

ISTA-Inspired Network for Image Super-Resolution

Yuqing Liu\textsuperscript{1}, Wei Zhang\textsuperscript{1,*}, Weifeng Sun\textsuperscript{2}, Zhikai Yu\textsuperscript{1}, Jianfeng Wei\textsuperscript{1} and Shengquan Li\textsuperscript{1}

\textsuperscript{1} Pengcheng Laboratory, Shenzhen, China.
\textsuperscript{2} Dalian University of Technology, Dalian, China.

Abstract—Deep learning for image super-resolution (SR) has been investigated by numerous researchers in recent years. Most of the works concentrate on effective block designs and improve the network representation but lack interpretation. There are also iterative optimization-inspired networks for image SR, which take the solution step as a whole without giving an explicit optimization step. This paper proposes an unfolding iterative shrinkage thresholding algorithm (ISTA) inspired network for interpretable image SR. Specifically, we analyze the problem of image SR and propose a solution based on the ISTA method. Inspired by the mathematical analysis, the ISTA block is developed to conduct the optimization in an end-to-end manner. To make the exploration more effective, a multi-scale exploitation block and multi-scale attention mechanism are devised to build the ISTA block. Experimental results show the proposed ISTA-inspired restoration network (ISTAR) achieves competitive or better performances than other optimization-inspired works with fewer parameters and lower computation complexity.

I. INTRODUCTION

Image super-resolution (SR), as one of the traditional image restoration tasks, has been widely investigated by researchers\textsuperscript{1, 2}. Given a low-resolution (LR) image, the task of image SR is to restore a corresponding high-resolution (HR) instance with more details. There are numerous applications considering the image SR, such as video deinterlacing\textsuperscript{3} and compression\textsuperscript{4}, remote sensing\textsuperscript{5–7}, EGG analysis\textsuperscript{8}, and spatiotemporal analysis\textsuperscript{9}.

Deep learning has demonstrated its amazing performance in image restoration. There are numerous convolutional neural networks (CNNs) specially designed for image SR. SR-CNN\textsuperscript{10} is the first CNN-based method for image SR. After that, deeper and wider networks show their effectiveness with better performance, such as VDSR\textsuperscript{11}, EDSR\textsuperscript{12}, RDN\textsuperscript{13} and RCAN\textsuperscript{14}. Recent image SR networks usually develop effective blocks for improving the network representation. IMDN\textsuperscript{15} and EFDN\textsuperscript{16} utilize information distillation mechanisms to build an efficient network for fast and accurate image SR. Cross-SRN\textsuperscript{17} builds an edge-preserving network with cross convolution. However, these works concentrate on the block designs but lack interpretation, which limits the performance.

Since image SR can be regarded as a classical optimization task\textsuperscript{18}, there are also works considering building the image SR network from the optimization perspective. IRCNN\textsuperscript{18} provides an iterative solution for the general image restoration task and designs a CNN-based network to solve the denoising prior. ISRN\textsuperscript{19} develops an iterative network with the help of the half-quadratic splitting (HQS) strategy. DPSR\textsuperscript{20} and USRNet\textsuperscript{21} also achieve good performance on image SR inspired by the HQS strategy. There are also works building the network by alternating direction method of multipliers (ADMM). Plug-and-Play ADMM\textsuperscript{22} regards the denoiser as a network prior for different image restoration tasks. ADMMNet\textsuperscript{23} provides an end-to-end network for the compression sensing task. PSRI-Net\textsuperscript{24} considers ADMM for SAR image SR. Although these works provide an interpretable network design, the CNN architectures just task the solution step as a whole, without giving an explicit optimization step on how to solve the denoising problem.

In this paper, we develop an unfolding network based on the iterative shrinkage thresholding algorithm (ISTA). Different from designing the CNN to directly solve the optimization step, ISTA blocks are specially designed to conduct the image restoration following the ISTA steps. In the ISTA block, CNNs are utilized to adaptively learn the functions in the feature space and speed up the optimization steps. To improve the network representation, multi-scale exploration (MSE) and multi-scale attention (MSA) mechanisms are utilized to build the ISTA block. An ISTA-inspired restoration network (ISTAR) is developed based on the ISTA block for effective image SR. Experimental results show the proposed ISTAR can achieve competitive or better performance than other works. Compared with other optimization-inspired methods, ISTAR achieves better performances with much fewer parameters and lower computation complexity. Figure 1 shows an example comparison among different image super-resolution methods. Compared with state-of-the-art methods, our proposed ISTAR can generate more satisfying textures that close to the HR image.

