

Lightweight image super-resolution via overlapping back-projection feedback network for embedded devices

Beibei Wang^a, Changjun Liu^a, Binyu Yan^{a,*}, Seunggil Jeon^b, Xiaomin Yang^a, Zhuoyue Zhang^a

^a College of Electronic and Information Engineering, Sichuan University, Chengdu, China

^b Samsung Electronics 129, Samseong-ro Yeongtong-gu Suwon-si, Gyeonggi-do 16677, South Korea

ARTICLE INFO

MSC:
00-01
99-00

Keywords:
Embedded devices
Image super-resolution
Lightweight
Feedback
Back-projection

ABSTRACT

Super-resolution (SR) technology is widely used in embedded devices because it can improve image quality. However, to achieve improved performance, SR networks usually take a massive memory because of their large number of parameters. They are not applicable for embedded devices with low power consumption. In this work, we propose an overlapping back-projection feedback network (LOBFN) for image SR, which is a lightweight network designed for embedded devices. First, a back-projection feedback block (PFB) and recursive concatenation are used to learn the hierarchical representations of the network. Second, an overlapping back-projection suitable for lightweight network is proposed to minimize the reconstruction errors. Finally, a fusion attention module (FAM) is proposed to perceive information-rich features. The final experiments proved that the proposed LOBFN significantly improved the SR performance of lightweight networks.

1. Introduction

Nowadays, the design of embedded devices is becoming increasingly important, for embedded devices are increasingly portable and miniaturized [1,2]. Low power consumption design of embedded systems is an issue that designers have to face, for embedded devices do not always have sufficient power supply [3,4]. The capabilities of embedded devices are enhanced with the help of artificial intelligent (AI) and big data. SR networks based on deep learning has attracted academic attention recently, and it is the best SR method at present. The single image super-resolution (SISR) aims at learning the hierarchical representations of low-resolution (LR) image data to restore high-resolution (HR) image, which is the focus of academic research at present. This paper focuses on the SISR network suitable for embedded devices with low power consumption.

The SISR has made great progress since the introduction of deep learning [5], and then many super-resolution methods have been proposed [6–15]. Thereafter, FSRCNN [7] was proposed to improve SRCNN [5], which up-sampled LR features using deconvolution on the last layer to reduce calculation amount. ESPCN [16] introduced sub-pixel convolution to up-sample LR features and reduce redundancy. Lai et al. progressively reconstructed HR images in LapSRN [12].

Many methods [17,18] have shown that deeper network is more expressive, but the training of deep networks is unstable. VDSR [19] is the

first deep SR network, which introduced residual learning to train deep SR network. Thereafter, ResNet [20] connected the layers by residual learning to enhance the gradient flow and reduce the degradation of the network. Then many methods [11,13,21] based on residual learning were proposed. However, deeper networks produce more parameters, which lead to more memory usage. Subsequently, recursive convolutional networks were proposed, such as the DRCN [22] and DRRN [23]. Then recursive fusion technique was used in the EBRN [24], which extracted features of different frequencies through different network depths and integrated them recursively. The recursive convolutional layers of these networks shared the same weights, so the networks can be deepened without increasing the number of parameters. Inspired by RNN, SRFBN [8] is proposed as a feedback network. DBPN [10] used error feedback to correct features in back-projection units.

The channel attention mechanism allows the network to perceive information-rich channels with very few parameters. It was initially proposed for image classification tasks, such as the SE module proposed by Hu et al. [25]. Then, the residual channel attention blocks (RCABs) for image SR was proposed in RCAN [6], which rescaled the features adaptively by considering the interdependent relationships between channels. Woo et al. proposed the channel and spatial attention using

* Corresponding author.

E-mail addresses: 690983790@qq.com (B. Wang), cjliu@scu.edu.cn (C. Liu), Yanby@scu.edu.cn (B. Yan), simon.sgJeon@gmail.com (S. Jeon), arielyang@scu.edu.cn (X. Yang), zhangzhuoyue@scu.edu.cn (Z. Zhang).

<https://doi.org/10.1016/j.micpro.2023.104777>

Received 5 September 2021; Received in revised form 12 December 2022; Accepted 23 January 2023

Available online 27 January 2023

0141-9331/© 2023 Elsevier B.V. All rights reserved.

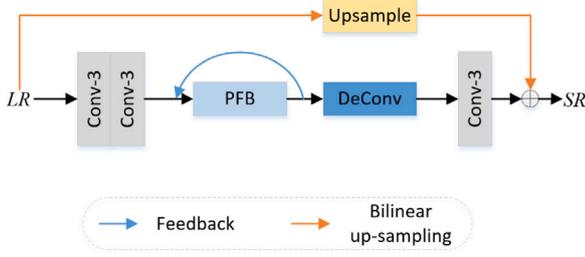
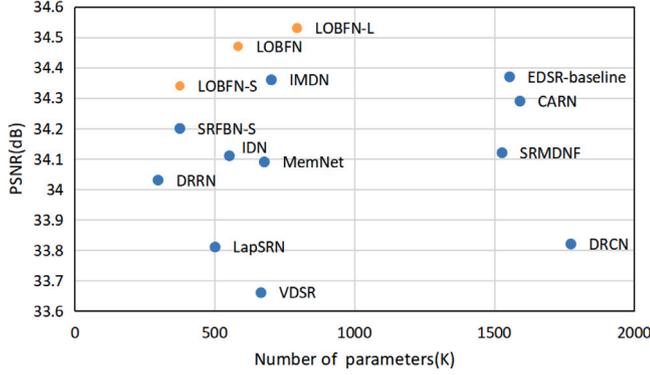


Fig. 1. The proposed LOBFN.

Fig. 2. PSNR (dB) vs. number of parameters ($K = 10^3$) with scaling factor of x3 on Set5 dataset.

both maximum and average-pooling in CBAM [26]. Then contrast-aware channel attention (CCA) module using both standard deviation and average-pooling was put forward in IMDN [27].

Since embedded devices are usually resource-constrained, lightweight applications with less parameters is getting increasingly important. Therefore, a lightweight feedback network via overlapping back-projection (LOBFN) is proposed, as shown in Fig. 1, which is inspired by SRFBN-S [8]. Our LOBFN is applicable for embedded devices with low power consumption, because the feedback mechanism and the overlapping back-projection units can reduce the required parameters and calculations. The PFB blocks of all iterations share the same weights as a feedback block, and the overlapping back-projection units reuse the operations and parameters from the previous projection unit. The proposed LOBFN has an outstanding performance compared with other classical methods, as shown in Fig. 2.

