Lightweight Super-Resolution Network with Information Distillation and Recursive Methods

Hee-Jo Woo, Ji-Woo Sim, Eung-Tae Kim
Department of Electronic Engineering
Korea Polytechnic University
Siheung, Republic of Korea
wq8336@gmail.com, 2020811010@kpu.ac.kr, etkim@kpu.ac.kr

***Abstract*—At single-image super-resolution, the number of parameters and computations required by deep networks increase, due to the excessive use of convolutional neural networks. So, deep networks could be difficult to use in real-time or low-power devices. To overcome this problem, we propose a lightweight recursive distillation super-resolution network (RDSRN) that uses recursive and information distillation methods to gradually extract hierarchical features, and creates more accurate high-frequency components using high-frequency residual refinement blocks (HFRRB). Experimental results show that the proposed method has better performance with fewer parameters, fewer computations, and faster processing than the conventional methods.**

Keywords*—*deep learning, lightweight network, super resolution, recursion, information distillation, high-frequency refinement

I. Introduction

Single-image super-resolution techniques have been widely used in various applications such as digital TV, mobile devices, and surveillance systems. The performance of deep-learning-based super-resolution algorithms is superior to that of conventional image processing methods, and various networks with different merits have been proposed. Zhang et al. proposed a residual dense network (RDN) [1], in which a residual dense block was employed to fully utilize the information of all the preceding layers to extract a feature map of images. The RDN has a strong representational ability, similar to that of deep networks, and its performance is superior to that of other state-of-the-art methods. However, it requires considerable storage space because of the large number of parameters and high computational complexity. Further, they have limitations in real-time or low-power devices because of the inefficiency in terms of processing time.

We propose a recursive distillation super-resolution network (RDSRN) to overcome these limitations by utilizing a small number of dense connections with the recursive method and the information distillation method proposed by Hui et al. [2].

II. Proposed Super-Resolution Method

The RDSRN is shown in Fig. 1. The network comprises a shallow feature extraction step, a feature extraction step with a recursive block (RB), an upscale step, and a high-frequency residual refinement block (HFRRB) and reconstruction step.

In the shallow feature extraction step, shallow features are extracted through two convolution operations from low-resolution image inputs before the RB.

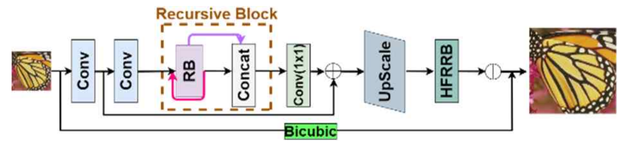

Fig. 1. Architecture of the proposed recursive distillation super-resolution network.

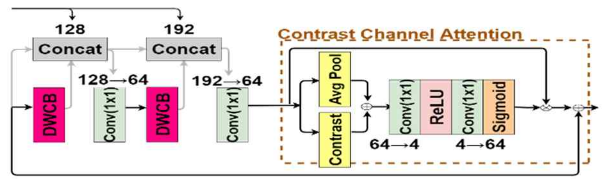

Fig. 2. Recursive block used in the recursive distillation super-resolution network structure.

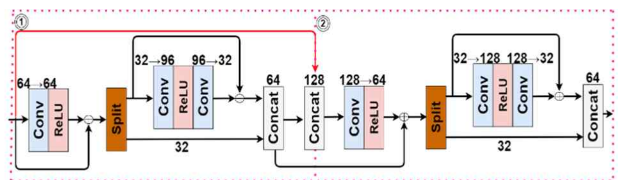

Fig. 3. Distillation wider channel block used in the recursive block.

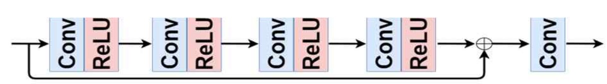

Fig. 4. Structure of high-frequency residual refinement block.

In the feature extraction step, the RB which comprises two distillation wider channel blocks (DWCBs) and a contrast channel attention (CCA) layer is used as shown in Fig. 2. CCA utilizes information from the entire region by assigning weights according to importance. The RB is used recursively and reduces the required storage space while increasing the network depth, thereby maintaining a compact model.

Fig. 3 shows the DWCB. A wider channel (WC) is applied to refined features that split via information distillation, as for the 2nd convolution in the 1st small block and the 5th convolution in the 2nd small block. This prevents the feature suppression caused by the activation function cascading in the deep network. The DWCB does not stack multiple convolution layers but uses the information distillation method. In this method, hierarchical features are progressively learned while maintaining preceding features using dense connections and channel splitting, such that subsequent layers can fully utilize the input information of the preceding layers. The input channels of existing convolution layers are reduced by setting the channel split ratio (CSR) in the range of 0–1. In this manner, the number of parameters and computation complexity can be decreased.

In the upscale step, upscaling is conducted using channel-specific shuffles through sub-pixel convolution. Then, for the features extracted through upscaling, additional training is performed using an HFRRB comprising four convolutions, as shown in Fig. 4. The high-frequency components with loss of detailed information are transformed into more accurate high-frequency components with additional details for improved network performance.