The contributions of this paper can be concluded as follows:
• We analyze the image super-resolution task from the optimization perspective and develop an ISTA block for image super-resolution.
• We develop the multi-scale exploration and multi-scale attention mechanism in the ISTA block, which improves the network representation and boosts the performance.
• Experimental results show the proposed network achieves competitive or better performance than other optimization-based works with much fewer parameters and lower computation complexity.

II. RELATED WORKS

A. Deep Learning for Image Super-Resolution

Deep learning has demonstrated its amazing performance on various computer vision tasks. There are numerous convolutional neural networks (CNNs) specially designed for image super-resolution (SR). SRCNN [10] is the first CNN-based image SR method composed of three convolution layers, which follows a sparse-coding manner. After SRCNN, deeper and wider networks have proposed to improve the restoration performance. FSRCNN [25] increases the network depth and improves the network capacity. ESPCN [26] provides a different upsampling strategy to restore the HR images, which is more effective than the deconvolution operation. Recently, researchers concentrate more on effective block design for better restoration performance. RDN [13] combines the residual connection [27] and densely connection [28], and develops a residual dense block for image SR. After that, the researchers introduce the residual-in-residual design with channel attention [29] for image SR and build an effective network termed RCAN [14]. RFANet [30] expands the residual connection and aggregates the residual features for better information transmission. IMDN [15] and RFDN [16] build the lightweight networks with the help of an information distillation mechanism. SHSR [31] and MSRN [32] utilize hierarchical exploration to further investigate the image features. These works usually concentrate on the effective block designs but neglect to analyze the image SR from the optimization perspective.

B. Optimization-Inspired Image Super-Resolution

There are also optimization-inspired networks for inter-operative image SR. ADMM-Net [23] provides a good example of dealing with the image restoration problem by the optimization strategy and develops a CNN-based denoiser for plug-and-play restoration. IRCNN [18] analyzes the image restoration with the help of the half-quadratic splitting (HQS) strategy and recovers the image with a CNN-based denoiser prior. After IRCNN, there are numerous HQS-based methods for effective image SR. DPSR [20] proposes a different observation model for image SR and uses kernel estimation and CNN denoiser for plug-and-play image SR. USRNet [21] develops an end-to-end network for different image SR tasks. ISRN [19] devises an effective network for image SR under the guidance of HQS and maximum likelihood estimation (MLE). HSRNet [33] also investigates the HQS strategy and develops a network for aliasing suppression image SR. However, these works just take the solution as a whole and calculate it directly by CNNs, without giving an explicit optimization step for each iteration.

III. METHODOLOGY

In this section, we first analyze the image super-resolution (SR) from the optimization perspective and propose an iterative solution with the help of ISTA. Then, we introduce the designed end-to-end network ISTAR. After that, we discuss the design of the ISTA block. Finally, the network settings are described in detail.

A. ISTA for Image Super-Resolution

Given an low-resolution (LR) image $I^{LR}$, the task of image SR is to find a corresponding image $I^{SR}$, satisfying

$$I^{SR} = \arg \min_{I^{SR}} \left| |D I^{SR} - I^{LR}|^2 + \lambda \|I^{LR}\|_1 \right|,$$

where $D$ is the down-sampling matrix, and $\lambda$ is a weighting factor. The prior term $\lambda \|I^{LR}\|_1$ is utilized to introduce the sparsity of the natural image.

To solve this function, we use ISTA to convert it into an iterative manner. Then, the solution is

$$I_{k+1}^{SR} = T_{\alpha_k} (I_k^{SR} - \alpha_k D^T (D I_k^{SR} - I^{LR})), \quad k \geq 0,$$

where $\alpha_k$ is the weighting factor for the $k$-th iteration and $T(\cdot)$ is the soft-thresholding operation.