Our contributions are as follows:

- To efficiently learn the hierarchical representations of the network, we use a feedback block to enhance the low-level representations, and use recursive concatenation to enhance the high-level representations. The feedback block can reduce the required parameters by sharing the same weights.
- To minimize the reconstruction errors efficiently, an overlapping back-projection suitable for lightweight networks is proposed. The projection errors are learned and fed back to correct intermediate features.
- To perceive information-rich features, a fusion attention module (FAM) is proposed. We argue that standard deviation, average- and max-pooling are all beneficial to learn information-rich features, so we fuse them to enhance feature expression ability of the network.

2. Method or methodology

In this section, the unfolded LOBFN is introduced in first. Then, the proposed overlapping back-projection groups (OBGs) and fusion attention module (FAM) in PFB are analyzed in detail.

2.1. Architecture of LOBFN

As shown in Fig. 3, we reconstruct SR images at the same scaling factor by T iterations (iteration $t \in (1, \dots, T)$), and take SR^T as the SR result of LOBFN. The LOBFN can be divided into three parts: LR feature extraction part, feedback part and reconstruction part.

We define F_{in} as the output of the first part, and it is one of the inputs of PFB. In the 1st iteration, PFB takes F_{in} as input. While in the other iterations, the input of PFB is the concatenation of F_{in} and the output of itself from last iteration F_{out}^{t-1} . Therefore, the output of PFB at the t^{th} iteration can be obtained by:

$$F_{out}^t = \begin{cases} f_{PFB}(F_{in}) & t = 1 \\ f_{PFB}([F_{in}, F_{out}^{t-1}]) & t \geq 2 \end{cases}, \quad (1)$$

where f_{PFB} denotes the PFB operations. $[]$ denotes the concatenation operation.

Reconstruction part contains a deconvolutional up-sampling layer and a convolutional layer for channel compression. In the t^{th} iteration, as the output of deconvolutional layer, F_{rb}^t can be obtained by:

$$F_{rb}^t = f_{UP}(F_{out}^t), \quad (2)$$

where f_{UP} denotes the operation of the deconvolutional layer.

Finally, as the SR result of the t^{th} iteration, SR^t can be obtained by:

$$SR^t = \begin{cases} f_c(F_{rb}^t) + f_{BI}(LR) & t = 1 \\ f_c([F_{rb}^t, F_{rb}^{t-1}, \dots, F_{rb}^1]) + f_{BI}(LR) & t \geq 2 \end{cases}, \quad (3)$$

where f_c denotes the convolution operation to compress feature channels for image reconstruction. f_{BI} denotes the operation of bilinear up-sampling operation.

2.2. Back-projection feedback block (PFB)

PFB is a feedback block, which feeds the output back to itself. As shown in Fig. 3, the PFB contains two parts: overlapping back-projection groups (OBGs) and fusion attention module (FAM). OBGs correct the intermediate features using error feedback, and FAM distributes different attention weights to features. The iteration of OBGs and FAM efficiently improved the expression ability of features.

2.2.1. Overlapping back-projection groups (OBGs)

Inspired by D-DBPN [10], we propose an overlapping back-projection. For better understanding, the proposed overlapping back-projection units are compared with the back-projection stages proposed in D-DBPN [10], as shown in Fig. 4. In D-DBPN [10], the two back-projection units are independent, but in our method, the two back-projection units overlap each other. Every projection unit except the first one, reuses the up-/down-sampling operation and parameters from the previous projection unit. In this ways, we reduced the calculation amount and parameters required for the back-projection units, making it suitable for lightweight networks.

Based on the proposed overlapping back-projection units, the structure of OBGs is described in detail in Fig. 5. OBGs contains G up-projection units, and the group g is from 1 to G . We define f_{cin} as the convolution operation for compressing the channels of input features. The input features L_0 can be obtained by:

$$L_0^t = \begin{cases} f_{cin}(F_{in}) & t = 1 \\ f_{cin}([F_{in}, F_{out}^{t-1}]) & t \geq 2 \end{cases}. \quad (4)$$

We define f_{UP} and f_{DOWN} as the up-sampling and down-sampling operations in overlapping back-projection groups. In the t^{th} iteration, we define H_g^t and L_g^t as the intermediate HR and LR features of the g -th up-projection unit, NH_g^t and NL_g^t as the error-corrected H_g^t and L_g^t . They can be obtained by:

$$H_g^t = \begin{cases} f_{UP}([L_0^t, \dots, L_{g-1}^t]) & g \leq 2 \\ f_{UP}([L_0^t, NL_1^t, \dots, NL_{g-2}^t, L_{g-1}^t]) & g \geq 3 \end{cases}, \quad (5)$$

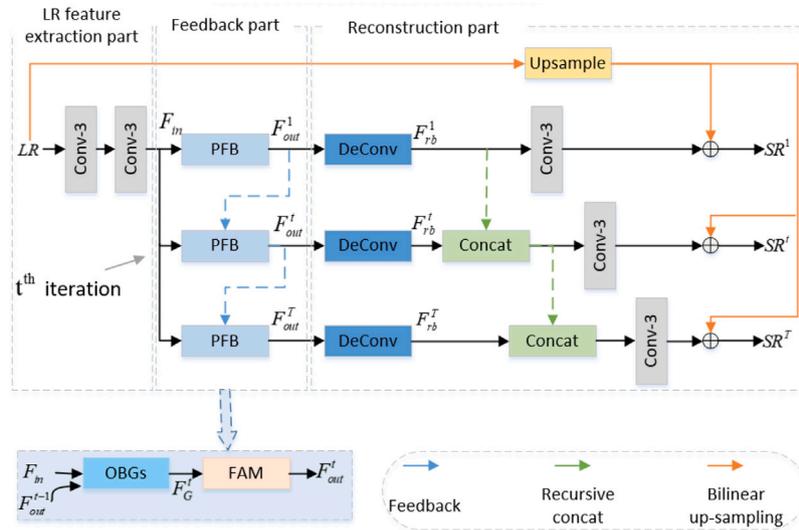


Fig. 3. The unfolded LOBFN.

