MCSNet: Multi-Channel Sharing Network for Single Image Super-Resolution

Wazir Muhammad†, Supavadee Aramvith*
and Watchara Ruangsang*

*Multimedia Data Analytics and Processing Unit, Department of Electrical Engineering,
Faculty of Engineering, Chulalongkorn University, Bangkok 10330, Thailand.
†Electrical Engineering Department BUET, Khuzdar, Pakistan.
E-mail: wazir.laghari@gmail.com
E-mail: supavadee.a@chula.ac.th
E-mail: watcharaknot@gmail.com

Abstract— **Due to the rapid development of convolutional neural networks (CNNs) architecture, single-image super-resolution (SISR) performance and processing speed increased. However, CNN-based architectures have a high computational cost, more processing time, and reconstructed results are still unsatisfactory. Furthermore, CNN architectures depend on the single-channel network architecture to reconstructing the high-resolution (HR) image, which also creates problems during the training, extracted features are not received at the last end layer. To report these issues, we suggested a Multi-Channel Sharing Network for Single Image Super-Resolution (MCSNet). In addition, the multi-channel sharing technique is used to extract and share the feature information through different routes, which is very important to resolve the SISR problem. Numerous quantitative and qualitative experimental findings demonstrate that our suggested method outperforms existing state-of-the-art approaches in terms of peak signal-to-noise ratio (PSNR), structural similarity index measure (SSIM), computing complexity, and perceptual quality. Especially on challenging enlargement scale factor 8×, our proposed method improves the overall average PSNR on all test datasets, including Manga109, Urban100, Set14, and Set5, by 2.04dB with the baseline method.**

I. Introduction

Single image super-resolution's goal is to reconstruct the damaged low-quality, low-resolution input image into a high quality, high-resolution (HR) output image [1-6]. The SISR is still a challenging task to reconstruct the HR image due to its ill-posed problem and more computational cost. Various image super-resolution (SR) methods have been developed to report above issues, including bicubic interpolations, reconstruction, and learning-based approaches. Deep CNN learning-based image SR methods have recently achieved tremendous performance due to the rapid improvement in the convolutional neural network (CNN) design. Currently, various image SR approaches using deep learning-based CNN techniques proposed by research community. Among them, Dong et al. [7] initially proposed the shallow-based architecture using three CNN layers to develop the direct relationships between an original LR input image and a high-resolution reconstructed output image known as a super-resolution convolutional neural network (SRCNN). These layers perform three basic operations: feature extraction, non-linear mapping, and feature restoration. This approach's main flaws are using the bicubic interpolation as a pre-processing step to upscale the LR image as an input and reconstruct the interpolated version of an HR output image. The author of SRCNN [7] modified the architecture and replaced the pre-upsampling bicubic interpolation technique with a post-upsampling learnable layer (transpose convolution) named Fast SR Convolutional Neural Network (FSRCNN) [8] This modification leads to a decrease in the number of parameters and memory consumption. To improve the shortcomings observed during the SRCNN [7] and FSRCNN [8] models, Kim et al. used the concept of the Visual Geometry Group (VGG) model, which is employed in the ImageNet classification task. They proposed the 20-layer network known as Very Deep Super Resolution (VDSR) [9]. The performance of VDSR is better than earlier approaches such as SRCNN [7] and FSRCNN [8] , with a large margin gap. Furthermore, VDSR [9] employed the residual learning technique to avoid the vanishing gradient problem and increase convergence speed. The relationship between SRCNN [7] and the sparse coding networks (SCN) methods confirms its acceptable performance. Wang et al. [10] replaced mapping layers with the set of sparse-based coding type networks, commonly known as SCN (sparse coding network). The deeper it is the better to follow this concept. Kim et al. [11] proposed the image SR method using a deeply recursive CN (DRCN). This architecture used a deep recursive type layer (16 times recursions). DRCNN [11] method same convolutional layers repeatedly apply several times as desired, but the computational cost does not increase although more times recursions are used. Likewise, Tai et al. suggested the concept of a two-way novel network, one for a deep recursive residual network (DRRN) [12], and the next used for a persistent memory network (MemNet) [13]. The earlier design is employed for the recursive type of learning to reduce the number of parameters. The final part of a model resolves the long-term dependency problem. However, these two methods used more processing time and extra huge memory consumption in testing and training. A robust Charbonnier loss function network architecture proposed by Lai et al. presented the Laplacian pyramid SR network (LapSRN) [14] to reconstruct the HR images gradually. Recently, deep CNN-based single image SR [15, 16] is to attain improved performance compared to conventional methods, but still faces some challenging issues [17]. Generally, the deep CNN model used a convolutional layer side-by-side to design a deeper framework but introduce the vanishing gradient problem. Later, end layers work as a dead layer. The interpolation method was also employed by certain well known deep CNN methods [7, 9] to upscale the LR image before extracting CNN features from the upscaled HR image. These approaches are unsuccessful because they introduce new noise in the models and use more burden on the model during the training. Additionally, deep CNN relies on two-way multi-scale channel sharing type network architecture, which does not extract the complete feature information through the previous layers. To address these problems, we have suggested the Multi-Channel Sharing Network for Single Image Super-Resolution, named MCSNet, which uses the Dense channel sharing ResNet blocks (DCSRB) with local and global skip connections. In summary, we establish a novel Multi-Channel Sharing Network for Single Image Super-Resolution, which improves PSNR/SSIM, and perceptual quality of the LR image. The overall main contribution of our MCSNet method can be summarized as follow:

- Inspired by Multi-scale Channel Sharing with ResNet architecture, we proposed a Multi-Channel Sharing Network for Single Image Super-Resolution method with a postupsampling technique to reconstruct the high-quality HR output image.
- We proposed a new multi-channel path scheme to boost the representations of features of the HR image excellently. Multi-channel path scheme consists of dense channel sharing ResNet block (DCSRB), which facilitates the feature extraction through each earlier and subsequent CNN layer.
- Conventional deep CNN methods used Rectified Linear Unit (ReLU) activation function, which introduces the vanishing-gradient problems in training, in case of deeper network architecture. In order to prevent the vanishing-gradient issues and more effectively train the model, we replaced the ReLU with a leaky ReLU activation function.

The remaining parts of the paper are divided into the following sections: Our proposed network architecture is presented in Section II. Section III discuss the experimental calculations. The conclusion section is reported in the last section.

II. PROPOSED METHOD

In this paper, we propose the Multi-Channel Sharing Network for Single Image Super-Resolution with the support of dense channel sharing Resnet block (DCSRB). First, we discuss the proposed network architecture of MCSNet in detail, and then discuss the DCSRB block with different CNN layers used in the main architecture, as shown in Fig. 1. Leaky Rectified Linear Unit (LeakyReLU) activation function is utilized after two CNN layers in the proposed MCSNet network. These CNN layers extracted the initial features and then added the four DCSRB blocks side-by-side and timely updated features information received through local and global skip connections. Finally, the cumulative LR features pass through the deconvolution layer to upscale and reconstruct the HR output image. The design of LReLU depends on the ReLU, and it provides two outputs, one for (0, x) and the other for (x, 0). Additionally, the earlier approach depends on a maximum of two-channel sharing residual blocks to extract the image features of different scales [18]. The reconstructed features information is satisfactory, but the information is lost in the case of deeper and denser network architecture. It cannot extract the detailed features, and sometimes information is stuck. In our proposed DCSRB block design in such a way to split the original features information with five different kernel sizes of multi-scale channels in the order of 1×1, 3×3, 5×5, 7×7, and 9×9. The first branch used 1×1, convolutional neural network layers to reduce the dimension of the input data and the computational cost of the model [19]. The other branches of different kernel sizes are used to extract the multi-channel information, as shown in Fig. 2. Finally, in the reconstruction phase, an upsampling operation is used by 3×3 deconvolution layer to generate the high-quality output HR image. To reduce the computational cost and size of the dimension, we used 1×1 bottleneck layer before the deconvolution layer.

