



Contents lists available at ScienceDirect

Optik

journal homepage: [www.elsevier.com/locate/ijleo](http://www.elsevier.com/locate/ijleo)



Original research article

## Multi-scale channel network based on filter pruning for image super-resolution



Deyun Wei <sup>\*</sup>, Zhaowu Wang

School of Mathematics and Statistics, Xidian University, Xi'an 710071, China

### ARTICLE INFO

**Keywords:**

Super-resolution  
Deep learning  
Multi-scale channel network  
Filter pruning

### ABSTRACT

Image super-resolution is an important image restoration technology, which is widely used in optical system, optical communication and other fields. Deep convolutional neural networks (CNNs) have made a great breakthrough in accuracy of single image super-resolution. Nevertheless, as the depth and width of the networks increase, CNNs methods face the challenge of memory consumption and computational complexity. To solve this problem, in this paper, we introduce a multi-scale channel network (MSCN) based on filter pruning to reconstruct the high-resolution (HR) image from the original low-resolution (LR) image. Firstly, we propose a multi-scale channel block (MSCB) to obtain the fine features based on the convolution kernels with different receptive fields. Channel split operator is used in MSCB to reduce the model parameters. The wider MSCN with more filters is trained to achieve a better performance. Then, we calculate the rank of feature maps and prune filters with low-rank feature information. Finally, the compressed multi-scale channel network with fewer parameters can achieve similar accuracies. The experimental results validate the effectiveness of our approaches.

### 1. Introduction

With the advances in modern biotechnology and modern medical technology, optical imaging system has received wide attention. Optical super-resolution breaks the diffraction limit and is widely used in optical imaging systems.

Single image super-resolution (SISR) aims to recover a high-resolution (HR) image from its corresponding low-resolution (LR) image, which is an ill-posed problem. To solve this problem, numerous SISR methods have been proposed, including interpolation-based methods [1–5], model-based methods [6–11], and learning-based methods [12–17].

The early interpolation-based methods simply estimate the gray value of pixel points based on the continuity of pixel changes. These methods have a high reconstruction efficiency. But they are powerless for the blurring or noising cases. In model-based methods, amount of prior information are used as a regularization term to limit the solution space. When prior information is inconsistent with the image degradation, these methods will generate blurry images and miss the high-frequency details.

With the development of deep learning, a number of learning-based methods have been proposed. In 2014, Dong et al. [18] first presented the three-layer SRCNN to solve the image SR problem, which achieved superior performance against traditional methods. Inspired by VGG-net [19], Kim et al. [20] increased the depth of network to 20 in VDSR. In order to make full use of shallow feature information of network, Tong et al. [21] presented SRDenseNet which putting dense skip connections into a deep network model. To

\* Corresponding author.

E-mail address: [dywei@xidian.edu.cn](mailto:dywei@xidian.edu.cn) (D. Wei).

further improve the performance, Lim et al. [22] removed unnecessary modules in conventional residual networks to stabilize the training procedure. With the presentation of the attention mechanism, some attention network are also proposed in SR field. Zhang et al. [23] introduced the residual channel attention network (RCAN) by using global average pooling as channel attention mechanism. Zheng et al. [24] proposed contrast-aware channel attention mechanism in IMDN.

Although CNN-based methods have achieved considerable improvements, the existing methods still have many disadvantages. First, CNN-based methods introduce larger scale and more network parameters, which is a hurdle to development in resource-constrained and real time system environments. Second, single scale filter cannot capture more precise feature information. To address these problems, in this paper, we introduce a multi-scale channel network (MSCN) based on filter pruning for SISR. Firstly, we design an effective multi-scale channel block to obtain the fine features with multi scale filter. We train the wider network to achieve better performance. Then, we put test image into MSCN and calculate the rank of each feature map. Inspired by HRank [25], high-rank feature maps are more informative and more important to preserve accuracy, which should be retain. After filters with low-rank feature information are pruned, we fine tune the compressed model. Finally, the compressed multi-scale channel network with fewer parameters can achieve similar accuracies.