It can be found that the right hand side of Equation 2 has two independent variables $I_k^{SR}$ and $I^{LR}$. To make it clear for understanding, we re-write Equation 2 as

$$I_{k+1}^{SR} = T_{\alpha_k} ((E - \alpha_k D^T D) I_k^{SR} - \alpha_k D^T I^{LR}), \quad k \geq 0,$$

where $E$ is the identity matrix.

In Equation 3, we can find that $D^T I^{LR}$ is shared for every iteration. In this point of view, we can calculate this term before the ISTA optimization, and regard it as an invariant to speed up the optimization.

B. Network Design

Figure 2 shows the entire network design of our ISTAR. Firstly, the input image $I^{LR}$ is converted into the feature space by one convolutional layer as

$$\hat{I}^{LR} = Conv(I^{LR}).$$

Then, two convolutional layers and one ReLU activation are utilized to calculate the $D^T \hat{I}^{LR}$ for ISTA steps, as shown in the figure. There are $K$ steps for ISTA optimization. For the $k$-th step, there is

$$\hat{I}_k^{SR} = \text{ISTABlock}(D^T \hat{I}^{LR}, \hat{I}_{k-1}^{SR}),$$

where $\text{ISTABlock}(\cdot)$ is the designed ISTA block.
After $K$ iterations, the output $\hat{I}_{SR}^K$ and $\hat{I}_{LR}$ are organized in a skip connection manner as

$$\hat{I}_{SR}^K = \hat{I}_{LR} + \text{Padding}(\hat{I}_{SR}^K),$$

where $\text{Padding}(\cdot)$ aims to introduce the non-linearity. The padding structure is composed of two convolutional layers and one ReLU activation. Finally, the restored image $I_{SR}$ is generated from the SR feature $\hat{I}_{SR}^K$ as

$$I_{SR} = \text{Upscaling}(\hat{I}_{SR}^K),$$

where $\text{Upscaling}(\cdot)$ is the upsampling module. The upsampling module is composed of one convolutional layer and a sub-pixel convolution.

### C. Design of ISTA Block

Figure 3 shows the ISTA block design. The inputs of $k$-th ISTA block are $I_{SR}^{k-1}$ and $D^T I_{LR}$, and the output of the block is $I_{SR}^k$. Multi-scale exploration (MSE) block and multi-scale attention (MSA) mechanism are utilized to generate $(E - \alpha_k D^T D) \hat{I}_{SR}^k$ from $\hat{I}$. Then, one $1 \times 1$ convolution combines the information from the $(E - \alpha_k D^T D) \hat{I}_{SR}^k$ and the $D^T I_{LR}$, and the soft-thresholding block (ST Block) jointly explores the features following the Equation $\ref{equation3}$. A padding structure is introduced to the ISTA block with skip connection for better gradient transmission. The padding structure is composed of two convolutional layers and a ReLU activation.

Figure 4 shows the block design of MSE. Multi-scale design has proved to be an effective structure for image SR. In this block, three different scales ($1 \times 1$, $3 \times 3$, and $5 \times 5$) are considered to explore the hierarchical information from the features. To make the exploration more efficient, the $5 \times 5$ exploration is separated into two $3 \times 3$ convolutions with a ReLU activation, which hold the same receptive field. After the multi-scale feature extraction, one ReLU activation and a convolutional layer concatenate the hierarchical features and generate the final output of the MSE.

Figure 5 shows the block design of multi-scale attention (MSA) mechanism. It can be found that the MSA has a similar multi-scale exploration design as MSE. Hierarchical
The loss function is chosen as \(\ell\) and the Upscaling module. There are convolutional layers are set with 64 except for the ST Block. The ST Block is composed of two convolutional layers and two ReLU activation operations. Then, the Sigmoid activation is introduced to generate the non-negative attention information.

Figure 6 shows the soft-thresholding block (ST Block) design. The ST Block is composed of two 1 \(\times\) 1 convolutional layers, one ReLU activation and a Sigmoid activation. The ST Block calculates the hyper parameters from the input features by the network design, and performs the soft-thresholding action with the learned parameters.