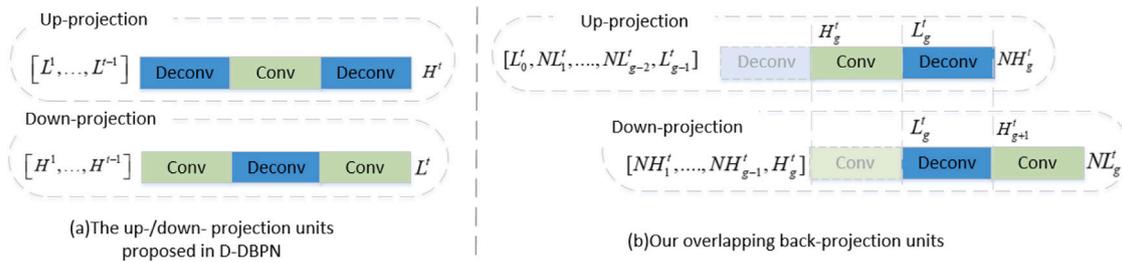


Fig. 4. The comparison of the back-projection proposed in D-DBPN [10] and our overlapping back-projection. The light colored up-/down-sampling blocks represent the reuse of operations and parameters from the previous back-projection unit. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

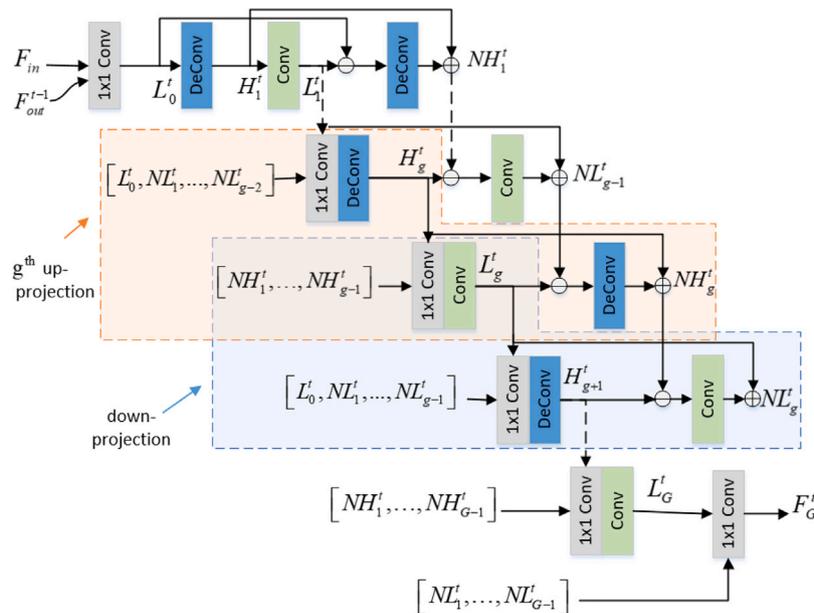


Fig. 5. The unfolded OBGs.

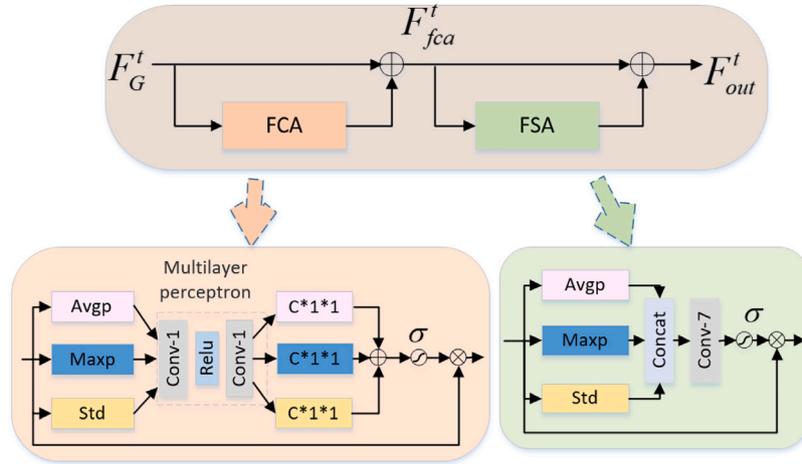


Fig. 6. Fusion attention module (FAM).

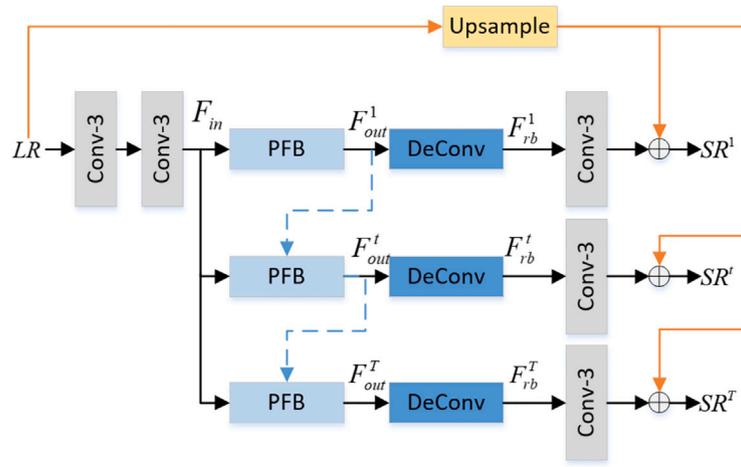


Fig. 7. The network architecture after the ablation of recursive concatenation.

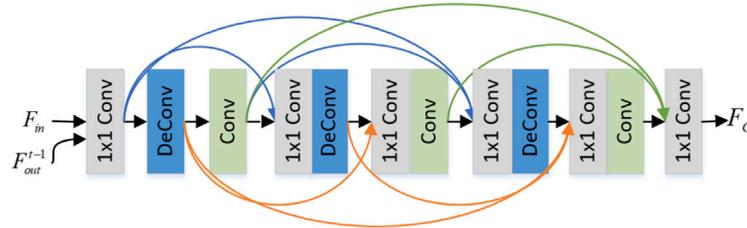


Fig. 8. Projection groups in SRFBN [8].

$$L_g^t = \begin{cases} f_{Down}(H_g^t) & g = 1 \\ f_{Down}([NH_1^t, \dots, NH_{g-1}^t, H_g^t]) & g \geq 2 \end{cases}, \quad (6)$$

$$NH_g^t = \begin{cases} H_g^t + f_{UP}(L_g^t - L_{g-1}^t) & g = 1 \\ H_g^t + f_{UP}(L_g^t - NL_{g-1}^t) & g \geq 2 \end{cases}, \quad (7)$$

$$NL_g^t = L_g^t + f_{DOWN}(H_{g+1}^t - NH_g^t) \quad g \geq 1. \quad (8)$$

Finally, F_G^t , as the output OBGs, can be obtained by:

$$F_G^t = f_{cout}([NL_1^t, \dots, NL_{G-1}^t, L_G^t]), \quad (9)$$

where f_{cout} is the convolutional operation for compressing the output feature channels.