## 2. Materials and methods

### 2.1. Multi-scale channel network

In this section, we will present MSCN for SISR. Its graphical depiction is shown in Fig. 1. Specifically, MSCN consists of shallow feature extraction, multi-scale deep feature extraction, feature fusion and upsampling. Furthermore, multi-scale deep feature extraction contains many multi-scale channel blocks (MSCB). In the following, we will give a detailed explanation of proposed model.

Given the input LR image  $I^{LR}$ , we first use one convolution layer to extract the shallow feature of  $F_0$ . This process can be formulated as,

$$F_0 = F_S(I^{LR}) \quad (1)$$

Where  $F_S(\cdot)$  is the shallow feature extraction operation,  $F_0$  is the shallow feature.

Multi-scale deep feature extraction contains some multi-scale channel blocks (MSCB). The structure of MSCB is shown in Fig. 2. One MSCB is composed of three multi-scale channel units (MSCU). In MSCU, convolution split operation is used to produce four-part features. Each part is sent to filters with different receptive fields.  $3 \times 3$  convolution is used to extract two-dimensional information,  $1 \times 3$  convolution can get one-dimensional horizontal information and  $3 \times 1$  convolution captures vertical features. The  $3 \times 3$

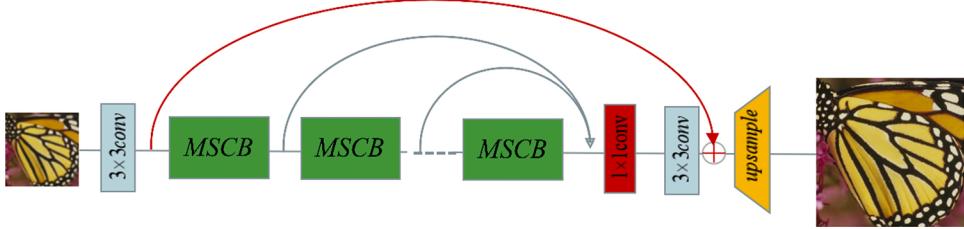


Fig. 1. The structure of the MSCN,  $\oplus$  is the element-wise addition.

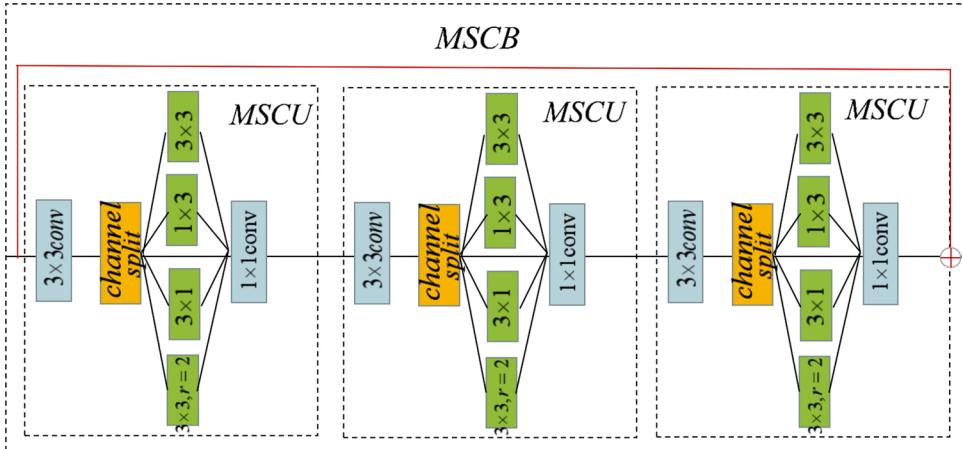


Fig. 2. The architecture of MSCB.