III. EXPERIMENTS

The experimental section explains, the training and testing datasets strategy for reconstructing the HR image. Our MCSNet used three different training datasets, such as Yang et al. [20], BSDS100 [21], and DIV2K [22] (including 50 images) for training purposes. The complete training dataset has been divided into two sections, like 80% used in training, remaining 20% in validation. Our MCSNet is tested on four benchmark datasets, including Set14 [20], Set5 [23], Urban100 [24], and Manga109 [25]. Set5 [23], and Set14 [20] used the images in the order of 5 and 14. In Urban100 and Manga109 [25] available images are in the order of 100 and 109. We evaluated our model with frequently used quality metrics is, the PSNR/SSIM, because these metrics directly related to pixel values. Windows 10 operating system with the supporting Graphics card of NVIDIA GeForce RTX2070 GPU model having RAM of 16 GB used for training and testing operations. Training performs on an enlargement factor 4× and 8× in Keras 2.15 and MATLAB2018a environment. The quantitative experimental results are in Table I, showing that the proposed method attains higher average PSNR values. Furthermore, computational costs in terms of number of parameters plays a vital role in real-time applications. Our proposed method used alternate strategy of multi-path approach to reduce the number of parameters more significantly as compared to state-of-the-art methods. Fig. 3, clearly shows that our model has a smaller number of parameters as well as improved PSNR (dB). Our parameters calculations performed on scale 4× overall test datasets (including Set5, Set14, Urban100, and Manga109). Additionally, the performance of proposed method evaluates with respect to perceptual quality using different publicly available methods, presents in the Fig. 4. Existing methods with proposed method initially used three images of butterfly, zebra, and PrayerHaNemurenai obtained from Set5, Set14, and Manga109 with scale factor 8×. The blurry texture generated output due to the use of bicubic interpolation. Other methods such as SRCNN, VDSR, and DRCN produce improved results than the baseline method. In our proposed method, the reconstructed output is a visually pleasing texture detail with a high value of PSNR/SSIM.

IV. CONCLUSION

Multi-Channel Sharing Network for Single Image Super-Resolution (MCSNet) using convolution, LeakyReLU, and dense channel sharing Resnet block (DCSRB) proposed in this paper for reconstructing the visually pleasing HR output image. To increase the quantitative and qualitative evaluation of the LR image, we employed the four DCSRB bocks continued by the LeakyReLU function. Furthermore, to improve the training convergence process and fix the dying ReLU problem in the dense CNN model, we replaced the ReLU activation function with LeakyReLU non-linear activation function. Moreover, the whole network architecture adopts local and global skip connections to detain the high-frequency feature information for single image SR restoration. Quantitative / qualitative presentation evaluated on four public test datasets demonstrates that the proposed MCSNet can provide superior performance in terms of reconstructed results.

TABLE I

THE AVERAGE PSNR COMPARISON WITH EXISTING ALGORITHMS USING FOUR TEST DATASETS ON CHALLENGING ENLARGEMENT FACTOR 4× AND 8×.

Method	Scale	SET5	SET14	URBAN100	MANGA109	Average
Bicubic	4×	28.42	26.00	23.14	24.89	25.61
SRCNN	4×	30.48	27.5	24.52	27.58	27.52
FSRCNN	4×	30.72	27.61	24.62	27.90	27.71
SCN	4×	30.41	27.39	24.52	27.39	27.43
VDSR	4×	31.35	28.01	25.18	28.83	28.34
DRCN	4×	31.53	28.02	25.14	28.93	28.41
MCSNet	4×	31.16	28.87	26.68	28.87	**28.90**
Bicubic	8×	24.40	23.10	20.74	21.47	22.43
SRCNN	8×	25.33	23.76	21.29	22.46	23.21
FSRCNN	8×	20.13	19.75	21.32	22.39	20.90
SCN	8×	25.59	24.02	21.52	22.68	23.45
VDSR	8×	25.93	24.26	21.70	23.16	23.76
DRCN	8×	25.93	24.25	21.71	23.20	23.77
MCSNet	8×	25.95	25.12	23.38	23.43	**24.47**

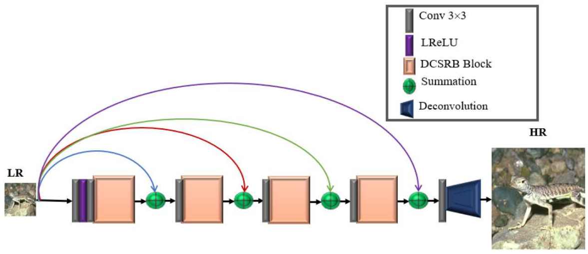

Fig. 1 Proposed Multi-Channel Sharing Network for Single Image Super-Resolution (MCSNet)