2.2.2. Fusion attention module (FAM)

The network with attention module can concentrate on information-rich features. In previous networks [6,25,26], average/max-pooling were widely used in attention module, for they can obtain information about distinctive objects to enhance image details. However, we argue that the standard deviation can obtain the information of textures, structures, and edges, so the features with large standard deviation are information-rich and should receive more attention. The standard deviation, average and maximum pooling are used to calculate channel and spatial attention, so our FAM contains two components: fusion channel attention (FCA) and fusion spatial attention (FSA). Furtherly, to enhance the gradient flow, we use residual learning in FAM, as shown in Fig. 6.

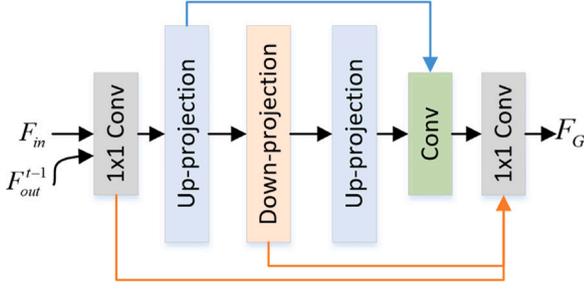


Fig. 9. Independent back-projection groups in D-DBPN [10].

We use FCA and FSA to improve the performance of PFB, and define the features after FCA as F_{fca}^t , which can be obtained by:

$$F_{fca}^t = F_G^t * \sigma(f_{mlp}(f_{std}^c(F_G^t)) + f_{mlp}(f_{avgp}^c(F_G^t)) + f_{mlp}(f_{maxp}^c(F_G^t))) + F_G^t, \quad (10)$$

where f_{std}^c , f_{avgp}^c and f_{maxp}^c are the standard deviation, average- and max-pooling operations of each channel. f_{mlp} is the operation of conv-relu-conv. σ is the sigmoid function.

Finally, in the t^h iteration, the output features of PFB can be obtained by:

$$F_{out}^t = F_{fca}^t * \sigma(f^{7 \times 7}(f_{std}^s(F_{fca}^t) + f_{avgp}^s(F_{fca}^t) + f_{maxp}^s(F_{fca}^t))) + F_{fca}^t, \quad (11)$$

where f_{std}^s , f_{avgp}^s and f_{maxp}^s are the operations of pixels in the same spatial location along the channel dimension. $f^{7 \times 7}$ is a conv-7 operation.

2.3. Loss function

We choose the L1 loss function and use the average loss values of T iterations to supervise the training of the LOBFN, so our loss function can be obtained by:

$$Loss = \frac{1}{T} \sum_{t=1}^T L_1(SR^t, HR). \quad (12)$$

3. Experimental results

3.1. Experimental details

We use DIV2K [28] dataset as the training set, and expand the number to 8000 by rotating and cropping the images. The LR images are generated by the bicubic down-sampling. We use the Adam optimizer and PyTorch framework. We set learning rate $lr = 0.0005$, and halved it every 200 epochs. The SR result of the last iteration is tested on the Set5, Set14, BSD100, Urban100 and Manga109 datasets. In this paper, we implement three models: LOBFN-S ($T = 4, G = 2$), LOBFN ($T = 4, G = 3$) and LOBFN-L ($T = 4, G = 4$).

3.2. Comparison with the baseline

The LOBFN we proposed is based on SRFBN-S [8]. To prove the improvements of our method, we compare our LOBFN-S with SRFBN-S [8], and they have almost the same number of parameters. The comparison results of them are shown in Table 1. LOBFN-S has a better performance than SRFBN-S [8], which proved our improvement. This is because we improved the feedback block using overlapping back-projection and fusion attention module, and we used recursive concatenation for SR images reconstruction.

3.3. Ablation study of the recursive concatenation

We use recursive concatenation for SR image reconstruction. It can be degenerated to the multi-reconstruction method used in SRFBN-S [8], if we remove the recursive concatenation, as shown in Fig. 7. To validate the efficiency of our recursive concatenation, we make a comparison as shown in Table 2, which demonstrate that the recursive concatenation operation is beneficial for improving the performance of the network. This is because recursive concatenation enhanced high-level representations of the HR features for SR reconstruction.

3.4. The effectiveness of overlapping back-projection

The overlapping back-projection groups (OBGs) can degenerate to the projection groups used in SRFBN-S [8], if we remove the error feedback, as shown in Fig. 8. We set $G = 5$ for the projection groups, which has almost the same number of parameters with LOBFN. Further, we compare OBGs with the independent back-projection groups proposed in D-DBPN [10], as shown in Fig. 9. We compare them with our overlapping back-projection, as shown in Table 3. The results demonstrate that the overlapping back-projection groups are beneficial for improving the performance of the network. This is because the error feedback can better guide the SR reconstruction, and the overlapping between them can learn the intermediate features more fully.

3.5. The effectiveness of FAM

We propose a fusion attention module (FAM) used after the projection groups in PFB, which are not used in SRFBN [8]. First, we remove it to validate the efficiency of FAM. Second, we change the order of FCA and FSA in FAM to obtain a better combining strategy. Finally, we compare our FAM with CBAM [26], which used max- and average-pooling to calculate channel and spacial attention, as shown in Fig. 10.

In Table 4, the comparison results demonstrate that the FAM module is beneficial for improving the performance of the network. This is because attention module enables the network to perceive information-rich features. Then, FCA + FSA has a better performance, so we use FCA + FSA in FAM. At last, FAM is better than CBAM [26], because the standard deviation of pixels in FAM helps to recover image details related to structures and textures.

3.6. Comparison with the classical lightweight SR methods

In this paper, we implement three models: LOBFN-S, LOBFN and LOBFN-L, among which LOBFN-L performs best. In Table 5, our methods are compared with the classical lightweight SR methods: the SRCNN [5], FSRCNN [7], VDSR [19], LapSRN [12], DRRN [23], MemNet [29], IDN [30], EDSR-baseline [13], SRMDNF [31], CARN [32] and IMDN [27]. The PSNR value of our LOBFN is not better than that of IMDN [27] at scaling factors of $\times 2$, but the parameters of LOBFN are less than that of IMDN [27]. For $3 \times SR$, the parameters of LOBFN are less than that of IMDN [27], but the PSNR values of our LOBFN are better than that of IMDN [27]. Therefore, our methods achieve outstanding performance, especially on large scaling factors.

The comparisons of the SR images with scaling factors of $\times 4$ are visually shown in Fig. 11, which demonstrate that our methods can better recover the detailed features. This proves the efficiency of the proposed LOBFN.

Table 1
Comparison of SRFBN-S [8] and LOBFN.