III. Experimental Results

The proposed network is trained using DIV2K images and tested using the Set5 and Urban100 datasets. The network is then compared with the RDN [1], super-resolution convolutional neural network (SRCNN) [3], and cascading residual network (CARN) [4].

When many recursion steps are used to prevent recursion loss, the network is no longer lightweight because the computational complexity rapidly increases, leading to a slow learning rate and large processing time. In addition, the CSR adjusts the number of channels that enter the convolution. Thus, the CSR has a significant impact on the factors that determine the network performance and whether the network is lightweight. The quantitative results obtained using different numbers of recursions and CSRs are summarized in Table I. The least increase in complexity and a high processing speed are achieved with three recursions and an CSR of 0.5, yielding the highest improvement in performance.

Fig. 5(a) and Fig. 5(b) compares the features learned, with and without the WC and HFRRB, respectively, in the second small block of the DWCB in the 3rd recursion. In the 31st channel, the results obtained with the WC and HFRRB show clearer learned features, such as high-frequency components of the eyes, nose, mouth, or edges, than those obtained without the WC and HFRRB.

Table II quantitatively compares the proposed network and other networks. Compared with the RDN, the PSNR of the proposed network is lower by 0.22 dB on average for all test datasets, the number of parameters is 32 times fewer, the computational complexity is 10 times lower, and the processing time is 3 times faster. Compared with the CARN, the PSNR of the proposed network is 0.22 dB higher on average for all datasets, the number of parameters is 2.2 times fewer, the computational complexity is 1.5 times higher, and the processing time is 1.3 times faster.

Fig. 6 visually compares the proposed network and other networks. The reconstructed image obtained using the RDN shows multiple diagonal lines in the direction opposite to that of the lines in the original image. The reconstructed image obtained using the CARN shows diagonal lines with wave patterns. The reconstructed image obtained using the proposed network shows diagonal lines in the direction most similar to that of the lines in the original image.

TABLE I. Quantitative comparisons with respect to number of recursions and channel split ratio for a scale factor X4 on Urban100 datasets.

Urban100, x4	**Recursions**			**Channel Split Ratio**		
	1	**3**	**9**	**0.25**	**0.5**	**0.75**
Params (M)	0.68	0.684	0.71	1.01	0.68	0.49
Multi-Adds (G)	77	136	223	192	136	102
Processing Time (s)	0.0085	0.0112	0.0320	0.0123	0.0112	0.0126
PSNR (dB)	25.95	26.29	26.46	26.36	26.29	26.07

(a) with WC and HFRRB **(b) w/o WC and HFRRB**

Fig. 5. Comparison of features learned with and without the wider channel and high-frequency residual refinement block at scale factor X4 for "baby" in Set5.

TABLE II. Quantitative comparisons of each network for a scale factor X4 on Set5 and Urban100 datasets.

Method	**Params (M)**	**Multi-Adds (G)**	**Processing Time (s)**	**Set5 PSNR (dB)**	**Urban100 PSNR (dB)**
SRCNN [3]	0.057	52.7	-	30.48	24.52
RDN [1]	22	1,309	0.0289	32.47	26.61
CARN [4]	1.5	90.9	0.0130	32.13	26.07
RDSRN	0.68	136	0.0099	32.34	26.29

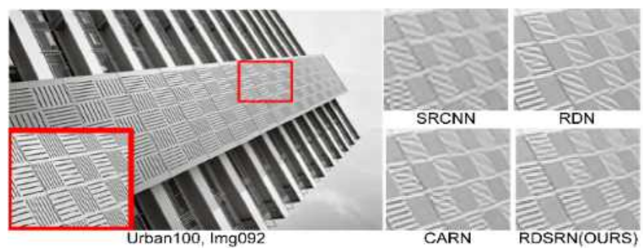

Fig. 6. Visual comparison of the proposed network with other super-resolution methods at scale factor X4 for "Img092" in Urban100.

IV. Conclusion

A lightweight RDSRN is proposed, in which retention and refinement features are split using information distillation and a DWCB with a WC. In addition, an HFRRB is employed to yield high-frequency components with improved accuracy. In the aspect of lightweight network, the experimental results have demonstrated that the performance of the proposed network is quantitatively and visually better than that of existing networks.

References

[1] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual dense network for image super-resolution," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2472–2481, 2018.

[2] Z. Hui, X. Wang, and X. Gao, "Fast and accurate single image super-resolution via information distillation network," IEEE Conference on Computer Vision and Pattern Recognition, pp. 723–731, 2018.

[3] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," IEEE Trans. Pattern. Anal. Mach. Intell., vol. 38, pp. 295–307, 2016.

[4] N. Ahn, B. Kang, and K. Sohn, "Fast, accurate, and lightweight super-resolution with cascading residual network," European Conference on Computer Vision, pp. 256–272, 2018.