dilated convolution has  $5 \times 5$  receptive fields. It can focus on the semantic information of local pixel blocks. Then, four-part features are fused by  $1 \times 1$  convolution. The MSCB can be described as:

$$F_i = F_{\text{MSCB}_i}(F_{i-1}) = F_{\text{MSCB}_i}^{\text{MSCU}_3}(F_{\text{MSCB}_i}^{\text{MSCU}_2}(F_{\text{MSCB}_i}^{\text{MSCU}_1}(F_{i-1}))) + F_{i-1} \quad (2)$$

Where  $F_{\text{MSCB}_i}(\cdot)$  is the  $i$ th MSCB operation. It receives previous MSCB information  $F_{i-1}$  and generates  $F_i$ .  $F_{\text{MSCB}_i}^{\text{MSCU}_j}(\cdot)$  is the  $j$ th MSCU operation in the  $i$ th MSCB. Local residual learning is used to accelerate model convergence.

The feature fusion part consists of two convolution layers, and its formula is expressed as follows,

$$\begin{aligned} F_{f1} &= F_{F1}(F_1, \dots, F_n) \\ F_{f2} &= F_{F2}(F_{f1}) + F_0 \end{aligned} \quad (3)$$

Where  $F_{F1}(\cdot)$  is the first feature fusion convolution, which receives feature maps of the previous MSCB.  $n$  is the number of MSCB.  $F_{F2}(\cdot)$  is the second feature fusion convolution. Global residual learning in identity branch is used in MSCN.

Finally, the super-resolution image  $I^{SR}$  is obtained through upsampling.

$$I^{SR} = F_{\text{Up}}(F_{f2}) \quad (4)$$

Where  $F_{\text{Up}}(\cdot)$  represents the upsample operation, which is sub-pixel convolution [26].

Given a training set  $\{X^{(i)}, Y^{(i)}\}_{i=1}^N$ , where  $N$  is the number of training patches and  $Y^{(i)}$  is the ground truth high-quality patch of the low-quality patch  $X^{(i)}$ , in training phase, the loss function of our network is

$$L = \frac{1}{N} \sum_{i=1}^N \|Y^{(i)} - F_{\text{MSCN}}(X^{(i)})\|_1 \quad (5)$$

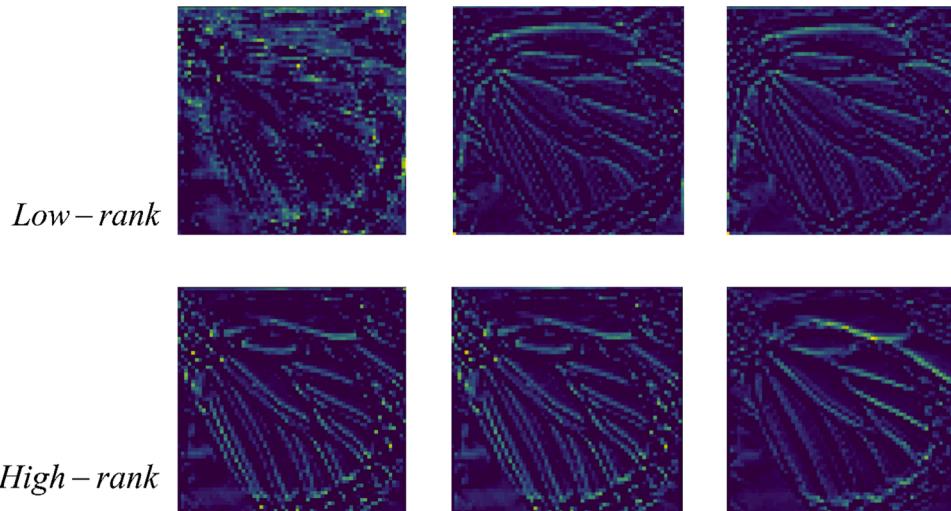
Where  $F_{\text{MSCN}}(\cdot)$  is the proposed MSCN method.