Methods	Scale	Params	Set5 PSNR/SSIM	Set14 PSNR/SSIM	BSD100 PSNR/SSIM	Urban100 PSNR/SSIM	Manga109 PSNR/SSIM
SRFBN-S [8]	×2	282k	37.78/0.9597	33.35/0.9156	32.00/0.8970	31.41/0.9207	38.06/0.9757
LOBFN-S		283k	37.86/0.9601	33.46/0.9165	32.06/0.8983	31.71/0.9241	38.20/0.9762
SRFBN-S [8]	×3	376k	34.20/0.9255	30.10/0.8372	28.96/0.8010	27.66/0.8415	33.02/0.9404
LOBFN-S		376k	34.34/0.9266	30.29/0.8416	29.04/0.8037	27.98/0.8487	33.45/0.9436
SRFBN-S [8]	×4	483k	31.98/0.8923	28.45/0.7779	27.44/0.7313	25.71/0.7719	29.91/0.9008
LOBFN-S		484k	32.11/0.8942	28.54/0.7814	27.52/0.7349	25.94/0.7807	30.40/0.9075

Table 2
Ablation study of recursive concatenation.

Recursive concatenation	Scale	Params	Set5 PSNR/SSIM	Set14 PSNR/SSIM	BSD100 PSNR/SSIM	Urban100 PSNR/SSIM	Manga109 PSNR/SSIM
×	×4	746k	32.20/0.8954	28.58/0.7820	27.57/0.7360	26.10/0.7864	30.62/0.9095
✓		753k	32.23/0.8955	28.62/0.7829	27.57/0.7362	26.11/0.7865	30.66/0.9103

Table 3
Comparison of projection groups, the independent back-projection groups and our overlapping back-projection groups.

Projection manner	Scale	Params	Set5 PSNR/SSIM	Set14 PSNR/SSIM	BSD100 PSNR/SSIM	Urban100 PSNR/SSIM	Manga109 PSNR/SSIM
Projection		605k	34.45/0.9274	30.39/0.8433	29.10/0.8053	28.23/0.8542	33.67/0.9451
Independent back-projection	×3	573k	34.40/0.9270	30.35/0.8426	29.07/0.8045	28.21/0.8516	33.51/0.9442
Overlapping back-projection		585k	34.47/0.9276	30.40/0.8433	29.10/0.8054	28.21/0.8537	33.70/0.9452

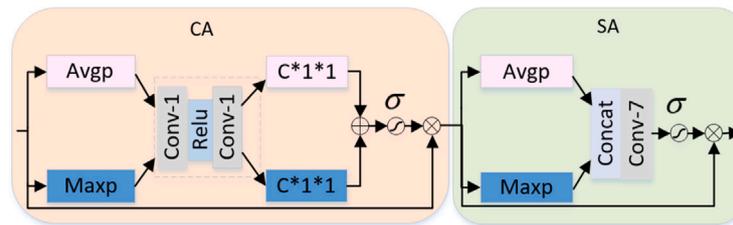


Fig. 10. The attention module of CBAM [26].

Table 4
The effectiveness the FAM module in OBGs.

Attention module	Scale	Params	Set5 PSNR/SSIM	Set14 PSNR/SSIM	BSD100 PSNR/SSIM	Urban100 PSNR/SSIM	Manga109 PSNR/SSIM
None		584k	34.42/0.9271	30.39/0.8429	29.09/0.8047	28.20/0.8531	33.64/0.9447
FCA + FSA	×3	585k	34.47/0.9276	30.40/0.8433	29.10/0.8054	28.21/0.8537	33.70/0.9452
FSA + FCA		585k	34.44/0.9272	30.39/0.8430	29.09/0.8049	28.23/0.8539	33.59/0.9448
CBAM [26]		584k	34.45/0.9274	30.37/0.8427	29.09/0.8049	28.19/0.8533	33.56/0.9446

4. Conclusion

In this paper, we propose an overlapping back-projection feedback network to achieve lightweight image super-resolution (LOBFN). We use a feedback block PFB and recursive concatenation to learn the hierarchical features of the network. In PFB, we propose an overlapping back-projection mechanism to minimize the reconstruction errors. Finally, we propose a fusion attention module (FAM) to perceive information-rich features. The final experiments demonstrate the superiorities of the proposed LOBFN. All the motivations we proposed are efficient and require few parameters, so it is suitable for embedded devices with low power consumption.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

Funding

This research was funded by Sichuan University and Yibin Municipal People's Government University and City strategic cooperation special fund project (Grant No. 2020CDYB-29); Science and Technology plan transfer payment project of Sichuan province (2021ZYF007); The Key Research and Development Program of Science and Technology Department of Sichuan Province (No. 2020YFS0575, No. 2021KJT0012-2021YFS0067); The funding from Science Foundation of Sichuan Science and Technology Department (2021YFH0119); The funding from Sichuan University (Grant No. 2020SCUNG205) and National Natural Science Foundation of China under Grant No. 62201370.