### D. Network Details

In the network, all convolution kernels are set with size 3 \(\times\) 3 except for the MSE and the ST Block. The convolutions in ST Block are set with kernel size as 1 \(\times\) 1. The filters of all convolutional layers are set with 64 except for the ST Block and the Upscaling module. There are \(K = 16\) ISTA blocks in the ISTAR. The loss function is chosen as \(\ell_1\)-norm between the SR and HR images.

### IV. Experiment

#### A. Settings

The network is trained with DIV2K [34] dataset, which contains 900 high resolution images. We choose first 800 images for training, and last 5 image for validation. Five common benchmarks are chosen for comparing the restoration effectiveness: Set14 [35], Set14 [36], B100 [37], Urban100 [38] and Manga109 [39]. We update our ISTAR for 1000 epochs by the Adam [40] optimizer with learning rate \(10^{-4}\). The learning rate is halved for every 200 epochs. The scaling factors are chosen as \(\times2, \times3\) and \(\times4\). The patch size for training is \(48 \times 48\) for LR images. All other settings are same as RDN [13]. The objective indicators are chosen as peak signal-to-noise ratio [41] (PSNR) and structural similarity [42] (SSIM).

#### B. Model Analysis

1) **Investigation on the Iteration Times:** According to the ISTA optimization, the result becomes more accurate with the increase of iteration times. To investigate the effectiveness of the iteration times, we compare the performances with different iteration time \(K\). Table I shows the PSNR/SSIM comparisons among different \(K\) with scaling factor \(\times4\). For a fair comparison, all the testing models are updated for 200 epochs under the same settings. In the table, we can find that the PSNR and SSIM rises with the increase of iteration time \(K\). When \(K = 16\), the network achieves the best performance. In this point of view, a deeper network leads to a better performance. When \(K\) increases from 12 to 16, the PSNR and SSIM gets small improvement. To balance the performance and the computation complexity, we choose \(K = 16\) to build the ISTAR.

2) **Investigation on the Multi-Scale Exploration:** In MSE, we utilize three different scales to explore the hierarchical information. To investigate the effectiveness of multi-scale exploration, we conduct the experiments with scaling combination as 1, (1 + 3) and (1 + 3 + 5). Table II shows the PSNR/SSIM comparisons among different scales combinations. In the figure, we can find that the multi-scale exploration brings 0.08 dB PSNR improvement on Urban100 dataset and 0.05 dB improvement on Manga109 dataset. Compared with \(S = 1\), the combination (1 + 3) brings 0.09 dB PSNR improvement on Manga109 dataset. In this point of view, the multi-scale exploration is an effective design for restoration.

3) **Investigation on the Multi-Scale Attention:** In ISTA block, MSA is considered for better network representation. To show the effectiveness of MSA, we compare the objective performances with and without MSA on different benchmarks. Table III shows the PSNR/SSIM comparisons after training 1000 epochs. In the figure, we can find that the network with MSA has 0.1 dB improvement on Manga109 dataset and 0.06 dB improvement on Set14 and Urban100 dataset. In this point of view, the MSA proves to be an effective component for image restoration and improves the network capacity.

#### C. Results

We compare our ISTAR with several traditional and recent CNN-based image SR works: SRCNN [10], FSR-CNN [25], VDSR [11], DRCN [43], CNF [44], LapSRN [45].
In the figure, we can find that our ISTAR achieves the best performance on all testing benchmarks with all scaling factors. Compared with Cross-SRN, our ISTAR achieves near 0.4 dB, 0.3 dB and 0.2 dB improvement on Urban100 dataset with scaling factor ×2, ×3 and ×4 separately. When scaling factor is ×4, ISTAR achieves 0.34 dB PSNR higher than Cross-SRN on Manga109 dataset. It should be noticed that Urban100 and Manga109 are two representative datasets with plentiful edges.
and lines. In this point of view, ISTAR can effectively restore the structural information than other works.