## 2.2. Filter pruning

In order to deploy the proposed MSCN on resource limited devices. We use HRank [25] to compress MSCN to get lightweight model.

In filter pruning, we need to remove less important filters. The rank of feature maps can determine the relative importance of filters. In this paper, we calculate the rank of each feature map and prune filters with low-rank feature information. First, we get a pre-trained MSCN with a set of  $K$  convolutional layers.  $C^i$  is the  $i$ th convolutional layer.  $W_{C_i} = \{w_1^i, w_2^i, \dots, w_{n_i}^i\} \in R^{(n_i \times n_{i-1} \times k_i \times k_i)}$  is a set of 3-D filter, where  $n_i$  is the number of filters in  $C^i$  convolutional layer.  $k_i$  is the kernel size.  $O^i = \{o_1^i, o_2^i, \dots, o_{n_i}^i\} \in R^{(n_i \times g \times h_i \times w_i)}$  represents feature maps. Where  $h_i$  and  $w_i$  are the height and width of the feature map. Then, we input test image  $I$  and get the feature map  $o_j^i(I, :, :)$ , which is generated by the filter  $w_j^i$ .

The specific filter pruning process is as follows: First, we input batch test images to MSCN and calculate the average rank of feature maps. Then, we get the rank set  $R^i = \{r_1^i, r_2^i, \dots, r_{n_i}^i\} \in R^{n_i}$ . Second, we re-rank the rank set in decreasing order  $\widehat{R}^i = \{r_{I_1}^i, r_{I_2}^i, \dots, r_{I_{n_i}}^i\} \in R^{n_i}$ .  $I_j$  is the index of the  $j$ th top value in  $R^n$ . Third, we determine the number of reserved filters and remove the less important filters.



**Fig. 3.** The comparison results of high-rank/low-rank feature maps.

Fig. 3 shows the comparison results of high-rank/low-rank feature maps. The pixel distribution of low-rank features is messy and has some noise interference. The opposite is true for high-rank feature maps. The high-rank feature maps have clear edge and structure.

### 3. Experimental results

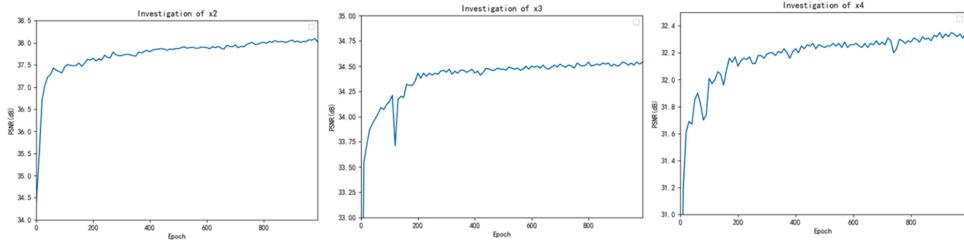
#### 3.1. Experiment settings and datasets

In this paper, we use DIV2K dataset [27] which contains 1000 images with 2K resolution. It consists of 800 training RGB images, 100 validation RGB images, and 100 test images. In order to improve the robustness of the proposed model, We augment the dataset with rotation operation and horizontal/vertical flipping. The batch size in this work is 16. The input image size is  $48 \times 48$ . We use four standard benchmark datasets for testing: Set5 [28], Set14 [29], B100 [30] and Urban100 [31].

For training MSCN, we use Adam optimizer with  $\beta_1 = 0.9$  and  $\beta_2 = 0.999$ . The activation function is LeakyReLU. The initial learning rate is 0.0002, and we set the decay step to decrease the learning rate every 200 steps, where the decay rate is set as 0.5. The proposed MSCN is trained over 1000 epoches.

For filter pruning, we use Stochastic Gradient Descent algorithm (SGD) with an initial learning rate of 0.01. The batch size, weight decay and momentum are set to 16, 0.0005 and 0.9. We choose 100 test images in DIV2k for pruning. For each layer, we retrain the network for 50 epochs after pruning.