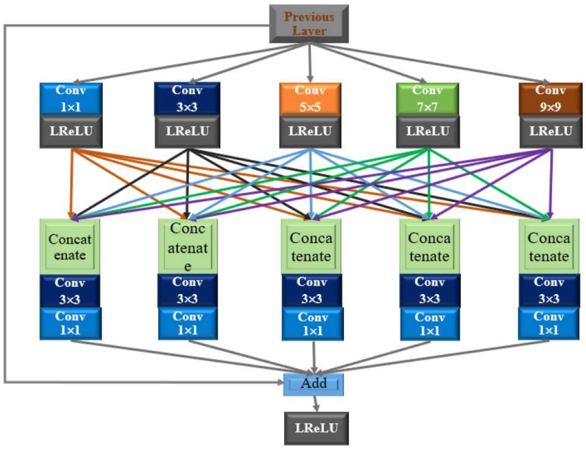

Fig. 2 Dense channel sharing ResNet Block (DCSRB)

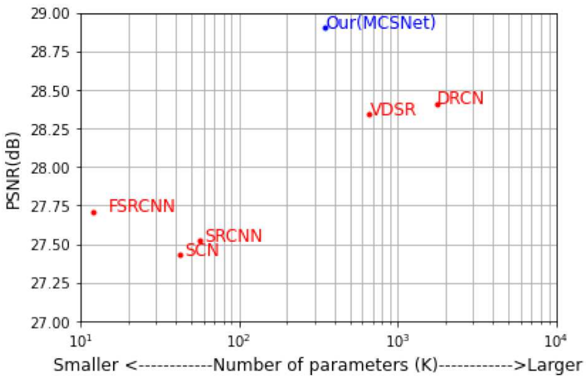

Fig. 3 Evaluate the computational complexity in terms of overall average PSNR versus model parameters on enlargement factor 4x (including Set5, Set14,

Urban100, and Manga109).

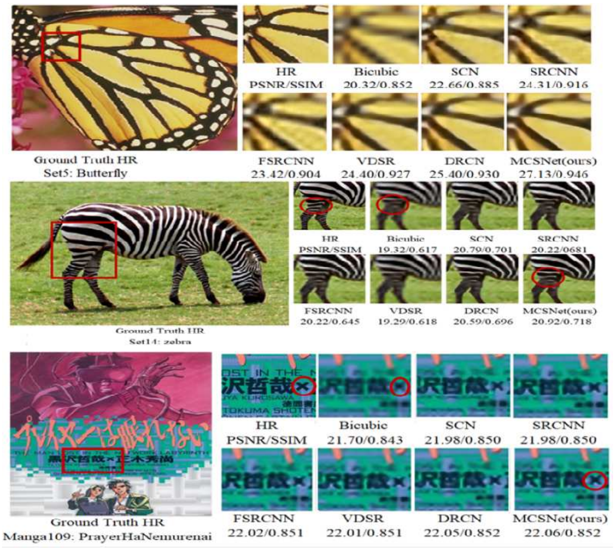

Fig. 4 Presents the perceptual quality comparison on enlargement factor 8× of various image SR methods using test datasets of Set5, Set14 and Manga109.

ACKNOWLEDGMENT

This work is supported by the Second Century Fund (C2F) Chulalongkorn University Bangkok Thailand, Electrical Engineering Department Chulalongkorn University Bangkok, Thailand, Thailand Science research and Innovation Fund Chulalongkorn University (CU_FRB65_ind-(9)157_21_23), Ratchadaphiseksomphot Endowment Fund (Multimedia Data Analytics and Processing Research Unit).

REFERENCES

[1] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael, "Learning Low-Level Vision," *International Journal of Computer Vision,* vol. 40, no. 1, pp. 25-47, 2000/10/01 2000, doi: 10.1023/A:1026501619075.

[2] J. Luo, L. Zhao, L. Zhu, and W. Tao, "Multi-scale receptive field fusion network for lightweight image super-resolution," *Neurocomput.,* vol. 493, no. C, pp. 314–326, 2022, doi: 10.1016/j.neucom.2022.04.038.

[3] X. Zhang *et al.*, "Remote Sensing Image Super-Resolution via Dual-Resolution Network Based on Connected Attention Mechanism," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 60, pp. 1-13, 2022, doi: 10.1109/TGRS.2021.3106681.

[4] X. Yang, T. Xie, Y. Guo, and D. Zhou, "Remote sensing image super-resolution based on convolutional blind denoising adaptive dense connection," *IET Image Processing,* vol. 15, no. 11, pp. 2508-2520, 2021.

[5] D. Qin and X. Gu, "Single-image super-resolution with multilevel residual attention network," *Neural Comput. Appl.,* vol. 32, no. 19, pp. 15615–15628, 2020, doi: 10.1007/s00521-020-04896-6.

