]	×3	1.33M	34.96 / 0.9623	33.05 / 0.9399	33.85 / 0.9588	34.67 / 0.9590	32.80 / 0.9347	33.69 / 0.9554
FSCWRN [6]	×3	3.50M	35.37 / 0.9653	33.24 / 0.9423	34.27 / 0.9618	35.30 / 0.9636	33.09 / 0.9390	34.34 / 0.9603
LIN [Ours]	×3	1.37M	35.39 / 0.9654	33.23 / 0.9421	34.26 / 0.9616	35.39 / 0.9642	33.25 / 0.9406	34.45 / 0.9609
LIN+ [Ours]	×3	1.37M	35.56 / 0.9661	33.44 / 0.9440	34.41 / 0.9627	35.59 / 0.9656	33.50 / 0.9429	34.63 / 0.9622
Bicubic	×4	/	29.13 / 0.8799	28.28 / 0.8312	27.86 / 0.8611	28.82 / 0.8713	27.96 / 0.8182	27.60 / 0.8511
NLM [2]	×4	/	30.27 / 0.9044	29.31 / 0.8655	28.85 / 0.8875	29.27 / 0.8906	28.68 / 0.8439	28.37 / 0.8718
SRCNN [3]	×4	24.5K	31.10 / 0.9181	29.90 / 0.8796	29.69 / 0.9052	30.52 / 0.9078	29.31 / 0.8616	29.32 / 0.8960
VDSR [4]	×4	0.67M	32.09 / 0.9311	30.57 / 0.8932	30.79 / 0.9240	31.69 / 0.9244	30.14 / 0.8818	30.51 / 0.9162
RecNet [5]	×4	1.33M	32.58 / 0.9378	30.86 / 0.9005	31.30 / 0.9310	32.16 / 0.9310	30.46 / 0.8900	31.03 / 0.9243
FSCWRN [6]	×4	3.50M	32.91 / 0.9415	30.96 / 0.9022	31.71 / 0.9359	32.78 / 0.9387	30.79 / 0.8973	31.71 / 0.9334
LIN [Ours]	×4	1.36M	32.94 / 0.9417	31.01 / 0.9033	31.72 / 0.9361	32.82 / 0.9391	30.88 / 0.8990	31.77 / 0.9339
LIN+ [Ours]	×4	1.36M	33.12 / 0.9432	31.28 / 0.9073	31.88 / 0.9376	33.03 / 0.9415	31.20 / 0.9041	31.96 / 0.9362



As can be seen from both quantitative and visual comparisons above, the proposed LIN model can achieve better SR performance with fewer model parameters and computational overhead.