We use Pytorch 1.0.0 framework to build our structure, use one NVIDIA GeForce GTX 1060 GPU as the training setup.



**Fig. 4.** Convergence analysis on MSCN. The curves for each method are based on the PSNR on Set5.

**Table 1**

Benchmark results of several state-of-the-art SISR methods. We compare the average PSNR(dB)/SSIM values with the scale factors  $\times 2$ ,  $\times 3$  and  $\times 4$  on Set5, Set14, B100 and Urban100. Red color indicates the best performance and blue color indicates the second best performance.

Algorithm	scale	Parameter	Set5	Set14	B100	Urban100
SRCNN	2	8K	36.66/0.9542	32.45/0.9067	31.36/0.8879	29.50/0.8946
FSRCNN	2	13K	37.00/0.9558	32.63/0.9088	31.53/0.8920	29.88/0.9020
VDSR	2	666K	37.63/0.9587	33.03/0.9124	31.90/0.8960	30.76/0.9140
DRCN	2	1774K	37.63/0.9588	33.04/0.9118	31.85/0.8942	30.75/0.9133
DRRN	2	297K	37.74/0.9591	33.23/0.9136	32.05/0.8973	31.23/0.9188
MemNet	2	677K	37.78/0.9597	33.28/0.9142	32.08/0.8978	31.31/0.9195
CARN	2	1592K	37.76/0.9590	33.52/0.9166	32.09/0.8978	31.92/0.9256
CBPN	2	1026K	<b>37.90/0.9590</b>	<b>33.60/0.9171</b>	<b>32.17/0.8989</b>	<b>32.14/0.9279</b>
MSCN	2	921K	<b>37.91/0.9593</b>	<b>33.59/0.9168</b>	<b>32.18/0.8990</b>	<b>32.10/0.9267</b>
SRCNN	3	8K	32.75/0.9090	29.30/0.8215	28.41/0.7863	26.24/0.7989
FSRCNN	3	13K	33.16/0.9140	29.43/0.8242	28.53/0.7910	26.43/0.8080
VDSR	3	666K	33.66/0.9213	29.77/0.8314	28.82/0.7976	27.14/0.8279
DRCN	3	1774K	33.82/0.9226	29.76/0.8311	28.80/0.7963	27.15/0.8276
DRRN	3	298K	34.03/0.9244	29.96/0.8349	28.95/0.8004	27.53/0.8378
MemNet	3	677K	34.09/0.9248	30.00/0.8350	28.96/0.8001	27.56/0.8376
CARN	3	1592K	<b>34.29/0.9255</b>	<b>30.29/0.8407</b>	<b>29.06/0.8034</b>	<b>27.38/0.8404</b>
MSCN	3	930K	<b>34.34/0.9263</b>	<b>30.30/0.8415</b>	<b>29.08/0.8043</b>	<b>28.04/0.8490</b>
SRCNN	4	8K	30.48/0.8628	27.49/0.7503	26.90/0.7101	24.52/0.7221
FSRCNN	4	13K	30.71/0.8657	27.59/0.7535	26.98/0.7150	24.62/0.7280
VDSR	4	666K	31.35/0.8838	28.01/0.7674	27.29/0.7251	25.18/0.7524
DRCN	4	1774K	31.53/0.8854	28.02/0.7670	27.23/0.7233	25.14/0.7510
DRRN	4	298K	31.68/0.8888	28.21/0.7720	27.38/0.7284	25.44/0.7638
MemNet	4	677K	31.74/0.8893	28.26/0.7723	27.40/0.7281	25.50/0.7630
CARN	4	1592K	32.13/0.8937	28.60/0.7806	<b>27.58/0.7349</b>	<b>26.07/0.7837</b>
CBPN	4	1197K	<b>32.21/0.8944</b>	<b>28.63/0.7813</b>	<b>27.58/0.7356</b>	<b>26.14/0.7869</b>
MSCN	4	942K	<b>32.18/0.8942</b>	<b>28.65/0.7817</b>	<b>27.55/0.7351</b>	26.04/0.7837