[6] J. Jiang, H. M. Kasem, and K. W. Hung, "A Very Deep Spatial Transformer Towards Robust Single Image Super-Resolution," *IEEE Access,* vol. 7, pp. 45618-45631, 2019, doi: 10.1109/ACCESS.2019.2908996.

[7] 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, doi: 10.1109/TPAMI.2015.2439281.

[8] C. Dong, C. C. Loy, and X. Tang, "Accelerating the Super-Resolution Convolutional Neural Network," in *Computer Vision – ECCV 2016*, Cham, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds., 2016// 2016: Springer International Publishing, pp. 391-407.

[9] 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)*, 27-30 June 2016 2016, pp. 1646-1654, doi: 10.1109/CVPR.2016.182.

[10] Z. Wang, D. Liu, J. Yang, W. Han, and T. S. Huang, "Deep Networks for Image Super-Resolution with Sparse Prior," *2015 IEEE International Conference on Computer Vision (ICCV),* pp. 370-378, 2015.

[11] J. Kim, J. K. Lee, and K. M. Lee, "Deeply-Recursive Convolutional Network for Image Super-Resolution," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 27-30 June 2016 2016, pp. 1637-1645, doi: 10.1109/CVPR.2016.181.

[12] Y. Tai, J. Yang, and X. Liu, "Image Super-Resolution via Deep Recursive Residual Network," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 2790-2798, 2017.

[13] Y. Tai, J. Yang, X. Liu, and C. Xu, "Memnet: A persistent memory network for image restoration," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 4539-4547.

[14] W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang, "Deep Laplacian Pyramid Networks for Fast and Accurate Super-Resolution," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 5835-5843, 2017.

[15] W. Muhammad, S. Aramvith, and T. Onoye, "Multi-scale Xception based depthwise separable convolution for single image super-resolution," *Plos one,* vol. 16, no. 8, p. e0249278, 2021.

[16] W. Muhammad and S. Aramvith, "Multi-Scale Inception Based Super-Resolution Using Deep Learning Approach," *Electronics,* vol. 8, no. 8, p. 892, 2019. [Online]. Available: https://www.mdpi.com/2079-9292/8/8/892.

[17] Y. Cai, G. Gao, Z. Jia, and H. Lai, "Image Reconstruction of Multibranch Feature Multiplexing Fusion Network with Mixed Multilayer Attention," *Remote Sensing,* vol. 14, no. 9, p. 2029, 2022. [Online]. Available: https://www.mdpi.com/2072-4292/14/9/2029.

[18] Y. Yan, L. Zhang, J. Li, W. Wei, and Y. Zhang, "Accurate spectral super-resolution from single RGB image using multi-scale CNN," in *Chinese Conference on Pattern Recognition and Computer Vision (PRCV)*, 2018: Springer, pp. 206-217.

[19] R. Lan, L. Sun, Z. Liu, H. Lu, C. Pang, and X. Luo, "MADNet: A Fast and Lightweight Network for Single-Image Super Resolution," *IEEE Transactions on Cybernetics,* vol. 51, no. 3, pp. 1443-1453, 2021, doi: 10.1109/TCYB.2020.2970104.

[20] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image Super-Resolution Via Sparse Representation," *IEEE Transactions on Image Processing,* vol. 19, no. 11, pp. 2861-2873, 2010, doi: 10.1109/TIP.2010.2050625.

[21] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE transactions on pattern analysis and machine intelligence,* vol. 33, no. 5, pp. 898-916, 2010.

[22] E. Agustsson and 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)*, 21-26 July 2017 2017, pp. 1122-1131, doi: 10.1109/CVPRW.2017.150.

[23] M. Bevilacqua, A. Roumy, C. Guillemot, and M.-L. Alberi-Morel, "Low-Complexity Single Image Super-Resolution Based on Nonnegative Neighbor Embedding," 09/01 2012, doi: 10.5244/C.26.135.

[24] J. B. Huang, A. Singh, and N. Ahuja, "Single image super-resolution from transformed self-exemplars," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 7-12 June 2015 2015, pp. 5197-5206, doi: 10.1109/CVPR.2015.7299156.

[25] Y. Matsui *et al.*, "Sketch-based manga retrieval using manga109 dataset," *Multimedia Tools and Applications,* vol. 76, no. 20, pp. 21811-21838, 2017/10/01 2017, doi: 10.1007/s11042-016-4020-z.