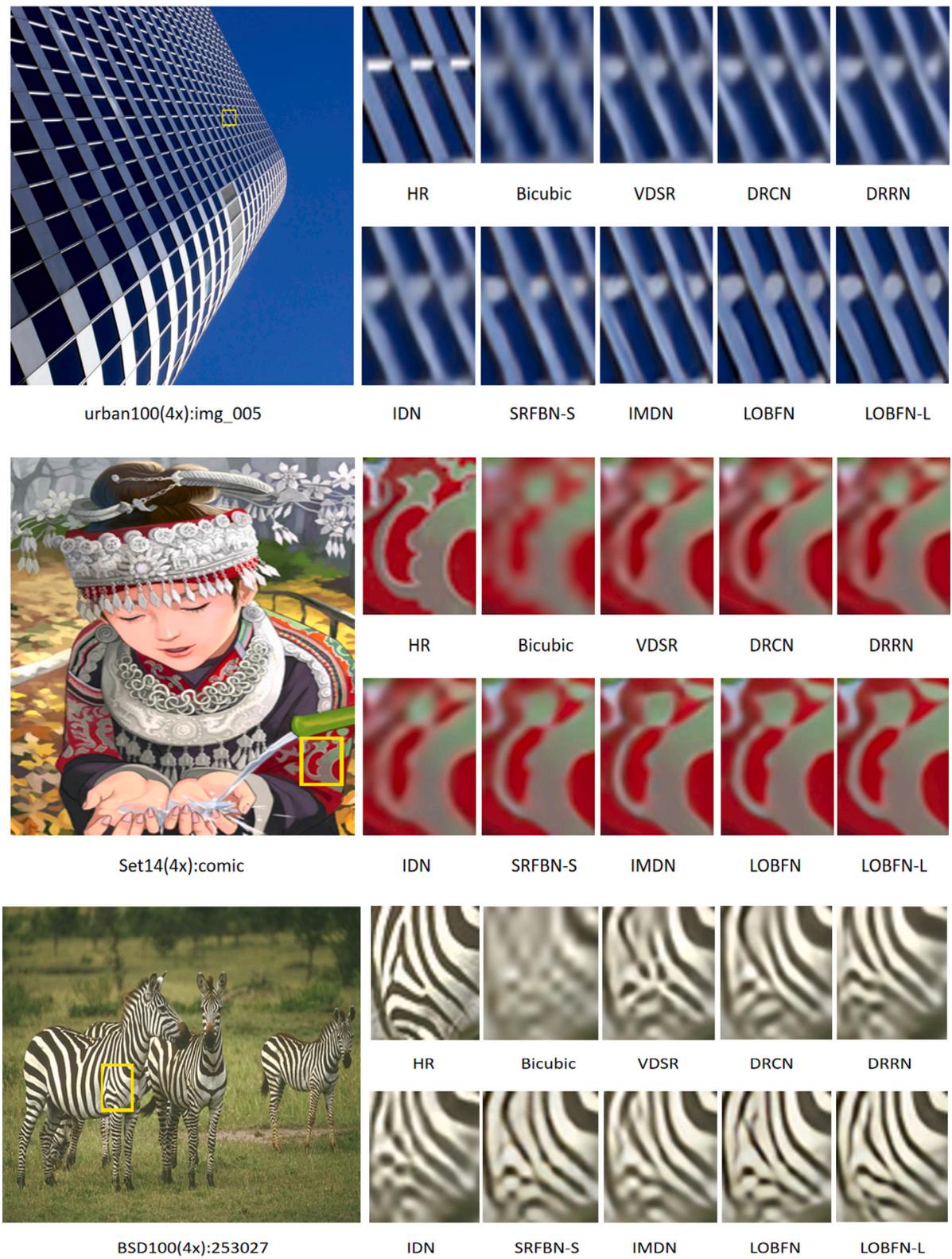


Fig. 11. Visual comparisons of LOBFN with other SR methods on Set14, Urban100 and BSD100 datasets.

Table 5

Comparison of the average PSNRs/SSIMs for different scale factors on the five benchmark datasets. Red represents the best results and blue represents the second-best results.

Methods	Scale	Params	Set5 PSNR/SSIM	Set14 PSNR/SSIM	BSD100 PSNR/SSIM	Urban100 PSNR/SSIM	Manga109 PSNR/SSIM
Bicubic	–	–	33.66/0.9299	30.24/0.8688	29.56/0.8431	26.88/0.8403	30.80/0.9339
SRCNN [5]	8k	8k	36.66/0.9542	32.45/0.9067	31.36/0.8879	29.50/0.8946	35.60/0.9663
FSRCN [7]	13k	13k	37.00/0.9558	32.63/0.9088	31.53/0.8920	29.88/0.9020	36.67/0.9710
VDSR [19]	666k	666k	37.53/0.9587	33.03/0.9124	31.90/0.8960	30.76/0.9140	37.22/0.9750
LapSRN [12]	251k	251k	37.52/0.9591	32.99/0.9124	31.80/0.8952	30.41/0.9103	37.27/0.9740
DRRN [23]	298k	298k	37.74/0.9591	33.23/0.9136	32.05/0.8973	31.23/0.9188	37.88/0.9749
MemNet [29]	678k	678k	37.78/0.9597	33.28/0.9142	32.08/0.8978	31.31/0.9195	37.72/0.9740
IDN [30]	553k	553k	37.83/0.9600	33.30/0.9148	32.08/0.8985	31.27/0.9196	38.01/0.9749
EDSR-baseline [13]	1370k	1370k	37.99/0.9604	33.57/0.9175	32.16/0.8994	31.98/0.9272	38.54/0.9769
SRMDNF [31]	1511k	1511k	37.79/0.9601	33.32/0.9159	32.05/0.8985	31.33/0.9204	38.07/0.9761
CARN [32]	1592k	1592k	37.76/0.9590	33.52/0.9166	32.09/0.8978	31.92/0.9256	38.36/0.9765
IMDN [27]	694k	694k	38.00/0.9605	33.63/0.9177	32.19/0.8996	32.17/0.9283	38.88/0.9774
LOBFN-S (Ours)	283k	283k	37.86/0.9601	33.46/0.9165	32.06/0.8983	31.71/0.9241	38.20/0.9762
LOBFN (Ours)	438k	438k	37.94/0.9602	33.57/0.9172	32.12/0.8988	31.95/0.9265	38.36/0.9765
LOBFN-L (Ours)	595k	595k	37.99/0.9604	33.64/0.9180	32.15/0.8990	32.10/0.9283	38.44/0.9765
Bicubic	–	–	30.39/0.8682	27.55/0.7742	27.21/0.7385	24.46/0.7349	26.95/0.8556
SRCNN [5]	8k	8k	32.75/0.9090	29.30/0.8215	28.41/0.7863	26.24/0.7989	30.48/0.9117
FSRCN [7]	13k	13k	33.18/0.9140	29.37/0.8240	28.53/0.7910	26.43/0.8080	31.10/0.9210
VDSR [19]	666k	666k	33.66/0.9213	29.77/0.8314	28.82/0.7976	27.14/0.8279	32.01/0.9340
LapSRN [12]	502k	502k	33.81/0.9220	29.79/0.8325	28.82/0.7980	27.07/0.8275	32.21/0.9350
DRRN [23]	298k	298k	34.03/0.9244	29.96/0.8349	28.95/0.8004	27.53/0.8378	32.71/0.9379
MemNet [29]	678k	678k	34.09/0.9248	30.00/0.8350	28.96/0.8001	27.56/0.8376	32.51/0.9369
IDN [30]	553k	553k	34.11/0.9253	29.99/0.8354	28.95/0.8013	27.42/0.8359	32.71/0.9381
EDSR-baseline [13]	1555k	1555k	34.37/0.9270	30.28/0.8417	29.09/0.8052	28.15/0.8527	33.45/0.9439
SRMDNF [31]	1528k	1528k	34.12/0.9254	30.04/0.8382	28.97/0.8025	27.57/0.8398	33.00/0.9403
CARN [32]	1592k	1592k	34.29/0.9255	30.29/0.8407	29.06/0.8034	28.06/0.8493	33.50/0.9440
IMDN [27]	703k	703k	34.36/0.9270	30.32/0.8417	29.09/0.8046	28.17/0.8519	33.61/0.9445
LOBFN-S (Ours)	376k	376k	34.34/0.9266	30.29/0.8416	29.04/0.8037	27.98/0.8487	33.45/0.9436
LOBFN (Ours)	585k	585k	34.47/0.9276	30.40/0.8433	29.10/0.8054	28.21/0.8537	33.70/0.9452
LOBFN-L (Ours)	795k	795k	34.53/0.9281	30.44/0.8442	29.12/0.8059	28.33/0.8562	33.80/0.9459
Bicubic	–	–	28.42/0.8104	26.00/0.7027	25.96/0.6675	23.14/0.6577	24.89/0.7866
SRCNN [5]	8k	8k	30.48/0.8628	27.50/0.7513	26.90/0.7101	24.52/0.7221	27.58/0.8555
FSRCN [7]	13k	13k	30.72/0.8660	27.61/0.7550	26.98/0.7150	24.62/0.7280	27.90/0.8610
VDSR [19]	666k	666k	31.35/0.8838	28.01/0.7674	27.29/0.7251	25.18/0.7524	28.83/0.8870
LapSRN [12]	502k	502k	31.54/0.8852	28.09/0.7700	27.32/0.7275	25.21/0.7562	29.09/0.8900
DRRN [23]	298k	298k	31.68/0.8888	28.21/0.7720	27.38/0.7284	25.44/0.7638	29.45/0.8946
MemNet [29]	678k	678k	31.74/0.8893	28.26/0.7723	27.40/0.7281	25.50/0.7630	29.42/0.8942
IDN [30]	553k	553k	31.82/0.8903	28.25/0.7730	27.41/0.7297	25.41/0.7632	29.41/0.8942
EDSR-baseline [13]	1518k	1518k	32.09/0.8938	28.58/0.7813	27.57/0.7357	26.04/0.7849	30.35/0.9067
SRMDNF [31]	1552k	1552k	31.96/0.8925	28.35/0.7787	27.49/0.7337	25.68/0.7731	30.09/0.9024
CARN [32]	1592k	1592k	32.13/0.8937	28.60/0.7806	27.58/0.7349	26.07/0.7837	30.47/0.9084
IMDN [27]	715k	715k	32.21/0.8948	28.58/0.7811	27.56/0.7353	26.04/0.7838	30.45/0.9075
LOBFN-S (Ours)	484k	484k	32.11/0.8942	28.54/0.7814	27.52/0.7349	25.94/0.7807	30.40/0.9075
LOBFN (Ours)	753k	753k	32.23/0.8955	28.62/0.7829	27.57/0.7362	26.11/0.7865	30.66/0.9103
LOBFN-L (Ours)	1025k	1025k	32.27/0.8960	28.69/0.7840	27.62/0.7375	26.29/0.7916	30.77/0.9116