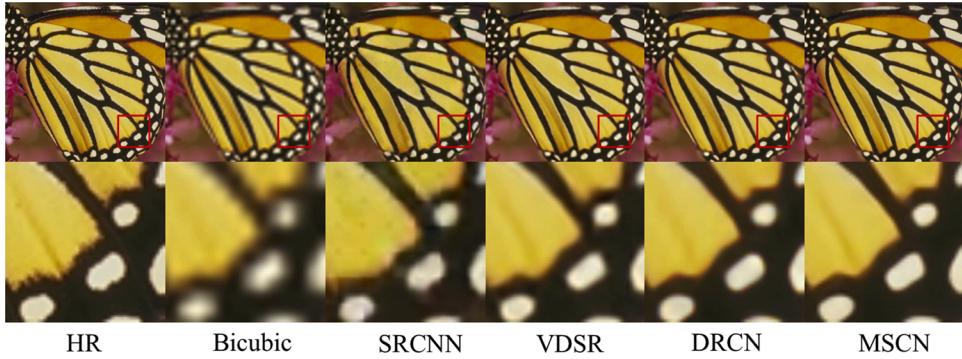
### 3.2. Filter number

MSCN contains 6 MSCB and each MSCB has 3 MSCU. In order to achieve better performance, we set the initial number of filters is 256. In MSCB, multi-scale filter has 64 filters. The convergence analysis on wider MSCN is shown in Fig. 4. After pruning filter, multi-scale filter has 16 filters and other number of filters is 64.

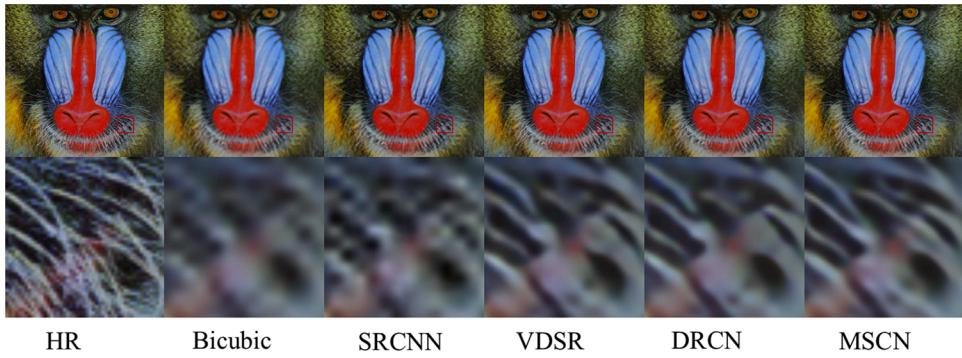
### 3.3. Comparison with other methods

In the following, we will compare with 9 learning-based methods in PSNR [32] and SSIM [33]: SRCNN [18], FSRCNN [34], VDSR [20], DRCN [35], DRRN [36], MemNet [37], CARN [38], CBPN [39] and pruned MSCN. We conduct subjective and objective experimental analysis in the same testing dataset. Table 1 shows quantitative comparisons for  $\times 2$ ,  $\times 3$ , and  $\times 4$  SR. In order to highlight the performance, we use red color indicates the best performance and blue color indicates the second best performance. It can find out that our MSCN performs favorably against other compared approaches on most datasets, especially at the scaling factor of  $\times 3$ .

The visual comparisons of pruned MSCN with other SR methods on Set5, Set14 and B100 datasets are shown in Figs. 5–7.