To further investigate the effectiveness of ISTAR, we compare the network with several optimization-inspired methods. Table V shows the PSNR/SSIM comparisons with different optimization-inspired networks. DBPN [59] is developed by the iterative back projection algorithm, while USRNet [21] and HSRNet [33] are inspired by the half-quadratic splitting (HQS) strategy. The MACs is calculated by the same method as the HSRNet. In the table, we can find that our ISTAR achieves better performance than USRNet and HSRNet. Compared with HSRNet, our method achieves 0.1 dB PSNR improvement on Set5 and Urban100 datasets. Furthermore, ISTAR achieves better performance than USRNet with near 29.7% parameters and 13.7% computation complexity. DBPN is one of the state-of-the-art image SR methods. Compared with DBPN, our method achieves competitive or better performance with much fewer parameters and lower computation complexity.

![Fig. 7. Visualization comparisons on Urban100 dataset with scaling factor ×4](image)

### TABLE V

| Method       | DBPN [59] | USRNet [21] | HSRNet [33] | ISTAR (Ours) |
|--------------|-----------|-------------|-------------|--------------|
| PSNR/SSIM    |           |             |             |              |
| B100         | 33.78/0.9259 | 32.29/0.9010 | 32.27/0.9000 | 32.55/0.9324 |
| Set5         | 38.09 / 0.9600 | 37.69 / 0.9595 | 38.07 / 0.9607 | 38.15 / 0.9610 |
| Set14        | 33.85 / 0.9190 | 33.43 / 0.9159 | 33.78 / 0.9197 | 33.79 / 0.9197 |
| B100         | 32.27 / 0.9000 | 32.09 / 0.9885 | 32.26 / 0.9006 | 32.29 / 0.9010 |
| Urban100     | 32.55 / 0.9324 | 31.78 / 0.9259 | 32.53 / 0.9320 | 32.68 / 0.9331 |

### REFERENCES

[1] L. Zhang, H. Dai, and Y. Sang, “Med-srnet: Gan-based medical image super-resolution via high-resolution representation learning,” *Computational Intelligence and Neuroscience*, vol. 2022, 2022.

[2] S. Lei, H. Zijian, Y. Jiebin, and F. Fengchang, “Super resolution image visual quality assessment based on feature optimization,” *Computational Intelligence and Neuroscience*, vol. 2022, 2022.

[3] Y. Liu, X. Zhang, S. Wang, S. Ma, and W. Gao, “Spatial-temporal correlation learning for real-time video deinterlacing,” in 2021 *IEEE International Conference on Multimedia and Expo (ICME)*, 2021, pp. 1–6.

[4] W. Gao, L. Tao, L. Zhou, D. Yang, X. Zhang, and Z. Guo, “Low-rate image compression with super-resolution learning,” in 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2020, pp. 607–610.

[5] X. Luo, Y. Xia, Y. Zhang, J. Zhu, Y. Zhang, Y. Huang, and J. Yang, “Super-resolution imaging for real aperture radar by two-dimensional deconvolution,” in 2021 *IEEE International Geoscience and Remote Sensing Symposium IGRSS*, 2021, pp. 6630–6633.

[6] L. Li, T. Xu, and Y. Chen, “Fuzzy classification of high resolution remote sensing scenes using visual attention features,” *Computational Intelligence and Neuroscience*, vol. 2017, 2017.

[7] W. Wang, C. Zhang, J. Tian, X. Wang, J. Ou, J. Zhang, and J. Li, “High-resolution radar target recognition via inception-based vgg (ivgg) networks,” *Computational Intelligence and Neuroscience*, vol. 2020, 2020.

[8] J. Stastný and P. Sovka, “High-resolution movement eeg classification,” *Computational intelligence and neuroscience*, vol. 2007, 2007.

[9] Q. Ma, J. Jiang, X. Liu, and J. Ma, “Deep unfolding network for spatiotemporal image super-resolution,” *IEEE Transactions on Computational Imaging*, vol. 8, pp. 28–40, 2022.
[10] C. Dong, C. C. Loy, K. He, and X. Tang. "Image super-resolution using deep convolutional networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38, no. 2, pp. 295–307, 2016.

[11] J. Kim, J. K. Lee, and K. M. 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.

[12] B. Lim, S. Son, H. Kim, S. Nah, and 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.

[13] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu. "Residual dense network for image restoration," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 43, no. 7, pp. 2480–2495, 2021.