References

- [1] H. Xiao, J. Qin, S. Jeon, B. Yan, X. Yang, MFEN: Lightweight multi-scale feature extraction super-resolution network in embedded system, *Microprocess. Microsyst.* 93 (2022) 104568, <http://dx.doi.org/10.1016/j.micpro.2022.104568>.
- [2] L. Chen, B. Fan, Y. Cai, Q. Shi, Application of iot medical image detection and prenatal genetic testing in obstetric clinic, *Microprocess. Microsyst.* 81 (2021) 103705, <http://dx.doi.org/10.1016/j.micpro.2020.103705>.
- [3] P. Ambalathankandy, S. Takamaeda, M. Masato, T. Asai, M. Ikebe, H. Kusano, Real-time HDTV to 4K and 8K-UHD conversions using anti-aliasing based super resolution algorithm on FPGA, *Microprocess. Microsyst.* 61 (2018) 21–31, <http://dx.doi.org/10.1016/j.micpro.2018.05.008>.
- [4] P. Cao, S. Zhang, Research on image recognition of wushu action based on remote sensing image and embedded system, *Microprocess. Microsyst.* 82 (2021) 103841, <http://dx.doi.org/10.1016/j.micpro.2021.103841>.
- [5] C. Dong, C.C. Loy, K. He, X. Tang, Learning a deep convolutional network for image super-resolution, in: *Computer Vision – ECCV 2014*, Vol. 8692, 2014, pp. 184–199, http://dx.doi.org/10.1007/978-3-319-10593-2_13.
- [6] Y. Zhang, k. Li, K. Li, L. Wang, B. Zhong, Y. Fu, Image super-resolution using very deep residual channel attention networks, in: *European Conference on Computer Vision*, Vol. 11211, 2018, pp. 294–310, http://dx.doi.org/10.1007/978-3-030-01234-2_18.
- [7] C.C.L. Chao Dong, X. Tang, Accelerating the super-resolution convolutional neural network, in: *Computer Vision – ECCV 2016*, 2016, pp. 391–407, http://dx.doi.org/10.1007/978-3-319-46475-6_25.
- [8] Z. Li, J. Yang, Z. Liu, X. Yang, G. Jeon, W. Wu, Feedback network for image super-resolution, in: *The IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, 2019.
- [9] Z. Hui, X. Wang, X. Gao, Fast and accurate single image super-resolution via information distillation network, in: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 723–731, <http://dx.doi.org/10.1109/CVPR.2018.00082>.
- [10] M. Haris, G. Shakhnarovich, N. Ukita, Deep back-projection networks for super-resolution, in: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 1664–1673, <http://dx.doi.org/10.1109/CVPR.2018.00179>.
- [11] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, W. Shi, Photo-realistic single image super-resolution using a generative adversarial network, in: *2017 IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, 2017, pp. 105–114, <http://dx.doi.org/10.1109/CVPR.2017.19>.
- [12] W.-S. Lai, J.-B. Huang, N. Ahuja, M.-H. Yang, Deep Laplacian pyramid networks for fast and accurate super-resolution, in: *IEEE Conference on Computer Vision and Pattern Recognition*, 2017, <http://dx.doi.org/10.1109/CVPR.2017.618>.
- [13] B. Lim, S. Son, H. Kim, S. Nah, K.M. Lee, Enhanced deep residual networks for single image super-resolution, in: *2017 IEEE Conference on Computer Vision*