**Fig. 5.** The ‘butterfly’ image on Set5 with an upscaling factor 4.



**Fig. 6.** The ‘baboon’ image on Set14 with an upscaling factor 4.



**Fig. 7.** The 37037 image on B100 with an upscaling factor 4.

Specifically, in Fig. 5, for the details of butterfly wings, we find that other compared methods would produce blurred edges. In contrast, MSCN method can recover sharper and clearer edges, more faithful to the ground truth. Some details of the baboon image of Set14 is shown in Fig. 6. The reconstructed word in baboon from other methods is more vague than MSCN. Fig. 7 shows the details of the aircraft tail, all the compared methods fail to recover it. From our result, MSCN can recover it better.

#### 4. Conclusion

In this paper, we introduce a multi-scale channel network (MSCN) based on filter pruning for SISR. The proposed multi-scale channel block in MSCN can catch the fine features based on the convolution kernels with different receptive fields. Channel split operator can extends the receptive field and reduces the model parameters. In order to further improve the performance, we first train the wider network and then compress it by removing the filters with low-rank feature information. Finally, the compressed MSCN with fewer parameters can achieve similar accuracies. Numerous experiments have validated the effectiveness of our approaches.

#### Conflict of interest

The authors declare no conflict of interest.

#### Declaration of Competing Interest

The authors report no declarations of interest.

#### Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 61971328.

#### References

- [1] M. Unser, A. Aldroubi, M. Eden, et al., Fast b-spline transforms for continuous image representation and interpolation, *IEEE Trans. Pattern Anal. Mach. Intell.* 13 (1991) 277–285.
- [2] D. Wei, Image super-resolution reconstruction using the high-order derivative interpolation associated with fractional filter functions, *IET Signal Process.* 10 (2016) 1052–1061.
- [3] H. Hou, H. Andrews, Cubic splines for image interpolation and digital filtering, *IEEE Trans. Acoust. Speech Signal Process.* 26 (1978) 508–517.
- [4] D. Wei, Y.-M. Li, Generalized sampling expansions with multiple sampling rates for lowpass and bandpass signals in the fractional fourier transform domain, *IEEE Trans. Signal Process.* 64 (2016) 4861–4874.
- [5] D. Wei, Y. Li, Sampling reconstruction of n-dimensional bandlimited images after multilinear filtering in fractional fourier domain, *Optics Commun.* 295 (2013) 26–35.
- [6] G. Pandey, U. Ghanekar, A compendious study of super-resolution techniques by single image, *Optik* 166 (2018) 147–160.
- [7] J. Jeon, J. Paik, Single image super-resolution based on subpixel shifting model, *Optik* 126 (2015) 4954–4959.
- [8] M. Gao, S. Qin, High performance super-resolution reconstruction of multi-frame degraded images with local weighted anisotropy and successive regularization, *Optik* 126 (2015) 4219–4227.
- [9] I. ul Haq, A.A. Mudassar, Geometric super-resolution using negative rect mask, *Optik* 168 (2018) 323–341.
- [10] X. Yang, T. Liu, D. Zhou, An adaptive super-resolution method based on regional pixel information and ringing artifacts suppression, *Optik* 125 (2014) 5962–5968.
- [11] X. Yang, T. Liu, D. Zhou, A multi-frame adaptive super-resolution method using double channel and regional pixel information, *Optik* 126 (2015) 5850–5858.
- [12] C. Deng, W. Tian, S. Wang, H. Zhu, W. Rao, S. Hu, Structural similarity based single image super-resolution with nonlocal regularization, *Optik* 125 (2014) 4005–4008.
- [13] T.-H. Kim, H. Oh, K. Kim, Y. Lee, Investigating single image super-resolution algorithm with deep learning using convolutional neural network for chest digital tomosynthesis, *Optik* 203 (2020) 164070.
- [14] Y. Tang, J. Zhang, M. Yue, Z. Qu, X. Wang, Y. Gui, X. Feng, Deep learning-based super-resolution images for synchronous measurement of temperature and deformation at elevated temperature, *Optik* 226 (2021) 165764.
- [15] M.Y. Abbas, Residual dense convolutional neural network for image super-resolution, *Optik* (2020) 165341.
- [16] L. Chen, Q. Kou, D. Cheng, J. Yao, Content-guided deep residual network for single image super-resolution, *Optik* 202 (2020) 163678.
- [17] J. Lu, B. Wu, Single-image super-resolution with joint-optimization of tv regularization and sparse representation, *Optik* 125 (2014) 2497–2504.
- [18] C. Dong, C.C. Loy, K. He, X. Tang, Image super-resolution using deep convolutional networks, *IEEE Trans. Pattern Anal. Mach. Intell.* 38 (2015) 295–307.
- [19] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-scale Image Recognition, 2014 (arXiv preprint), arXiv:1409.1556.
- [20] J. Kim, J. Kwon Lee, K. Mu Lee, Accurate image super-resolution using very deep convolutional networks, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016) 1646–1654.
- [21] T. Tong, G. Li, X. Liu, Q. Gao, Image super-resolution using dense skip connections, *Proceedings of the IEEE International Conference on Computer Vision* (2017) 4799–4807.
- [22] B. Lim, S. Son, H. Kim, S. Nah, K. Mu Lee, Enhanced deep residual networks for single image super-resolution, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (2017) 136–144.
- [23] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, Y. Fu, Image super-resolution using very deep residual channel attention networks, *Proceedings of the European Conference on Computer Vision (ECCV)* (2018) 286–301.
- [24] Z. Hui, X. Gao, Y. Yang, X. Wang, Lightweight image super-resolution with information multi-distillation network, *Proceedings of the 27th ACM International Conference on Multimedia* (2019) 2024–2032.
- [25] M. Lin, R. Ji, Y. Wang, Y. Zhang, B. Zhang, Y. Tian, L. Shao, Hrank: filter pruning using high-rank feature map, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2020) 1529–1538.
- [26] 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, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016) 1874–1883.
- [27] E. Agustsson, R. Timofte, Ntire 2017 challenge on single image super-resolution: dataset and study, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (2017) 126–135.