[14] Y. K. Li, L. Wang, B. Zhong, and Y. Fu. "Image super-resolution using very deep residual channel attention networks," in Computer Vision - ECCV 2018 - 15th European Conference, Munich, Germany, September 8-14, 2018, Proceedings, Part VII, ser. Lecture Notes in Computer Science, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds., vol. 11211. Springer, 2018, pp. 294–310. [Online]. Available: https://doi.org/10.1007/978-3-030-01254-2_18

[15] Z. Hui, X. Gao, Y. Yang, and X. Wang. "Lightweight image super-resolution with information multi-distillation network," in ACM International Conference on Multimedia (MM), 2019, p. 2024–2032.

[16] J. Liu, J. Tang, and G. Wu. "Residual feature distillation network for lightweight image super-resolution," in Computer Vision - ECCV 2020 Workshops - Glasgow, UK, August 23-28, 2020, Proceedings, Part III, ser. Lecture Notes in Computer Science, A. Bartoli and A. Fusiello, Eds., vol. 12531. Springer, 2020, pp. 41–55. [Online]. Available: https://doi.org/10.1007/978-3-030-67070-2_2

[17] Y. Liu, Q. Jia, X. Fan, S. Wang, S. Ma, and W. Gao. "Cross-srm: Structure-preserving super-resolution network with cross convolution," IEEE Transactions on Circuits and Systems for Video Technology, pp. 1–1, 2021.

[18] K. Zhang, W. Zuo, S. Gu, and L. Zhang. "Learning deep CNN denoiser prior for image restoration," in 2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017. IEEE Computer Society, 2017, pp. 2808–2817. [Online]. Available: https://doi.org/10.1109/CVPR.2017.300

[19] Y. Liu, S. Wang, J. Zhang, S. Wang, S. Ma, and W. Gao. "Iterative network for image super-resolution," IEEE Transactions on Multimedia, pp. 1–1, 2021.

[20] K. Zhang, W. Zuo, and L. Zhang. "Deep plug-and-play super-resolution for arbitrary blur kernels," in IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019. Computer Vision Foundation / IEEE, 2019, pp. 1671–1681. [Online]. Available: http://openaccess.thecvf.com/content_CVPR_2019/html/Zhang_Deep_Plug_And_Play_Super-Resolution_for_Arbitrary_Blur_Kernels_CVPR_2019_paper.html

[21] K. Zhang, L. Van Gool, and R. Timofte. "Deep unfolding network for image super-resolution," in 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2020, Seattle, WA, USA, June 13-19, 2020, pp. 2356–2365.

[22] Y. Liu, X. Zhang, S. Wang, S. Ma, and W. Gao. "Sequential hierarchical learning with distribution transformation for image super-resolution," ACM Trans. Multimedia Comput. Commun. Appl., apr 2022.

[23] J. Li, F. Fang, K. Mei, and G. Zhang. "Multi-scale residual network for image super-resolution," in European Conference on Computer Vision (ECCV), 2018, pp. 527–542.

[24] Y. Liu, Q. Jia, J. Zhang, X. Fan, S. Wang, S. Ma, and W. Gao. "Hierarchical similarity learning for aliasing suppression image super-resolution," CoRR, vol. abs/2006.03361, 2022.

[25] E. Agustsson and R. Timofte. "Ntire 2017 challenge on single image super-resolution: Dataset and study," in IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2017, pp. 1122–1131.

[26] M. Bevilacqua, A. Roumy, C. Guilleminot, and M. line Alberi Morel, "Low-complexity single-image super-resolution based on nonnegative neighbor embedding," in British Machine Vision Conference (BMVC), 2012, pp. 135.1–135.10.

[27] R. Zeyde, M. Elad, and M. Protter. "On single image scale-up using sparse-representations," in International Conference on Curves and Surfaces. Springer, 2010, pp. 711–730.

[28] D. Martin, C. Fowlkes, D. Tal, and J. Malik. "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in IEEE International Conference on Computer Vision (ICCV), vol. 2, 2001, pp. 416–423 vol.2.

[29] J. Huang, A. Singh, and N. Ahuja. "Single image super-resolution from transformed self-exemplars," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 5197–5206.