- and Pattern Recognition Workshops, CVPRW, 2017, pp. 1132–1140, <http://dx.doi.org/10.1109/CVPRW.2017.151>.
- [14] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, Y. Fu, Residual dense network for image super-resolution, in: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 2472–2481, <http://dx.doi.org/10.1109/CVPR.2018.00262>.
- [15] T. Tong, G. Li, X. Liu, Q. Gao, Image super-resolution using dense skip connections, in: 2017 IEEE International Conference on Computer Vision, ICCV, 2017, pp. 4809–4817, <http://dx.doi.org/10.1109/ICCV.2017.514>.
- [16] W. Shi, J. Caballero, F. Huszár, J. Totz, A.P. Aitken, R. Bishop, D. Rueckert, Z. Wang, Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 2016, pp. 1874–1883, <http://dx.doi.org/10.1109/CVPR.2016.207>.
- [17] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, in: Conference on Computer Vision and Pattern Recognition, CVPR, 2014.
- [18] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 2015, pp. 1–9, <http://dx.doi.org/10.1109/CVPR.2015.7298594>.
- [19] J. Kim, J. Lee, K. Lee, Accurate image super-resolution using very deep convolutional networks, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 2016, pp. 1646–1654, <http://dx.doi.org/10.1109/CVPR.2016.182>.
- [20] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 2016, pp. 770–778, <http://dx.doi.org/10.1109/CVPR.2016.90>.
- [21] G. Huang, Z. Liu, L. van der Maaten, K. Weinberger, Densely connected convolutional networks, 2017, <http://dx.doi.org/10.1109/CVPR.2017.243>.
- [22] J. Kim, J.K. Lee, K.M. Lee, Deeply-recursive convolutional network for image super-resolution, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 2016, pp. 1637–1645, <http://dx.doi.org/10.1109/CVPR.2016.181>.
- [23] Y. Tai, J. Yang, X. Liu, Image super-resolution via deep recursive residual network, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 2017, pp. 2790–2798, <http://dx.doi.org/10.1109/CVPR.2017.298>.
- [24] Y. Qiu, R. Wang, D. Tao, J. Cheng, Embedded block residual network: A recursive restoration model for single-image super-resolution, in: 2019 IEEE/CVF International Conference on Computer Vision, ICCV, 2019, pp. 4179–4188, <http://dx.doi.org/10.1109/ICCV.2019.00428>.
- [25] J. Hu, L. Shen, G. Sun, Squeeze-and-excitation networks, in: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 7132–7141, <http://dx.doi.org/10.1109/CVPR.2018.00745>.
- [26] S. Woo, J. Park, J.-Y. Lee, I. Kweon, CBAM: Convolutional Block Attention Module: 15th European Conference, Munich, Germany, September 8–14, 2018, Proceedings, Part VII, 2018, pp. 3–19, http://dx.doi.org/10.1007/978-3-030-01234-2_1.
- [27] Z. Hui, X. Gao, Y. Yang, X. Wang, Lightweight image super-resolution with information multi-distillation network, in: Proceedings of the 27th ACM International Conference on Multimedia, 2019, pp. 2024–2032, <http://dx.doi.org/10.1145/3343031.3351084>.
- [28] E. Agustsson, R. Timofte, NTIRE 2017 challenge on single image super-resolution: Dataset and study, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops, CVPRW, 2017, pp. 1122–1131, <http://dx.doi.org/10.1109/CVPRW.2017.150>.
- [29] Y. Tai, J. Yang, X. Liu, C. Xu, MemNet: A persistent memory network for image restoration, in: 2017 IEEE International Conference on Computer Vision, ICCV, 2017, pp. 4549–4557, <http://dx.doi.org/10.1109/ICCV.2017.486>.
- [30] Z. Hui, X. Wang, X. Gao, Fast and accurate single image super-resolution via information distillation network, in: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 723–731, <http://dx.doi.org/10.1109/CVPR.2018.00082>.
- [31] K. Zhang, W. Zuo, L. Zhang, Learning a single convolutional super-resolution network for multiple degradations, in: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 3262–3271, <http://dx.doi.org/10.1109/CVPR.2018.00344>.
- [32] B.K. Namhyuk Ahn, K.-A. Sohn, Fast, accurate, and lightweight super-resolution with cascading residual network, in: Computer Vision – ECCV 2018, Vol. 11214, 2018, pp. 256–272, http://dx.doi.org/10.1007/978-3-030-01249-6_16.



Beibei Wang was born in Taian, China, in 1988. She received B.S. degree in electronics and information engineering from Shandong Normal University, Jinan, China, in 2010 and the M.S. degree in detection technology and automatic equipment from the University of Electronic Science and Technology, Chengdu, China, in 2013. She is currently pursuing the Ph.D. degree in electronics and communication engineering at Sichuan University, Chengdu, China. Her current research interests include wireless power transmission, microwave plasma sources, super-resolution, and deep learning. Email: 690983790@qq.com.



Changjun Liu is currently a professor in College of Electronics and Information Engineering, Sichuan University. He received his BS degree from Hebei University, and received his MS and Ph.D. degrees from Sichuan University. His research interests are microwave circuits, microwave wireless energy transmission, numerical simulation and other directions. Email: cjliu@scu.edu.cn.



Binyu Yan is currently an associate professor in College of Electronics and Information Engineering, Sichuan University. He received his BS degree from Sichuan University, and received his MS and Ph.D. degrees in communication and information system from Sichuan University. His research interests are image process and pattern recognition. Email: Yanby@scu.edu.cn.



Seunggil Jeon received the B.S. and M.S. degrees from Konkuk University, and Ph.D. degree from Hanyang University (2008). He is a Principal Engineer at Samsung Electronics, Suwon, Korea. Email: simon.sgJeon@gmail.com.



Xiaomin Yang received the B.S. and Ph.D. degrees in communication and information system from Sichuan University, Chengdu, China, in 2002 and 2007, respectively. She held a postdoctoral position with the University of Adelaide. She is currently a Professor with the College of Electronics and Information Engineering, Sichuan University. Her research interests are image processing and pattern recognition. Email: arielyang@scu.edu.cn.



Zhuoyue Zhang was born in Dezhou, China, in 1987. He received the B.S. degree in electronic science and technology from Southwest Jiaotong University, Chengdu, China, in 2009, the M.S. degree in microwave theory and technology from the University of Electronic Science and Technology, Chengdu, China, in 2013, and the Ph.D. degree in information and communication engineering from Sichuan University, Chengdu, China, in 2020. He was a faculty of the University of Electronic Science and Technology, Chengdu, China, from 2013 to 2017. He has been an assistant professor with the West China Medical Center and the College of Electronics and Information Engineering, Sichuan University, Chengdu, China, since 2021. His current research interests include microwave engineering, wireless power transmission, spintronics, microwave antenna. Email: zhangzhuoyue@scu.edu.cn.