- [28] M. Bevilacqua, A. Roumy, C. Guillemot, M.L. Alberi-Morel, Low-Complexity Single-Image Super-Resolution Based on Nonnegative Neighbor Embedding, 2012.
- [29] R. Zeyde, M. Elad, M. Protter, On single image scale-up using sparse-representations, International Conference on Curves and Surfaces (2010) 711–730.
- [30] D. Martin, C. Fowlkes, D. Tal, J. Malik, A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics, in: Proceedings Eighth IEEE International Conference on Computer Vision, ICCV 2001, vol. 2, IEEE, 2001, pp. 416–423.
- [31] J.-B. Huang, A. Singh, N. Ahuja, Single image super-resolution from transformed self-exemplars, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2015) 5197–5206.
- [32] Q. Huynh-Thu, M. Ghanbari, Scope of validity of psnr in image/video quality assessment, Electron. Lett. 44 (2008) 800–801.
- [33] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (2004) 600–612.
- [34] C. Dong, C.C. Loy, X. Tang, Accelerating the super-resolution convolutional neural network, European Conference on Computer Vision (2016) 391–407.
- [35] J. Kim, J.K. Lee, K.M. Lee, Deeply-recursive convolutional network for image super-resolution, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2016) 1637–1645.
- [36] Y. Tai, J. Yang, X. Liu, Image super-resolution via deep recursive residual network, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2017) 3147–3155.
- [37] Y. Tai, J. Yang, X. Liu, C. Xu, Memnet: A persistent memory network for image restoration, Proceedings of the IEEE International Conference on Computer Vision (2017) 4539–4547.
- [38] N. Ahn, B. Kang, K.-A. Sohn, Fast, accurate, and lightweight super-resolution with cascading residual network, Proceedings of the European Conference on Computer Vision (ECCV) (2018) 252–268.
- [39] F. Zhu, Q. Zhao, Efficient single image super-resolution via hybrid residual feature learning with compact back-projection network, Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops (2019).