[30] Y. Matsui, K. Ito, Y. Aramaki, A. Fujimoto, T. Ogawa, T. Yamashita, and Y. Suzawa. "Sketch-based manga retrieval using manga109 dataset," Multimedia Tools and Applications (MTA), vol. 76, no. 20, pp. 21 811–21 838, 2017.

[31] D. P. Kingma and J. Ba. "Adam: A method for stochastic optimization," in International Conference on Learning Representations, 2015.

[32] A. Horé and D. Ziou. "Image quality metrics: Psnr vs. ssim," in International Conference on Pattern Recognition (ICPR), 2010, pp. 2366–2369.

[33] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600–612, 2004.

[34] J. Kim, J. K. Lee, and K. M. Lee. "Deeply-recursive convolutional network for image super-resolution," in IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 1637–1645.

[35] H. Ren, M. El-Khamy, and J. Lee. "Image super resolution based on fusing multiple convolution neural networks," in IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2017, pp. 1050–1057.

[36] W. Lai, J. Huang, N. Ahuja, and M. Yang. "Deep laplacian pyramid networks for fast and accurate super-resolution," in IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 5835–5843.

[37] Y. Tai, J. Yang, and X. Liu. "Image super-resolution via deep recursive residual network," in IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 2790–2798.

[38] Y. Fan, H. Shi, J. Yu, D. Liu, W. Han, H. Yu, Z. Wang, X. Wang, and T. S. Huang. "Balanced two-stage residual networks for image super-resolution," in IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2018, pp. 1157–1165.

[39] Y. Tai, Y. Yang, X. Liu, and C. Xu. "Memnet: A persistent memory network for image restoration," in IEEE International Conference on Computer Vision, 2017, pp. 4549–4557.
[49] J. Choi and M. Kim, “A deep convolutional neural network with selection units for super-resolution,” in 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2017, pp. 1150–1156.

[50] N. Ahn, B. Kang, and K. Sohn, “Fast, accurate, and lightweight super-resolution with cascading residual network,” in European Conference on Computer Vision, vol. 11214, 2018, pp. 256–272.

[51] Y. Wang, L. Wang, H. Wang, and P. Li, “Resolution-aware network for image super-resolution,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 29, no. 5, pp. 1259–1269, 2019.

[52] C. Xie, W. Zeng, and X. Lu, “Fast single-image super-resolution via deep network with component learning,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 29, no. 12, pp. 3473–3486, 2019.

[53] F. Li, H. Bai, and Y. Zhao, “Filternet: Adaptive information filtering network for accurate and fast image super-resolution,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 30, no. 6, pp. 1511–1523, 2020.

[54] Z. He, Y. Cao, L. Du, B. Xu, J. Yang, Y. Cao, S. Tang, and Y. Zhuang, “MRFN: multi-receptive-field network for fast and accurate single image super-resolution,” IEEE Transactions on Multimedia, vol. 22, no. 4, pp. 1042–1054, 2020.

[55] F. Fang, J. Li, and T. Zeng, “Soft-edge assisted network for single image super-resolution,” IEEE Transactions on Image Processing, vol. 29, pp. 4656–4668, 2020.

[56] W. Yang, J. Feng, J. Yang, F. Zhao, J. Liu, Z. Guo, and S. Yan, “Deep edge guided recurrent residual learning for image super-resolution,” IEEE Transactions on Image Processing, vol. 26, no. 12, pp. 5895–5907, 2017.

[57] Y. Liu, Q. Jia, X. Fan, S. Wang, S. Ma, and W. Gao, “Cross-srn: Structure-preserving super-resolution network with cross convolution,” IEEE Transactions on Circuits and Systems for Video Technology, pp. 1–1, 2021.

[58] T. Tong, G. Li, X. Liu, and Q. Gao, “Image super-resolution using dense skip connections,” in IEEE International Conference on Computer Vision. IEEE Computer Society, 2017, pp. 4809–4817.

[59] M. Haris, G. Shakhnarovich, and N. Ukita, “Deep back-projection networks for single image super-resolution,” IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), pp. 1–1, 2020.